Wavelet Design by Means of Multi-Objective GAs for Motor Imagery EEG Analysis

Javier Asensio, Edgar Galvan, Ramaswamy Palaniappan, John Q. Gan

School of Computer Science and Electronic Engineering, University of Essex, United Kingdom
jasens@essex.ac.uk;edgar.galvan@gmail.com;palani,jqgan@essex.ac.uk

Abstract

Wavelet-based analysis has been broadly used in the study of brain computer interfaces(BCI), but in most cases these wavelet functions have not been designed taking into account the requirements of this field. In this study we propose a method to automatically generate wavelet-like functions by means of genetic algorithms. Results strongly indicate that it is possible to generate (evolve) wavelet functions that improve the classification accuracy compared to other well-known wavelets (e.g. Daubechies and Coiflets).

1 Introduction

Wavelets are wave-like oscillations used to extract relevant information from a given signal. They have useful features such as the time-frequency resolution, which make them suitable for electroencephalography (EEG) signals unlike the Fourier Transform (FT) that only gives information in the frequency domain [1] (although related methods like the short time FT provides time-frequency resolution). Several related approaches have already been used in the BCI research such as the continuous wavelet transform [2], the discrete wavelet transform(DWT) [3] and wavelet packages [4].

The main goal of this paper is to explore the possibility of automatically creating a wavelet that can be at least as good as other well-defined wavelets. Inspired by other works (e.g., [5, 6, 7]), we will explore the possibility to automatically evolve wavelets by means of Genetic Algorithms (GAs). It should be noted that the main difference between this work and the works mentioned previously is that we use GAs in order to evolve directly the filter coefficients for a one-way DWT.

This paper is organised as follows. In the next section we describe our approach. In Section 3, we present and discuss the results. In Section 4 we draw some conclusions.

2 Approach

2.1 Data Description

Each of the subjects, three in total, sat in a comfortable arm chair, one metre away from a 19" screen. EEG signals were recorded from five channels using bipolar electrode positions with respect to the international 10-20 system: FC3-PC3, FC1-PC1, Cz-Pz, FC2-PC2, FC4-PC4. The recording was made with a 16-channel EEG amplifier from g.Tech and sampled at 250 Hz. Each subject went through four sessions and 30 trials were recorded in each one. In each trial an arrow was shown from $t = 3s$ to $8s$, with the arrow pointing left if the user was meant to imagine left hand movement, right for right hand and down for feet. Only two classes are taken into account in this study, left hand and feet, making a total of 80 trials for each subject.

2.2 Wavelet Processing

To assess the performance of the wavelet function obtained by means of GAs, first it is necessary to perform an equivalent experiment using a standard wavelet. For this purpose, we selected Daubuchies and Coiflets wavelets as they have been widely used in other studies and their accuracy has been well assessed [4] [3] [8]. For every cued instant in the dataset, a second of the signal from a specific point is transformed using the DWT. This procedure is repeated for each of the five channels, and all the outputs are joined and labeled with its class to build an input pattern. Then a sliding window is applied, moving a 1/8 of a second ahead to take the next second of the signal to build the next pattern and so on, until the cued data is completely consumed.

A multi-frequency decomposition (MFD) is performed to obtain the most interesting coefficients from the wavelet transform. As the original data was acquired with a sampling frequency of 250 Hz the wavelet decomposition is done down to the 5^{th} level, by discarding D0 and D1 the coefficients for A5(0-3.90Hz), D5(3.90-7.81Hz), D4(7.81-15.62Hz), D3(15.62-31.25Hz) and D2(31.25-62.5Hz) are obtained. We selected this range as it covers from delta to gamma bands, letting the GA decide which elements are more important.

Dimensionality of the output from the MFDs is too high for classification purpose. Thus, Davis-Bouldin Index (DBI) [9] is used to measure the features' separability and then the best 25 are selected. The patterns obtained are classified using Fisher's linear discriminant (LDA) [10]. This model was used due to its simplicity and efficiency, it was also shown during BCI competitions 2003 and 2005 that LDA performs as well as Neural Networks or Support Vector Machines in terms of classification accuracy [11].

The classification process uses a five-fold cross-validation model. Also, the classification rates are calculated using trial by trial validation,i.e. a trial is classified as c if the majority of its samples are classified as c.

2.3 Evolving Wavelet Filters

Often, wavelets must fulfil some constraints such as the admissibility condition [1], so that the inverse transform can be calculated. In our case, there is no requirement to have a proper wavelet as we are not interested in applying the inverse DWT to reconstruct the original signal.

So, the aim is to generate a wave-like oscillation which increases the separability of the data when the wavelet transform is applied, but ignoring those restrictions that make the operation reversible. This, in consequence, makes the design of the fitness function used in our GA (described later in this section) easier than other scenarios (where the wavelet inverse transform is required).

The discrete wavelet transform can be described as a filtering and down-sampling waterfall process, the approximation and detail function for each level within the down sampling process is defined by a filter of length L and is given by $\phi(x) = \sum_n^L h_0(n)\phi(2x-n)$ and $\psi(x) = \sum_n^L h_1(n)\phi(2x-n)$, where h_0 and h_1 are the high-pass and low-pass filters, respectively.

2.3.1 Multi-Objective Genetic Algorithm

The aim of the proposed study, as stated previously, is to automatically evolve wavelets. Thus, it is necessary to evolve the coefficients of both filters h_0 and h_1. The filter length varies depending on the wavelet order. Also, these coefficients are defined by real numbers. The length of each chromosome in the population set is 20. Each chromosome is composed by two parts: the first half is designed to obtain the best value for h_0 and the second part aims at finding the best value for h_1. The original population is generated with random filter coefficient values in the range $[-1, 1]$.

The experiments were conducted using GA with tournament selection (size = 7), bit mutation (0.01) and two-point crossover (0.7). To obtain more reliable results, we performed 20 independent runs. Runs were stopped when the maximum number of generations (120) was reached.

As a first approach, we used a fairly simple fitness function based on the individual accuracy (raw fitness). A wavelet function is generated using the coefficients defined by each individual in the population; this function is then applied to every signal from the data set. Next, the fitness value is calculated as the outcome from the FDA. To assess a correct behaviour in the generated wavelets, the data set is divided into two different subsets (of 40 trials each): the first is used during the GA execution and the second is used to validate the performance of this newly generated function against the Daubechies and Coiflets wavelets.

We also tested a more robust approach by considering Multi-Objective (MO) GAs, in specific using a well proved algorithm called Non-Dominated Sorting GA [12]. Thus, the fitness function is formed by two elements: (a) the classification accuracy from FDA, as discussed earlier, and (b) the minimisation of the standard deviation in the classification accuracy among the different classes. For the latter objective, a weight of 0.6 has been applied in order to decrease its importance in the selection of the 'best so far' solution. In the following section, we present and discuss our findings.

Table 1: GA Mean result. SF stands for Single Fitness and MO for Multi-objective

Subject	SF GA	SF Validation	MO GA	MO Validation
1	0.612 +/- 0.046	0.624 +/- 0.047	0.696 +/- 0.045	0.658 +/- 0.045
1 best	0.725 +/- 0.0447	0.680 +/- 0.057	0.762 +/- 0.0406	0.757 +/- 0.037
2	0.597 +/- 0.045	0.664 +/- 0.043	0.709 +/- 0.040	0.695 +/- 0.038
2 best	0.762 +/- 0.042	0.817 +/- 0.031	0.817 +/- 0.038	0.795 +/- 0.031
3	0.786 +/- 0.035	0.698 +/- 0.041	0.767 +/- 0.038	0.725 +/- 0.042
3 best	0.892 +/- 0.029	0.782 +/- 0.041	0.852 +/- 0.030	0.785 +/- 0.051

Table 2: Benchmark comparison

Subject	Coif5	Db10	PSD	Generated Wavelet
1	0.688 +/- 0.116	0.690 +/- 0.130	0.730 +/- 0.120	**0.752 +/- 0.037**
2	0.667 +/- 0.162	0.633 +/- 0.110	**0.827 +/- 0.080**	0.690 +/- 0.130
3	0.665 +/- 0.140	0.690 +/- 0.150	0.697 +/- 0.147	**0.767 +/- 0.130**

3 Results and Discussions

The GA was run for 20 different rounds applying both single and multi-objective. The results are shown in Table 1. This table shows the mean of ten executions for the best individual in those rounds. The GA column values are the averaged accuracy of the best individuals against the set used for evolving the wavelets, whereas the validation column values show the averaged accuracy of the best individuals for the validation set. For each subject the results for the best wavelet are shown. In every case, the MO produced better results compared to the single fitness approach. This indicates that encouraging the GA to assure a balanced rate among class through classification accuracy leads to obtain a more robust wavelet against unseen trials. During the early experiments, an over-fitting problem arose and this was addressed by selecting only a random subset of every test fold during the evaluation. In most cases the population converged before the 120th generation and was stable for at least 20 generations, therefore we can assume that the number of iterations was sufficient to allow the algorithm to evolve.

In Table 2 a comparison among the generated wavelets and other techniques is displayed. The columns *Db10* and *Coif5* shows the result for the process described in Section 2.2. The column *PSD* is the results of applying *Power Spectral Density* from 0 Hz to 62 Hz where the mean values for bands with width of 2 Hz each are computed, ending with 155 features per pattern. The DBI is applied to select the most 25 discriminating features. The column named *Generated Wavelet* shows the result of the best wavelet obtained by the GA and applied to the validation set, therefore the decision on which wavelet to use is not biased by the results against the validation set as shown in Table 1. Notice that the results shown in Table 2 are the average accuracies for the validation set from ten different runs where the data was randomly shuffled.

The results obtained show that the generated wavelet always perform better than the Daubechies and Coiflets. When compared to PSD the generated wavelet performed better for subjects 1 and 3 but worse in case of subject 2.

Frequency response study of the best generated filters in the Table 1 shows that all of them are stable as indicated by their phase responses. The Daubechies wavelet filters are designed in such a way that they are high-pass and low-pass filters. However, the behaviour found in the generated wavelets is different. All the filters obtained show a multi-band behaviour with the pass bands different for each h_0 and h_1. Due to the analytic complexity of the DWT, to state what each multi-band filtering implies is out of the scope for this study.

It is clear that the DBI feature selection strongly affects the evolutionary process but it allows us to draw out the most important frequency bands and channels involved in the classification. Studying the selected features, we can observe that for every subject the importance of each bipolar electrode varies. E.g. for the first subject, 5% of the selected features come from FC4-PC4, whereas for the third subject, this ratio was up to 35%.

If we focus our attention on the frequency bands, we find that for any user the wavelet decomposition level D2 corresponding to the 31.25 to 62.5 Hz frequency range occurs in 40% of the features selected by DBI (although beta band features tend to rank higher in the DBI). This result supports the outcome presented in [13] where it is stated that gamma band contains highly useful information for motor imagery classification. This behaviour is consistent among the generated functions but not in the Daubechies wavelet where every different subject presents a different distribution in the gamma range.

4 Conclusions

The results obtained in this paper show that the evolved wavelet performs better than Daubechies and Coiflets wavelets. This implies that using out-of-the-box wavelets might not be the best approach when dealing with EEG data. Thus, other evolved functions could improve the performance in wavelet based solutions, as shown in this study. It should be noted that one disadvantage of this approach is its efficiency, as it generates a different wavelet for each user. Future work will address this issue.

References

[1] P. S Addison. *The Illustrated Wavelet Transform Handbook.* Institute of Physics Publ., 2002.

[2] V. Bostanov. BCI competition 2003-data sets ib and IIb: feature extraction from event-related brain potentials with the continuous wavelet transform and the t-value scalogram. *Biomedical Engineering, IEEE Transactions on*, 51(6):1057–1061, 2004.

[3] R. Scherer, S. P Zanos, K. J Miller, R. P.N Rao, and J. G Ojemann. Classification of contralateral and ipsilateral finger movements for electrocorticographic brain-computer interfaces. *Neurosurgical Focus*, 27(1):12–4967, 2009.

[4] J. Sherwood and R. Derakhshani. On classifiability of wavelet features for EEG-based brain-computer interfaces. In *Neural Networks, 2009. IJCNN 2009. International Joint Conference on*, pages 2895–2902. IEEE, 2009.

[5] M. M. Lankhorst, M. Van Der Laan, and W. A. Halang. Wavelet-based signal approximation with genetic algorithms. *Systems Analysis Modelling Simulation*, 43(11):1503–1528, 2003.

[6] E. Jones, P. Runkle, N. Dasgupta, L. Couchman, and L. Carin. Genetic algorithm wavelet design for signal classification. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pages 890–895, 2001.

[7] D. Farina, O. F. Do Nascimento, M. F Lucas, and C. Doncarli. Optimization of wavelets for classification of movement-related cortical potentials generated by variation of force-related parameters. *Journal of Neuroscience Methods*, 162(1-2):357–363, 2007.

[8] B. G. Xu and A. G. Song. Pattern recognition of motor imagery EEG using wavelet transform. *J. Biomedical Science and Engineering*, 1:64–67, 2008.

[9] J. C Bezdek and N. R Pal. Some new indexes of cluster validity. *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, 28(3):301–315, 1998.

[10] R. A Fisher. The use of multiple measurements in taxonomic problems. *Annals of Eugenics*, 7:179–188, 1936.

[11] B. Blankertz, K. R Muller, D. J Krusienski, G. Schalk, J. R Wolpaw, A. Schlogl, G. Pfurtscheller, J. R Millan, M. Schroder, and N. Birbaumer. The BCI competition III: validating alternative approaches to actual BCI problems. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 14(2):153–159, 2006.

[12] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan. A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Transactions on Evolutionary Computation*, 6(2):182–197, 2002.

[13] J. Ginter Jr, K. J. Blinowska, M. Kaminski, P. J. Durka, G. Pfurtscheller, and C. Neuper. Propagation of EEG activity in the beta and gamma band during movement imagery in humans. *Methods Inf Med*, 44(1):106–113, 2005.