ORIGINAL ARTICLE

Year : 2016
Volume : 3
Issue Number : 2
Doi Number : 10.5455/JNBS.1454666348

Article history:
Received 05 February 2016
Received in revised form 15 February 2016
Accepted 02 May 2016

COMPARISON OF WAVELET FAMILIES FOR MENTAL TASK CLASSIFICATION

ZİHİNSEL GÖREV SINIFLANDIRMA İÇİN DALGACIK DÖNÜŞÜMÜ FONKSİYONLARININ KARŞILAŞTIRMASI

Çağlar Uyulan*[1], Türker Tekin Ergüzel[2]

Abstract

Wavelet theory is a widely used feature extraction method for raw electroencephalogram (EEG) signal processing. The nature of the EEG signal is non-stationary, therefore applying wavelet transform on EEG signals is a valuable process for extraction promising features. On the other hand, determining the proper wavelet family is a challenging step to get the best fitted features for high classification accuracy. In this paper, therefore, we focused on a comparative study of different Discrete Wavelet Transform (DWT) methods to find the most convenient wavelet function of wavelet families for a non-stationary EEG signal analysis to be used to classify mental tasks. For the classification process, four different mental tasks were selected to and we grouped each with another one to set dual tasked sets including all possible combinations. Feature extraction steps are performed using wavelet functions haar, coiflets (order 1), biorthogonal (order 6.8), reverse biorthogonal (order 6.8), daubechies (order 2) and, daubechies (order 4). Later, a specific feature reduction formula is applied to the extracted feature vector. Generated feature vector is then split into train and test data before the classification. Artificial neural network was used for classification of the extracted feature sets. From the result of the repeated analysis for each DWT methods, Coiflets performed relatively better compared to other wavelet families.

Keywords: Discrete Wavelet Transform, mental task classification, coiflet wavelet, daubechies wavelet, wavelet families

Özet

Dalgacık dönüşümü, ham EEG (elektroensefalografi) verilerinden öznitelik çıkartma yöntemi olarak yaygın şekilde kullanılmaktadır. EEG sinyalleri doğası gereği durağan değildir, dolayısıyla dalgacık dönüşümü, sınıflandırma performansına katkıda bulunacak özniteliklerin çıkartılması sürecinde oldukça etkili bir yöntemdir. Diğer taraftan, uygun dalgacık fonksiyonunun seçimi de en iyi sınıflandırma performansını elde edebilmek için önem arz etmektedir. Bu sebepten dolayı, bu çalışmada, ayrık dalgacık dönüşümü yöntemlerinin karşılaştırılması üzerinde durarak, zihinsel görevlerin sınıflandırılmasına ilişkin en iyi sınıflandırma performansını gösteren dalgacık fonksiyonunu bulmayı amaçladık. Sınıflandırma süreci için dört farklı zihinsel görev seçildi ve her birinin, diğerleri ile ikili-üçlü kombinasyonları ve tüm durumlara ilişkin karşılaştırılmalı sonuçları elde edildi. Öznitelik çıkartma aşamalarında sırasıyla, haar, coiflets (seviye 1), biortogonal (seviye 6.8), ters biortogonal (seviye 6.8), daubechies (seviye 2) ve daubechies (seviye 4) kullanılmıştır. Sonrasında, elde edilen öznitelik kümesine, öznitelik indirgeme formülü uygulanmış ve elde edilen öznitelik vektörü, eğitim ve test veri kümesi olarak sınıflandırma öncesinde ayrılmıştır. Çıkartılan öznitelik kümeleri, yapay sinir ağı ile sınıflandırılmıştır. Ayrık dalgacık dönüşümü fonksiyonlarından coiflets'in, diğer fonksiyonlara göre daha iyi sonuç verdiği gözlenmiştir.

Anahtar Kelimeler: *Ayrık dalgacık dönüşümü, zihinsel görev sınıflandırma, coiflet dalgacık, daubechies dalgacık, wavelet fonksiyonları*

*[1] Corresponding Author: Istanbul Technical University, Graduate School of Science, Engineering and Technology, Department of Mechatronics Engineering, Istanbul –TURKEY
[2] Uskudar University, Faculty of Engineering and Natural Sciences, Department of Computer Engineering, Istanbul – TURKEY

JNBS 2016 Published by Üsküdar University www.jnbs.org

THE JOURNAL OF NEUROBEHAVIORAL SCIENCES
NÖRODAVRANIŞ BİLİMLERİ DERGİSİ

JNBS 2016 Published by Üsküdar University www.jnbs.org

1.Introduction

The wavelet transform were used to be a powerful and efficient time-frequency analysis method for analyzing non-stationary signals like EEG. In order to determine the features in the frequency bands of related potential recorded by EEG, various methods are used for the spectral and spatial analysis including Wavelet Filter Bank (Robinson et al., 2012). Wavelet filter bank decomposes the transient EEG signal into different frequency bands and every frequency band is figured out by their scaling function (Gandhi et al., 2011). By using this method, the most important step is to choosing a proper wavelet family including the mother wavelet function for signal characterization. It is also important to decide the optimal wavelets and the appropriate number of decomposition levels (Sonia et al.,2013).

A set of features, which contain sequence of wavelet coefficient vectors can be obtained after this process, which decomposes the signal into its wavelets at the specific sub-band frequencies. By using this method, called as feature extraction, the discrimination of the mental tasks becomes applicable. The dimension of the extracted features should also be reduced and made in compact form. The reduced feature vectors are evaluated as the inputs to the classification stage. In this study, an artificial neural network (ANN) with back propagation algorithm was employed. In order to acquire a satisfactory classification performance, the design process and parameters such as the number of neurons in each layer and the number of layers, of the classifier is important to be taken into consideration. The main objective of this paper is determining the most appropriate wavelet function to extract features from raw EEG signals for mental task classification combining a learning algorithm to be able to make classification prediction. The outline of this study is as follows; in section 2.1, we will explain about the experimental system for data collection and filtering stage. In section 2.2, the theory and application of feature extraction method used is reviewed followed by the concepts of discrete wavelet transforms. In section 2.3, the theoretical basis of the wavelet mother functions of DWT are explained in details. In section 2.4, the classification stage using ANN is described. The analysis of the experiments done and results obtained in Section 3 and conclusions are given in the last section.

2.Material and Methods

2.1. Data Acquisition and Preprocessing Stage

All training and testing data were collected from a healthy subject, 28 year old male. The subject focused on each mental task for 10 consecutive seconds. The sampling time of the neuroheadset is 128 samples/second for all channels. 40 different epochs were assigned for training and 10 epochs were used for test data collection. Therefore, we had in total 1600 train and 400 test dataset matrix in association with 4 mental tasks for each 14 sensor channels respectively. After performing the feature extraction process, a valuable reduced feature is evaluated. The dimension of the input training matrix and input test matrix for each feature was set as 14x1600 and 14x400 before the training process. Here, we propose a discrete wavelet transform based feature extraction method by using various wavelet mother functions to investigate the dominant frequency band and timing in EEG signals and compare with each other functions classification accuracy for all combinations of 2 mental task groups of 4 pre-defined mental tasks which are; a) Reciting the alphabet backwards, b) Imagination of rotation of a cube, c) Imagination of left arm movements (open/close) and d) Mentally performing mathematical operation. All dataset were collected from the Emotive EPOC Neuroheadset that it is available to save the EEG signals from 14 channels of the Emotiv-Headset (AF3-F7-F3-FC5-T7-P7-O1-O2-P8-T8-FC6-F4-F8-AF4). The headset samples from all channels at 128 samples/second. The EEG signals were filtered with band pass filter between 0.5 and 45 Hz. using a 6^{th} order butterworth band pass filter to remove the artifacts.

2.2.Discrete Wavelet Transform

The wavelet term indicates to a wave based window function of the main frequency f_0. The wavelets are classified with its wavelet window. The continuous wavelet transform is given as:

$$W_c(\tau,\mu) = \frac{1}{\sqrt{\mu}} \int_{-\infty}^{\infty} x(t)\psi(t)\left\{\frac{t-\tau}{\mu}\right\}dt$$

where $\psi(t)$ refers to the mother wavelet function, the factor $\frac{1}{\sqrt{\mu}}$, normalizes the energy of the signal.

The scale factor μ is the inverse of frequency. If this scaling factor has a small value, the wavelet corresponds to high frequencies of the EEG signal. If this scaling factor is larger, then the wavelet is expanded and refers to low frequencies. τ is the translation factor and corresponds to the position of the center of window while it is shifted by the signal (Tobin, 2007).

The CWT is inefficient for ANN classification, because of the generation of redundant information. For this reason, discrete wavelet transform (DWT) was used which is implemented using sub-band coding method as filter bank (Kannan et al., 1996). The multi-rate filter bank has a series of high-pass and low-pass FIR filter and decimation factors are shown in Figure 1.

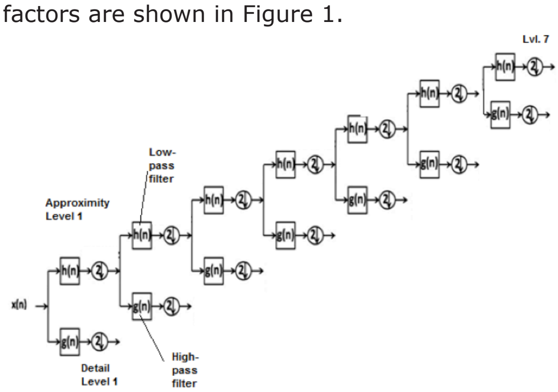


Figure 1: Filter bank representation up to seventh level.

The discrete wavelet transform is defined by the following equation (Mallat, 2008).

$$W_d(j,k) = \sum_j \sum_k x(k) 2^{-j/2} \psi\left(2^{-j}n - k\right)$$

In DWT, the original signal passes through two complementary filters, named low-pass and high-pass filters, and emerges as two signals, called approximation coefficients and detail coefficients (Weeks, 2010). DWT is convenient for processing signals like EEG, since it is very effective in time-frequency localization and multi-scale resolution (Sonia et al., 2013). The low frequency components, h[n] "approximations" are most important than high frequency components, g[n] "detail" in characterizing EEG signals. The consecutive low and high pass filtering can be evaluated by the following equations:

$$Y_{high}[k] = \sum_n x[n] g[2k - n]$$

$$Y_{low}[k] = \sum_n x[n] h[2k - n]$$

Where, Y_{high} and Y_{low} are the outputs of high pass and low pass filters respectively. The filters have a function of sub-sampling the input signal by 2. DWT has varying window size at low and high frequencies, which scans both spatial and spectral domains in order to resolve all frequencies optimally (Mallat, 2008). The wavelet coefficients (or scales) displayed in Figure 2 is in expanded time format. This is difficult format to interpret from a time-frequency point of view since both are embedded in the display. Each of the levels are concatenated in time and displayed as amplitude versus time. Applying the DWT to EEG signals yields the frequency spectrum for each sub-band.

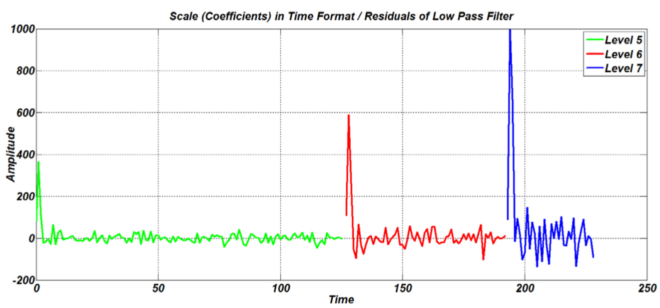


Figure 2: DWT coefficients in time format

$$\beta_i = sqrt\left(abs\left(sum\left(diff\left(fft\left(\lambda_i\right)\right)\right)\right)\right)$$

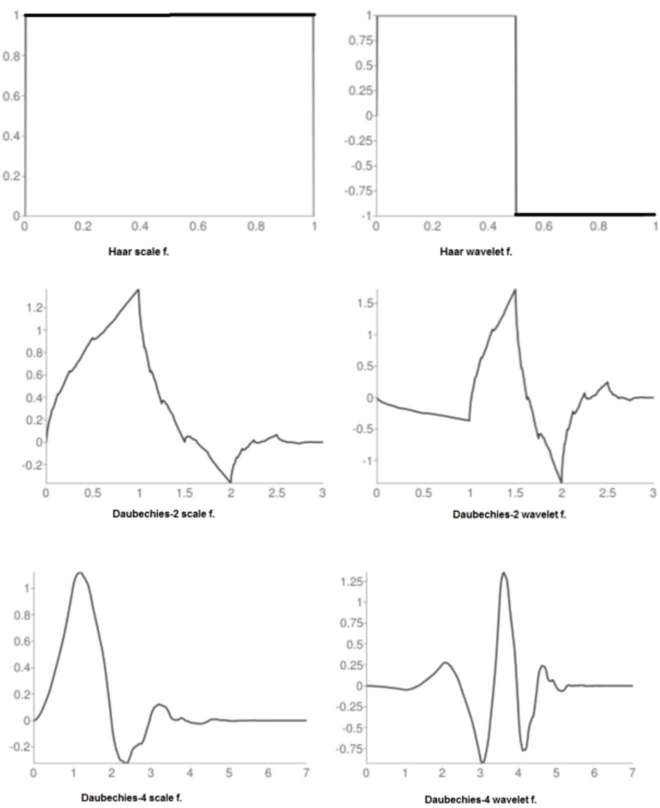
Here, λ_i denotes the node vector, which is the output of the i.th level low-pass filter and β_i is the specific feature for this level.

2.3.Neural Network Classifier

A neural network contains highly interconnected, complex nodes in order to model biological neurons. It works as a parallel processor composed of simple processing units in order to deal with uncertain, fuzzy data sets. Each neuron acquires a weighted input vector or matrix and produces an output vector (Freeman, 2006). Multi-Layer Perceptron NN supported by back-propagation training algorithm are very convenient for brain computer interface (BCI) and pattern recognition applications. MLP's are designed with an input, one or more hidden and an output layer. Multi-layer feed forward NN was applied to our data set for the classification process. The weights on the network are adjusted applying deeply training, the error is minimized based on the gradient descent algorithm. Hereafter the NN was tested with the test dataset by means of performance criteria. After making some heuristic trials, we have set the optimal configuration for the NN parameters as; number of hidden layers (1 or 2), number of neurons in each layer (50) and the maximum number of iterations in the learning process (1000), the learning rate 0.035, the momentum rate 0.2, performance criteria "mse", training algorithm "scaled conjugate gradient".

2.4 Explanation of DWT Basis Functions

Unfolding the information, localized within the signal, is based upon the structural basis function. The information hidden in the signal can be obtained through dilation and shifting procedure. It is essential to select the correct and efficient wavelet function for specific applications. In this paper, we explain a frequently used DWT basis function analysis (Kuzu et al., 2013). The scale and mother wavelet functions of the haar, daubechies-2, daubechies-4, coiflets-1, biorthogonal-6.8, reverse biorthogonal-6.8 are depicted in Figure 3, respectively.



THE JOURNAL OF NEUROBEHAVIORAL SCIENCES
NÖRODAVRANIŞ BİLİMLERİ DERGİSİ

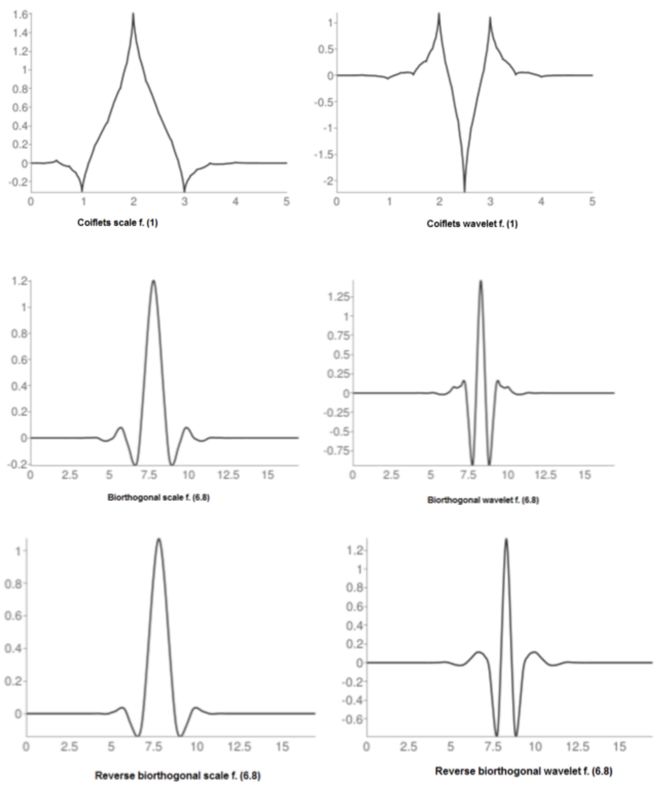


Figure 3: Wavelet shapes of different wavelet functions and their scale functions

2.4.1.Haar

Haar wavelet is a row of rescaled "square-shaped" functions. Take that $\phi(t)$ is a scaling function (Daniel, 1994; Stoloescu et al., 2010).

$$\phi(t) = \begin{cases} 1 \to 0 \le t \le 1 \\ 0 \to otherwise \end{cases}$$

The haar mother wavelet function can be obtained by the following function:

$$\psi(t) = \begin{cases} 1 \to 0 \prec t \le \frac{1}{2} \\ -1 \to \frac{1}{2} \prec t \le 1 \\ 0 \to otherwise \end{cases}$$

Haar wavelet is orthogonal to its own translations and dilations and not continuous.

2.4.2.Daubechies

Daubechies wavelets are similar to the haar wavelet transform by evaluating the averages and difference through the scalar production with scaling and wavelets (Mohammed et al., 2009). The orthonormal wavelets are established with arbitrary number N of vanishing wavelet moments and minimal length of support 2N-1 (Cerna et al., 2008). Daubechies wavelet can be evaluated by using following mother and scaling functions:

$$\psi(t) = \sqrt{2}\sum_{m=0}^{2N-1}(-1)^m h_{2N-1-m}\phi(2t-m)$$

$$\phi(t) = \sqrt{2}\sum_{m=0}^{2N-1} h_m\phi(2t-m)$$

Where $h_0, h_1, h_2, \ldots, h_{2N-1}$ are the constant coefficients of the filter.

2.4.3.Coiflets

Coiflet wavelet function and its scaling function have 2N and 2N-1 moments equal to 0, respectively. The two functions have a support of length 6N-1. The main indicative feature of coiflet wavelet is to have highest number of vanishing moments for both scaling and wavelet function for any given support width (Majumdar et al., 2013). The approximation properties depend on the number of vanishing wavelet moments (Cerna et al., 2008). Let $P_k f$ be an approximation of $f \in L^2(\mathbb{R})$ on level k.

$$P_k f = \sum_{q\in\mathbb{Z}} \langle f, \phi_{k,q}\rangle \phi_{k,q}$$

and for J < k

$$P_k f = \sum_{q\in\mathbb{Z}} \langle f, \phi_{J,q}\rangle \phi_{J,q} + \sum_{l=J}^{k-1}\sum_{q\in\mathbb{Z}} \langle f, \psi_{l,q}\rangle \psi_{l,q}$$

where

$$\phi_{l,q} = 2^{l/2}\phi(2^l . - q)$$

$$\psi_{l,q} = 2^{l/2}\psi(2^l . - q)$$

The wavelet coefficients are evaluated by following formula:

$$\langle f, \psi_{l,q}\rangle = \int_{-\infty}^{\infty} f(t) 2^{l/2}\psi(2^l t - q)\,dt$$

2.4.4.Biorthogonal

Two wavelets are used for decomposition and reconstruction. Biorthogonal wavelets are not based on vahishing moment and all wavelets referred to its family have a symmetric structure. For orthogonal wavelets, the scaling function and mother wavelet are presented by the recursive relationship (Fritz, 1994).

$$\psi(t) = \sqrt{2}\sum_m g_m\phi(2t-k)$$

$$\phi(t) = \sqrt{2}\sum_m h_m\phi(2t-m)$$

JNBS 2016 Published by Üsküdar University www.jnbs.org

Their scaled translates are denoted by;

$$\phi_m^n(t) = 2^{n/2}\phi\left(2^n t - m\right)$$

$$\psi_m^n(t) = 2^{n/2}\psi\left(2^n t - m\right)$$

2.4.5.Reverse biorthogonal

Reverse biorthogonal wavelet family is obtained from the biorthogonal wavelet coupled. Reverse biorthogonal wavelet families are guided by biorthogonal spline wavelets, therefore the symmetrical condition and reconstruction can be assured (Varuneshkumar et al., 2015).

3.Results and Discussion

Feature vectors were created from the extracted nodes of decomposed wavelet coefficients of EEG signals at the 7th level. Our approach is to find out a proper mother wavelet function based on extracted feature set to get satisfactory classification accuracy. We investigated the effect of various wavelet functions whose filter lengths are different from each other. The EEG signals collected for 4 mental tasks were decomposed into coarse approximation and detailed information. DWT employs its set of scaling functions and wavelet functions, which are associated with low-pass and high-pass filters respectively. The EEG signals collected from each electrode channel were decomposed up the 7th level in the case of wavelet filter bank decomposition. Extracted feature vectors from both the methods were fed into the ANN for classification step. We divided the feature vector set into three sets, 70% of which is the data is used for training, 15% for validation and 15% for the testing processes respectively. ANN uses one input layer, one hidden layer and one output layer. Working with this network structure, the feature vector set obtained were first trained and then their performance were tested accordingly. The corresponding accuracies of each mental task were evaluated after testing processes. The classification performance results obtained using DWT for different wavelet functions are listed in table 1.

Table 1: The classification accuracies obtained for 4 mental task subsets by using different wavelet basis function as feature extraction method on test data

Class \ Mother Wavelet	hear	coiflets 1	biorthogonal 6.8	reverse biorthogonal 6.8	db2	db4
Alphabet-Cube	91.4%	99.6%	95.6%	95.4%	97.8%	99.4%
Alphabet-LeftArm	60.3%	64.2%	61.1%	61.0%	62.5%	63.5%
Alphabet-MathOp	62.5%	68.2%	65.1%	65.4%	67.0%	68.0%
Cube-LeftArm	90.3%	98.5%	94.3%	94.1%	96.5%	98.0%
Cube-MathOp	91.0%	99.7%	95.3%	95.2%	97.5%	99.3%
LeftArm-MathOp	73.0%	80.0%	76.5%	76.4%	78.3%	79.6%
Alphabet-Cube-LeftArm-MathOp	77.5%	85.3%	81.5%	81.2%	83.2%	84.5%

Since there are different mother wavelets of different wavelet families available, the choice of the wavelet family and the mother wavelet plays an important role in terms of classification accuracies. The results clearly underline that coiflets1 type of mother wavelet performs better than the other members of its wavelet family with its 85.3% classification performance for 4 mental task classification and higher accuracies for all other dual task classifications. Dual task performance of the proposed methods verifies the better performance of coiflet1. Besides it is also possible to deduce that, the performance of each method is good enough to work on. For the following studies, especially real time BCI design, real time response is crucially important and the response time is to be considered and has much more importance.

4. Conclusion

In this paper, a comparative study of wavelet based feature extraction methods such as discrete wavelet transform based wavelet filter bank decomposition are performed. These methods are combined with neural networks for classification purpose. The performance of both these techniques are tested and evaluated. Both the techniques are found to be efficient in EEG signal processing. The most suitable mother wavelet for feature extraction and classification of EEG signals was found. The features were extracted from the 4 different mental task performed after the decomposition by each of the wavelet family and ANN was employed to classify the cases. Based on the classification accuracy rate obtained, it was found that Coiflet of order 1 mother wavelet function, whose general classification performance for 4 mental tasks is 85.3%, is the best wavelet family for analysis of EEG signal. The computational complexity and the feature vector size were also reduced by using DWT. Provided that the experimental results done, a wavelet transform is an elegant tool for the analysis of non-stationary signals like EEG. The experimental results show that this hybrid architecture using DWT and ANN could effectively extract the features from the EEG signal for various applications. Coiflets1 as a feature extraction method achieve higher classification accuracy compared to other wavelet functions.

References

Cerna D., Finek V., Nazjar K., (2008). On the exact values of coefficients of coiflets, Central European Journal of Mathematics 6(1) pp. 159-169.

Daniel T.L., Akio Y., (1994). Wavelet Analysis: Theory and Applications, Hewlett Packard Journal.

Freeman J.A., Skapura D.M., (2006). Neural Networks Algorithm Application and Programming Techniques, Pearson Education.

Fritz Keinert, (1994) Biorthogonal Wavelets for Fast Matrix Computations, Applied and Computational Harmonic Analysis, Vol 1, Issue 2, Pages 147-156,

Gandhi T., Panigrahi B.K., Ananad S., (2011). A comparative study of wavelet families for EEG signals classification, Neurocomputing, 74, 3051-3057.

Kannan R., Vetterli M., Herley C., (1996). Wavelets, Subband Coding, and Best Basis, proceedings of the IEEE, Vol. 84, No. 4.

Kuzu A., Baran, E.A., Bogosyan S., Gokasan M., Sabanovic A., (2013). WPT based compression for bilateral control," in Industrial Electronics Society, IECON 2013 - 39th Annual Conference of the IEEE , pp.5686-5691.

JNBS 2016 Published by Üsküdar University www.jnbs.org

Majumdar. S., Dixit. A. (2013). Comparative Analysis of Coiflet and Daubechies Wavelets Using Global Threshold For Image Noising, International Journal of Advances in Engineering&Technology.

Mallat S., (2008). A Wavelet Tour of Signal Processing 3rd ed.. Academic Press.

Mohammed A.S., Nivin G., Beate M., (2009). Daubechies Versus Biorthogonal Wavelet for Moving Object Detection in Traffic Monitoring Systems, http://edoc.hu-berlin.de/series/informatik-berichte.

Robinson, N., Vinod, A.P., Cuntai Guan, (2012). A modified Wavelet-Common Spatial Pattern method for decoding hand movement directions in brain computer interfaces, in Neural Networks (IJCNN), The 2012 International Joint Conference on.

Sonia S., David P.S., Poulose J., (2013). A Comparative Study of Wavelet Based Feature Extraction Techniques in Recognizing Isolated Spoken Words, International Journal of Signal Processing Systems.

Stolojescu C., Railean I., Moga S., Isar A., (2010). Comparison of wavelet families with application to WiMAX traffic forecasting, in Optimization of Electrical and Electronic Equipment (OPTIM), 12th International Conference on, pp.932-937.

Tobin P., (2007), PSpice for Digital Signal Processing, Morgan & Claypool.

Varuneshkumar, M., Anil K., and Jaiswal A.K., (2015). Performance Comparison of Daubechies, Biorthogonal and Haar Transform for Grayscale Image Compression. International Journal of Computer Applications 126(9):40-42.

Weeks M., (2010). Digital Signal Processing Using MATLAB & Wavelets, 2nd. Edition, Jones and Barlett Publishers Inc.