Neurocomputing 74 (2011) 3051–3057

Contents lists available at ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

# A comparative study of wavelet families for EEG signal classification

Tapan Gandhi [a,*], Bijay Ketan Panigrahi [b], Sneh Anand [a]

[a] Brain Dynamics and Cognitive Engineering Laboratory, Centre for Biomedical Engineering, Indian Institute of Technology, All India Institute of Medical Sciences, New Delhi, India
[b] Machine Intelligence Research Laboratory, Department of Electrical Engineering, Indian Institute of Technology, Delhi, India

ARTICLE INFO

*Article history:*
Received 9 March 2010
Received in revised form
4 April 2011
Accepted 13 April 2011
Communicated by V. Palade
Available online 6 June 2011

*Keywords:*
Wavelet transform
Feature extraction
Epileptic EEG
Probabilistic Neural Network (PNN)
Support Vector Machine (SVM)

ABSTRACT

Over the past two decades, wavelet theory has been used for the processing of biomedical signals for feature extraction, compression and de-noising applications. However the question as to which wavelet family is the most suitable for analysis of non-stationary bio-signals is still prevalent among researchers. This paper attempts to find the most useful wavelet function among the existing members of the wavelet families for electroencephalogram signal (EEG) analysis. The EEGs considered for this study belong to both normal as well as abnormal signals like epileptic EEG. Important features such as energy, entropy and standard deviation at different sub-bands were computed using the wavelet functions—Haar, Daubechies (orders 2–10), Coiflets (orders 1–10), and Biorthogonal (orders 1.1, 2.4, 3.5, and 4.4). Feature vectors were used to model and train the Probabilistic Neural Network (PNN) and the classification accuracies were evaluated for each case. The results obtained from PNN classifier were compared with Support Vector Machine (SVM) classifier. From the statistical analysis, it was found that Coiflets 1 is the most suitable candidate among the wavelet families considered in this study for accurate classification of the EEG signals. In this work, we have attempted to improve the computing efficiency as it selects the most suitable wavelet function that can be used for EEG signal processing efficiently and accurately with lesser computational time.



## 1. Introduction

Processing and analysis of bio-signals using software techniques have come into play since the early 1960s providing physicians with fast and accurate means toward more precise diagnosis [1]. Feature extraction and classification of the signal (required for diagnostic purposes), however, have always been the two most critical problems encountered in time domain analysis [2]. The sole purpose of feature extraction, is to extract salient characteristics from digitized data collected from the data acquisition phase [3] followed by classification based on the extracted features [4–8]. The feature of the signal is derived from its linear expansion coefficients where the most common linear expansion method used is Fourier transform [9]. Since the early days of digital processing, Fourier transform has been most commonly applied for signal representation. However, bio-signals frequently characterized by a non-stationary time behavior if processed with Fourier transform, would not yield the best result. Hence, for such transient signals, a time–frequency representation is highly desirable, with an aim to derive meaningful features [10].

From the variety of approaches available [2,3,7–9,11,12], the Wavelet transform was found to be an effective time–frequency analysis tool for analyzing the transient signals, as this method unifies different tools that have been developed for processing application till now. The feature extraction and representation properties can be used to evaluate various transient events in biological signals [13]. Several wavelet families are available for signal characterization and selection of appropriate wavelet is very important for the analysis of signals. Depending on the type of bio-signal to be analyzed, the mother wavelet is chosen according to the convenience and the requirement of the experimenter. The research work done till date for bio-signal classification using wavelet technique has been carried out mostly using the Daubechies family of order mostly 2 or 4 [3,14–16]. Moreover, the automated diagnostic system designed for detection purposes gives an accuracy of 70–90% (depending on the bio-signal classified) [14–16]. The research presented in this paper deals with the selection of the most suitable wavelet function for signal analysis of EEG signature in particular. Here, the wavelet techniques were used to decompose the epileptic/normal EEGs for feature extraction followed by classification of the signals using SVM and PNN with an impressive diagnostic accuracy of about 99.3% [7,8]. The reason for selection of epileptic EEG signal for this work is its variation in morphologies like sharp, spikes, and slow waves, each of which is in a different frequency range [17].

* Corresponding author.
*E-mail addresses:* gandhitk@gmail.com (T. Gandhi), bkpanigrahi@ee.iitd.ac.in (B.K. Panigrahi), sneha@cbme.iitd.ac.in (S. Anand).

0925-2312/$ - see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.neucom.2011.04.029

The property of wavelet theory can be best explored in such type of transient signal as it decomposes the epochs into different frequency bands (spike—13.5–50 Hz; sharp—5–12.5 Hz, and slow—1–2.5 Hz) which can then be analyzed depending upon their scaling function [18].

Till date, several soft-computing methods have been proposed in the literature for the diagnosis of the epileptic activity in EEG signal [19–26]. They include template matching [27],time domain, frequency domain [27,28] and time–frequency domain [27–29] which are very few. There is no standard method for selecting the best wavelet for processing EEG signals [11,14]. The choice of wavelet has significant impact on the quality of results with regard to the classifier, which takes the wavelet coefficients as input features. Using efficient classification tool, precise learning ability and processing capacity of neural network can be found out to analyze EEGs efficiently in minimal time for reliable diagnosis.

The goal of this present work is to find out the most suitable wavelet function which can be used to extract features from EEG signals for various applications like brain machine interfacing or to design expert systems for diagnosis of epileptic activity efficiently.

## 2. Data collection and analysis

Two categories of data were selected for the present study. The first category of EEG data used for this study was recorded on a Grass Telefactor EEG Twin3 machine with sampling frequency 400 Hz available at Sir Ganga Ram Hospital, New Delhi. EEG recordings of twenty one subjects were collected from the hospital's epileptic seizure database. Supracranial data was acquired using 16 gold electrodes on the scalp according to international 10–20 system of electrode placement. Two sets of data were selected for the study. The first set of data was epileptic signals having 300 epochs of 5 s duration. The second set of data was background activity of the same person when signals were seizure-free (300 epochs of 5 s duration). The data was selected under supervision of experienced neurologists from the large epileptic database of the hospital.

The second category of EEG data used for this study is publicly available, described in Ref. [27]. The detailed description of the signals can be obtained from reference as mentioned. The complete data set consists of five sets (denoted A–E), each containing 100 single channel EEG signals of 23.6 s duration. Sets A and B have been taken from surface (extracranial) EEG recordings of five healthy volunteers with eye open and closed, respectively. Signals from sets D and C have been measured in seizure-free intervals from five subjects in the epileptic zone and from the hippocampal formation of the opposite hemisphere of the brain. Set E comprises of epileptic signals recorded during seizure (ictal) from all recording sites. Sets C–E have been recorded intracranially.

### 2.1. Wavelet based feature extraction and parameter estimation

Performance of the expert system depends on the signal analysis, feature selection and classification methods used. Wavelet decomposition was employed and features were extracted from EEG (normal/abnormal) signals. However, the output of wavelet transform can be significantly affected by the choice of the mother wavelet (the basic wave shape) with which the signal is analyzed [30–33].

### 2.2. Wavelet decomposition

The wavelet is a smooth and quickly vanishing oscillating mathematical function with good localization both in frequency and time [33]. A wavelet family $\psi_{a,b}(t)$ is a set of elementary function generated by dilations and translations of a unique admissible mother wavelet $\psi(t)$

$$\Psi_{a,b}(t) = \frac{1}{\sqrt{|a|}}\psi\left(\frac{t-b}{a}\right) \tag{1}$$

where $a$, $b \in R$, $a \neq 0$, $a$, $b$ are the scaling (dilation) and translation parameters, respectively, and '$t$' is the time. The scale parameter will decide the oscillatory frequency and the length of the wavelet, the translation parameter will decide its shifting position.

The WT can be implemented with a specially designed pair of FIR filters called quadrature mirror filter (QMFs) pairs. QMFs are distinctive because the frequency responses of the two FIR filters separate the high- and low-frequency components of the input signal. The dividing point is usually half-way between 0 Hz and half of the data sampling rate (the Nyquist frequency). The outputs of the QMF filter pair are decimated (or de-sampled) by a factor of two. The low-frequency (low-pass) filter output is fed into another identical QMF filter pair. This operation can be repeated recursively as a tree or pyramid algorithm, yielding a group of signals that divides the spectrum of the original signal into octave bands with successively coarser measurements in time as the width of each spectral band narrows and decreases in frequency. The tree or pyramid algorithm can be applied to the WT by using the wavelet coefficients as the filter coefficients of the QMF filter pairs. In WT multi-resolution algorithm (MRA), same wavelet coefficients are used in both low-pass (LP) and high-pass (HP) filters. The LP filter coefficients are associated with the scaling function, and the HP filter is associated with the wavelet function. Fig. 1 shows the tree algorithm of a multi-resolution WT for a discrete EEG signal sampled at 400 Hz.

The outputs of the LP filters are called the approximations (A), and the outputs of the HP filters are called the details (D). In MRA, any time series can be completely decomposed in terms of the approximation and detail coefficients based on the level of decomposition as shown in Fig. 1. Application of DWT on raw signal produces a multi-resolution analysis (MRA) of various statistical and non-statistical parameters across time and frequency. The subsets of the wavelet coefficients of the decomposition tree were selected as input vectors to the classifier.

### 2.3. Parameters for feature extraction

Feature sets were constructed using MRA analysis shown in Fig. 1 and coefficients were stored for further processing.

In order to reduce the feature dimension, few statistical and non-statistical parameters were considered instead of directly training the classifier with detail ($D_i(t)$) and approximation ($A_i(t)$) coefficients. For this purpose, we have selected Energy (EDA), Entropy (ENT), and standard deviation (SD) as parameters based on which the PNN was trained for the classification of signals.

The energy at each decomposition level was calculated using the following equations:

$$ED_i = \sum_{j=1}^{N} |D_{ij}|^2, \quad i = 1,2,\ldots,l$$

$$EA_i = \sum_{j=1}^{N} |A_{ij}|^2 \tag{2}$$

The entropy at each decomposition level was calculated using the following equation.