A Minimal Channel Set for Individual Identification with EEG Biometric Using Genetic Algorithm

K.V.R. Ravi[1] and R. Palaniappan[2]

[1]*School of Information & Communications Technology, Republic Polytechnic, 9 Woodlands Ave 9, 738964, Singapore*.

[2]*Biosignal Analysis Group, Dept. of Computing and Electronic systems, Colchester, CO4 University of Essex, Vivenhoe Park, 3SQ, United Kingdom.*

[1]*kvr_ravi@rp.sg;* [2]*rpalan@essex.ac.uk*

Abstract

In this paper, we explore the use of genetic algorithm (GA) to select a minimum number of channels that identifies individuals based on brain signals i.e. electroencephalogram (EEG). The fusion of GA with linear discriminant classifier shows that the identification performance of EEG signals from 40 subjects does not degrade when using 23 selected channels as compared to all the available 61 channels as studied previously. As the channel identification method by GA is general, it could be used in any feature reduction application.

1. Introduction

The standard method for identifying an individual is through the use of fingerprints [1] but in recent years, there has been significant interest in using other biometrics for identifying individuals. These include techniques that rely on:- DNA, hand geometry, palm print, face (both optical and infrared), iris, retina, signature, ear shape, odor, keystroke entry pattern, gait, and voice [2]. Other emerging biometrics such as ear force fields [3], heart signals [4], and brain signals [5-7] have also been proposed in recent years. As signal recording from the brain is rather complicated, biometrics based on brain signals has not been studied extensively though it is one of the most fraud resistant biometrics.

There are only a handful of studies that have utilized this brain signal based biometric. These include results by Paranjape et al [6] who studied that autoregressive (AR) modeling of electroencephalogram (EEG) in combination with discriminant analysis and achieved a classification accuracy ranging between 49% and 85%, while Poulus et al [7] studied the problem of distinguishing an individual from the rest using a set of EEG recordings. Their method was based on AR modeling of EEG signals and Linear Vector Quantization (LVQ) neural network (NN), which gave 72-80% classification accuracy. However, this method was not tested on the task of recognition of individual subjects.

The objective of this paper is to provide further perspective on the use of EEG biometric by minimizing the number of required channels. The approach here is an extension to the one proposed in [5], where individual identification was achieved using features from 61 channels.

A problem encountered in the method proposed in [5] is the determination of channels or electrodes that carry significant information for identification purposes. This is especially true with modern EEG measuring instruments of many electrodes, where it is often preferable to use signals from certain channels. This is since some channels carry significant information while the other channels either impair or do not influence the identification results. Therefore, the identification of suitable channels would minimize the number of required channels and may even increase the identification performance.

GA to select features for EEG classification of a Brain Computer Interface has been investigated in [8]. This method requires two classifiers, a k-nearest neighbor classifier to evaluate the GA population fitness and LVQ3 algorithm to classify the different mental thought processes represented by EEG. Similarly, the method in [9] used two neural network classifiers, Fuzzy ARTMAP (FA) and multi-layer perceptron (MLP) trained by the backpropagation (BP) algorithm. However, the use of neural networks is computationally expensive especially when used with GA.

Another method to select relevant electrodes for EEG classification of hand movements using principal component analysis (PCA) has been proposed in [10]. However, PCA maximizes signal representation with minimum features. This might not necessarily maximize classification performance, which is however the advantage of using GA.

In this study, a reduction in the number of required channels for identifying individuality using brain signals is sought using genetic algorithm (GA) fused with a single linear discriminant classifier (LDC).

2. Data

EEG signal data recorded non-invasively from the scalp were used. EEG signals are electrical potentials exhibited by neuronal excitations in the cortex [11].

To obtain EEG signals in gamma frequency range, filtering was performed, and the energies of these filtered signals were used as a set of features (after some pre-processing) to be classified by the simple LDC.

The subjects (totalling 40) were seated in a reclining chair located in a sound attenuated RF shielded room. Measurements were taken from 61 active channels placed on the subject's scalp, sampled at 256 Hz. The electrode positions were according to the extension of Standard Electrode Position Nomenclature, recommended by the American Encephalographic Association.

The EEG signals were recorded from subjects while being exposed to a stimulus, which consist of drawings of objects chosen from Snodgrass and Vanderwart picture set [12]. These pictures represent common black and white objects, such as, for instance, airplane, banana, and ball. These were chosen according to a set of rules that provides consistency of pictorial contents. They have been standardised based on the variables of central relevance to memory and cognitive processing. These objects had definite verbal labels, i.e. they could be named.

The subjects were asked to remember or recognise the stimulus. Stimulus duration of every picture was 300 ms with an inter-trial interval of 5.1s. All the stimuli were shown using a display located 1 meter away from the subjects. One-second EEG measurements after each stimulus onset were stored. Figure 1 illustrates a stimulus presentation. This data set used is a subset of a larger experiment designed to study the short-term memory [13].

EEG signals contaminated with eye blink artifacts were not considered in the classification, and were detected using a 100 μV threshold. This is a common threshold value in EEG studies, and is used since blinking produces 100-200 μV potential lasting 250 milliseconds [14]. A total of 40 artifact free trials were considered for every subject, to make a total 1600 EEG data sets.

The EEG signals were filtered using a forward and reverse Elliptic band-pass digital filter, to obtain zero phase distortion. The 3-dB pass-band was chosen to be between 30 and 50 Hz, whereas the stop-band was fixed at 28 and 52 Hz. The minimum stop-band attenuation was set at 20 dB.

To form the EEG features, the energy of the EEG signal from each channel was computed and normalised according to the total energy from all 61 channels.

3. Methodology

GA is a family of computational models inspired by evolution and is based on genetic processes of biological organisms. They are adaptive methods, which may be used to solve search and optimization problems. Over many generations, natural populations evolve according to the principles of natural selection and “survival of the fittest” [15].

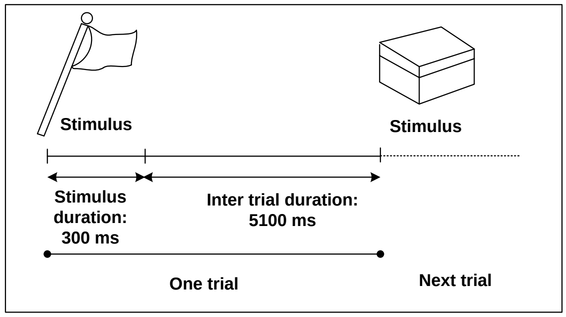

Figure 1. Example of visual stimulus presentation

GA requires fitness or objective function, which provides a measure of performance of the population individuals. The evaluation function must be relatively fast since GA incurs the cost of evaluating the population of potential solutions. This is why we have used LDC classification to evaluate the fitness function and not other types of neural networks like MLP-BP or FA.

The dataset of 40 patterns from each subject is split randomly into four non-overlapping sets with each consisting of 10 patterns, i.e. each dataset consist of 400 patterns. GA uses datasets 1 and 2. The other two sets are not used here to ensure unbiasness in the ability of GA to select optimal channels.

Initially, a set of populations is generated as random binary strings (a sequence of 1’s and 0’s) with a certain number of bits used to represent the active/inactive state of the channel. A value of 1 denotes the activation of the channel feature (i.e. the channel feature is used) and a value of 0 denotes deactivation of the channel feature (i.e. the channel feature is not used). In our case, we have 61 channels; therefore we need 61 bits to represent each chromosome. Following this convention we generate 100 chromosomes. Figure 2 illustrates this situation.

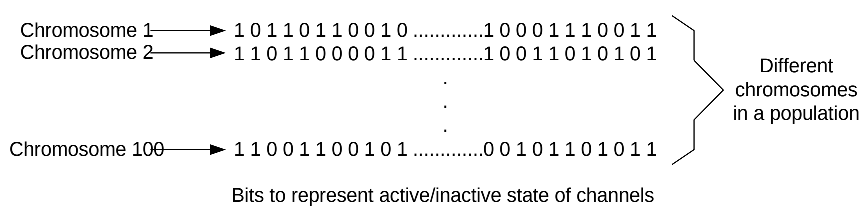

Figure 2. Initial GA population

Using this population, features from EEG pattern of the active channels from dataset 1 are fed into LDC to be trained. Since GA requires LDC classification performance as a measure of fitness of the population, we need to evaluate the performance of this population. EEG features of the same active channels from the dataset 2 data are now used to evaluate the LDC performance in identifying the identity of the subjects. This process of training and evaluation is repeated for all the chromosomes in the population. The fitness function for each population is

$$fitness_{population} = EEG_{correct} / EEG_{total} + 0.5 * channels_{inactive} / channels_{total}, \quad (1)$$

where $EEG_{correct}$ equals the correctly classified EEG patterns and EEG_{total} equals the total number of EEG patterns in dataset 2; $channels_{inactive}$ represents the inactive channels (represented by 0 in the chromosome) and the value of $channels_{total}$ is 61 to represent the total 61 channels. The weight of 0.5 is used to give more weight to improved classification performance rather than minimization of channels.

GA uses the performance from this evaluation step to generate the populations in the next generation using selection, crossover, mutation and inversion operators. Three selection operators were used here: tournament, elite, and roulette wheel. Tournament selection is applied during reproduction from a pool of 25 chromosomes chosen randomly among the total populations and the best chromosome (i.e. with the highest fitness) is stored. This is repeated 33 times to obtain 33 offspring chromosomes.

Elite method is used to keep the good parent chromosomes where the best 33 chromosomes are duplicated as 33 offspring chromosomes next, roulette wheel method is used to generate another 34 offspring chromosomes.

A two-point crossover is used since they are able to wrap around at the end of the string and therefore better than a single point crossover. Two chromosomes are chosen randomly and crossover is performed if a random number generated exceeds the crossover probability. Similarly, an inversion is performed between two selected points in a randomly chosen chromosome if a random number generated exceeds the inversion probability. A mutation of a randomly selected bit in a randomly selected parent is performed if a random number generated exceeds mutation probability. The initial crossover probability is set at 0.5 while the mutation is set at a lower probability of 0.1 to reduce excessive random perturbations. The inversion probability is set to a very low value of 0.01 to avoid severe damages to the fitness value that is possible with inversion operator. These probability values are reduced as the generations increase by

$$probability = initial_{probability} * (1 - generation / \max\_gen), \quad (2)$$

This entire cycle is then iterated for 100 generations and the best chromosome is stored. Figure 3 illustrates this operation.

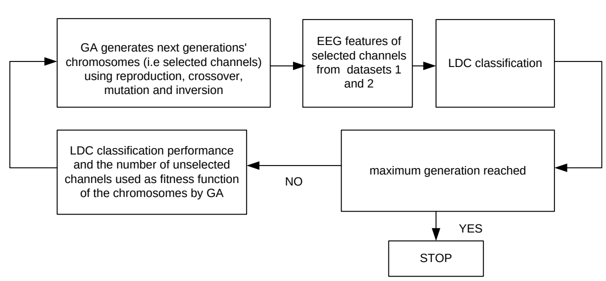

Figure 3. GA method to select optimal channels

As the initial search space does determine the final maximum point and to obtain one unique result, the GA procedure described above is repeated 50 times. Mean of the 50 chromosomes is obtained and the channel is considered selected if the mean value is above a certain threshold, T. The higher the T, the smaller the number of selected channels. Using three different values of T: 0.1, 0.3 and 0.5, we obtained 40, 23 and 13 selected channels.

4. LDC Results

LDC is used with datasets 3 and 4, which GA has not seen earlier. Classification is performed using a 20 fold equal class cross validation procedure. The 20 classification results of using all 61 channels and the selected 13, 23 and 40 channels are shown in Figure 4.

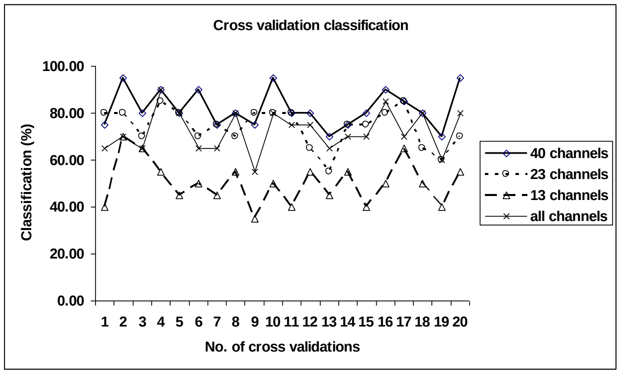

Figure 4. LDC cross validation results using different number of channels

Using Student's t-test and the 20 classification results from the cross validation procedure, it was determined that the performance of using 23 channels was similar to all the channels ($p=0.26$), while using 13 channels gave a lower performance ($p=1.39e-8$) and using 40 channels gave superior performance ($p=0.0004$). Hence, it could be concluded that the use of 23 channels gives similar performance to the use of all 61 channels.

Table 1 gives the 23 channels selected by GA, while Figure 5 shows the location of these channels. The locations are of some importance but it is beyond the scope of this present study as the aim of this study is only on the reduction of the number of channels.

Table 1. 23 Channels selected by GA

FP1 F8 AF1 F3 FC6 FC5 FC1 CZ PO2 PO1 O2 AF7 FT7 FT8 FC3 TP7 P6 C2 PO7 PO8 POZ P1 CPZ

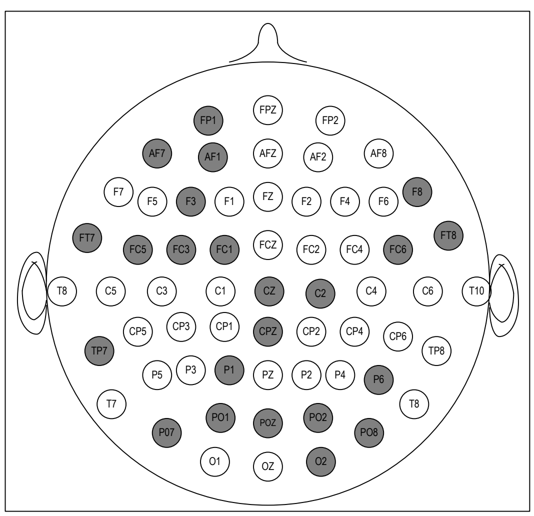

Figure 5. The locations of the 23 channels selected by GA

5. Conclusion

We have proposed a method to select channels or electrodes that are discriminatory to minimize the number of channels while maintaining similar individual identification performance using EEG biometric. This method uses GA combined with LDC. The classification results show that the use of the selected optimal channels would significantly reduce computational time and hardware/experimental set-up complexity while maintaining the classification performance. This is since the proposed method can pick up the discriminatory channels that are vital for classification from channels that impair or do not influence classification. Since the method is general, it could be used for any feature reduction in classification applications.

We hope that this study will stimulate and encourage further exploration on the rather neglected but promising EEG biometric.

Acknowledgement

The authors thank the late Prof. Henri Begleiter at the Neurodynamics Laboratory at the State University of New York Health Centre at Brooklyn, USA who generated the raw EEG data and Mr. Paul Conlon, of Sasco Hill Research, USA for sending the data.

References

[1] Uludag, U., Pankanti, S., Prabhakar, S., and Jain A.K., "Biometric cryptosystems: Issues and challenges," *Proceedings of the IEEE*, vol. 92, no.6, pp.948-960, June 2004.

[2] Pankanti, S. Bolle, R.M. and Jain, A., "Biometrics, the future of identification," *IEEE Computer (Special Issue on Biometrics)*, vol. 33, issue 2, pp.46-49, 2000.

[3] Hurley, D., Nixon, M., and Carter, J., "Force field feature extraction for ear biometrics," *Computer Vision and Image Understanding*, vol. 98, no. 3, pp. 491-512, 2005.

[4] Biel, L., Pettersson, O., Philipson, L., and Wide, P., "ECG analysis: a new approach in human identification," IEEE Transactions on Instrument and Measurement, vol. 50, no.3, 2001, pp. 808-812.

[5] Ravi, K.V.R., and Palaniappan, R. "Neural network classification of late gamma band electroencephalogram features," *Soft Computing*, vol. 10, no.2, pp. 163-169, 2006.

[6] Paranjape, R.B., Mahovsky, J., Benedicenti, L., and Koles, Z., "The electroencephalogram as a biometric," in *Proceedings on Canadian Conference on Electrical and Computer Engineering*, vol. 2, pp. 1363-1366, 2001.

[7] Poulos, M., Rangoussi, M., Chrissikopoulos, V., and Evangelou, A., "Person identification based on parametric processing of the EEG," in *Proceedings on IEEE International Conference on Electronics, Circuits, and Systems*, vol. 1, pp. 283-286, 1999.

[8] Muller, T., Ball, T., Kristeva-Feige, R., Mergner, T. and Timmer, J., "Selecting relevant electrode positions for classification tasks based on the electro-encephalogram," *Medical and Biological Engineering and Computing*, vol. 38, pp. 62-67, 2000.

[9] Palaniappan, R., Raveendran, P., and Omatu, S., "EEG optimal channel selection using genetic algorithm for neural network classification of alcoholics," IEEE Transactions on Neural Networks, pp. 486-491, vol. 13, no. 2, 2002.

[10] Flotzinger, D., Pregenzer, M. and Pfurtscheller, G., "Feature Selection with Distinction Sensitive Learning Vector Quantization and Genetic Algorithm," *Proceedings of IEEE International Conference on World Congress on Computational Intelligence*, vol. 6, pp. 3448-3458, 1994.

[11] Misulis, K.E., *Spehlmann's Evoked Potential Primer: Visual, Auditory and Somatosensory Evoked Potentials in Clinical Diagnosis*, Butterworth-Heinemann, 1994.

[12] Snodgrass, J.G., and Vanderwart, M., "A standardized set of 260 pictures: norms for name agreement, image agreement, familiarity, and visual complexity," *Journal of Experimental Psychology: Human Learning and Memory*, pp. 174-215, vol. 6, no. 2, 1980.

[13] Zhang, X.L., Begleiter, H., Porjesz, B., Wang, W., and Litke, A., "Event related potentials during object recognition tasks," *Brain Research Bulletin*, pp. 531-538, vol. 38, no. 6, 1995.

[14] A.M. Halliday (ed.), *Evoked Potentials in Clinical Testing*, Churchill Livingstone, 1993.

[15] Goldberg, D.E., *Genetic Algorithm in Search, Optimization and Machine Learning,* Reading Mass., Addison–Wesley, 1989.