International Journal of Knowledge-based and Intelligent Engineering Systems 9 (2006) 1–8 1
IOS Press

Simplified fuzzy ARTMAP classification of individuals using optimal VEP channels

K.V.R. Ravi[a,*], Ramaswamy Palaniappan[c] and S.-H. Heng[b]
[a]*School of Information & Communications Technology, Republic Polytechnic, 738964, Singapore*
[b]*Faculty of Information Science and Technology, Multimedia University, Melaka, 75450, Malaysia*
[c]*Department of Computer Science, University of Essex, Colchester, CO4 3SQ, United Kingdom*

Abstract. In previous studies, identification of individuals using 61 channel Visual Evoked Potential (VEP) signals from the brain has been shown to be feasible. These studies used neural network classification of gamma band spectral power of VEP signals from 20 individuals. This paper explores our continuing work in this area to include more subjects in the experiment and to reduce the number of required channels using Fisher Discriminant Ratio function. The experimental study showed that 27 optimal channels were sufficient to yield an average classification rate of 90.97% across 800 test VEP patterns from 40 subjects. Being fewer in number than 61 channels, it is less cumbersome, requires lower computational time, design complexity and cost. This was achieved without loss of performance as 61 channels gave an average classification result of 89.11%. The positive results obtained here showed that the neural activity during perception of visual stimulus was different across individuals. This method could be explored further as a biometric tool to identify individuals as the brain signals are difficult to be forged.

Keywords: Fisher Discriminant Ratio, gamma band power, neural network, individual identification, optimal channels, Visual Evoked Potential

1. Introduction

Human identification is a challenging avenue in the biometric research area. There are many biometrics such as fingerprint [11,12], iris [2], palmprint [3], hand geometry [4] and face [15] to identify the individuals. But some of these methods have their drawbacks like individuality problem [12] and forgery. The individuality problem is the problem that the fingerprints are not proven scientifically to be unique to each individual though it is the most widely used biometric. Therefore, it is important to come up with newer biometric methods that could augment or replace the existing biometrics.

There are not many published work on the use of brain signals as biometric to identify individuals. Paranjape et al. [13] proposed an approach for EEG classification using neural networks (NN) and discriminant analysis. Poulus et al. [14] proposed a method using autoregressive (AR) modelling of EEG signals and Linear Vector Quantisation (LVQ) NN to classify an individual as distinct from other individuals.

Visualising a picture evokes perception and memory and it is our assumption that this level of neural activity between individuals is different. The study in [10] indicated that the levels of neural activity, measured as EEG, were different across subjects even for similar mental activity. Using this assumption, the previous studies [8,9] have used spectral powers in gamma band range of 30–50 Hz computed from the VEP signals recorded during a visual stimulus. These spectral powers were then used to classify the 20 individuals using either Multilayer Perceptron with Backpropagation training (MLP-BP) or Simplified Fuzzy ARTMAP

*Corresponding author: K.V.R. Ravi, Academic Staff, School of Information & Communications Technology, Republic Polytechnic, 9 Woodlands Ave 9, 738964, Singapore. Tel.: +65 65104097; Fax: +65 64151310; E-mail: kvr_ravi@rp.sg. Co-authors E-mails: rpalan@essex.ac.uk; shheng@mmu.edu.my.

ISSN 1327-2314/06/$17.00

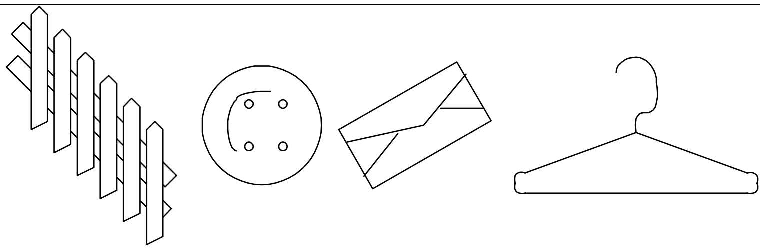

Fig. 1. Examples of pictures from Snodgrass and Vanderwart picture set.

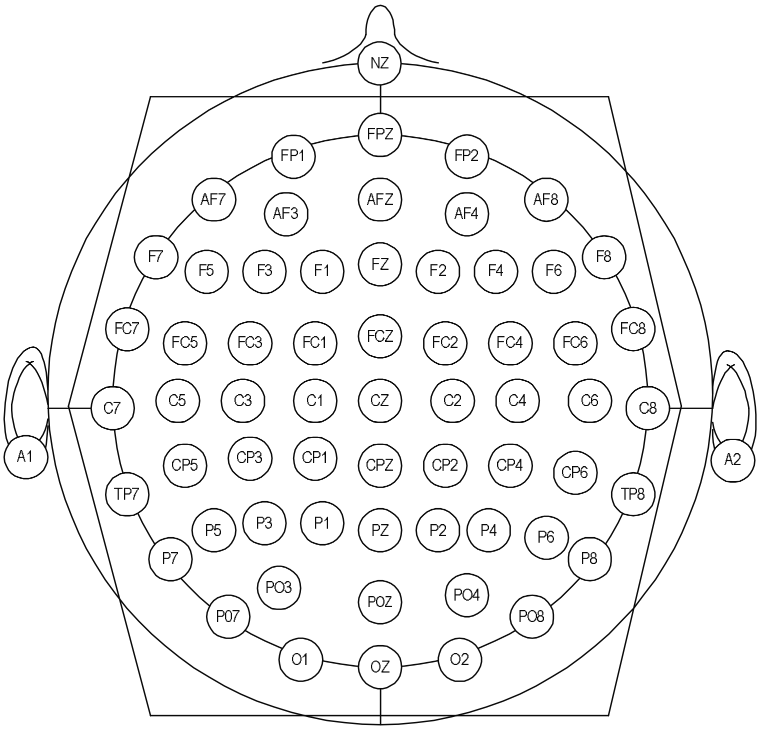

Fig. 2. Electrode locations.

(SFA) neural networks (NNs). These studies used 61 VEP channels.

The aim of this research work is to include more subjects in the experimental study and to reduce the number of required VEP channels to identify individuals, without degradation in the classification performance. These optimal channels, once located will reduce the computational time, design complexity and cost due to being fewer in number.

To achieve this, VEP signals from 61 channels during a visual stimulus were recorded. The VEP signals were subjected to preprocessing using Principal Component Analysis (PCA) for noise reduction and gamma band powers (GBPs) were obtained using discrete-time bandpass Butterworth filter and Parseval's time-frequency equivalence theorem. Fishers Discriminant Ratio (FDR) function was used to compute the discriminatory power of each channel. Next, SFA classifications were carried out using the optimal channels.

2. Methodology

The recording of VEP data, pre-processing, extracting the feature and its classification is similar to the studies in [8,9]. In addition, FDR is used to explore the optimal channels with the help of additional 20 subjects VEP data.

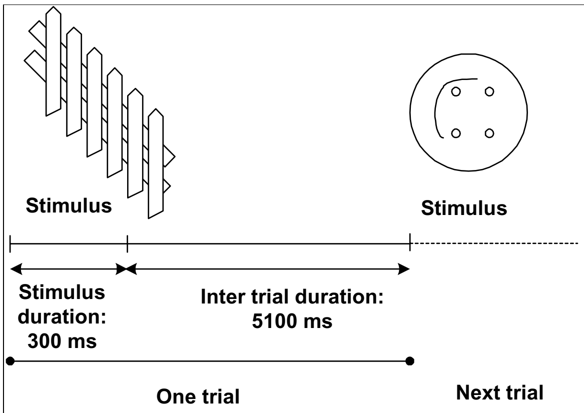

Fig. 3. Example of stimulus presentation.

2.1. VEP data

Data from 40 subjects were used in the experimental study. The average age and SD of the subjects were 23.7 $\pm$ 2.4 years. Half of them were male while the rest half female. The subjects were subjected to visual stimuli by showing some pictures. These pictures were collections from Snodgrass and Vanderwart (SV) picture set [17]. The SV picture set consists of black and white drawings of common objects, such as, airplane, banana, and ball. These were chosen according to a set of rules that provides consistency of pictorial contents. They were physically different but standardised with respect to demands on memory and cognitive processing, for instance, objects can be named (definite verbal labels). Figure 1 illustrates some examples of this picture set.

When the subject viewed the object, brain signals would be generated and these were captured using 61 active electrodes on the scalp. The location of these electrodes followed the nomenclature of Standard Electrode Position, American Encephalographic Association. The electrode locations are shown in the Fig. 2. The active channels are enclosed within hexagon. The three electrode locations outside the hexagon are reference channels.

The subjects were positioned in a reclining chair located in a sound attenuated RF shielded room. The signals were hardware band-pass filtered between 0.02 and 50 Hz and sampled at 256 Hz. The subjects were asked to recognize and remember the stimulus. The stimulus was shown for 300 ms using a computer display unit located at one meter from the subject's eyes. There was an inter-trial interval of 5.1 s and one-second VEP measurements after each stimulus onset were stored.

Eye blink contamination is a common problem associated with brain signal recordings. Here, eye blink VEP signals that exceeded 100 μV were assumed to be contaminated by eye blinks and were discarded. A total of 40 eye-blink free VEP signals for each subject were stored giving a total of 1600 VEP signals. Actually, each subject looked at a total of 90 pictures, where the pictures were not repeated. However, only 40 picture-trials from each subject were used as some of the recordings were erroneous or contaminated by eye blinks. Figure 3 illustrates an example of the stimulus presentation.

The data that we have used are actually a subset of the data recorded by Zhang et al. [18] for their experiments on studying the low frequency c247 component of VEP as a marker for visual short-term memory. The data are available at http://kdd.ics.uci.edu/databases/eeg/eeg.html.

2.2. Pre-processing

PCA [6] was used to reduce noise from the VEP signals. Note that PCA was not used to reduce the dimensionality of the features. The original signal consists of signal and noise. PCA will separate the noise from the signal using the fact that, noise subspace will constitute of principal components (PCs) with eigenvalues chosen below the threshold. Eigenvalues with PCs above this threshold represent signal subspace.

PCA method is as follows. Let matrix $\mathbf{z}$ denote the extracted VEP signal, and let $\mathbf{R}$ be the covariance of $\mathbf{z}$, given by

$$\mathbf{R} = \mathbf{E}(\mathbf{z}\mathbf{z}^{\mathbf{T}}), \tag{1}$$

where $\mathbf{E}(\bullet)$ denotes the statistical expectation operator. Let $\mathbf{F}$ be the orthogonal matrix of eigenvectors of $\mathbf{R}$, and $\mathbf{D}$ the diagonal matrix of eigenvalues of $\mathbf{F}$, that is, $\mathbf{D} = diag(d_1, \ldots, d_n)$. Principal components (PCs) were computed as

$$\mathbf{y} = \mathbf{F}^{\mathbf{T}}\mathbf{z}^{\mathbf{T}}. \tag{2}$$

Next, the VEPs with reduced noise were reconstructed using

$$\tilde{x} = \tilde{\mathbf{F}}\hat{\mathbf{y}}, \tag{3}$$

where $\tilde{\mathbf{F}}$ and $\hat{\mathbf{y}}$ denote respectively the eigenvectors and PCs which correspond to eigenvalues whose values were greater than unity.

Using PCA, vector space projections are performed along the directions of the components that describe most of the signal variance (power). Based on this prin-

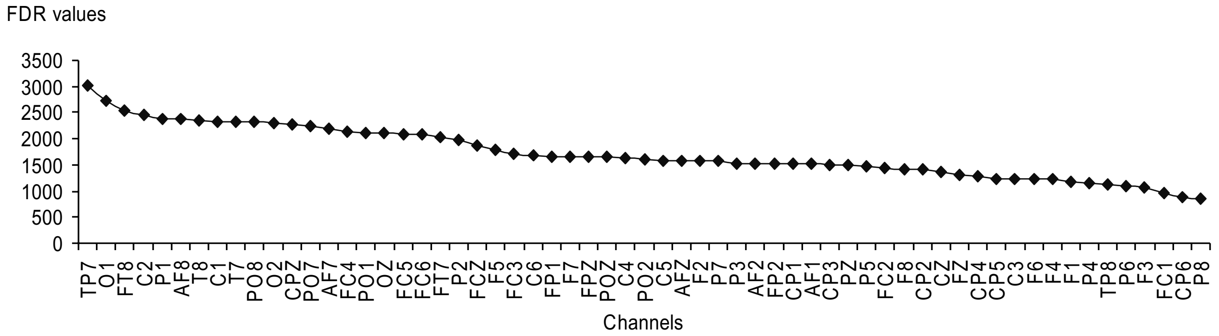

Fig. 4. FDR values for 61 channels.

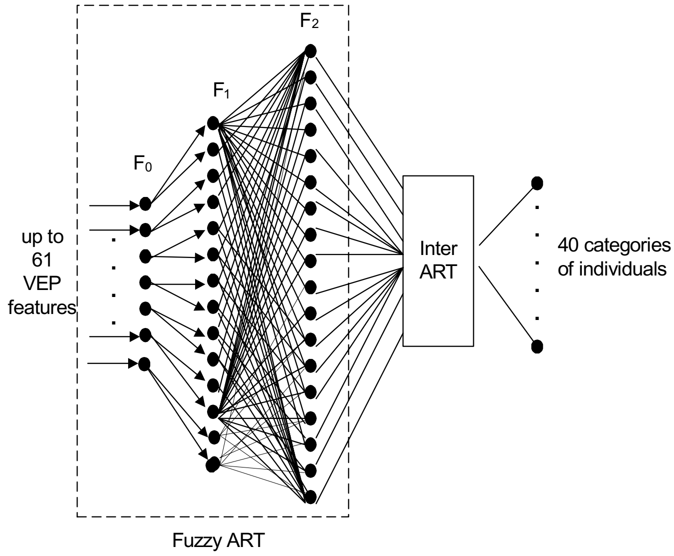

Fig. 5. Architecture of SFA.

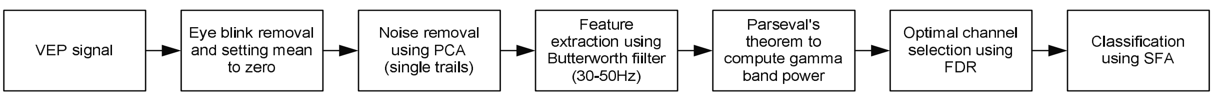

Fig. 6. Methodology of this study.

ciple, the first few PCs account for a large proportion of the VEP variance, while the rest can be attributed to the noise. In this work, Kaiser's rule [6] was used to separate the signal and noise; PCs with eigenvalue more than 1.0 were automatically selected as part of the VEP subspace. On average, 28 PCs (from a total of 61) were required to reconstruct the noise reduced VEP signals.

2.3. *Feature extraction*

A 10^{th} order forward and 10^{th} order backward Butterworth digital filter was used to extract the VEP in the 3 dB passband of 30 to 50 Hz, which is in gamma band range. The power line frequency was 60 Hz. To achieve zero phase response from the non-linear Butterworth filter, forward and reverse operations were used. The signals were subjected to the Parseval's time-frequency relation to get the gamma band power (GBP) of the signal $\tilde{x}$ using

$$GBP(\text{channel i}) = \frac{1}{N}\sum_{n=1}^{N}\left[\tilde{x}_i(n)\right]^2, \quad (4)$$

where N was the total number of data in the filtered signal. This power was normalised with the sum of power

Table 1
Average SFA performance of 27 optimal channels and all 61 channels

Channels	Classification (%)	Train time (s)	Test time (s)	Cluster size
27	90.97	0.014	0.0009	82.05
61	89.11	0.018	0.002	105.61

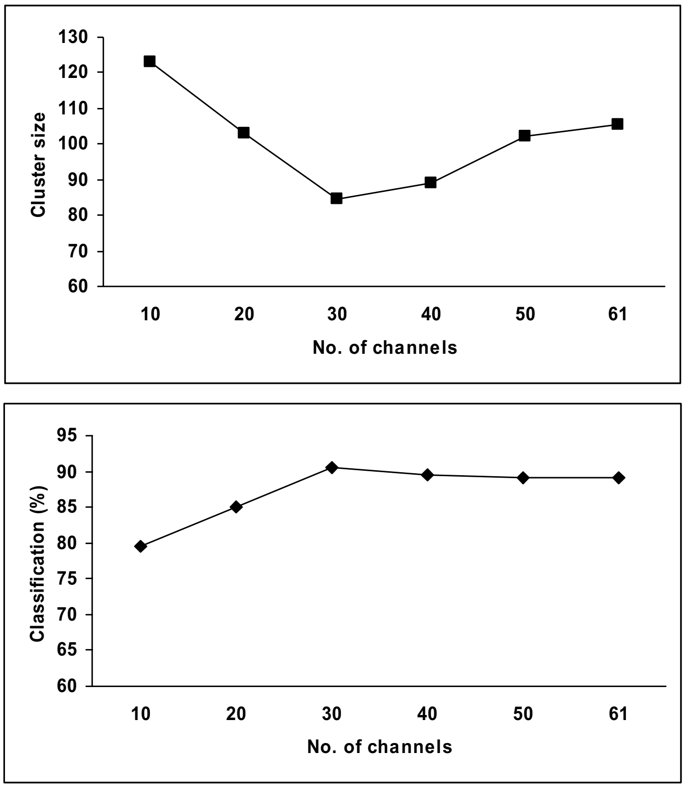

Fig. 7. Averaged SFA cluster size and classification performance (10–61 channels).

from all 61 channels to give the normalised gamma band power (NGBP):

$$NGBP \text{ (channel i)} = \frac{GBP(i)}{\sum_{k=1}^{61} GBP(k)}, \quad (5)$$

These NGBP values from 61 channels were concatenated into one feature array representing the particular VEP pattern.

2.4. Optimal channel selection

In this study, FDR was used to order the discriminating channels from the 61 VEP channels. The higher FDR values denote channels that were more discriminatory between the classes. The following relation shows the FDR function:

$$FDR_k = \sum_{i=1}^{N} \sum_{j=i+1}^{N} \frac{\left(\mu_k^i - \mu_k^j\right)^2}{\left(\sigma_k^i\right)^2 + \left(\sigma_k^j\right)^2}, \quad (6)$$

where N is the number of classes i.e. 40 subjects, μ_k^i is the mean of k^{th} feature for i^{th} class, σ_k^i is the standard deviation of the k^{th} feature for i^{th} class and i is varied from 1 to 40 subjects. Figure 4 shows the FDR values of the channels (from high to low).

2.5. Simplified fuzzy ARTMAP classifier

SFA [7] was chosen instead of MLP or the newer Support Vector Machine (SVM) classifiers due to its incremental supervised learning ability, i.e. more subjects could be added for identification without having to retrain the whole system again. Further, it is much faster to train than MLP or SVM. During supervised

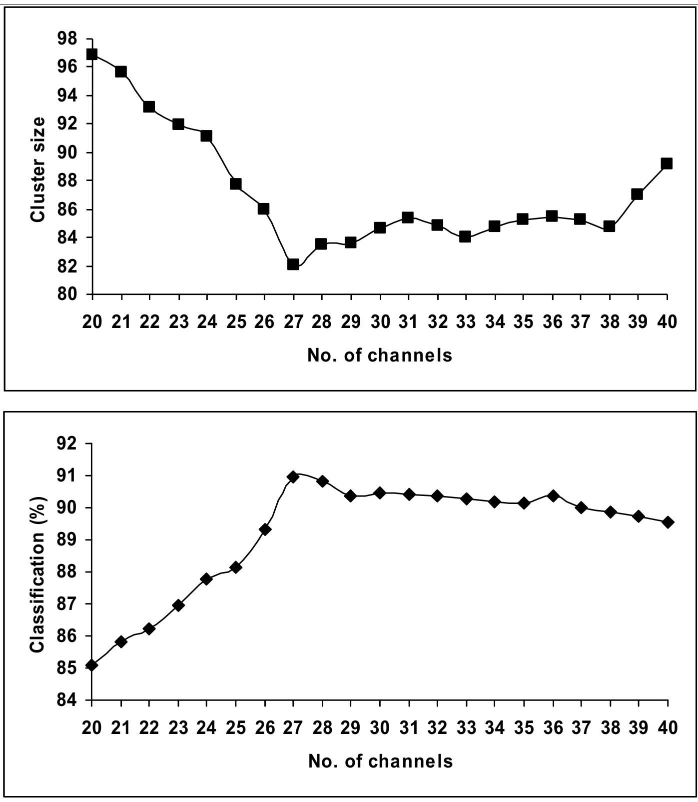

Fig. 8. Averaged SFA cluster size and classification performance (20–40 channels).

learning, Fuzzy ART receives a stream of input features representing the pattern that map to the output classes in the category layer. Vigilance parameter, ρ calibrates the minimum confidence that Fuzzy ART must have in an input vector in order for Fuzzy ART to accept that category, rather than search for a better one through an automatically controlled process of hypothesis testing. Lower values of ρ denote increased generalisation ability and lower category formations, while larger values of ρ induce more categories [1]. There is few difference between SFA and Fuzzy ARTMAP when the latter is used as a classifier. The only significant difference is that the SFA architecture consists of Fuzzy ART and Inter ART modules, while Fuzzy ARTMAP has the additional Fuzzy ART module, which is not suitable for classification. Figure 5 depicts the architecture of SFA as used here.

The input features at layer F_0 were the 61 NGBP values. Complement coding was performed (at layer F_1), which doubled the inputs to 122. Training was performed, which mapped F_1 nodes to F_2 nodes. The cluster size represents the number of F_2 nodes, which is variable depending on the training. Each F_2 node mapped to one of the 40 categories that represent the individuals.

During testing stage, inputs were fed into layer F_0 and the outputs were obtained in the range of [0, 1] in layer F_2. The maximum value of the outputs was the winner node and this node matched one of the 40 categories that represent the individuals. The matching node was then compared to determine if SFA predicted the correct individual or not.

In all the experiments, half of the available VEP patterns (i.e. 20 from each subject) were used for SFA training while the rest half were used for SFA testing. As such, 800 VEP patterns were used in training, while the rest 800 VEP patterns were used in testing. The selection of VEP signals for the training and testing datasets were conducted randomly and were fixed for the experiment. SFA was trained in the fast learning mode. Since the order of input patterns affects the classification performance, voting strategy as suggested in [1] from 20 simulations were used. Further details of SFA could be found from [1,7]. Figure 6 shows the sequential stages of the proposed methodology for this study.

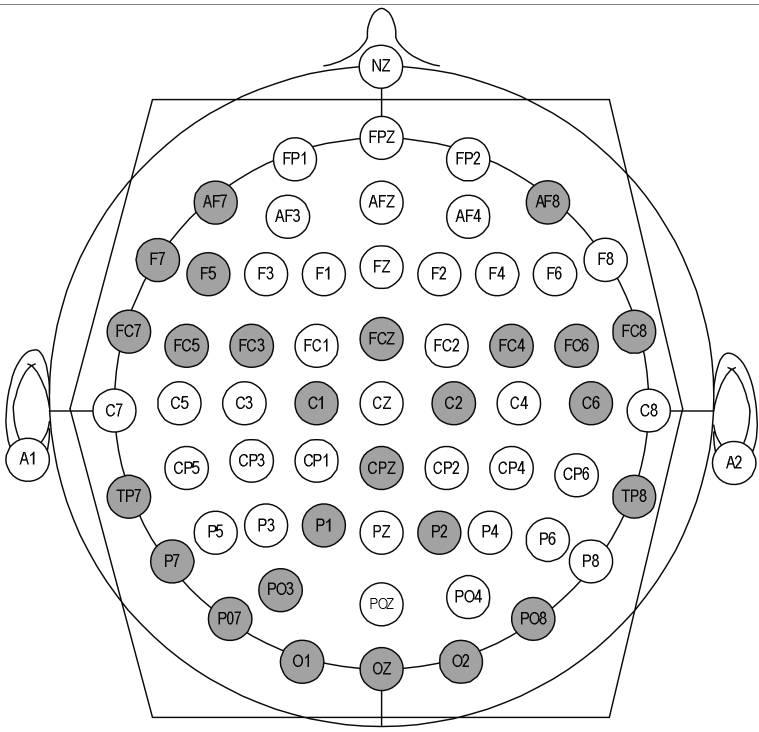

Fig. 9. 27 optimal channels.

3. Results and discussion

The results of SFA classifications are shown in Fig. 7 with different number of channels in increments of 10, i.e. 10, 20, 30, 40, 50 and 61 channels. These channels were selected as highest 10, highest 20, *etc* using the FDR values. Also shown in the figure are the results of Fuzzy ART cluster sizes. The SFA was run with vigilance parameter values ranging from 0.0 to 0.9 in steps of 0.1 but to save space, only the averaged results are shown in Fig. 7. From this figure, it could be noticed that the optimal channels are in the range of 20–40 channels. Therefore, additional SFA classifications were carried out using 20–40 channels, in steps of 1, i.e. the highest 20, highest 21, *etc* using the FDR values. These results are shown in Fig. 8.

It could be seen that the optimal channels were 27 channels, which gave the best classification performance. These channels are as depicted in Fig. 9. Ten channels were located on the right hemisphere, while 14 were on the left. Three were in midline. The locations are of some importance but a study of it is beyond the aim of this work.

From Table 1, which compares the average SFA performance between 27 optimal channels and 61 channels, it could be seen that 27 optimal channels improved the SFA classification percentage from 89.11% to 90.97%. The lower number of channels would reduce the hardware requirement. In addition, the use of optimal channels also reduced the Fuzzy ART cluster size, training and testing times. The reduction in average Fuzzy ART cluster size was from 105.61 to 82.05, which would be a further reduction in design complexity. The average training time for each pattern was reduced from 0.018 s to 0.015 s, while the average testing time was reduced from 0.002 s to 0.0009 s. We chose to compare the execution times, rather than the number of operations needed, which was done for convenience. Since the algorithms were run on the same code, using Matlab (Mathworks, Inc.), this still provided a fair comparison of the complexity of using the different number of channels.

It is known that low frequency parameters of VEP like P3 and N4 varies with time, stimulus type, etc. Though there is the possibility that GBP VEP might be variable over time but from the experimental results, we believe that VEP-based biometric is worth exploring further. Recent study by others [16] has shown that gamma band oscillations have some relationship with higher brain functions like feature binding ability during visual stimulus perception. The key point proposed in that study is that the gamma band oscillations do not represent information itself, but rather provides a temporal structure for correlations in the neurons that do encode specific information. Another study [5] speculated that the function of gamma band oscillations is to provide a reference clock to control the firing of the excitatory neurons. Therefore, it is our assumption that GBP VEP may not change drastically over time and therefore suitable for use in biometric applications.

4. Conclusion

In this paper, we have used FDR to select the optimal channels for Fuzzy ARTMAP classification of individuals using GBP of VEP signals. The optimal channels gave improved performance with lower design complexity and computational time. Though the classification improvement was small, it was achieved with lower training and testing times and lower network size (lower cluster size and lower number of channels). Our next challenge would be to improve the classification performance and to investigate the variability of GBP of VEP signals over longer periods of time.

Acknowledgement

We thank Prof. Henri Begleiter at the Neurodynamics Laboratory at the State University of New York Health Centre at Brooklyn, USA whose group recorded the raw VEP data and Mr. Paul Conlon, of Sasco Hill Research, USA for sending the data to us.

References

[1] G.A. Carpenter, S. Grossberg and J.H. Reynolds, A fuzzy ARTMAP nonparametric probability estimator for nonstationary pattern recognition problems, *IEEE Transactions on Neural Networks* **6**(6) (1995), 1330–1336.

[2] J. Daugman, Recognizing persons by their iris patterns, in: *Biometrics: Personal Identification In Networked Society*, A.K. Jain, R. Bolle and S. Pankanti, eds, Kluwer Academic, 1999.

[3] N. Duta, A.K. Jain and K.V. Mardia, Matching of palmprints, *Pattern Recognition Letters* **23**(4) (2002), 477–485.

[4] A.K. Jain, A. Ross and S. Pankanti, *A prototype hand geometry-based verification system*, Proceedings of 2nd International Conference on Audio and Video-Based Biometric Person Identification, March 22–24, 1999, 166–171.

[5] J.G.R. Jefferys, R.D. Traub and M.A. Whittington, Neuronal networks for induced "40 Hz" rhythms, *Trends in Neurosciences* **19** (1996), 202–208.

[6] I.T. Jolliffe, *Principal Component Analysis*, Springer-Verlag, 1986.

[7] T. Kasuba, Simplified Fuzzy ARTMAP, *AI Expert* **8**(11) (1993), 19–25.

[8] R. Palaniappan, Method for identifying individuals using VEP signals and neural network, *IEE Proceedings – Science, Measurement and Technology* **151**(1) (2004), 16–20.

[9] R. Palaniappan and K.V.R. Ravi, A new method to identify individuals using signals from the brain, Fourth International Conference on Information Communications and Signal Processing and Fourth Pacific-Rim Conference on Multimedia, Singapore, Dec. 15–18, (vol. 3), 2003, 1442–1445.

[10] R. Palaniappan, P. Raveendran, S. Nishida and N. Saiwaki, A new brain-computer interface design using fuzzy ARTMAP, *IEEE Transactions on Neural Systems and Rehabilitation Engineering* **10**(3) (2002), 140–148.

[11] S. Pankanti, R.M. Bolle and A. Jain, Biometrics: the future of identification, *IEEE Computer* (*Special Issue on Biometrics*) **33**(2) (2000), 46–49.

[12] S. Pankanti, S. Prabhakar and A.K. Jain, On the individuality of fingerprints, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **24**(8) (2002), 1010–1025.

[13] R.B. Paranjape, J. Mahovsky, L. Benedicenti and Z. Koles, The electroencephalogram as a biometric, *Proceedings of Canadian Conference on Electrical and Computer Engineering* **2** (2001), 1363–1366.

[14] M. Poulos, M. Rangoussi, V. Chrissikopoulos and A. Evangelou, Person identification based on parametric processing of the EEG, *Proceedings of the 6th IEEE International Conference on Electronics, Circuits, and Systems* **1** (1999), 283–286.

[15] Samal and P. Iyengar, Automatic recognition and analysis of human faces and facial expressions: a survey, *Pattern Recognition* **25**(1) (1992), 65–77.

[16] W. Singer and C.M. Gray, Visual feature integration and the temporal correlation hypothesis, *Annual Review of Neuroscience* **18** (1995), 555–586.

[17] J.G. Snodgrass and M. Vanderwart, A standardized set of 260 pictures: norms for name agreement, image agreement, familiarity, and visual complexity, *Journal of Experimental Psychology: Human Learning and Memory* **6**(2) (1980), 174–215.

[18] X.L. Zhang, H. Begleiter, B. Porjesz, W. Wang and A. Litke, Event related potentials during object recognition tasks, *Brain Research Bulletin* **38**(6) (1995), 531–538.