

Fisher Linear Discriminant Based Person Identification Using Visual Evoked Potentials

A. Yazdani, A. Roodaki, S. H. Rezatofighi, K. Misaghian, and S. K. Setarehdan^{*}
*Control and Intelligent Processing Centre of Excellence, Faculty of ECE, University of Tehran,
Tehran, Iran*
E-mail: ksetareh@ut.ac.ir

Abstract

Biometrics is the technique of uniquely recognizing a person among a group of people. It is usually performed based on one or more of human's intrinsic physical or behavioral traits. One such trait is the electroencephalogram (EEG) signal. In this paper, the feasibility of Visual Evoked Potential (VEP) in the gamma band of EEG signal, as a physiological trait, is studied, and used to identify individuals in a group of 20 people. To this end, the parameters of the AR model together with the peak of the power spectrum density (PSD) of the gamma band VEP signal (GMVEP) are considered as main features. Next, the Fisher's Linear Discriminant (FLD) is used to reduce the feature vector dimensions. Finally, the K Nearest Neighbor (KNN) technique is employed to classify the data and the leave-one-out cross validation method is used for accuracy assessment. A correct classification rate of 100% is achieved.

1. Introduction

Person identification has applications in many access control systems where the high level of accuracy and security is essential. Methods such as using ID cards or entering passwords may not be beneficial in some person identification applications due to their low level of security. To overcome this shortage, different biometric traits such as fingerprints, iris, speech, face, etc. were used for person identification in the past [1-5]. These traits, however, can be altered or manipulated either intentionally or accidentally. For this reason, researchers have been investigating for new biometrics in the last couple of years which can satisfy the important conditions of (a) being unique for any individual; (b) not changed or be forged easily; and (c) being not so complex for analyzing. Considering these conditions, such

biometrics as electrocardiogram (ECG) [6] and electroencephalogram (EEG) [10]-[15] has been proposed for person identification.

Several studies have shown the relationship between the EEG signals collected from the scalp of a person and his/her genetic code [7]-[9]. The genetic code is a unique characteristic of each person; hence, it is possible that the EEG signal is specific enough to be utilized as a safe biometric trait in person identification systems. Different EEG signal analyzing methods for person identification were proposed in the past. The authors in [10] extracted Auto-Regressive (AR) model parameters of the EEG signal and used them as features for classification using a Learning Vector Quantization network (LVQ). They reported the maximum classification accuracy of 84%. In [11] the AR model parameters are used to estimate the second order statistics of the EEG signal and a classification accuracy of 80% was achieved. In [12] the gamma band spectral power of the Visual Evoked Potential of the EEG signal is used within a Simplified Fuzzy ARTMAP (SFA) neural network framework and an average classification accuracy of 94.18% was reported.

In this paper, the VEP signal continuation in the gamma band (GBVEP), i.e. the frequency contents between 30Hz and 50Hz, is used as a biometric trait. The AR model parameters of different orders together with the peak value of the Power Spectral Density (PSD) of the signal are extracted and used as features. The Fisher Linear Discriminant (FLD) is then used for feature dimension reduction. The features are then classified using the K-Nearest Neighbor (KNN) method. For validation of the proposed algorithm the leave-one-out cross validation procedure is employed. A classification accuracy of 100% is achieved on a data set obtained from 20 healthy subjects.

The rest of the paper is organized as follows. Section 2 describes the proposed algorithm together

with the EEG dataset used in this study. Section 3 provides the results and a comparative study considering the results of other previously reported methods. Finally section 4 concludes the paper.

2. Methodology

2.1. The Dataset

The EEG signals used in this research are taken from a large dataset developed to examine the correlations of EEG signal with the genetic predisposition to alcoholism. This study was performed at the Neurodynamics Laboratory of the State University of the New York Health Center at Brooklyn. Two groups of subjects were assessed in this study: alcoholic and control. Each subject was exposed to either a single stimulus (S1) or to two stimuli (S1 and S2) which were pictures of objects chosen from the 1980 Snodgrass and Vanderwart picture set. When two stimuli were shown, they were presented in either a matched condition where S1 was identical to S2 or in a non-matched condition where S1 was different from S2.

The signal acquisitions were carried out by placing 61 active electrodes on the scalp of the subjects. Each signal has a time length of 1 second and was sampled with a sampling frequency of 256 Hz.

A total number of 1788 EEG signals, each of length one second, was selected from the dataset. In this study the signals obtained from twenty of the normal subjects are used for person identification.

2.2. Feature extraction

2.2.1. VEP extraction from the EEG signal. The signal used in this research for person identification is the VEP signal which is extracted from the background EEG in the frequency range 30Hz to 50Hz. This is carried out using the band-pass filter. Three typical examples of the EEG signals together with the corresponding extracted GBVEP signals are shown in Figure 1.

2.2.2. Autoregressive model. Due to the short length (1 second) of each EEG signal in the data set, it can be considered stationary. The AR model of a signal $X(t)$ is as follows:

this matrix can be considered as a separability measure. By applying the SVD to the separability matrix and discarding the insignificant eigenvalues and

$$X(t) = \sum_{k=1}^P a(k).X(t-k) + e(t) \quad (1)$$

where P is the order of the model, $e(t)$ is an independent white noise, and the $a(k)$ s are the model parameters. These parameters can be estimated via several methods. In this research the Burg's algorithm [16] is used.

2.2.3. Power Spectrum Density. The power spectrum density (PSD) describes how the power (or variance) of a time series is distributed with respect to frequency. Mathematically, it can be defined as the Fourier transform of the autocorrelation sequence of the time series. It can also be computed using the AR model parameters of the time series. In this paper the estimation of the PSD were performed by means of AR modeling and the Burg method. The peaks of the PSD of GBVEP signals of different electrodes were considered as another feature of the GBVEP signal.

2.3. Feature reduction

The aim of this stage is to reduce the dimension of the feature vectors and at the same time to improve the classification accuracy by using the Fisher's criterion which is defined as follows:

$$FDR = \frac{(\mu_1 - \mu_2)^2}{\sigma_1^2 + \sigma_2^2} \quad (2)$$

where μ_i and σ_i represent the mean and standard deviation of each class, respectively. Let us assume that there are L different classes of C_1, C_2, \dots, C_L for which the number of data points are N_1, N_2, \dots, N_L respectively. In order to use the Fisher's criterion, the so-called "within-class scatter matrix" and "between-class scatter matrix" are defined as follows:

$$WCS = S_w = \sum_{k=1}^L S_k \quad (3)$$

$$BCS = S_b = \sum_{k=1}^L N_k (\mu_k - \mu)(\mu_k - \mu)^t \quad (4)$$

where

$$S_k = \sum_{x \in C_k} (x - \mu_k)(x - \mu_k)^t \quad (5)$$

In (5), x is the feature vector and μ is the mean of all datapoints.

Now the Fisher's criterion can be defined as $S_w^{-1}S_b$ which is called the separability matrix. The trace of

their corresponding columns in the eigenvector matrix, the linear transform matrix and the new dimension-reduced feature vectors are obtained. It can

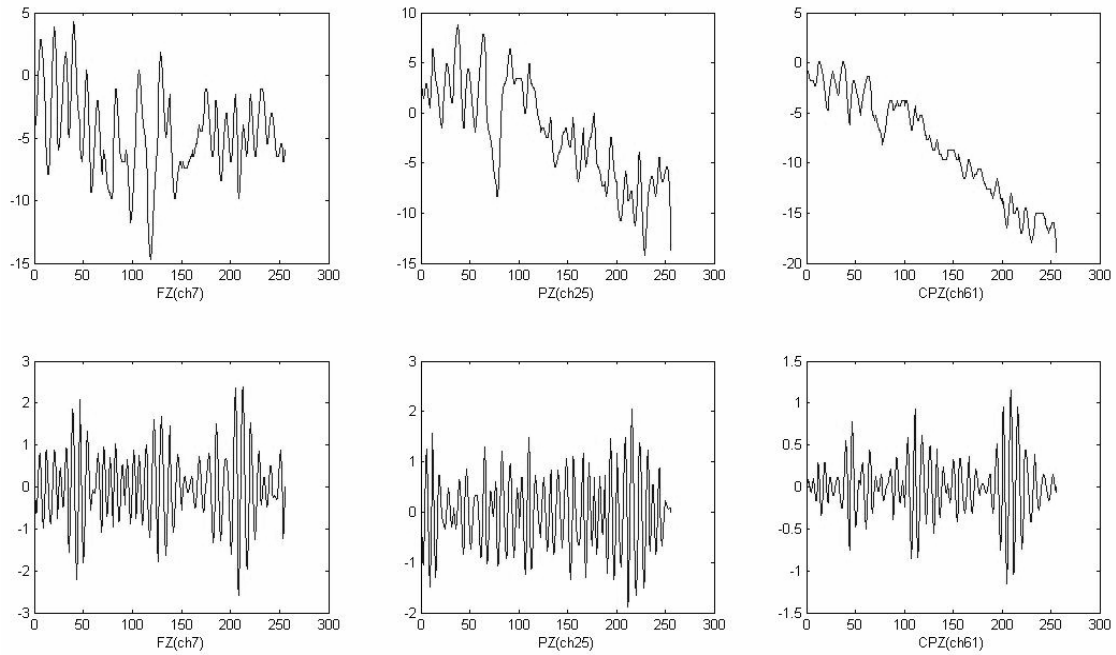


Figure 1 - Three typical EEG signals from different channels (upper row) and the extracted GBVEP signal (lower row).

be readily shown that the discarded eigenvalues do not have a significant role in the trace of the separability matrix. The remaining eigenvalues and their corresponding eigenvectors would serve as a linear transform of the feature matrix. It can be shown that the final dimension of feature vectors after this stage would be L-1.

Since the data used in the current study were from 20 different classes, hence applying the aforementioned method, the dimension of the feature vectors were reduced to 19.

2.4. Classification

The K-Nearest Neighbor (KNN) majority voting based on Euclidean distance classifier is used here in order to categorize the 20 subjects. For this purpose, the distances between the input test feature vector and all other vectors in the database are calculated, and then the K nearest feature vectors are selected. The unknown feature vector is then assigned to the class which has the maximum number of feature vectors among the K-nearest neighbors. This algorithm was implemented for different number of neighbors (K) from K=1 to 100. An optimal value of K=5 is obtained for classification of the dataset using the reduced feature vectors.

3. Result and discussion

As mentioned earlier, the total of 1788 feature vectors for 20 subjects were used in the experimental study. At this stage, the leave-one-out cross validation was used to analyze the performance of the classifier. The Euclidean distances between a feature vector and the other vectors were computed, and the vector is assigned to the class which has more nearest feature vectors among the winners.

The results are shown in Figure 2 for different AR orders. In these Box Plots, the median and quartile of the results of the classifier for each AR order and for different values of K (ranging from 1 to 100) is demonstrated. It can be seen from this figure that an accuracy of 100% is achieved for any AR model with an order more than 14.

4. Conclusion

In this paper, the feasibility of the VEP signal continuation in the gamma band extracted from the EEG signal, for person identification was studied. The KNN classifier based on the Euclidean distance could effectively identify persons based on the AR model parameters extracted from VEP signals of the brain activities. This was accomplished using the Fisher Linear Discriminant (FLD) method applied on the original feature vectors.

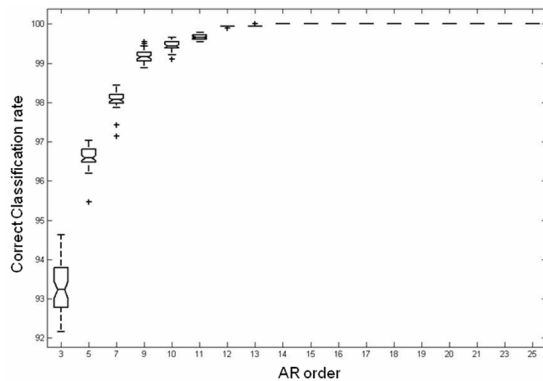


Figure 2 - Box plot of different correct classification rates versus AR orders for K ranges from 1 to 100.

There are some practical limitations regarding to this technique, such as the necessity to record the EEG signal of the subjects which can cause some inconvenience for the subjects and needs some time for setting up the electrodes.

In addition, EEG equipments are still much more expensive than other biometric sensors commercially available (such as cameras for face or eye patterns, microphone for voice or the several kinds of fingerprint sensors). However, this method could be used in situations where other methods are not very reliable. In those cases physiological and genetic tests should be carried out. Such tests could be very time consuming, expensive, and may take lot longer than the method presented in this paper. Evaluation of the method on larger data sets in future will help to study the feasibility of this method in more depth.

6. References

- [1] Z. M. Kovacs -Vajna, "A fingerprint verification system based on triangular matching and dynamic time warping," *IEEE Trans. Pattern. Anal.*, vol. 22, no. 11, 2000.
- [2] P. A. R. Rodriguez, and J. D. L. Silva, "Biometric identification by dermatoglyphics," *International Conference on Image Processing*, vol. 1, pp. 16–19, 1996.
- [3] R. Kaushik and B. Prabir, "An iris recognition method based on zigzag collarette area and asymmetrical Support Vector Machines," *Proceedings of the ICSMC conference*, vol. 1, pp. 861–865, 2006.
- [4] N. A. Fox, R. Gross, J. F. Cohn, and R. B. Reilly, "Robust biometric person identification using automatic classifier fusion of speech, mouth, and face experts," *IEEE Trans. Multimedia.*, vol. 9, no. 4, 2007.
- [5] L. Chen, H. Y. M. Liao, and J. Lin, "Person identification using facial motion," *International Conference on Image Processing*, 2001.
- [6] A. D. C. Chan, M. M. Hamdy, A. Badre, and V. Badee, "Person identification using electrocardiograms,"

Canadian Conference on Electrical and Computer Engineering, 2006.

- [7] A. P. Anokhin, V. Muller, U. Lindenberger, A. C. Heath, and E. Myers, "Genetic influences on dynamic complexity of brain oscillations," *Neurosci. Lett.*, vol. 397, pp. 93–98, 2006.
- [8] B. P. Zietsch, J. L. Hansen, N. K. Hansell, G. M. Geffen, N. G. Martin, and M. J. Wright, "Common and specific genetic influences on EEG power bands delta, theta, alpha, and beta," *Biol. Psychol.*, vol. 75, pp. 154–164, 2007.
- [9] Y. Tang, S. O'connor, L. Bauer, D. B. Chorlian, M. Rangaswamy, M. Schuckit, S. Kuperman, and J. Rohrbaugh, "Genetic influences on bipolar EEG power spectra," *Int J Psychophysiol.*, vol. 65, pp. 2-9, 2007.
- [10] M. Poulos, M. Rangoussi, V. Chrissikopoulos, and A. Evangelou, "Person identification based on parametric processing of the EEG," *In: Proc. 6th IEEE International Conference on Electronics, Circuits, and Systems*, vol. 1, pp. 283–286, 1999.
- [11] R. B. Paranjape, J. Mahovsky, L. Benedicenti, and Z. Koles, "The electroencephalogram as a biometric," *Int Proc. Canadian Conference on Electrical and Computer Engineering*, vol. 2, pp. 1363–1366, 2001.
- [12] R. Palaniappan, and K. V. R. Ravi, "A new method to identify individuals using signals from the brain," *Proceedings of the Conference on Information, Communications and Signal Processing*, vol.3, pp. 1442–1445, 2003.
- [13] R. Palaniappan and K.V.R. Ravi, "Improving visual evoked potential feature classification for person recognition using PCA and normalization," *Pattern Recogn. Lett.*, vol. 27, pp. 726–733, 2006.
- [14] R. Palaniappan and D. P. Mandic, "Biometrics from brain electrical activity: A machine learning approach," *IEEE Trans. Pattern Anal.*, vol. 29, no. 4, 2007.
- [15] G. Mohammadi, P. Shoushtari, B. Molaei Ardekani, and M. B. Shamsollahi, "Person identification by using AR model for EEG signals," *IEEE Trans. Eng. Comput. Technol.*, vol. 11, pp. 281–285, 2006.
- [16] R. Bos, S. de Waele, and P. M. T. Broersen, "Autoregressive spectral estimation by application of the Burg algorithm to irregularly sampled data," *IEEE Trans. Instrum. Meas.*, vol. 51, no. 6, pp. 1289–1294, 2002.