35th Annual International Conference of the IEEE EMBS
Osaka, Japan, 3 - 7 July, 2013

Optimal sampling frequency in wavelet-based signal feature extraction using particle swarm optimization

C. Guarnizo[1], A. A. Orozco[2] and M. A. Alvarez[2]

***Abstract*— A methodology for optimum sampling frequency selection for wavelet feature extraction is presented. We show that classification accuracy is enhanced by adequately selecting the parameters: number of decomposition levels, wavelet function and sampling rate. A novel approach for selecting the parameters based on particle swarm optimization (PSO) is presented. Experimental results conducted on two different datasets with support vector machine (SVM) classifiers confirm the superiority and advantages of the proposed method. It is shown empirically that the proposed method outperforms significantly the existing methods in terms of accuracy rate.**

I. INTRODUCTION

Discrete wavelet transform (DWT) has been widely used for feature extraction from biomedical signals, e.g., electrocardiograms (ECG) [1], phonocardiograms (PCG) [2], electroencephalograms (EEG) [3], microelectrode recordings (MER) [4], among others. Ranging from applications in denoising, compression and classification. In several works have been found that classification accuracy depends on the wavelet function [5], the decomposition levels [6] and the sub-bands selection [7].

Wavelet function can be selected from a list of previously designed wavelet functions with different orders, such as, Daubechies, Symlet or Coiflet [5]. Another approach is to customized the wavelet function to the problem at hand, by means of stochastic optimization algorithms [4], [1].

On the other hand, the decomposition levels are related to the number of features extracted from wavelet transform and sub-band division of spectral content. Finally, features extracted from the wavelet transform can be optimized by a process of feature selection to avoid redundant or useless features. An adequate selection of these parameters must be performed to ensure a high classification accuracy. For all possible combinations of these parameters a different feature space is obtained. Also, the feature space depends on the kind of measures used as features, like, statistical moments [8], information measures [9], among others. Features extracted from wavelet transform are highly problem-dependent.

In this paper, we show how the sampling rate in the DWT feature extraction stage, greatly affects signal classification performance. Thus, sampling rate is included in the DWT feature extraction for signal classification. Sampling rate has been previously used to enhance signal denoising and compression. The former is presented in [10], an optimal sampling rate is selected to improve the reconstruction of Partial Discharge signals. The latter is presented in [11], optimal sampling frequency is selected for different examples of signal compression.

Additionally to sampling rate optimization, the number of decomposition levels and the wavelet function are optimized by particle swarm optimization (PSO). PSO has successfully been used in wavelet transform parameter selection and wavelet function customization [1].

This paper is organized as follows: Section II outlines brief review of theoretical methods. The proposed method is presented in section III. Finally, in section IV, results and final discussion are given.

II. METHODS

A. Discrete Wavelet Transform

Discrete wavelet transform (DWT) is calculated performing the discrete convolution between the signal (x) and two filters, one filter is a high pass discrete function g_n known as mother wavelet and the other one is a low pass filter h_n known as father wavelet. Coefficients calculated from each convolution are downsampled [12], thus, coefficients contain half frequency content from its predecessor. Another convolution between coefficients from the original signal and father wavelet is used to obtain a narrow frequency representation on the signal (see Fig. 1). Where d_k^l and c_k^l represent

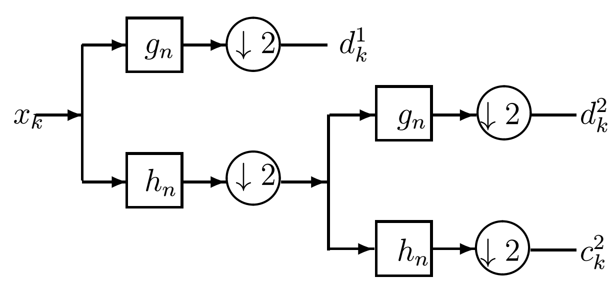


Fig. 1. Discrete wavelet transform with 2 decomposition levels

the detail and approximation coefficients obtained from the convolution of the signal with filters g_n and h_n, respectively. For any wavelet basis with 2 decomposition levels (as shown in Fig. 1) the partition of frequency content is given by the following sub-bands $\{[0\ fs/8], [fs/8\ fs/4], [fs/4\ fs/2]\}$, where fs is the sampling rate. Filter response for different orders of Daubechies wavelet is shown in Fig. 2, all wavelet functions have the same cut off frequency. The slope around the cut off frequency is proportional to wavelet order, thus,

[1]C. Guarnizo is a Doctoral student at Faculty of Electrical Engineering, Universidad Tecnológica de Pereira, Colombia cdguarnizo@gmail.com
[2]A. A. Orozco and M. A. Alvarez are with the Department of Electrical Engineering, Universidad Tecnológica de Pereira, Colombia aaog,malvarez at utp.edu.co

a higher order allows a better separation between the sub-bands.

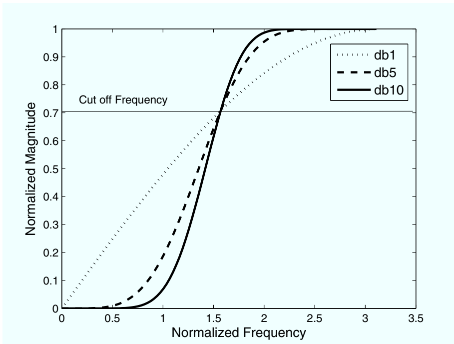


Fig. 2. Filter response for different orders of Daubechies wavelet.

B. Feature Extraction

Feature extraction from discrete wavelet transform is performed by taking measures from the coefficients of each sub-band. In this paper the standard deviation was used, it has been successfully used as feature in [6], [13]. For L decomposition levels with $l = 1, \ldots, L$ the feature extraction process is calculated by (1) and (2).

$$STD_l = \sqrt{\frac{1}{N-1}\sum_{k=1}^{N}\left(d_k^l - \mu^l\right)^2} \qquad (1)$$

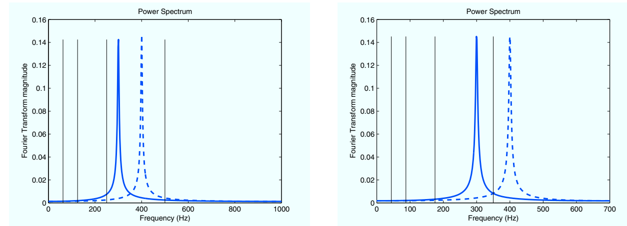
$$STD_{L+1} = \sqrt{\frac{1}{N-1}\sum_{k=1}^{N}\left(c_k^L - \mu^{L+1}\right)^2} \qquad (2)$$

where μ^l and μ^{L+1} are the mean value of d_k^l and c_k^L, respectively. N is the total number of wavelet coefficients at the analyzed subband.

C. Re-sampling effect

In [11] an optimal signal compression approach using discrete wavelet transform is presented. Signal compression value is enhanced by varying the sampling frequency in the range $[1, 2) \cdot fs$. The main aim is to adapt the signal energy distribution to wavelet tree decomposition. We believe that selecting an adequate value for the sampling rate will improve the classification rate in pattern recognition tasks. An example is shown in Fig. 3(a), where the mean energy distribution of two different classes lay in the same sub-band. If we change the sampling rate from 2KHz to 1.4KHz we are able to separate the classes by using the energy from sub-bands, as shown in Fig. 3(b).

To change the original sampling frequency to another, polyphase re-sampling method based on fractional decimation [14] was adopted. Fractional decimation approach is shown in Fig. 4. Zero values are added in the input sequence x_k by the expander P. The expanded version of the input is fed to the interpolation filter h_n. Then, the interpolation filter averages the missing samples. Finally, output from the filter is downsampled by a factor of Q to produce the output signal y_k. The discrete filter h_k is low pass with passband edge at $\min(\pi/P, \pi/Q)$. Rational factor (RF) between P and Q is defined in (3).



(a) (b)

Fig. 3. Mean energy distribution in frequency band along with 4-decomposition level wavelet bank filter (a) before re-sampling and (b) after re-sampling .

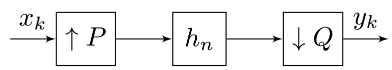


Fig. 4. Re-sampling by Fractional decimation scheme

$$RF = \frac{P}{Q} \qquad (3)$$

In this work, RF is varied by fixing the Q value at 100 and P takes values in the integer range $[1, 190]$. Thus, sampling frequency variation is defined in the range $[0.01, 1.9] \cdot fs$.

D. Particle Swarm Optimization

Like most stochastic optimization methods, PSO is a population based algorithm, one member of the population is known as particle. Each particle has a position (p_m) and a velocity (v_m). Particle's position indicates the parameters that are going to be optimized, thus, $p_m \in R^d$, where d is the number of parameters. The aim of velocity v_m is to update the position of each particle, with the aim of trying to reach a global optimum [15]. Global optimum is achieved from the combination of the general best particle gb (best position from population) and the best local position pb_m (best position historically reached by particle p_m). For particle p_m at iteration i, the velocity is calculated as

$$v_m^i = w \cdot v_m^{i-1} + c_1 \cdot r_1 \cdot (pb_m^i - p_m^i) + c_2 \cdot r_2 \cdot (gb^i - p_m^i) \qquad (4)$$

where w is inertia weight used as a tradeoff between global and local exploration capabilities of the swarm [16]. Large values of w allow better global exploration, while small values lead to a fine search in the solution space. r_1 and r_2 are random variables drawn from a uniform distribution in the range [0,1], they provide a stochastic weighting of the different components participating in the particle velocity definition [1]. c_1 and c_2 are constants that regulate the relative velocities with respect to the best local and global positions, respectively. Then particle p_m is updated for next iteration $i + 1$ with

$$p_m^{i+1} = p_m^i + v_m^i \qquad (5)$$

A more detailed explanation of the application of PSO for wavelet transform parameter selection is given in the following section.

III. THE PROPOSED METHOD

As discussed in section II-A, different parameters from wavelet transform require to be optimized. However, there is no rule to decide what is the best value for each parameter. Similar problems have been addressed by empirically comparing results from different values of each parameter. Therefore, we propose a wavelet parameter selection method based on PSO framework, where wavelet function (WF), rational factor (RF) and number of decomposition level (DL) were optimized according to the classification accuracy. In this sense, the coordinates of the particles of the swarm encode RF, WF and DL parameters $\{\text{RF,WF,DL}\}$. RF was computed by varying the value of P from (3) in the range $[1, 190]$, with $P \in \mathcal{Z}$. WB was selected from a list D of conventional wavelet functions. Dictionary D was built with orthonormal wavelet functions as shown in (6).

$$D = \{\text{Db1}, \ldots, \text{Db10}, \text{Coif1}, \ldots, \text{Coif5}, \ldots, \text{Sym10}\} \quad (6)$$

D contains up to 24 different wavelet functions. Meanwhile, parameter DL was set to vary in the integer range $[1, 10]$, thus, a maximum of 10 decomposition levels can be performed. Concerning the fitness function, we use the classification accuracy of a SVM classifier achieved by cross-validation (CV) on the training set [1]. The main steps of the proposed wavelet parameter selection method are described in the following:

1) **Initialization:**
   - Initialize swarm population by generating for each particle p_m a position vector $\{\text{RF,WB,DL}\}$ of integer random values uniformly distributed. Velocity vector of each particle is initially set to zero.
2) **Particle evaluation:**
   - Compute the fitness function value for each particle p_m through the following steps:
     - Resample each database signal with rational factor RF ($p_m^i(1)$) as in (3).
     - Apply DWT with wavelet function WF ($p_m^i(2)$) and number of decomposition level DL ($p_m^i(3)$) to each resampled signal;
     - extract $DL + 1$ features from wavelet coefficients using (1) and (2);
     - train an SVM classifier by feeding it with the generated wavelet features. Set the fitness F_m^i function value of each particle by computing its cross-validation accuracy.
     - Store the best position of each particle in the best local position vector. From best local position vector, save the particle with largest fitness function value as the best global position.
3) **Updating particle positions:**
   - Update each velocity vector using (4);
   - then, update each particle position as in (5).
4) **Convergence Check:** The algorithm is stopped if the number of generations has reach its maximum, or if the best global position does not vary significantly.

IV. EXPERIMENTAL RESULTS

In this section, we present accuracy values using the proposed method on two different datasets. Accuracy values are calculated from the mean of the result of a SVM classifier (Guassian kernel) trained using a 5-fold cross validation scheme. When cross validation is performed, Gaussian kernel parameters C and γ are optimized using fixed ranges $[2^{-5}, \ldots, 2^{15}]$ and $[2^{-15}, \ldots, 8]$, respectively. The proposed method is performed on EEG and MER data sets. The former is used to diagnose seizure events, while the latter is used to recognize Subthalamic Nucleus (STN) for deep brain stimulation applications.

A. Results in EEG analysis

The database used here is described in detail in [17]. Complete data set consists of 5 groups (A-E), where each group contains 100 single-channel registers. Data were digitalized using 12 bit resolution and sampled at 173.61 Hz. Sets A and B consist of segments recorded from five healthy subjects with open eyes and closed eyes, respectively. Sets C, D and E consist of segments recorded from five patients whom achieved complete seizure control of epileptogenic zone. Segments in set C were recorded from the epileptogenic zone, and those in set D from the hippocampal formation of the opposite hemisphere of the brain. While set E contained seizure activity, sets C and D contained only seizure free intervals. In this work, five different classification problems are created from the above dataset in order to compare the performance of our method in different scenarios. Table I shows a brief description of the five classification problems. The proposed method was used to find an adequate value of parameters involved in wavelet transform feature extraction. A comparison result is shown in Table II. From the results shown in Table II, it can be concluded that: for A and E classification problem, all EEGs segments in the test data are correctly classified. Additionally, a reduction of sampling rate value can be included. In all classification problems a better accuracy is achieved by the proposed

TABLE I

EEG CLASSIFICATION PROBLEMS

Classification Problem	Classes	EEG segments
A,E	Normal (A)	100
	Seizure (E)	100
(A,C,D),E	Non-seizure (A,C,D)	300
	Seizure(E)	100
(A,B,C,D),E	Non-seizure (A,B,C,D)	400
	Seizure (E)	100
(AB),(CD),E	Surface Normal (AB)	200
	Intracranial Seizure Free (CD)	200
	Seizure (E)	100
A,B,C,D,E	Dataset description	100/class

TABLE II

EEG CLASSIFICATION ACCURACY COMPARISON FOR ORIGINAL AND OPTIMIZED WAVELET TRANSFORM

Classification Problem	RF	DL	WF	Accuracy (%)
A,E	1	1	Db1	100.00
	≥ 0.25	**1**	**Db1**	**100.00**
(A,C,D),E	1	2	Db1	98.50
	0.35	**1**	**Db1**	**99.00**
(A,B,C,D),E	1	4	Db7	98.50
	0.22	**1**	**Sym10**	**99.00**
(AB),(CD),E	1	4	Coif5	98.00
	0.81	**6**	**Db7**	**98.20**
A,B,C,D,E	1	6	sym6	85.40
	0.35	**5**	**Sym10**	**86.00**

TABLE III

MER CLASSIFICATION ACCURACY COMPARISON FOR ORIGINAL AND OPTIMIZED WAVELET TRANSFORM

RF	DL	WF	Accuracy (%)
1	4	Coif1	85.48
0.34	**3**	**Sym10**	**87.10**

method. Also, benefits such as feature space and sampling frequency reduction are obtained, except for (AB),(CD),E classification problem, where the number of decomposition level is increased to obtain a slightly higher accuracy value.

B. Results in MER analysis

Dataset for analysis corresponds to five interventions carried out locally in the city of Pereira. All the subjects gave their informed consent allowing the use of the neural signals recorded to research. The acquisition equipment used is the ISIS MER of Inomed, neural signals were labeled by two specialists in neurosurgery and neurophysiology; the sampling rate was 24 kHz and 16-bit resolution. There are 160 neural signals divide in two groups, 80 STN signals and 80 non-STN signals. The proposed method is applied in MER classification problem, results are shown in Table III. From the results shown in Table II, we conclude that accuracy rate value is increased by optimizing sampling rate. While accuracy value is increased, the number of decomposition level is reduced, this, feature space is less complex for the classifier. Sampling rate is changed from 24kHz to 8.16kHz, also reducing the requirement of a high rate data acquisition system for signal digitalization.

V. CONCLUSIONS

In this paper, we proposed a novel wavelet parameter optimization procedure based on PSO. Sampling rate was optimized with the number of decomposition levels and the wavelet function, to adjust wavelet feature extraction in discrimination enhancement. Discrimination capability was measured through an empirical estimate of the SVM classifier accuracy. The experimental results show that it achieves better classification accuracies compared to different wavelet functions without resampling. Its main drawback is that the method tends to select high order wavelet functions, because they have better frequency separation capabilities than low order wavelet functions.

ACKNOWLEDGMENT

This research is developed under the projects: "Desarrollo de un sistema automático de mapeo cerebral y monitoreo intraoperatorio cortical y profundo: aplicación neurocirugía", and "Desarrollo de un sistema efectivo y apropiado de estimación del volumen de tejido activo cerebral para el mejoramiento de los resultados terapéuticos en pacientes con enfermedad de Parkinson intervenidos quirúrgicamente", financed by Colciencias with codes 111045426008 and 111056934461, respectively.

REFERENCES

[1] A. Daamouche, L. Hamami, N. Alajlan, and F. Melgani, "A wavelet optimization approach for ecg signal classification," *Biomedical Signal Processing and Control*, vol. 7, no. 4, pp. 342–349, 2012.

[2] L. H. Cherif, S. Debbal, and F.Bereksi-Reguig, "Choice of the wavelet analyzing in the phonocardiogram signal analysis using the discrete and the packet wavelet transform," *Expert Systems with Applications*, vol. 37, pp. 913–918, 2010.

[3] L. Guo, D. Rivero, J. Dorado, J. R. Rabu, and A. Pazos, "Automatic epileptic seizure detection in eegs based on line length feature and artificial neural networks," *Journal of Neuroscience Methods*, vol. 191, pp. 101–109, 2010.

[4] R. Pinzon-Morales, A. Orozco-Gutierrez, and G. Castellanos-Dominguez, "Novel signal-dependent filter bank method for identification of multiple basal ganglia nuclei in parkinsonian patients," *Journal of Neural Engineering*, vol. 8, p. 036026, 2011.

[5] J. Rafiee, M. Rafiee, N. Prause, and M. Schoen, "Wavelet basis functions in biomedical signal processing," *Expert Systems with Applications*, vol. 38, pp. 6190–6201, 2011.

[6] I. Guler and E. Ubeyli, "Adaptive neuro-fuzzy inference system for classification of eeg signals using wavelet coefficients," *Journal of Neuroscience Methods*, vol. 148, no. 2, pp. 113–121, October 2005. [Online]. Available: http://dx.doi.org/10.1016/j.jneumeth.2005.04.013

[7] H. Ocak, "Optimal classification of epileptic seizures in eeg using wavelet analysis and genetic algorithm," *Signal Processing*, vol. 88, pp. 1858–1867, 2008.

[8] R. Carreno and M. I. Vuskovic, "Wavelet transform moments for feature extraction from temporal signals," in *2nd International Conference in Control, Automation and Robotics (ICINCO 2005)*, 2005.

[9] M. Murugappana, N. Ramachandran, and Y. Sazali, "Classification of human emotion from eeg using discrete wavelet transform," *Journal of Biomedical Science and Engineering*, vol. 3, pp. 309–396, 2010.

[10] X. Zhou, C. Zhou, B. Stewart, and A. Nesbitt, "Optimal sampling rate for wavelet-based denoising in pd measurement," in *Conference Record of the 2004 IEEE International Symposium on Electrical Insulation*, 2004.

[11] R. L. Kirlin and R. M. Dizaji, "Optimum sampling frequency in wavelet based signal compression," in *Proceedings of the 2000 Canadian Conference on Electrical and Computer Engineering*, vol. 1, 2000, pp. 327–330.

[12] S. Mallat, *A Wavelet Tour of Signal Processing, Second Edition (Wavelet Analysis & Its Applications)*. Academic Press, September 1999.

[13] E. D. Übeyli, "Combined neural network model employing wavelet coefficients for eeg signals classification," *Digit. Signal Process.*, vol. 19, no. 2, pp. 297–308, 2009.

[14] P. P. Vaidyanathan, "Multirate digital filters, filter banks, polyphase networks, and applications: A tutorial," *Proceedings of the IEEE*, vol. 78, no. 1, pp. 56–93, 1990.

[15] H. J. Escalante, M. Montes, and L. E. Sucar, "Particle swarm model selection," *Journal of Machine Learning Research*, vol. 10, pp. 405–440, 2009.

[16] F. Melgani and Y. Bazi, "Classification of electrocardiogram signals with support vector machines and particle swarm optimization," *IEEE Transactions of Information Technology in Biomedicine*, vol. 12, pp. 667–677, 2008.

[17] R. G. Andrzejak, K. Lehnertz, F. Mormann, C. Rieke, P. David, and C. E. Elger, "Indications of nonlinear deterministic and finite dimensional structures in time series of brain electrical activity: Dependence on recording region and brain state," *Physical Review E*, vol. 64, p. 061907, 2001.