IntechOpen

Discrete Wavelet Transforms

Biomedical Applications

Edited by Hannu Olkkonen

DISCRETE WAVELET TRANSFORMS - BIOMEDICAL APPLICATIONS

Edited by **Hannu Olkkonen**

Discrete Wavelet Transforms - Biomedical Applications
http://dx.doi.org/10.5772/1818
Edited by Hannu Olkkonen

Contributors

Daniel Barbosa, Carlos Lima, Dalila Roupar, Osslan Osiris Vergara Villegas, Vianey Guadalupe Cruz Sanchez, Dora Maria Ballesteros, Andres Eduardo Gaona, Luis Fernando Pedraza, Mario Orlando Oliveira, Arturo Bretas, Tilendra Shishir Sinha, Gautam Sanyal, Irene Lena Hudson, Gustavo Adolfo Alonso, Jean-Louis Marty, Roberto Munoz, Juan Manuel Gutiérrez, Abdel-Rahman Al-Qawasmi, Khaled Daqrouq, Khalooq y. Al Azzawi, Tariq Ali Abu Hilal, M Murugappan, Ramachandran Nagarajan, Sazali Yaacob, Gang-Ding Peng, Allan Wong, Raquel Cervigon, Francisco Castells, S. M. Debbal

Notice

First published in Croatia, 2011 by INTECH d.o.o.
eBook (PDF) Published by IN TECH d.o.o.
Place and year of publication of eBook (PDF): Rijeka, 2019.
IntechOpen is the global imprint of IN TECH d.o.o.
Printed in Croatia

Legal deposit, Croatia: National and University Library in Zagreb

Additional hard and PDF copies can be obtained from orders@intechopen.com

Discrete Wavelet Transforms - Biomedical Applications
Edited by Hannu Olkkonen

p. cm.

ISBN 978-953-307-654-6

eBook (PDF) ISBN 978-953-51-5599-7

Meet the editor

Hannu Olkkonen received his Ph.D. degree in year 1982 from the University of Kuopio, Finland. Since then he has worked as a Professor of Physics and Electronics at the Department of Applied Physics, University of Kuopio, Finland. He is currently a leader of the Biomedical Instrumentation Group (BMIG). He is authored more that 260 reviewed journal articles and served as an editor of several scientific books and book chapters. His research interests include wide range of digital signal processing applications in research and industry, wireless multisensor networks and discrete wavelet transform algorithms for biomedical signal analyses.

Contents

Preface

The discrete wavelet transform (DWT) has an established role in multi-scale processing of biomedical signals, such as EMG and EEG. Since DWT algorithms provide both octave-scale frequency and spatial timing of the analyzed signal. Hence, DWTs are constantly used to solve and treat more and more advanced problems. The DWT algorithms were initially based on the compactly supported conjugate quadrature filters (CQFs). However, a drawback in CQFs is due to the nonlinear phase effects such as spatial dislocations in multi-scale analysis. This is avoided in biorthogonal discrete wavelet transform (BDWT) algorithms, where the scaling and wavelet filters are symmetric and linear phase. The biorthogonal filters are usually constructed by a ladder-type network called lifting scheme. Efficient lifting BDWT structures have been developed for microprocessor and VLSI environment. Only integer register shifts and summations are needed for implementation of the analysis and synthesis filters. In many systems BDWT-based data and image processing tools have outperformed the conventional discrete cosine transform (DCT) -based approaches. For example, in JPEG2000 Standard the DCT has been replaced by the lifting BDWT.

A difficulty in multi-scale DWT analyses is the dependency of the total energy of the wavelet coefficients in different scales on the fractional shifts of the analysed signal. This has led to the development of the complex shift invariant DWT algorithms, the real and imaginary parts of the complex wavelet coefficients are approximately a Hilbert transform pair. The energy of the wavelet coefficients equals the envelope, which provides shift-invariance. In two parallel CQF banks, which are constructed so that the impulse responses of the scaling filters have half-sample delayed versions of each other, the corresponding wavelet bases are a Hilbert transform pair. However, the CQF wavelets do not have coefficient symmetry and the nonlinearity disturbs the spatial timing in different scales and prevents accurate statistical analyses. Therefore the current developments in theory and applications of shift invariant DWT algorithms are concentrated on the dual-tree BDWT structures. The dual-tree BDWTs have appeared to outperform the real-valued BDWTs in several applications such as denoising, texture analysis, speech recognition, processing of seismic signals and multiscale-analysis of neuroelectric signals.

This book reviews the recent progress in DWT algorithms for biomedical applications. The book covers a wide range of architectures (e.g. lifting, shift invariance, multi-scale analysis) for constructing DWTs. The book chapters are organized into four major parts. Part I describes the progress in implementations of the DWT algorithms in biomedical signal analysis. Applications include compression and filtering of biomedical signals, DWT based selection of salient EEG frequency band, shift invariant DWTs for multiscale analysis and DWT assisted heart sound analysis. Part II addresses speech analysis, modeling and understanding of speech and speaker recognition. Part III focuses biosensor applications such as calibration of enzymatic sensors, multiscale analysis of wireless capsule endoscopy recordings, DWT assisted electronic nose analysis and optical fibre sensor analyses. Finally, Part IV describes DWT algorithms for tools in identification and diagnostics: identification based on hand geometry, identification of species groupings, object detection and tracking, DWT signatures and diagnostics for assessment of ICU agitation-sedation controllers and DWT based diagnostics of power transformers.

The chapters of the present book consist of both tutorial and highly advanced material. Therefore, the book is intended to be a reference text for graduate students and researchers to obtain state-of-the-art knowledge on specific applications. The editor is greatly indebted to all co-authors for giving their valuable time and expertise in constructing this book. The technical editors are also acknowledged for their tedious support and help.

Hannu Olkkonen, Professor
University of Eastern Finland,
Department of Applied Physics
Kuopio,
Finland

Part 1

Biomedical Signal Analysis

1

Biomedical Applications of the Discrete Wavelet Transform

Raquel Cervigón
Universidad de Castilla-La Mancha
Spain

1. Introduction

In eighties wavelets came up as the time-frequency revolution in signal processing. In 1989 Mallat proposed the fast Discrete Wavelet Transform (DWT) algorithm to decompose a signal using a set of quadrature mirror decomposition filters, and which have respective band-pass and low-pass properties specific to each mother wavelet (Mallat, 1999). Since this period Wavelets have been applied in a variety of fields including fluid dynamics, engineering, finance geophysics, study of musical tones, image compressionand de-noising just to name few. In addition, it has been extensively used in medicine because of the irregularities inherent to biological signals.

In the discrete wavelet analysis the information stored in the wavelets coefficient is not repeated, it allows the complete regeneration of the original signal without redundancy. This property has motivated much of the effort for development of wavelet-based signal compression algorithms, particularly for ECG signals compression techniques are important to enlarge storage capacity an improve methods of ECG data transmission. DWT removes redundancy in the signal and provides a high compression ratio and high quality reconstruction of ECG signal.

The bioelectric signals contain noise originated by devices or interference of the network that hardly can be eliminated by conventional analogous filters. DWT is a technique to filtrate signals with low distortion to eliminate noise. This process can be applied to different physiology signals, where signals with additive noise are decomposed using the DWT and a threshold is applied to each of the detail coefficient levels. All coefficients with an absolute value greater than the threshold are thought to be part of information and those below the threshold are presumably derived from noise. The noise coefficients can be set to zero and a noise-free signal can then be reconstructed and used for signal detection. Recently, several wavelet-based methods have been used for unsupervised de-noising and detection of data with low signal-to-noise ratio. In particular, DWT has been applied in the quantification of human sympathetic nerve signal activity to discriminate action potentials. Wavelet decomposition effectively filters the nerve signal into several frequency sub-bands while preserving its temporal structure. Each sub-band of wavelet processing decorrelates successive noise-related values and compares progressively more dilated versions of a general spike shape to each point in the signal. This process can make easier the detection of action potentials by separating the signal and noise using their distinct time-frequency signatures.

Discrete Wavelet analysis corresponds to windowing in a new coordinate system, in which space and frequency are simultaneously localized; this property can be helpful in pattern extraction. Wavelets as an alternative tool to analyze non-stationary signal have been applied to ECG delination, to detect accurately the different waves forming the entire cardiac cycle, especially in areas of limited perfomance of of current techniques like QT and ST intervals, P and T-wave recognition, and to clasify ECG waves in different cardiopatologies, identifying ECG waveforms from different arrhythmias, or discriminating between normal and anormal cardiac pattern. In addition, DWT is able to detect specific detailed time-frequency components of ECG signal, for instance, the registers which are sensitive to transient ischemia and eventual restoration of electrohysiological funtion of the myocardial tissue. Moreover, methods for analysing heart rate variability using wavelet transform can be used to detect transient changes without losing frequency information. Several authors have successfully demonstrated the utility of the DWT in time-varying spectral analysis of heart-rate variability during dynamic cardiovascular stimulation.

2. Discrete Wavelet Transform

DWT is a fast algorithm for machine computation, like the Fast Fourier Transform (FFT), it is linear operation that operates on a data vector, transforming it into a numerically different vector of the same length. Also like the FFT, the DWT is invertible and orthonormal. In the FFT, the basis functions are sines and cosines. In the DWT, they are hierarchical set of "wavelet functions" that satisfy certain mathematical criteria (Daubechies, 1992; Mallat, 1989b) and are all translations and scalings of each other.

There is an even faster family of algorithms based on a completely different idea, namely that of multiresolution analysis, or MRA (Mallat, 1989a), then the whole construction may be transcripted into a pair of quadrature mirror filters, defined from the underlying wavelet function, and both are applied to the signal and down-sampled by a factor of two. This process splits the signal into two component, each of half the original length, with one containing the low-frequency or "smooth" information and the other the high-frequency or "difference" information. The process is performed again on the smooth component, breaking it up into "low-low" and "high-low" components and it is repeated several times.

DWT achieves a multiresolution decomposition of x_n on J octaves labelled by $j = 1, \ldots, J$. It is precisely this requirement for a multresolution-hence hierarchical- structure that makes fast computation possible. The requirement for a multiresolution computation can be stated as follows: Given some signal, at scale j, one decomposes it in a sum or details, at scale $j+1$ (the true wavelet coefficients), plus some residual, representing the signal at resolution $j+1$ (twice as coarse). A further analysis at coarser scales involves only the residual.

Consider two filter impulse responses $g(n)$ (corresponding to some low-pass interpolating filter-the scaling function) and $h(n)$ (corresponding to some a high-pass filter-the discrete wavelet) (eq. 1 and 2). The downsampled outputs of first low-pass and high-pass filters provide the approximation, and the detail, respectively. The first approximation is further decomposed and this process is continued until all wavelet coefficients are determined.

$$h_{j+1}(n) = \sum_k h_j(k) g[n - 2k] \tag{1}$$

$$g_{j+1}(n) = \sum_k g_j(k) h[n - 2k] \tag{2}$$

The wavelets and scaling sequences are obtained iteratively as i.e., one goes from one octave j to the next $(j+1)$ by applying the interpolation operator

$$f(n) \rightarrow \sum_k f(n)g(n-2k) \tag{3}$$

Which should be thought of as the discrete equivalent to the dilation $f(t) \rightarrow 2^{-1/2}f(t/2)$. Consider, for example, the computation of wavelet coefficients $c_{j,k}$, for a fixed j, the coefficient is the result of filtering the input signal by $h_j(n)$ and decimating the output by the suppression of one every $2^j th$ sample. Now the z-transform of filter $h_j(n)$ can be easily deduced from equation 1, which reads $H_{j+1}(z) = H_j(z^2)G(z)$ in z-transform notation. We obtain:

$$H_{j+1}(z) = G(z)G(z^2)\dots G(z^{2^{j-1}})H(z^{2^j}) \tag{4}$$

and similarly for $g_j(n)$,

$$G_{j+1}(z) = G(z)G(z^2)\dots G(z^{2^j}) \tag{5}$$

The computations of a DWT are easily reorganized in form of binary tree, where the decomposition may also be truncated at any level of the process before an average signal of length of one sample is reached. In any event, the dyadic DWT consists of the set of detail signals generated at each level of the transform, together with the average signal generated at the highest level (shortest length signals) of the transform.

A remarkable feature of many useful wavelet transforms, is that they obey a perfect reconstruction theorem. That is the dyadic DWT may be inverted to recover the original signal exactly. The inversion process is carried out first by upsampling (or expanding) the highest level detail and average signals. Upsampling is carried out by inserting zeros between samples of the signal to be upsampled. Then, the upsampled average and detail signals are run through synthesis filters and added together. The sum signal is the average signal for the next lowest level of the wavelet transform. This process is carried out at each lower level until the original signal is recovered at the lowest level as the zero level average signal (Kaiser, 1994; Mallat, 1998; Strang & Nguyen, 1997).

The computed wavelet coefficients provide a compact representation that shows the energy distribution of the signal in time and frequency. We assume that the signals are stationary within each short segment in time. Thus within the segment, the variance of the wavelet transform $wx(t)$ and the wavelet function $\psi(t)$ can be considered as a value unrated to t, written as,

$$E[x\psi(t)]^2 = E[(x*\psi)^2(t)] = \sigma_x^2 \tag{6}$$

And in the frequency domain,

$$\sigma_x^2 = \frac{1}{2\pi}\int_{-\infty}^{\infty} S_x(\omega)|\psi(\omega)|^2 d\omega \tag{7}$$

Furthermore, the spectral components of interest may be located anywhere in the frequency axis, even in the neighborhood of the cross-point between two adjacent frequency bands. At this location, the spectral component is assigned with a small gain signalling, and low detection sensitivity. This problem can be approached by considering the cross-correlation between $wx_j(t)$ and $wx_{j+1}(t)$, $Rwx_j, wx_{j+1} = E[wx_j(t+1)wx_{j+1}(t)]$ and the autocorrelation of the signal.

3. Discrete Wavelet Transform in biomedical research

Wavelet Transform has been proposed as an alternative way to analyze the non-stationary biomedical signals, which expands the signal onto the basis functions. The wavelet method act as a mathematical microscope in which we can observe different parts of the signal by just adjusting the focus.
A conventional application of wavelet methods to processing of a medical waveform uses a wavelet transform based on the application of a single wavelet, rather than a basis set constructed from a family of mathematically related wavelets. Again, the choice of a wavelet with appropriate morphological characteristics relative to the physiological signal under consideration is crucial to the success of the application. In the following sections will be introduced different uses of DWT in cardiology research, with interesting applications such as de-noising and compression of medical signals, electrocardiogram (ECG) segmentation and feature extraction, analysis of heart rhythm variability, and the analysis of different cardiac arrhythmias.

4. Signal compression

The compressibility of a sampled signal is the radio of the total area of time-frequency plane (N, for a signal sampled at N) divides by the total area of the information cells. It is possible to automatically analyze signals by expanding them in the best basis, then drawing the corresponding time-frequency plane representation.
The DWT is both "complete" and has "zero redundancy", which means that all the signal information is contained in the resulting transform and none is duplicated between transform coefficients. By converting the signal into its DWT coefficients and then removing all except those containing the most pertinent signal information, the resulting transform is much smaller in size, which provides a good way of compressing a signal. Performing an "inverse transform" on the remaining components recreates a signal that very nearly matches the original. This is the basis of compression algorithms that can be applied to biomedical images and signals, such as in the development of effective ECG data compression. Increasing use of computerized ECG processing systems requires effective ECG data compression techniques which aim to enlarge storage capacity and improve methods of ECG data transmission over internet lines. Moreover ECG signals are collected both over long periods of time and at high resolution. This creates substantial volumes of data for storage and transmission. The fundamental reason that ECG compression is regarded as a difficult problem is that the ECG waveform contains clinically significant information on a wide variety of time scales. Data compression seeks to reduce the number of bits of information required to store or transmit digitized ECG signals without significant loss of signal quality. Moreover, some ECG compression algorithms have been used only for strictly limited diagnostic objectives, as in Holter monitors. Another objective is to develop a high-fidelity compression algorithm that would not impair later physician diagnoses.
An early paper suggested the wavelet transform as a method for compressing both ECG and heart rate variability data sets (Crowe et al., 1992). Thakor et al. compared two methods of data reduction on a dyadic scale for normal and abnormal cardiac rhythms, detailing the errors associated with increasing data reduction ratios (Thakor et al., 1993). Using DWT and Daubechies D10 wavelets, Chen et al. compressed ECG data sets resulting in compression ratios up to 22.9:1 while retaining clinically acceptable signal quality, with an adaptive quantization strategy which allows a predetermined desired signal quality to be

achieved (Chen & Itoh, 1998). Miaou et al. (Miaou & Larn, 2000) also propose a quality driven compression methodology based on Daubechies wavelets and later on biorthogonal wavelets (Miaou & Lin, 2002), this algorithm adopts the set partitioning of hierarchical tree (SPIHT) coding strategy.

5. Wavelet Transform based filtering "De-noising"

The noise present in the signal can be removed by applying the wavelet shrinkage de-noising method while preserving the signal characteristics, regardless of its frequency content.
Wavelets have the added advantage that the resulting expansions are orthogonal or energy preserving, allowing to compare an adapted expansion to signals in order to minimize the cost of representation. Such adapted decompositions perform compression and analysis simultaneously. It is possible to design an idealized graphical presentation of the time-frequency information obtained by such a best adapted wavelet analysis, and for such presentation is possible to recognize and extract transient features. The small components in the analysis may be treated as noise and discarded, where an iterative algorithm always produces the best decomposition, at the cost of many more iterations plus more work for each iteration. Mallat's stopping criterion is to test the amplitude ratio of successive extracted amplitudes; this is a method of recognizing residuums which have the statistics of random noise.
Consider the standard univariate regression: $y_i = f(x_i) + \epsilon_i$, where $i = 1, ..., n$, and ϵ_i are independent $N(0, \sigma^2)$ random variables; and f is the "true" function. We can reformulate the problem in terms of wavelet coefficients: $\hat{w}_{jk} = w_{jk} + \epsilon_{jk}$, where j is the level ($j = 0, ..., j-1$), and k, the displacement ($k = 0, ..., 2^j$). It is often reasonable to assume that only a few large coefficients contain information about the underlying function, while small coefficients can be attributed to noise. *Shrinkage* consists in attenuating or eliminating the smaller wavelet coefficients and reconstructing the profile using mainly the most significant wavelet coefficients and all the scaling coefficients. Several shrinkage approaches have been proposed. For example, the "hard" threshold approach selects coefficients using a keep or kill policy, nevertheless using "soft" thresholding, if the magnitude of the wavelet coefficient is greater than (less than, respectively) the threshold, the coefficient is shrunk toward zero by an amount that depends on how large the magnitude of the coefficient is (set to zero, respectively). Donoho and Johnstone proposed the "universal" threshold, $\lambda = \sigma\sqrt{2log n}$, and showed that it performs very well in both hard and soft thresholding. Thresholds can also be chosen based on the data using a hypothesis testing procedure (Alshamali & Al-Fahoum, 2003; Donoho & Johnstone, 1994). Data-adaptive thresholds might become very important in analyzing molecular biological data because hypothesis testing procedures can be used to test the appropriateness of various thresholds to the data under different biological assumptions (Lio, 2003). Finally, it is worth mentioning that several authors have proposed Bayesian thresholds and have reported interesting results (Abramovich et al., 2009).
This evolution in electrocardiographic start with the algorithms for noise reduction in ECG signals using the dyadic wavelet transform with wavelet-based and wavelet packet-based thresholding methods for removing noise from the ECG (Kishimoto et al., 1995; Tikkanen, 1999).
More recently, Nikolaev et al have suppressed electromyogram (EMG) noise in the ECG using a method incorporating wavelet transform domain Wiener filtering (Nikolaev et al., 2001), this method resulted in an improvement in signal-to-noise ratio of more than 10 dB.

In addition, the non-invasive blood pressure artifact removal algorithm makes use of DWT. The system used in most patient monitors measures the small fluctuations in pressure in a blood pressure cuff (applied to one of the patient's limbs) to obtain a determination of the patient's systolic and diastolic pressure. Usually the mean arterial pressure and pulse rate are obtained as well. These pressure fluctuations are usually termed "oscillometric pulses" (Geddes & Badylak, 1991). The wavelet-based artifact elimination algorithm is based on the observation that the dyadic DWT puts the physiologic oscillometric waveform in a very different region of the transform plane than the signal components attributable to artifact. The modified DWT may then be inverted to yield a reconstruction of the oscillometric signal with artifact substantially reduced. The reconstructed oscillometric signal may then be used as an input to a pressure determination algorithm in the usual way for the measurement of desired patient pressure values.

6. ECG signal parameter extraction

The ECG registers a measure of the electrical activity associated with the heart. The ECG is measured at the body surface and results from electrical changes associated with activation first of the two small heart chambers, the atria, and then of the two larger heart chambers, the ventricles. The contraction of the atria manifests itself as the *P* wave in the ECG and contraction of the ventricles produces the feature known as the *QRS* complex. The subsequent return of the ventricular mass to a rest state repolarization produces the *T* wave. Repolarization of the atria is, however, hidden within the dominant QRS complex. Analysis of the local morphology of the ECG signal and its time varying properties has produced a variety of clinical diagnostic tools.

To use ECG signals as identity verification, a real-time detection of the ECG characteristics is needed. With the real-time extraction of ECG characteristics, we could verify different individual. The basic objects of the analysis are a P-wave, a QRS-complex, a T-wave, a P-Q interval, a S-T segment, and a Q-T interval (see Fig. 1).

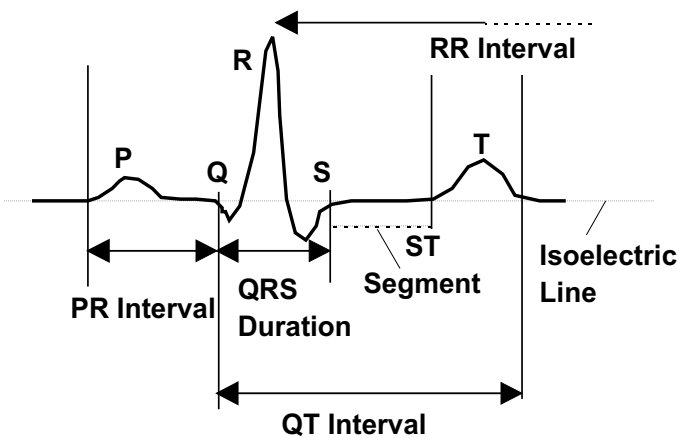


Fig. 1. Normal ECG delineation

Producing an algorithm for the detection of the *P* wave, *QRS* complex and *T* wave in an ECG is a difficult problem due to the time varying morphology of the signal subject to physiological conditions, moreover the localization of wave onsets and ends is much more difficult, as

the signal amplitude is low at the wave boundaries and the noise level can be higher than the signal itself. A number of wavelet-based techniques have been proposed to detect these features. Senhadji et al (1995) compared the ability of three different wavelets transforms (Daubechies, spline and Morlet) to recognize and describe isolated cardiac beats (Senhadji & Wendling, 2002). Sahambi et al. employed a first-order derivative of the Gaussian function as the wavelet for the characterization of ECG waveforms (Sahambi et al., 1997a;b). Moreover, wavelet-based QRS detection methods have been suggested by a variety of groups including Li et al (1995) who proposed a method based on finding the modulus maxima larger than a threshold obtained from the pre-processing of preselected initial beats, this threshold can be updated during the analysis to obtain a better performance (Li et al., 1995).
Kadambe et al. have described an algorithm which finds the local maxima of two consecutive dyadic wavelet scales, and compared them in order to classify local maxima produced by R waves and by noise. Kadambe et al. report a sensitivity of 96.84% and a positive predictive value of 95.20% when tested on a limited data set (four 30 min tapes acquired from the American Heart Association (AHA) database) (Kadambe et al., 1999). Other work has been undertaken by Park et al (1998) using a wavelet adaptive filter to minimize the distortion of the ST-segment due to baseline wanderings. In a subsequent paper by Park et al (2001), a wavelet interpolation filter (WAF) is described for the removal of motion artefacts in the ST-segment of stress ECGs (Park et al., 2001). Furthermore, Martinez et al (2004) also utilize the algorithm of Li et al applying a dyadic wavelet transform to a robust ECG delineation system which identifies the peaks, onsets and offsets of the QRS complexes, and P and T waves. The QRS detector obtains a sensitivity and a positive predictivity of 99.8% in a very well-known MIT-BIH Arrhythmia Database (Martinez et al., 2004).

7. Heart rate variability

Rather than consider the local morphology of the whole ECG signal, many researchers have focused on the longer term temporal variability of the heartbeat, the analysis of which allows an assessment of autonomous nervous system activity. The analysis of heart rate variability (HRV) requires the sequence of timing intervals between beats, taken between each R point on the QRS complex. This RR interval can be plotted against time to give the RR time series. In normal practice, however, ectopic beats are removed from the RR series leaving only normal sinus beats: the NN time series. It is this modified time series that is used in the analysis of HRV. The minute fluctuations present in the NN intervals are used for assessing the influence of the autonomic nervous system components on the heart rate. Long range correlations and power law scaling have been found through the analysis of heartbeat dynamics. The heart rate and rhythm is largely under the control of the autonomic nervous system. Traditional spectral analysis of HRV has been reported to aid the understanding of the modulatory effects of neural mechanisms on the sinus node. There are three main spectral components in a traditional spectral calculation, they are generally classed as: very low frequency (VLF) ranging from 0.003 to 0.04 Hz, low frequency (LF) ranging from 0.04 to 0.15 Hz and high frequency (HF) ranging from 0.15 to 0.4 Hz components. In addition, sometimes an ultra low frequency (ULF) is defined as spectral components with frequencies less than 0.003 Hz. The relative contribution of vagal and sympathetic modulation of the heart rate is attributed to the distribution of spectral power in these bands. The most common of the techniques rely on the accurate determination of the temporal location of the R wave based on signal matched filters or time-frequency decomposition methods. Over recent years, a number of groups have attempted to use wavelet-based methods to gain additional insight

into the mechanisms controlling heart rate variability. The wavelet transform is partitioned into the HF, LF and VLF regions whereby temporal-spectral characteristics of the surface may then be investigated. Thurner et al have employed both Daubechies D10 and Haar wavelets in the analysis of human heartbeat intervals. They found that, at distinct wavelet scales, corresponding to the interval 16-32 heartbeats, the scale-dependent standard deviations of the wavelet coefficients could differentiate between normal patients and those with heart failure. Significantly, they report 100% accuracy for a standard 27 patient data set (Thurner et al., 1998). Further development of the technique is detailed in subsequent papers (Heniford et al., 1998; Seidensticker et al., 1998; Wiklund et al., 2011). Ivanov et al investigated the ECG signals acquired from subjects with sleep apnea, by sampling at an a scale equivalent to 8 heartbeats, they performed a local smoothing of the high-frequency variations in the signal in order to probe patterns of duration in the interval 30-60 s. The authors used the data to characterize the nonstationary heartbeat behaviour and elucidate phase interactions (Ivanov et al., 1996). Furthermore, this type of analysis has been applied to study myocardial ischemia, where a method for analysing HRV signals using wavelet transform was applied to obtain a time-scale representation for VLF, LF and HF bands using the orthogonal multiresolution pyramidal algorithm. Comparing a normality zone against the ischaemic episode, it was found a statistical significant increase in the LF and HF bands in the ischaemic episode, this index can be useful for the assessment of dynamic changes and patterns of HRV during myocardial ischaemia (Gamero et al., 2002).

8. Cardiac arrhythmias

A number of wavelet-based techniques have been proposed for the identification, classification and analysis of arrhythmic ECG signals. In 1997, Govindan described an algorithm for classifying bipolar electrograms from the right atrium of sheep into four groups: normal sinus rhythm, atrial flutter, paroxysmal atrial fibrillation and chronic atrial fibrillation. In this method, it was used a Daubechies D6 wavelet to preprocess the ECG data prior to classification using an artificial neural network. They found paroxysmal AF the most difficult to classify with a $77\% \pm 9\%$ average success rate and normal sinus rhythm the easiest, achieving $94\% \pm 8\%$ (Govindan et al., 1997). Using a raised cosine wavelet transform, Khadra et al undertoke a preliminary investigation of three arrhythmias: ventricular fibrillation (VF), ventricular tachycardia (VT) and atrial fibrillation (AF) Khadra et al. (1997), they developed an algorithm based on the scale-dependent energy content of the wavelet decomposition to classify the arrhythmias, distinguishing them from each other and normal sinus rhythm. Zhang et al proposed a novel arrhythmia detection method, based on a wavelet network, for use in implantable defibrillators, their system, originally developed as a model to identify relationships between concurrent epicardial cell action potentials and bipolar electrogram, detects the bifurcation point in the ECG where normal sinus rhythm degenerates into a pathological arrhythmia (ventricular fibrillation) (Zhang et al., 1999). Al-Fahoum and Howitt proposed a radial basis neural network for the automatic detection and classification of arrhythmias which employs preprocessing of the ECG using the Daubechies D4 wavelet transform, they reported 97.5% correct classification of arrhythmia from a dataset of 159 arrhythmia files from three different sources, with 100% correct classification for both ventricular fibrillation and ventricular tachycardia (al Fahoum & Howitt, 1999). Moreover, it has been already shown its potential for the detection of ventricular late potentials (Dickhaus et al., 1994; Khadrea et al., 1993; Meste et al., 1994).

Morlet et al presented a Morlet wavelet-based method for the discrimination of patients prone to the onset of ventricular tachycardias (VTs), they found that the detection of strings of local maxima of the wavelet transform vector at or after 98 ms after the QRS onset point was a reasonable criterion for VT risk stratification in post-infarction patients. They reported achieving 85% specificity at 90% sensitivity for their patient group (Morlet et al., 1993). Englund et al studied the predictive value of wavelet decomposition of the signal averaged ECG in identifying patients with hypertrophic cardiomyopathy at increased risk of malignant ventricular arrhythmias or sudden death (Englund et al., 1998), wavelet analysis used in their study was undertaken subsequent to signal averaging of the beats. Thus intermittent local or transient aspects of the ECG can be lost to its interrogation. A later study by this group evaluated a number of wavelet decomposition parameters for their potential for risk stratification of patients with idiopathic dilated cardiomyopathy (Yi et al., 2000). They found that wavelet analysis was superior to time domain analysis for identifying patients at increased risk of clinical deterioration.

In addition, different wavelet analysis have been applied to Atrial fibrillation (AF). It is the most frequently found sustained cardiac arrhythmia in clinical practice. It is the most common cause of embolic stroke, and is associated with a doubling of overall mortality and morbidity from cardiovascular disease (Benjamin et al., 1998; Kannel et al., 1982). AF is characterized by an abnormal excitation of the atria, where the normal and regular atrial activation is substituted by several coexisting wavefronts that continuously depolarize the atrial cells (Allessie et al., 1995; Fuster et al., 2006). As a result, atrial activation is chaotic and disorganized, and consequently the atria are not able to be contracted in a regular rhythm. On the surface electrocardiogram (ECG), *P* waves are no longer visible, being replaced by rapid oscillations or fibrillatory waves that vary in size, shape, and timing (Allessie et al., 1995; Bollmann et al., 1999). The ventricular response depends on the electrophysiological properties of the atrioventricular node, what results in an irregular and rapid ventricular rhythm. Fig. 2 represents an example of normal sinus rhythm and AF episodes.

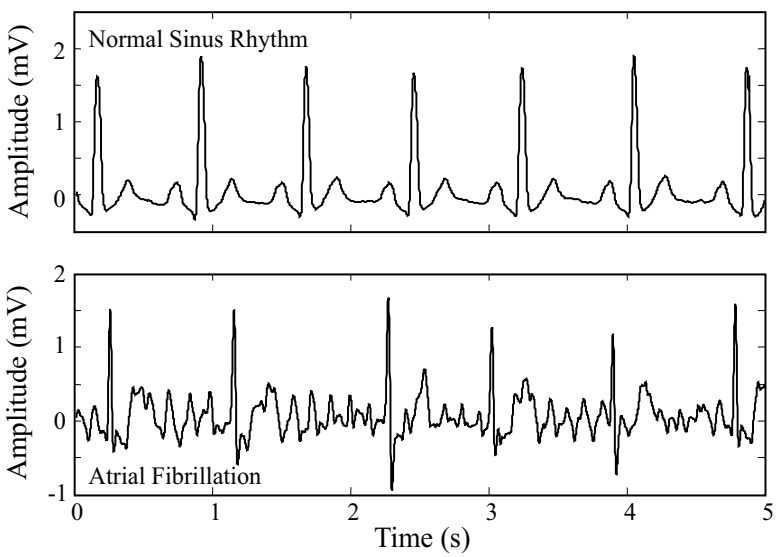


Fig. 2. Examples of normal sinus rhythm and AF episodes.

Duverney et al have developed a combined wavelet transform-fractal analysis method for the automatic detection of AF from heart rate intervals (Roche et al., 2002). After training their method on healthy sinus rhythm and chronic AF ECGs, they achieved 96.1% sensitivity and 92.6% specificity for discriminating AF episodes in paroxysmal AF. A technique for the explanation of AF from within an ECG signal using a modulus maxima de-noising technique (Addison et al., 2000), where the modulus maxima lines, at this scale with a high proportion of the total energy within this scale are selected, are followed across scales and subtracted to leave a residual signal associated with both system noise and, more importantly, atrial activity. This time-frequency partitioning of the signal results in two components: one (1) containing combined low and high frequency components that correspond to large scale features in the signal, and a second (2) containing the remaining high frequency components that correspond to small scale AF features and noise. In practice, most applications are concerned with signal de-noising and hence the retention of component (1).
Furthermore, a study was conducted to analyze ECG signals from patients with persistent AF in order to extract reliable parameters to predict early AF recurrence after successful electrical cardioversion. The technique employed for ECG analysis was based on the wavelet transform, which have been successfully employed to solve other ECG problems. DWT analysis with biorthogonal family was applied, and the energy from different scales of detail coefficients of the descomposition was evaluated, the wavelet coefficients output at each subband may provide important information of the ECG signal, and they could be used in combination with appropriate statistical analysis tools in order to predict the risk of AF recurrence after successful electrical cardioversion. From this analysis, standard deviations of the coefficients in each subband were obtained, but its significance was lower than the cited parameter. The calculus of the ratio of the energy between different scales of the decomposition resulted statistically significant, however its capacity of prediction resulted lower than the continuous wavelet transform analysis, and the higher differences were obtained in the variable energy (eq. 6) in relation to some detail coefficients and the ratio between some scales of the decomposition (Cervigón et al., 2007). In addition, the effect of anaesthetic agents in restoration rhythm procedures during AF has not been investigated. It was evaluated the effects of a widely used anaesthetic agent (propofol) in the fibrillation patterns. Intra-atrial recordings belong patients diagnosed with AF were analyzed “before” (baseline) and “during” anaesthetic infusion. The goal of this study is to characterize the variation in atrial properties along the atria in both states. The wavelet variance of a time series on a scale by scale basis along the DWT decomposition, hence has considerable appeal when physical phenomena are analyzed in terms of variations operating over a range of different scales. As mother wavelet was used the haar wavelet and discrete wavelet transform partitioned the variance of a signal over 7 scales. The proposed methodology provide an additional approach to the understanding of the role of the anaesthetic, showing a decrease in the variance inter-scales during the anaesthetic infusion in the right atrium, with the opposite effect in the left atrium (i.e. a increase in the organization degree) (Cervigón et al., 2008).

9. Conclusions

Signal processing of the ECG has been already demonstrated its effectiveness to solve some clinical problems. In that sense, wavelet transform has emerged over recent years as a key time-frequency analysis and coding tool for the ECG. Indeed, its ability to localize simultaneously local spectral and temporal information within a signal. In addition, the fact that the wavelet transform exhibits different window sizes depending on the frequency

band —broad at low frequencies and narrow at high frequencies— leads to an optimal time-frequency resolution in all frequency ranges. The coefficients output by the wavelet transform at each subband may provide important information of the ECG signal, and they could be used in combination with appropriate statistical analysis tools in order to predict different arrhythmias. It has been already shown its potential for feature extraction and discrimination between normal and abnormal cardiac patterns, detection of ventricular late potentials, characterization of beat-to-beat fluctuations in the heart rate under diverse physiological conditions, study of cardias arrhytmias, such as he risk of AF recurrence after successful electrical cardioversion etc.

In addition, its discrete form, the DWT provide the basis of powerful methodologies for partitioning pertinent signal components which serve as a basis for potent compression strategies.

The DWT has interesting mathematics and fits in with standard signal filtering and encoding methodologies. It produces few coefficients, where it is possible to recover the original signal, during the inverse transform process, without any loosing of energy. However, it exhibits non-stationarity and coarse time-frequency resolution.

DWT analysis of different signals has made possible the identification of pertinent features within the transform difficult, if not practically impossible. The non-stationarity of the DWT can also cause problems in terms of repeatability and robustness of the analysis, unless it particularly lends itself to an ensemble averaged method.

In conclusion, wavelet transform can be a helpful instrument to know more about the mechanisms of biological structure, it has been shown that inside biomedical signals, such as ECG signal contains hidden information that a tool such as wavelet transform could extract.

10. References

Thurner S., Markus F., & Malvin T., (1998). Multiresolution Wavelet Analysis of Heartbeat Intervals Discriminates Healthy Patients from Those with Cardiac Pathology *Physical Review Letters* 7(80): 1544–9.

Abramovich, F., Grinshtein, V., Petsa, A. & Sapatinas, T. (2009). On bayesian "testimation" and its application to wavelet thresholding.

Addison, P. S., Watson, J. N., Clegg, G. R., Holzer, M., Sterz, F. & Robertson, C. E. (2000). Evaluating arrhythmias in ecg signals using wavelet transforms, *IEEE Eng Med Biol Mag* 19(5): 104–9.

al Fahoum, A. S. & Howitt, I. (1999). Combined wavelet transformation and radial basis neural networks for classifying life-threatening cardiac arrhythmias, *Med Biol Eng Comput* 37(5): 566–73.

Allessie, M., Konings, K. & Wijffels, M. (1995). *Atrial arrhythmias: State of the art*, DiMarco, J.P. and Prytowsky, E.N. Eds., Futura Publishing Company, Armonk, NY, chapter Electrophysiological mechanisms of atrial fibrillation, pp. 155–161.

Alshamali, A. & Al-Fahoum, A. (2003). Comments on "an efficient coding algorithm for the compression of ecg signals using the wavelet transform", *Biomedical Engineering, IEEE Transactions on* 50(8): 1034–1037.

Benjamin, E., Wolf, P., RB, R. D., Silbershatz, H., Kannel, W. & Levy, D. (1998). Impact of atrial fibrillation on the risk of death, *Circulation* (98): 946–952.

Bitzen, Alexander, Sternickel, Karsten, Lewalter, Thorsten, Schwab, Otto, J., Yang, Alexander, Schrickel, Wilko, J., Linhart, Markus, Wolpert, Christian, Jung, Werner, David, Peter, Luderitz, Berndt, Nickenig, Georg, Lickfett & Lars (2007). Automatic p wave analysis

over 24 hours in patients with paroxysmal or persistent atrial fibrillation, *Annals of Noninvasive Electrocardiology* 12(4): 306–315.

Bollmann, A., Sonne, K., Esperer., H., Toepffer, I., Langberg, J. & Klein, H. (1999). Non-invasive assessment of fibrillatory activity in patients with paroxysmal and persistent atrial fibrillation using the holter ECG, *Cardiovascular Research* 44: 60–66.

Cervigón, R., Castells, F., Mateo, J., Moreno, J. & Sánchez, C. (2008). Variance differences in atrial fibrillation during anaesthetic effect, *Proc Conf IEEE Comput Cardiol*.

Cervigón, R., Sánchez, C., Castells, F., Blas, J. & Millet, J. (2007). Wavelet analysis of electrocardiograms to characterize recurrent atrial fibrillation, *J Franklin Instit* 344: 196–211.

Chen, J. & Itoh, S. (1998). A wavelet transform-based ecg compression method guaranteeing desired signal quality, *IEEE Trans Biomed Eng* 45(12): 1414–9.

Crowe, J. A., Gibson, N. M., Woolfson, M. S. & Somekh, M. G. (1992). Wavelet transform as a potential tool for ecg analysis and compression, *J Biomed Eng* 14(3): 268–72.

Daubechies, I. (1992). *Ten Lectures on Wavelets (CBMS-NSF Regional Conference Series in Applied Mathematics)*, 1 edn, SIAM: Society for Industrial and Applied Mathematics.

Dickhaus, H., Khadrea, L. & Brachmann, J. (1994). Time-frequency analysis of ventricular late potentials, *Methods of Inform in Med* 33(2): 187–195.

Donoho, D. & Johnstone, J. (1994). Ideal spatial adaptation by wavelet shrinkage, *Biometrika* 81(3): 425–455.

Englund, A., Hnatkova, K., Kulakowski, P., Elliot, P. M., Malik, M. & McKenna, W. J. (1998). Use of spectral turbulence analysis for the identification of patients at high risk for ventricular fibrillation and sudden death in patients with hypertrophic cardiomyopathy, *Cardiology* 90(2): 79–82.

Fuster, V., RydÈn, L. E., Cannom, D. S., Crijns, H. J., Curtis, A. B., Ellenbogen, K. A., Halperin, J. L., Heuzey, J.-Y. L., Kay, G. N., Lowe, J. E., Olsson, S. B., Prystowsky, E. N., Tamargo, J. L., Wann, S., Smith, S. C., Jacobs, A. K., Adams, C. D., Anderson, J. L., Antman, E. M., Halperin, J. L., Hunt, S. A., Nishimura, R., Ornato, J. P., Page, R. L., Riegel, B., Priori, S. G., Blanc, J.-J., Budaj, A., Camm, A. J., Dean, V., Deckers, J. W., Despres, C., Dickstein, K., Lekakis, J., McGregor, K., Metra, M., Morais, J., Osterspey, A., Tamargo, J. L., Zamorano, J. L., of Cardiology/American Heart Association Task Force on Practice Guidelines, A. C., of Cardiology Committee for Practice Guidelines, E. S., Association, E. H. R. & Society, H. R. (2006). Acc/aha/esc 2006 guidelines for the management of patients with atrial fibrillation: a report of the american college of cardiology/american heart association task force on practice guidelines and the european society of cardiology committee for practice guidelines (writing committee to revise the 2001 guidelines for the management of patients with atrial fibrillation): developed in collaboration with the european heart rhythm association and the heart rhythm society., *Circulation* 114(7): e257–e354.

Gamero, L., Vila, J. & Palacios, F. (2002). Wavelet transform analysis of heart rate variability during myocardial ischaemia, *Medical and Biological Engineering and Computing* 40: 72–78.

Geddes, L. A. & Badylak, S. F. (1991). Power capability of skeletal muscle to pump blood, *ASAIO Trans* 37(1): 19–23.

Govindan, A., Deng, G. & Power, J. (1997). Electrogram analysis during atrial fibrillation using wavelet and neural network techniques, *Proc. SPIE 3169* pp. 557–562.

Heniford, B. T., Senler, S. O., Olsofka, J. M., Carrillo, E. H. & Bergamini, T. M. (1998). May-thurner syndrome: management by endovascular surgical techniques, *Ann Vasc Surg* 12(5): 482–6.

Ivanov, P. C., Rosenblum, M. G., Peng, C. K., Mietus, J., Havlin, S., Stanley, H. E. & Goldberger, A. L. (1996). Scaling behaviour of heartbeat intervals obtained by wavelet-based time-series analysis, *Nature* 383(6598): 323–7.

Kadambe, S., Murray, R. & Boudreaux-Bartels, G. F. (1999). Wavelet transform-based qrs complex detector, *IEEE Trans Biomed Eng* 46(7): 838–48.

Kaiser, G. (1994). *A Friendly Guide to Wavelets*, Birkhauser.

Kannel, W., Abbott, R. & Savage, D. (1982). Epidemiological features of chronic atrial fibrillation and the risk of stroke. the framingham study, *New England Journal of Medicine* (306): 1018–1022.

Khadra, L., al Fahoum, A. S. & al Nashash, H. (1997). Detection of life-threatening cardiac arrhythmias using the wavelet transformation, *Med Biol Eng Comput* 35(6): 626–32.

Khadrea, L., Dickhaus, H. & Lipp, A. (1993). Representation of ecg-late potentials in the time frequency plane, *J.Med Eng and Techcnol.* 17(6): 228–231.

Kishimoto, K., Inoue, H., Hamada, M. & Shibuya, T. (1995). Time frequency analysis of dispersive waves by means of wavelet transform, *Journal of Applied Mechanics* 62(4): 841–846.

Li, C., Zheng, C. & Tai, C. (1995). Detection of ecg characteristic points using wavelet transforms, *IEEE Trans Biomed Eng* 42(1): 21–8.

Lio, P. (2003). Wavelets in bioinformatics and computational biology: state of art and perspectives, *Bioinformatics* 19(1): 2–9.

Mallat, S. (1989a). Multiresolution approximations and wavelet orthonormal bases of l2(r), *Transactions of the American Mathematical Society* 315(1).

Mallat, S. (1989b). A theory for multiresolution signal decomposition - the wavelet representation, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 11: 674–693.

Mallat, S. (1998). *A Wavelet Tour of Signal Processing*, Academic Press.

Mallat, S. (1999). *A Wavelet Tour of Signal Processing, Second Edition (Wavelet Analysis & Its Applications)*, 2 edn, Academic Press.

Martinez, J., Almeida, R, O. S., Rocha, A. & Laguna, P. (2004). A wavelet-based ecg delineator: evaluation on standard databases, *IEEE Transactions on Biomedical Engigneering* 51: 570–581.

Meste, O., Rix, H., Caminal, P. & Thakor, N. (1994). Ventricular late potentials characterization in time -frequency domain by means of a wavelet transform, *IEEE Tras Biomed. Eng.* 1: 625–634.

Miaou, S. G. & Larn, J. H. (2000). Adaptive vector quantisation for electrocardiogram signal compression using overlapped and linearly shifted codevectors, *Med Biol Eng Comput* 38(5): 547–52.

Miaou, S.-G. & Lin, C.-L. (2002). A quality-on-demand algorithm for wavelet-based compression of electrocardiogram signals, *IEEE Trans Biomed Eng* 49(3): 233–9.

Morlet, D., Peyrin, F., Desseigne, P., Touboul, P. & Rubel, P. (1993). Wavelet analysis of high-resolution signal-averaged ecgs in postinfarction patients, *J Electrocardiol* 26(4): 311–20.

Nikolaev, N., Gotchev, A., Egiazarian, K. & Nikolov, Z. (2001). Suppression of electromyogram interference on the electrocardiogram by transform domain denoising, *Med Biol Eng Comput* 39(6): 649–55.

Park, K. L., Khil, M. J., Lee, B. C., Jeong, K. S., Lee, K. J. & Yoon, H. R. (2001). Design of a wavelet interpolation filter for enhancement of the st-segment, *Med Biol Eng Comput* 39(3): 355–61.

Roche, F., Duverney, D., Court-Fortune, I., Pichot, V., Costes, F., Lacour, J.-R., Antoniadis, J. A., Gaspoz, J.-M. & Barthelemy, J.-C. (2002). Cardiac interbeat interval increment for the identification of obstructive sleep apnea, *Pacing Clin Electrophysiol* 25(8): 1192–9.

Sahambi, J., Tandon, S. & Bhatt, R. (1997a). Quantitative analysis of errors due to power-line interference and base-line drift in detection of onsets and offsets in ecg using wavelets, *Medical and Biological Engineering and Computing* 35(6): 747–751.

Sahambi, J., Tandon, S. & Bhatt, R. (1997b). Using wavelet transforms for ecg characterization. an on-line digital signal processing system, *IEEE Engineering in Medicine and Biology Magazine* 16(1): 77–83.

Seidensticker, D., Wilcox, J. & Gagne, P. (1998). Treatment of may-thurner syndrome with catheter-directed thrombolysis and stent placement, complicated by heparin-induced thrombocytopenia, *Cardiovasc Surg* 6(6): 607–13.

Senhadji, L. & Wendling, F. (2002). Epileptic transient detection: wavelets and time-frequency approaches., *Neurophysiologie clinique = Clinical neurophysiology* 32(3): 175–192.

Strang, G. & Nguyen, T. (1997). *Wavelets and filter banks*, Wellesley-Cambridge Press.

Thakor, N. V., Guo, X. R., Sun, Y. C. & Hanley, D. F. (1993). Multiresolution wavelet analysis of evoked potentials, *IEEE Trans Biomed Eng* 40(11): 1085–94.

Tikkanen, P. (1999). Nonlinear wavelet and wavelet packet denoising of electrocardiogram signal., *Biol Cybern* 80(4): 259–267.

Wiklund, K., Fernández-Varea, J. M. & Lind, B. K. (2011). A monte carlo program for the analysis of low-energy electron tracks in liquid water, *Phys Med Biol* 56(7): 1985–2003.

Yi, G., Hnatkova, K., Mahon, N. G., Keeling, P. J., Reardon, M., Camm, A. J. & Malik, M. (2000). Predictive value of wavelet decomposition of the signal-averaged electrocardiogram in idiopathic dilated cardiomyopathy, *Eur Heart J* 21(12): 1015–22.

Zhang, X. S., Zhu, Y. S., Thakor, N. V., Wang, Z. M. & Wang, Z. Z. (1999). Modeling the relationship between concurrent epicardial action potentials and bipolar electrograms, *IEEE Trans Biomed Eng* 46(4): 365–76.

2

Discrete Wavelet Transform in Compression and Filtering of Biomedical Signals

Dora M. Ballesteros[1], Andrés E. Gaona[2] and Luis F. Pedraza[3]
[1]*Faculty of Engineering, University Military Nueva Granada*
[2]*Faculty of Engineering, University Francisco José de Caldas*
[3]*Faculty of Technology, University Francisco José de Caldas*
Colombia

1. Introduction

Biomedical signals are a kind of signals that are measured from a specific part of the body, for example from the hearth (electrocardiography: ECG), muscles (electromyography: EMG) and brain (electroencephalography: EEG). This kind of signals have a no-stationary behavior, it means the behavior through the time is changing every time window. For this reason, the pre-processing, processing, and analysis should be different of the deterministic and stationary signals. One of the methods used in the last years to examine biomedical signals is the Discrete Wavelet Transform (DWT), it represents both time and frequency the signal's characteristics in a multi-resolution mode.

In this chapter, we are going to present two applications of the DWT in biomedical signals, it known as filtering and compression. When you have a device that measures the body's signals, it is desired that the information stored or transmitted have high quality and low redundancy; this corresponds to apply a filter and compress the signal. These two blocks (filtering and compression) are added once the signal is acquired and processed by digital signal processing methods. The goal of using the DWT in an algorithm of filtering and compression biomedical signals is the possibility of choosing the signal's coefficients with a significant energy and discards the others that have a very low percentage of all energy. This is possible because in every level of decomposition, the energy of different frequencies and time position is related to a specific coefficient.

In the first part, we present one model of filtering of biomedical signals based on Discrete Wavelet Transform. We analyze the different parameters in the model and its relation to the quality of the new signal. Every parameter affects in low or high manner the quality of the filtered signal and we present the most common test to probe the signal's distortion when the coefficients with low energy have been removed. Additionally, we present some results with one real EMG signal with different configuration of the parameters.

In the second part, we extend the model of filtering to include the stage of compression; we explain the encoding block, which is added to the compression model. Two lossless encoding methods are explained and compared. The compression of some records of ECG is presented.

Finally, in the third part, the architecture of the Discrete Wavelet Transform on a FPGA is shown. The convolution and sub-sampling processes are modeled using VHDL and its performance is simulated using a CAE tool.

2. Filtering technique for background noise

The basic idea of filtering technique is to improve the signal to noise ratio, in fact to reduce the background noise in the biomedical signal. Because noise can affect the reading and interpretation of the signal, a pre-processing step is desirable before the computer analysis. Because the external noise does not have a specific band and its frequency is commonly superposed to the biomedical frequency, it is necessary to design an intelligent model which can be adaptable to different kind of signals. It is possible with the Discrete Wavelet Transform.

The classic technique (Donoho & Johnstone, 1994) includes three important stages: the decomposition of the signal; the identification of low energy coefficients and its rejection (thresholding); and finally, the reconstruction of the new coefficients. It is shown below;

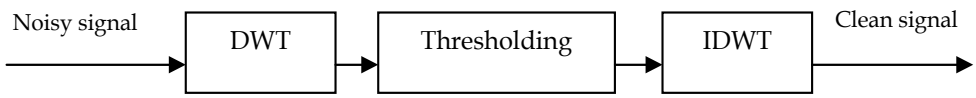


Fig. 1. The filtering technique.

The selection of the DWT is due to the simultaneous representation of the signal and noise in time and frequency. The technique is applied for the model with additive noise, according to:

$$ns = s + an \tag{1}$$

In the expression above *ns* is the noisy signal, s is the biomedical signal and *an* is the additive noise. Because the model corresponds to a lineal system, the wavelet coefficients of the *ns* are equal to the sum of the wavelet coefficients of *s* and the wavelet coefficients of *an*, according to:

If

$$NS = DWT\{ns\}; \qquad S = DWT\{s\}; \qquad AN = DWT\{an\}$$

Then

$$NS = S + AN \tag{2}$$

If the external signal corresponds to white noise, its energy is sparse with low amplitude. Then, the wavelet coefficients of *ns* with low amplitude correspond to the noise of the signal. The noise can be eliminated if the coefficients below a threshold are turned to zero (thresholding).

Every stage of the Figure 1 has parameters related to the performance of filtering. Specifically, the decomposition and reconstruction involve the base wavelet and the number of levels; and the thresholding involves the threshold and the rule.

2.1 Parameter selection

The time and frequency characteristics can vary from signal to signal, then it is necessary to establish a method to identify the best conditions for each specific type of signal (normal,

with pathology, of woman, of man,…). In this section, we present the methodology of selection of the main parameters in the filtering model.
To evaluate in objective form the performance of every combination, there are measurements of the quality of the filtered signal. One of the most used is the PRD (Percentage of RMS), which is calculated according to:

$$PRD = \sqrt{\frac{\sum (ns - f)^2}{\sum (ns)^2}} \tag{3}$$

f represents the filtered signal. If PRD is high, the filter could have eliminated important components of the signal; while, if the PRD is very low, the filter could have not eliminate the noise.

2.1.1 Wavelet family

In Table 2, the most common families are presented. This list is supported by Matlab ©. The index is related to the length of the filter, for example, for sym4 the length is eight. We suggest selecting the base according to the similarity with the biomedical signal. At the most it looks like, better is the representation.
In relation to the length of the filter, it is not recommended to use long filters for short time signals, for example, if the time is 10ms, sym10 is better than sym45.

Wavelet Family	Wname
Daubichies	db1 or haar, db2, ... db10, ... db45
Coifltes	coif1,…, coif45
Symlets	sym2,..sym4,..sym45
Discrete Meyer	Dmey
Biorthogonal	bior1.1, bior1.3, bior1.5, bior2.2, bior2.4, bior2.6, bior2.8, bior3.1, bior3.3, bior3.5, bior3.7, bior3.9, bior4.4, bior5.5, bior6.8,

Table 1. Wavelet Family

2.1.2 Levels of decomposition

The number of levels of decomposition (N) depends on the relation between the sampling frequency and the bandwidth of the signal. A big N is required if the relation is high; an initial rank can be 3 to 10 levels.

2.1.3 Thresholding rule

There are two important thresholding rules applied in most of papers of biomedical signal denoising: soft and hard threshold. However, other rules have been proposed (Quian, 2000). Soft threshold is defined by (Donoho, 1995):

$$g(x)=\mathrm{sgn}(x)(|x|-|th|) \quad \text{if } |x|\geq|th|;$$

$$\text{else} \quad g(x)=0 \tag{4}$$

Where th is the threshold, x is the initial amplitude (input) and g(x) is the result (output). In Figure 2 the function is represented.

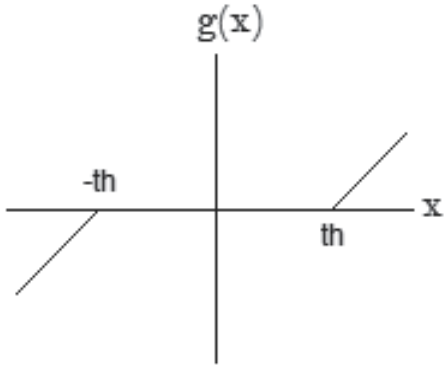


Fig. 2. Soft threshold.

Hard threshold is defined by (Mallat, 1998):

$$g(x) = x \quad \text{if } |x| \geq |th|;$$

$$\text{else} \quad g(x) = 0 \tag{5}$$

The difference between the soft and hard rules is the output when the input exceeds the threshold. In both cases, the output is zero when the input is less than the threshold. The function is presented in Figure 3.

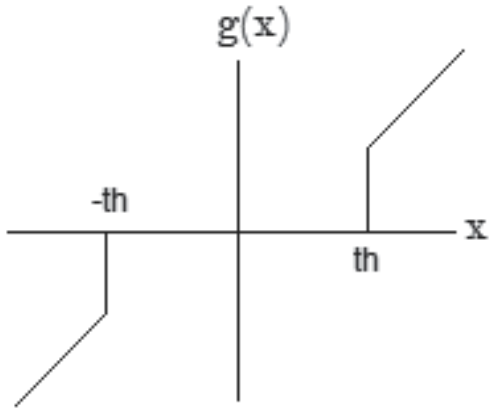


Fig. 3. Hard threshold.

2.2 Results: A case of study of filtering SEMG

The electromyography signal is one of the biomedical signals, which correspond to record of the muscle activity. Because in a typical record, the activity and non-activity regions are acquired and transmitted, it is desirable to clean the signal for improvement the contrast between the two regions.

We analyze the relation of the triangle: threshold, energy retained and percentage of zeros. When the threshold increases the percentage of zero increases too, but, the energy decreases. The appropriate point of the triangle corresponds to the maximum percentage of zeros with the maximum energy retained. Experimentally, we found that the 95% of energy retained is adequate for a right interpretation of the signal (Ballesteros & Gaona, 2007, 2008). In table 2, the results of our study are presented.

Parameter	PRD		Zeros %		Energy Retained %	
	sym6	sym8	sym6	sym8	sym6	sym8
Th1 & hard	0.14	0.16	71.7	72.3	98	96
Th1& soft	0.25	0.27	71.7	72.3	74	71
Th2 & hard	0.22	0.25	77.5	81.1	95	92
Th2 & soft	0.36	0.38	77.5	77.5	83	81

Table 2. Results of filtering: N=3.

Figure 4 presents the results with different parameters. In the left side, the sym6 base was used in the decomposition, threshold was equal to 0.26 and the hard thresholding was applied. It obtained the 95% of the energy retained for the 77.5% of the wavelet coefficients set to zero. In the right, the rule was soft. In this case, the 83% of the energy was conserved with the same number of wavelet coefficients set to zero.
According to table 2 and figure 4, the best combination is th2&hard with the sym6 base. This satisfies the balance criteria.

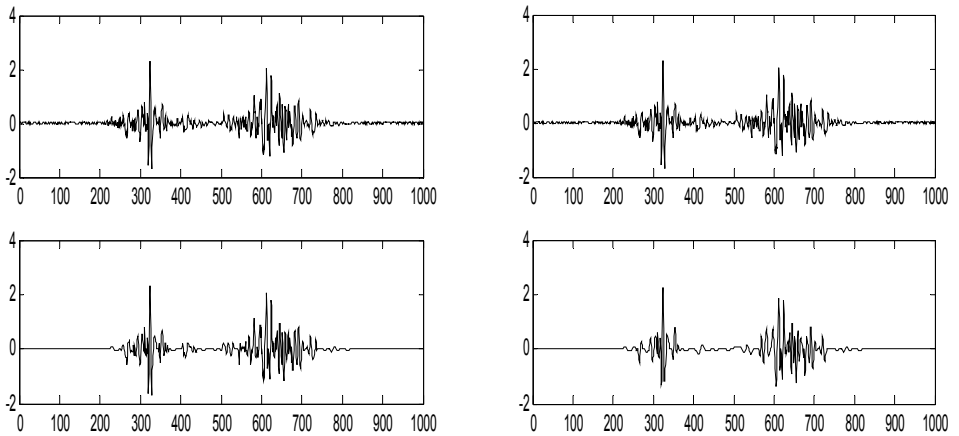


Fig. 4. SEMG and its filtered version.

3. Compression model

The basic idea of the compression model is to reduce the amount of information. Although in some applications the quality in the compressed signal is not important, in the case of biomedical signals the difference must to be the minimum. The purpose is to find the redundancy in the information and eliminate it.
In addition to the three stages of filtering model, an encoding block should be used to improve the Compression Relation (CR). The compression model is presented in Figure 5.

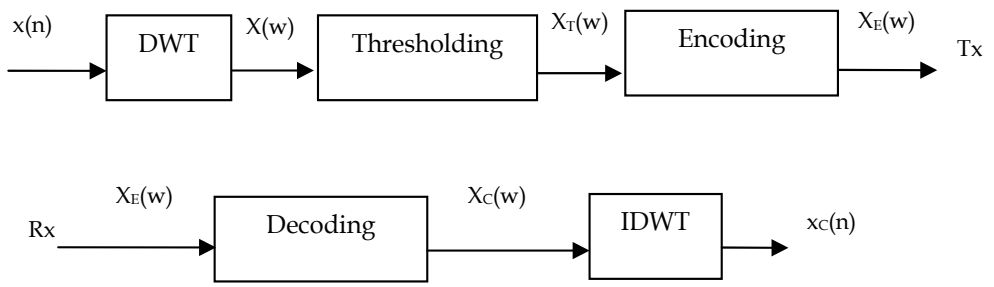


Fig. 5. Compression Model: transmission channel and reception channel.

The encoding can be by lossless or lossy methods (Hankerson et al., 2005). First, the data in the receiver are the same that in the transmitter; while in the second, a part of the information is lost in the process. Because in biomedical signal compression model is appropriate to retain the non-redundant information, we suggest methods without loss of information.

3.1 Lossless methods

We present two of the lossless methods: Run-Length and Huffman. Both take advantage of the thresholding stage to increase the compression ratio (CR).

3.1.1 Run-Length encoding (RL)

This method is used when a number (commonly zero) is repeated many times in a sequence, and then the data in the original stream is replaced by the number and its repetition. (Smith, 2003). The length of the new data decreases when the quantity of zeros increases.
Suppose you have the following data stream:

1 5 0 0 0 0 0 0 0 10

The output stream with run-length encoding is:

1 5 0 7 10

And the CR is 10/5=2.

3.1.2 Huffman encoding

Huffman encoding defines the codebook according to the repetition of every data. It uses more bits in the no-frequent data and fewer bits for the data with higher occurrence (Huffman, 1952). An important feature of Huffman code is that no code can be the header of another; the decoding of data is unique.
The steps for creating the code are:

1. Sort the data from high to low level of repetition.
2. Grouped in pairs of minor repetition. Reapply the first step.
3. Repeat second step until all data have been combined.
4. Draw the Huffman tree with branches of two nodes, where data sets with higher levels of repetition are located to the left of the tree and the lowest level on the right. Assign

'1' to the data of the left and a '0' to the right. Huffman code is read from top to bottom of the tree.

Suppose we have the following list of repetitions into a stream:

Date	A	B	C	D
Repetition	60	15	22	3

Table 3. Date example Huffman code.

The tree according to the four previous steps is:

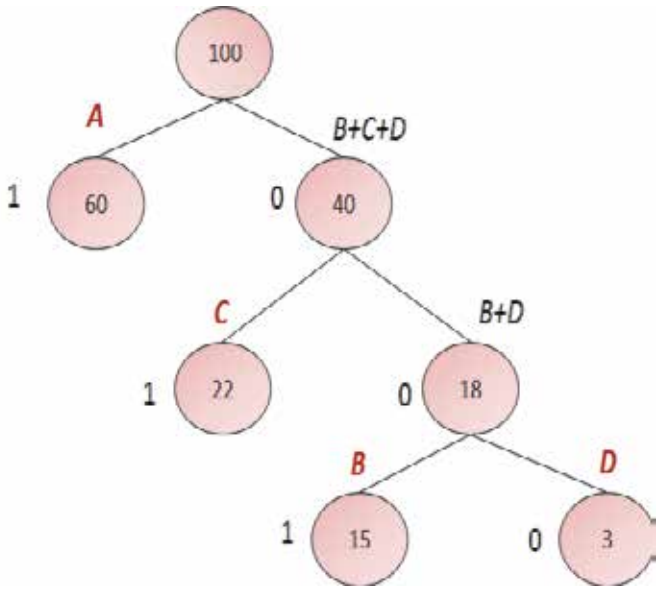


Fig. 6. Huffman tree.

And the codebook is presented in Table 4:

Date	A	B	C	D
Huffman Code	1	001	01	000

Table 4. Codebook.

3.2 Results: A case of study in compression of ECG

The first step to compress a biomedical signal is to analyze its characteristics and determine the parameters for decomposition and the thresholding stage (section 2.1) No matter what encoding method, the output coefficients of the threshold block must retain much of its energy.

3.2.1 Compression with DWT and Run-length/Huffman encoding

Wavelet Transform followed by run length encoding has been used in the last decade in biomedical signal compression. Chen et al (2006) consider that the Huffman code is less

robust against the different statistical characteristics, because it is necessary a prior knowledge of its symbol statistics; but we develop a solution with a static unique code which can be used in different type of ECG.

Ballesteros & Gaona (2009, 2010) developed a compression model for ECG. This work had the following parameters:

- Family: Daubichies, Symlets and BiorSplines (db6, sym6, bior5.5)
- Levels of decomposition: 3 and 4 levels.
- Amplitude of the threshold: it is calculated by

$$th_p=0.0075*\max(w_p) \quad (6)$$

  th_p is the threshold of level P, w_p is the wavelet coefficients of the detail/coarse of level *P* and max(.) is the maximum function.
- Rule of thresholding: hard.
- Encoding method: run-length and Huffman.
- Time: 2, 4, 5 and 10 seconds of the ECG.
- ECG: records 100, 101, 104, 107, 108 and 200 of the MIT Database.

Run length encoding: it used the consecutive zeros to form the output sequence. The wavelet coefficients of four levels of decomposition are the input of the encoding algorithm: the input (D) is composed of detail coefficients (d4, d3, d2, d1) and coarse coefficients (c4). The value of consecutives zero is assigned to the output.

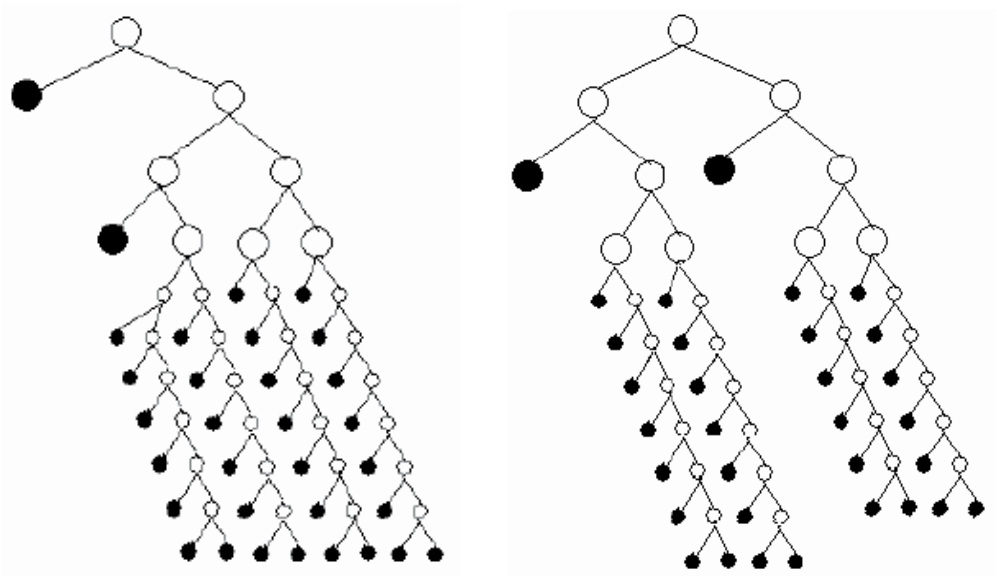


Fig. 7. Huffman tree: detail (left) and coarse (right).

Huffman encoding: two codebooks for the first and second level of decomposition were computed. Additionally, in the third level, one codebook was calculated to the coarse coefficients and other to the detail coefficients. Every codebook was composed by thirty-two codes. Because most of detail coefficients are zero after the output of the threshold block and the coarse coefficients are non-zero, two different kind of tree were estimated. For detail

codebook, the average of all detail coefficients of the six signals was computed. The average signal is divided in thirty-two ranges and the Huffman code was estimated.
The coarse codebook is calculated according to the distribution. The thirty-two codes are assigned according to the coarse amplitude: bigger amplitude has shorter code and smaller amplitude has longer code.

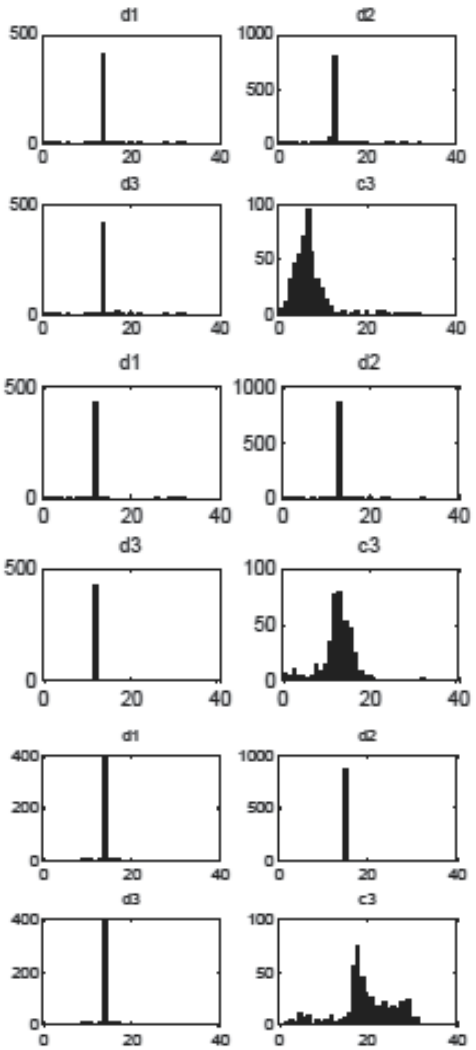


Fig. 8. Histograms: 101 (top); 104 (middle) and 107 (bottom).

The previous picture presents the histograms of the 104 and 107 records, which have different characteristics in time and frequency, but, significant similarities in its histograms. According to Figure 8, the detail coefficients (d1, d2, d3) have a histogram with a significant concentration in one bar; while, the coarse coefficients (c3) have the energy distributed in many amplitudes.
The length of the output stream is theoretically calculated as:

$$Length = \sum (Pi*HCLi) \quad \text{for i=1,…32} \tag{7}$$

Pi is the probability in the range I; HCLi is the Huffman Code Length in the range i.
The results of the 101 record with the base sym6 are presented in Figure 9. It obtained PRD=1.35 and CR=9.24 for run-length encoding and PRD=0.98 and CR=9.32 for Huffman encoding. The figure 10 presents the result for the base db6. Run-length encoding obtained PRD=1.1 and CR=9.11; while Huffman encoding obtained PRD=0.91 and CR=9.4.

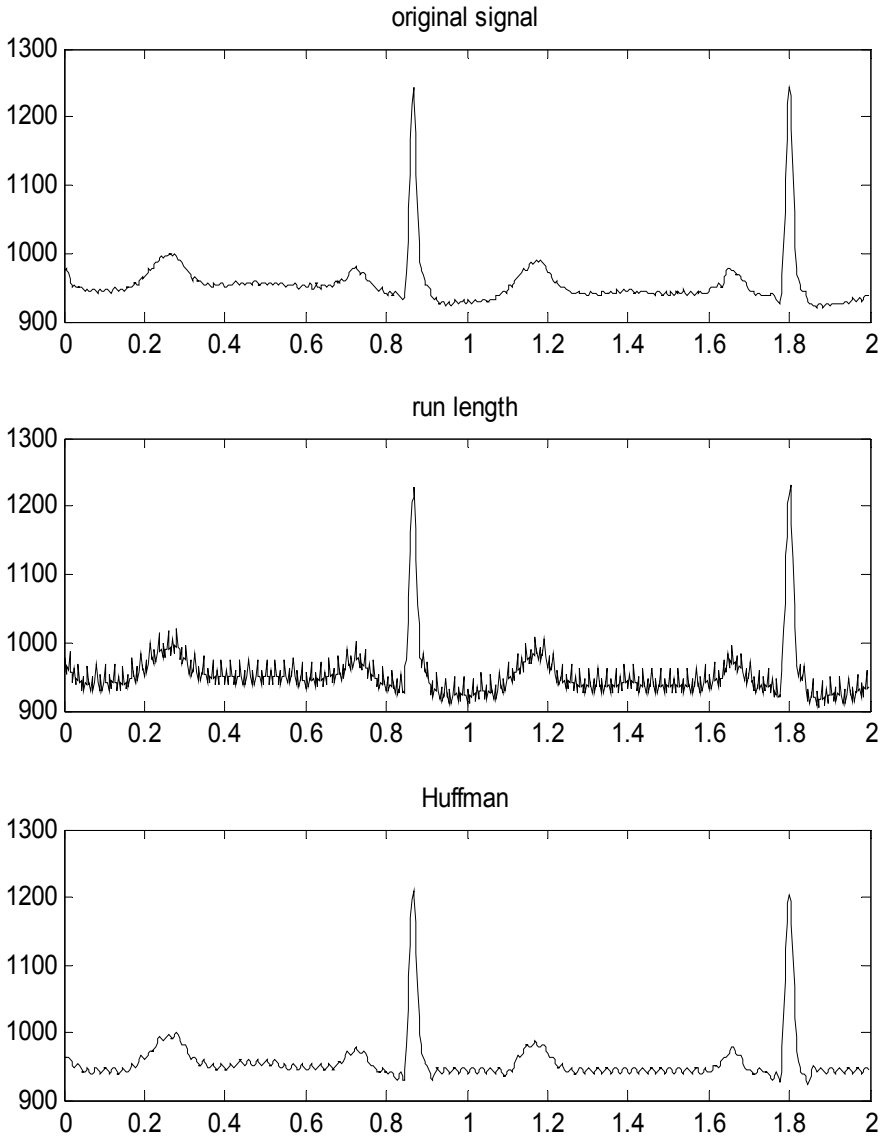


Fig. 9. Record 101: case 1.

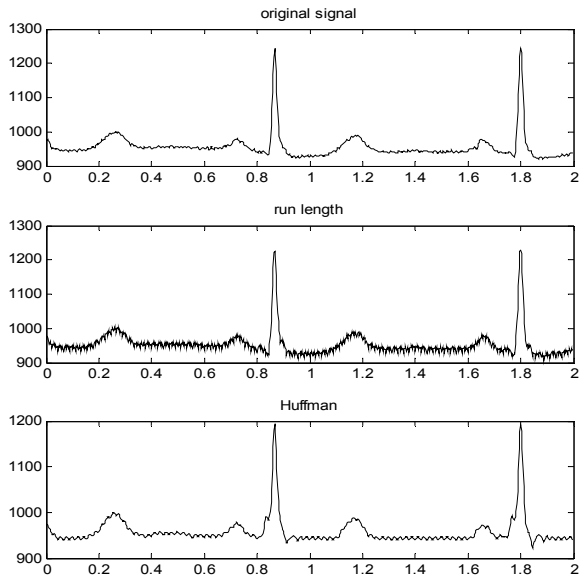


Fig. 10. Record 101: case 2.

Although the record 104 has a pathological behavior opposed to the record 101, it is possible to compress the signal with the same codebook. According to the figure 11, the quality in the signal with RL encoding is better than Huffman encoding. But, the CR is better in the second method

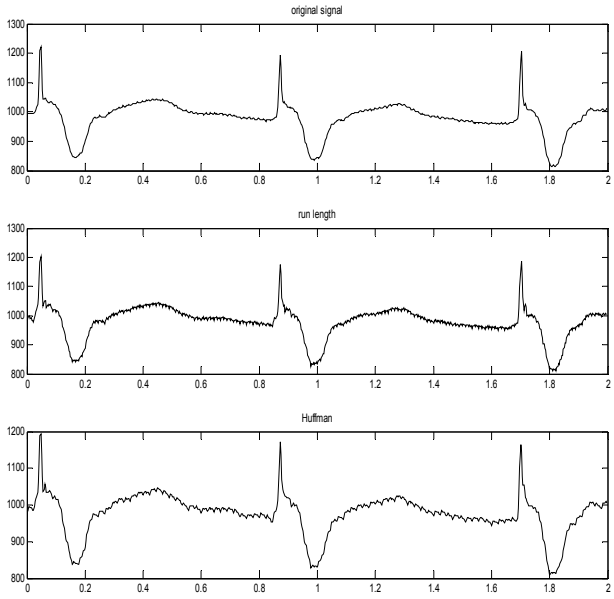


Fig. 11. Record 104. Run length: PRD=0.901 and CR=8.78; Huffman: PRD=1.08 and CR=9.08.

The review results for the records 101, 104 and 107 are presented in table 5.

Record	Method	CR			PRD		
		db6	sym6	bior5.5	db6	sym6	bior5.5
101	Huffman	9,32	9,28	9,32	0,76	0,93	0.98
	RL	8,98	6,91	9,24	1,25	1,25	1.35
104	Huffman	8,98	9,06	9,08	1,41	1,52	1,08
	RL	8,53	8,84	8,78	0,73	0,77	0,90
107	Huffman	7,62	7,78	7,88	1,64	1,85	1,38
	RL	8,22	8,42	9,78	0,99	0,94	0,89

Table 5. Compression results

4. Hardware implementation

In previous works, the authors have developed strategies of hardware processing in the areas of filtering and compression. The first approach presented the Wavelet Transform architecture in decomposing process for denoising of electroencephalographic (EEG) signals (Gaona & Ballesteros, 2005). The second, it presented the comparison in hardware implementation between FIR and IIR (Corredor & Pedraza, 2009). The last project organized the above results in the architecture for biomedical compression based on Discrete Wavelet Transform (DWT) and Run Length encoding (Ballesteros et al., 2010).

4.1 Model of Discrete Wavelet Transform

The Discrete Wavelet Transform is composes by two stages: the convolution of the input signal by the wavelet base and the subsampling process. The convolution is performed by the equations:

$$y_1[n] = \sum_{i=1} h_i x[n-k] \text{ for } \quad \text{k=1,2,.....m-1} \tag{9}$$

and

$$y_2[n] = \sum_{i=1} g_i x[n-k] \text{ for } \quad \text{k=1,2,.....m-1} \tag{10}$$

In the above equations, $y_1[n]$ and $y_2[n]$ are the outputs of the FIR filters, h_i is the impulse response of the lowpass filter, g_i is the impulse response of the highpass filter and $x[n]$ is the input (signal). The expression [n-k] corresponds to the delay in the input. The value of m depends of the length of the wavelet base. For example, if the base is sym6, then m is equal to twelve.

The subsampling is calculated, according to:

$$y_3[n] = y_1[2n] \tag{11}$$

$$y_4[n] = y_2[2n] \tag{12}$$

To implement the equations (9) to (12), we propose the following blocks:

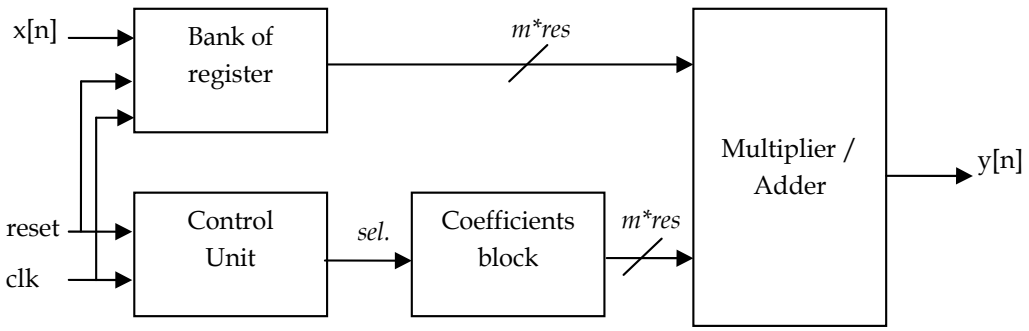


Fig. 12. Architecture of Discrete Wavelet Transform.

- Bank of register: This block stores in a *m*-position register the samples x[n] in an orderly way, depending of its arrival time. At the output is showed the current input value and the last *m-1* inputs. If the reset signal is active, all the outputs are set to zero.
- Coefficients Block: Two *m*-positions memories keep the value of the wavelet filters: *hi* and *gi*.
- Control Unit: it generates the control signal to select hi or gi. Because the subsampling process eliminates the half of operations, it is most efficient to calculate only the half of operations, it means every two cycles of clock.
- Multiplier / Adder: it computes the mathematics operations of the equations (9)-(10) and (11)-(12).

In figure 12, *m* corresponds to the length of the FIR filter; *sel* is the bit of selection between the low-pass or high-pass filter and *res* is the number of bits of every input data.

4.2 Model of thresholding and encoding

To complete the compression model, two stages are added to the previous one: the thresholding and the encoding.

- Thresholding block: A comparison between a threshold value and the data from the low-pass or high-pass filter is doing. It computes the hard rule and the data with a lower value than the threshold are modified to zero.
- Encoding block: It performs the run-length method. At the output of this block there are two signals that go to a memory. The first signal indicates the position in the memory where the second signal ought to be stored.

4.3 Results: Compression and filtering on FPGA

In Figure 13 we present the RTL of the Multiplier/Adder for eight coefficients (m=8). The eight multipliers and seven adders were modelled.

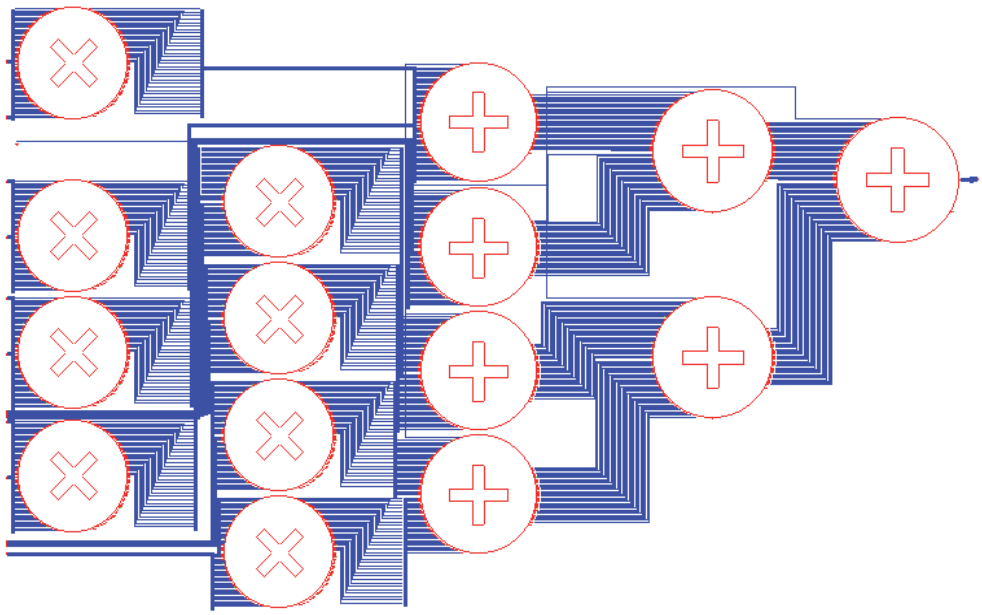


Fig. 13. RTL: Multiplier/Adder block.

In Figure 14, the RTL of the thresholding stage is presented. The comparator is the main structure of the thresholding block is. Its input is the output of the Multiplier/Adder block.

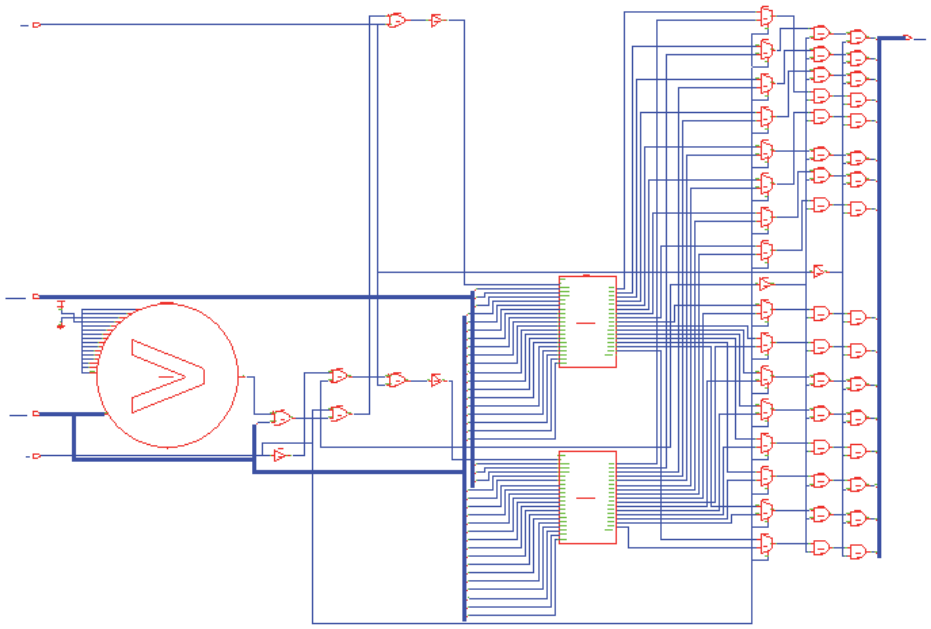


Fig. 14. RTL: Thresholding block.

The architecture was modeled in a FPGA Spartan 3XC3S200. It obtained the following performance:

Maximum Frequency	% Slices	% IOBs
142MHz	0.05%	7%

Table 5. Performance.

5. Conclusion

In this chapter we present the basic idea about the filtering and compression of biomedical signals with the Discrete Wavelet Transform. Although many works have been develop in this theme, we suggest a simple way to select the best parameters and we propose a novel static Huffman encoding for the compression of ECG signals.

For the younger researchers, the methodology in the selection of decomposition-thresholding parameters has been presented with a real case of SEMG signal. For the researchers with more experience in this area, we describe a novel method to generate static codebooks that it can be used in signals with different characteristics; this theme could be explored further.

Finally, the real time processing with the Discrete Wavelet Transform in filtering and compression of biomedical signals is conceived on FPGAs. The advantage of hardware solutions over software is the time of response which is lower in the first. The easy math of the DWT allows very rapid prototyping.

6. Acknowledgment

This work was supported in part by the University Military Nueva Granada under Grant ING290 and ING641.

7. References

Ballesteros, D.M. ; Gaona, A.E. (2008). Deteccion de actividad muscular en registros EMG superficiales en aplicaciones de compresion de dato, *IFMBE Proceedings IV Latin American Congress on Biomedical Engineering 2007*, Vol. 1, Part 1, pp. 25-29, Springer-Verlag Berlin Heidelberg.

Ballesteros, D.M .; Gaona, A.E. (2008). Modelo de Compresion de Señales SEMG para DSP utilizando análisis multi-resolucion, *Proceedings IEEE ANDESCON*, Cusco, Peru, Oct. 15-17, 2008.

Ballesteros, D.M .; Gaona, A.E. (2009). Lossless encoders in Compression of Arrhythmia signals, *Proceedings of the 12th IASTED International Conference on Intelligent Systems and Control*, ISBN 978-0-88986-814-4, Boston, USA, Nov 2-4, 2009.

Ballesteros, D.M .; Gaona, A.E. (2010). Multi-resolution analysis and lossless encoders in the compression of electrocardiographic signals. Rev. Vision Electronica, Vol. 5, 2010, ISSN 1909-9746.

Ballesteros, D.M.; Moreno, D.M.; Gaona, A.E. (2010). Compression of Biomedical Signals on FPGA on by DWT and Run-length, *Proceedings IEEE ANDESCON*, ISBN 978-1-4244-6740-2, Bogota, Colombia, Sep. 15-17, 2010.

Burrus, C. ; Gopinath, R. ; Guo, H. (1998). Introduction to Wavelets and Wavelet Transforms. Prentice Hall. 1998. Pags: 1-40.

Corredor, O.F.; Pedraza, L.F.; Hernandez, C.A. Design and Implementation of Digital Filters. Rev. Vision Electronica, Vol. 3, 2009, ISSN 1909-9746.

Chen, J.; Ma, J.; Zhang, Y.; Shi, X. A Wavelet-Based ECG Compression Algorithm Using Golomb Codes, *Proceedings of International Conference on Communications, Circuits and Systems*, pp. 130-133, ISBN 0-7803-9584-0, June, 2006.

Donoho, D.L. ; Johnstone, I.M. (1994). Threshold selection for wavelet shrinkage of noisy data, *Proceedings of the 16th Annual International Conference of the IEEE,* pp. A24-A25, ISBN 0-7803-2050-6/94, Baltimore, MD, USA, Nov 3-6, 1994

Donoho, D.L. (1995). De-noising by soft-thresholding. *IEEE Transactions of Information Theory,* Vol. 41, Issue: 3, (May 1995), pp. 613-627, ISSN 0018-9448

Gaona, A.E.; Ballesteros, D.M. (2005). Architecture for denoising EEG signals based on Discrete Wavelet Transform, *Proceedings 12th International Workshop on Systems, Signals & Image Processing*, ISBN 0-907776-20-5, Chalkida, Greece, Sep. 22-24, 2005.

Hankerson, D.C.; Harris, G. & Johnson, P. (2005*). Introduction to Information Theory and Data Compression*, Taylor & Francis e-Library, ISBN 1-58488-313-8.

Huffman, D.A. (1952). A Method for the Construction of Minimum- Redundancy Codes, *Proceedings of the I.R.E*, Vol. 40(9), pp. 1098-1101, Sep, 1952. Available from http://compression.ru/download/articles/huff/huffman_1952_minimum-redundancy-codes.pdf

Mallat, S. (1998). *A Wavelet Tour of Signal Processing,* Elsevier, ISBN 0-12-466606-X, USA.

Quian J. (2000). Denoising by Wavelet Transform. Rice University, Department of Electrical Engineering. Available from http://www.daimi.au.dk/~pmn/spf02/CDROM/pr1/Litteratur/Denoising%20by%20wavelet%20transform.pdf

Smith, S. (2003). *Digital Signal Processing : A practical guide for engineers and scientists*, Elsevier, ISBN 0-75067444-X, USA.

3

Discrete Wavelet Transform Based Selection of Salient EEG Frequency Band for Assessing Human Emotions

M. Murugappan, R. Nagarajan and S. Yaacob
Intelligent Signal Processing (ISP) Research Cluster,
Universiti Malaysia Perlis (UniMAP)
Seriab, Kangar, Perlis
Malaysia

1. Introduction

In recent years, assessment of human emotions from Electroencephalogram (EEG) has one of the active research areas in developing intellectual man-machine interfaces. In this work, we considered two different frequency ranges of EEG signals such as alpha to gamma (8 Hz – 64 Hz) and delta to gamma (0 Hz – 64 Hz) for classifying five emotions (disgust, happy, surprise, fear, and neutral). EEG signals are collected using 64 channels from 20 subjects in the age group of 21~39 years. The raw EEG signals are preprocessed using *Surface Laplacian* filtering and a set of conventional and proposed statistical features are derived using Discrete Wavelet Transform (DWT). Two linear classifiers (K Nearest Neighbor (KNN) and Linear Discriminant Analysis (LDA)) are used to map the statistical features into corresponding emotions. One of the proposed features derived by using "db8" wavelet function under alpha to gamma band gives the maximum average classification rate of 86.55% using KNN and 82.70% using LDA compared to delta to gamma band. In addition, the proposed features gives higher classification rate compared to conventional features. Finally we present the average classification accuracy and subsets of discrete emotion classification rates for justifying the performance of our emotion recognition system.

The nonverbal communication through emotions, intentions and affective states are the current key areas of research on developing intellectual man-machine systems. Manifestations of emotional states are normally straightforward to detect and understand by humans, as these are reflected in both voice and body languages (Adler and Rodman 2003; Pease and Pease 2004). In recent years, the research efforts in Human Computer Interaction (HCI) are focused on empowering computers to understand human emotions. Most of the efforts have been dedicated to the design of user-friendly and ergonomic systems by means of innovative interfaces such as voice, vision, and gestures. Many literature works have been reported on emotion recognition using facial expressions and speech modalities (Chen and Huang 2000; Daabaj 2002; Hongo et al. 2000; Massaro 2000). These conventional methods of assessing emotions thorough the speech and the facial expressions of a subject are purposefully expressed and it can be more easily concealed by other subjects (Takahashi 2004). Indeed, it cannot be used for those people who have suffered from severe motor disabilities, amyotrophic lateral sclerosis, paralysis, and introverted characters.

Another possible approach for emotion recognition is done by using physiological signals (EEG (Electroencephalogram), ECG (Electrocardiogram), EMG (Electromyogram), SCR (Skin Conductance Resistance), ST (Skin Temperature), HR (Heart Rate) and RR (Respiration Rate)). There are many works reported on physiological signals based emotion recognitions (Hai-Rong et al. 2008; Takahashi and Tsukaguchi 2003; Wagner et al. 2005; Yongjin and Ling 2005). Furthermore, the physiological response of individual subjects will not be concealed by the physiological responses of other subjects. Compared to all these physiological signals, EEG plays a major role on detecting the emotion directly from the brain at higher temporal and spatial resolution. Furthermore, the brain activity is naturally expected to precede the muscular and vascular activities. Several approaches have been reported by different researchers on assessing the emotional changes from EEG signals (Takahashi and Tsukaguchi 2003; Teixera et al. 2009). More details on the automatic emotion recognition using physiological signals and EEG as well as more complete list of reference can be found in (Murugappan et al. 2010).

2. Related work

Rhythmic activity is a fundamental property of neural elements and is organized in complex patterns depending on the state of the brain. The oscillatory nature of EEG signals varies according to the role in variety of brain operations, including the aspects of emotions, perception, cognition and action. The EEG signals are primarily categorized into 5 different frequency bands: delta (0-4) Hz, theta (4-8) Hz, alpha (8-16) Hz, beta (16-32) Hz, and gamma (32-64) Hz. It is very difficult to find the specific region on skull where the brain activity is sufficiently high to detect an emotional state. Tekell, et.al (Min et al. 2005) reports that, most of the meaningful information about emotional changes is found in the frequency below 30 Hz EEG signals.

Table 1 shows the list of previous works on emotion recognition using different EEG frequency bands features. Several studies have yielded results on the alpha frequency band, which is more prominent rhythm for determining the emotional states in terms of amplitude and frequency characteristics of EEG (Choppin 2000; Robert et al. 2008; Yuan-Pin et al. 2007). The first work on classifying discrete emotions (joy, anger, sad, and neutral) using alpha band power is reported in (Yuan-Pin et al. 2007). Normally, the high frequency EEG waves namely beta and gamma frequency ranges play a vital role on two-way (agree (valence) – disagree (arousal)) classification of emotions (Teixeira and Vinhas 2008; Teixera et al. 2009). The gamma band of oscillation is used for accessing information about emotional consciousness of the person (Mu and Bao-Liang 2009).

In (Gunsel et al. 2006), they have considered the information of all the five frequency bands for classifying emotions in two dimensional (calm-exciting) and three dimensional (calm-exciting-neutral) emotions and achieved an average classification rate of 60% for two and 50 % for three dimensions respectively.

The depression of alpha rhythm on right hemisphere is found on “positive” emotions (pleasant) and the frontal electrodes in left hemisphere are responding to “negative” emotions (unpleasant) under beta rhythm (Gunsel et al. 2006). Indeed, the investigation on theta band also carried out for both positive and negative emotions. On the other hand, four frequency bands such as: delta, theta, alpha and beta band powers are used for classifying discrete emotions (Khalili and Moradi 2009; Schaaff and Schultz 2009a; Schaaff and Schultz 2009b). Heraz, et.al., has developed the emomental (emotional & mental) agent for *Intelligent*

Tutoring System and used the energy of four different frequency bands (delta, theta, alpha, and beta) in classifying 8 emotional-mental states such as: anger, boredom, confusion, contempt, curious, disgust, eureka and frustration (Heraz et al. 2007). Among all the previous works, the maximum emotion classification rate of 93.25% is reported on classifying two emotions (happy and sad) (Mu and Bao-Liang 2009).

Reference	No of Subjects	No of Electrodes	Stimuli	Two Dimensional Emotions		Max % CR
				No's	Types	
Alpha Frequency Band						
(Choppin and 2000)	31	4	Visual	2	Valence - Arousal	64
(Yuan-Pin et al. 2007)	5	32	Audio	4	Joy, Anger, Pleasure, Sad	69.69
(Robert et al. 2008)	10	64	Visual	2	Valence – Arousal	71
Other Frequency Bands						
(Murugappan et al. 2010)	- -	3	Visual	3	Joy, Sad, Neutral	74
(Mu and Bao-Liang 2009)	10	62	Visual	2	Happy, Sad	93.25
(Khalili and Moradi 2009)	5	10	Visual	3	Calm, Positively excited and negatively excited	76.67
(Schaaff and Schultz 2009b)	5	4	Visual	3	Pleasant, Neutral, and Unpleasant	66.7
(Schaaff and Schultz 2009a)	5	4	Visual	3	Pleasant, Neutral, and Unpleasant	47.11
(Heraz et al. 2007)	17	3	Visual	8	4 emotional and 4 mental states	82.27
(Petrantonakis and Hadjileontiadis 2009)	16	4	Visual	6	Happy, Anger, Fear, Disgust, Sad, and Surprise	84.72
Five Frequency Bands (Delta to Gamma)						
(Jacko et al. 2009)	28	3	Visual	2	Valence- Arousal	74.11

Table 1. List of previous works on emotion recognition using EEG signals with its classification rate

Determining the rhythmic brain activity for different emotional stimulus is a promising area of research for providing an in-depth picture of how the brain frequencies vary for different emotions. Most of the previous studies are focused on assessing the two dimensional human emotion (valence-arousal) through different frequency ranges of EEG signals. Only few works have been proposed on discrete human emotion classification. Furthermore, there is no defined range of frequency nominated for improving the emotion classification rate. Hence, it is necessary to determine the range of frequency which gives the maximum classification accuracy on assessing discrete emotions. The maximum mean emotion classification rate of 84.72% is achieved using alpha and beta band for classifying six emotions (Petrantonakis and Hadjileontiadis 2009). In addition, most of the previous works have considered less number of subjects for developing the emotion recognition system.

In our earlier work, we have considered alpha band features for classifying the discrete emotions. The maximum mean emotion classification rate of 78.043% is achieved using KNN (Murugappan et al. 2009b). In this work, we consider two different frequency bands (alpha to gamma and delta to gamma) of EEG signals for classifying discrete emotions using a set of conventional and proposed features.

3. Research methodology

3.1 EEG data acquisition

3.1.1 Emotion elicitation method

This section describes the acquisition of EEG signals for emotion assessment experiments. Emotions can be induced by one of the following ways: (a) visual (images/pictures) (Yongjin and Ling 2005) (b) audio-visual (film clips/video clips) (Takahashi 2004) (c) recalling of past emotional events (d) audio (songs/sounds) (Wagner et al. 2005). Most of researchers are using visual stimuli and audio-visual stimuli for evoking emotions. In our previous work, we have used both visual and audio-visual stimuli for evoking discrete emotions. The result of this study confirms that, audio-visual stimulus performs superior in evoking emotions than visual stimulus (Murugappan et al. 2009a). The main advantage of this method resides in the strong correlation between induced emotional states and the physiological responses. Hence, we have designed an audio-visual induction based protocol for eliciting the discrete emotions in this present work. The structural overview of emotion recognition system using EEG signals is shown in Fig 1. The audio-visual stimulus protocol for Trail 1 of our experiment and self-assessment questionnaires is shown in Fig. 2 and Fig 3. The orders of the emotional video clips are changed in a random manner for other trials (Table 2). Time durations of video/film clips for five trials is given in Table 3. X1 to X5 denote time periods of selected video clips. The time duration of video clips vary from one another. All the video clips are short in time duration and with more dynamic emotional content. The selection of video clips is based on self assessment questionnaires given in Fig 3. Between each emotional stimulus (video clips), a blank screen is shown for 10 sec duration to bring the subject to their normal state and to experience a calm mind. As a result of the self-assessment report of subjects, totally five trials for three emotions (happy, surprise and disgust) and four trials for two emotions (fear and neutral) are considered for all our future analysis.

3.1.2 Subjects

A pilot panel study is conducted on 25 university students to select any 5 video clips (trials) for each emotion from 115 emotional video clips including from the international standard emotional clips*. The subjects who have undergone for this panel study does not take part in the data collection experiment. Three females and seventeen males in the age group of 21-39 years were employed as subjects in our experiment. Once the consent forms were filled-up, the subjects were given a simple introduction about the research work and the various stages of experiment.

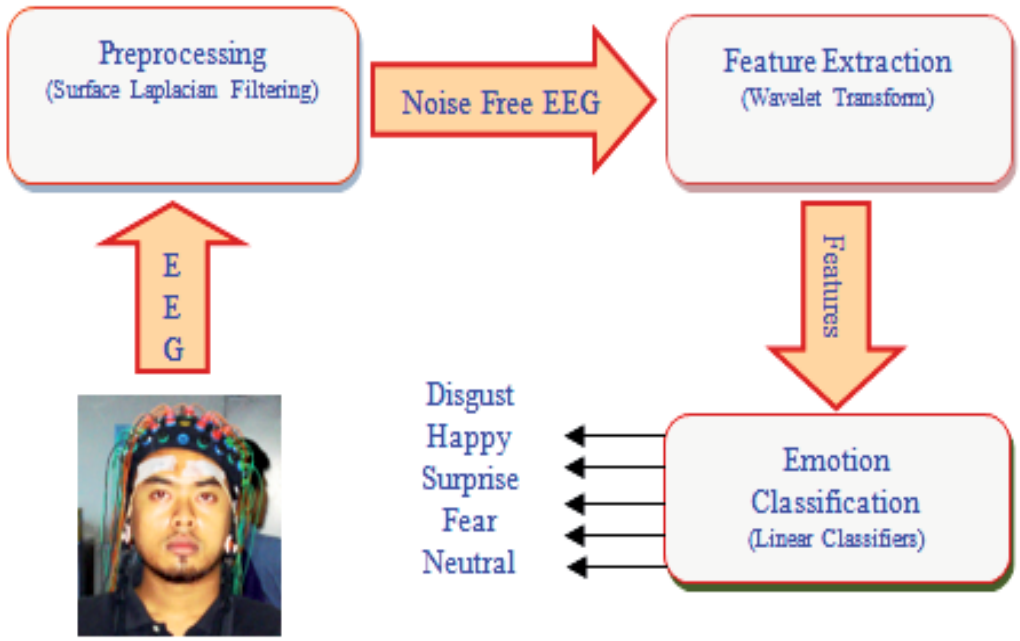


Fig. 1. Emotion Recognition System overview

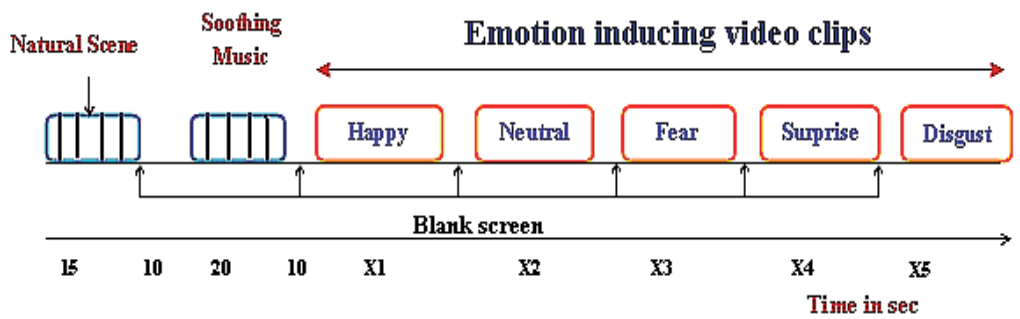
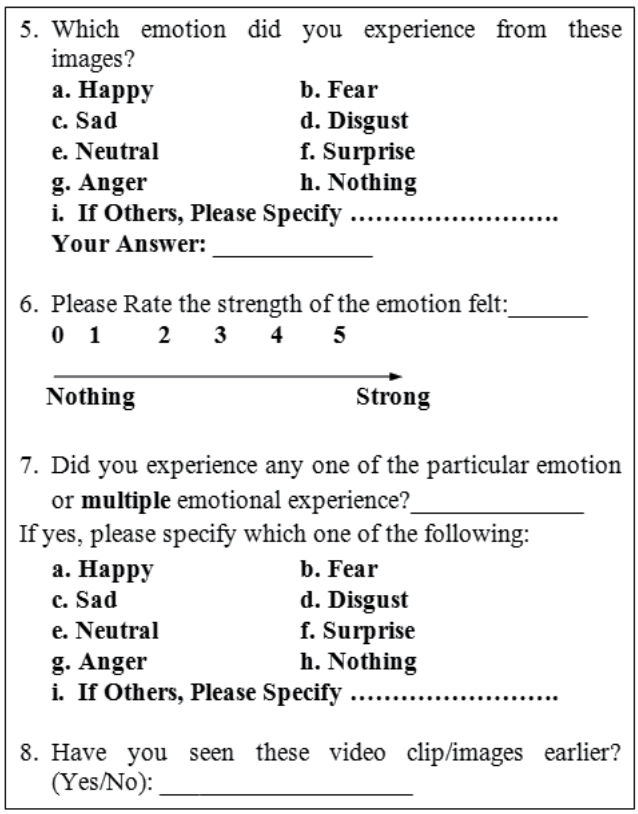


Fig. 2. EEG data acquisition protocol using audio-visual stimuli


5. Which emotion did you experience from these images?
a. Happy **b. Fear**
c. Sad **d. Disgust**
e. Neutral **f. Surprise**
g. Anger **h. Nothing**
i. If Others, Please Specify …….
Your Answer: _____

6. Please Rate the strength of the emotion felt:_____
0 1 2 3 4 5
Nothing **Strong**

7. Did you experience any one of the particular emotion or **multiple** emotional experience?_____
If yes, please specify which one of the following:
a. Happy **b. Fear**
c. Sad **d. Disgust**
e. Neutral **f. Surprise**
g. Anger **h. Nothing**
i. If Others, Please Specify …….

8. Have you seen these video clip/images earlier? (Yes/No): _____

Fig. 3. Self-assessment questionnaires used in emotion recognition experiment

3.1.3 Experimental setup

The recording of EEG signal is done through Nervus EEG, Iceland with 64 channel electrodes at a sampling frequency of 256 Hz and band-pass filtered between 0.05 Hz and 70 Hz. There are totally 62 active electrodes plus one electrode for ground (Oz) and one for reference (AFz). In addition, we collected the recording of eye blink rate by two EOG (Electroocculogram) electrodes (EOG_L & EOG_R), which are placed above the right and left eyes of the subjects.

Trials	Order of Emotions				
1	Happy	Neutral	Fear	Surprise	Disgust
2	Neutral	Surprise	Happy	Disgust	Fear
3	Disgust	Neutral	Surprise	Happy	Fear
4	Surprise	Happy	Disgust	Fear	Neutral
5	Disgust	Happy	Fear	Neutral	Surprise

Table 2. Order of emotional stimuli used over five trials in emotion recognition experiment

Trials	Duration of emotional stimuli in Minute (min)				
1	0.13	0.58	0.14	0.10	0.59
2	0.59	0.14	0.23	0.10	0.22
3	0.56	0.24	0.12	0.13	0.30
4	0.20	0.10	0.58	0.25	0.44
5	0.59	0.09	0.35	0.57	0.15
Statistical Computations					
	Happy	Neutral	Fear	Surprise	Disgust
Mean	0.136	0.484	0.252	0.142	0.484
Std. Deviation	0.056	0.149	0.080	0.038	0.215

Table 3. Time duration of each emotional stimulus used in emotion recognition experiment

All the electrodes are placed over the entire scalp using International standard 10-10 system (Böcker et al. 1994) (Fig 4). The impedance of the electrodes is kept below 5 kΩ. Between each emotional video clips, under self assessment section, the subjects were informed to answer the emotions they have experienced (Abu Osman et al. 2008). Finally, 5 trials for disgust, happy and surprise emotions and 4 trials for fear and neutral emotions are considered for further analysis. All the signals are collected without much discomfort to the subjects.

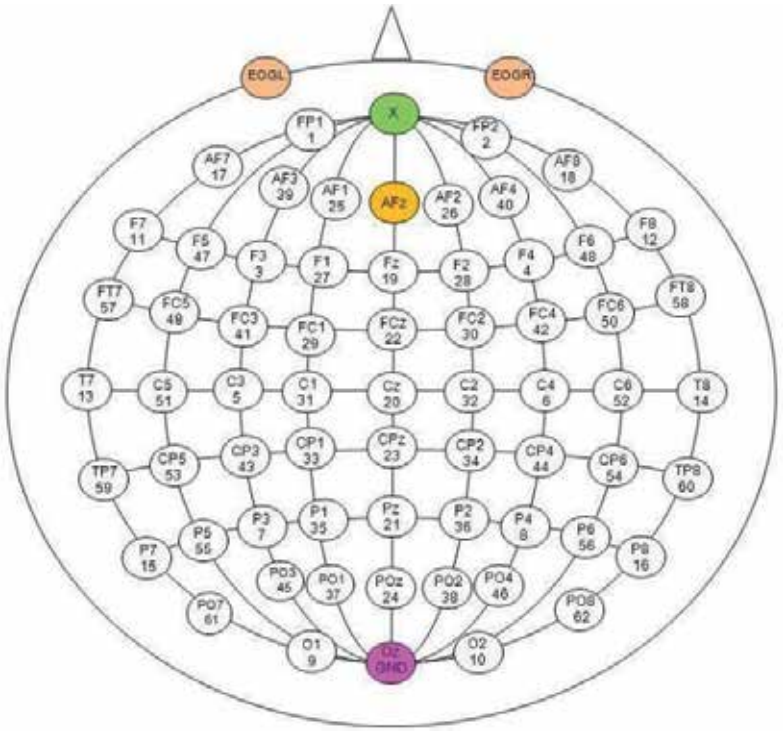


Fig. 4. Placement of electrodes according to International 10-10 system

3.2 Preprocessing and normalization

The recorded EEG signals are usually contaminated with noises (due to power line fluctuations and due to external interferences) and artifacts (due to eye blinks, eye ball rotations and muscular movements). The complete removal of artifacts will also remove some of the useful information of EEG signals. This is one of the reasons why considerable experience is required to interpret EEGs clinically (Gott et al. 1984; Jung et al. 2000). A couple of methods are available in the literature to avoid artifacts in EEG recordings. However, removing artifacts entirely is impossible in the existing data acquisition process. The research methodology of emotion recognition using EEG is shown Fig 5.

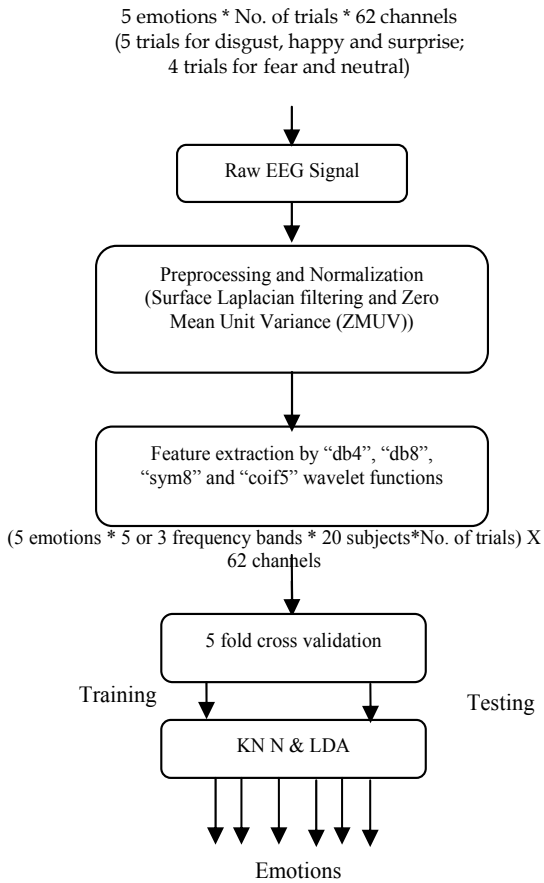


Fig. 5. Overview of emotion recognition procedure

In this work, we used Surface Laplacian (SL) filter for removing the noises and artifacts. The SL filter is used to emphasize the electric activities that are spatially close to a recording electrode, filtering out those that might have an origin outside the skull (Gin-Shin et al. 2005). In addition, it also attenuates the EEG activity which is common to all involved channels in order to improve the spatial resolution of the recorded signal. The neural activities generated by the brain, however, contain various spatial frequencies. Potentially

useful information from the middle frequencies may be filtered out by the analytical Laplacian filters (Xinyi et al. 2008). Hence, the signal "pattern" derived from SL filters is similar to "spatial distribution of source in the head".
The mathematical modeling of Surface Laplacian filter is given as

$$X_{new}(t) = X(t) - \frac{1}{N_E}\sum_{i=1}^{N_E} X_i(t) \quad (1)$$

where X_{new} : filtered signal ; $X(t)$: raw signal ; N_E: number of neighborhood electrodes
The preprocessed signals are normalized using Zero Mean Unit Variance (ZMUV) method. Normalization is used to reduce the individual difference between the subjects and also used to reduce the computational complexity.

3.3 Feature extraction

EEG signals are often quantified based on their frequency domain characteristics. Typically the spectrum is estimated using Fast Fourier Transform (FFT). A fundamental requirement in the FFT based spectral analysis is that the signal is to be stationary. Indeed, the EEG signals cannot be considered as stationary even under short time duration, since it can exhibit considerable short-term nonstationarities (Anderson and Sijerčić 1996). In the EEG based emotion recognition research, the non-parametric method of feature extraction based on multi-resolution analysis of Wavelet Transform (WT) is quite new. The joint time-frequency resolution obtained by WT makes it a good candidate for the extraction of details as well as approximations of the signal which cannot be obtained either by Fast Fourier Transform (FFT) or by Short Time Fourier Transform (STFT) (Mallat 1989; Merzagora et al. 2006). Hence, this method of feature extraction is adopted in this present work.
The non-stationary nature of EEG signals allow us to expand basis functions created by expanding, contracting and shifting a single prototype function ($\Psi_{a,b}$, the mother wavelet), specifically selected for the signal under consideration
The mother wavelet function $\Psi_{a,b}(t)$ is given as

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}}\psi\left(\frac{t-b}{a}\right) \quad (2)$$

where a, b $\in$ R, a>0, and R is the wavelet space. Parameters 'a' and 'b' are the scaling factor and the shifting factor respectively. The only limitation for choosing a prototype function as mother wavelet is to satisfy the admissibility condition (Eqn. 3),

$$C_\psi = \int_{-\infty}^{\infty} \frac{|\Psi(\omega)|^2}{\omega} d\omega < \infty \quad (3)$$

where $\psi(\omega)$ is the Fourier transform of $\psi_{a,b}(t)$.
The time-frequency representation is performed by repeatedly filtering the signal with a pair of filters that cut the frequency domain in the middle. Specifically, the discrete wavelet transform decomposes the signal into an approximation coefficients (CA) and detailed coefficients (CD). The approximation coefficient can be subsequently divided into new approximation and detailed coefficients. This process can be carried out iteratively

producing a set of approximation coefficients and detail coefficients at different levels of decomposition (Parameswariah and Cox 2002).
In this work, four different wavelet functions: "db4", "db8", "sym8" and "coif5" are used for decomposing the EEG signals into five different frequency bands (delta, theta, alpha, beta, and gamma). These wavelet functions are chosen due to their near optimal time-frequency localization properties. Moreover, the waveforms of these wavelets are similar to the waveforms to be detected in the EEG signal. Therefore, extraction of EEG signals features are more likely to be successful (Glassman 2005). In Table 3, A5, D5, D4, D3, and D2 represent the five EEG frequency bands. Table 4 also presents the bandwidth and the frequencies corresponding to different levels of decomposition of EEG signal with a sampling frequency f_s=256 Hz (Parameswariah and Cox 2002).

Frequency Range (Hz)	Decomposition Level	Frequency Bands	Frequency Bandwidth (Hz)
0 - 4	A5	Delta	4
4 – 8	D5	Theta	4
8 – 16	D4	Alpha	6
16 – 32	D3	Beta	18
32 – 64	D2	Gama	32
64 - 128	D1	Noises	64

Table 4. Decomposition of EEG signals into different frequency bands with a sampling frequency of 256 Hz

3.3.1 Frequency band selection

From the literature study in Section 2, most of the researchers have attempted to classify the two dimensional emotions (valence-arousal) by considering the three different frequency ranges: alpha band (8 Hz -16 Hz), beta and gamma band (16 Hz – 64 Hz), and delta to gamma band (0 Hz – 64 Hz). In this work, a first set of experiments are carried out with four conventional features (power, standard deviation, entropy and variance) for classifying distinct emotions (Table 5). We have computed the average classification accuracy and individual classification rate for each feature over two different frequency bands. The result of classification accuracy is reported in Section 4.
In order to improve the classification rate of emotion assessment, we proposed a new feature related to frequency band power ratio called as "*Recoursing Power Efficiency*" (RPE) (Eqn 6) for classifying the emotions. In addition, we also extended the RPE called *Logarithmic Recoursing Power Efficiency* (LRPE) (Eqn 7) and *Absolute Logarithmic Recoursing Power Efficiency* (ALRPE) (Eqn 8) for classifying emotions. Eqn 6 to Eqn 8 show the mathematical equation of computing statistical features of RPE of gamma band using five frequency bands. Similarly, these equations are extended to remaining frequency bands for calculating the statistical features in classifying emotions. One of the major limitations on this area of research is the lack of international standard data base. Hence, in this present work, we compared the efficacy of proposed features with the efficacy of a set of conventional features for estimating the classifiability. Similarly, we can derive the proposed features for three frequency bands by replacing $P_{total\text{-}5b}$ as $P_{total\text{-}3b}$ in Eqn (6) to Eqn (8).

Features	Formula to Compute	Description
Standard Deviation	$SD_j = \sqrt{\frac{\sum_k (d_j(k) - \overline{d_j(k)})}{k-1}}$ $\overline{d_j(k)}$ is the mean value of the wavelet coefficient $d_j(k)$ is the detail wavelet coefficient	Measures the deviation of electrodes potential from its mean value over different emotional EEG signals.
Variance	$V_j = \frac{1}{N}\sum_{k=1}^{N}(d_j(k) - \overline{d_j(k)})^2$	Measures the value variation in electrical potential for various emotions.
Power	$P_j = \frac{1}{N}\sum_{k=1}^{N}(d_j(k)^2)$	Measures the squares of the amplitude of EEG signal
Entropy	$H = -\sum_{j=1}^{M} p_j \log p_j$; where $p_j = \frac{E_j}{E_T}$ E_j = Energy at j[th] frequency band of decomposition E_T = Energy of all frequency band of decomposition	Measures the useful information (nonlinearity) about the EEG signal for emotion from the intrusive noise.
J = level of wavelet decomposition; k = No of wavelet coefficients varies from 1 to N		

Table 5. Conventional statistical features used for emotion recognition and their description

$$P_{total-5b} = P_{delta} + P_{theta} + P_{alpha} + P_{beta} + P_{gamma} \tag{4}$$

$$P_{total-3b} = P_{alpha} + P_{beta} + P_{gamma} \tag{5}$$

$$RPE_{gamma} = \frac{P_{gamma}}{P_{total-5b}} \tag{6}$$

$$LRPE_{gamma} = \log 10\left[\frac{P_{gamma}}{P_{total-5b}}\right] \tag{7}$$

$$ALRPE_{gamma} = abs\left(\log 10\left[\frac{P_{gamma}}{P_{total-5b}}\right]\right) \tag{8}$$

where P= frequency band power; $P_{total\text{-}5b}$= total power of 5 frequency bands (delta to gamma; $P_{total\text{-}3b}$= total power of 3 frequency bands (alpha to gamma)

3.3.1.1 Motivation for new statistical feature proposal

In conventional analysis, researchers have computed the total wavelet energy by considering the complete range of EEG signal frequencies irrespective to the selected frequency bands. For example, if the researchers considered the alpha to gamma frequency (8 Hz – 64 Hz) bands for their analysis means, the total wavelet energy is computed between delta to gamma (0 Hz – 64 Hz) compared to the total energy of the selected frequency range (8 Hz – 64 Hz).

According to the literature, the authors certainly believes that, the low frequency details (0 Hz – 8 Hz) gives an useful information for some special applications such as sleep stage classification (0.5 Hz – 4 Hz) and mental fatigue analysis and driver drowsiness detection (0.5 Hz – 8 Hz). Therefore, the consideration of entire frequency range may yield poor classification results due to the additional and non-informative details than considering the selected frequency ranges. Thus, this present work has adopted the selected frequency range for computing the total wavelet energy for deriving the statistical features further than entire frequency range. Therefore, the effectiveness of newly proposed statistical features on two different combinations of frequency bands is considered for classifying the emotions.

3.4 Emotion classification

In this work, we have employed two simple classifiers such as Linear Discriminant Analysis (LDA) and K Nearest Neighbor (KNN) for classifying the discrete emotions. Classification accuracy, representing the percentage of correctly classified instances, has been adopted to quantify the performance of KNN and LDA.

3.4.1 Linear Discriminant Analysis (LDA)

Among these two classifiers, LDA provides extremely fast evaluations of unknown inputs performed by the calculations of distances between a new sample and mean of training data samples in each class weighed by their covariance matrices. Indeed, LDA is of very simple but elegant approach to classify various emotions. A linear discriminant analysis tries to find an optimal hyper plane to separate five classes (here, disgust, happy, surprise, fear and neutral emotions). Besides the training and testing samples, LDA does not require any external parameters for classifying the discrete emotions.

3.4.2 K Nearest Neighbour (KNN)

KNN is also a simple and intuitive method of classifier used by many researchers typically for classifying the signals and images. This classifier makes a decision on comparing a new labeled sample (testing data) with the baseline data (training data). In general, for a given unlabeled time series X, the KNN rule finds the K "closest" (neighborhood) labeled time series in the training data set and assigns X to the class that appears most frequently in the neighborhood of k time series. There are two main schemes or decision rules in KNN algorithm, the similarity voting scheme and majority voting scheme (Chaovalitwongse et al. 2007).

In our work, we used the majority voting for classifying the unlabeled data. It means that, a class (category) gets one vote, for each instance, of that class in a set of K neighborhood samples. Then, the new data sample is classified to the class with the highest amount of votes. This majority voting is more commonly used because it is less sensitive to outliers. However, in KNN, we need to specify the value of "K" closest neighbor for emotions classification. In this experiment, we try different "K" values ranging from 2 to 6. The use of linear classifiers may perform well over the non-linear classifiers on classifying the EEG signals. Thus, this work uses aforementioned two linear classifiers for classifying the human emotions.

4. Experimental results and discussions

The main motivation on this present work is to improve the emotion classification rate by selecting the salient EEG frequency bands. Among all twenty subjects, we sample and preprocess the total of 460 EEG epochs from five discrete emotions (20 subjects x 5 emotions (5 trials for 3 emotions and 4 trials for 2 emotions). The number of data points in each epoch depends on the time duration of video clips. In our experiment, the time duration of video clips vary from one another. The next stage is to train the KNN classifier with a best value of K while LDA classifier directly works for classifying the emotions. The classification ability of a statistical feature set can be measured through classification accuracy by averaging five times over a 5 fold cross-validation. The basic stages of 5 fold cross-validation include: (a) total number of samples are divided into 5 disjoint sets (b) 4 sets are used for training and 1 set is used for testing (c) repeat stage (b) for five times and each time the data set is permuted differently. This method reduces the possibility of deviations in the results due to some special distribution of training data and test data. This also ensures that the system is tested with different samples that which are not used in training.

The major limitation of the research on emotion recognition using EEG is the lack of international standard database for comparing the efficacy of emotion classification through different approaches. Hence, most of the researchers have used the dissimilar conventional statistical features for classifying emotions through EEG. Currently, the researchers are focusing on determining new statistical features for achieving maximum emotion classification rate. For this purpose, we have proposed the feature based on frequency band power ratio called *Recoursing Power Efficiency (RPE)*. We also expanded this statistical feature into another two forms namely: *Logarithmic Recoursing Power Efficiency (LRPE)* and *Absolute Logarithmic Recoursing Power Efficiency (ALRPE)*.

In general, these features, measure the ratio of power of selected frequency band to the total power of frequency bands. Table 6 to Table 9 shows the average classification accuracy of emotions using conventional and proposed features on two different frequency bands using KNN and LDA respectively. From Table 6 & Table 8, we found that, the conventional feature (entropy) performs better over proposed features on 5 frequency bands. In addition, KNN outperforms LDA by giving a maximum average classification accuracy of 83.04% than the proposed feature (ALRPE) of 79.39%.

In this frequency rage, "sym8" wavelet function captures the chaotic nature of EEG signal under different emotions. The individual classification accuracy for the conventional and proposed feature corresponding to the delta to gamma band feature which gives the maximum average classification accuracy on KNN and LDA based classification (Table 6 and Table 8) are presented in Table 10. Basically, entropy is a non-linear feature and it

measures the amount of non-linearity present in an EEG signal. The non-linearity is probably due to the noise but may also due to the complex dynamics.

Wavelet	K	KNN based Classification of Emotions using Delta to Gamma band (0 Hz – 64 Hz) in %						
		Conventional Features				Proposed Features		
		ENT	**POW**	**STD**	**VAR**	**RPE**	**LRPE**	**ALRPE**
db4	6	82.35± 1.14	68.13± 1.36	71.91± 3.08	68.91± 1.23	67.47± 1.80	78.91± 1.30	78.47± 1.28
db8	6	81.96± 2.09	68.30± 1.21	73.52± 1.46	68.04± 1.13	68.21± 1.09	77.87± 1.41	78.60± 1.46
sym8	6	**83.04± 1.56**	67.78± 1.39	72.65± 2.46	68.78± 2.18	67.34± 0.74	78.00± 1.48	**79.39± 0.97**
coif5	6	82.52± 1.09	67.61± 1.51	76.61± 1.31	68.83± 1.44	65.91± 0.77	76.87± 2.05	78.08± 2.58

Table 6. KNN based classification of emotions using conventional and proposed features on delta to gamma frequency bands

Wavelet	K	KNN based Classification of Emotions using Alpha to Gamma Bands (8 Hz – 64 Hz) in %						
		Conventional Features				Proposed Features		
		ENT	**POW**	**STD**	**VAR**	**RPE**	**LRPE**	**ALRPE**
db4	5	69.93± 2.27	55.87± 2.79	62.46± 1.76	56.01± 3.06	82.32± 2.41	83.04± 2.13	82.25± 1.97
db8	5	70.15± 2.11	55.44± 2.94	62.10± 3.75	55.58± 3.31	82.03± 1.79	**86.55± 0.75**	81.23± 0.87
sym8	5	71.16± 3.04	57.03± 1.42	62.17± 1.60	56.59± 2.29	81.45± 0.83	81.30± 2.89	81.67± 1.04
coif5	5	**81.01± 0.60**	55.07± 3.17	65.58± 3.42	55.94± 4.18	82.97± 1.10	82.97± 1.11	82.75± 2.63

Table 7. KNN based classification of emotions using conventional and proposed features on alpha to gamma frequency bands

Wavelet	LDA based Classification of Emotions using Delta to Gamma Bands (0 Hz – 64 Hz) in %						
	Conventional Features				Proposed Features		
	ENT	**POW**	**STD**	**VAR**	**RPE**	**LRPE**	**ALRPE**
db4	**81.44± 2.14**	52.17± 4.91	65.09± 2.74	51.87± 1.65	71.00± 1.95	78.04± 0.84	77.04± 1.88
db8	80.52± 1.61	53.74± 4.76	63.44± 0.28	51.87± 3.04	71.00± 1.16	77.87± 0.96	77.48± 0.11
sym8	79.74± 2.04	51.35± 2.76	64.78± 2.26	50.04± 2.89	70.91± 1.03	77.39± 1.11	77.57± 1.97
coif5	80.30± 1.29	51.52± 1.82	63.83± 1.94	50.35± 3.30	71.04± 1.65	77.96± 0.92	**78.26± 0.61**

Table 8. LDA based classification of emotions using conventional and proposed features on delta to gamma frequency bands

Wavelet	LDA based Classification of Emotions using Alpha to Gamma Bands (8 Hz – 64 Hz) in %						
	Conventional Features				Proposed Features		
	ENT	**POW**	**STD**	**VAR**	**RPE**	**LRPE**	**ALRPE**
db4	72.90±1.41	63.41±3.10	71.74±2.15	64.57±1.88	76.88±2.05	75.58±1.92	75.73±2.40
db8	72.68±2.38	63.91±2.63	71.59±1.50	62.39±2.29	77.03±1.96	76.38±2.11	75.65±2.13
sym8	**73.55±1.51**	63.19±3.76	71.67±2.27	64.57±4.22	77.61±2.71	76.81±2.38	76.01±2.14
coif5	70.51±2.65	63.04±2.03	71.88±1.21	63.70±2.39	**82.70±1.96**	74.93±0.59	76.74±2.36

Table 9. LDA based classification of emotions using conventional and proposed features on alpha to gamma frequency bands

Classifier	Method	Feature	Wavelet	K	Disgust	Happy	Surprise	Fear	Neutral
KNN	Conventional	ENT	sym8	5	92	85	67	68.75	96.25
	Proposed	ALRPE	sym8	5	93	84	71	70	78.75
LDA	Conventional	ENT	db4		94	83	61	72.5	86.67
	Proposed	ALRPE	coif5		92	87	68	68.75	85

Table 10. Individual classification rate of emotions corresponds to the maximum classification rate of conventional and proposed features on delta to gamma frequency bands

Classifier	Method	Feature	Wavelet	K	Disgust	Happy	Surprise	Fear	Neutral
KNN	Conventional	ENT	coif5	5	90	85	71.67	66.67	95.83
	Proposed	LRPE	db8	5	91.67	87.67	75	72.92	97.92
LDA	Conventional	ENT	sym8		83.33	61.667	66.67	52.083	85.417
	Proposed	RPE	coif5		93.33	63.33	81.67	70.83	93.75

Table 11. Individual classification rate of emotions corresponds to the maximum classification rate of conventional and proposed features on alpha to gamma frequency bands

However, the proposed feature, the Logarithmic Recoursing Power Efficiency (LRPE) performs better over the conventional features in 3 frequency band analysis on KNN and Recoursing Power Efficiency (RPE) on LDA. This statistical parameter achieves a maximum of 86.55% classification rate than 81.01% on conventional feature (entropy). Furthermore, LDA based classification gives the maximum classification rate of 82.70% on proposed feature (RPE) compared to the conventional features. The individual classification accuracy for the conventional and proposed feature corresponding to the alpha to gamma band feature which gives the maximum average classification accuracy on KNN and LDA based classification (Table 7 and Table 9) are presented in Table 11. The classification rate of this emotion recognition method shows that, the EEG data collected under audio-visual stimuli environment has useful and enough information to assess discrete emotions from it.
In addition, the proposed statistical feature (LRPE) derived from "db8" wavelet function achieve the maximum average classification rate of 86.55% compared to that of all other statistical features. Therefore, these results confirm that, the characteristic pattern of EEG under different emotional states are seems to be similar to the characteristic pattern of "db8" wavelet function. Hence, retrieval of information from the EEG signal from "db8" is highly possible than those of other wavelet functions. Though the level of improved classification rate of proposed feature is smaller than the conventional features, it confirms that the future work on deriving new statistical features based on "power" can improve the emotion recognition rate to a remarkable level. All these analysis are performed in offline using MATLAB 7 software.

5. Conclusion

A discrete emotion recognition system to recognize emotions from selected frequency range of EEG signals using new statistical features is presented in this paper. The range of frequency selected by the newly proposed feature gives a maximum average and individual classification rate compared to other conventional features. Therefore the extracted features successfully capture the emotional changes of the subject through their EEG signals regardless of the user's cultural background, race, and age. Thus, the combination of wavelet features and non-linear classifier greatly improved the emotion classification rate of the proposed system over previous works. In addition, the accuracy of emotion classification reported on alpha to gamma band is higher compared to the other frequency bands considered by previous researchers with a maximum number of subjects with larger category of emotions (Table 1). This study is ongoing to involve different classification

algorithms in order to track the emotional status of brain activation during audio-visual stimuli environment.

6. Future work

This work can be extended to localize a brain region related to different emotional states. Thereby, the removal of inactive electrodes can be considered for developing more reliable/portable emotion recognition system. If so, the emotion recognition algorithm will be implemented into real life human-computer interaction applications. In addition, the dimensionality of the feature can be reduced by using the artificial intelligence methods to further improve the emotion classification rate.

7. Acknowledgement

This work is supported by Ministry of Higher Education, Malaysia and Universiti Malaysia Perlis (UniMAP) Cluster Fund, Malaysia. Grant Code: 9001-00191 & 9003-00214.

8. References

Adler, R. B., and Rodman, G. (2003). "Understanding Human Communication." Oxford University Press.

Anderson, C. W., and Sijerčić, Z. "Classification of EEG signals from four subjects during five mental tasks." *Proceedings of the Conference on Engineering Applications in Neural Networks (EANN'96)*, 407-414.

Böcker, K. B. E., van Avermaete, J. A. G., and van den Berg-Lenssen, M. M. C. (1994). "The international 10–20 system revisited: Cartesian and spherical co-ordinates." *Brain Topography*, 6(3), 231-235.

Chaovalitwongse, W. A., Ya-Ju, F., and Sachdeo, R. C. (2007). "On the Time Series K-Nearest Neighbor Classification of Abnormal Brain Activity." *Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on*, 37(6), 1005-1016.

Chen, L. S., and Huang, T. S. "Emotional expressions in audiovisual human computer interaction." *Multimedia and Expo, 2000. ICME 2000. 2000 IEEE International Conference on*, 423-426 vol.1.

Choppin, A., and (2000). "EEG-based human interface for disabled individuals: emotion expression with neural networks ", Tokyo Institute of Technology, Yokohama, Japan.

Daabaj, Y. "An evaluation of the usability of human-computer interaction methods in support of the development of interactive systems." *System Sciences, 2002. HICSS. Proceedings of the 35th Annual Hawaii International Conference on*, 1868-1877.

Gin-Shin, C., Chih-Wei, C., Ming-Shaung, J., Chou-Ching, L., and Chih-Cheng, L. "Portable Active Surface Laplacian EEG Sensor for Real-Time Mu Rhythms Detection." *Engineering in Medicine and Biology Society, 2005. IEEE-EMBS 2005. 27th Annual International Conference of the*, 5424-5426.

Glassman, E. L. (2005). "A wavelet-like filter based on neuron action potentials for analysis of human scalp electroencephalographs." *Biomedical Engineering, IEEE Transactions on*, 52(11), 1851-1862.

Gott, P. S., Hughes, E. C., and Whipple, K. (1984). "Voluntary control of two lateralized conscious states: Validation by electrical and behavioral studies." *Neuropsychologia*, 22(1), 65-72.

Gunsel, B., Jain, A., Tekalp, A., Sankur, B., Chanel, G., Kronegg, J., Grandjean, D., and Pun, T. (2006). "Emotion Assessment: Arousal Evaluation Using EEG's and Peripheral Physiological Signals." Multimedia Content Representation, Classification and Security, Springer Berlin / Heidelberg, 530-537.

Hai-Rong, L., Zhong-Lin, L., Wen-Jun, Y., and Jin, D. "Emotion recognition based on pressure sensor keyboards." *Multimedia and Expo, 2008 IEEE International Conference on*, 1089-1092.

Heraz, A., Razaki, R., and Frasson, C. "Using machine learning to predict learner emotional state from brainwaves." *Advanced Learning Technologies, 2007. ICALT 2007. Seventh IEEE International Conference on*, 853-857.

Hongo, H., Ohya, M., Yasumoto, M., and Yamamoto, K. "Face and hand gesture recognition for human-computer interaction." *Pattern Recognition, 2000. Proceedings. 15th International Conference on*, 921-924 vol.2.

Jacko, J., Bamidis, P., Frantzidis, C., Konstantinidis, E., Luneski, A., Lithari, C., Klados, M., Bratsas, C., Papadelis, C., and Pappas, C. (2009). "An Integrated Approach to Emotion Recognition for Advanced Emotional Intelligence." Human-Computer Interaction. Ambient, Ubiquitous and Intelligent Interaction, Springer Berlin / Heidelberg, 565-574.

Jung, T.-P., Makeig, S., Humphries, C., Lee, T.-W., Mckeown, M. J., Iragui, V., and Sejnowski, T. J. (2000). "Removing electroencephalographic artifacts by blind source separation." *Journal of Psychophysiology*, 37(2), 163-178.

Khalili, Z., and Moradi, M. H. (2009). "Emotion recognition system using brain and peripheral signals: using correlation dimension to improve the results of EEG." Proceedings of the 2009 international joint conference on Neural Networks, IEEE Press, Atlanta, Georgia, USA.

Mallat, S. G. (1989). "A theory for multiresolution signal decomposition: the wavelet representation." *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 11(7), 674-693.

Massaro, D. W. "Perceptual interfaces in human computer interaction." *Multimedia and Expo, 2000. ICME 2000. 2000 IEEE International Conference on*, 563-566 vol.1.

Merzagora, A. C., Bunce, S., Izzetoglu, M., and Onaral, B. "Wavelet analysis for EEG feature extraction in deception detection." *Engineering in Medicine and Biology Society, 2006. EMBS '06. 28th Annual International Conference of the IEEE*, 2434-2437.

Min, Y.-K., Chung, S.-C., and Min, B.-C. (2005). "Physiological Evaluation on Emotional Change Induced by Imagination." *Applied Psychophysiology and Biofeedback*, 30(2), 137-150.

Mu, L., and Bao-Liang, L. "Emotion classification based on gamma-band EEG." *Engineering in Medicine and Biology Society, 2009. EMBC 2009. Annual International Conference of the IEEE*, 1223-1226.

Murugappan, M., Juhari, M. R. B. M., Nagarajan, R., and Yaacob, S. (2009a). "An investigation on visual and audiovisual stimulus based emotion recognition using EEG." *International Journal of Medical Engineering and Informatics*, 1(3), 342 - 356.

Murugappan, M., Nagarajan, R., and Yaacob, S. "Appraising human emotions using Time Frequency Analysis based EEG alpha band features." *Innovative Technologies in Intelligent Systems and Industrial Applications, 2009. CITISIA 2009*, 70-75.

Murugappan, M., Rizon, M., Nagarajan, R., and Yaacob, S. (2010). "Inferring of Human Emotional States using Multichannel EEG." *European Journal of Scientific Research*, 48(2), 281-299.

Murugappan, M., Rizon, M., Nagarajan, R., Yaacob, S., Hazry, D., and Zunaidi, I. (2008). "Time-Frequency Analysis of EEG Signals for Human Emotion Detection." 4th Kuala Lumpur International Conference on Biomedical Engineering 2008, R. Magjarevic, ed., Springer Berlin Heidelberg, 262-265.

Parameswariah, C., and Cox, M. (2002). "Frequency characteristics of wavelets." *Power Delivery, IEEE Transactions on*, 17(3), 800-804.

Pease, A., and Pease, B. (2004). *The definitive book of body language- How to read others' thoughts by their gestures*, Peace International.

Petrantonakis, P. C., and Hadjileontiadis, L. J. "EEG-based emotion recognition using hybrid filtering and higher order crossings." *Affective Computing and Intelligent Interaction and Workshops, 2009. ACII 2009. 3rd International Conference on*, 1-6.

Robert, H., Dragos, D., and Leon, J. M. R. (2008). "Emotion recognition using brain activity." Proceedings of the 9th International Conference on Computer Systems and Technologies and Workshop for PhD Students in Computing, ACM, Gabrovo, Bulgaria.

Schaaff, K., and Schultz, T. "Towards an EEG-based emotion recognizer for humanoid robots." *Robot and Human Interactive Communication, 2009. RO-MAN 2009. The 18th IEEE International Symposium on*, 792-796.

Schaaff, K., and Schultz, T. "Towards emotion recognition from electroencephalographic signals." *Affective Computing and Intelligent Interaction and Workshops, 2009. ACII 2009. 3rd International Conference on*, 1-6.

Takahashi, K. "Remarks on Emotion Recognition from BioPotential Signals." *2nd Int. Conf. on Autonomous Robots and Agents*, 186-191.

Takahashi, K., and Tsukaguchi, A. "Remarks on emotion recognition from multi-modal bio-potential signals." *Systems, Man and Cybernetics, 2003. IEEE International Conference on*, 1654-1659 vol.2.

Teixeira, J., and Vinhas, V. "Multimodal Emotion Assessment Framework: Gender and High Frequency Electroencephalography as Key-Factors." *International Conference on Informatics and Control*, 233-236.

Teixera, J., Vasco, V., Luis, P. R., and Eugenio. "Automatic Emotion Recognition Induction and Assessment Framework: Enhancing User Interfaces by Interpreting Users

Multimodal Signals." *International Conference on Bio-Inspired Systems and Signal Processing*, 487-490.

Wagner, J., Kim, J., and Andre, E. "From Physiological Signals to Emotions: Implementing and Comparing Selected Methods for Feature Extraction and Classification." *Multimedia and Expo, 2005. ICME 2005. IEEE International Conference on*, 940-943.

Xinyi, Y., Ward, R. K., and Birch, G. E. "Facial EMG contamination of EEG signals: Characteristics and effects of spatial filtering." *Communications, Control and Signal Processing, 2008. ISCCSP 2008. 3rd International Symposium on*, 729-734.

Yongjin, W., and Ling, G. "Recognizing human emotion from audiovisual information." *Acoustics, Speech, and Signal Processing, 2005. Proceedings. (ICASSP '05). IEEE International Conference on*, ii/1125-ii/1128 Vol. 2.

Yuan-Pin, L., Chi-Hong, W., Tien-Lin, W., Shyh-Kang, J., and Jyh-Horng, C. "Multilayer perceptron for EEG signal classification during listening to emotional music." *TENCON 2007 - 2007 IEEE Region 10 Conference*, 1-3.

4

Discrete Wavelet Transform Algorithms for Multi-Scale Analysis of Biomedical Signals

Juuso T. Olkkonen and Hannu Olkkonen
VTT Technical Research Centre of Finland
Department of Applied Physics, University of Eastern Finland, Kuopio
Finland

1. Introduction

The discrete wavelet transform (DWT) algorithms have a firm position in multi-scale processing of biomedical signals, such as EMG and EEG. The DWT algorithms were initially based on the compactly supported conjugate quadrature filters (CQFs) (Smith & Barnwell, 1986; Daubechies, 1988). However, a drawback in CQFs is due to the nonlinear phase effects such as spatial dislocations in multi-scale analysis. This is avoided in biorthogonal discrete wavelet transform (BDWT) algorithms, where the scaling and wavelet filters are symmetric and linear phase. The biorthogonal filters are usually constructed by a ladder-type network called lifting scheme (Sweldens, 1988; ITU-T, 2000). Efficient lifting BDWT structures have been developed for microprocessor and VLSI environment (Olkkonen et al. 2005; Olkkonen & Olkkonen, 2008). Only integer register shifts and summations are needed for implementation of the analysis and synthesis filters.

A severe obstacle in multi-scale DWT analysis is the dependence of the total energy of the wavelet coefficients in different scales on the fractional shifts of the analysed signal. If we have a discrete-time signal $x[n]$ and the corresponding time shifted signal $x[n-\tau]$, where $\tau \in [0,1]$, there occurs a notable difference in the energy of the wavelet coefficients as a function of the time shift. Kingsbury (2001) proposed a nearly shift invariant method, where the real and imaginary parts of the complex wavelet coefficients are approximately a Hilbert transform pair. The energy (absolute value) of the wavelet coefficients equals the envelope, which provides smoothness and approximate shift-invariance. Selesnick (2002) observed that using two parallel CQF banks, which are constructed so that the impulse responses of the scaling filters have half-sample delayed versions of each other: $h[n]$ and $h[n-0.5]$, the corresponding wavelet bases are a Hilbert transform pair. Selesnick (2002) proposed a spectral factorization method based on the half delay all-pass Thiran filters for design of the scaling filters. However, the scaling filters do not owe coefficient symmetry and the nonlinearity interferes with the spatial timing in different scales and prevents accurate statistical correlations.

In this book chapter we review the shift invariant DWT algorithms for multi-scale analysis of biomedical signals. We describe a dual-tree DWT, where two parallel CQF wavelet sequences form a Hilbert pair, which warrants the shift invariance. Next we review the construction of the shift invariant BDWT, which is based on the novel design of the Hilbert

transform filter. Finally, we describe the FFT based computation of the analytic signal and the implementation of the shift invariant quadrature mirror filter (QMF) bank.

2. Shift invariant CQF bank

In the following we describe a shift invariant DWT algorithm based on two parallel real-valued CQF banks. The conventional CQF DWT bank consists of the $H_0(z)$ and $H_1(z)$ analysis filters and $G_0(z)$ and $G_1(z)$ synthesis filters for N odd (Fig. 1)

$$\begin{aligned} H_0(z) &= (1+z^{-1})^K P(z) \\ H_1(z) &= z^{-N} H_0(-z^{-1}) \\ G_0(z) &= H_1(-z) \\ G_1(z) &= -H_0(-z) \end{aligned} \tag{1}$$

where $P(z)$ is a polynomial in z^{-1}. The scaling filter $H_0(z)$ has the Kth order zero at $\omega = \pi$. The wavelet filter $H_1(z)$ has the Kth order zero at $\omega = 0$, correspondingly. The filters are related via the perfect reconstruction (PR) condition

$$\begin{bmatrix} H_0(z) & H_1(x) \\ H_0(-z) & H_1(-z) \end{bmatrix} \begin{bmatrix} G_0(z) \\ G_1(z) \end{bmatrix} = \begin{bmatrix} 2z^{-N} \\ 0 \end{bmatrix} \tag{2}$$

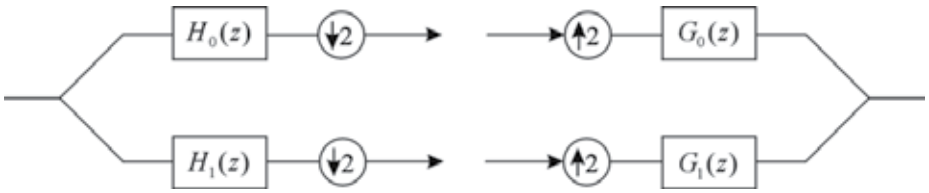


Fig. 1. The analysis and synthesis parts of the real-valued CQF DWT bank.

Let us denote the frequency response of the z-transform filter as

$$H(z) = \sum_n h_n z^{-n} \Rightarrow H(\omega) = \sum_n h_n e^{-j\omega n} \tag{3}$$

Then we obtain the relations

$$\begin{aligned} H(-z) &\Rightarrow H(\omega - \pi) \\ H(-z^{-1}) &\Rightarrow H^*(\omega - \pi) \end{aligned} \tag{4}$$

where * denotes complex conjugation. The tree structured implementation of the two parallel real-valued CQF filter banks is described in Fig. 2. In M-stage CQF tree the frequency response of the wavelet sequence is

$$W_M(\omega) = H_1\left(\frac{\omega}{2}\right) \prod_{k=2}^{M} H_0\left(\frac{\omega}{2^k}\right) \tag{5}$$

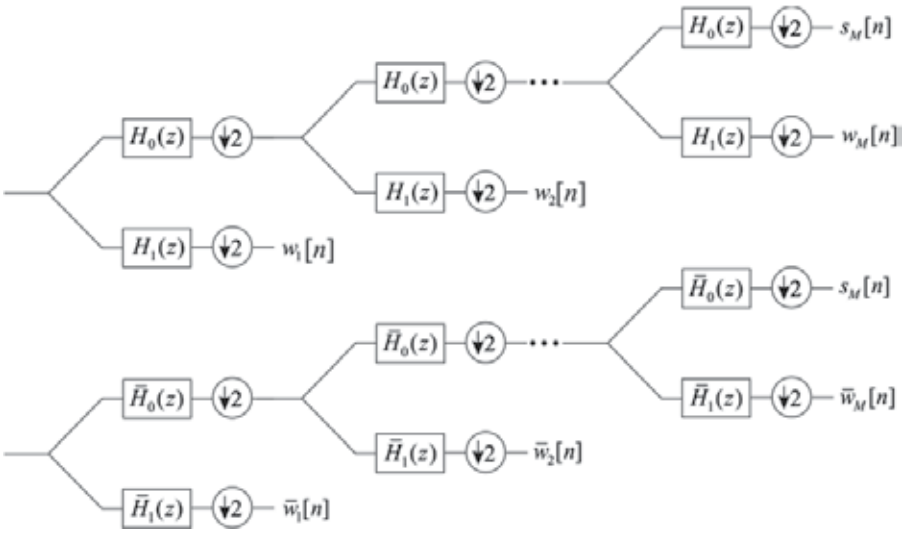


Fig. 2. The implementation of two parallel real-valued CQF banks, which yields the wavelet sequences $w_1[n], w_2[n]\ldots$ and $\bar{w}_1[n], \bar{w}_2[n]\ldots$

Next we construct a phase shifted parallel CQF filter bank consisting of the scaling filter $\bar{H}_0(z)$ and the wavelet filter $\bar{H}_1(z)$ (Fig. 2). Let us suppose that the scaling filters in parallel CQF trees are related as

$$\bar{H}_0(\omega) = e^{-j\phi(\omega)} H_0(\omega) \tag{6}$$

where $\phi(\omega)$ is a 2π periodic phase function. Correspondingly, the CQF filters are related as

$$H_1(\omega) = e^{-j\omega N} H_0^*(\omega - \pi) \tag{7}$$

We have

$$\bar{H}_1(\omega) = e^{-j\omega N} \bar{H}_0^*(\omega - \pi) = e^{-j\omega N} e^{j\phi(\omega-\pi)} H_0^*(\omega - \pi) = e^{j\phi(\omega-\pi)} H_1(\omega) \tag{8}$$

We may note that the phase shifted CQF bank (6,8) obeys the PR condition (2). The frequency response of the M-stage CQF wavelet sequence is

$$\begin{aligned} \bar{W}_M(\omega) &= \bar{H}_1\left(\frac{\omega}{2}\right)\prod_{k=2}^{M}\bar{H}_0\left(\frac{\omega}{2^k}\right) = e^{j\phi\left(\frac{\omega}{2}-\pi\right)} H_1\left(\frac{\omega}{2}\right)\prod_{k=2}^{M} e^{-j\phi\left(\frac{\omega}{2^k}\right)} H_0\left(\frac{\omega}{2^k}\right) \\ &= e^{j\phi\left(\frac{\omega}{2}-\pi\right)} e^{-j\sum_{k=2}^{M}\phi\left(\frac{\omega}{2^k}\right)} H_1\left(\frac{\omega}{2}\right)\prod_{k=2}^{M} H_0\left(\frac{\omega}{2^k}\right) = e^{j\theta} W_M(\omega) \end{aligned} \tag{9}$$

where the phase function

$$\theta = \phi\left(\frac{\omega}{2} - \pi\right) - \sum_{k=2}^{M}\phi\left(\frac{\omega}{2^k}\right) \tag{10}$$

By selecting the phase function $\phi(\omega)$ in (6) as

$$\phi(\omega) = \frac{\omega}{2} \tag{11}$$

the scaling filters (6) are half-sample delayed versions of each other. By inserting (11) in (10) we have

$$\theta = \frac{\frac{\omega}{2} - \pi}{2} - \omega \sum_{k=2}^{M} \frac{1}{2^{k+1}} = -\frac{\pi}{2} + \frac{\omega}{2^{M+1}} \tag{12}$$

The wavelet sequences (5,9) yielded by the CQF bank (1) and the phase shifted CQF bank (6,8) can be interpreted as real and imaginary parts of the complex wavelet sequence

$$W_{MC}(\omega) = W_M(\omega) + j\bar{W}_M(\omega) \tag{13}$$

The requirement for the shift-invariance comes from

$$\bar{W}_M(\omega) = \mathcal{H}\{\psi_M(\omega)\} \tag{14}$$

where $\mathcal{H}$ denotes the Hilbert transform. The frequency response of the Hilbert transform operator is defined as $\mathcal{H}(\omega) = -j\,\mathrm{sgn}(\omega)$, where $\mathrm{sgn}(\omega) = 1$ for $\omega \geq 0$ and $\mathrm{sgn}(\omega) = 0$ for $\omega < 0$. In this work we apply the Hilbert transform operator in the form

$$\mathcal{H}(\omega) = e^{-j\pi/2}\,\mathrm{sgn}(\omega) \tag{15}$$

The result (12) indicates that if the scaling filters are the half-sample delayed versions of each other, the resulting wavelet sequences are not precisely Hilbert transform pairs. There occurs a phase error term $\omega / 2^{M+1}$, which depends both in frequency and the stage M of the wavelet sequence. However, the error term can transferred in front of the CQF tree by using the equivalence described in Fig. 3. Then the error term reduces to $\omega / 2$ and the phase error term can be simply eliminated by prefiltering the analyzed signal by the half-sample delay operator, which has the frequency response $D(\omega) = e^{-j\omega/2}$. The total phase function is then $\theta(\omega) = \angle D(\omega) - \pi/2 + \omega/2 = -\pi/2$, which implies that the M-stage CQF wavelet sequence and the phase error corrected sequence are a Hilbert transform pair.

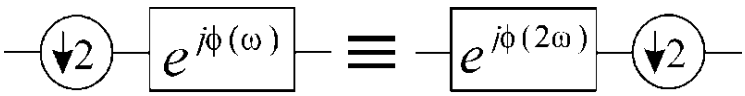


Fig. 3. The two equivalents for moving the phase function in front of the phase shifted CQF tree.

The two parallel BDWT trees can be considered to form a complex wavelet sequence by defining the Hilbert transform operator

$$\mathcal{H}_a(z) = 1 + j\,\mathcal{H}(z) \tag{16}$$

By filtering the real-valued signal $x[n]$ by the Hilbert transform operator results in an analytic signal

$$x_a[n] = x[n] + j\,\mathcal{H}\{x[n]\} \tag{17}$$

whose magnitude response is zero at negative side of the frequency spectrum

$$X_a(\omega) = \begin{cases} 2\,X(\omega) & 0 \le \omega < \pi \\ 0 & -\pi \le \omega < 0 \end{cases} \tag{18}$$

Let us consider the complex wavelet sequence at the first stage (Fig. 6).The wavelet sequence is obtained by decimation of the high-pass filtered analytic signal

$$W_1(\omega) = \left[X_a(\omega)H_1(\omega)\right]_{\downarrow 2} = W_a(\omega)_{\downarrow 2} = \frac{1}{2}X_a\left(\frac{\omega}{2}\right)H_1\left(\frac{\omega}{2}\right) \tag{19}$$

The frequency spectrum of the undecimated wavelet sequence $W_a(\omega)$ contains frequency components only in the range $0 \le \omega < \pi$, but the frequency spectrum of the decimated analytic signal has the frequency band $0 \le \omega < 2\pi$. Hence, the decimation does not produce overlapping and leakage (aliasing) to the negative frequency range.

A key feature of the dual-tree wavelet transform is the shift invariance of the decimated analytic wavelet coefficients. The frequency spectrum of the decimated wavelet sequence of the fractionally delayed signal $x[n-\tau]$ is $0.5e^{-j\omega\tau/2}W_a(\omega/2)$. The energy of the decimated wavelet coefficients is $0.5\left|W(\omega/2)\right|$, which does not depend on the fractional delay.

2. Shift invariant BDWT filter bank

The two-channel BDWT filter bank is of the general form

$$\begin{aligned} H_0(z) &= (1+z^{-1})^L Q(z) \\ H_1(z) &= (1-z^{-1})^M R(z) \\ G_0(z) &= H_1(-z) \\ G_1(z) &= -H_0(-z) \end{aligned} \tag{20}$$

where the scaling filter $H_0(z)$ has the Lth order zero at $\omega = \pi$. The wavelet filter $H_1(z)$ has the Mth order zero at $\omega = 0$, correspondingly. $Q(z)$ and $R(z)$ are polynomials in z^{-1}. The low-pass and high-pass reconstruction filters $G_0(z)$ and $G_1(z)$ are defined as in the CQF bank. For two-channel BDWT filter bank the PR relation is

$$\begin{bmatrix} H_0(z) & H_1(x) \\ H_0(-z) & H_1(-z) \end{bmatrix}\begin{bmatrix} G_0(z) \\ G_1(z) \end{bmatrix} = \begin{bmatrix} 2z^{-N} \\ 0 \end{bmatrix} \tag{21}$$

An essential result is related to the modification of the BDWT bank (Olkkonen & Olkkonen, 2007a).

Lemma 1: If the scaling filter $H_0(z)$, the wavelet filter $H_1(z)$ and the reconstruction filters $G_0(z)$ and $G_1(z)$ in BDWT filter bank (20) have a perfect reconstruction property (21), the following modified BDWT filter bank obeys the PR relation

$$\begin{aligned}\bar{H}_0(z) &= F(z)H_0(z)\\ \bar{H}_1(z) &= F^{-1}(-z)H_1(z)\\ \bar{G}_0(z) &= F^{-1}(z)G_0(z)\\ \bar{G}_1(z) &= F(-z)G_1(z)\end{aligned} \tag{22}$$

where $F(z)$ is any polynomial in z^{-1}. Proof is yielded by direct insertion (22) to PR condition (21).

In the following we apply **Lemma 1** for constructing the shift invariant BDWT filter bank. We describe a novel method for constructing the Hilbert transform filter $\mathcal{H}(z)$ based on the half-sample delay filter $D(z) = z^{-0.5}$. The classical approach for design of the half-sample delay filter $D(z)$ is based on the Thiran all-pass interpolator

$$D(z) = z^{-0.5} = \prod_{k=1}^{p} \frac{c_k + z^{-1}}{1 + c_k z^{-1}} = \frac{z^{-N} C(z^{-1})}{C(z)} = \frac{c_N + c_{N-1} + \cdots + z^{-N}}{1 + c_1 z^{-1} + \cdots + c_N z^{-N}} \tag{23}$$

where the c_k coefficients are optimized so that the frequency response follows approximately $D(\omega) = e^{-j\omega/2}$. Correspondingly, the quadrature mirror filter $D(-z)$ has the frequency response

$$D(\omega - \pi) = e^{-j(\omega-\pi)/2} \tag{24}$$

The Hilbert transform filter is then obtained as

$$\mathcal{H}(\omega) = \frac{D(\omega)}{D(\omega - \pi)} = e^{-j\omega/2} e^{j(\omega-\pi)/2} = e^{-j\pi/2} \Rightarrow \mathcal{H}(z) = \frac{D(z)}{D(-z)} \tag{25}$$

The Hilbert transform filter is inserted in the BDWT bank using the result of **Lemma 1** (22). The modified prototype BDWT filter bank is

$$\begin{aligned}\bar{H}_0(z) &= \mathcal{H}(z)H_0(z)\\ \bar{H}_1(z) &= \mathcal{H}^{-1}(-z)H_1(z)\\ \bar{G}_0(z) &= \mathcal{H}^{-1}(z)G_0(z)\\ \bar{G}_1(z) &= \mathcal{H}(-z)G_1(z)\end{aligned} \tag{26}$$

A highly simplified BDWT filter bank can be obtained by noting that in (25) $\mathcal{H}^{-1}(-z) = \mathcal{H}(z)$. We have

$$\begin{aligned}\bar{H}_0(z) &= \mathcal{H}(z)H_0(z)\\ \bar{H}_1(z) &= \mathcal{H}(z)H_1(z)\\ \bar{G}_0(z) &= \mathcal{H}(-z)G_0(z)\\ \bar{G}_1(z) &= \mathcal{H}(-z)G_1(z)\end{aligned} \tag{27}$$

The modified BDWT filter bank (27) can be realized by the Hilbert transform filter $\mathcal{H}(z)$, which works as a prefilter for the analysed signal (Fig. 4). The Hilbert transform filter $\mathcal{H}(-z)$ works as a postfilter in the reconstruction stage, respectively.

Fig. 4. The realization of the Hilbert transform filter.

An integer-valued Hilbert transform filter can be constructed by the B-spline transform (see details Olkkonen & Olkkonen, 2011b). The frequency response of the Hilbert transform filter shows a maximally flat magnitude spectrum. The phase spectrum corresponds to an ideal Hilbert transformer (15).
The Hilbert transform filter in Fig. 4 can be replaced by the Hilbert transform operator (16), which yields an analytic signal. This avoids the need for two parallel filter banks. In the following we describe a FFT based method for computation of the analytic signal and the implementation of the shift invariant quadrature mirror filter (QMF) bank.

3. FFT based computation of analytic signal

The fast Fourier transform of the signal $x[n]$, n = 0, 1, 2, …,N-1 is of the form

$$FFT_N\{x[n]\} = Y_k = \sum_{n=0}^{N-1} x[n]\,W_N^{nk} \quad k = 0,...,N-1 \tag{28}$$

where $W_N = \mathrm{e}^{-2\pi j/N}$. The FFT coefficients Y_k (k=N/2,…,N-1) represent the values in the negative frequency band $(-\pi \le \omega < 0)$. By zeroing those coefficients, the inverse fast Fourier transform (IFFT) yields an analytic signal. A more accurate result is obtained by weighting the FFT coefficients by a window

$$w_k = \begin{cases} 2 & k = 1,2,...,N/2-1 \\ 0 & k = N/2+1,...,N-1 \\ 1 & k = 0 \; and \; N/2 \end{cases} \tag{29}$$

The analytic signal is then computed using the inverse FFT transform

$$x_a[n] = \frac{1}{N}\sum_{k=0}^{N-1} w_k Y_k W_N^{-nk} = IFFT_N\{w_k Y_k\} \tag{30}$$

The weighting sequence in (29) can be eliminated by writing

$$x_a[n] = \frac{1}{N/2}\sum_{k=0}^{N/2-1} Y_k W_N^{-nk} - \frac{Y_0}{N} + (-1)^n \frac{Y_{N/2}}{N} \tag{31}$$

Now, for even n we have

$$x_a[2n] = \frac{1}{N/2}\sum_{k=0}^{N/2-1} Y_k W_{N/2}^{-nk} - mean\{x[2n+1]\} = IFFT_{N/2}\{Y_k\} - mean\{x[2n+1]\} \tag{32}$$

and for odd n

$$x_a[2n+1] = \frac{1}{N/2}\sum_{k=0}^{N/2-1} W_N^{-k} Y_k W_{N/2}^{-nk} - mean[x[2n]\} = IFFT_{N/2}\{W_N^{-k} Y_k\} - mean\{x[2n]\} \tag{33}$$

For zero mean signal $mean\{x[n]\} = 0$, which yields $mean\{x[2n+1]\} = -mean\{x[2n]\}$. If the even points of the analytic signal are known, the FFT coefficients are solved from (32)

$$Y_k = FFT_{N/2}\{x_a[2n] - mean\{x[2n]\}\} \quad k = 0,...,N/2-1 \tag{34}$$

The odd points of the analytic signal are then computed from (33). We call this as the reconstruction property of the zero mean analytic signal. In the following we present a novel shift invariant QMF bank, which utilizes the reconstruction property of the analytic signal.

4. Shift invariant QMF bank

In QMF bank the scaling and wavelet filters obey the relation $H_1(z) = H_0(-z)$, i.e. their frequency response is symmetric with respect to $\omega = \pi/2$. In this work we define the scaling and wavelet filters as half band QMFs

$$\begin{aligned} H_0(z) &= \frac{1}{2} + z^{-1}A(z^2) \\ H_1(z) &= \frac{1}{2} - z^{-1}A(z^2) \end{aligned} \tag{35}$$

The shift invariant tree structured QMF DWT is described in Fig. 5. The FFT based Hilbert transform operator $\mathcal{H}_a(z)$ produces an analytic signal, which is fed to the scaling $H_0(z)$ and wavelet $H_1(z)$ filters and decimated. If the original zero mean signal is $x[n]$, the decimated scaling and wavelet coefficients $s[n]$ and $w[n]$ are obtained from

$$\begin{aligned} s[n] &= \{h_0[n] * x_a[n]\}_{\downarrow 2} \\ w[n] &= \{h_1[n] * x_a[n]\}_{\downarrow 2} \end{aligned} \tag{36}$$

where $*$ denotes convolution. From (35) we have

$$H_0(z) + H_1(z) = 1 \quad \Rightarrow \quad h_0[n] + h_1[n] = \delta_n \tag{37}$$

The reconstruction consists of the summation of the decimated signals. We obtain

$$s[n] + w[n] = \{(h_0[n] + h_1[n]) * x_a[n]\}_{\downarrow 2} = \{\delta[n] * x_a[n]\}_{\downarrow 2} = x_a[2n] \tag{38}$$

i.e. the summation of the decimated signals produces the even points $x_a[2n]$ of the analytic signal. The odd points $x_a[2n+1]$ of the analytic signal are then reconstructed from the even points $x_a[2n]$ via our results (32)-(34). The original signal is obtained from $x[n] = real(x_a[n])$.

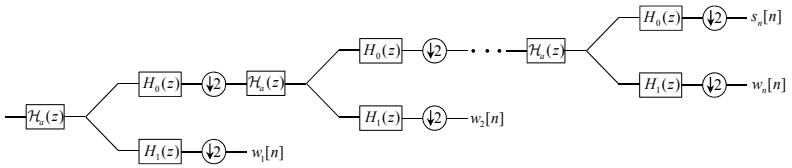


Fig. 5. The shift invariant tree structured QMF DWT.

5. Conclusion

The dual-tree DWT algorithms have appeared to outperform the real-valued DWTs in several applications such as denoising, texture analysis, speech recognition, processing of seismic signals and multiscale-analysis of neuroelectric signal analysis (Olkkonen et al. 2006; Olkkonen et al. 2007b, Olkkonen & Olkkonen, 2010, Olkkonen & Olkkonen 2011a).

Selesnick (2002) noted that a half-sample time-shift between the scaling filters in parallel CQF banks yields a nearly shift invariant DWT, where the wavelet bases form a Hilbert transform pair. However, the multi-scale analyses of neuroelectric signals have revealed that the first stages of wavelet sequences are quite poorly shift invariant. We reanalysed the condition and observed a phase-error term $\omega / 2^{M+1}$ (12) compared with the ideal phase response $\theta(\omega) = -\pi / 2$. The phase error attains s highest value at high frequency range and small stage M of the wavelet sequence. Fortunately the phase error term can be cancelled by adding a half-delay prefilter in front of the CQF chain. For this purpose the half-delay filter constructed by the B-spline transform (Olkkonen & Olkkonen, 2011b) is well suited. In addition, there exists many other design methods for half-delay filters (see e.g. Laakso et al. 1996; Johansson & Lowenborg, 2002; Pei & Tseng, 2003; Pei & Wang, 2004; Tseng, 2006).

In this book chapter we described a novel shift invariant dual-tree BDWT (27) based on *Lemma 1* (22) and the Hilbert transform filter (25). In many respects the shift invariant BDWT bank outperforms the previous nearly shift invariant DWT approaches. The Hilbert transform filter assisted BDWT yields precisely shift invariant wavelet sequences, which permits the statistical analyses between scales in multi-scale analyses of biomedical signals such as EMG and EEG.

The Hilbert transform filter in Fig. 4 can be replaced by the Hilbert transform operator (16), which yields an analytic signal. This avoids the need for two parallel filter banks. In this work we described a FFT based method for computation of the analytic signal and the implementation of the shift invariant QMF bank. As a clear advantage of the half-band QMF structure is that the frequency responses of the scaling and wavelet filters are mirror symmetric with respect to $\omega = \pi / 2$. Hence, they split the energy of the signal precisely to the low-pass and high-pass fractions. The energy preservation property is of utmost importance in automated statistical signal processing of the multi-scale signals. In tree structured multi-scale analysis the linear phase of the QMFs is advantageous since the timing information of the wavelet coefficients in different scales is preserved. Without an exact timing of the subscale signals the statistical comparison of the wavelet coefficients in different scales is not relevant and may lead to misleading results. For example in EEG signal the neuroelectric discharge contains fast repetitive transients with related timing and

overlapping waveforms. In multi-scale analysis different components can be separated due to their different timing and scale related intensification.

6. Acknowledgements

This work was supported by the National Technology Agency of Finland (TEKES).

7. References

Daubechies, I. (1988). Orthonormal bases of compactly supported wavelets. *Commmun. Pure Appl. Math.,* Vol. 41, 909-996.

ITU-T (2000) Recommend. T.800-ISO DCD15444-1: *JPEG2000 Image Coding System. International Organization for Standardization,* ISO/IEC JTC! SC29/WG1.

Johansson, H. & Lowenborg, P. (2002). Reconstruction of nonuniformy sampled bandlimited signals by means of digital fractional delay filters, *IEEE Trans. Signal Process.*, Vol. 50, No. 11, pp. 2757-2767.

Kingsbury, N.G. (2001). Complex wavelets for shift invariant analysis and filtering of signals. *J. Appl. Comput. Harmonic Analysis.* Vol. 10, 234-253.

Laakso, T., Valimaki, V., Karjalainen, M. & Laine, U.K. (1996). Splitting the unit delay. Toolsfor fractional delay filter design, *IEEE Signal Processing Magazine*, pp. 30- 80.

Olkkonen, H., Pesola, P. & Olkkonen, J.T. (2005). Efficient lifting wavelet transform for micro-processor and VLSI applications. *IEEE Signal Process. Lett.* Vol. 12, No. 2, 120-122.

Olkkonen,H.,Pesola, P.,Olkkonen, J.T. & Zhou, H. (2006). Hilbert transform assisted complex wavelet transform for neuroelectric signal analysis. *J. Neuroscience Meth.* Vol. 151, 106-113.

Olkkonen, H. & Olkkonen, J.T. (2007a). Half-delay B-spline filter for construction of shift-invariant wavelet transform. *IEEE Trans. Circuits and Systems II.* Vol. 54, No. 7, 611-615.

Olkkonen, H., Olkkonen, J.T. & Pesola, P. (2007b). FFT-based computation of shift invariant analytic wavelet transform. *IEEE Signal Process. Lett.* Vol. 14, No. 3, 177-180.

Olkkonen, H. & Olkkonen, J.T. (2008). Simplified biorthogonal discrete wavelet transform for VLSI architecture design. *Signal, Image and Video Process.* Vol. 2, 101-105.

Olkkonen, H. & Olkkonen, J.T. (2010). Shift-invariant B-spline wavelet transform for multi-scale analysis of neuroelectric signals. IET Signal Process. Vol. 4, No. 6, 603-609.

Olkkonen, J.T. & Olkkonen, H. (2011a). Shift invariant biorthogonal discrete wavelet transform for EEG signal analysis. Book chapter in: *Discrete Wavelet Transforms - Theory and Applications*, Edited by Juuso T. Olkkonen, Intech, pp. 169-178.

Olkkonen, J.T. & Olkkonen, H. (2011b). Complex Hilbert transform filter. *J. Signal and Information Process. Vol.2, 112-116.*

Pei , T. S.C. & Tseng, C.C. (2003). An efficient design of a variable fractional delay filter using a first-order differentiator, *IEEE Signal Processing Letters*, Vol. 10, No. 10, pp. 307-310.

Pei , S.C. &and Wang, P.H. (2004). Closed-form design of all-pass fractional delay, *IEEE Signal Processing Letters*, Vol. 11, No. 10, pp. 788-791.

Selesnick, I.W. (2002). The design of approximate Hilbert transform pairs of wavelet bases. *IEEE Trans. Signal Process.* Vol. 50, No. 5, 1144-1152.

Smith, M.J.T. & Barnwell, T.P. (1986). Exaxt reconstruction for tree-structured subband coders. *IEEE Trans. Acoust. Speech Signal Process.* Vol. 34, 434-441.

Sweldens, W. (1988). The lifting scheme: A construction of second generation wavelets. *SIAM J. Math. Anal.* Vol. 29, 511-546.

Tseng,C.C. (2006). Digital integrator design using Simpson rule and fractional delay filter, *IEEE Proc. Vision, Image and Signal Process.*, Vol. 153, No. 1, pp. 79-85.

5

Computerized Heart Sounds Analysis

S.M. Debbal
Genie -Biomedical Laboratory (GBM),
Department of electronic
Faculty of science engineering
university Aboubekr Belkaid
Algeria

1. Introduction

Heartbeat sound analysis by auscultation is still insufficient to diagnose some heart diseases. It does not enable the analyst to obtain both qualitative and quantitative characteristics of the phonocardiogram signals [1],[2]. Abnormal heartbeat sounds may contain, in addition to the first and second sounds, S1 and S2, murmurs and aberrations caused by different pathological conditions of the cardiovascular system [2]. Moreover, in studying the physical characteristics of heart sounds and human hearing, it is seen that the human ear is poorly suited for cardiac auscultation [3]. Therefore, clinic capabilities to diagnose heart sounds are limited.

The sound emitted by a human heart during a single cardiac cycle consist of two dominant events, known as the first heart sound S1 and the second heart sound S2 (Figure1), S1 relates to the closing of the mitral and tricuspid valves whilst S2 is generated by the halting of the aortic and pulmonary valves leaflets [1]. S1 corresponds in timing to the QRS complex in ECG (Electrocardiogram) and S2 follows the systolic pause in the normal cardiac cycle. Heartbeat sound analysis by auscultation only is still insufficient to diagnose some heart diseases. It does not enable the analyst to obtain both qualitative and quantitative characteristics of S1 and S2 of the phonocardiogram [1-2]. Moreover, in studying the physical characteristics of heart sounds and human hearing, it is seen that the human ear is poorly suited for cardiac auscultation [3]. Therefore, clinic capabilities to diagnose heart sounds are limited. In this paper we are interested in the study of the physical characteristics of **the second heart sound S2 which consist of two major components in the spectrum of the signal** . One of these components A2 is due to the closure of the aortic valve and the other P2 is due to the closure of the pulmonary valve.

The aortic component is loudest than the pulmonic component. It is discernible at all the auscultation sites.It is best heard at the right base, with the diaphragm of the chest piece firmly pressed, whereas the pulmonic on a may only be audible at the left base, with the diaphragm of the chest piece firmly pressed.

The aortic component has higher frequency contents and generally precedes the pulmonary component because in normal heart activity the aortic valve closes before the pulmonary valve. The difference of time occurrence between these valves activities is known in medical

community [4 - 6] as split. However in many diseases [1] this order of time occurrence may be reverse and its delay varies.

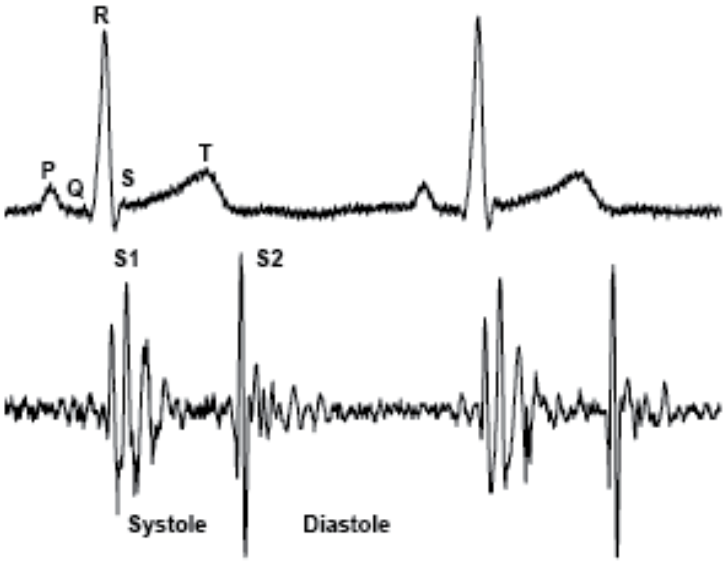


Fig. 1. Correlation between the phonocardiogram signal (PCG) and the electrocardiogram signal (ECG)

The characteristics of the PCG signal and other features such as heart sounds S1 and S2 location; the number of components for each sound; their frequency content; their time interval; all can be measured more accurately by digital signal processing techniques.
The FFT (Fast Fourier Transform) can provide a basic understanding of the frequency contents of the heart sounds. However, FFT analysis remains of limited values if the stationary assumption of the signal is violated. Since heart sounds exhibit marked changes with time and frequency, they are therefore classified as non - stationary signals. To understand the exact feature of such signals, it is thus important, to study their time – frequency characteristics.
In this paper the wavelet transform is used to analyse both the normal and abnormal heart sound in both time and frequency domains. This technique has been shown to have a very good time resolution for high-frequency components. In fact the time resolution increases as the frequency increases and the frequency resolution increases as the frequency decreases [4],[5].
Furthermore, the wavelet transform has demonstrated the ability to analyse the heart sound more accurately than other techniques STFT or Wigner distribution [6] in some pathological cases.
In fact the spectrogram (STFT) cannot track very sensitive sudden changes in the time direction. To deal with these time changes properly it is necessary to keep the length of the time window as short as possible. This however, will reduce the frequency resolution in the

time-frequency plane. Hence, there is a trade-off between time and frequency resolutions [6].
However the Wigner distribution (WD) and the corresponding WVD (Wigner Ville Distribution) have shown good performances in the analysis of non-stationary signals. This comes from the ability of the WD to separate signals in both time and frequency directions. One advantage of the WD over the STFT is that it does not suffer from the time-frequency trade-off problem. On the other hand, the WD has a disadvantage since it shows cross-terms in its response. These cross-terms are due to the nonlinear behaviour of the WD, and bear no physical meaning. One way to remove these cross-terms is by smoothing the time-frequency plane, but this will be at the expense of decreased resolution in both time and frequency [7].
The WD was applied to heart sound signal it shows no succes in displaying or separating the signal components in both the time and frequency direction [6], although it provides high time-and frequency- resolution in simple monocomponent signal analysis[8].
To overcome these difficulties with the STFT and the WD an alternative way to analyse the non-stationary signals is the wavelet transform (WT). It expand the signal some basis functions. The basis functions can be constructed by dilation, contractions and shifts of a unique function called the "wavelet prototype" or "wavelet mother". The WT act as "mathematical microscope" in which we can observe different parts of the signal by just adjusting the focus.
The wavelet Transform is a technique in the domain of time-frequency distributions. The main idea of this method is the representation of an arbitrary signal as a superposition of basic signals, "atoms", located in time and frequency. These atoms may be derived by means of a special operation on a single parent atom. Parent atoms and derivation operation are usually chosen such as to enable the construction of an orthonormal system [9].
The study of the decomposition of the signal in " atoms " was first carried out by Gabor however, it was quickly abandoned be cause of :

1. the no simultaneous Representation in time and frequency
2. grid made up of rectangular cells is not a flexible device
3. the mathematical theory of the phenomenon is badly structured.

The representation time-scale of WT based on a dyadic paving appears more flexible. It a mathematical structure governed by a formula of exact inversion [10] making possible the existence of orthonormal basis. This makes the wavelet to be a simultaneous function of time and frequency.
In this paper the continuous wavelet transform (CWT) is applied to analyse pathological PCG signals. The CWT is more appropriate than the discrete wavelet transform (DWT) , since we are interested in the analysis of non-stationary signals and not signal coding where DWT is found to be more useful

2. Theoretical background

2.1 Fourier Transform (FT)

In 1882, Joseph Fourier discovered that any periodic function could be represented as an infinite sum of periodic complex exponential functions [10]. The inclusive property of only periodic functions was later extended to any discrete time function. The Fourier transform (FT) [as regular Fourier Transform] converts a signal expressed in the time domain to a signal expressed in the frequency domain. The FT is widely used and usually implemented

in the form of FFT algorithm (fast Fourier transform). The mathematical definition of the FT is given below.

$$X(f) = \int x(t) e^{-j2\Pi ft} dt \tag{1}$$

Where t and f are respectively the time and frequency parameters. The time domain signal $x(t)$ is multiplied by a complex exponential at a frequency f and integrate over all time. In other words, any discrete time signal may be represented by a sum of sines and cosines, which are shifted and are multiplied by a coefficient that changes their amplitude. $X(f)$ are the Fourier coefficients which are large when a signal contains a frequency component around the frequency f .

The peaks in a plot of the FT of a signal correspond to dominant frequency components of the signal. Fourier analysis is simply not effective when used on non stationary signals because it does not provide frequency content information localized in time. Most real world signals exhibit non stationary characteristics (such as heart sound signals). Thus, Fourier analysis is not adequate.

2.2 Short-time Fourier Transform (STFT)

The STFT is obtained by calculating the Fourier transform of a sliding windowed version of the time signal s(t). The location of the sliding window adds a time dimension and one gets a time-varying frequency analysis.

The mathematical representation of STFT is :

$$S(t, f) = \int_{-\infty}^{+\infty} s(\tau) w(\tau - t) e^{-j2\pi \ ft} d\tau \tag{2}$$

Where $w(\tau - t)$ it is the sliding window applied to the signal $s(t)$, f is the frequency and t is the time.

The length of the window is chosen so that to maintain signal stationary in order to calculate the Fourier transform. To reduce the effect of leakage (the effect of having finite duration), each sub-record is then multiplied by an appropriate window and then the Fourier transform is applied to each sub-record. As long as each sub-record does not contain rapid changes the spectrogram will give an excellent idea of how the spectral composition of the signal has changed during the whole time record. However, there exist many physical signals whose spectral content is so rapidly changing that finding an appropriate short-time window is problematic, since there may not be any time interval for which the signal is stationary. To deal with these time changes properly it is necessary to keep the length of the time window as short as possible. This, however, will reduce the frequency resolution in the time-frequency plane. Hence, there is a trade-off between time and frequency resolutions.

2.3 Wgner Distribution function (WD)

In contrast to the STFT, which is resolution limited either in time or in frequency (dictated by the window function), and suffers from smearing and side lobe leakage, the WD offers excellent resolution in both the frequency and time domains. The WD of two signals, $x(t)$, $y(t)$, is defined via,

$$W(t,\omega) = \int_{-\infty}^{+\infty} x(t+\frac{\tau}{2}) y^*(t-\frac{\tau}{2}) e^{-j\omega\tau} d\tau \quad (3)$$

The auto WD is obtained when $x(t) = y(t)$=s(t), it is a bilinear function of the signal s(t). the WD, and can also be expressed as

$$S(t,\omega) = \int_{-\infty}^{+\infty} s(t+\frac{\tau}{2}) s^*(t-\frac{\tau}{2}) e^{-j\omega\tau} d\tau \quad (4)$$

Where t and ω are respectively, the time and frequency variables, and * denotes the complex conjugate.

The WD had shown good applications in the analysis of non-stationary signal [7]. This comes from the ability of this method to separate signals in both time and frequency directions. The WD has a disadvantage that it is limited by the appearance of cross-terms. These cross-terms are due to the nonlinear behaviour of the WD, and bear no physical meaning. One way to remove these cross-terms is by smoothing the time-frequency plane [7], but this will be at the expense of decreased resolution in both time and frequency.

2.4 Continuous Wavelet Transform (CWT)

The continuous wavelet transform was developed as a method to obtain simultaneous, high resolution time and frequency information about a signal. The term 'wavelet' was first mentioned in 1909 in a thesis by Alfred Haar [M.Misiti.Y.Misiti.G.Oppenheim.J.-M.Poggi.Wavelet Toolbox:For use with MATLAB.The Math Works Inc.1996], although the progress in the field of wavelet has been relatively slow until the 1980's when scientist and engineers from different fields realized they were working on the same concept and began collaborating .

The CWT rather than the STFT uses a variable sized window region .Because the wavelet may be dilated or compressed, different features of the signal are extracted. While a narrow wavelet extracts high frequency components, a stretched wavelet picks up on the lower frequency components of the signal.

The CWT is computed by correlating the signal s(t) with families of time-frequency atoms g(t), it produce a set of coefficients C(a,b) given by :

$$C(a,b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} s(t) g^*(\frac{t-b}{a}) dt \quad (5)$$

$$= \sqrt{a} \int_{-\infty}^{+\infty} G(aw) S(w) e^{jbw} dw \quad (6)$$

Where

- b is the time location
- a is called scale factor and it is inversely proportional to the frequency (a > 0)
- *dénotes a complexe conjugate.
- g(t) is the analysing wavelet .
- S(w) and G(w) are, respectively, the Fourier transforms of s(t) and g(t).

The analyzing wavelet function g(t) should satisfy some properties. The most important ones are continuity, integrability, square integrability, progressivity and it has no d.c component. Moreover, the wavelet g(t) has to be concentrated in both time and frequency as much as possible. It is well known that the smallest time-bandwidth product is achieved by the Gaussian function [3]. Hence the most suitable analyzing wavelet for time-frequency analysis is the complex exponential modulated Gaussian function of the form :

$$g(t) = \exp(-\frac{t^2}{2} + j\, w_0\, t) \tag{7}$$

If we choose the analyzing wavelet that has the following Fourier transform (FT):

$$G(w) = A\exp[-\left(w - w_0\right)^2 / 2] + \varepsilon \tag{8}$$

When ε is a small correction term, theoretically necessary to satisfy the admissibility conditions of wavelets; w_o is chosen large enough so that the correction term is negligible and can be ignored.
This is known as the Gabor wavelet. It was shown [5] that w_0 = 5.33, which is enough to make the correction term negligible and gives an optimal time-bandwidth product.
In a continuous wavelet transform, the wavelet corresponding to the scale and the time location b is given by

$$g_{a,b}(t) = 1/\sqrt{|a|}\ \ g(\tfrac{t-b}{a}) \tag{9}$$

Where g(t) is the wavelet "prototype" or mother which can be thought of as a band pass function. The factor $/a/^{-1/2}$ is used to ensure energy preservation [5].

2.5 Discrete Wavelet Transform (DWT)

Wavelet series (WS) coefficients are sampled CWT coefficients. Time remains continuous but time-scale parameters (b,a) are sampled on a so-called "dyadic" grid in the time-scale plane (b,a) [11]. A common definition is :

$$C_{jk} = \mathrm{CWT}\{s(t);\ a = 2^j,\ b = k\,2^j\} \quad \text{with } j,k \in Z \tag{10}$$

The wavelets are in this case :

$$\Psi_{jk}(t) = 2^{-j/2}\ \ \Psi(2^{-j}\, t - k) \tag{11}$$

The discrete wavelet transform (DWT) has been recognized as a natural wavelet transform for discrete-time signals. Both the time and time-scale parameters are discrete.
The discretization process partially depends upon the algorithm chosen to perform the transformation.
The $C_{j,k}$ could be well approximated by digital filter banks. By using Mallat's [12] remarkable fast pyramid algorithms which involve use of low-pass and high-pass filters. The Mallat algorithm is in fact a classical scheme known in the signal processing community as two-channel subband coder. The wavelet analysis permits to decompose the

phonocardiogram signal in filter banks. The signal to be analyzed is decomposed in approximation an in detail while using two filter h and g. The h filter is a low - pass filter with a band pass [0,Π/2], it generates the approximation signal "A".
The g filter is a high – pass filter of band pass [0, Π], it generates the detail signal "D". The filtered signals are under-sampled (decimation) according to the rule of Nyquist. This leads to a considerable reduction in computing time. Every approximation is decomposed all over again in approximation and detail. In this case the signal is decomposed into several frequency bands instead of two bands. The number of bands depends on decomposition level.
Figure2 shows the discrete wavelet decomposition on three levels, as well as the note associated filter bank for each level. We note that the width of the filter banks band-pass is veritable according to the decomposition level. The wavelet analysis permits to decompose the phonocardiogram signal in filter banks. The signal to be analyzed is decomposed in approximation an in detail while using two filter h and g. The h filter is a low - pass filter with a band pass [0,Π/2], it generates the approximation signal "A".

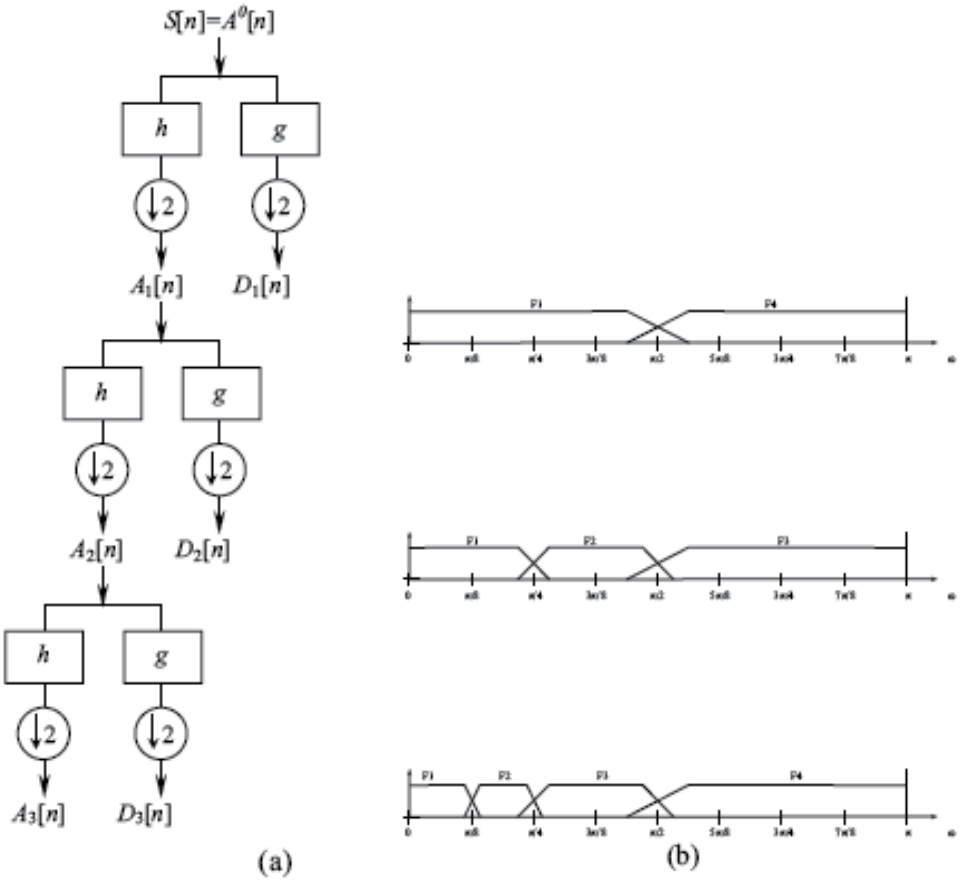


Fig. 2. a) Exemple of the discrete wavelet decompostion for three levels; b) Filter banks of the decomposition of each level

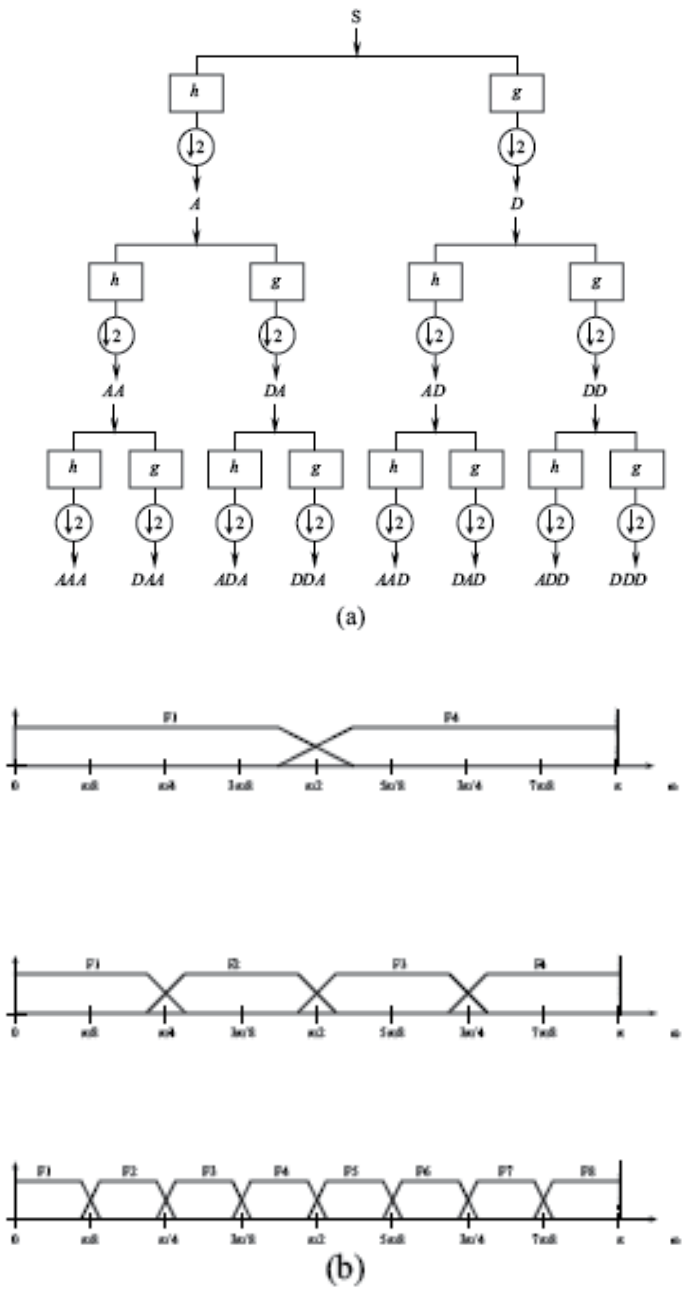


Fig. 3. a) Exemple of the packet wavelet decompostion for three levels; b) Filter banks of the decomposition of each level

2.6 Wavelet Packet Transform (WPT)

Wavelet packet analysis is an extension of the discrete wavelet transform (DWT) and it turns out that the DWT is only one of the much possible decomposition that could be performed

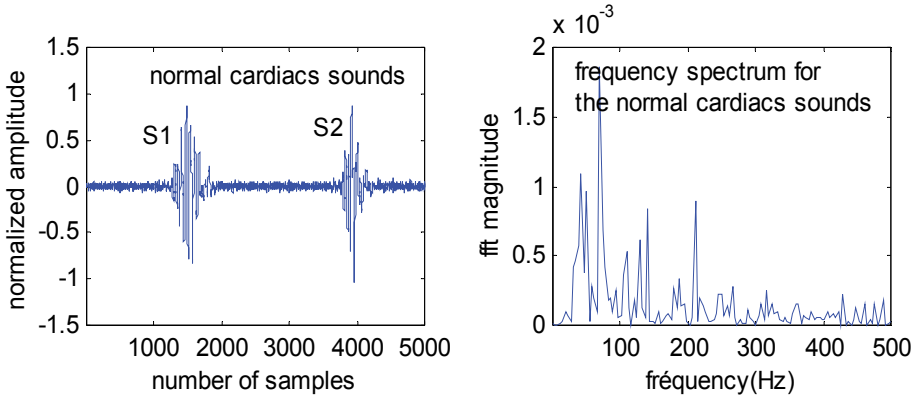
on the signal. Instead of just decomposing the low frequency component, it is therefore possible to subdivide the whole time-frequency plane into different time-frequency pieces as can be seen from. The advantage of wavelet packet analysis is that it is possible to combine the different levels of decomposition in order to achieve the optimum time-frequency representation of the original [13].
To obtain a uniform filter bank we will use the wavelet packet analysis which is a generalization of the discrete wavelet transform analysis. The decomposition is also made of the details signal "D". Figure3 shows the wavelet packet decomposition on three levels as well as the filter banks associated for each uniform level.

3. Results and discussion

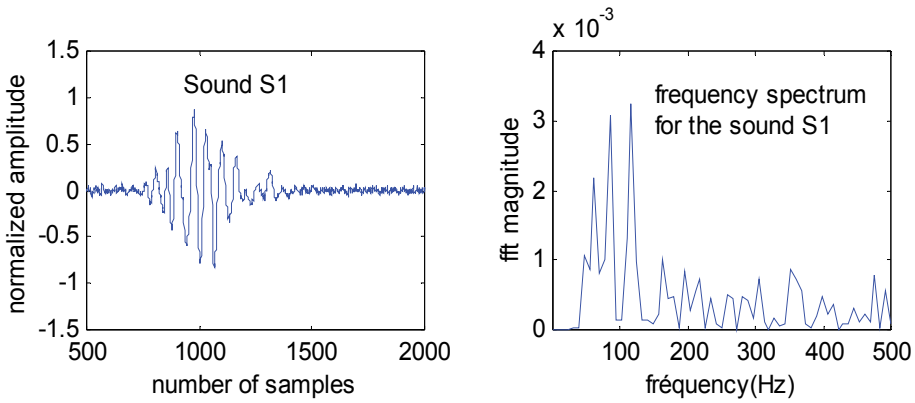
3.1 Fast Fourier Transform application

An FFT algorithm is first applied to the PCG signal given in Figure1. The frequency spectrum illustrated in Figure4a shows that the normal PCG signal has a frequency content varying from around 40Hz up to 200Hz. The FFT can be applied to the first part of this signal to analyse the frequency content of S1 as shown in Figure 4b and then to the second half to analyse the frequency content of S2 as shown in Figure4c. A 512 points FFT is applied to S1 and S2**.** At this stage the sound S1 or S2 cannot be separated.
In fact; the application of the FFT on heart sounds S1 and S2 after their separation or identification show that the basic frequency components are obviously detected by the Fourier transform .
The spectrum of S1 has reasonable values in the range 10-180Hz. The spectrum is distinctly resolved in time into two majors components (M1 and T1) while the most of the energy of these sounds appears to be concentrated.
The two components A2 (due to the closure of the aortic valve) and P2 due to the closure of the pulmonary valve) of the second sound S2 are obvious in Figure4c. The spectrum of the sound S2 has reasonable values in the range 50-250Hz. The spectrum for this sound is distinctly resolved in time into two majors components (A2 and P2) as shown in Figure4c.
However the FFT analysis of S2 cannot tell neither which of A2 and P2 precedes the other, nor the value of the time delay known as the "split" which separate them. For a normal heart activity usually A2 precedes P2 and the value of the split is lower than 30ms. This time delay between A2 and P2 is very important to detect some pathological cases. The sound S2 seem to have higher frequency content than that of S1 as shown in Figure4b and Figure4c.
The FFT is applied also to analyse fourth PCG signals one normal and three different marked pathological cases (the aortic-insufficiency, the aortic-stenosis and the mitral-stenosis). These are illustrated in Figure5 along with the normal PCG signal. The basic frequency content is obviously different from that of the normal PCG signal. It is clearly shown that there is great loss of frequency component in each of the pathological case with respect to normal case. In addition except the aortic- insufficiency case where we note the apparition of frequency component higher than 200Hz , the other cases (mitral-stenosis and aortic-stenosis) present a frequency spectrum limited to 200Hz.
The aortic-insufficiency and the aortic-stenosis are two pathological cases resulting from a severe organic attack, which generally involves a disappearance of the aortic component A2 of the sound S2. This shown in their corresponding PCG frequency responses illustrated in Figure5, where we notice a lack in frequency contents in the range under 100Hz compared to the normal case, where there is much more frequency component in this range. On the

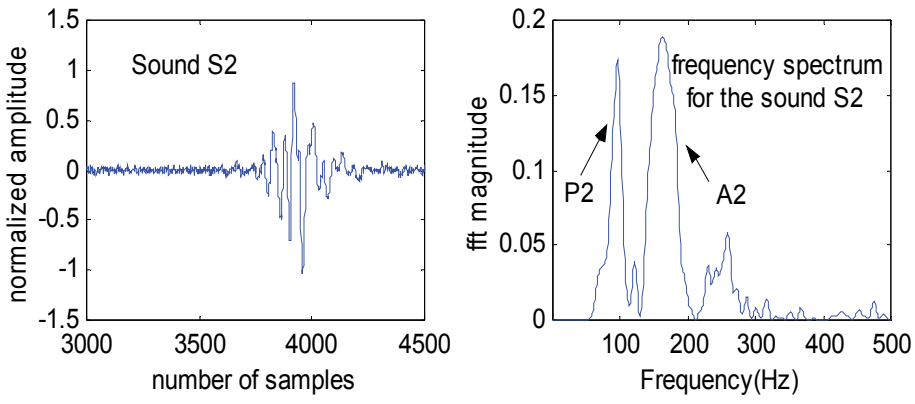
other hand the mitral-stenosi is rather a severe attack of the mitral valves thus involving a presystolic reinforcement as well as a bursting of the sound S1.



a)



b)



c)

Fig. 4. Frequency spectrum for the normal cardiac sounds and the sounds S1 and S2

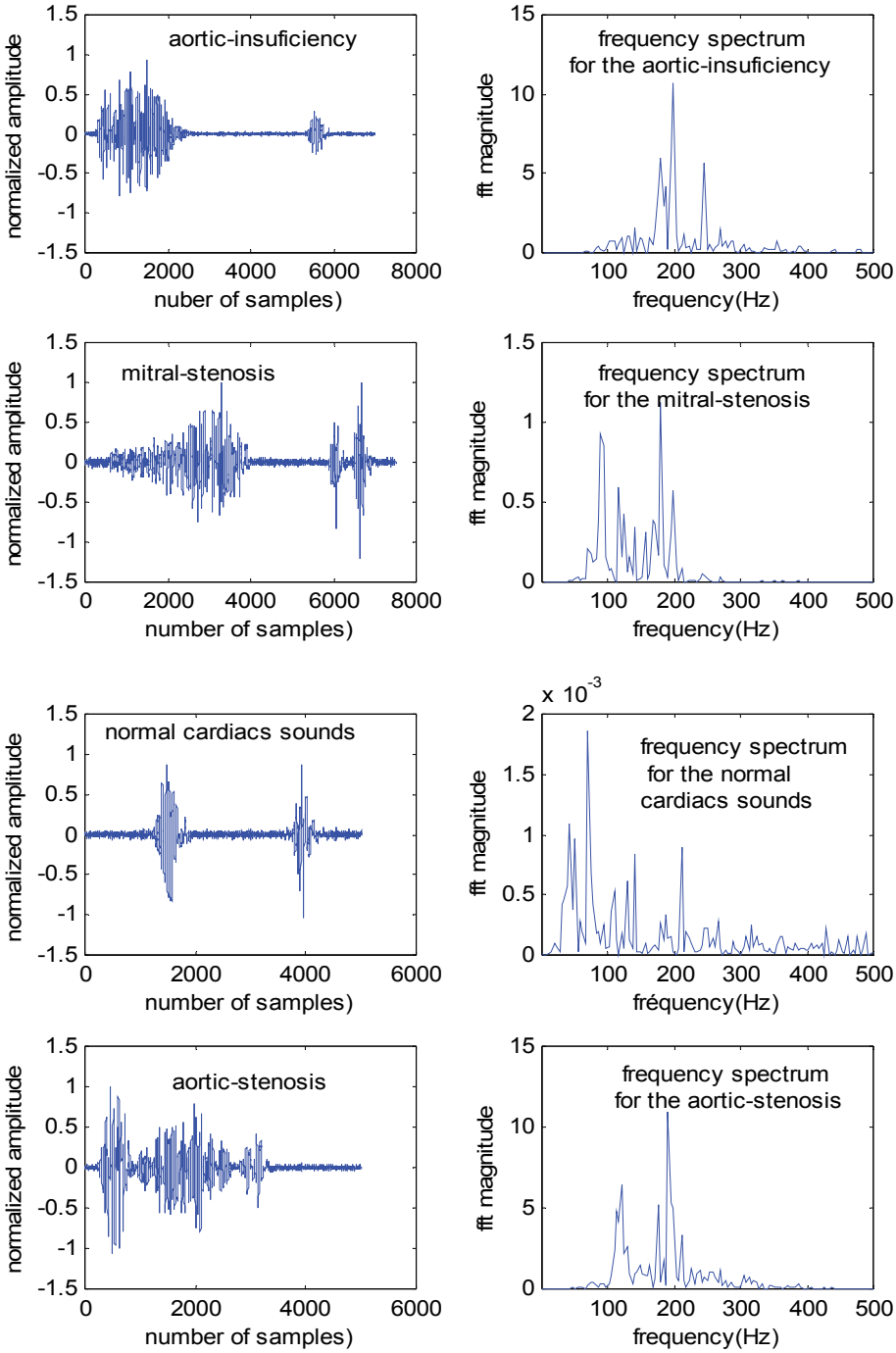


Fig. 5. Normal and abnormals cardiacs sounds and their frequency spectrum respectively

As the frequency extent of the sound S1 is less important than that of the sound S2, the spectral response of the PCG signal related to this pathological case is not much affected comparad to that of the normal case as was the case in the aortic -insufficiency and aortic-stenosis

In conclusion, and by applying the spectral analysis to different PCG signals, we can affirm which of the sounds S1 or S2 is directly concerned by the pathology, and more precisely which component of these sounds is affected.

With regard to normal PCG the basic frequency components are obviously detected by the FFT but not the time delay between these components. In fact as it was shown for example in Figure4c, the components A2 and P2 of the second sound S2 are obvious. However the FFT analysis of S2 cannot tell what is the value of the time delay between A2 and P2. It is thus essential to look for a transform which will describe a kind of "time-varying" spectrum.

3.2 Short-time Fourier Transform application

The normal phonocardiogram signal (Figure 6a) and the coarctation of the aorta (Figure 6a) are analyzed in this section. The coarctation of the aorta has been deliberately chosen here to evaluate the performance of the STFT analysis for it is very similar to the normal case. Figures 6a illustrate thus such a signal where we can notice that the temporal representation is almost similar to that of normal case given in Figure 6a.

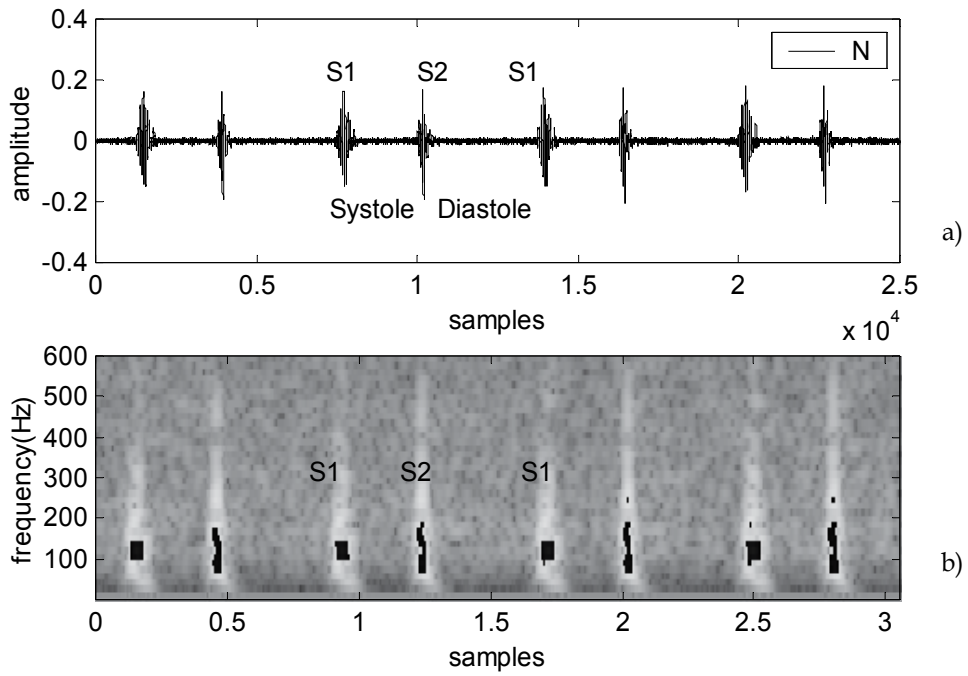


Fig. 6. The STFT application of the normal PCG (N); a) the normal PCG signal (N); b) the spectrogram of the normal signal (N)

Figure 6b and Figure 7b provide respectively the STFT application of the normal and the CA case. From these Figures we can see the difference of the time-frequency features between

them. For the two cases (N and CA), the second sound (S2) is shown to have higher frequency content than of the first sound (S1) [14,15]. This expected since the amount of blood present in the cardiac chambers is smaller [16].
We consider here two examples of the phonocardiogram signals with murmur: the pulmonary stenosis (PS :Figure 8a) with a systolic murmur and the aortic regurgitation (AR : Figure 9a) with a diastolic murmur.
Figure 8b and Figure 9b shows respectively the STFT application of the PS and AR signals. We can notice that the frequency extent of the diastolic murmur of the AR case is highly (About 600Hz) than the systolic murmur of the PS case (about 400 Hz).

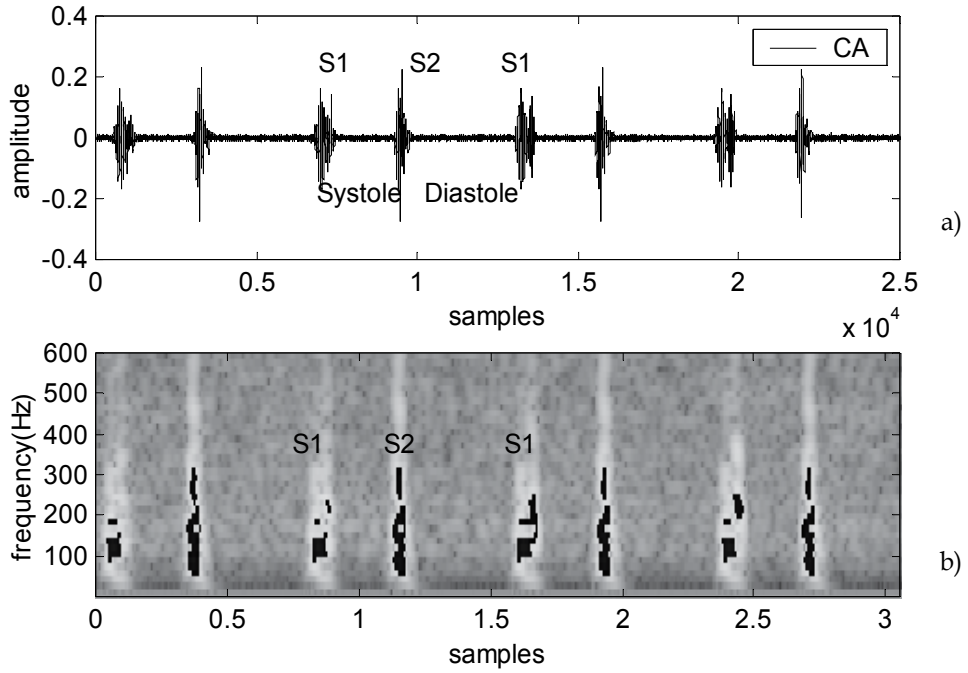


Fig. 7. The STFT application of the Coarctation of the aorta (CA); a) the Coarctation of the aorta signal (CA); b) the spectrogram of the signal CA

In this section we presents the experimental results of the short-time frequency transform application of the three followings groups of the PCG signals used.

- Group 1 : normal (N) or similar morphological signal (CA) ;
- Group 2 : opening snap (OS) and ejection click (EC);
- Group 3: PCG signal with width murmur (PS and AR).

Figure 10 provides a better representation of the results obtained concerning the frequency contents of the sounds and murmurs analysed. If under the normal conditions (N) or in the presence of similar signals (CA) the frequency content of the sound S2 is more significant than that of the sound S1.
We noted that the light murmurs (OS, EC.) can influence the time-frequency content of the principal sounds S1 and S2 and have a frequency extent more significant than them.

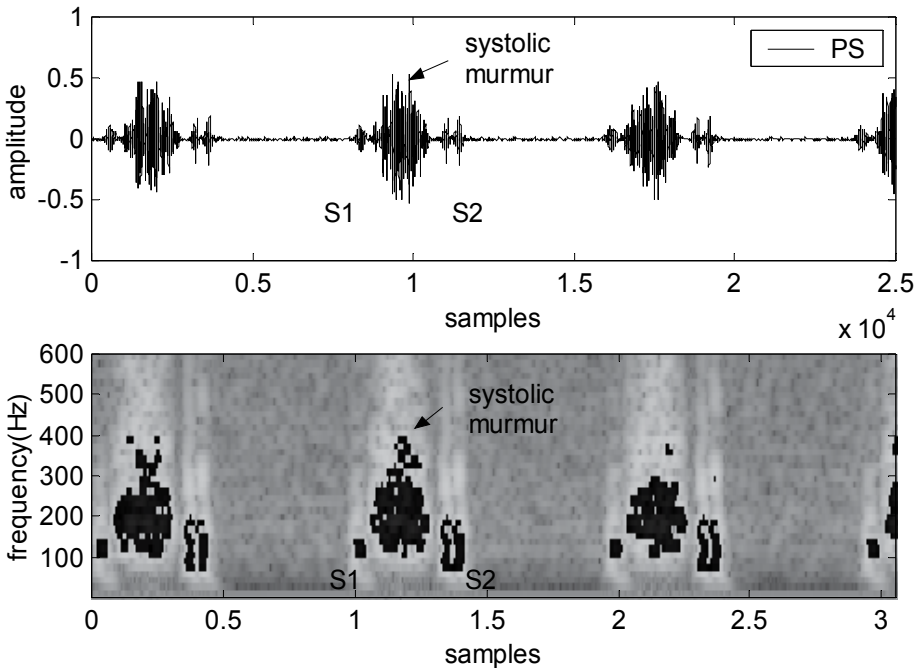


Fig. 8. The STFT application of the pulmonary stenosis; a) The pulmonary stenosis signal (PS); b) The spectrogram of the signal PS

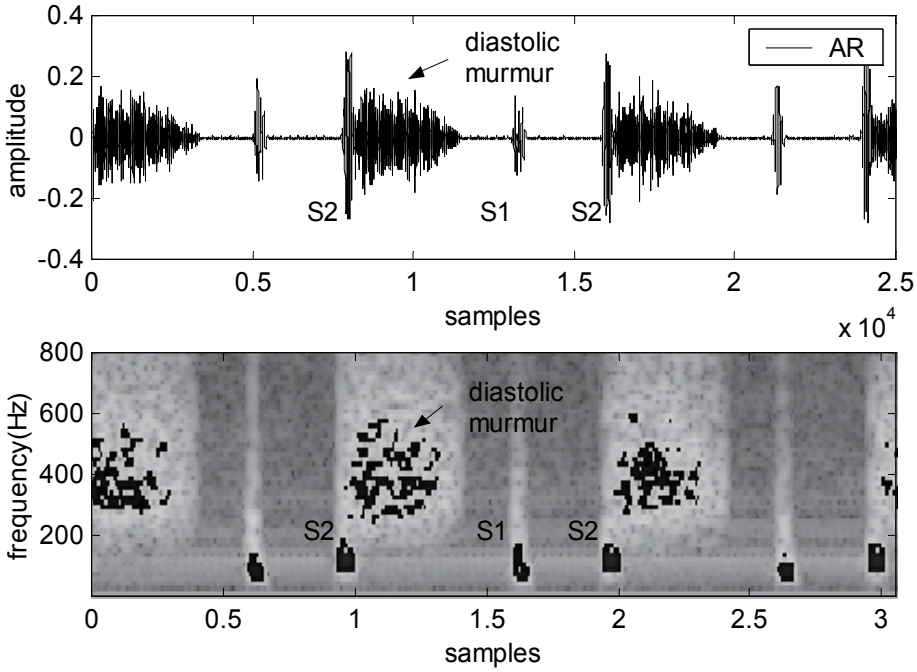


Fig. 9. The STFT application of the aortic of the regurgitation; a)The aortic of the regurgitation signal (AR); b)The spectrogram of the signal AR

Finally in fact the width murmurs (PS and AR cases) present a frequency extent very significant. Discrimination between the systolic and diastolic murmurs can be made starting from this frequency extent, diastolic murmurs thus having a frequency extent more significant than the systolic murmurs. In more these murmurs seem not too much not to affect the time-frequency content of the sounds S1 and S2

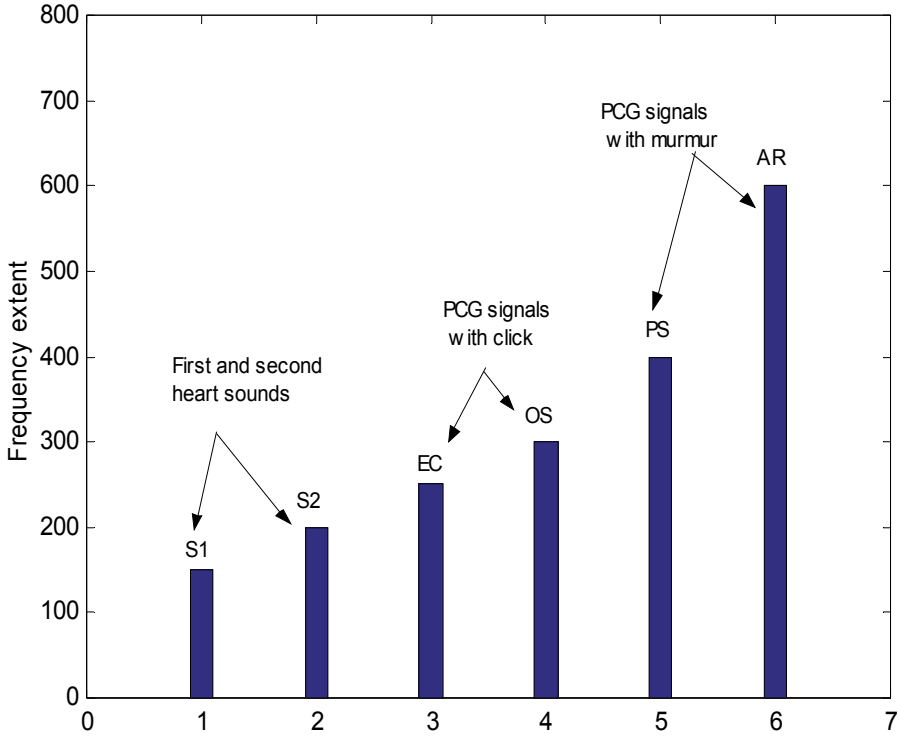


Fig. 10. Frequency extent of the three groups of the PCG signals

3.3 Wigner Distribution application

Figure11 shows the WD application of the cardiac sound of the normal case (11a), the coarctation of the aorta (11b) and the innocent murmur (11c). One can notice here that the two principal components (A2 and P2) start to appear in the presence cross-terms. The WD results may be improved by increasing the sampling rate of original signal,, but it still suffers from the cross-terms problem because of the nonlinearity of the WD.

However the Wigner distribution (WD) have shown good performances in the analysis of non-stationary signals. This comes from the ability of the WD to separate signals in both time and frequency directions. One advantage of the WD over the STFT is that it does not suffer from the time-frequency trade-off problem. On the other hand, *the WD has a disadvantage since it shows cross-terms in its response. These cross-terms are due to the nonlinear behaviour of the WD, and bear no physical meaning*. One way to remove these cross-terms is by smoothing the time-frequency plane, but this will be at the expense of decreased resolution in both time and frequency

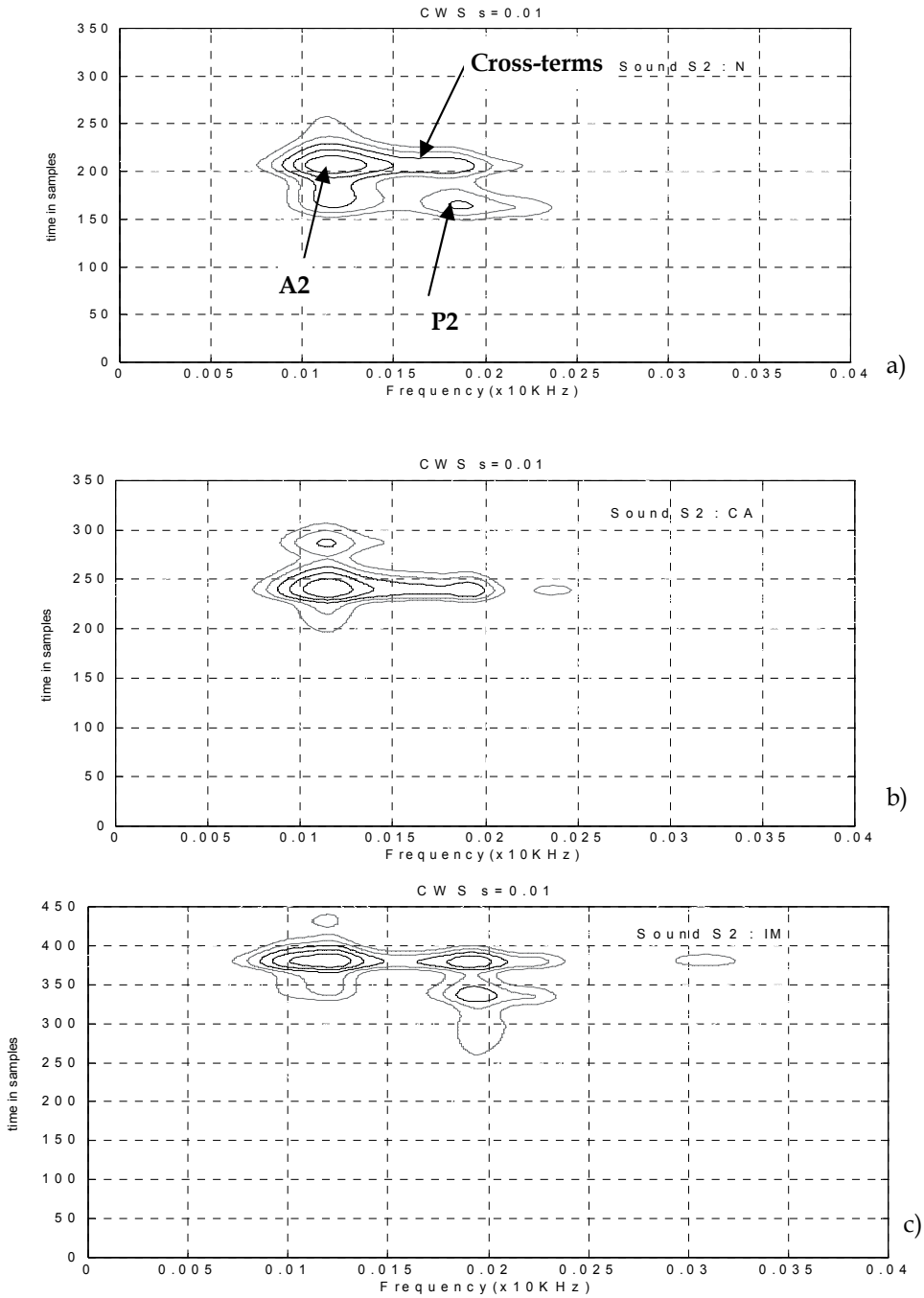


Fig. 11. The wigner distribution of the second cardiac sound S2 : a) the normal case , b) the coarctation of the aorta case (CA), c) the innocent murmur case (IM)

3.4 Continuous Wavelet Transform application

An algorithm under MATLAB environnement of the Continuous Wavelet Transform is developed then applied to analyse the different PCG signals. First the normal PCG is studied as illustrated by Figure12.a. The two heart sounds are clearly shown in dark color. They are spaced with 2500 samples corresponding to 0.312 secondes. The continuous wavelet transforms of S1 and S2 are also displayed separately in Figure12.b and Figure12.c respectively. The sound S2 is shown to have higher frequency content than that of the S1. This is expected since the amount of blood present in the cardiac chambers is smaller [1],[16]. On Figure 12b we can see that S1 is clearly resolved in two major component (M1 and T1). On the figure 10c the sound S2 is resolved also into two major's components (A2 and P2). The time delay between A2 and P2 can be easily measured with the use of the wavelet coefficients (Figure12c). This delay is measured to be 13ms. It is smaller than the 30ms [17] as foreseen in the normal conditions of the PCG signal. Pathological conditions could cause this time difference to narrow or widen. Moreover, the order of occurrence of A2 and P2 may be reversed. The wavelet transform allows measurement and determination of this time difference, and thus allows a diagnosis process regarding this important parameter to be produced. Table I resumes the differences observed between the components A2 and P2.

	Localisation of the Delay "d" (ms)	Minimal frequency (in scale)	Maximal frequency (in scale)	Frequency Extent
A2	13	19	124	105
P2	13	18	116	98

Table I. Temporal and frequential measurements related to the componentsA2 and P2.

It can be concluded for the normal PCG that :

1. The component A2 precedes in time the component P2.
2. The component A2 have higher frequency content than P2.
3. The amplitude of A2 is more important than that of P2.

These parameters, particularly the frequency, make it possible to see a difference between A2 and P2.

Moreover the ability of the wavelet transform in heart disease diagnosis can be studied by applying the CWT algorithm on different marked cases. The result of this application are illustrated in Figure13b (aortic-insufficiency), Figure13c (aortic-stenosis) Figure13d (mitral-stenosis). The coefficients of the CWT allow us to clearly discern the frequency range of each signal. It also shows the major components according to the temporary variation ; the maximal amplitude is characterised by a darker color than those of the small amplitudes.

3.5 Discrete Wavelet Transform application

The multiresolution analysis based on the discrete wavelet transform (DWT) is a powerful tool in and filtering, separating and identification of the internal components and murmurs of the various analyzed signals (Figure14).

Figure15a shows the application of the discrete wavelet transform of one cycle of the normal PCG. Levels d1 and d2 represent the high frequency variations of the base line of PCG signal. Levels d3 and d4 emphasize clearly the side of high frequency content of the sound S2 compared to S1The component A2 (the most predominant in sound S2) appears better on levels d5 and d6. Level d7 represents the two principal components of the sound S1 (M1, T1).

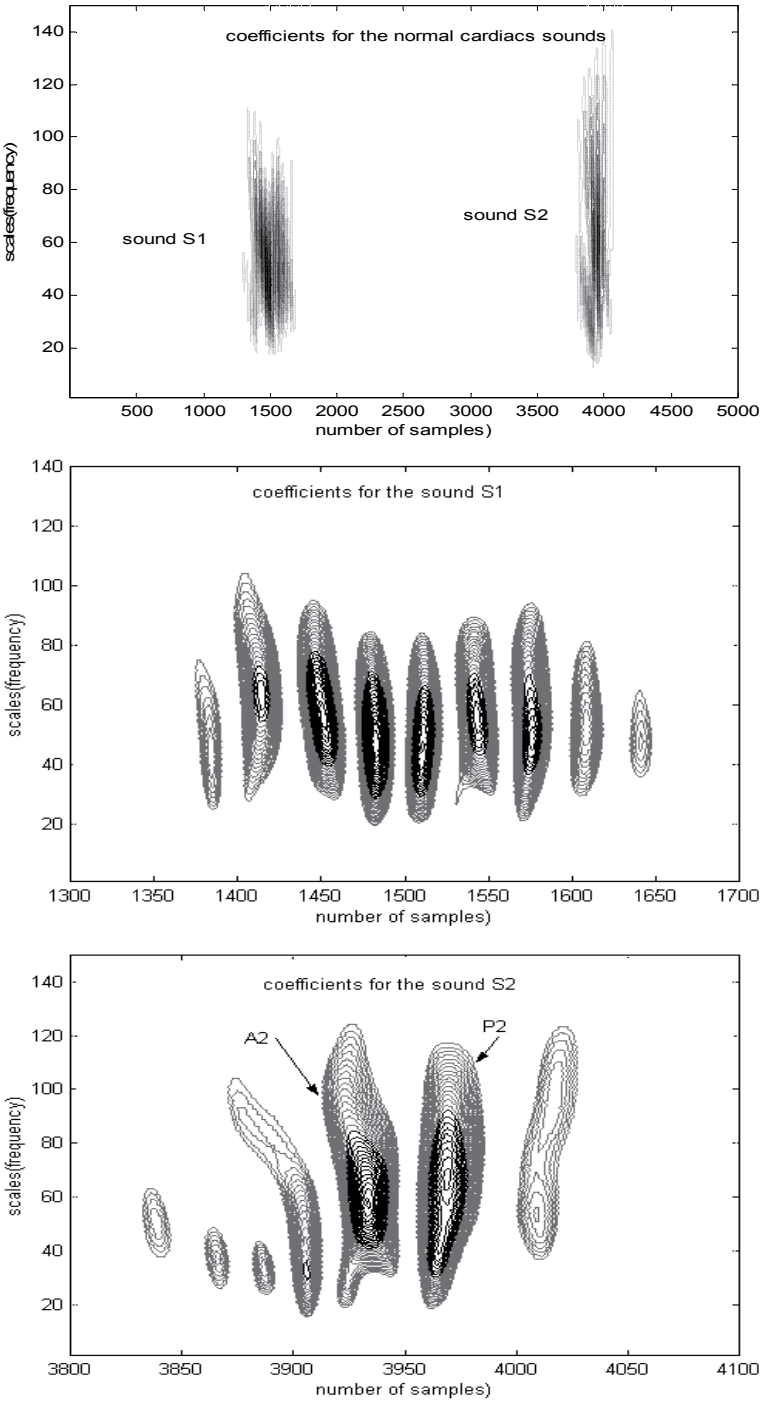


Fig. 12. Continuous wavelet transform analysis (CWT) for the normal cardiac sounds S1 and S2

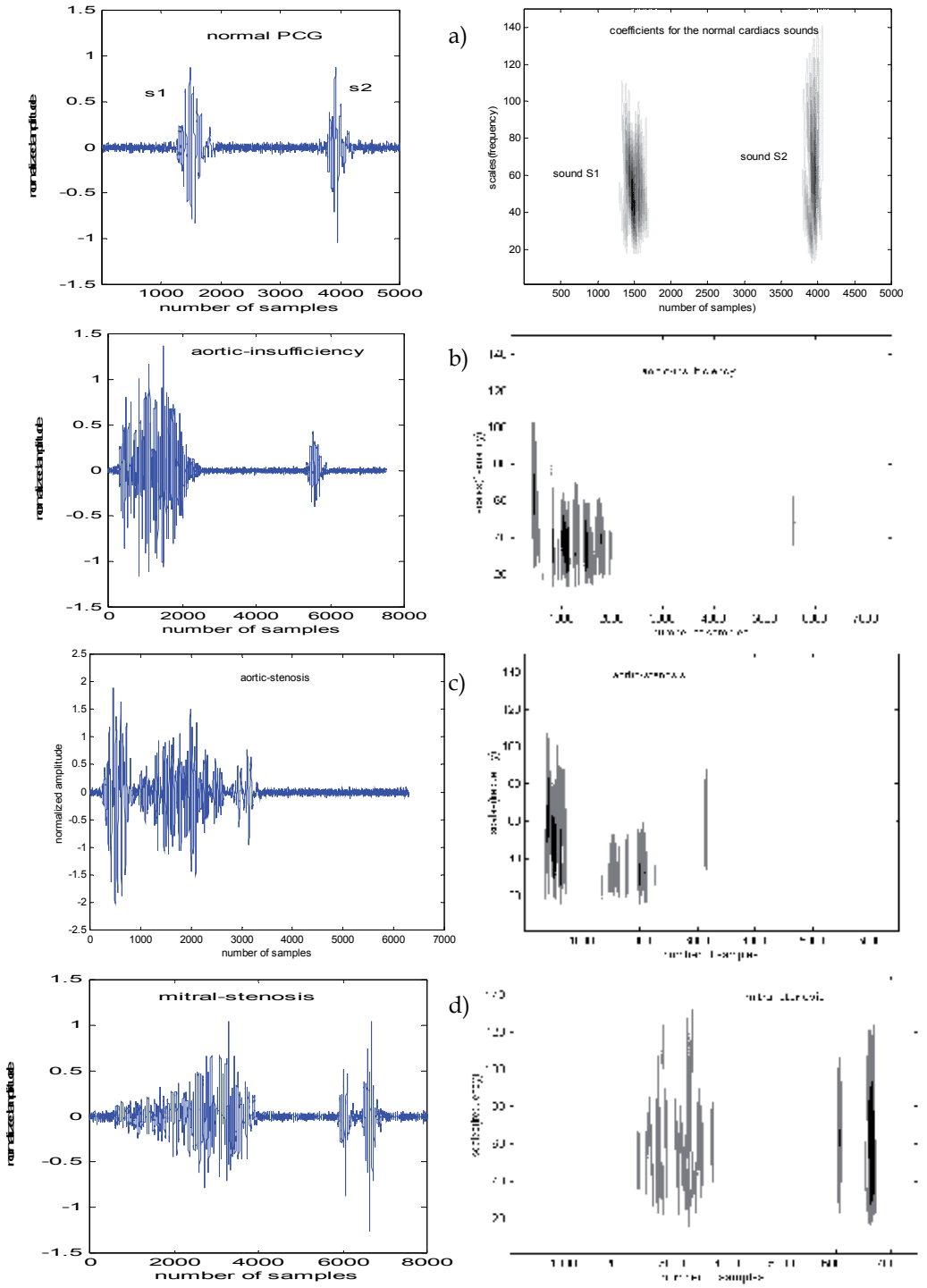


Fig. 13. Continuous wavelet analysis (CWT) for the normal PCG and abnormal PCG.

The same analysis by using the discrete wavelet transform concerning the coarctation of the aorta and the mitral stenosis is also applied (Figures16b and Figure16c). Figure 13b, in level d7, provide a well representation of the third component "C" added a the two majors components of the sound S2 (A2 and P2).

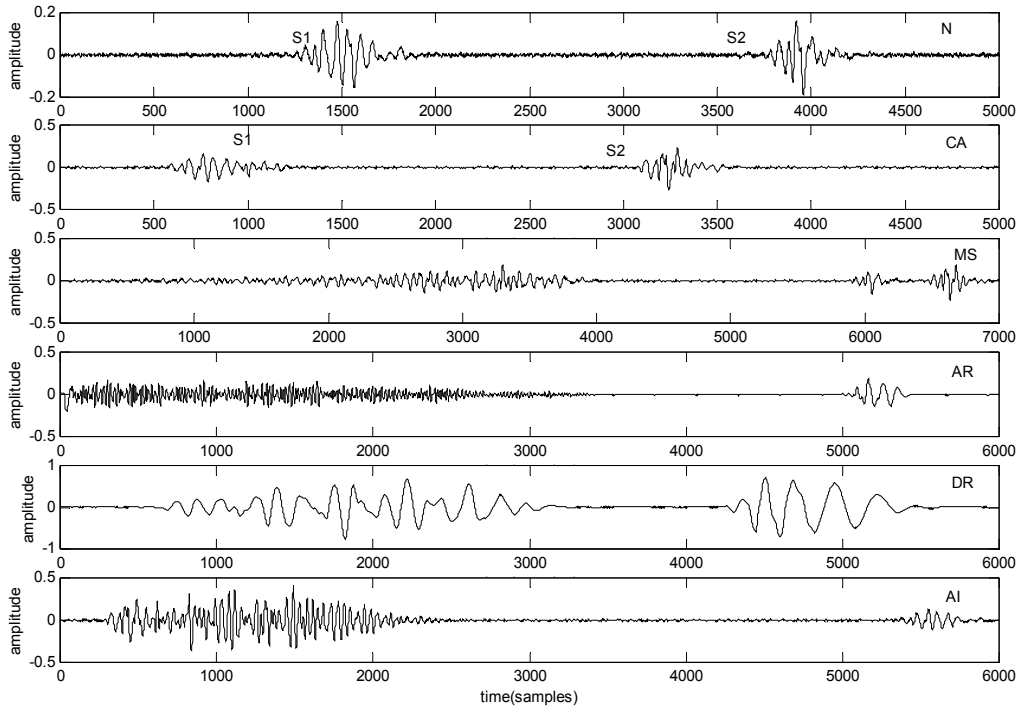


Fig. 14. PCG signals used a) normùal (N), b) the coarctation of the aorta (CA); c) the mitral stenosis (MS), d) the aortic regurgitation (AR), e) the diastolic ruble (DR) , f) the aortic insufficiency (AI)-

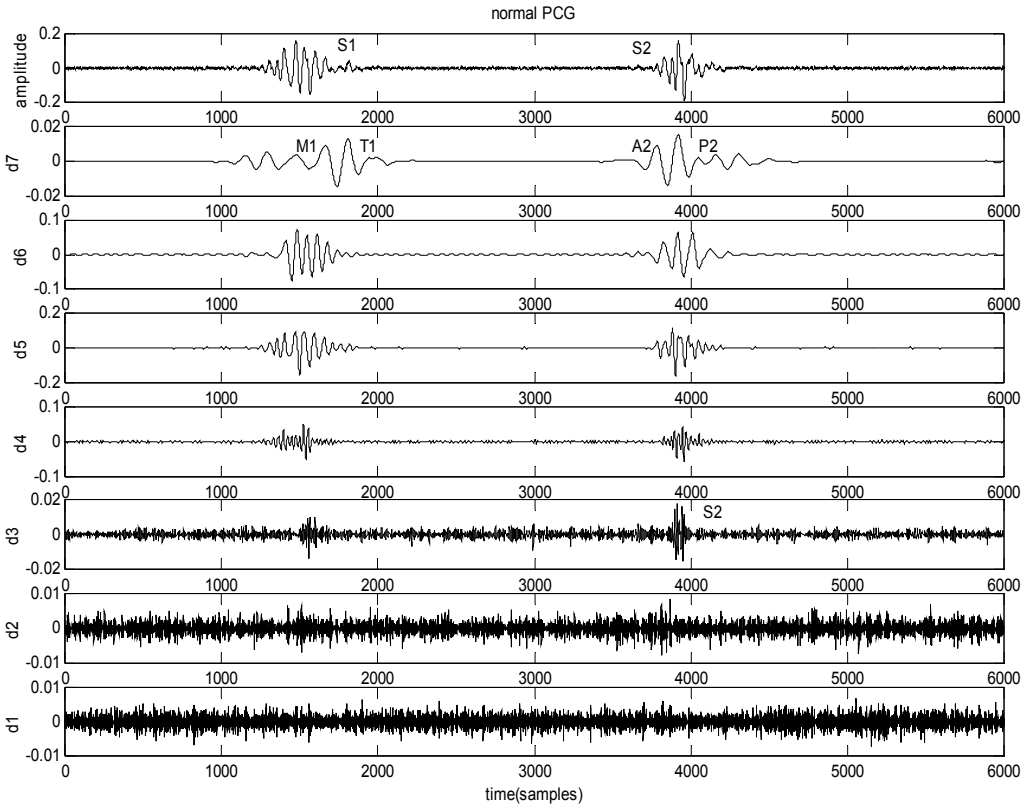
3.6 Packet wavelet transform application

The used of the wavelet packet transform is exactly the same as those developed in the discrete wavelet transform. The only difference is that wavelet packets offer a more complex and flexible analysis, because in wavelet packet analysis, the details as well as the approximations are split (Figure3). Single wavelet packet decomposition gives a lot of bases from which you can look for the best representation with respect to a design objective. The *wavelet packet* method is thus a generalization of wavelet decomposition that offers a richer range of possibilities for signal analysis. In wavelet analysis, a signal is split into an approximation and a detail. The approximation is then itself split into a second-level approximation and detail, and the process is repeated.

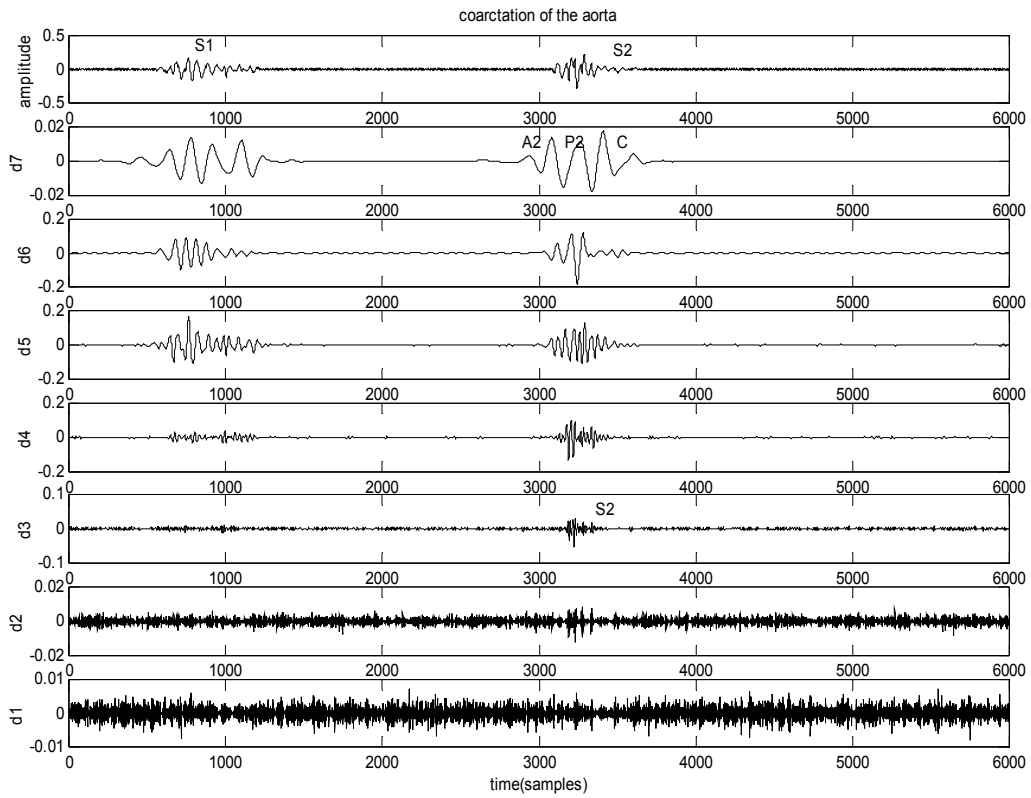
The wavelet packet transform analysis in this paper give important features of the extent frequency of the heart sounds (S1 or S2) and cardiac murmur. These features can be help clinician in their diagnosis or recognizing pathological conditions concerning the recording PCG signals.

Figure 16 provide a time frequency (TF) representation of one cardiac cycle of the heart sounds concerning the normal PCG signal (Figure 16a), the coarctation of the aorta case (Figure 16b) and the mitral stenosis case (Figure 16c). These figures shown the frequency range of each component or murmur of the PCG signal studied. Thus we can observe the component "C" added of the two majors component for the sound S2 (A2 and P2). The same result have been find in the section 3.5 (Figure 14b.detail d7)

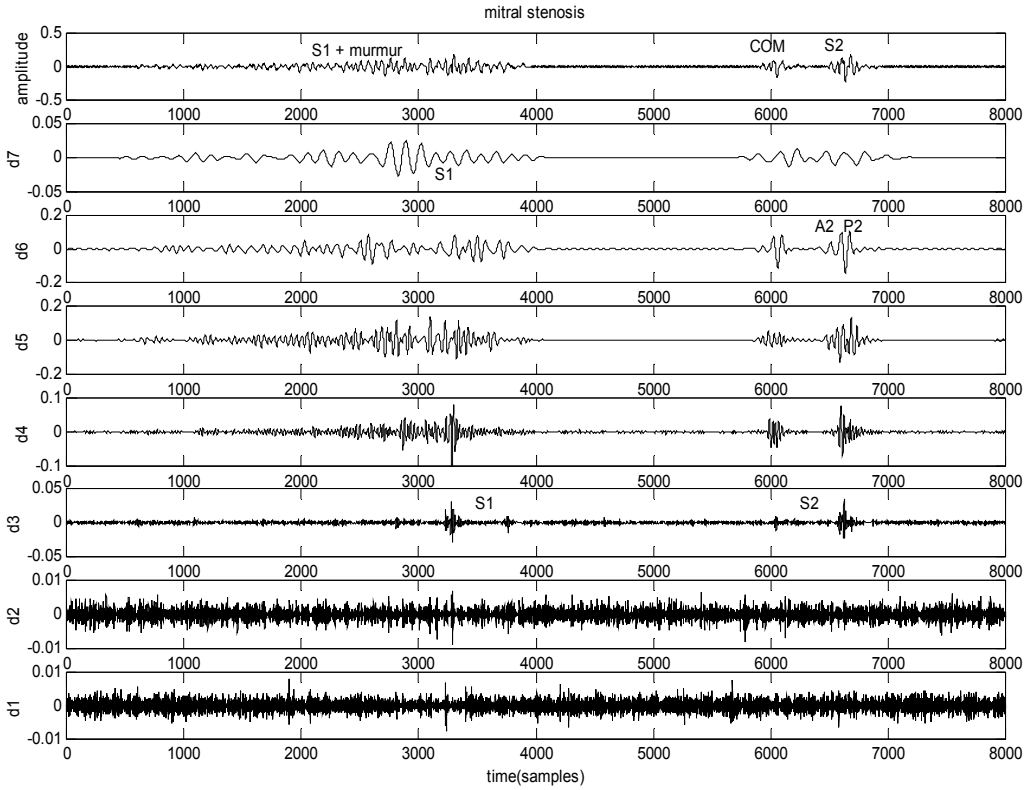
Figure 17 has the advantage of presenting at the same time the frequency extent of the various components of the cardiac sound like their frequency site one compared to the other. We can shown clearly that the Diastolic ruble (DR) case have high frequency content that the aortic insufficiency case (AI) or the aortic regurgitation case (AR).



a)

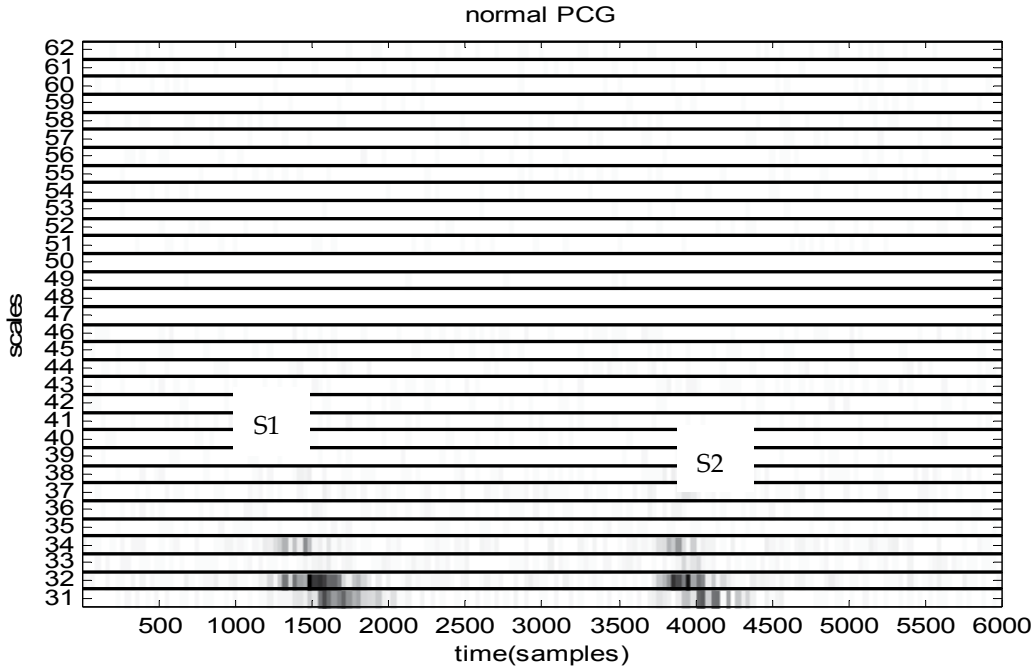

coarctation of the aorta
amplitude
S1
S2
A2
P2
C
d7
d6
d5
d4
d3
d2
d1
time(samples)

b)

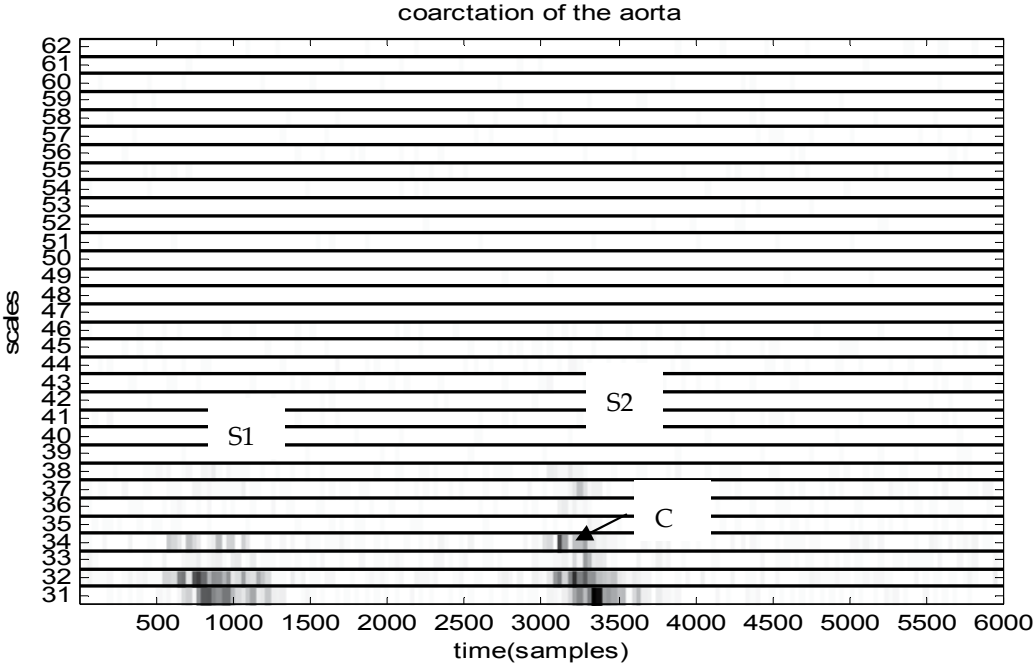


c)

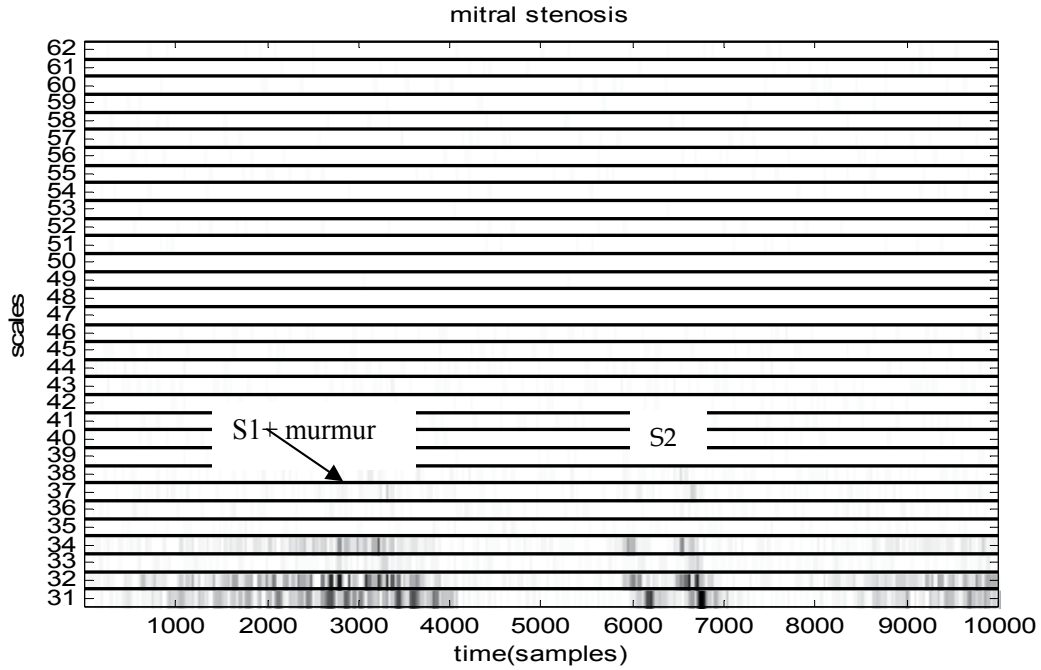
Fig. 15. Discrete wavelet transform (DWT) analysis for: a) the normal PCG, b) the Coarctation of the aorta, c) the mitral stenosis


normal PCG
scales
time(samples)
S1
S2

a)

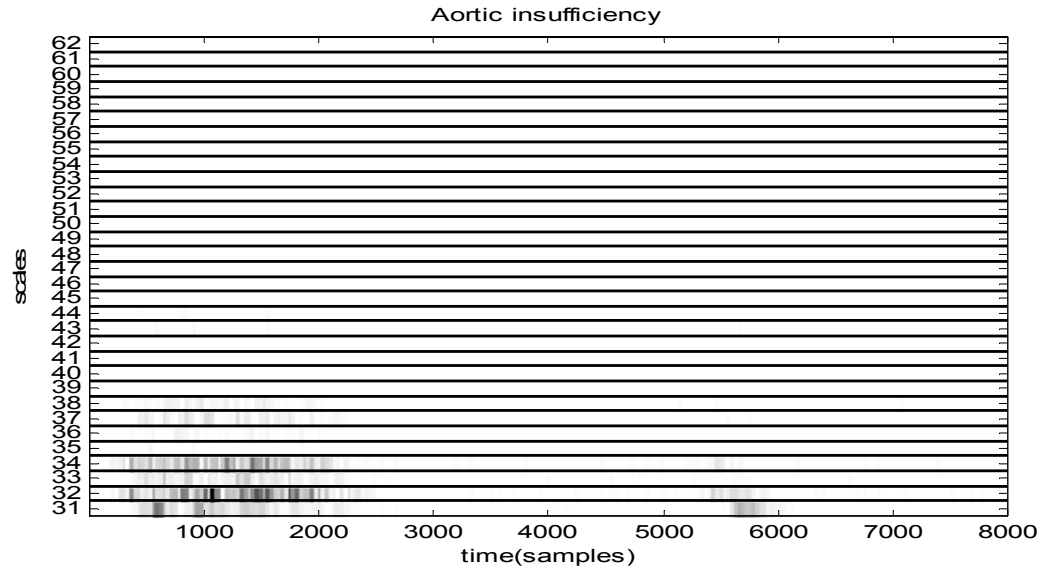

coarctation of the aorta
scales
time(samples)
S1
S2
C

b)

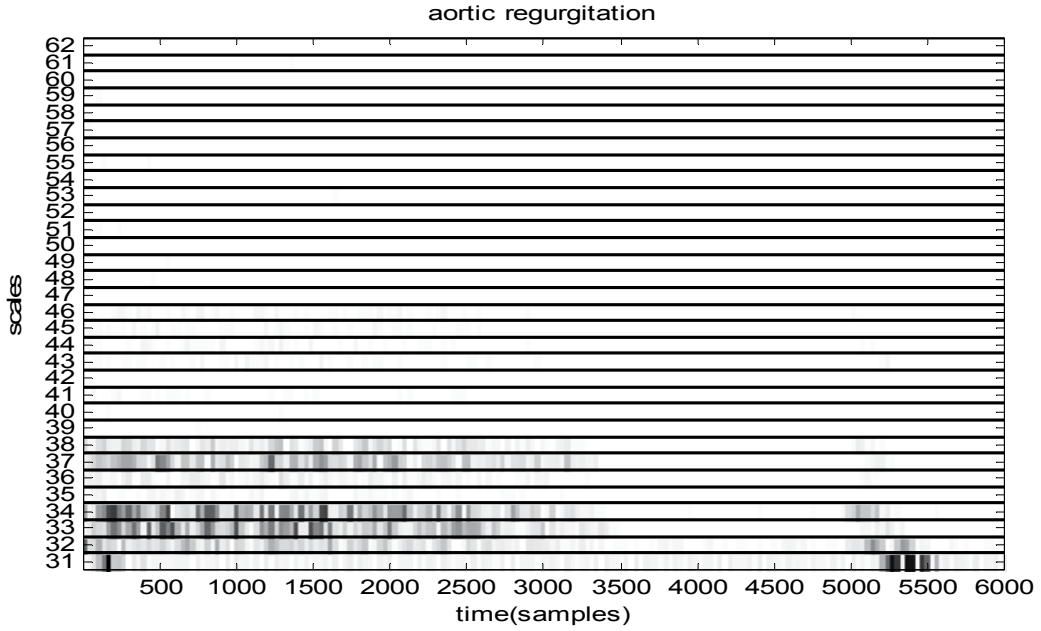


c)

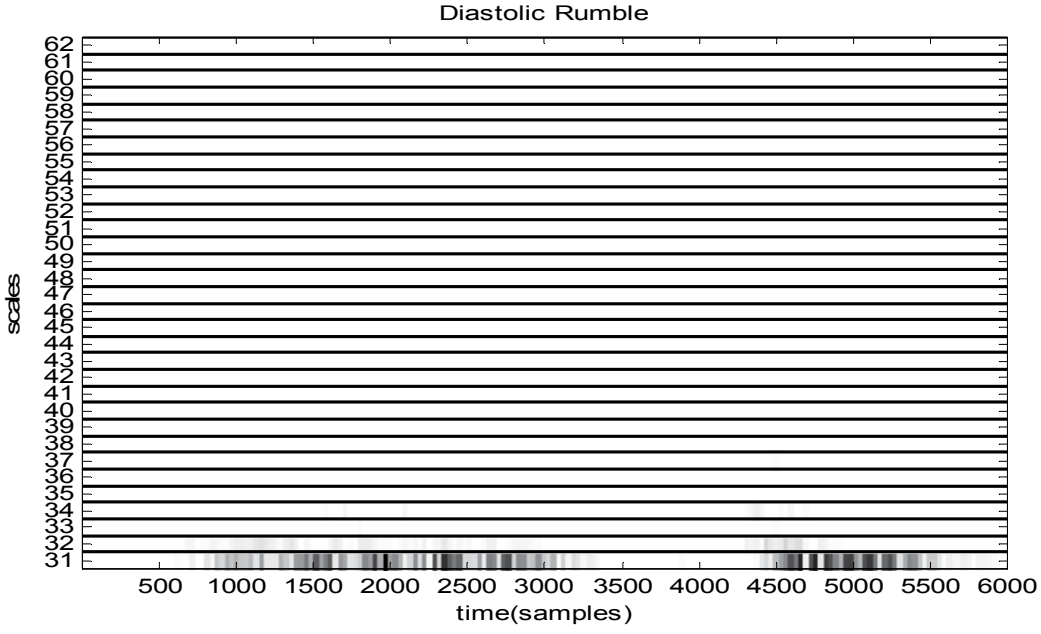
Fig. 16. Wavelet packet transform (WPT) analysis for: a) the normal PCG, b) the coarctation of the aorta , c) the mitral stenosis



a)



b)



c)

Fig. 17. Wavelet packet Transform (WPT) analysis for: a) the aortic Insufficiency (Ai), b) the aortic regurgitation (AR), c) the Diastolic Rumble (DR)

4. Conclusion

The cardiac (heartbeat sound) cycle of phonocardiogram (PCG) is characterized by transients and fast changes in frequency as time progresses. It was shown that basic frequency content of PCG signal can be easily provided using FFT technique. However, time duration and transient variation cannot be resolved; the CWT wavelet transform therefore i is a suitable technique to analyse such a signal. It was also shown that the coefficients of the continuous wavelet transform give a graphic representation that provides a quantitative analysis simultaneously in time and frequency. It is therefore very helpful in extracting clinically useful information.

The measurement of the time difference between the A2 and P2 components in the sound S2, the number of major components of the sounds S1 and S2 and the frequency range and duration for all these components and sounds can be accurately achieved for the CWT simultaneously as was clearly illustrated.

It is found that the wavelets transform is capable of detecting the four major components of the first sound S1 and the two components (the aortic valve component A2 and the pulmonary valve component P2) of the second sound S2 of a normal PCG signal. These components are not accurately detectable using the STFT or WD. However the standard FFT can display the frequencies of the components A2 and P2 but cannot display the time delay between them.

The application of the STFT in the analysis of the PCG signals made it possible to obtain appreciable information on the time-frequency content of the sounds S1, S2 and of the added murmurs (OS,EC or width murmurs).

If under the normal conditions (N) or in the presence of similar signals (CA) the frequency content of the sound S2 is more significant than that of the sound S1. We noted that the light murmurs (OS, EC.) can influence the time-frequency content of the principal sounds S1 and S2 and have a frequency extent more significant than them.

Finally in fact the width murmurs (PS and AR cases) present a frequency extent very significant. Discrimination between the systolic and diastolic murmurs can be made starting from this frequency extent, diastolic murmurs thus having a frequency extent more significant than the systolic murmurs. In more these murmurs seem not too much not to affect the time-frequency content of the sounds S1 and S2.

The two version of analysis of the wavelet transform (DWT and PWT) make it possible to gather time-frequency information concerning the characteristics of the cardiac sounds.

It is shown that the FFFT, the STFT, the WD and the WT techniques provides more information of the PCG signals with murmurs that will help physicians to obtain qualitative and quantitative measurements of the time and the time-frequency PCG signal characteristics and consequently aid to diagnosis.

5. References

[1] Rangayyan,R.M and Lehner ,R.J (1988). A review.CRC Critical Reviews in Biomedical Engineering 15 (3), 211-236.

[2] Rangayyan,R.M and Lehner ,R.J (1988). Phonocardiogram signal analysis: a review. CRC Critical Reviews in Biomedical Engineering 15 (3), 211-236.

[3] Feigen , L.P (1971). Physical characteristics of sound and hearing. American journal of Cardiology, 28 (2), 130-133 .

[4] Tuteur, F.B (1988). Wavelet Transforms in signal detection. IEEE ICASSP, CH2561-9, 1435-1438.

[5] Grossmann.A; Holschneider Kronland-Martinet.R and Morlet,J (1987). Detection of abrupt Changes in sound signal with the help of the wavelet transform. In: Inverse problemes: An interdisciplinary study. Advances in Electronics and Electron physics. Supplement 19 [New York, Academic], 298-306.

[6] Obaidat.M.S. Phonocardiogram signal analysis: techniques and performance comparison. Journal of Medical Engineering & technologie, vol 17, No 6 (November-December 1993), 221-227.

[7] Boashas, B (1993). Time-frequency signal analysis. In advances in spectrum Estimation, edited by S.Haykin, (NJ: Prentice-Hall).

[8] William J.Williams (1997). Time-frequency and wavelets in Biomedical Signal Processing. Edited by Metin Akay. IEEE Press Serie in BME. 3-8.

[9] Olivier.R and Duhamel.P. Fast algorithms for Discrete and Continuous Wavelet Trnsforms. IEEE Transactions Information Theory, vol38, No.2, (march 1992), 569-586

[10] Harris, A., Sutton, G. And Towers, M. (1976). Physiological and clinical aspects of cardiac auscultation. Medicine Ltd, London, Uk.

[11] Y.Meyer (1990). Ondelettes et Opérateurs. Tome1. Paris: Hermann.

[12] S.Mallat (1989). A theory for multiresolution signals decomposition: The wavelet representation. IEEE Trans.Pattern Anal.Machine Intell, vol 11, pp 674-693.

[13] P.I.J. Keeton, F.S. Schlindwein, Application of Wavelets in Doppler Ultrasound, MCB University Press, Emerald Sensor Review, Vol. 17(1) 1997, pp. 38–45.

[14] S.M Debbal and F.Bereksi-Reguig. The Fast Fourier and the wavelet transforms analysis of the cardiacs sound. *Physical and Chemical News (PCN)*.vol 15, pp 54-59, 2004.

[15] S M Debbal – F Bereksi Reguig: « *The fast Fourier transform and the continuous wavelet transform analysis of the phonocardiogram signal* »; Journal of Mechanics in Medicine and Biology (JMMB); ISSN: 0219-5194; vol.4, No 3, September 2004 pp 257-272.

[16] Bentley P.M, Grant A, Mc Donnel JTE. Time-frequency and Time-scale for the classification of native and Bioprosthetic valve sounds. *IEEE Trans. Biomed. Eng. 45(1), pp 125-128, 1998.*

[17] Djebbari.A (1999). Synthèse des méthodes d'analyse temporelle, spectrale et spectro-temporelle du signal phonocardiogramme. Electronic magister thesis of signals and systems. Departement of electronics, faculty of science engineerig, university Aboubekr belkaid Tlemcen, Algeria, 18-24

Part 2

Speech Analysis

6

Modelling and Understanding of Speech and Speaker Recognition

Tilendra Shishir Sinha[1] and Gautam Sanyal[2]
[1]*Computer Science & Engineering Department, DIMAT*
Raipur, Chhattisgarh State
[2]*Computer Science & Engineering Department, NIT*
Durgapur, West Bengal
India

1. Introduction

The main goal of automatic speech and speaker recognition (ASSR) is to transcribe natural speech and recognize its speaker. Recognizing a spoken sentence is obviously a knowledge-intensive process, which must take into account all variable information about the speech communication process. Out of several approaches, artificial intelligence approach has been observed to give remarkable results by visualizing, analysing and finally making a decision on the acoustic patterns. The present chapter deals with two basic processes for ASSR: *modelling* process and *understanding* process. The modelling process involves three stages: enhancement, segmentation and pre-processing. The understanding phase involves the recognition of the speech and the speaker through the known knowledge-based model. It has been observed from the literature that, in any speech processing system, because of channel coupling, many source signals mix together. Thus there may be a chance of variability in speech signals affecting the performance of the speech systems due to some factors like: background and channel noise, electrical noise from different sources, meaningless sounds (a sneeze) or filler words ('uh' or 'um') between words, different speaking rate, mood and styles of speakers. Due to this some signals may not be observed and may result to out-of-vocabulary (OOV) word problem during recognition or understanding process. Thus speech enhancement has played a vital role in the development of a perfect acoustic model or noise – free artificial word model (AWM) and vowel diphthong model (VDM). To process for speech enhancement in frequency – domain, the speech signal has to be split into frames using blind signal separation (BSS) method (Yilmaz and Richard 2004; Yamashita and Hirai 2004; Araki et al 2002; Murata et al 2001; G. J. Jang and Lee 2003; F. Bach and Jordan 2005; B. A. Pearl Mutter and Olsson 2006). Further the spectrum of the background noise has to be estimated by subtracting the noise spectrum from the spectrum of the frame (M. Berouli et al 1979; Yasser Ghanbari and Md. Reza 2004). It has been observed that some of the noise remains in the spectrum when the value of noise is greater than its mean. At the same time, some of the speech spectrum gets removed when noise is greater than the actual value of noise. In the spectrum this produces negative values and have to be set to zero. The overall effect puts a noise in the output signal known as

residual noise. To reduce the level of residual noise further subtraction has to be carried out by adjusting the 'over-subtraction' or 'over-estimation' factor with respect to signal to noise ratio as suggested by Berouli et al (1979). Before processing the speech any further, the wavelet (a real valued function of time) in a noisy signal (if any) has to be detected, using discrete wavelet transform (DWT) method. Further, analysis has to be done for the speech segmentation, which means partitioning an entire speech into isolated sub-words with optimal boundaries (Shriberg 2000; Abdulla 2002; Delacourt and Wellekens 2000; Shafron and Rose 2003). In segmental modelling, speech parameters are represented by trajectories, that is, sequences of points in the parameter space. The speech trajectories have to be characterized using the mean, varian and shape of the particular segment. The shape of the signal has to be obtained using the wavelet coefficient as estimated earlier. Constructing a probability density function and using adaptive vector quantization over the training vector data set speech segmentation has to be done. Later on, loss-less compression methods: discrete cosine transform (DCT) and principle component analysis (PCA) have to be employed. Next the pre-processing (that is extraction of speech features) has to be done using hybrid approach of soft-computing techniques. Here the hybrid approach resemblances to artificial neural network (ANN) and genetic algorithm (GA). The hybrid approach has to be applied in a well-defined way in the present chapter that has been illustrated in the subsequent paragraphs.

2. Modelling of AWM and VDM

For the formation of AWM and VDM (as shown in figure 1), good distortion-free features have to be extracted. If this fails then the recognition process also degrades or suffers.

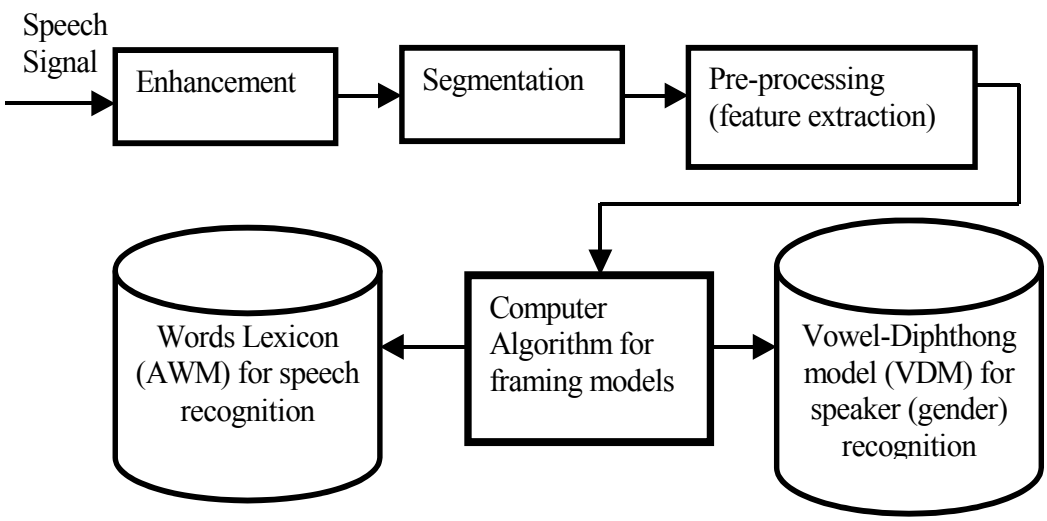


Fig. 1. General outline for framing models (AWM and VDM)

The most important speech features that have to be extracted are: pitch or fundamental frequency, formant frequencies or energy values, speaking rate, speech duration and so on, for the formation of models (that is AWM and VDM). Based on these features, twenty-two parameters have to be extracted for further analysis. The pitch related parameters that have

to be extracted are: mean, median, standard deviation, minimum and maximum range of pitch. The energy related parameters that have to be extracted are: mean, median, standard deviation, minimum and maximum range of loudness (energy). The duration related features that have to be extracted are: the ratio of voiced and unvoiced speech, speech rate. The speech rate has to be calculated by taking the ratio of duration of voiced speech to the total number of words uttered. All the above features have to be extracted using cepstral analysis, particle-filtering (Monte-Carlo) method from the spectrogram. The spectrogram based pitch and formant detection, have to be carried out (Hue et al 2001; Gustafson 2002; Arulampalam 2002; Vermaak 2002; Welling et al 1998). The formant frequency plays a vital role in giving the details about the vocal tract shape and its movements in various pronunciations. More illustrations have been done in the next subsequent subsections of this chapter.

2.1 Extraction of speech features

Before the extraction of speech features and relevant parameters, first, frequency bands corresponding to the analysed formants have to be extracted using forward-backward dynamic programming (FBDP) method. Next a particle-filtering method has to be applied to locate formants in every formant area based on the posterior probability density function (pdf) described by a set of support points with associated weights. As per the work done by Acero (1999) and Watanabe (2001), it has been found that capturing and tracking formants accurately from natural speech are very difficult task because of the variety of speech sounds. Typically, formant-tracking algorithms have three phases: *pre-emphasis*, *frame-dependent formant candidates*, and *generation-and-tracking*. For the first two phases, hypothesis testing and cepstral analysis has to be adopted. The third phase has to be carried out using the spectrogram-based particle-filtering method, because it is well known that the horizontal bands in grey-scale spectrogram with higher energy show the formant positions, which can be easily tracked in the spectrum. The mathematical analysis for this has been illustrated in the subsequent sections of this chapter. The main idea behind an acoustic model, that has been depicted in figure 2, with a trained data set, arranged in word-map, has to be categorized to estimate the best parameters that define the distribution (namely the mean, variance, shape and so on), and are represented as 'w%' where % = 1 to maximum size of the vocabulary.

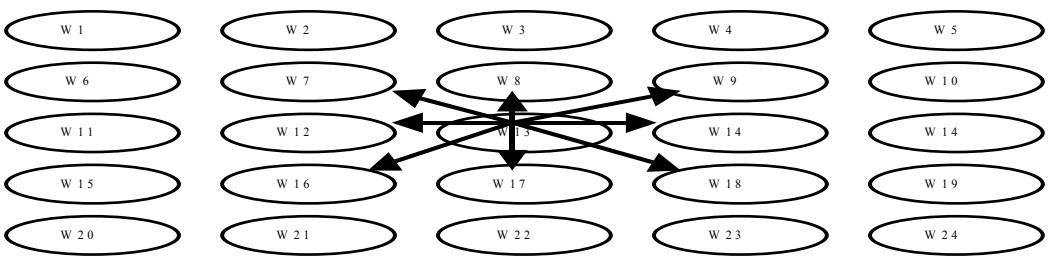


Fig. 2. A noise-free artificial word model (AWM)

The arrows are being designated as directions of constraints. In this model each phrase or word has to be linked with eight other corresponding word cell. Each word cell has to be trained and stored using forward-backward dynamic programming (FBDP) method. The

mathematical analysis has been also illustrated for three constraints in the subsequent sections of this chapter.

2.1.1 Formation of AWM and VDM with mathematical analysis

The speech production process is the initial step in the human speech communication system. This process is very complex and involves many components. During normal conversation, speaker stops or pauses for a while, because that time his or her brain might be busy in searching for an appropriate word from the vocabulary. If the brain has to be trained properly, every word cell in the brain gets activated and the speaker speaks continuously. If the training has not been done properly then a little amount of word cell gets activated and the speaker speaks with some pause. In a similar manner, the present experimental set-up has to be done, by forming an intermediate transient speech (ITRANS) table or master table.

To explain this formation of AWM in more illustrative way, initially the speech signal has to be captured through the microphone and background noise has to be removed successfully, using blind signal separation and spectral subtraction method. Then word boundaries have to be computed not only by using traditional zero-crossing measurement (ZCM) but also by using discrete cosine transform (DCT), because it has a strong energy compaction property. The main property is that it considers real-values and provides better approximation of a signal with fewer coefficients. With reference to figure 2, the model is being composed of many simple non-linear processors called neurons connected in parallel. Each neuron has an input and output characteristics and performs a computation or function of the form:

$$O_i = f(S_i) \text{ and } S_i = W^T X \quad (1)$$

where $X = (x_1, x_2, x_3, \ldots, x_m)$ is the input vector to the neuron and W is the weight matrix with w_{ij} being the weight (connection strength) of the connection between the j^{th} element of the input vector and i^{th} neuron. The f (.) is an activation or nonlinear function (usually a sigmoid), O_i is the output of the i^{th} neuron and S_i is the weighted sum of the inputs. A single neuron, as shown in figure 3, by itself is not a very useful tool for AWM formation.

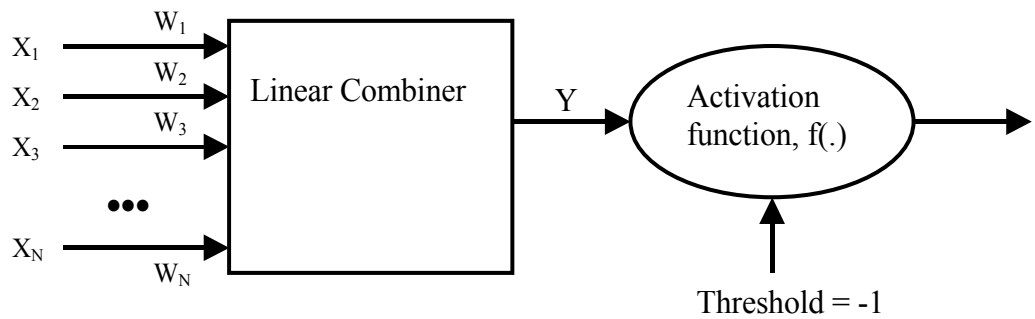


Fig. 3. An artificial neuron

The real power comes when a single neuron is combined into a multi-layer structure called neural networks (as shown in figure 4). The neuron has a set of nodes that connect it to the inputs, output or other neurons called synapses. A linear combiner is a function that takes

all inputs and produces a single value. Let the input sequence be $\{X_1,X_2,\ldots,X_N\}$ and the synaptic weight be $\{W_1,W_2,W_3,\ldots,W_N\}$, so the output of the linear combiner, Y, yields,

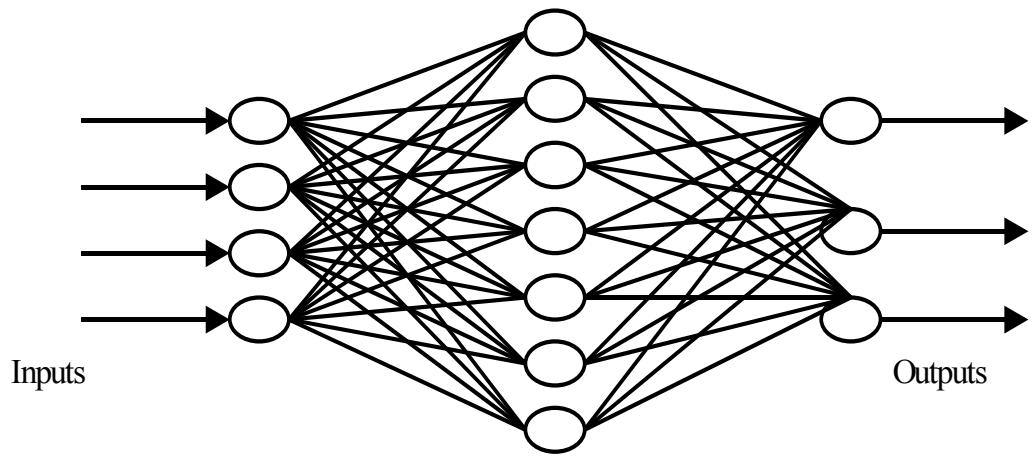


Fig. 4. A simple neural network

$$Y = \sum_{i=1}^{N} X_i W_i \tag{2}$$

An activation function will take any input from minus infinity to infinity and squeeze it into the range –1 to +1 or between 0 to 1 intervals. Usually an activation function being treated as a sigmoid function that relates as:

$$f(Y) = \frac{1}{1+e^{-Y}} \tag{3}$$

The threshold defines the internal activity of the neuron. This has to be kept fixed to –1. In general, for the neuron to fire or activate the sum should be greater than the threshold value. The learning capability is a result of the ability of the network to modify the weights through usage of a learning rule. Here, feed-forward network has to be used as a topology and backpropagation as a learning rule. A simple neural-network has been shown in figure 4, with inputs as speech feature values and a hidden and output layer.

The extracted speech feature values of each of the training sets have fed as input to the neural network. If 'ℓ' features have to be fed as input nodes then '2ℓ' nodes have to be used for the hidden layer. Each output neuron represents a word, thus only one output node has been treated. Here, twenty-two speech parameters have been extracted, hence twenty-two nodes in the input layer and forty-four nodes at the hidden layer and a single node at the output layer have been used. In order to minimize the error between the inputs and outputs of the neural-network, the weights of the threshold input has been adjusted using backpropagation algorithm (Wang et al 1999; Cybenko 1989; Hornik et al 1989; Ooyen et al 1992; Ismail et al 2004), which has also been depicted below (algorithm – 1):

Algorithm – 1: Backpropagation algorithm

Initialization: Initial weights w_i set to small random values;

Learning rate, $\eta = 0.1$

Repeat

For each training data set (x,y)

Calculate the outputs using the sigmoid function,

$O_j = \sigma(S_j) = \frac{1}{1+e^{-S_j}}$, where $S_j = \sum_{i=0}^{d} w_{ij} O_j$

$O_k = \sigma(S_k) = \frac{1}{1+e^{-S_k}}$, where $S_k = \sum_{i=0}^{d} w_{ik} O_k$

Compute the benefit β_k at the nodes 'k' in the output layer:

$\beta_k = O_k(1 - O_k)\,[Y_k - O_k]$

Compute the changes for weights j → k on connections to nodes in the output layer:

$\Delta w_{jk} = \eta \beta_k O_j$

$\Delta w_{ok} = \eta \beta_k$

Compute the benefit β_j for the hidden layer 'j' with the formula:

$\beta_j = O_j(1 - O_j)\,[\sum_k \beta_k w_{jk}]$

Compute the changes for the weights I → j on connections to nodes in the hidden layer:

$\Delta w_{ij} = \eta \beta_j O_i$

$\Delta w_{ok} = \eta \beta_j$

Update the weights by the computed changes: $w = w + \Delta w$,

Until termination condition is satisfied

Based on the assumption that the original spectral is additive with noise. To compute the approximate shape of the wavelet (i.e., Any real valued function of time possessing some structure), in a noisy signal and also to estimate its time of occurrence, two methods are available, first one is a simple structural analysis and the second one is the template matching technique. For the detection of wavelets in noisy signal, assume a class of wavelets, $S_i(t)$, I = 0,2,…..N-1, all having some common structure. Based on this assumption, consider a speech signal s(n) has to be corrupted by stationary additive noise d(n), to produce a noisy speech signal x(n) and has been modeled by the equation

$$x(n) = s(n) + G\,d(n) \tag{4}$$

where s(n) is the clean speech signal, d(n) is the noise and G is the term for signal-to-noise ratio control. Next windowing the signal and assuming G = 1, equation (4) becomes:

$$x_w(n) = s_w(n) + d_w(n) \tag{5}$$

Fourier transform of both sides of equation (5), yields:

$$X_w(e^{j\omega}) = S_w(e^{j\omega}) + D_w(e^{j\omega}) \tag{6}$$

Where $X_w(e^{j\omega})$, $S_w(e^{j\omega})$ and $D_w(e^{j\omega})$ are the Fourier transforms of windowed noisy, speech and noise signals respectively. To simply further, the notation the 'w' subscript has been dropped and multiplying both sides by their complex conjugates, it yields:

$$\left|X(e^{j\omega})\right|^2 = \left|S(e^{j\omega})\right|^2 + \left|D(e^{j\omega})\right|^2 + 2\left|S(e^{j\omega})\right|\left|D(e^{j\omega})\right|\cos(\Delta\theta) \tag{7}$$

where $\Delta\theta$ is the phase difference between speech and noise. Thus $\Delta\theta = \Delta\theta_S - \Delta\theta_D$.
The histograms for $\Delta\theta_S$, $\Delta\theta_D$, $\Delta\theta$ and cosine ($\Delta\theta$) has been shown in figure 5, figure 6, figure 7 and figure 8 respectively. For further analysis some assumptions have to be made: noise and speech magnitude spectrum values are independent of each other and also the phase of noise and speech are independent of each other. Taking the expected value of both sides of equation (7), and substituting E{cos($\Delta\theta$)} = 0, a power spectrum of the speech has been obtained which has been given in equation (8):

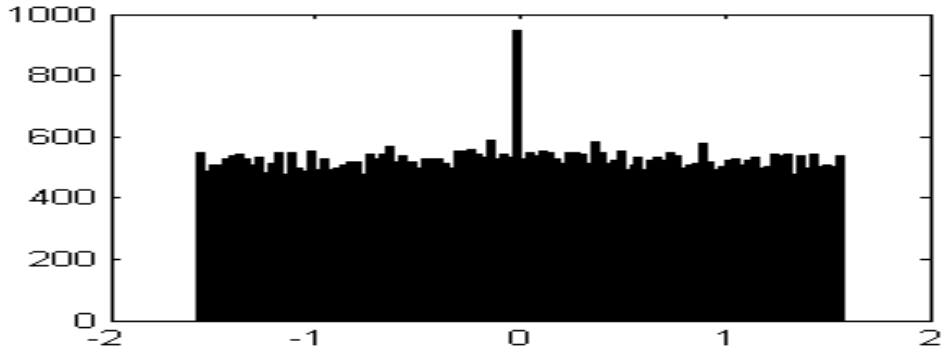


Fig. 5. Histogram of $\Delta\theta_S$ (speech magnitude spectrogram)

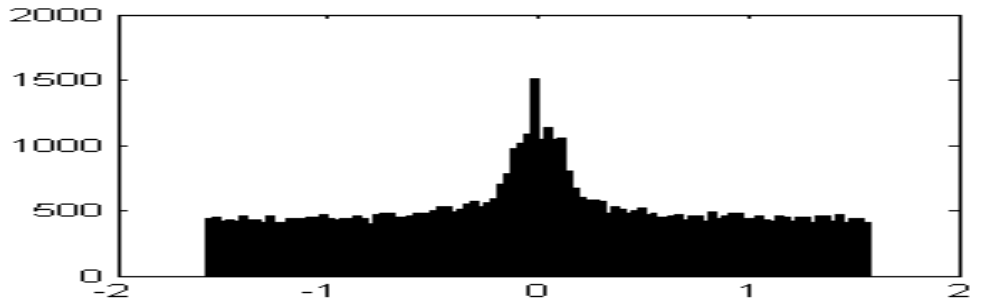


Fig. 6. Histogram of $\Delta\theta_D$ (noise magnitude spectrogram)

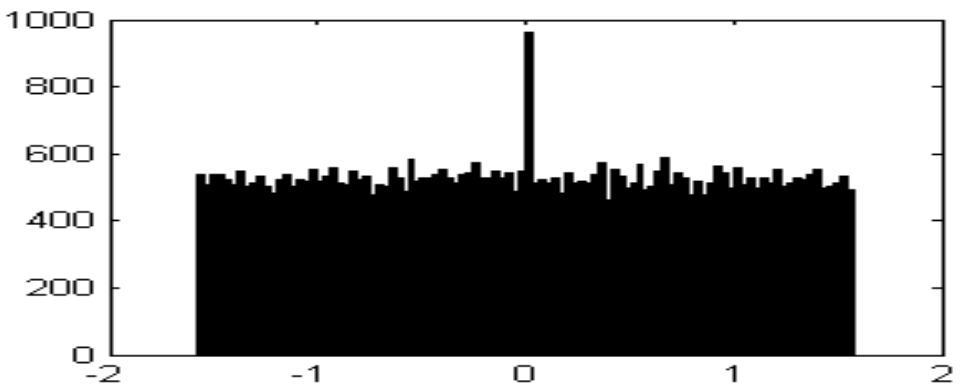


Fig. 7. Histogram of $\Delta\theta = \Delta\theta_S - \Delta\theta_D$

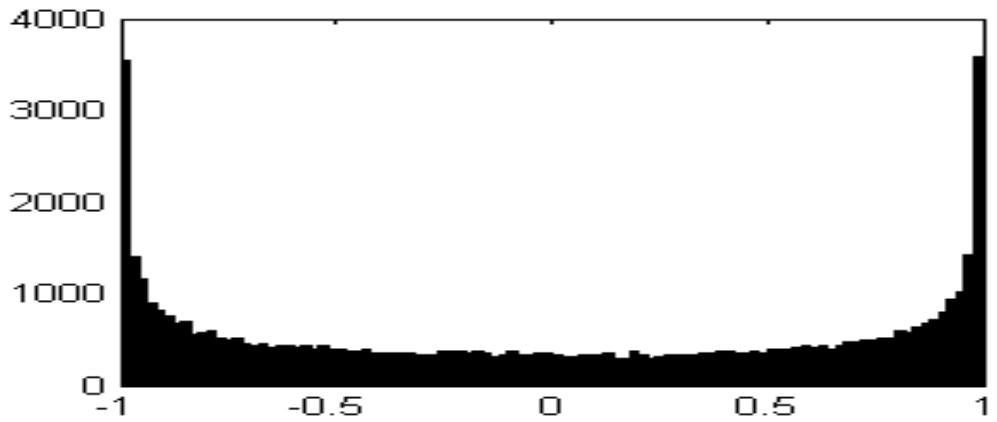


Fig. 8. Histogram of cosine ($\Delta\theta$) = cosine ($\Delta\theta_S$ - $\Delta\theta_D$)

$$\left|S(e^{j\omega})\right|^2 = \left|X(e^{j\omega})\right|^2 - E\left\{\left|D(e^{j\omega})\right|^2\right\} \tag{8}$$

Similarly, the magnitude spectrum of the speech, has to be estimated by substituting E {cos ($\Delta\theta$)} = 1, in equation (7) and hence computing the expected value of it. Thus it yields:

$$\left|S(e^{j\omega})\right| = \left|X(e^{j\omega})\right| - E\left\{\left|D(e^{j\omega})\right|\right\} \tag{9}$$

The histogram has been plotted for the equations (8) and (9) and has been depicted in figure 9. It has been observed from figure 9, that there remains some noise in the spectrum, as shown in narrow bands. This occurs due to the presence of negative values in the spectrum that are to be removed, hence these negative values has to be set to zero (Berouti et al. (1979)). To reduce the level of residual noise further subtraction has to be done. Thus equation (9) has been further simplified and it yields:

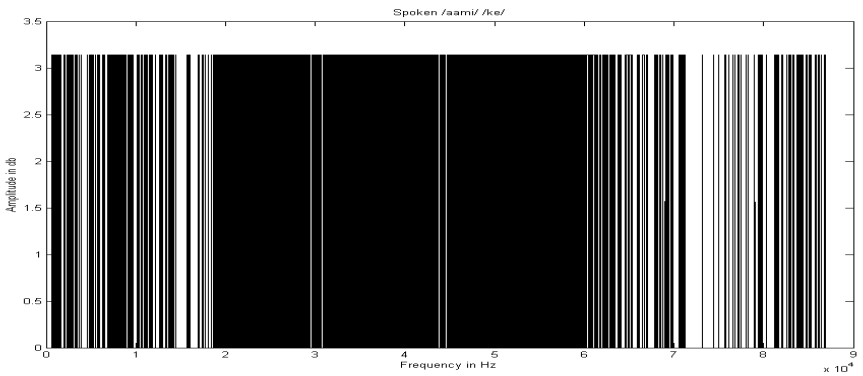


Fig. 9. Fluctuations of noise spectrogram for a speech uttered through microphone

$$\left|R(e^{j\omega})\right| = \left|X(e^{j\omega})\right| - \alpha E\left\{\left|D(e^{j\omega})\right|\right\} \tag{10}$$

where 'α' in the over-subtraction factor whose value should be more than unity. The residual noise has to be practically categorized and has been shown in figure 10.

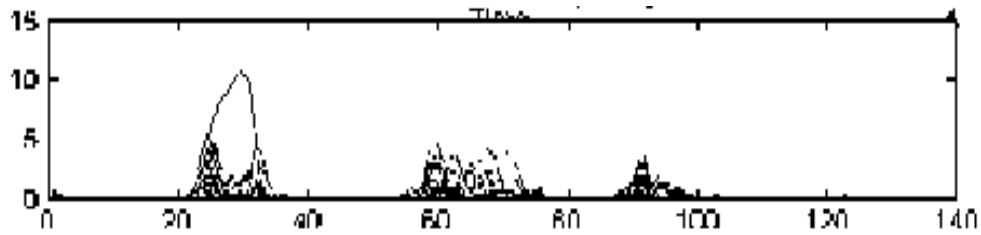


Fig. 10. Noise spectrum without negative values

To de-noise further this noisy speech signal, discrete wavelet transform has to be applied. Let the mother wavelet or basic wavelet be ψ(t), which yields:

$$\psi(t) = \exp(j2\pi ft - t^2/2) \tag{11}$$

Further as per the definition of Continuous wavelet transform CWT (a,τ), the relation yields to:

$$\text{CWT}(a, \tau) = (1/\sqrt{a}) \int x(t)\, \psi\{(t-\tau)/a\}\, dt \tag{12}$$

The parameters obtained in equation (12) has to be discretized, using discrete parameter wavelet transform, DPWT (m, n), by substituting $a = a_0^m$, $\tau = n\,\tau_0\, a_0^m$. Thus equation (12) in discrete form results to equation (13):

$$\text{DPWT}(m, n) = 2^{-m/2} \sum_k x(k)\, \psi(2^{-m}k - n) \tag{13}$$

where 'm' and 'n' are the integers, a_0 and τ_0 are the sampling intervals for 'a' and 'τ', x(k) is the speech signal. The wavelet coefficient has been computed from equation (13) by substituting $a_0 = 2$ and $\tau_0 = 1$.

Once the residual noise has been lowered, further analysis has to be done for the segmentation stage of speech processing. In classification problems with two or more classes, it is often required to choose a subset of 'd' speech features out of the given 'n' speech features (d < n). To do this, a measure of class separability has to be done. In the present work, scatter matrices have been used to form a separability criterion. A criterion for separability can be any criterion which is proportional to the between scatter matrix and also proportional to the inverse of the within scatter matrix. Maximization of such a criterion will ensure that while maximizing the distance between classes, there is no sufficient amplification in the scatter of the classes, thus causing no improvement to the separability. This scatter matrix has to be computed by finding the covariance matrix of the speech features in a given class. This has to be computed on the application of Fishers linear discriminant analysis (FLDA) which is a transformation that reduces the dimensionality of the feature vector from 'n' into d = M – 1 (where M is the number of classes involved), while

optimally preserving the separability between classes. The idea behind the Fisher's linear discriminant is the projection of n dimensional feature vectors onto a lower dimensional surface. The surface has to be chosen in such a way that the separation between classes must be kept at a minimum distance of the regression line. In order to find the optimal surface to project onto, a measure of separability must be done.
Consider a two-class problem with 'N' known samples, 'X_i'. 'N_1' of which belong to class 'w_1' and 'N_2' of which belong to class 'w_2'. Consider 'Y_i' be a linear combination of the features 'X_i' :

$$Y_i = \rho^T X_i \tag{14}$$

The 'n' dimensional vector, ρ, can be considered a line in the n dimensional space, then Y_i is the projection of X_i on this line (scaled by $\|\rho\|$). Let 'μ_I' be the mean of the 'N_i' samples of class 'w_i' in the 'n' dimensional space:

$$\mu_i = \frac{1}{N_i}\sum X_i \tag{15}$$

and the mean of the projected points Y_i on the line ρ, $\overline{\mu_i}$, is the projection of μ_i :

$$\overline{\mu_i} = \frac{1}{N_i}\sum Y_i = \frac{1}{N_i}\sum \rho^T X_i = \rho^T\mu_i \tag{16}$$

The separation of the means on ρ is given by,

$$\left|\overline{\mu_1} - \overline{\mu_2}\right| = \left|\rho^T(\mu_1 - \mu_2)\right| \tag{17}$$

Since the separation of the two classes must include the variance of three samples. Defining the n x n scatter matrix,

$$w_i = \sum_{\beta \in w_i} (X - \mu_i)(X - \overline{\mu_i})^T \tag{18}$$

where w_i is the estimation of the covariance of the i^{th} class in the n-dimensional feature space. It represents a measure of the dispersion of the signals belonging to w_i. Thus the total matrix is defined as:

$$W = w_1 + w_2 \tag{19}$$

Consider now the variance between the means of the various classes. Denote the matrix, B, represents the dispersion between the means of the various classes.

$$B = (\mu_1 - \mu_2)\left(\overline{\mu_1} - \overline{\mu_2}\right)^T \tag{20}$$

The auto-covariance of the speech signal has also to be computed using the relation:

$$C_{xx}(g) = E\{[x(nT) - x'(nT)][x(nT) - x'(nT)]\} \tag{21}$$

Then the power spectrum density has to be calculated from equation (21) and it yields to,

$$P_E(f) = \sum_{m=0}^{N-1} C_{xx}(m) W \exp(-j2\pi fm) \tag{22}$$

where $C_{xx}(m)$ is the auto-covariance function with 'm' sample. The data compression has to be performed using discrete cosine transform as shown below,

$$X_c(k)=\mathrm{Re}[X(k)]= \sum_{n=0}^{N-1} x(n)\cos\left(\frac{k2\pi n}{N}\right), k=0,1,..,N-1 \tag{23}$$

Further reduction in the dimensionality of the feature vector has to be carried out using principal component analysis. For this first discrete fourier transformation (DFT) method has to be employed, it yields to the equation of the form:

$$X(k)=F_D[x(nT)]= \sum_{n=0}^{N-1} x(nT)\exp\ (-jk(2\pi / N)n) \tag{24}$$

where k = 0, 1, 2, ……..,N-1. If $W_N = \exp(-j2\pi/N)$, then equation (24) becomes,

$$X(k) = F_D[x(nT)] = \sum_{n=0}^{N-1} x(nT) W_N^{kn} \tag{25}$$

Further for the computation of principal components (i.e., eigen values and the corresponding eigen vectors), a pattern vector $\overline{p_n}$, which can be represented by another vector $\overline{q_n}$ of lower dimension, has to be formulated using (25) by linear transformation. Thus,

$$\overline{p_n} = [M]\ \overline{q_n} \tag{26}$$

where $[M] = [X(k)]$ for k = 0 to N-1 and $\overline{q_n}$ = min([M]), such that $\overline{q_n} > 0$

Taking the covariance of equation (26), it yields, the corresponding eigen vector,

$$\overline{P} = \mathrm{cov}(\overline{p_n}) \tag{27}$$

And thus,

$$\overline{P}\ .\ M_i = \lambda_i\ .\ M_i \tag{28}$$

where 'λ_i' are the corresponding eigen values.

One of the fundamental problems that arise when computing two patterns is that of time scaling. Most of the researchers assume that both pattern and template (reference) to be compared share the same time base. This is not always correct, especially in speech analysis. It has been found that when a speaker utters the same word several times, generally each utterance he / she does so with different time bases. Each word is spoken such that parts of it are uttered faster, and parts are uttered slower. The human brain, it seems, can easily overcome these differences and recognize the word. But machine finds this a severe difficulty for the recognition. Due to this there may be a chance of occurring out-of-vocabulary (OOV) word problem. In order to sort out such problem, dynamic time warping

(DTW) method has to be adopted. Mathematically, this has been analyzed in the subsequent paragraphs.

Assume two speech signals, say $x(t_i)$ and $x(t_j)$ are defined, each with its own time base, t_i and t_j. Also assume that the beginning and end of the speech signal are known, denoted as (t_{is}, t_{if}) and (t_{js}, t_{jf}) respectively. If both the signals are sampled at the same rate, then both signals begin t sample i = j = 1, that occurs without any loss of generality. Thus, the mapping function, i = j . (i / j), is linearly related. Since speech signals are non-linear, so non-linear time warping functions must be calculated, with several assumptions. Let the warping function, w(k), be defined as a sequence of points: c(1), c(2), …..,c(k), where c(k) = (i(k), j(k)) is the matching of the point i(k) on the first time-base and the point j(k) on the second time-base. This has been summarized in figure 11, below. From figure 11, the warping, w(k), only allows to compare the appropriate parts of $x(t_i)$ with that of $x(t_j)$. Setting the monotonic and continuity conditions on the warping function, it restricts to the relations between two consecutive warping points, c(k) and c(k-1).

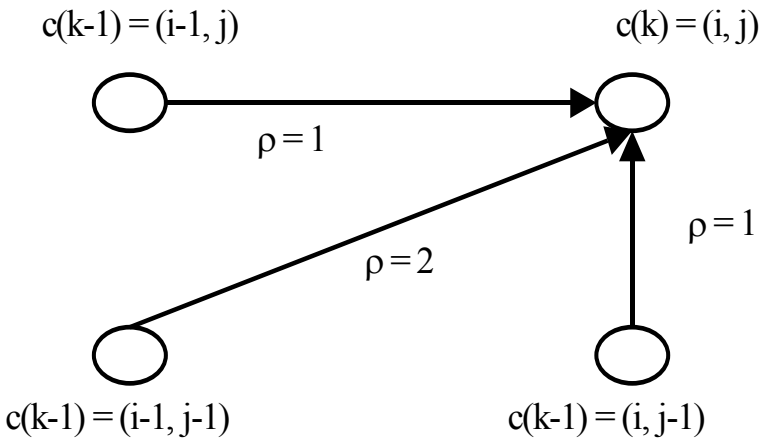


Fig. 11. Constraints on the dynamic time warping (DTW)

Thus from figure 11, there are three ways to get to the point c(i,j), which is given below,

$$c(k) = c(i,j) \tag{29}$$

and

$$c(k\text{-}1) = \begin{cases} (i(k), j(k)-1) \\ (i(k)-1, j(k)-1) \\ (i(k)-1, j(k)) \end{cases} \tag{30}$$

Thus the boundary condition is,

$$c(k) = (i,j) \tag{31}$$

By the boundary condition, matching of the beginning and end of the speech signal has to be done using dynamic programming method. As the measures have not been stored then it is difficult to trace the optimal route in an artificial word model (AWM) for the best match of the word. To overcome such problem, forward-backward dynamic programming (FBDP) method has to be adopted.

Next for the formation of AWM, the extracted speech feature values have to be fed as input to the neural network. Let the input sequences are $\{x_1, x_2, \ldots, x_n\}$, which takes real values within the range (-n, n). The weights $w_1, w_2, \ldots, w_n$, correspond to the synaptic strengths of the neuron. They serve to increase or decrease the effects of the corresponding 'x_i' input values. The sum of the products $x_i * w_i$, i = 1 to n, serve as the total combined input to the node.

So to perform the computation of the weights, assume the training input vector be 'G_i' and the testing vector be 'H_i' for i = 1 to n. The weights of the network have to be re-calculated iteratively comparing both the training and testing data sets so that the error is minimized.

The weight matrix $W = \{w_{ij}\}$ has to be computed through the relation,

$$w_{ij} = \sum_{r=1}^{n} G^T(r)H(r) \tag{32}$$

To compute the net input to the output units, the delta rule for pattern association is employed, which is given by the relation,

$$y_{\text{-inj}} = \sum_{i,j=1}^{n} x_i w_{ij} \tag{33}$$

where '$y_{\text{-inj}}$' is the output pattern for the input pattern 'x_i' and j = 1 to n.

Thus the weight matrix for the auto-associative memory neural network has to be calculated from equation (33) and the responses have to be checked by the trained input patterns. The output vector 'y' gives the pattern associated with the input vector 'x'. An activation function (usually a unipolar sigmoid) will take any input from minus infinity to infinity and squeeze it into the range –1 to +1 or between 0 to 1 intervals. A unipolar sigmoid function relates as:

$$f(y_{\text{-inj}}) = \frac{1}{1+e^{-y}} \tag{34}$$

Thus the neural network model can learn from the input / output training data pairs. Once the training has been done, it can be used as a function simulator. Similarly for framing vowel-diphthong model (VDM), the same (above discussed) concept for AWM formation, has to be followed. The constraint is that only vowel sounds are to be captured through microphone from male and female subjects. An algorithm called SCB_AWM_VDM (soft-computing based artificial word model and vowel – diphthong model) has been developed, which has been depicted in algorithm – 2 below.

Algorithm - 2: SCB_AWM_VDM

1. Record a speech through a microphone and store it in a file with extension wav.
2. Find the length of the speech signal, say N.
3. Create the row vectors 'n' and 'k' such that $0 \le n, k \le N - 1$.

4. Employ BSS and SS methods
5. Employ DFT and compute the auto-correlation coefficients. Find the minimum, maximum, mean value of the amplitude and pitch of the analyzed signal, and PSD function and probability distribution function. Compress the data further by employing DCT and also find the principal component values.
6. Find the length of the compressed speech signal. Apply AVQ technique and create a token of words
7. Count the number of frames. Pass it to a file, say atm.dat with the fields of the database as frame-1, frame-2,……,frame-N.
8. Compute the range of parameters using the relation as,
   UB = upper bound = (((mmax – mmean) / 2)*A) + mmean
   LB = lower bound = (((mmean – mmin) / 2)*A) + mmin
   where 'A' is the pre-emphasis coefficient
9. Using each of the frames extract the parameters and store in a master file as a template, thus forming a noise-free artificial word model and vowel-diphthong model.

2.2 Performance measures of developed algorithm for modelling: A case study

The number of training samples per word has to be kept sufficient for improving the accuracy of pattern matching and hence increases the performance factor. With this small condition, the developed algorithm called SCB_AWM_VDM has to be applied for the formation of a noise-free artificial word model (AWM) taking into consideration 22 speakers of varying age groups. The vocabulary has to be limited to 215 Bengali (Indian Language) words. In this work, nine male and six female adults of age group 30 – 40 years, five male adults of age group 40 – 50 years and two male adults above 50 years of age have to be selected for the testing of developed algorithm. Each phrase has to be uttered five times by each speaker leading to a total size of (22 x 215 x 5) 23650 Bengali words. Some of the speech samples have to be collected through e-mail through attachments and tested with the proposed algorithm for the formation a large Bengali vocabulary. Here, Bengali (an Indian) language has to be adopted as a case study. The developed algorithm has to be used as a tool for the formation of AWM and can also be used independently with some natural language. The worst-case time complexity is Oh (N * log(N)) and the worst-case space complexity is Oh (N*P) of the developed algorithm, where 'N' means total number of words and 'P' means the number of speech features (here P = 22). The complexities of the developed algorithm have been shown in figure 12.

From figure 12, it has been observed that accuracy increases almost exponentially. Similarly, vowel-diphthong model (VDM) has to be formed using the developed algorithm taking into consideration 10 speakers of varying age groups. Here, for VDM formation in Bengali (an Indian) language, six male and four female adults of age groups 30 – 40 years have to be selected for the testing of developed algorithm.

More discussions based on the practical implementation of the developed algorithm for the formation of a noise-free AWM and VDM can be made. Figure 13a, shows the original speech signal uttered /bondo/ /koro/ in Bengali means /close/ /it/ in English, has to be segmented using adaptive vector quantization (AVQ) method. The segmented result has been shown in figure 13b. After segmenting the words uttered, the features have to be extracted, using discrete cosine transform (DCT) and principal component analysis (PCA).

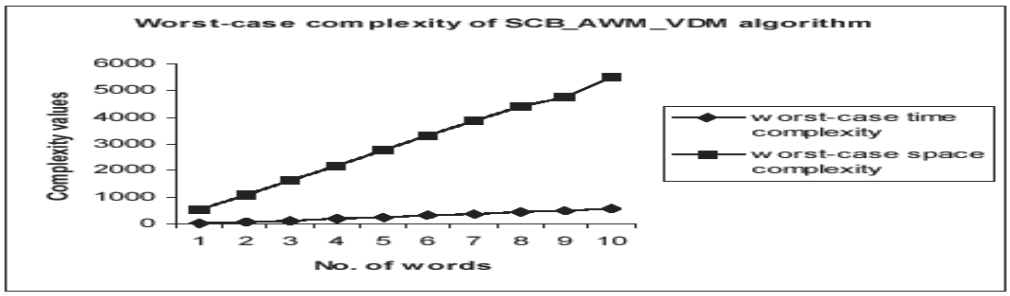


Fig. 12. Complexities of developed SCB_AWM_VDM algorithm

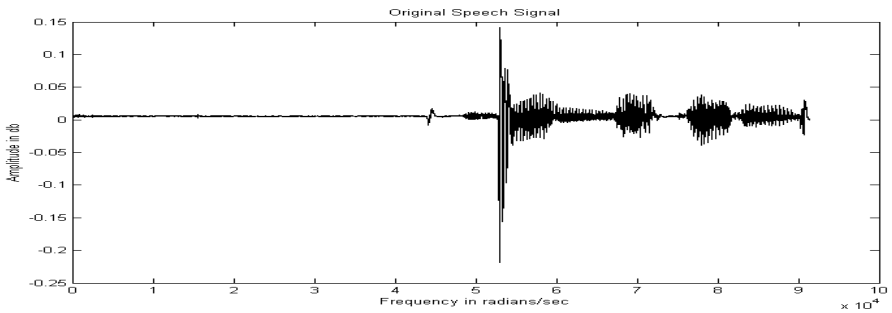


Fig. 13a. Original speech signal spoken /bondo/ /koro/

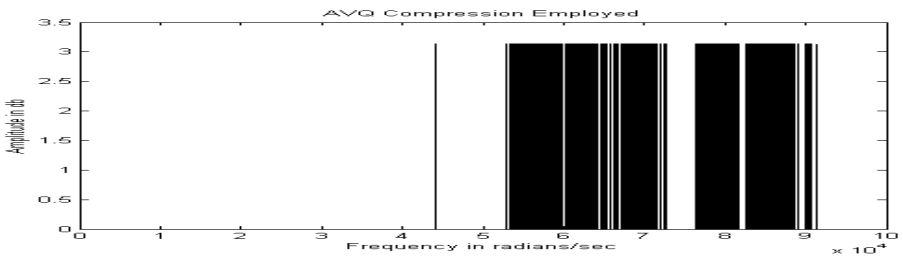


Fig. 13b. Segmentation using adaptive vector quantization (AVQ) method

The vocal tract model frequency response correspond to the position of the formants of the speech signal along frequency axis has been shown in figure 14.

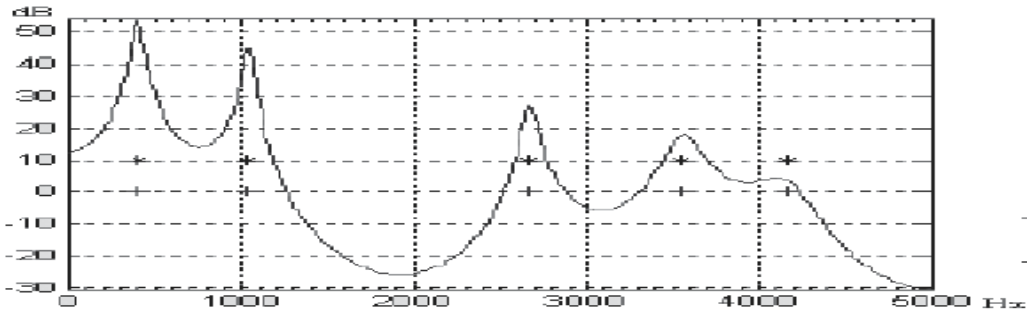


Fig. 14. Vocal tract model frequency response where + and * correspond to the position of the formants of the speech signal along frequency axis.

In figure 14, + and * shows the positions of the formants of the speech signal. Some patterns has been shown related to segmentation of the voiced pitch of the speech signal using AVQ method. Figure 15 and figure 16 shows the proper voiced pitch features using AVQ technique. It has been observed that segmentation is more accurate using AVQ technique as compared to traditional usage of zero-crossing measurement (ZCM).

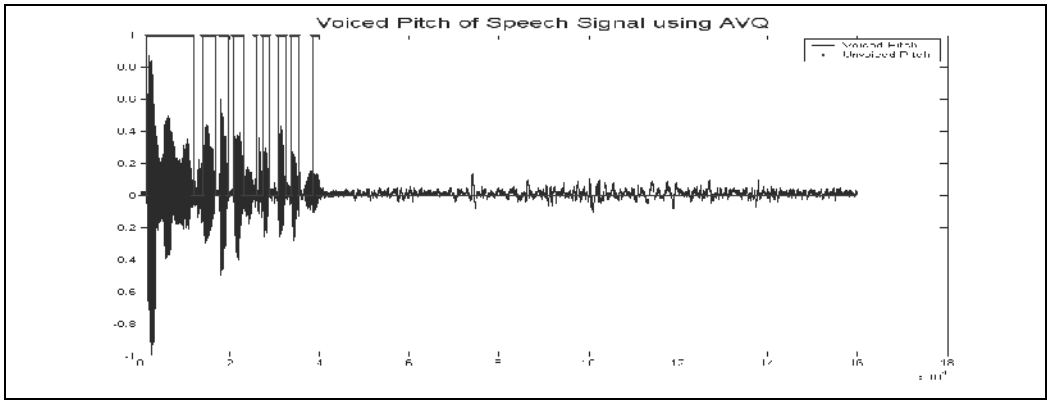


Fig. 15. Voiced pitch of the speech signal segmented using AVQ method

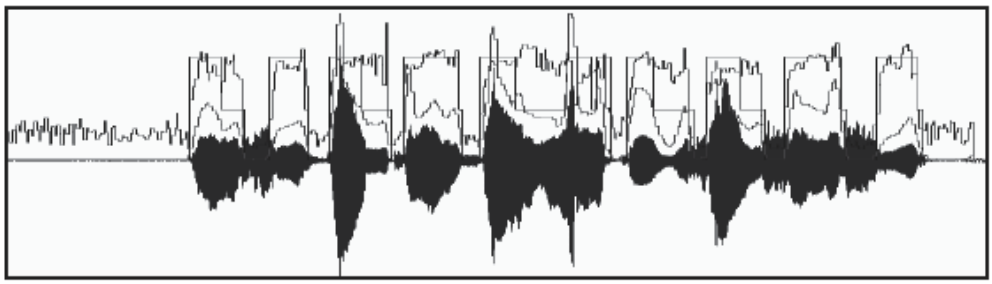


Fig. 16. Proper voiced pitch segmentation of the speech signal using AVQ method

To track the formats of the speech signal, in the present work, five formants (F_1, F_2, F_3, F_4 and F_5) have to be tracked that has been shown in figure 17.

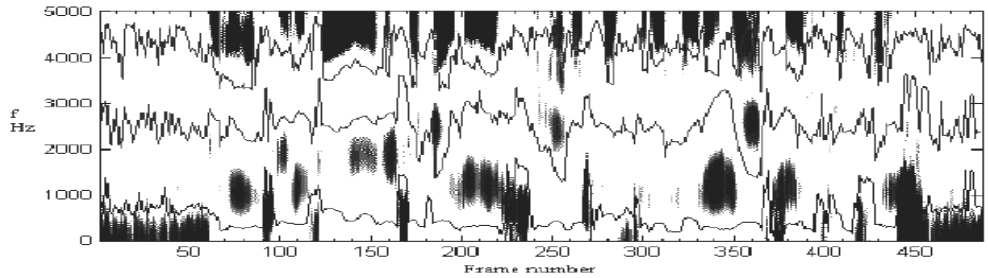


Fig. 17. Five formants have to be tracked on uttering /aakashe/ /pakhi/ /oodche/ in Bengali by a male speaker means /A/ /bird/ /is/ /flying/ /in/ /the/ /sky/ in English.

The twenty-two speech features related to pitch, energy or loudness, duration, formants and speaking rate have been depicted in table 1 and table 2.

Features	lb	μ	ub	m	σ
F_0 or pitch	-41.64	-5.71	30.23	-5.69	0.3966
Energy	-0.07	0.03	0.10	0.95	0.3962
Duration	0.00	1.16	3.19	--	--
Formants	0.0164	0.0167	0.0171	0.0128	0.4376
Speaking rate	0.00	0.91	0.83	0.88	--

Table 1. Enhanced parameters extracted from a speech signal (sample #1)

Features	lb	μ	ub	m	σ
F_0 or pitch	-41.64	-5.71	30.23	-5.69	0.3966
Energy	-0.07	0.03	0.10	0.95	0.3962
Duration	0.00	1.16	3.19	--	--
Formants	0.0164	0.0167	0.0171	0.0128	0.4376
Speaking rate	0.00	0.91	0.83	0.88	--

Table 2. Enhanced parameters extracted from a speech signal (sample #2)

3. Understanding of AWM and VDM

In this section of the chapter, hybrid approach of soft-computing techniques has to be used for the understanding of AWM and VDM for the recognition of speech and speaker. The

important speech features $K_a = \{k_1,k_2,\ldots, k_m\}$ extracted from speech signal has to be modeled as a function of the super-flame information using a neural network of soft computing techniques. This model has to be used then in conjunction with a genetic algorithm to obtain optimized super-flame information resulting in low K_a values from the model (AWM). This has to be carried out by estimating, the mapping-function between inputs and outputs of the model. Such functions are usually highly non-linear and have to be computed using adaptive vector quantization (AVQ) (Tolvi (2004); Hongwei et al (2005)) based unidirectional temporary associative memory (UTAM) of neural-network, as depicted in figure 18.

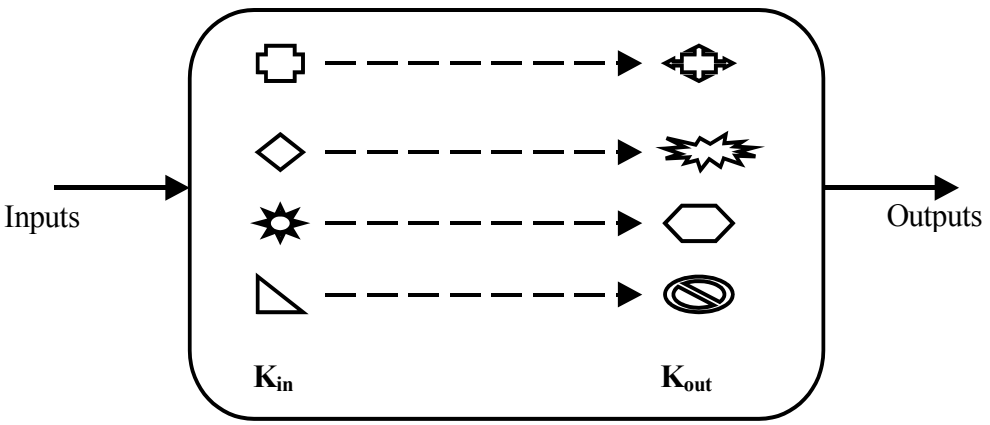


Fig. 18. Unidirectional temporary associative memory (UTAM)

In figure 18, the inputs K_{in} and the outputs K_{out} components have been shown. The term unidirectional has to be used because each K_{in} component is mapped with K_{out} component with one-to-one relationship. Each component has to be designated with a unique codeword. The set of codewords is called a codebook. The concept of UTAM has to be employed in the present work, as mapping-function for two different cases:

1. Distortion measure between unknown and known speech signals
2. Locating codeword between unknown and known speech feature

To illustrate these cases, Let $K_{in} = \{I_1,I_2,\ldots,I_n\}$ and $K_{out} = \{O_1,O_2,\ldots\ldots,O_m\}$ consisting of 'n' and 'm' input and output codeword respectively. The values of 'n' and 'm' are the maximum size of the vocabulary set. In the recognition stage, an unknown speech signal, represented by a sequence of feature vector, $U = \{U_1,U_2,\ldots,U_u\}$, has to be compared with a known speech signal stored in the form of model (AWM), represented by a sequence of feature vector, $K_{database} = \{K_1,K_2,\ldots,K_q\}$. Hence to satisfy the unidirectional associatively condition, i.e., $K_{out} = K_{in}$, an artificial word model (AWM) has to be utilized for proper matching of words. The matching of words, have to be performed on computing the distortion measure. The word with lowest distortion has to be chosen. This yield to, the relation,

$$C_{found} = \underset{1<=q<=n}{\arg\min}\left\{S(U_u, K_q)\right\} \tag{35}$$

The distortion measure has to be computed by taking the average of the Euclidean distance

$$S(U,K_i) = \frac{1}{Q}\sum_{i=1}^{Q} d(u_i, C_{\min}^{i,q}) \tag{36}$$

where $C_{\min}^{i,q}$ denotes the nearest word in the template or AWM and d(.) is the Euclidean distance. Thus, each feature vector in the sequence 'U' has to be compared with the codeword in AWM, and the minimum average distance has to be chosen as the best-match codeword. If the unknown vector is far from the other vectors, then it is very difficult to find the word from the AWM, resulting to out-of-vocabulary (OOV) word problem. Assigning weights to all the codewords in the database (called weighting method) has eliminated the OOV word problem. So instead of using a distortion measure a similarity measure that should be maximized are considered. Thus it yields,

$$S_w(U,K_i) = \frac{1}{Q}\sum_{i=1}^{Q}\frac{1}{d(u_i, C_{\min}^{i,q})} w(C_{\min}^{i,q}) \tag{37}$$

Dividing equation (36) by equation (37), it yields,

$$\gamma = \text{recognition rate} = \frac{S(U,K_i)}{S_w(U,K_i)} = \frac{unweighted}{weighted} \tag{38}$$

The procedure for computing the weights, has been depicted in an algorithm – 3 below:

Algorithm – 3: procedure to compute weight (S)

```
  for each C_I in S do
      for each C_J in C_I do
              sum = 0
              for each C_K and K != I, in S do
                 d_min = distancetonearest(C_J,C_K);
                      sum = sum + 1 / d_min;
               endfor;
          w(C_IJ) = 1 / sum;
     endfor
endfor
return weights = w(C_IJ)
```

Next for locating the codeword, hybrid approach of soft computing has to be applied in the well-defined way in the present chapter. The hybrid approach of soft computing techniques utilizes some bit of concepts from forward-backward dynamic programming and some bit of neural-networks. The work done by Fernando Bacao et al (2005), S. H. Ling et al (2007), H. Sakoe et al (1997), R. S. Chang et al (1978), and C. Y. Chang et al (1973), have been extended by considering eight constraints for searching the AWM using concepts of genetic algorithm (GA), for the best match of the uttered phrase. In general, for an optimal solution, GA is the best search algorithm based on the mechanics of natural selection, crossover and mutation. It combines survival of the fittest among string structures with a structured yet randomized information exchange. In every generation, new sets of artificial strings are created and hence tried for a new measure. It efficiently exploits historical information to speculate on new search points with expected improved performance. In other words genetic algorithms are theoretically and computationally simple and thus provide robust and optimized search methods in complex spaces. The selection operation has to be performed by selecting the speech signal as chromosomes from the population with respect to some probability

distribution based on fitness values. The crossover operation has to be performed by combining the information of the selected chromosomes (speech signal) and generates the offspring. The mutation operation has to be utilized by modifying the offspring values after selection and crossover for the optimal solution. Here in the present chapter, an acoustic model signifies the population of genes or speech parameters. Using genetic algorithm along with associative memory technique a similar type of work has been done by Tilendra Shishir Sinha et al (2006) for the recognition of speech and speaker using proposed GSAMTSS (genetic search with associative memory technique for the speech and speaker) algorithm. The methodology adopted was different in classification and recognition process and the work has been further modified by them and has been highlighted in the present part of the book using soft computing techniques of genetic and artificial neural network.
In many applications, speech signals have been either stored for later use or transmitted over some media. In both cases, interest lies in reducing the size of the signal because of cost, time and certain other benefits. As an application of the previous work done by Tilendra Shishir Sinha et al (2006), it has been discussed properly how the speech parameters can be embedded within images for promoting global cyber security through steganalysis on a standalone machine. Further some of the results reported by Yuexi Ren et al (2004), have described the semantic analysis for proper speech user interface in an intelligent tutoring system. In the work carried out earlier by Tilendra Shishir Sinha et al (2006), regarding the recognition of the gender of the speaker, the threshold values of the genders of the speakers have to be assumed which creates a void and the present work has attempted to fill in this void using mathematical analysis and employing known algorithms like polar algorithm. For the recognition of a speech and speaker, experimental setup has been summarized in figure 19.

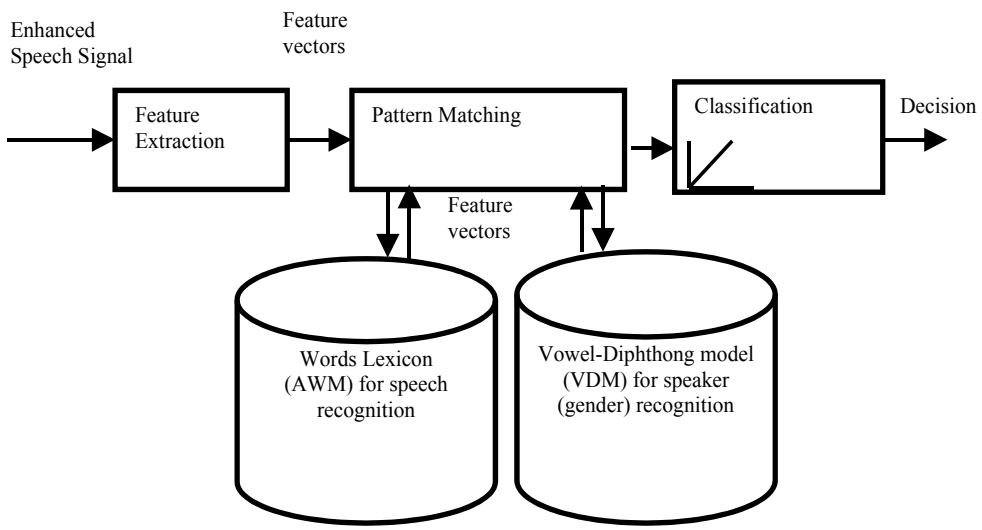


Fig. 19. General framework for recognition process

For the recognition of speech, first the speech signals (test signal) have to be made free from noise using blind signal separation (BSS) and spectral subtraction (SS) method. Hence segmented and features have to be extracted and computed the distance measures between trained data set and test data set. Hence searched for the codeword from the Bengali

vocabulary (master database or template) depending upon the minimum distance measures. If test and trained pattern matches then the decision has to be taken with: *acceptance* or *rejection*.
Generally, gender or speaker recognition is being carried out with prior knowledge of vowel uttered by the speakers (Doddington 2001). Using two vowels {o, e} the present work has to be carried out, because most of the Bengali sentences end either with 'o' or with 'e'. Extracting formant based speech features from the word model has carried this part of recognition for the gender. Depending upon the best matches decision has to be taken whether the speaker is 'male' or 'female'. The analysis has to be done using a two-class (male and female) problem. The above theoretical study has to be implemented experimentally and mathematical analysis related to the work has been also discussed.
For the best match of the words spoken in Bengali language, the developed algorithm called RCGSTSS (real-coded genetic search technique for the speech and speaker recognition) has been employed (Tilendra Shishir Sinha and Gautam Sanyal 2009).

Algorithm – 4: RCGSTSS algorithm

1. Capture a test speech signal through microphone and store it in a file with extension wav. Find the length of the test speech signal, say, M. Create the row vectors 'n' and 'k' such that $0 \le n, k \le M-1$
2. Assign spch_wer = 0, spch_acc = 0
3. Randomly generate an initial population $X(o) = (X_1, X_2, \ldots, X_M)$
4. Read the size of the trained model, say, Q, and set the counter, say, q = 0
5. Do while ($q \le Q$)
   Compute the fitness $f(X_i)$ of each individual X_i of the current population.
   Generate an intermediate population $X_r(t)$ applying the reproduction operator.
   Generate X(t+1) applying other operators to $X_r(t)$.
   Increment t = t + 1
   Increment the counter, q = q + 1
   If matched then
   spch_text = popup_from_UTAM(spch_code_word)
   spch_text = concatenate(spch_text)+concatenate(space(2))
   else
   spch_wer = spch_wer + 1
   End if
   If q = Q then
   spch_vowel = popup_from_UTAM(spch_code_word)
   spch_vowel = max(spch_vowel)
   if spch_vowel >= threshold then
   speaker = 'Female'
   else
   speaker = 'Male'
   End if
   else
   Display 'Not matched'
   End if
   End do
1. Display 'Spoken words in: ' + spch_text + 'Word error rate: '+spch_wer
2. Display 'Gender of Speaker: '+speaker + 'Recognition accuracy:'+ spch_acc

4. Conclusion

In order to recognize certain Bengali phrases spoken in speaker independent environment a noise-free artificial word model and vowel diphthong model has to be developed using blind signal separation method, spectral subtraction method, and adaptive vector quantization. It has been also observed that the residual noise occurs because of the channel coupling, which has to be minimized using spectral subtraction method by further adjusting the 'over subtraction' coefficient 'α'. Due to channel coupling most of the valuable information in speech signals goes un-observed. This problem has to be also rectified using blind signal separation method and filtered further using loss-less compression technique called discrete cosine transform and also employed principal component analysis for further analysis. The developed algorithm called SCB_AWM_VDM (soft computing based artificial word model and vowel-diphthong model) has to be implemented to train the system considering 22 speakers of varying age groups for AWM formation and 10 speakers of varying age groups for VDM formation. The vocabulary has to be limited to 215 Bengali words, considering, nine male and six female adults of age group 30 – 40 years, five male adults of age group 40 – 50 years and two male adults above 50 years of age. Each phrase has to be uttered five times by each speaker leading to a total size of (22 x 215 x 5) 23650 Bengali words for overcoming the out-of-vocabulary (OOV) word problem during recognition process of the speech and the speaker of Bengali language. Hence for simultaneous automatic speech and speaker recognition (ASSR) in combination with AWM another model called VDM has to be used. These models have to be used properly by adopting the methods forward-backward dynamic programming (FBDP) and genetic algorithm (GA). The divergence has to be calculated and also comparison has to be made with the observation that GA converges with optimal solution, thus improving the performance of the complete dialogue system. The developed algorithm (RCGSTSS) has to be successfully tested with necessary experimental data.

5. References

Yilmaz O, Rickard S., 2004, *Blind separation of speech mixtures via time-frequency masking*, IEEE Transactions on Signal Processing, 52(7): 1830-1847.

Yamashita J, Hirai Y, 2004, *Blind source separation using orientation histograms in joint mixture distributions*, In the Proceedings of Neural Network Computing and Intelligence, pp 152-157.

Araki S, Makino S, Mukai R, Hinamoto Y, Nishikawa T, Saruwatari H, 2002, *Equivalence between frequency domain blind source separation and frequency domain adaptive beam forming*, In *the* Proceedings of ICASSP, vol II, pp 1785-1788.

Murata N, Ikeda S, Ziehe A, 2001, *An approach to blind source separation based on temporal structure of speech signals*, Neuro Computing 41: 1-24.

Jang, G. J. and Lee, T.W., 2003, *A maximum likelihood approach to single channel source separation*, in the Proceedings of JMLR, vol. 4, pp. 1365-1392.

Bach, F. and Jordan, M.I., 2005, *Blind one-microphone speech separation: A spectral learning approach*, in the Proceedings of NIPS, pp. 65-72.

Pearl Mutter, B. A. and Olsson, R. K., 2006, *Algorithms differentiation of linear programs for single-channel source separation*, in the Proceedings of MLSP.

Berouti, M., Schwartz, R., and Makhoul, J., 1979, *Enhancement of speech corrupted by acoustic noise*, in *the* Proceedings of the IEEE Conference ASSP, April, pp. no. 208-211.

Yasser Ghanbari and Mohammad Reza Karami, 2004, *Spectral subtraction in the wavelet domain for speech enhancement*, International Journal of Softwares and Information Technologies, vol 1, no.1, August, pp. no. 26-29.

Shriberg, E., Stoleke, A., Hakkani-Tur and Tur, G., 2000, *Prosody-based automatic segmentation of speech into sentences and topics*, Speech Communication, vol 32 (1-2), pp 127-154.

Abdulla, W.H., 2002, *HMM-based techniques for speech segments extraction*, in the Journal of Scientific Programming, vol 10(3), pp. 221-239.

Delacourt, P., and Wellekens, C. J., 2000, *DISTBIC: A speaker based segmentation for audio data indexing*, in the Proceedings of Speech Communication, vol. 32, pp. 111-126.

Shafran I., and Rose, R., 2003, *Robust speech detection and segmentation for real-time ASR applications*, in Proceedings of International Conference on Acoustics, Speech and Signal Processing (ICASSP'03), vol. 1, pp. 432-435, Hong Kong, April.

Hue, C., Cadre, J., and Pervez, P., 2001, *A particle filter to track multiple objects*, IEEE Workshop on Multi Object Tracking.

Gustafson, F., 2002, *Particle filters for positioning, navigation and tracking*, IEEE Transaction on Signal Processing, pp. 425-437.

Arularnpalam, M., 2002, *A tutorial on particle filters for online nonlinear / non-Gaussian Bayesian tracking*, IEEE Transaction on Signal Processing, pp. 174-188.

Vermaak, J., 2002, *Particle methods for Bayesian modelling and enhancement of speech signals*, IEEE Transaction on Speech and Audio Processing, pp. 173-185.

Welling, L., and Ney, H., 1998, *Formant estimation for speech recognition*, IEEE Transaction Speech and Audio Processing, pp. 36-48.

Acero, A., 1999, *Formant analysis and synthesis using HMMs*, in the Proceedings of Eurospeech.

Watanabe, A., 2001, *Formant estimation method using inverse-filter control*, IEEE Transaction on Speech and Audio Processing, pp. 317-326.

Wang, C., and Principe, J. C., 1999, *Training Neural Networks with additive noise in the desired Signal,* IEEE Transaction on Neural Network., vol. 10, no. 6, pp. 1511 – 1517.

Cybenko, G., 1989, *Approximation by Superposition of a Sigmoid function,* Mathematics of Controls, Signals and Systems, vol. 2, pp. 303 – 314.

Hornik, K., Stinchcombe, M., and White, H., 1989, *Multilayer feed forward networks are universal approximators,* Neural Networks, vol. 2, pp. 359 – 366.

Ooyen, V., and Nienhuis, 1992, *Improving the convergence of the Backpropogation algorithm*, Neural Networks, vol. 5, pp. 465 – 471,.

Ismail, S., Abdul Manan Bin Ahmad, *2004, Recurrent Neural Network with Backpropogation through time algorithm for Arabic Recognition,* in the Proceedings of 18th European Simulation Multiconference, Europe.

Tolvi, J., 2004, *Genetic algorithms for outlier detection and variable selection in linear regression models*, Soft Computing, 8: 527-533, Springer-Verlag, London

Hongwei Sun, Kwok-Yan Lam, Siu-Leung Chung, Weiming Dong, Ming Gu, Jiaguang Sun, 2005, *Efficient vector quantization using genetic algorithm, Neural – Computing and applications*, 14 : 203 – 211, Springer – Verlag London

Fernando Bacao, Victor Lobo, Marco Painbo, 2005, *Applying genetic algorithms to zone design*, *Soft Computing*, 9 :341 – 348, Springer – Verlag London

Ling, S. H., and Leung, F.H.F., 2007, *An improved genetic algorithm with average-bound crossover and wavelet mutation operations*, Soft computing, 11 : 7-31, *Springer – Verlag London*

Sakoe, H., and Chiba, S., 1997, *Dynamic programming optimization for spoken word recognition*, in the proceedings of ICASSP, vol. 26, no. 1, pp. 43-49.

Chang, R. S and Eisentein, B. A., 1978, *Feature selection via dynamic programming for text-independent speaker identification*, IEEE Transaction Acoustic, Speech Signal Processing, vol. 26, no. 397

Chang, C. Y., 1973, *Dynamic programming as applied to feature subset selection in pattern recognition system*, IEEE transaction system man cybernetics, vol. 3, no. 166.

Sinha, Tilendra Shishir, Sanyal, Gautam and Mukherji, Abhijit, 2006, *Some aspects of modelling and simulation for the recognition of speech and speaker of Bengali language using proposed GSAMTSS Algorithm*, International Journal of Systemics, Cybernetics and Informatics, pp. no. 69 – 75.

Sinha, Tilendra Shishir and Sanyal, Gautam, 2006, *Neuro – Genetic based speech processing for promoting global cyber security using steganography technique*, *Proceedings of IEEE INDICON, September 15-17, 2006 Conference on Emerging Trends in ICT*, New Delhi, India.

Yuexi Ren, Mark Hasegawa Johnson, Stephen E. Levinson, 2004, *Semantic analysis for a speech user interface in an intelligent tutoring system*, ACM, *pp. 13 – 16.*

Doddington, G., 2001, *Speaker recognition based on idiolectal differences between speakers*, In the proceedings of EUROSPEECH, Aalborg, Denmark, pp. no. 2521-2524.

Sinha, Tilendra Shishir., and Sanyal, Gautam, 2008, *Modelling and Simulation of Speech and Speaker Recognition of a Language,* International Journal of Tomography and Statistics (IJTS), Fall 2009, Vol. 12, No. F09, pp. 19-38.

Sinha, Tilendra Shishir., and Sanyal, Gautam, 2008, *Understanding of speech and speaker model for recognition of a language,* International Journal of Artificial Intelligence (IJAI), Autum 2009, Vol. 2, No. A09, pp. 107-125.

7

Discrete Wavelet Transform & Linear Prediction Coding Based Method for Speech Recognition via Neural Network

K.Daqrouq[1], A.R. Al-Qawasmi[1], K.Y. Al Azzawi[2] and T. Abu Hilal[3]
[1]*Philadelphia University/Communications and Electronics Department*
[2]*University of Technology*
[3]*Dhofar University*
[1]*Jordan*
[2]*Iraq*
[3]*Oman*

1. Introduction

In the proposed work, the techniques of wavelet transform (WT) and neural network were introduced for speech based text-independent speaker identification and Arabic vowel recognition. The linear prediction coding coefficients (LPCC) of discrete wavelet transform (DWT) upon level 3 features extraction method was developed**. F**eature vector fed to probabilistic neural networks (PNN) for classification. The functions of features extraction and classification are performed using the wavelet transform and neural networks (DWTPNN) expert system**.** The declared results show that the proposed method can make an powerful analysis with average identification rates reached 93**.** Two published methods were investigated for comparison. The best recognition rate selection obtained was for framed DWT. Discrete wavelet transform was studied to improve the system robustness against the noise of 0dB. Our investigation of speaker-independent Arabic vowels classifier system performance is performed via several experiments depending on vowel type. The declared results show that the proposed method can make an effectual analysis with identification rates may reach 93%**.**

In general, a speaker identification system can be implemented by observing the voiced/unvoiced components or through analyzing the energy distribution of utterances. A number of digital signal processing algorithms, such as LPC technique (Adami & Barone, 2001; Tajima, Port, & Dalby, 1997), Mel frequency cepstral coefficients (MFCCs) (Mashao & Skosan, 2006; Sroka & Braida, 2005; Kanedera, Arai, Hermansky & Pavel, 1999; Daqrouq & Al-Faouri, 2010), DWT (Fonseca, Guido, Scalassara, Maciel, & Pereira, 2007) and wavelet packet transform (WPT) (Lung, 2006; Zhang & Jiao, 2004) are extensively utilized. In the beginning of 1990s, Mel frequency cepstral technique became the most widely used technique for recognition purposes due to its aptitude to represent the speech spectrum in a compacted form (Sarikaya & ansen, 2000). Actually, MFCCs simulate the model of umans' auditory perception and have been proven to be very effective in automatic speech recognition system and modeling the individual frequency components of speech signals.

ESI has been under research by a large number of researches for about four decades (Reynolds, Quatieri, & Dunn, 2000). From a commercial point of view, ESI is a technology with potentially large market due to the applications of frequently ranges from automation of operator- helped service to speech-to-text aiding system for hearing impaired individuals (Reynolds et al., 2000).

Artificial neural network performance is depending mainly on the size and quality of training samples (Visser, Otsuka, & Lee, 2003). When the number of training data is small, not representative of the possibility space, standard neural network results are poor (Kosko & Bart, 1992). Incorporation of neural fuzzy or wavelet techniques can improve performance in this case, particularly, by input matrix dimensionality decreasing (Nava & Taylor, 1996). Artificial neural networks (ANN) are known to be excellent classifiers, but their performance can be prevented by the size and quality of the training set. Fuzzy theory has been used successfully in many applications (Gowdy & Tufekci, 2000). This applications show that fuzzy theory can be used to improve neural network performance.

In this study, authors improve effective feature extraction method for text-independent system, taking in consideration that the size of ANN input is very crucial issue. This affects quality of the training set. For this reason, the presented features extraction method offers a reduction of dimensionality of features comparing with conventional methods. LPCC of DWT in conjunction is utilized. For classification of features extraction coefficients, PNN is proposed.

In this paper, an expert system for speaker identification was proposed for the investigation of the speech signals using pattern identification. The speaker identification performance of this method demonstrated on the total 59 individual speakers (39 male speakers and 20 female speakers). LPCC in conjunction with DWT upon level seven features extraction method were developed. For performing the classification process PNN was investigated. The function of feature extraction and classification is performed using the DWPN expert system. The declared results show that the proposed method can make an effectual analysis.. The average identification rates were 94.89, better than other methods published before. It was found that the recognition rates enhanced upon increasing the number of feature sets (by higher DWT levels). Nevertheless, the improvement implies a tradeoff between the recognition rate and extracting time. The proposed method can offer a significant computational advantage by reducing the dimensionality of the WT coefficients by means of LPCC. DWT approximation Sub-signal via several levels instead of original imposter had good performance on real noise facing, particularly upon level 3 and 4.

2. Discrete Wavelet Transform

The DWT indicates an arbitrary square integrable function as a superposition of a family of basis functions called wavelet functions. A family of wavelet basis functions can be produced by translating and dilating the mother wavelet related to the family (Mallat, 1989). The DWT coefficients can be generated by taking the inner product between the input signal and the wavelet functions. Since the basis functions (wavelet functions) are translated and dilated versions of each other, a simpler algorithm, known as Mallat's pyramid tree algorithm, has been proposed in (Mallat, 1989).

The DWT can be treated as the multiresolution decomposition of a sequence. It takes a length N sequence a(n) as input and produces a length *N* sequence as the output. The

output has values at the highest resolution (level 1) and $N/4$ values at the next resolution (level 2), and so on. Let $N = 2^m$, and let the number of frequencies, or resolutions, be m, we are bearing in mind $m = \log N$ octaves [18]. So that, the frequency index k varies as 1, 2,…, m corresponding to the scales $2^1, 2^2, ..., 2^m$. In
As described by Mallat pyramid algorithm (Fig.1), the DWT coefficients of the previous stage are expressed as follows (Souani et al., 2000):

$$W_L(n,k) = \sum_i W_L(i,k-1)h(i-2n), \tag{1a}$$

$$W_H(n,k) = \sum_i W_L(i,k-1)g(i-2n), \tag{1b}$$

Where $W_L(p,q)$ is pth scaling coefficient at the qth stage, $W_H(p,q)$ is the pth wavelet coefficient at the qth stage, and $h(n), g(n)$ are the dilation coefficients relating to the scaling and wavelet functions, respectively.
For computing the DWT coefficients of the discrete-time data (signal), it is assumed that the input data represents the DWT coefficients of a high resolution stage. Equations (1a) and (1b) may be used for obtaining DWT coefficients of subsequent stages. In practice, this decomposition is used only for a few stages. We note that the dilation coefficients $h(n)$ stand for a low-pass filter, where the corresponding $g(n)$ stands for a high-pass filter. In order that, DWT takes out information from the signal at different scales. The first level of wavelet decomposition extracts the details of the signal (high frequency parts), while the second and all subsequent wavelet decompositions take out progressively coarser information (lower frequency parts). Each step of retransforming the low-pass output is called dilation. A schematic of three stages DWT decomposition is shown in Fig. 1. H presents the High pass filter and
L denotes the low pass filter. At the output of each filter the result is down sampled (decimated) by taking one coefficient and leave other (Souani et al., 2000).
So as to reconstruct the original data, the DWT coefficients are up sampled (insertion of a zero between two samples) and passed through another set of low- and high-pass filters, which are expressed as

$$W_L(n,k) = \sum_p W_L(p,k+1)h'(n-2p) + \sum_l W_H(l,k+1)g'(n-2l), \tag{2}$$

where $h'(n)$ and $g'(n)$ are the low- and the high-pass synthesis filter , respectively. It is observed from Eq. (2) that the kth level DWT coefficients may be obtained from $(k+1)$th level DWT coefficients. Efficiently supported wavelets are generally used in various applications.
In the last decade, there has been a huge increase in the applications of wavelets in various scientific disciplines. Typical applications of wavelets include signal processing, image processing, security systems, numerical analysis, statistics, biomedicine, etc. Wavelet transform tenders a wide variety of useful features, on the contrary to other transforms, such as Fourier transform or cosine transform. Some of these are as follows:

- Adaptive time-frequency windows,
- Lower aliasing distortion for signal processing applications,

- Computational complexity of O(N) , where N is the length of data;
- Inherent scalability;
- Efficient Very Low Scale Integration implementation

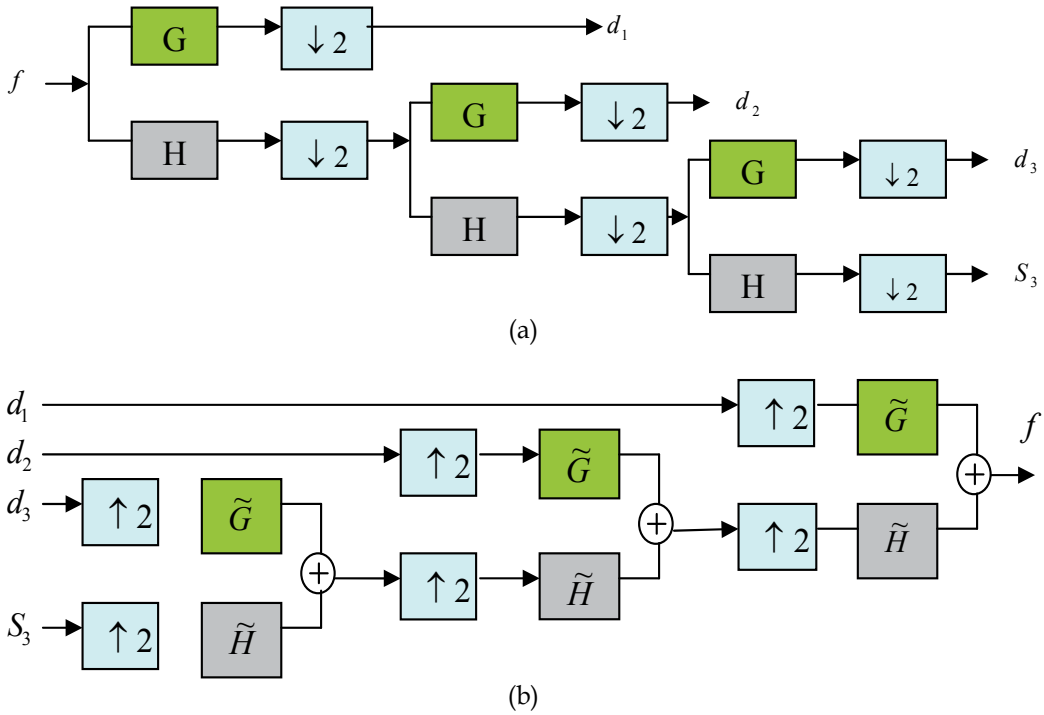


Fig. 1. a. DWT-tree by Mallat's Algorithm; b. IDWT by Mallat's Algorith

3. The use of DWT for feature extraction

Before the stage of features extraction, the speech data are processed by a silence removing algorithm followed by the application of a pre-processed by applying the normalization on speech signals to make the signals comparable regardless of differences in magnitude. In this study three feature extraction methods based on discrete wavelet transform are discussed in the following part of the paper.

3.1 DWT method with LPC

For an orthogonal wavelet function, a library of DWT bases is generated. Each of these bases offers a particular way of coding signals, preserving global energy and reconstructing exact features. The DWT is used to extract additional features to guarantee higher recognition rate. In this study, DWT is applied at the stage of feature extraction, but these data are not proper for classifier due to a great amount of data length. Thus, we have to seek for a better representation for the speaker features. Previous studies proposed that the use of LPC of DWT as features in recognition tasks is competent. (Adami & Barone, 2001; Tajima, Port, & Dalby, 1997) Suggested a method to calculate the LPC orders of wavelet transform for speaker recognition.

In this method the LPC is obtained from DWT Sub signals. The DWT at level three is generated and then 30 LPC orders are obtained for each sub signals to be combined in one feature vector. The main advantage of such sophisticated feature method is to extract different LPC impact based on multi resolution of DWT capability. LPC orders sequence will contain distinguishable information as well as wavelet transform. Fig.3 shows LPC orders calculated for DWT at depth 3 for three different utterances for the same person. We may notice that the feature vector extracted by DWT and LPC is appropriate for speaker recognition.

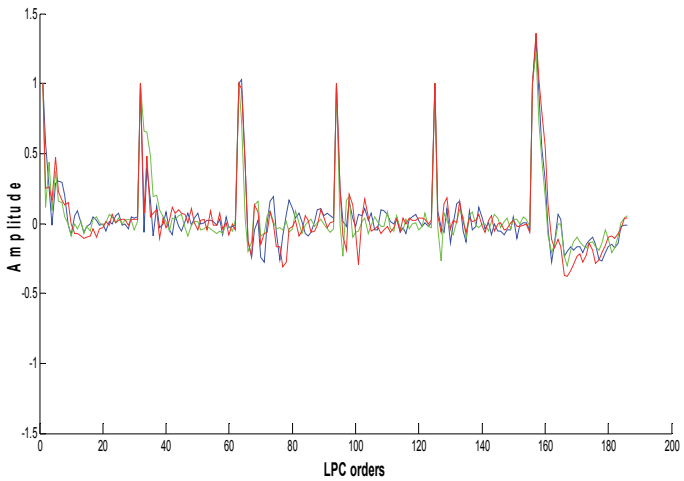


Fig. 3. LPC orders calculated for DWT at depth 3 for three different utterances for the same person

3.2 DWT method with entropy

Turkoglu et al., (2003) Suggested a method to calculate the entropy value of the wavelet norm in digital modulation recognition. [16] Proposed features extraction method for speaker recognition based on a combination of three entropy types (sure, logarithmic energy and norm). Lastly, (Daqrouq, 2011) investigated a speaker identification system using adaptive wavelet sure entropy.

As seen in above studies, the entropy of the specific sub-band signal may be employed as features for recognition tasks. This is possible because each Arabic vowel has distinct energy (see Fig.3). In this paper, the entropy obtained from the DWT will be employed for speaker recognition. The features extraction method can be explained as follows:

- Decomposing the speech signal by wavelet packet transform at level 7, with Daubechies type (db2).
- Calculating three entropy types for all 256 nodes at depth 7 for wavelet packet using the following equations:

Shannon entropy:

$$E1(s) = -\sum_i s_i^2 \log(s_i^2) \tag{3}$$

Log energy entropy:

$$E1(s) = \sum_i \log(s_i^2) \tag{4}$$

Sure entropy:

$$|s_i| \le p \Rightarrow E(s) = \sum_i \min(s_i^2, p^2) \tag{5}$$

Where s is the signal, s_i are the DWT coefficients and p is a positive threshold. Entropy is a common concept in many fields, mainly in signal processing. Classical entropy-based criterion describes information-related properties for a precise representation of a given signal. Entropy is commonly used in image processing; it posses information about the concentration of the image. On the other hand, a method for measuring the entropy appears as a supreme tool for quantifying the ordering of non-stationary signals. Fig.3 shows the three entropies calculated for DWT at depth 3 for three different utterances for the same person. We may notice that the feature vector extracted by DWT and entropy is appropriate for speaker recognition. This conclusion has been obtained by interpretation the following criterion: the feature vector extracted should possess the following properties Vary widely from class to class. 2) Stable over a long period of time. 3) Should not have correlation with other features (see Fig.4).

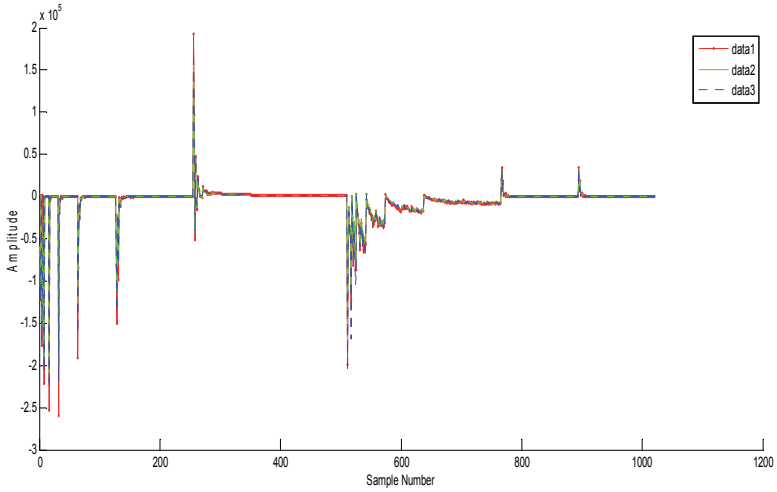


Fig. 4. Entropy calculated for DWT at depth 3 for three different utterances for the same person

4. Proposed probabilistic neural networks algorithm

We create a probabilistic neural network algorithm for classification problem (see Fig.5 and Fig.6):

$$Net = PNN(P, T, SPREAD),$$

where P is $4x2^{q+1}x24$ matrix of 24 input vowel feature vectors for net training, of 2^{q+1} (minus 2, repeated original node) WP nodes number;

$$P = \begin{bmatrix} WR_{11} & WR_{12}, ..., WR_{124} \\ WR_{21} & WR_{22}, ..., WR_{224} \\ . & . & . \\ . & . & . \\ . & . & . \\ WR_{4x2^{q+1}1} & WR_{4x2^{q+1}2}, ..., WR_{4x2^{q+1}24} \end{bmatrix}, \quad (6)$$

T is the target class vector

$$T=[1,2,3, ...,24], \quad (5)$$

and SPREAD is spread of radial basis functions. We employ a SPREAD value of 1 because that is a typical distance between the input vectors. If SPREAD is near zero the network acts as a nearest neighbor classifier. As SPREAD becomes larger the designed, network will take into account several nearby design vectors.

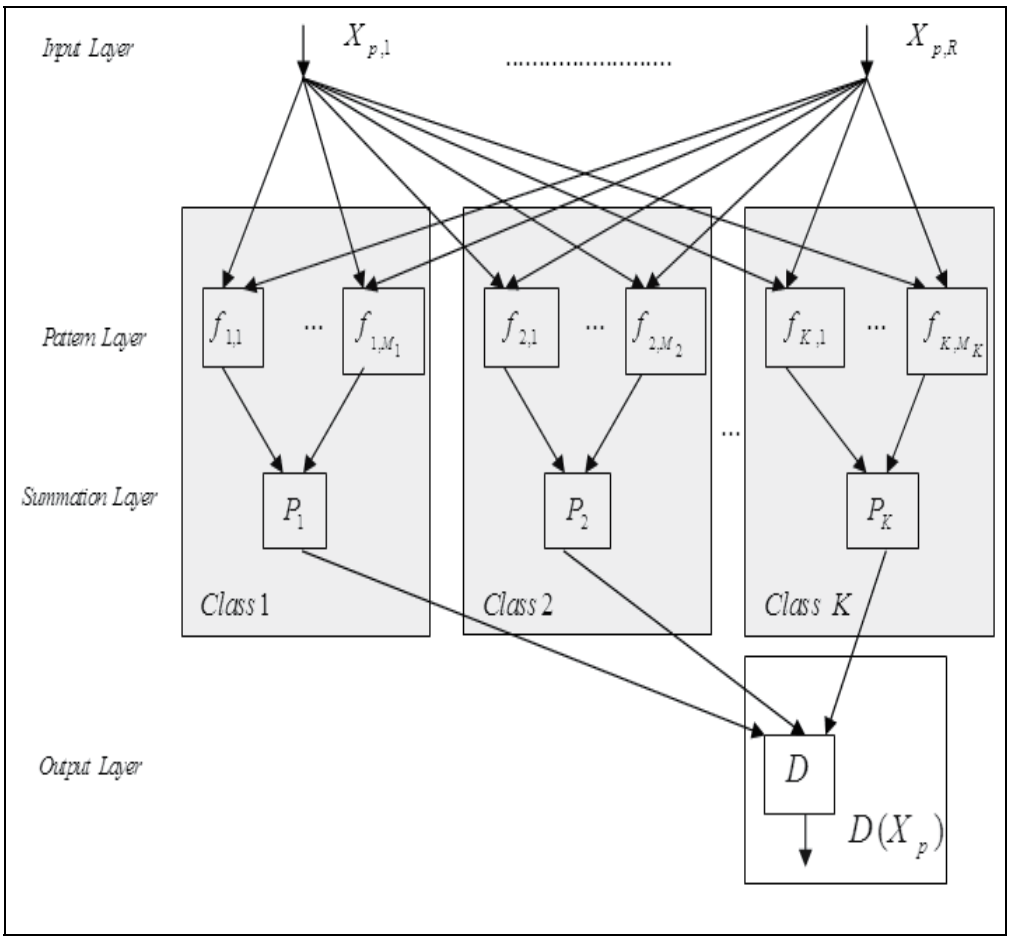


Fig. 5. Structure of the original probabilistic neural network

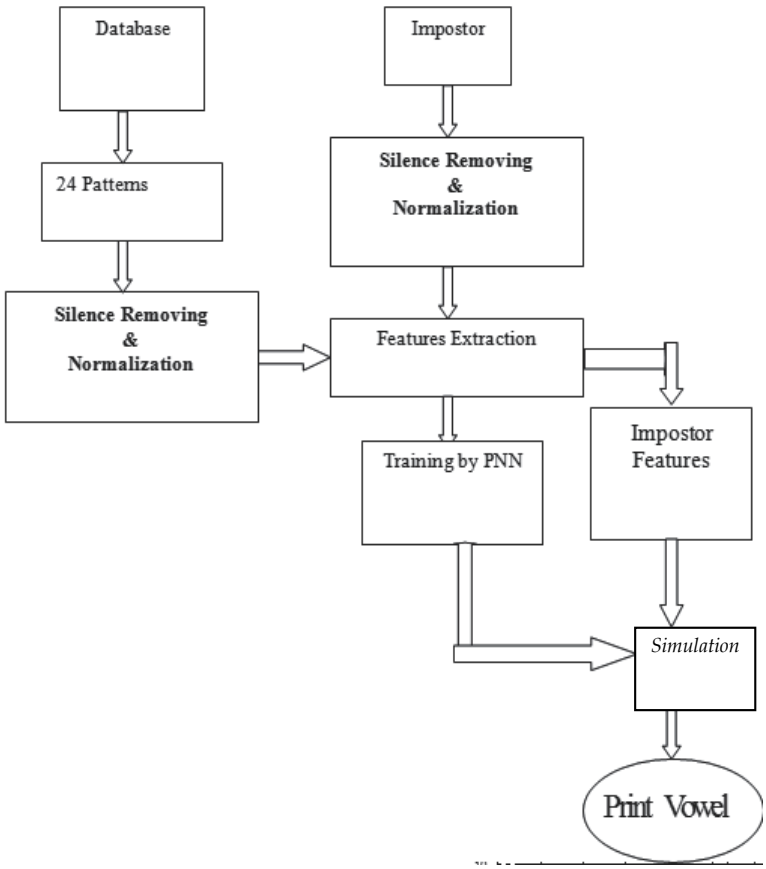


Fig. 6. Flow chart for proposed expert system

5. Results and discussion

5.1 Speaker identification by DWTLPC

A testing database was produced from Arabic language. The recording environment is a normal office environment through PC-sound card, with frequency 4 KHz and sampling frequency 16 KHz.

These utterances are Arabic spoken words. Total 47 individual speakers (19 to 40 years old) who are 31 individual male and 16 individual female spoken these Arabic words for training and testing phases. The total number of tokens considered for training and testing was 653.

It were performed experiments using total 653 the Arabic utterances of total 47 individual speakers (31 male speakers and 16 female speakers). For each of these speakers, up to 15 speech signals were used. 6 of these signals were used for training and from 4 to 9 of these signals (depends of recordings signals for each speaker) were used for testing the expert system (Fig.6). In this experiment, 93.26% correct classification was obtained by means of DWTLPC among the 47 different speaker signal classes. Testing results are tabulated in Tab.1. It, clearly, indicates the usefulness and the trustworthiness of the proposed approach for extracting features from speech signals gender identification system.

Speaker	Number of Signals	Recognized Signals	Recognition Rate [%]
Sp.1	9	9	100
Sp.2	9	8	88.88
Sp.3	9	9	100
Sp.4	9	8	88.88
Sp.5	9	9	100
Sp.6	9	6	66.66
Sp.7	9	9	100
Sp.8	9	9	100
Sp.9	9	9	100
Sp.10	9	9	100
Sp.11	9	8	88.88
Sp.12	9	6	66.66
Sp.13	9	9	100
Sp.14	9	9	100
Sp.15	9	9	100
Sp.16	9	9	100
Sp.17	8	7	87.5
Sp.18	8	8	100
Sp.19	8	7	87.5
Sp.20	4	4	100
Sp.21	4	4	100
Sp.22	4	4	100
Sp.23	4	4	100
Sp.24	4	4	100
Sp.25	8	8	100
Sp.26	8	8	100
Sp.27	8	8	100
Sp.28	8	5	62.5
Sp.29	8	7	87.5
Sp.30	8	8	100
Sp.31	8	8	100
Sp.32	8	8	100
Sp.33	8	8	100
Sp.34	8	7	87.5
Sp.35	8	7	87.5
Sp.36	8	8	100
Sp.37	8	8	100
Sp.38	8	8	100
Sp.39	8	8	100
Sp.40	8	8	100
Sp.41	8	7	87.5
Sp.42	8	7	87.5

Speaker	Number of Signals	Recognized Signals	Recognition Rate [%]
Sp.43	8	8	100
Sp.44	8	7	87.5
Sp.45	7	7	100
Sp.46	8	6	75
Sp.47	8	5	62.5
Total	371	346	93.26

Table 1. DWTLPC Identification Rate results

Table 2 shows the experimental results of different approaches used in the experimental investigation for comparison. Modified DWT with proposed feature extraction method (MDWTLPC), framing DWTLPC (FDWTLPC) illustrated in Fig.8, where LPC orders are obtained from six frames of each DWT sub signal and proposed method DWTLPC were investigated for comparison. The recognition rate of MDWTLPC reached the lowest value. The best recognition rate selection obtained was 93.53% for FDWTLPC.

Identification Method	Identification System	Number of Signals	Identification Rate [%]
DWTLPC	Text-independent	653	93.26
MDWTLPC	Text-independent	653	92.66
FDWTLPC	Text-independent	653	93.53

Table 2. Comparison of different classification approaches

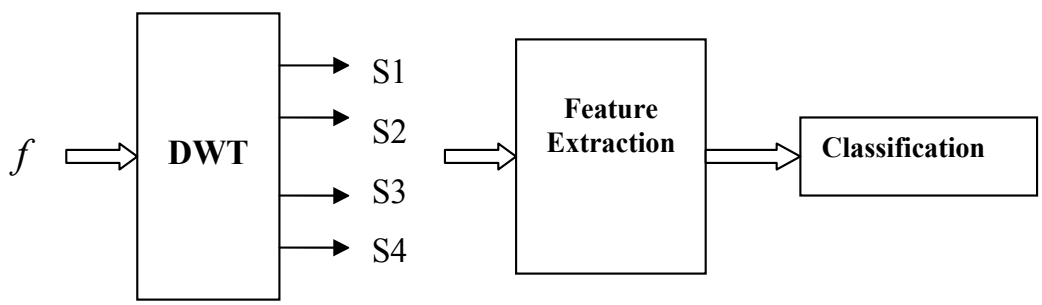
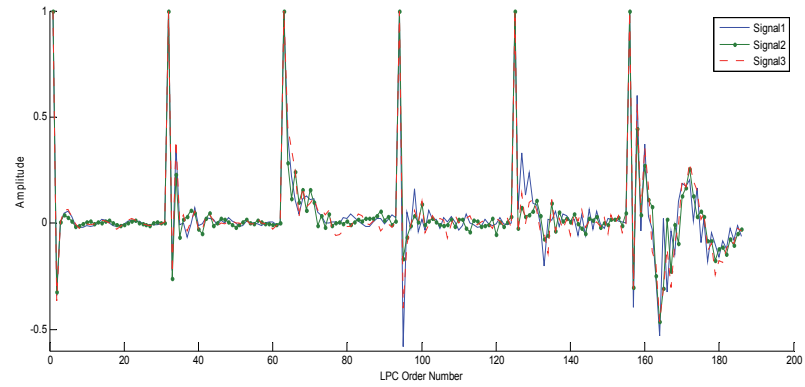
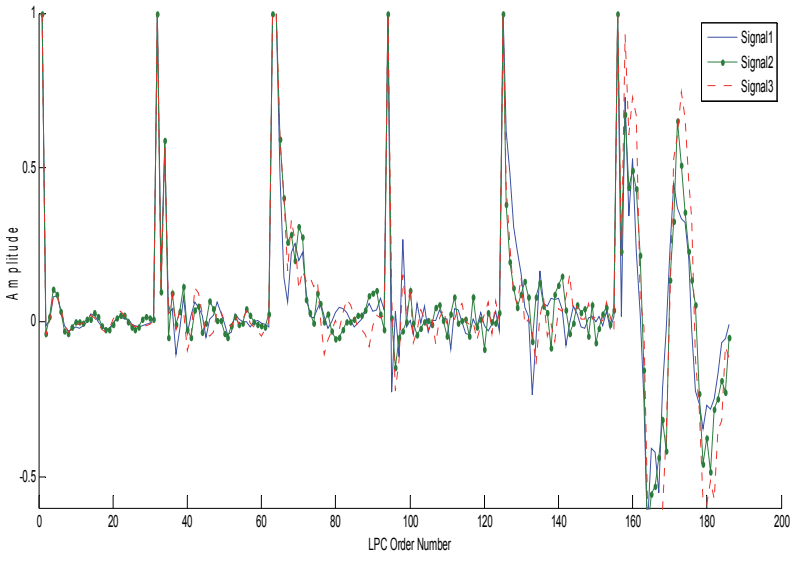


Fig. 7. Proposed system performance by using DWT approximation sub-signals (at level 1 to 4).

To improve the robustness of DWTLPC to additive white Gaussian noise (AWGN), same wavelet decomposition process was applied to DWT approximation Sub-signal via several levels instead of original imposter (Daqrouq, 2011). Afterwards, the features extraction was applied to each of the obtained wavelet decomposition sub-signals (see Fig.7). After performing proposed classification mechanism for each sub-signal of distinct DWT level, we can notice that at level 3 and 4 the highest recognition rate was achieved (see Tab.4). In this experiment it was found that the recognition rates were not improved upon increasing the DWT level more than four.



(a)



(b)

Fig. 8. Feature extraction vectors for three signals of same speaker obtained by a. FDWTLPC. b.DWTLPC

Speaker	Number of Signals	Recognized Signals [100%]			
		Level 1	Level 2	Level 3	Level 4
Sp.1	24	0	0	25	40
Sp.2	24	0	22	87	10
Sp.3	24	0	25	25	25
Sp.4	24	0	0	50	100
Sp.5	24	10	55	60	40
Sp.6	24	25	12	25	100

Table 4. DWTLPC Identification Rate results through DWT with SNR= 0dB

5.2 Arabic vowel classification by using DWTLPC

In recent times, Arabic language became one of the most significant and broadly spoken languages in the world, with an expected number of 350 millions speakers distributed all over the world and mostly covering 22 Arabic countries. Arabic is Semitic language that characterizes by the existence of particular consonants like pharyngeal, glottal and emphatic consonants. Furthermore, it presents some phonetics and morpho-syntactic particularities. The morpho-syntactic structure built, around pattern roots (CVCVCV, CVCCVC, etc.) (Zitouni and Sarikaya, 2009). The Arabic alphabet consists of 28 letters that can be expanded to a set of 90 by additional shapes, marks, and vowels. The 28 letters represent the consonants and long vowels such as ى and آ (both pronounced as/a:/), ي (pronounced as/i:/), and و (pronounced as/u:/). The short vowels and certain other phonetic information such as consonant doubling (shadda) are not represented by letters directly, but by diacritics. A diacritic is a short stroke located above or below the consonant. Table 1 shows the complete set of Arabic diacritics. We split the Arabic diacritics into three sets: short vowels, doubled case endings, and syllabification marks. Short vowels are written as symbols either above or below the letter in text with diacritics, and dropped all together in text without diacritics. We get three short vowels: fatha: it represents the /a/ sound and is an oblique dash over a letter, damma: it represents the /u/ sound and has shape of a comma over a letter and kasra: it represents the /i/ sound and is an oblique dash under a letter as reported in Table 1.

In this work, speech signals were obtained via PC-sound card, with a sampling frequency of 16000 Hz. The Arabic vowels were recorded by 27 speakers: 5 females, along with 22 males. The recording process was provided in normal university office circumstances. Our study of speaker-independent Arabic vowels classifier system performance is performed via several experiments depending on vowel type. In the following three experiments the used feature extraction method is DWTLPC.

Experimental-1

We experimented 200 long Arabic vowels آ (pronounced as/a:/) signals, 400 long Arabic vowels ي (pronounced as/e:/) signals and 90 long Arabic vowels و (pronounced as/u:/) signals. The results indicated that 96% were classified correctly for Arabic vowels آ, 90% of the signals were classified correctly for Arabic vowel ي, and 94% of the signals were classified correctly for Arabic vowel و. Tab.5 shows the results of recognition rates.

Long Vowels	Number of Signals	Accepted Signals	Not Recognized Signals	Recognition Rate [%]
Long A آ	200	192	8	96
Long E ي	400	360	40	90
Long O و	90	85	5	94
			Avr. Recognition Rate	**93.33**

Table 5. The recognition rate results for long vowels

Experimental-2

In this experiment we study the recognition rates for long vowels connected with other consonant such ل (pronounced as/l/) and ر (pronounced as/r/). Tab.6, reported the recognition rates. The results indicated 88.5% average recognition rate.

Long Vowels	Number of Signals	Recognized Signals	Not Recognized Signals	Recognition Rate [%]
La لا	60	57	3	95
Le لي	60	60	0	100
Lo لو	60	42	18	70
Ra را	60	54	6	90
Re ري	60	57	3	95
Ro رو	60	49	11	81
			Avr. Recognition Rate	**88.5**

Table 6. The recognition rate results for long vowels connected with other letters

Probabilistic neural network based speech recognition system is presented in this work. This system was performed using a wavelet feature extraction method. In this work, effective feature extraction method for Arabic vowels system is developed, taking in consideration that the computational complexity is very crucial issue. The experimental results on a subset of recorded database showed that feature extraction method proposed in this work is suitable for Arabic recognition system. Our study of speaker-independent Arabic vowels classifier system performance is performed via two experiments depending on vowel type. The declared results show that the proposed method can make an effective analysis with identification rates may reach 93%.

The proposed future work of this study is to improve the capability of proposed system to work in real time. This may be performed by modifying the recording apparatus and a data acquisition system (such as NI-6024E), and interfacing online with written Matlab code that simulates the expert system.

6. Conclusion

In this work, an expert system for speaker identification was investigated for the analyzing of the speech signals using pattern identification. The speaker identification performance of this method demonstrated on the total 47 individual speakers (31 male speakers and 16 female speakers). LPC in conjunction with framed DWT upon level three features extraction method was developed. For performing the classification process PNN was proposed. The stated results show that the proposed method can make an powerful analysis. The performance of the intelligent system was given in Table 1 and Table 2. The average identification rates were 93.26%, better than other methods. Our investigation of speaker-

independent Arabic vowels classifier system performance is performed via several experiments depending on vowel type. The declared results show that the proposed method can make an effectual analysis with identification rates may reach 93%.

7. References

Acero, (1999), Formant analysis and synthesis using hidden markov models, *in Proc. Eur. Conf.*

Adami, A. G., & Barone, D. A. C. (2001). A speaker identification system using a model of artificial neural networks for an elevator application. *Information Sciences*, 138, 1–5.

Avci, D. (2009), An expert system for speaker identification using adaptive wavelet sure entropy, *Expert Systems with Applications*, 36, 6295–6300.

Avci, E. (2007), A new optimum feature extraction and classification method for speaker recognition: GWPNN, *Expert Systems with Applications* 32, 485–498.

Bachorowski JA, Owren MJ, (1999). Acoustic correlates of talker sex and individual talker identity arepresent in a short vowel segment produced in running speech. J *Acoust Soc Am.*106(2):1054–1063.

Bennani Y., Gallinari P. (1995), Neural networks for discrimination and modelization of speakers, *Speech Communication*, 17 (1995) 159-175.

Chen, C.-T. Lung, S.-Y. . Yang, C.-F. & Lee, M.-C. (2002), Speaker recognition based on 80/20 genetic algorithm, in: IASTED *International Conference on Signal Processing, Pattern*

Cherif, A., Bouafif, Dabbabi, T. (2001), Pitch detection and formants analysis of Arabic speech processing, *Applied Acoustcs*, 62, 1129-1140.

Daqrouq K., Abu-Isbeih, I., Daoud, O. & Khalaf, E., (2010), An investigation of speech enhancement using wavelet filtering method, Int J *Speech Technol.*, 13: 101–115.

Daqrouq K., Al-Faouri, M. (2010), Spoken Arabic Digits Classifier via Sophisticated Wavelet Transform Features Extraction Method, *Information Sciences and Computer Engineering*, Vol. 1, No. 1, 16–25

Daqrouq, K. (2011)., Wavelet entropy and neural network for text-independent speaker identification. *Engineering Applications of Artificial Intelligence* (2011), doi:10.1016/j.engappai.2011.01.001.

Deng L., Bazzi I., and Acero A. (2003), Tracking vocal tract resonances using an analytical nonlinear predictor and a target guided temporal constraint, in Proc. *Eur. Conf. Speech Communication Technology.*

Deng L., Lee L., Attias H., and Acero A., (2004), "A structured speech model with continuous hidden dynamics and prediction residual training for tracking vocal tract resonances," in *IEEE ICASSP*,.

Derya, A. (2009), An expert system for speaker identification using adaptive wavelet sure entropy, *Expert Systems with Applications* 36, 6295–6300

Engin, A. (2007), A new optimum feature extraction and classification method for speaker recognition: GWPNN, *Expert Systems with Applications* 32, 485–498.

Evangelista, G. (1993). Pitch-synchronous wavelet representations of speech and music signals. *IEEE Transactions on Signal Processing*, 41(12).

Evangelista, G. (1994). Comb and multiplexed wavelet transforms and their application to speech processing. *IEEE Transactions on Signal Processing*, 42(2).

Farooq, O. & Datta, S. (2003), Phoneme recognition using wavelet based features, *Information Sciences* 150, 5–15

Fonseca, E. S., Guido, R. C., Scalassara, P. R., Maciel, C. D., & Pereira, J. C. (2007). Wavelet timefrequency analysis and least squares support vector machines for the identification of voice disorders. *Computers in Biology and Medicine*, 37, 571–578.

Ganchev T., Fakotakis N., Kokkinakis G. (2005), Comparative evaluation of various MFCC implementations on the speaker verification task, in: *Proceedings of the SPECOM*-2005, vol. 1, pp. 191–194.

Gelfer M. P. and Mikos V.A. (2007), The Relative Contributions of Speaking Fundamental Frequency and Formant Frequencies to Gender Identification Based on Isolated Vowels, *Journal of Voice*, Vol. 19, No. 4, pp. 544–554,

Gowdy, J. N., & Tufekci, Z. (2000). Mel-scaled discrete wavelet coefficients for speech recognition. In Proceedings of the acoustics speech and signal processing, ICASSP '00. *IEEE international conference*. Istanbul.

Haydar, A. , Demirekler, M. & Yurtseven, M.K. (1998), Speaker identification through use of features selected using genetic algorithm, *Electron. Lett*. 34, 39–40.

Huang X., Acero A., and Hon H.-W. (2001), Spoken Languag Processing, Prentice Hall PTR

Hyeon, J. &. Bang, S.Y., Feature selection for multi-class classification using pairwise class discriminatory measure and covering concept, *Electron. Lett*. 36 (2000) 524–525.

Kadambe, S., & Boudreaux-Bartels, G. F. (1992). Application of the wavelet transform for pitch detection of speech signals. *IEEE Transactions on Information Theory*, 32(March), 712–718.

Kadambe, S., & Srinivasan, P. (1994). Applications of adaptive wavelets for speech. *Optical Engineering*, 33(7), 2204–2211.

Kanedera, N., Arai, T., Hermansky, H., & Pavel, M. (1999). On the relative importance of various components of the modulation spectrum for automatic speech recognition. *Speech Communication*, 28, 43–55.

Kosko & Bart (1992). Neural networks and fuzzy systems: A dynamical approach to machine intelligence. *Englewood Cliffs*, NJ: Prentice Hall.

Lung, S.-Y. (2004), ,Applied multi-wavelet feature to text independent speaker identification, *IEICE Trans. Fundam*. E87-A (4) 944–945.

Lung, S-Y. (2007), Efficient text independent speaker recognition with wavelet feature selection based multilayered neural network using supervised learning algorithm, *Pattern Recognition* 40, 3616 – 3620.

Malkin J., Li X., Bilmes J., A Graphical Model for Formant Tracking, SSLI Lab, *Department. Of Electrical Engineering, University of Washington, Seattle,* This work was supported by NSF grant ITS-0326382.

Mallat, S. A. (1989). Theory for multiresolution signal decomposition: The wavelet representation, *IEEE Transactions Pattern Analysis Machine Intelligent*, 31, 674–693.

Mashao, D. J., & Skosan, M. (2006). Combining classifier decisions for robust speaker identification, *Pattern Recognition*, 39, 147–1.

Nathan, K.S. & Silverman, H.F. (1994), Time-varying feature selection and classification of unvoiced stop consonants, *IEEE Trans. Speech Audio Process*, 2 (1994) 395–405.

Nava, P., & Taylor, J. (1996). Voice recognition with a fuzzy neural network. *In Proceedings of the fifth IEEE international conference on fuzzy systems* (pp. 2049–2052).

Reynolds, D. A., Quatieri, T. F., & Dunn, R. B. (2000). Speaker verification using adapted Gaussian mixture models. *Digital Signal Processing*, 10(1-3), 19–41.

Sarikaya, R., & Hansen, J. H. L. (2000). High resolution speech feature parametrization for monophone-based ressed speech recognition. *IEEE Signal Processing Letters*, 7(7), 182–185.

Souani C, Abid M, Torki K., VLSI design of 1-D DWT architecture with parallel Filters, *the VLSI journal 29* (2000) 181}207.

Sroka, J. J., & Braida, L. D. (2005). Human and machine consonant recognition. *Speech Communication*, 45, 401–423.

Tajima, K., Port, R., & Dalby, J. (1997). Effects of temporal correction on intelligibility of foreign accented English. *Journal of Phonetics*, 25, 1–24.

Turkoglu, I., Arslan, A., & Ilkay, E. (2003). An Intelligent system for diagnosis of the heart valve diseases with wavelet packet natural Networks. *Computer in Biology and Medicine*, 33, 319–331.

Visser, E., Otsuka, M., & Lee, T. (2003). A spatio-temporal speech enhancement scheme for robust speech recognition in noisy environments. *Speech Communication*, 41, 393–407.

Wu, J.-D. & Lin B.-F. (2009), Speaker identification using discrete wavelet packet transform technique with irregular decomposition, *Expert Systems with Applications* 363136–3143.

Xia, K. and.Espy-Wilson, C. (2000), A new strategy of formant tracking based on dynamic programming, in Proc. *Int. Conf. on Spoken Language Processing*.

Zitouni I. and Sarikaya R. (2009), Arabic diacritic restoration approach based on maximum entropy models, *Computer Speech and Language*, 23 257–276

Part 3

Biosensors

8

Implementation of the Discrete Wavelet Transform Used in the Calibration of the Enzymatic Biosensors

Gustavo A. Alonso[1,2], Juan Manuel Gutiérrez[1],
Jean-Louis Marty[2] and Roberto Muñoz[1]
[1]*Centro de Investigación y de Estudios Avanzados del instituto Politécnico Nacional*
[2]*Université de Perpignan Via Domitia, IMAGES EA4218*
Centre de Phytopharmacie
[1]*France*
[2]*Mexico*

1. Introduction

The extensive use of insecticides in modern agriculture has raised serious public concern regarding the environment and food safety, and considerable efforts have been devoted to the development of highly sensitive detection methods (Bachmann et al., 2000; Istamboulie et al., 2007). Biosensors have been described for many years (Andreescu and Marty, 2006a) as good candidates to replace or complement conventional analytical methods, as they can provide real-time qualitative information about the composition of a sample with minimum treatment. In recent years, biosensors have played an important role in the determination of pesticides, because the compounds used in its composition have anticholinergic properties and therefore strongly inhibit cholinesterase enzymes. Evaluation of cholinesterase activity is the crucial factor in the construction of biosensors, however, in the case of multi-component samples (more common in reality) is absolutely essential include appropriate data processing tools to find relationships between the biosensor responses and the measured data. In most cases, it is necessary a first data pretreatment step in order to explore and validate these obtained information (Ehrentreich, 2002).

Many applications related with the use of biosensor responses entail data interpretation problem related to: (1) noisy records due to temperature changes; (2) data acquisition noise present in records, (3) presence of interference signals in the biosensor response mainly contaminated by signals coming for the electrochemical equipment i.e. potentiostats, magnetic stirrings and thermostats (Cai and Harrington, 1998; Zanchettin and Ludermir, 2007): (4) according with the inhibition method the responses can be slow which implies signals with information in low frequencies and a large number of samples per essay (Arduini et al., 2010).

In this way, different data processing strategies have been proposed in order to achieve better interpretation models and discard irrelevant content coming from original data. During the last decade, Wavelet Transform (WT) has been widely employed in signal processing analysis were denoising and compression is an important step in the data mining

process. WT has been proved to be fast and reliable than traditional methods such as Fourier Transform (FT) and Savitzky-Golay (SG) (Alsberg et al., 1997; Cai and Harrington, 1998; Gorry, 1990).

Wavelet transform is a well localized in both the time and frequency domain. Therefore it may decompose a signal retaining the information of both domains. Unlike the more traditional filtering methodologies, wavelet transforms have the ability to preserve the temporal locality of sharp transitions within time domain signals (Alsberg et al., 1997). This property makes WT useful to represent different features of the signal, especially non-linearities or discontinuities.

Since 90´s decade, WT has attracted interest in fields of chemistry for signal processing. Analytical techniques such as chromatography, infrared spectroscopy, mass spectrometry, nuclear magnetic resonance spectroscopy, ultraviolet-visible spectroscopy and electroanalysis, have also been benefited from the properties of wavelet processing for data processing tasks mainly related to compression, noise removal, base-line correction, zero crossing, regression and multi-resolution of overlapping signals (Leung et al., 1998a; Leung et al., 1998b; Nie et al., 2001).

The viability of the use of wavelet analysis in electrochemical applications has been reported since two decade ago. Bos and Hoogendam proposed the use of WT to minimized the effect of noise and base-line drift to determine peak intensities in flow-injection analysis (Bos and Hoogendam, 1992). In that work, obtained results indicate that wavelet analysis applied to peak shapes (with appropriate signal-to-noise ratio), may improve the detection limit obtained with traditional signal processing methods in which the peak heights and peak areas are determined. In this sense, Morlet wavelet function was utilized to transform the FIA signal into a two dimensional time-frequency form.

One of the first reports related to the analysis of electrochemical signals was presented by Zou and Mo (1997), they employed wavelet multi-frequency channel decomposition (WMCD) in records coming from linear scan voltametry, in order to extract useful information from voltamograms with high levels of noise. Selecting optimal wavelet basis and frequency scales the absolute values of the peak have relative errors less than 2% (Zou and Mo, 1997). Other interesting work developed by Aballe et al. (1999), used wavelet transform to analyse the electrochemical noise and its fluctuations in order to establish either the rate or the mechanism of a corrosion process (Aballe et al., 1999). This kind of analysis allowed obtaining complementary information about these parameters where other techniques do not work properly.

Since the wavelet analysis is carried out by a digital system, we assumed that the analysis is in time-discrete and depends of discrete parameters, such as rate of acquire analogical signal, numbers of operations per second and resolution of the data. These parameters are not taken into account when the wavelet transforms are performed by personal computer and specialized mathematician software, but when the implementations are oriented towards specific applications or/and the necessity to work with out a computer (digital robust system) hardware implementation emerges as an interesting alternative.

With the advancement of the microelectronics, new trends are oriented to develop wavelet transforms implemented in portable systems with high accuracy, low-cost, short-time response and easy programmed to be a suitable option in electrochemistry calibrations (Alonso et al., 2010).

In the last decade, different works of implementations of discrete wavelet transforms (DWTs) based on Digital Signal Processors (DSPs), Field Programmable Gate-Array (FPGA)

and CMOS technology were reported such as an alternative of implementation of wavelet transforms into dedicated systems. The first dedicated devices used to compute the DWT were the DSPs, which have high process-computing power, high speed and normally the manufactures proved full support. DSPs provide some special hardware units, such as Multiply-Acumulate (MAC) to improve the performance of discrete wavelet transform Nevertheless, these devices use to be expensive, not compatible with other hardware and with the possibility to use only one processing core; so tasks must be programmed sequentially.

One of the first authors to report the use of these devices were Bahoura et al. (Bahoura M. et al.,1997), they develop an algorithm based on wavelet transform suitable for real time implementation. This algorithm was implemented in a SPROC-1400 device with a 50 MHz frequency clock and was used to detect ECG arrhythmia characteristics (Bahoura et al., 1997). An application related to DWT and the JPEG2000 image compression standard was described by Gnavi et al. (2002). The authors developed wavelet engines on a DSP platform, based on the so-called lifting scheme. Using this approach, authors presented a performance comparison between traditional convolution and lifting scheme, proving that the later were faster and computationally less demanding (Gnavi et al., 2002). After this work, Jichang et al (2003), described an algorithm to be implementable in DSP TMS3320C3X using the instruction parallel multiply-accumulate with circular addressing programmed in assemble language program (Jichang et al., 2003).

All the works mentioned have been implemented successfully integer discrete wavelet transforms but do not develop algorithms that compute the floating point DWT, when floating point representation of DWT coefficients is required. In this sense, although it is possible to implement floating-point operations on DSP fixed point architectures, the main disadvantage of this approach is related to the time-consuming transfer data to and from memory (Smith, 1997).

In order to provide flexibility for the DWT implementation based on wavelet filter length and decomposition structure, implementations based on FPGA were proposed. The first recorded work, describing a real-time application of discrete wavelet transform for audio and video compression is reported by Motra et al. (2004). In this work, the reported architecture was programmed in Verilog-HDI in a FPGA and supports higher hardware utilization and the latter scheme speed up the clock rate of DWT (Motra et al., 2004). Another similar work is presented by Zhang and Hu. Here the authors proposed a DWT algorithm based on pyramidal structural data coding and it was programmed in VHDL language (Zhang and Hu, 2004). Latest work related to this topic is presented by Knowles who reported approach of simple modular and recursive hardware implementation of DWT using basic units: input delay, filter, register bank, and control unit. The implementation based on VLSI performs both high- and low- pass filter with just one set of multipliers (Knowles, 2008).

This brief overview of works illustrates the recent trends of DWT implementation, carried out by different research groups. On the one hand, DSP implementations have been widely used because of its easy way to programming and its high precision computing. Nevertheless, disadvantages such as non standardized compiler languages and high hardware costs have promoted in recent years the use of FPGA devices.On the other hand, FPGA implementations have advantages to allow preserving parallel architecture using programmable gates on a single chip, and the source code can be modified by the user with

relative simplicity. A well known disadvantage of these systems, is the need to coupled with other peripherals, becoming difficult to construct a cheap specific system (Shahbahrami et al., 2010).
Recently, the development of low-cost programmable devices has created a new family of microcontrollers with high simplicity and flexibility of programming and wide hardware compatibility. These devices called dsPIC® (from Microchip®) combines the best features of both microcontrollers and DSPs in a single core. These new features combined with high process speed promise to be the key to new implementation strategies of the DWT on a dedicated device.
In this chapter, we will attempt to describe the nature of discrete wavelet transform to build our implementation model in any low-cost dsPIC chip for denoising and compression biosensor responses. The knowledge needed to reach our main goal will be summarized along different sections to give the reader comprehensive and specialized information of the practical use of implemented discrete wavelet transform such as data pre-treatment tool.

2. Fundamentals of Wavelet Transform

In the 80's decade, Morlet et al. (Morlet et al., 1982) described the concept *wavelets* which used to decomposition signals without the necessity of windowed them as Gabor suggested before (Gabor, 1946). In wavelet treatment all basis functions $\psi_{s,t}(x)$ can be derived from a mother wavelet Ψ(x) (eq. (1) through the following translation and dilation process.

$$\psi_{s,t}(x) = \frac{1}{\sqrt{|s|}}\psi\left(\frac{x-t}{s}\right) \tag{1}$$

Where s and t, are respectively, the scale and translation parameter expressed in real numbers R. $1/\sqrt{|2|}$ is an energy normalization factor for all sub wavelets functions. The basic idea of WT is to represent any arbitrary function $f(x)$ as a superposition of wavelets. The continuous wavelet transform is given by eq. (2).

$$W_f(s,t) = \frac{1}{\sqrt{|s|}}\int_{-\infty}^{\infty} f(x)\overline{\psi}\left(\frac{x-t}{s}\right)dx = \left\langle f(x),\psi_{s,t}(x)\right\rangle \tag{3}$$

With $s > 0$ and t having arbitrary values.
In figure 1, a variety of mother wavelets are shown. The more common wavelets are; Haar wavelet, Mexican Hat, Morlet, Daubechies. The main idea of wavelet transform is the overlapping of a wavelet mother over any signal see figure 2(b).

2.1 Discrete Wavelet Transform

Croisier et al. (Croisier et al., 1976) created a technique to analyze digital signals by the decomposition of them. In the same year Crochiere et al. (Crochiere et al., 1977) reported a similar work to code audio signals. This technique was named sub-band coding. Vetterli and Le Gall, improved the analysis of sub-bands decreasing the redundant data from pyramidal algorithm (Vetterli and Le Gall, 1989).
Mallat presented a multi-resolution representation for analyzing images. The decomposition was defined as an orthogonal multi-resolution representation called wavelet representation.

It was computed by a pyramidal algorithm which was based on convolutions with quadrate mirror filters. The analysis was applied to image compression, and texture discrimination (Mallat, 1989).

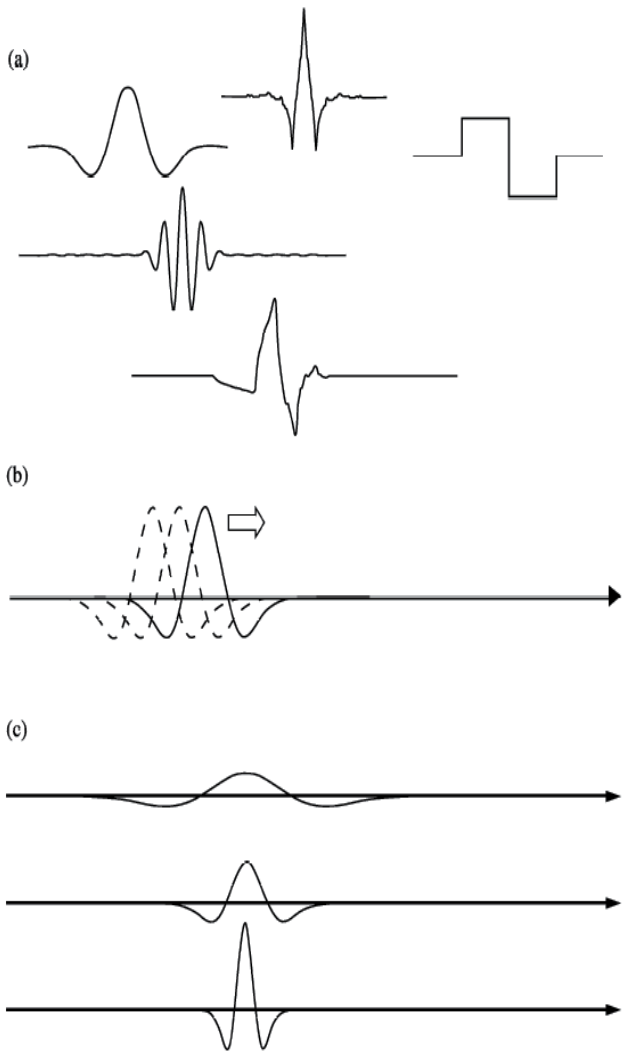


Fig. 1. Mother wavelets and positioning, scale concepts. (a) Different wavelets Mexican Hat, Harr, Morlet and Daubechies. (b) Positioning of a wavelet trough time. (c) Scale of a wavelet.

According with (Tim, 1991) the discrete wavelet transform is an orthogonal function which can be applied to an infinite group of data. Functionally, it is like the discrete Fourier Transform, which is based on a orthogonal function to apply the transformation. A signal passed twice through the transformation which is unchanged, and the input signal is assumed to be a set of discrete-time samples. Both transforms are convolutions.

Whereas the transform function of DFT is a sinusoid. The wavelet basis is a set of functions which are defined by a recursive function.

$$\varphi(n) = \sum_{k=0}^{N-1} h(k)\varphi(n-k) \tag{3}$$

Eq. (3) is orthogonal to its translations and to its dilations; the range of the summation is determined by the specified number of nonzero coefficients N. The number of nonzero coefficients is arbitrary, and will be referred to as the order of the wavelet. The value of the coefficients is, not arbitrary and it is determined by specific constraints of normalization and orthogonality. These parameters are; area under the wavelet curve, normalization.

$$\sum_k c_k = 2 \tag{4}$$

Normalization requires:

$$\sum_k c_k c_k + 2k' = \begin{cases} 2 \text{ If k'} = 0 \\ \\ 0 \text{ otherwise} \end{cases} \tag{5}$$

Which means the above sum is zero for all coefficients no equal to zero, and the sum squares of all coefficients is two. Another important equation which can be derivated from the above conditions is:

$$\sum_{k=0}^{N_k-1} (-1)^k c_k k^m = 0 \tag{6}$$

One way to solve eq. (3) is to construct a matrix of coefficients values. This is a square NxN matriz where N is the number of nonzero coefficients. The matrix always has the eigenvalue equal to 1, and its corresponding (normalized) eigenvector contains, as its components, the value of the φ function at integer values of x.

This class of wavelet functions is constrained, by definition, to be zero outside of a small interval. This is what makes the wavelet transform able to operate on a finite set of data.(Addison, 2002) The wavelet functions which are normally used to perform transforms consist of sets of a well-chosen coefficients resulting in a function with a discernible shape. Two of these functions are Harr and Daubechies wavelets see figure 1(a).

2.2 Discrete Wavelet Transform DWT algorithm

The Mallat algorithm also know as pyramid algorithm (Mallat, 1989) is a computational efficient method of implementing the wavelet transform, and is used as a basis for hardware implementations (Motra et al., 2004; Zhang and Hu, 2004; Zhilu et al., 2002).

The pyramid algorithm operates on a finite set of input data, where N is not necessary to be a power of two as we will explain in section 2.2.1. The value of the number of input points will be named as the *input block size*. The data are passed trough two convolutions functions; each creates an output stream that is half the length of the original input. These convolutions functions are filters; one half of the output is produced by the *low-pass* filter function eq.(7).

$$y_{low-pass}[k] = \sum_n x[n] h_0[2k-n] \tag{7}$$

And the other half is produced by the *high-pass* filter function eq. (8).

$$y_{high-pass}[k] = \sum_n x[n] h_1[2k-n] \tag{8}$$

Where N is the input block size, h_{0or1} is the wavelet function with well-chosen coefficients (filters) $x(n)$ is the input function and $y_{low\text{-}pass}$, $y_{high\text{-}pass}$ are respectively the low-pass and the high-pass outputs. In many situations, the low-pass output contains most of the *information content* of the original input signal. In general, higher-order wavelets (those with more non-zero coefficients) tend to put more information into the *low-pass* output and less into the *high-pass* output. If the average amplitude of the high output is low enough, then the high output may be discarded without greatly affecting the quality of the reconstructed signal. The high output is named also as detail output and the low output is know as approximation output, with this is logical to work with the approximation of the signal than with the details of the signal (Addison, 2002; Tim, 1991). A scheme of this decomposition is presented in figure 2.

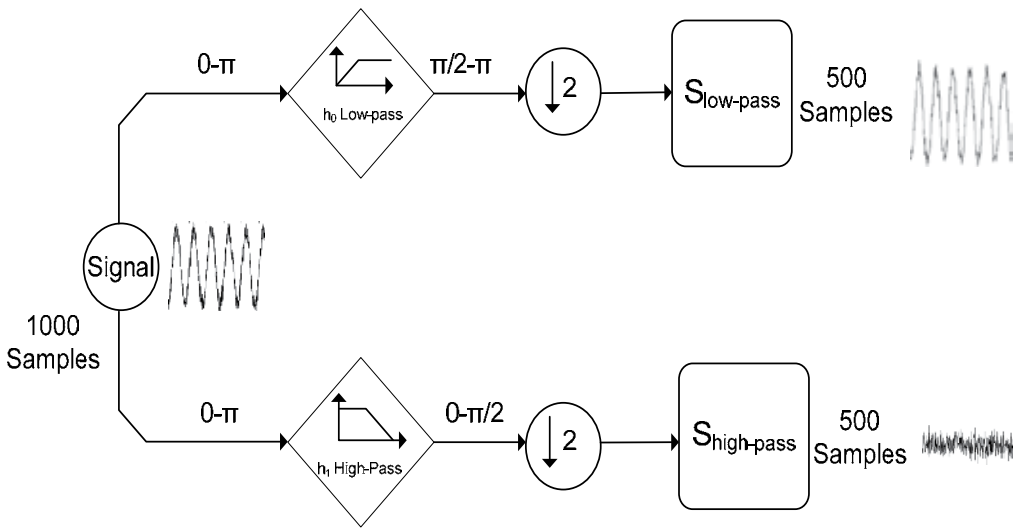


Fig. 2. Decomposition of a signal by the pass of the signal trough two filters.

Since most of information is in the *low- pass* output, it is easy imaging to continue transforming we have to rename the output as input and pass it trough h_0 and h_1, again to get two new sets of data, each one quarter of the size of the original block size, if again less information is contained in output filter, we can discarded and take just low-pass output.
Each step of retransforming the *low-pass* output is called *dilation* (Tim, 1991), and the maximum of dilations can be performed will result in a single *low-pass* value and a single *high-pass* value.
The algorithm developed allows the implementation of a DWT and it is described as follows: we suppose that have an input sequence of 10 samples of size, with the values of low-pass coefficients we compute the eq (7). The calculation for the first value of the low-

pass output in a first level is carried out by the matching the coefficients in the initial input sequence sample, after the low-pass coefficients are positioned two index positions and matching again with the input sequence, applying the eq. (7) again we will have the second value of the low-pass output for the first level. We repeat this until match the last sample of the input sequence with the first coefficient of the low-pass filter. At the end we will have a half input signal expressed by low-pass output, we need to do the same with high-pass coefficients to get the high-pass output. As we have described before, we can keep just the low-pass signal which has the most representative information of the total signal. For more decomposition levels we have to rename the low-pass output as input signal and repeat the complete algorithm. A scheme of this algorithm is presented in figure 3. It is important notice that before matching the coefficients and the input signal we have to add some values (number of coefficients -1) at the beginning and at the end of the sequence to avoid the bound effects.

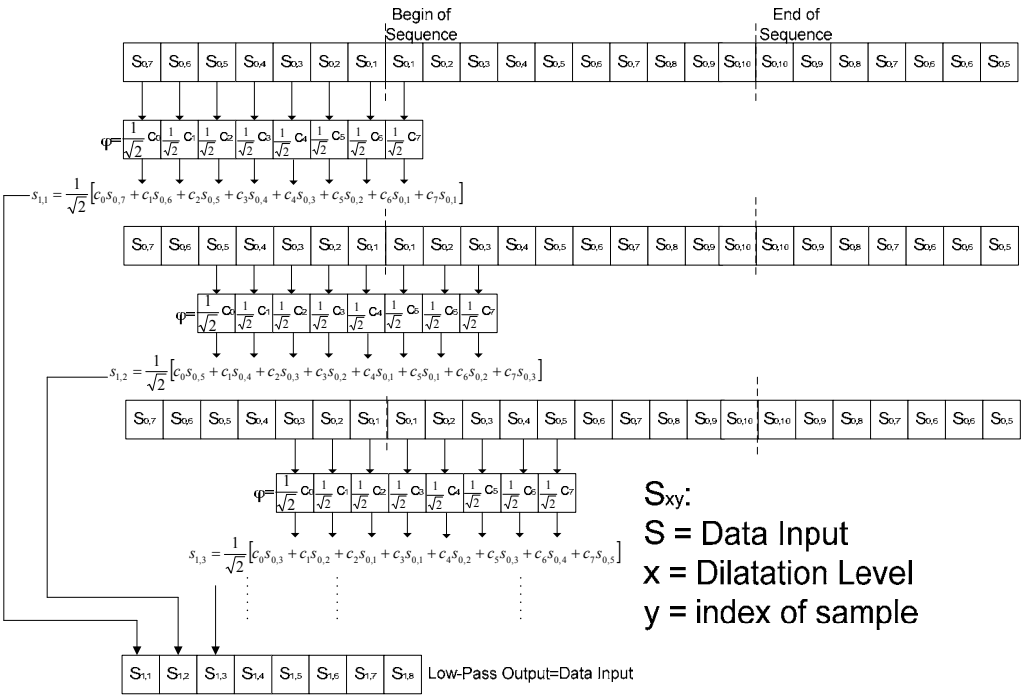


Fig. 3. Low-pass coefficients passing through an input data for a first dilation (first decomposition level) (Addison, 2002).

We can establish that successive dilations represent lower and lower frequency information by halves. It also clears that high rates of compression may require large block sizes of the input, so that more dilations can be made, and so that lower frequencies can be represented in the decomposition.

It must be notice that in some applications a pre compression level test can be done to establish the best rate of compression-reconstructions level for a specific application (Moreno-Baron et al., 2006).

2.2.1 Boundary effects

The boundary effect is present when DWT is applying to a set of discrete data, and takes place when the wavelet function is positioning at the beginning and/or end of the stream data. When the convolution is in process with the first sample of the input data and no samples have been defined to avoid boundary effects, the result of the convolution are unpredictable due to the operation can be performed with data which are not in memory space giving as a result either an over load of memory or the modification of the output signal. In this sense, if there is more than one decomposition level, the reconstruction will be inaccuracy.

There are many ways to avoid the boundary effects, these approximations include: (a) make the convolution equal to zero at the beginning and the end of the transformation, this solves the problem of the load memory errors but makes an inaccuracy transform. (b) another way to avoid the boundary effect is adding to the input data N-1 samples, where N is the number of coefficients of wavelet function, with non zero values. The main disadvantage is in the implementation of DWT the value of these non zeros coefficients must be calculated a priori according with the noise of the signals. This technique is no able for signals with different noise rates. (c) the most usual technique to avoid the boundary effect is by the reflection of the last samples of the input data, due to the added samples are from the same signal the accuracy of the reconstruction signal is not compromised. The reflection scheme can be seen in figure4.

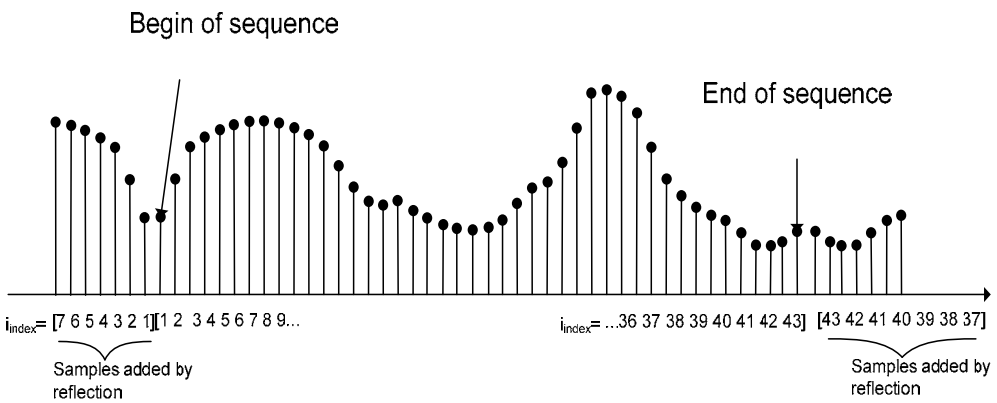


Fig. 4. Reflection of a signal to avoid the boundary effect.

2.3 Implementing the wavelet transform in a digital system

The implementation of digital wavelet transform involves some tasks which must be carried out before implementation. (1) Identify if the application needs; filtering, decomposition and/or reconstruction due to some applications do not need all the previous features of a wavelet transform, more features applied involve more hardware resources, memory, speed, large number of variable etc. (2) set the decomposition level needed to satisfied pre-treatment conditions, this can be done by a previous test of decomposition-reconstruction analysis. (3) Chose the type and number of filter coefficients. This can be done by the simplicity of implementation and/or by the order of filtering.

As we describe previously, in many chemical applications the signals from biosensors have high frequency noise, and imply a large number of samples, for this, we can use the wavelet transform to denoise and compress biosensor signals.
The design of the implementation of the Discrete Wavelet Transform (DWT) can be carried out as follows.

1. Compute and simulate the filter coefficients, one option can be Matlab® from Mathworks® software which includes toolbox with the most known wavelets and some practical algorithms to compute them. To denoise and compress, it is necessary calculate the coefficients for low-pass filter, the number of coefficients is correlated with the value of noise, this value should be measured before by an analysis of rate signal-noise. The level of the decomposition should be set by the previous compress-reconstruction analysis.
2. Confirms the function of the filter and level of decomposition with one real signal by simulation software.
3. Chose the language programming to write the code of DWT that can be after implemented in a programmable device, one of the best options is C code which can work with the IEEE754 format and it is supported by several dedicated devices.
4. Define fixed memory for the constants values of wavelet coefficients, level of decomposition and reserve data memory for two arrays, the first one has the same size of input data plus 2*length filter coefficients, second one has half size of input data plus 2*length filter coefficients (temporal array).
5. Develop a normalization function due the some application are based on concentrations of compounds which means the value of the data are represented in scientific format, for computation tasks is recommended rescale values from 0 to 1 see eq. (8).

$$N = (\text{Data-minimum})/(\text{maximum-minimum}) \tag{8}$$

6. Reflect the input sequence at the beginning and at the end of the input array.
7. Compute the convolution based on figure 3 and storage the result in the temporal array.
8. Rename the temporal array as input data and the input data as temporal array.
9. Repeat step 6 and compute the convolution again and storage the result in the temporal array.
10. Repeat steps 6,7,8, 9 until reach the level of compression set.

2.3.1 DsPIC30F6014 microcontroller

The dsPIC® microcontrollers are 16-bit Digital Signal Controllers (DSC) from Microchip®. They preserve the compatibility from classical PIC microcontrollers. They have some special characteristics such as 144 Kbytes of program memory; a large size of memory which makes possible implemented mathematical tools. The dsPICs can execute more than 30 millions instruction per second (30 MIPS) one of the fastest microcontroller from Microchip®. dsPICs have an optimized C compiler and 8-K of random access memory (RAM) and 4-K of read only memory (ROM).

3. Denoising and compression of biosensor based on screen-printed electrodes response using DWT

During years, the large interest in inhibition biosensors has been focused on the kind of responses and the quantitative relationships obtained using immobilized enzyme. A variety

of linear, non-linear, logarithmic responds between inhibition percentage and either inhibition concentration or incubation time were reported (Evtugyn et al., 1999; Fennouh et al., 1997; Hart et al., 1997). Concerning biosensors based on acetylcholinesterase inhibition, a theoretical model was reported based on diffusion limited inhibitor transport, taking into account the heterogeneous nature of enzyme inhibition that results from immobilization at the transducer surface (Zhang et al., 2001) :

3.1 Biosensor preparation

The biosensor was constructed as follows: Screen-printed three-electrode system with Cobalt-phthalocyanine-modified carbon as working electrode, graphite as counter electrode and Ag/AgCl as a reference electrode were fabricated using a DEK248 screen-printing system (Weymounth, UK). The biosensor based on the genetically-modified enzyme (B394) was mixed with polymer polyvinyl alcohol (PVA) in a rate of percentage of 30/70 (v/v). The enzyme was immobilized directly on the working electrode surface by entrapment in PVA-AWP polymer. For this purpose, the homogeneous mixture of B394, and PVA-AWP polymer was prepared and 3µL were carefully applied on the graphite working electrode surface. The final enzyme amount deposited on each electrode was 1mIU. The electrodes were exposed 5 h under a neon lamp (15 W) at 4°C to carry out photopolymerization and were ready to use after drying for 48 h (Andreescu and Marty, 2006b; Andreescu et al., 2002; Cortina et al., 2008; Silva Nunes et al., 2004; Valdés-Ramírez et al., 2009).

3.2 Procedures

The mechanism of inhibition of AChE by organophosphate compounds is well known, it can be summarized by the following reactions (Charpentier et al., 2000):

$$E + PX \underset{}{\overset{kd}{\leftrightarrow}} E^*PX \xrightarrow{k2} EP + X$$

with E = Enzyme, PX = Organophosphate insecticide, X = Leaving group.
The inhibitor phosphorylates the active-site, and the inhibition can be considered as irreversible in the first 30 min. This scheme can be simplified with the bimolecular rate constant ki = k2/Kd. However, the immobilization of the enzyme on the electrode surface does not allow the calculation by direct equations of all the constant rates involved in an enzyme-catalyzed reaction (Walsh et al., 2010).
The reactions that take place in an acetylcholinesterase-based sensor are described in scheme 1 (Silva Nunes et al., 2004).

$$\text{Acetytiocholine} \xrightarrow{\text{AChE}} \text{Thiocholine} + CH3COOH$$

$$\text{Thiocholine} + 2\text{Co-PC(ox)} \rightarrow \text{Thiocholine-oxidized form} + 2\text{Co-PC(red)}$$

$$\text{Co-PC(red)} \rightarrow \text{Co-PC(ox)} + 2e^-$$

The flow of electrons is proportional to the rate of acetylthiocholine hydrolysis, which decreases upon phosphorylation by OP of a serine present in the enzyme active site. As can be noticed in the previous scheme, the mediator Co-PC transforms the ionic current into

flow of electrons, the main advantage of Co-PC allows the oxidation-reduction using low voltages, (100mV) and giving a more stable and higher signal compared with other mediators (Shaidarova et al., 2004). The use of Cobalt-phthalocyanine involves higher current signals but also high levels of noise.

Inhibition protocol is described as follows: the biosensor is vertically inserted into an analytical cell containing phosphate buffer (10mL) under constant magnetic stirring at constant temperature (30°C). ATChCl solution (final concentration in the cell 1 mM) is added in the cell and the signal obtained at steady-state current was recorded. This step is repeated to ensure the stability of the biosensor. Inhibition is measured after addition at the steady-state step a known concentration of pesticide solution, as described in fig. 5.

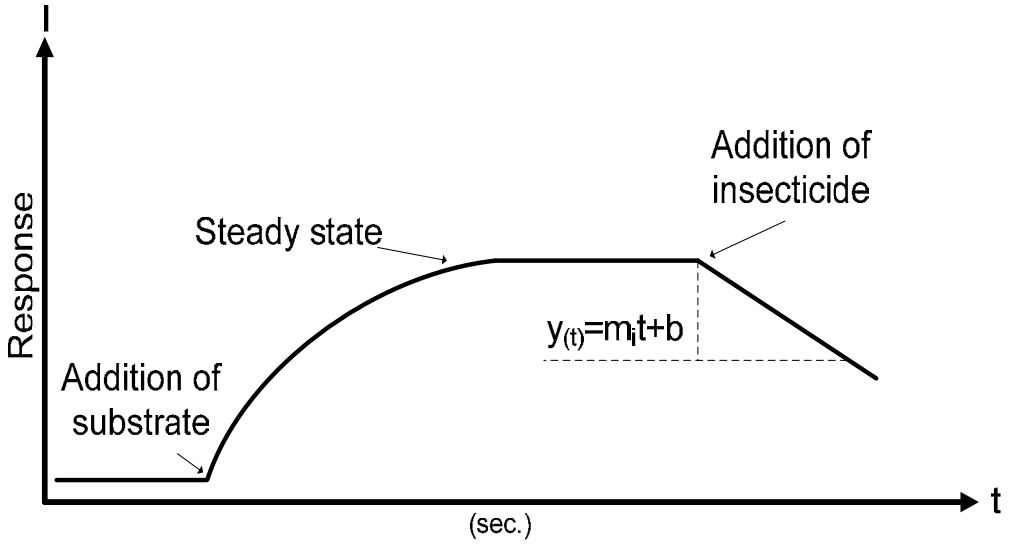


Fig. 5. Protocol of measurement of enzyme inhibition, typical response

From figure 5 we can see that to detect the pesticide, the part of the signal which has the analytical information is the signal generated by the addition of pesticide (the slope) to analyze *y(t)* we have to wait until the previous process (addition of substrate and steady-state) are completed. When the concentration of pesticides is very low, the time need to reach a slope can take several minutes, and to approximate the slope data to the model eq. (10) due to the noise and size of the signal sometime is not an easy tasks

$$y_{(t)} = - m_i t + b \tag{10}$$

The analog signal is acquired by the Analog/Digital converter peripheral of the dsPIC30F6014 microcontroller which computes the DWT and sends the digital data to a PC by RS232 interface to storage the inhibition curves. The protocol of the acquisition is shown in figure 6(a).

The DWT was implemented by Daubechies (Db) coefficients which were used such as banks of filters. Daubechies filters are recommended for discrete applications (Addison, 2002; Tim, 1991). The calculus of the Db coefficients and the level of decomposition were computed a priori in Matlab®.

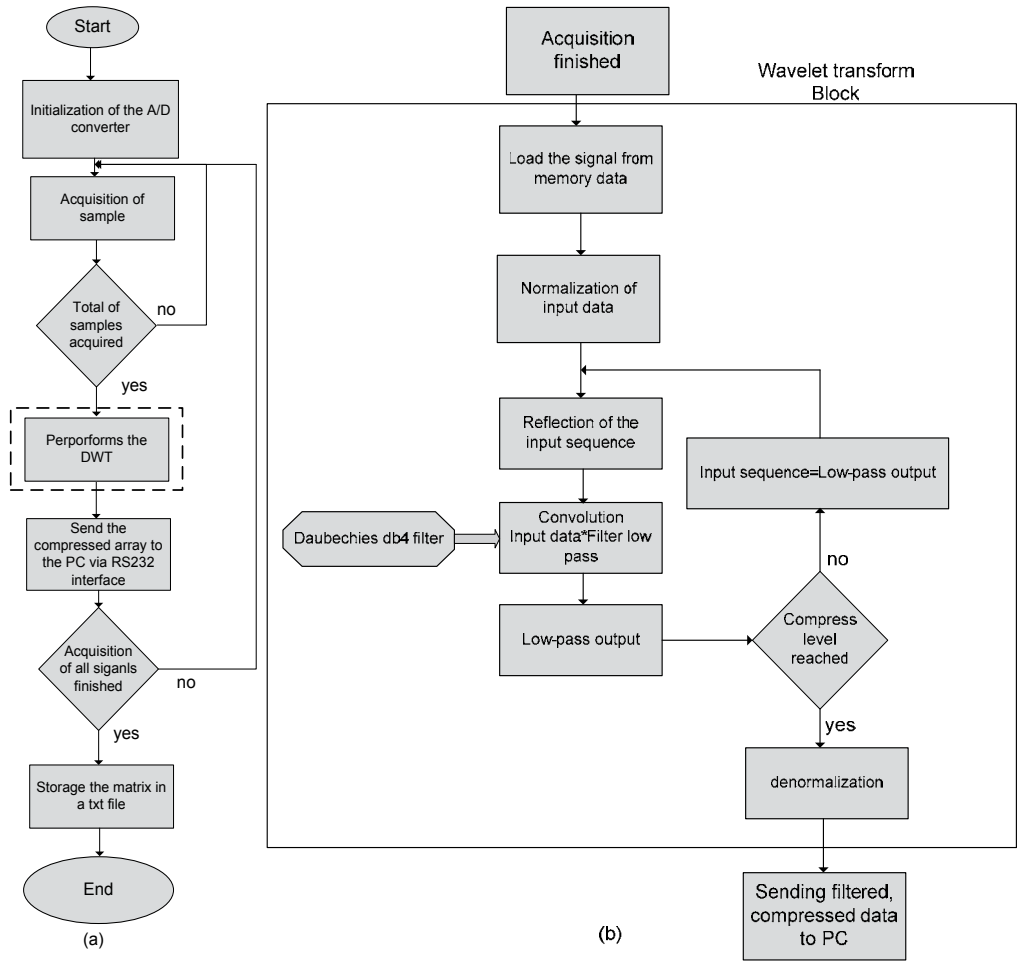


Fig. 6. Protocol of (a) Acquisition signals and (b) Computation of the DWT by dsPIC30F6014®.

4. Obtained results

The recorded signals correspond to inhibitions carried out as follows: the activity of biosensors is tested two times to prove the stability, after, when the activity has reached the *plateau* for second time, a known concentration of pesticide is added into the electrochemical cell (batch) and a decreasing of activity is observed(slope due to inhibition). The slope of inhibition is correlated with the concentration of pesticide. This process was repeated three times to ensure the reproducibility of the measurement. Three pesticides were used, Chlorpyruphos oxon (CPO), Chlorfenvinphos (CFV) and Azynphos methyl-oxon (AZMO) which are in the list of priority substances in the field of water policy of European community (Decision, 2001).

A total of 105 voltamograms (with 800 samples each one) were digitalized sequentially (with acquisition rate of 1Hz), sending to PC via rs232 interface and storaged in the dsPIC memory (see Figure 7). The input vector was reflected in both sides to avoid the boundary

effects, then, the vector was filtered by the Db filter, this was the first decomposition level, and the low-pass output was named as input vector for the next decomposition level. The algorithm repeats the above mentioned process until gets the decomposition level after the compressed signal is also sent it to a PC via rs232-interface.

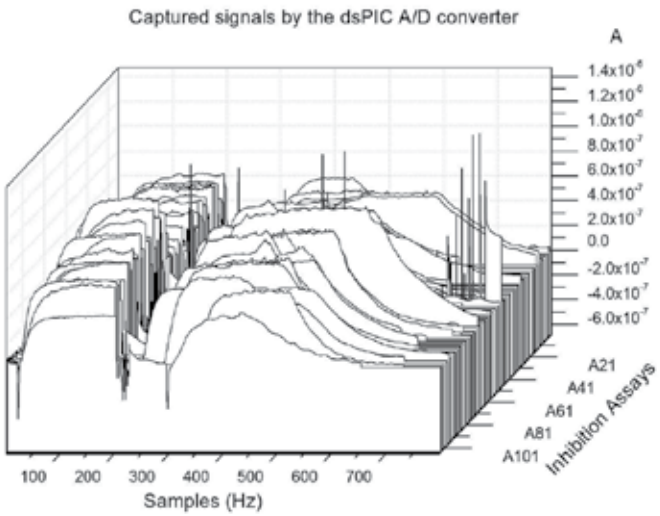


Fig. 7. The matrix of the voltammograms captured by the ADC of dsPIC30F6014.

A priori test was made to optimize the level of the composition and the order of wavelet used in the implementation, the test consisted in the use of different orders of Daubechies and different levels of compression to determine the ratio of similarity of the signal constructed from low-pass outputs. The Daubechies 4 and a 4th level (see figure 8) of compression were used to fixed parameters in the hardware implementation.

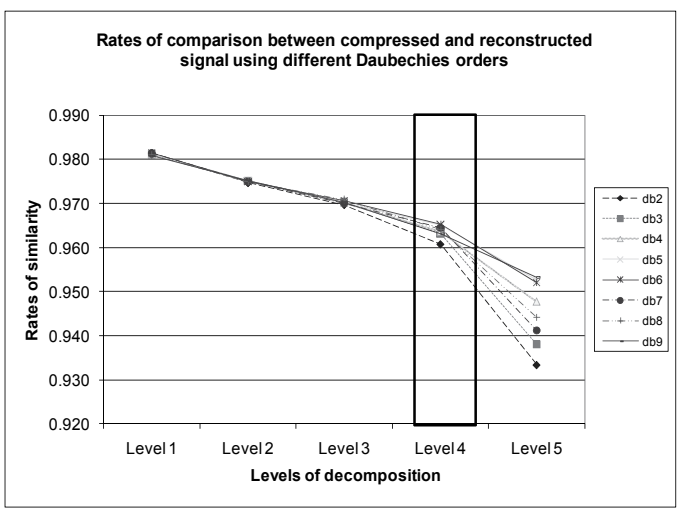


Fig. 8. Ratios of reconstructed signals using different orders of Daubechies wavelets and different levels of decompositions.

Daubechies coefficients were computed in Matlab® and saved in the dsPIC mamory as constant vector. The values of Daubichies coefficients are shown in table 1.

index k	Daubechies coefficients
0	0.32580343
1	1.010954572
2	0.89220014
3	-0.03957503
4	-0.26450717
5	0.0436163
6	0.0465036
7	-0.01498699

Table 1. Coefficients Daubechies used as filter low-pass.

The time needed to compute the DWT by the dsPIC6014 was 68.87ms. The total time is the time to acquire a complete signal (1 sample/second) plus the time to perform the DWT.

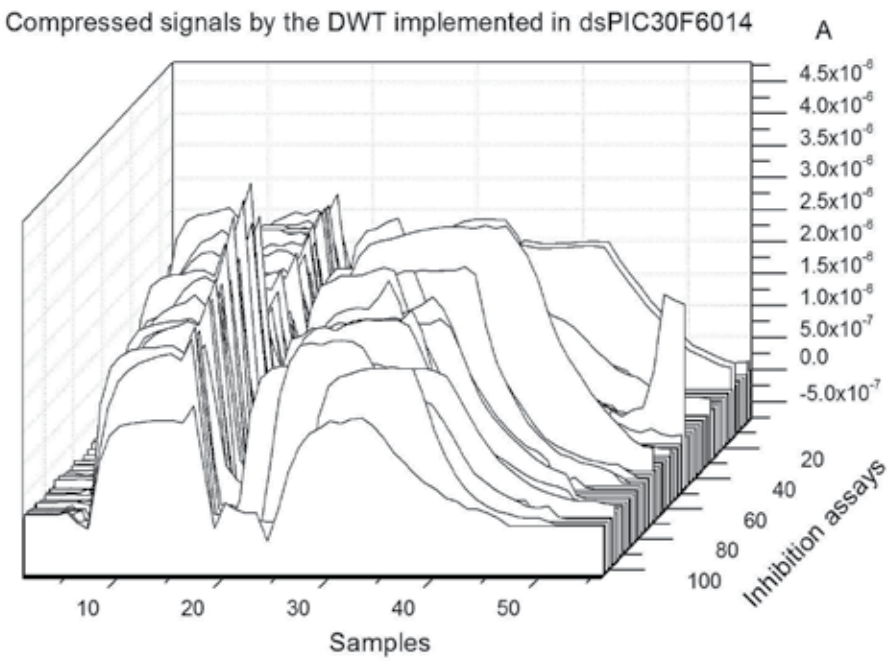


Fig. 9. The matrix of the voltammograms compressed by the dsPIC30F6010.

The analysis of the slopes was made in Matlab® by the function *polyfit* to determine the slope and interception point eq (10).
The values of the slopes (m_i) were correlated with the concentration of three different pesticides to determine the inhibition calibration curve for CPO, CFV and AZMO insecticides.

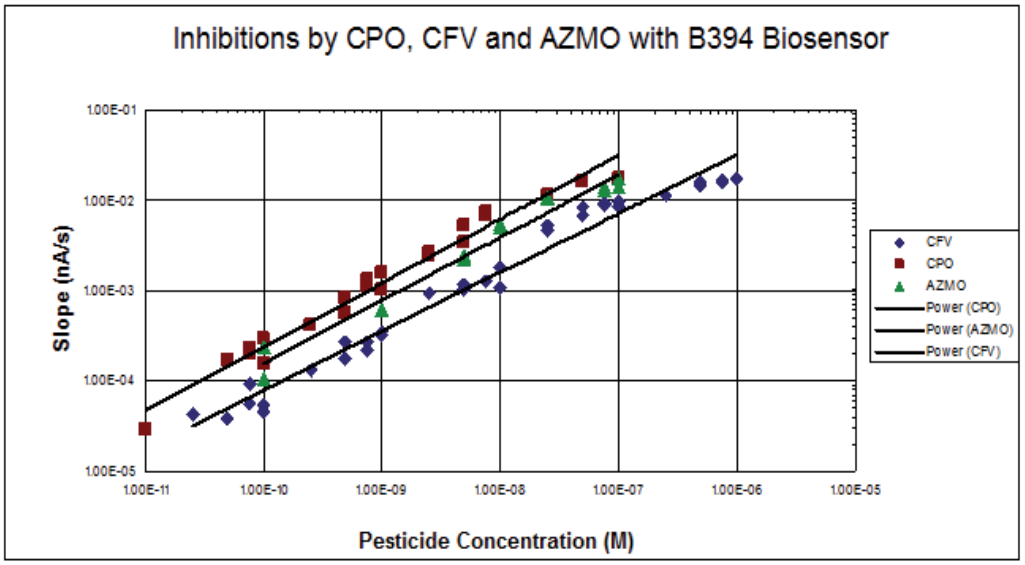


Fig. 10. Single Detections of CPO, CFV and AZMO insecticides by the Biosensor B394.

The response of the biosensor toward each pesticide was different as it is dependent on the sensitivity of the enzyme to the inhibitor. The limit of detection (LOD) was defined according to (Armbruster et al., 1994): considering the standard deviation of the slope σ observed in absence of inhibitor, the LOD corresponds to the lowest pesticide concentration inducing a slope increase equivalent to 3σ. In the same manner, the limit of quantification (LOQ) was defined as the pesticide concentration inducing a response equal to 10σ. In this case, the response considered for calculating the LOD was 3x10-4 nA/sec. The values of LOD and LOQ obtained for the used bio-sensor are summarized in table 2.

	B394		
Pesticide	Correlation coefficient	LOD (M)	LOQ (M)
CPO	0.9722	$1.664x10^{-10}$	$1.9256x10^{-10}$
CFV	0.9765	$6.81x10^{-10}$	$8.02x10^{-10}$
AZMO	0.9281	$3.3464x10^{-10}$	$3.7749x10^{-10}$

Table 2. Limit of detection and limit of quantification for CPO, CFV and AZMO using the three different biosensors.

5. Conclusions

In this wok a novel processing tool based on DWT is presented. DWT was implemented in a low-cost dedicated system able to analyse voltammeric signals.. The main advantages of this implementation are related to the simplicity that can be modified characteristics such as the compaction level and the choice of wavelet coefficients for specific applications. The

algorithm described in this work, can be implemented in any digital system with out disadvantages of fixing the size of input data to power of two, the resolution of the data was solved by the supporting of the dsPIC® programming language which supports IEEE754 format.

A complete case study of signal treatment in electrochemical analysis has been shown in detail. Based on its results, it is possible to observe that the implementation of DWT was achieved successfully employing as a pre-treatment tool to denoise and compress voltamograms. The use of genetically modified enzyme allow to achieve low detections limits and performing inhibitions in short time without incubation stage. In this sense, the proposed methodology is well suited for *in-situ* applications.

6. Acknowledgements

The authors greatly acknowledge the PCP Franco-Mexicain J100.116/2009, CONACyT-MAE for financial support through the project "Détection sélective et sensible d'organophosphorés par des biocapteurs utilisant des enzymes génétiquement modifiées" and the Consejo Nacional de Ciencia y Tecnología (CONACYT, México), for the scholarship granted to Gustavo Alonso.

7. References

Addison P.S. (2002) The illustrated wavelet transform handbook. *institute of Physics,* Publishing London, May 2002.

Alonso G., Ramirez A., Munoz R. (2010) Hardware implementation of the wavelet transform coupled with Artificial Neural Network for quantification purposes, *IEEE proccedings in Conference of Panamerican Healt Care Exchanges*, pp. 32-37.

Alsberg B.K., Woodward A.M., Kell D.B. (1997) An introduction to wavelet transforms for chemometricians: a time-frequency approach. *Chemometrics and Intelligent Laboratory Systems,* vol.37, pp. 215-239.

Andreescu S., Marty J.L. (2006a) Twenty years research in cholinesterase biosensors: from basic research to practical applications. *Biomolecular engineering,* vol.23, pp. 1-15.

Andreescu S., Barthelmebs L., Marty J. (2002) Immobilization of acetylcholinesterase on screen-printed electrodes: comparative study between three immobilization methods and applications to the detection of organophosphorus insecticides. *Analytica Chimica Acta*, Vol. 464, pp. 171-180.

Arduini F., Amine A., Moscone D., Palleschi G. (2010) Biosensors based on cholinesterase inhibition for insecticides, nerve agents and aflatoxin B₁ detection (review). *Microchimica Acta,* vol.170, pp. 193-214.

Armbruster D., Tillman M., Hubbs L.M. (1994) Limit of detection (LQD)/limit of quantitation (LOQ): comparison of the empirical and the statistical methods exemplified with GC-MS assays of abused drugs. *Clinical chemistry,* vol.40, pp. 1233.

Bachmann T., Leca B., Vilatte F., Marty J., Fournier D., Schmid R. (2000) Improved multianalyte detection of organophosphates and carbamates with disposable multielectrode biosensors using recombinant mutants of Drosophila acetylcholinesterase and artificial neural networks. *Biosensors and Bioelectronics,* vol.15, pp. 193-201.

Bahoura M., Hassani M., Hubin M. (1997) DSP implementation of wavelet transform for real time ECG wave forms detection and heart rate analysis. *Computer methods and programs in biomedicine,* vol.52, pp. 35-44.

Cai C., Harrington P.B. (1998) Different discrete wavelet transforms applied to denoising analytical data. *Journal of Chemical Information and Computer Sciences*, vol.38, pp. 1161-1170.

Charpentier A., Menozzi P., Marcel V., Villatte F., Fournier D. (2000) A method to estimate acetylcholinesterase-active sites and turnover in insects. *Analytical Biochemistry,* vol.285, pp. 76-81.

Cortina M., del Valle M., Marty J.L. (2008) Electronic Tongue Using an Enzyme Inhibition Biosensor Array for the Resolution of Pesticide Mixtures. *Electroanalysis* vol.20, pp. 54-60.

Crochiere R.E., Flanagan J.L., Webber S.A. (1977) Digital speech communication system for minimizing quantizing noise, *Google Patents.*

Croisier A., Esteban D., Galand C. (1976) Perfect channel splitting by use of interpolation/decimation/tree decomposition techniques. pp. 443-446.

Decision. (2001) 2455/2001/EC of the European Parliament and of the Council of 20 November 2001 establishing the list of priority substances in the field of water policy and amending Directive 2000/60/EC (Text with EEA relevance). *Official Journal of the European Communities* vol.15, pp1-5.

Ehrentreich F. (2002) Wavelet transform applications in analytical chemistry. *Analytical and Bioanalytical chemistry* vol.372, pp. 115-121.

Evtugyn G.A., Ivanov A.N., Gogol E.V., Marty J.L., Budnikov H.C. (1999) Amperometric flow-through biosensor for the determination of cholinesterase inhibitors. *Analytica Chimica Acta* vol.385, pp.13-21.

Fennouh S., Casimiri V., Burstein C. (1997) Increased paraoxon detection with solvents using acetylcholinesterase inactivation measured with a choline oxidase biosensor. *Biosensors and Bioelectronics* vol.12, pp97-104.

Gabor D. (1946) Theory of communication. Part 1: The analysis of information. Electrical Engineers-Part III. J*ournal of the Institution of Radio and Communication Engineering*, vol.93, pp. 429-441.

Gnavi S., Penna B., Grangetto M., Magli E., Olmo G. (2002) Wavelet kernels on a DSP: a comparison between lifting and filter banks for image coding. EURASIP *Journal on Applied Signal Processing* 2002, pp. 981-989.

Gorry P.A. (1990) General least-squares smoothing and differentiation by the convolution (Savitzky-Golay) method. *Analytical Chemistry,* vol.62, pp. 570-573.

Hart A.L., Collier W.A., Janssen D. (1997) The response of screen-printed enzyme electrodes containing cholinesterases to organo-phosphates in solution and from commercial formulations. *Biosensors and Bioelectronics,* vol.12, pp. 645-654.

Istamboulie G., Andreescu S., Marty J., Noguer T. (2007) Highly sensitive detection of organophosphorus insecticides using magnetic microbeads and genetically engineered acetylcholinesterase. *Biosensors and Bioelectronics,* vol.23, pp.506-512.

Jichang G., Yan L., Xiangping L., Qiang L., Yaqi Z. (2003) The realizations of fast wavelet transform algorithms based on DSP, *IEEE transaction on Electrical and Computer Engineering, 2003*. pp. 2005-2008.

Leung A.K.-m., Chau F.-t., Gao J.-b. (1998a) A review on applications of wavelet transform techniques in chemical analysis: 1989-1997. *Chemometrics and Intelligent Laboratory Systems*, vol.43, pp. 165-184.

Mallat S.G. (1989) A theory for multiresolution signal decomposition: The wavelet representation. *IEEE transactions on pattern analysis and machine intelligence,* vol.11, pp. 674-693.

Moreno-Baron L., Cartas R., Merkoci A., Alegret S., Del Valle M., Leija L., Hernandez P., Munoz R. (2006) Application of the wavelet transform coupled with artificial neural networks for quantification purposes in a voltammetric electronic tongue. *Sensors and Actuators B*, vol.113, pp.487-499.

Morlet J., Arens G., Fourgeau E., Glard D. (1982) Wave propagation and sampling theory—Part I: Complex signal and scattering in multilayered media. *Geophysics*, vol.47, pp. 203.

Motra A., Bora P., Chakrabarti I. (2004) An efficient hardware implementation of DWT and IDWT, *Conference on Convergent Technologies for Asia-Pacific Region*, pp. 95-99.

Nie L., Wu S., Wang J., Zheng L., Lin X., Rui L. (2001) Continuous wavelet transform and its application to resolving and quantifying the overlapped voltammetric peaks. *Analytica Chimica Acta*, vol.450, pp. 185-192.

Shahbahrami A., Ahmadi M., Wong S., Bertels K. (2010) A New Approach to Implement Discrete Wavelet Transform using Collaboration of Reconfigurable Elements, *International Conference on Reconfigurable Computing and FPGAs*, pp. 344-349.

Shaidarova L.G., Ziganshina S.A., Medyantseva E.P., Budnikov G.K. (2004) Amperometric Cholinesterase Biosensors with Carbon Paste Electrodes Modified with Cobalt Phthalocyanine. *Russian Journal of Applied Chemistry* vol.77, pp. 241-248.

Silva Nunes G., Jeanty G., Marty J. (2004) Enzyme immobilization procedures on screen-printed electrodes used for the detection of anticholinesterase pesticides:: Comparative study. *Analytica Chimica Acta*, vol.523, pp. 107-115.

Tim T.E. (1991) Discrete wavelet transforms: Theory and implementation, *technical report*, Stanford University.

Valdés-Ramírez G., Gutiérrez M., Del Valle M., Ramírez-Silva M., Fournier D., Marty J. (2009) Automated resolution of dichlorvos and methylparaoxon pesticide mixtures employing a Flow Injection system with an inhibition electronic tongue. *Biosensors and Bioelectronics,* vol.24, pp. 1103-1108.

Walsh R., Martin E., Darvesh S. (2010) A method to describe enzyme-catalyzed reactions by combining steady state and time course enzyme kinetic parameters. *Biochimica et Biophysica Acta* vol.1800, pp.1-5.

Zanchettin C., Ludermir T. (2007) Wavelet filter for noise reduction and signal compression in an artificial nose. *Applied Soft Computing,* vol. 7, pp.246-256.

Zhang S., Zhao H., John R. (2001) A Theoretical Model for Immobilized Enzyme Inhibition Biosensors. *Electroanalysis* vol.13, pp. 1528-1534.

Zhang Z., Hu J. (2004) FPGA implementation of 4 samples DWT based on the model of pyramidal structural data coding, *Fourth International Conference on Computer and Information Technology,* pp. 819-823.

Zhilu W., Guanghui R., Yaqin Z. (2002) A study on implementing wavelet transform and FFT with FPGA, *IEEE Proceedings. 4th International Conference on ASIC*, 2001 . pp. 486-489.

Zou X., Mo J. (1997) Spline wavelet analysis for voltammetric signals. *Analytica Chimica Acta*, vol.340, pp.115-121.

9

Multiscale Texture Descriptors for Automatic Small Bowel Tumors Detection in Capsule Endoscopy

Daniel Barbosa[1], Dalila Roupar[2] and Carlos Lima[2]
[1]*Katholieke Universiteit Leuven*
[2]*Universidade do Minho*
[1]*Belgium*
[2]*Portugal*

1. Introduction

Conventional endoscopic exams do not allow the entire visualization of the gastrointestinal (GI) tract. Push enteroscopy (PE) is an effective diagnostic and therapeutic procedure, although it only allows exploration of the proximal small bowel (Pennazio et al., 1995). Simultaneously, convetional colonoscopy is limited at the terminal ileum. Therefore, prior to the wireless capsule endoscopy era, the small intestine was the conventional endoscopy's last frontier, because it could not be internally visualized directly or in it's entirely by any method (Herrerías & Mascarenhas-Saraiva, 2007). The small intestine accounts for 75% of the total length and 90% of the surface area of the gastrointestinal tract. In adults it measures about 570 cm at *post mortem*, which is substantially longer than conventional video endoscopes (100-180 cm) (Swain & Fritscher-Ravens, 2004). Intraoperative enteroscopy is the most complete but also the most invasive means of examining the small bowel (Gay et al., 1998).
Given the technical and medical improvements introduced on the assessment of the gastrointestinal (GI) tract, Capsule Endoscopy (CE) is considered as the first major technological innovation in GI diagnostic medicine since the flexible endoscope (Kaffes, 2009). More recently, a new technique, the double-balloon enteroscopy (DBE), has been introduced into clinical practice (Yamamoto & Kita, 2006). DBE has the potential to examine the entire length of the small bowel with biopsy and therapeutic capability. Nevertheless, it is a time consuming procedure that requires specialist training for the operating physician. We should note that DBE and CE are complementary tools and not competitive (Chen et al., 2007). Hence, the diagnostic ease of CE can be complemented with a targeted and often therapeutic DBE (Kaffes, 2009). Therefore, CE can be used as a first line diagnosis method, while DBE can be used as a confirmatory or therapeutic modality for lesions first visualized by CE (Pennazio, 2006).
The endoscopic capsule is a pill-like device, with only 11mm x 26 mm, and includes a miniaturized camera, a light source and circuits for the acquisition and wireless transmission of signals (Iddan et al., 2000). As the capsule moves through GI tract, propelled exclusively by peristalsis, it acquires images at a rate of two per second and sends them to a hard disk receiver that is worn in the belt of the patient, in a wireless communication scheme. The acquisition

of images is limited by the battery life of the device, usually around eight hours, which imply that in a single CE exam more than 50000 images are acquired. If no complications arise, the capsule should be in the patient's stool, usually within 24-48 h, and not reused (Pennazio, 2006). Capsule endoscopy has evolved in a few short years to become a first-line, noninvasive diagnostic technique for the small bowel. CE is now being utilized worldwide to assess patients for obscure gastrointestinal bleeding, possible Crohn's disease, celiac disease and small bowel tumors (Lee & Eisen, 2010). It is now available in over 4500 practice sites around the world (Munoz-Navas, 2009).

The time required to a physician to analyze the resulting video is, on average, 40-60 min (Pennazio, 2006). The reading time and interpretation of CE exams is very time consuming given that more than 50,000 images have to be reviewed (Delvaux & Gay, 2006; Mergener et al., 2007), which contributes to the high cost of a CE exam (Westerhof et al., 2009). Thus, a computer assisted diagnosis tool to help the physicians to evaluate CE exams faster and more accurately is an important technical challenge and an excellent economical opportunity.

After the introduction of CE, it was discovered that the prevalence and malignancy rates for small bowel tumors are much higher than previously reported and that the early use of CE can lead to earlier diagnoses, reduced costs and, hopefully, prevent cancer (Herrerías & Mascarenhas-Saraiva, 2007).

The application of texture analysis techniques to classify capsule endoscopic frames is feasible and presents promising results. Kodogiannis *et al.* proposed two different schemes to extract features from texture spectra in the chromatic and achromatic domains (Kodogiannis et al., 2007). The first is a structural approach based in the theory of formal languages, while the second is a statistical approach, where statistical texture descriptors are calculated from the histograms of the RGB and HSV color spaces of CE video frames. Recently, Mackiewicz *et al.* proposed an automatic capsule endoscopy segmentation algorithm based in color and texture features to determine the topological division of capsule endoscopic videos (Mackiewicz et al., 2008). Several other CE image processing methodologies were reviewed recently by Karargyris and Bourbakis (Karargyris & Bourbakis, 2010).

In the present chapter, several multiscale texture descriptors are extracted from both wavelet and curvelet domains and their classification performance is assessed. In section 2, we present a brief introduction to multiscale image representation, namely through Discrete Wavelet Transform and Discrete Curvelet Transform. In section 3, the algorithm used to extract the multiscale texture descriptors is described, as well as modelling techniques that allow to extract statistical dependence of textural descriptors taken in different color channels. In section 4, the implementation details of the method are discussed. In section 5, the performance of the proposed methods is assessed and discussed. Section 6 resumes the key findings and presents future research orientation.

2. Multiscale representation of image information

It is known for a long time that human perception of texture is based in a multi-scale analysis of patterns, which can be modeled by multi-resolution approaches. In fact, the multi-resolution ability of the Discrete Wavelet Transform (DWT) has been vastly explored in several fields of image processing such as compression, denoising and classification. However, the directional information of the DWT is limited, which might not be enough to capture all the complex texture patterns within an image.

Introduced in 2000, the Continuous Curvelet Transform (CCT) is based in an anisotropic notion of scale and high directional sensitivity in multiple directions (Candès & Donoho, 2000). While wavelets are certainly suitable for dealing with objects where the interesting phenomena, e.g., singularities, are associated with exceptional points, they are ill-suited for detecting, organizing, or providing a compact representation of intermediate dimensional structures. Given the significance of such intermediate dimensional phenomena, there has been a vigorous effort to provide better adapted alternatives by combining ideas from geometry with ideas from traditional multi-scale analysis (Candès et al., 2006). Therefore, this tool can be used as a multi-resolution and multi-directional representation of the information within an image. The Discrete Curvelet Transform (DCT) coefficients are accurate representations of the original image with different detail, given by the different frequency content in each scale, but also with different detail in multiple directions, overcoming the directional limitations of the Discrete Wavelet Transform. This might be well suited for the analysis of complex spatio-frequency patterns as texture.

2.1 Wavelet Transform

The scientific breaktrough achieved with the introduction of the Wavelet Transform has inequivocally changed the research direction on the biomedical signal/image processing. Indeed, since the seminal works of Daubechies (Daubechies, 1988) and Mallat (Mallat, 1989), more than 9000 papers and 200 books were published between the late eighties and 2003, with a significant part being focused in biomedical applications (Unser et al., 2003). While not being exhaustive, and in order to provide to the interested reader an idea of some applications of wavelets in medical imaging, the DWT has been applied in several problems such as:

- Image denoising
- Compression of medical images
- Feature extraction and image classification
- Tomographic reconstruction
- Image Enhancement

Since the main goal of this section is to provide to the reader the key concepts regarding the DWT, the mathematical details will be kept to a minimum. The Continuous Wavelet Transform (CWT) is a signal representation in a scale-time space, and the CWT coefficients of a time-varying signal $x(t)$ are given through:

$$X_{\Psi}(\tau, s) = \int_{-\infty}^{+\infty} x(t)\Psi^{*}\left(\frac{t-\tau}{s}\right) dt \tag{1}$$

where Ψ is the mother wavelet function and * stands for the complex conjugate.
Analogously, signal can be recovered from its wavelet coefficients through the Inverse Continuous Wavelet Transform (ICWT):

$$x(t) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} X_{\Psi}(\tau, s)\Psi\left(\frac{t-\tau}{s}\right) d\tau ds \tag{2}$$

By varying both the scale and the translation shift parameters, s and τ respectively, we can obtain a family of daughter wavelets from the mother wavelet function Ψ:

$$\Psi_{s,\tau}(t) = \frac{1}{\sqrt{s}}\Psi\left(\frac{t-\tau}{s}\right), \tag{3}$$

Thus, for a fixed value of the scale parameter s, the CWT, which is now a function of the continuous shift parameter τ, can be written as a convolution equation where the filter corresponds to a rescaled and time-reversed version of the wavelet as shown by equation (1) setting t=0. Combining the variation of the scale parameter s and the time shift parameter τ, the CWT provides a complete scale-time representation of a signal. Furthermore, a given CWT coefficient can be simply seen as the inner product:

$$X_{\Psi}(\tau, s) = \langle x(t), \Psi_{s,\tau} \rangle. \tag{4}$$

Each $\langle x(t), \Psi_{s,\tau} \rangle$ can be seen as a quantom of information that is represented as a rectangle in the time-frequency plane. However, it can be shown through the uncertainty principle theorem that this rectangle has a minimum surface that limits the joint time-frequency resolution. The ultimate consequence of this limit is the trade-off between temporal and frequency resolution. Thus, an increase in the frequency resolution would lead to a decrease in the temporal resolution. In limit, we could perfectly identify the frequency content of the signal and simultaneously lose all the localization details of these frequency components, as in the case of the typical Fourier transform. However, and given the relevance of the temporal localization of abnormal frequency content phenomena, there is the need of a better use of the joint time-frequency resolution. A possible solution could be the use of Short Time Fourier Transform (STFT). However, the constant window size implies a uniform tilling of the time-frequency plane. The wavelet can tackle this problem very elegantly, as will be shown ahead.

As can be easily understood, by varying both the scale s and the time shift τ parameters, the CWT representation of the signal $x(t)$ is highly redundant. Indeed, the CWT maps the information within a one-dimensional signal to a two-dimensional time-scale joint representation. This implies a heavy computational burden which reduces its application potential to daily life problems. Nonetheless, this problem can be tackled by limiting the continuous scalability and translatability of the daughter wavelets $\Psi_{s,\tau}$. Using such scheme, the mother wavelet function Ψ generates a smaller family of daughter wavelets:

$$\Psi_{j,k}(t) = \frac{1}{\sqrt{s_0^j}} \Psi \left(\frac{t - k\tau_0 s_0^j}{s_0^j} \right) \tag{5}$$

where j and k are integers and s_0>1 a fixed dilation step. The discretization of the time-scale plane is usually achieved using a dyadic sampling, where s_0=2 and τ_0=1.

A wavelet atom $\Psi_{j,k}(t)$ is localized around the point $2^j k$ and has a support size proportional to the scale 2^j. Using this approach, the scale index j corresponds to the level of focus from the which the signal is viewed, which is related to the frequency range involved. Indeed, a lower j corresponds to the high frequency contents, which can be easily deducted from the support size.

In the time-frequency plane, the Heisenberg resolution box of $\Psi_{j,k}(t)$ is a dilation by 2^j and translation by $2^j k$ of the Heisenberg box of Ψ. This leads to a perfect tiling of the time-frequency plane illustrated in Fig. 1. The height and width of the Heisenberg boxes in the time-frequency plane represents the resolution on the frequency and time domains respectively. Note that there is an intrinsic trade-off between good temporal localization of high frequency content and good frequency resolution of low frequency content in the signal, as in opposition to the homogeneous time-frequency resolution of the STFT. This is the central advantage of the Wavelet Transform when compared to the Short Time Fourier Transform.

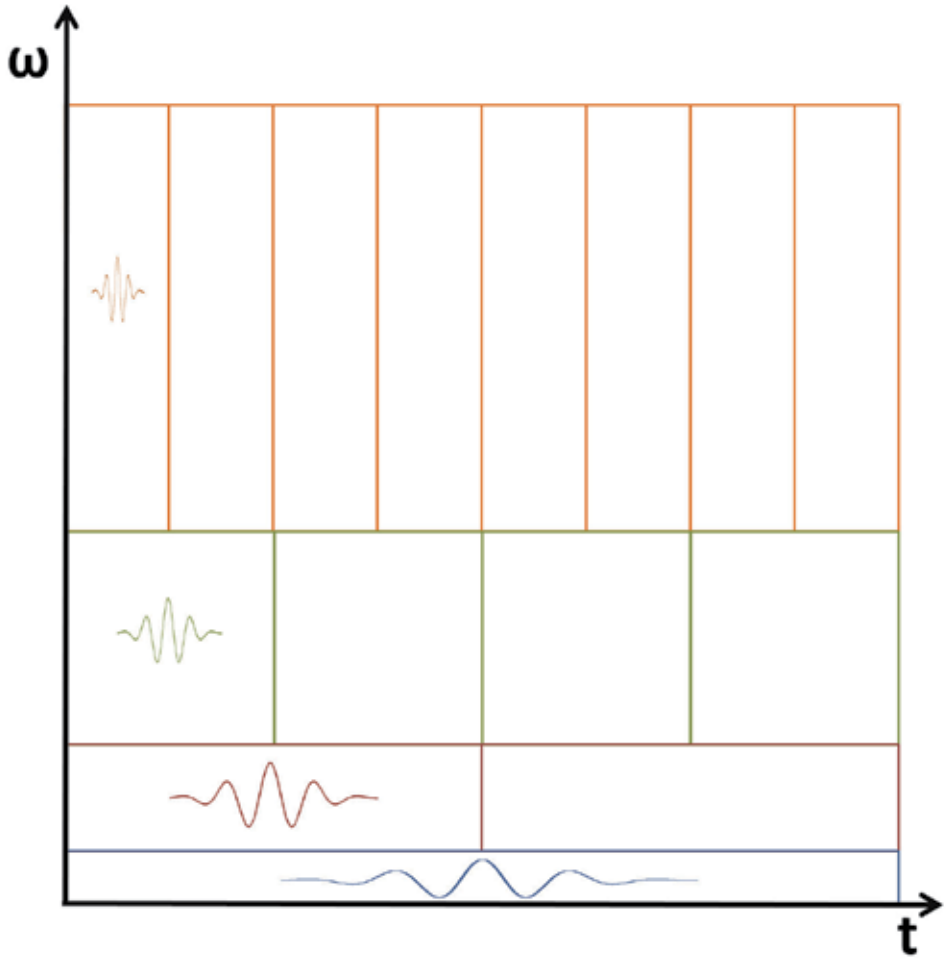


Fig. 1. Inhomogeneous dyadic tilling of the time-frequency plane with wavelet atoms.

From the previous figure, another property of the wavelet atoms $\Psi_{j,k}(t)$ becomes evident: their frequency spectrum has a band-pass nature. Indeed, from Fourier theory, we know that the dilatation operation in the time domain corresponds to a compression in the frequency domain and a shift towards the zero frequency. However, this would imply that an infinite number of scales would be needed to cover the entire frequency spectrum, as can be seen in Fig. 2. Since this would not be a plausible solution, Mallat (Mallat, 1989) introduced the scalling function to cover the spectrum not spanned by the wavelet atoms. Indeed, when analyzing a signal using a combination of a scalling function and wavelets, the scalling function will represent the signal information in the spectrum covered by all the wavelet atoms up to a scale j, while the remaining spectrum is analyzed by wavelets. This can be observed in Fig. 3.
When regarding the schematic representation of the Wavelet Transform in Fig. 3, one can easily observe that the tilling of the spectrum with scaling and wavelet functions is similar to consecutive low-pass and band-pass filter operations. Due to its low-pass nature, the scalling function will allow to extract coefficients that express an approximate version of the signal

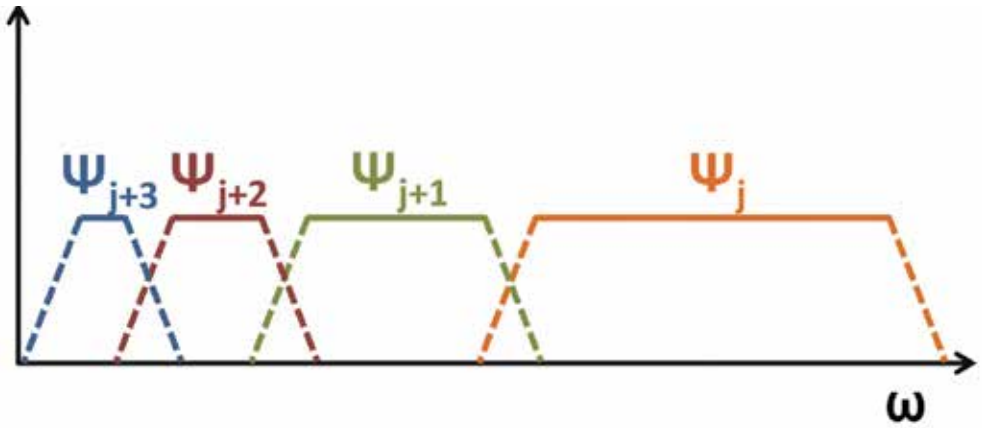


Fig. 2. Wavelet atoms spectrum resulting from dyadic scaling of the mother wavelet Ψ.

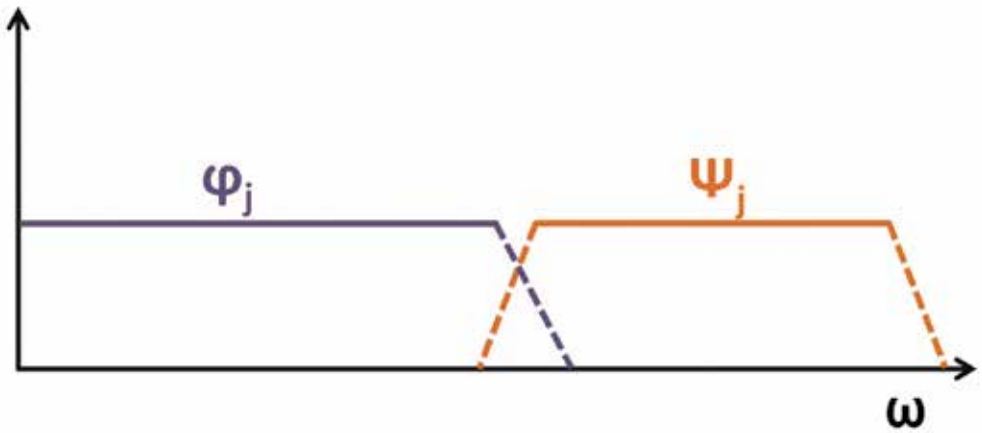


Fig. 3. Scaling function (ϕ) and wavelet atoms (Ψ) spectra. Note how the scaling function (ϕ_j) covers the spectral region spanned by the wavelet atoms until $j+1$.

content, whereas the wavelet coefficients correspond to detail information at different levels of focus. Although this concept, commonly designated as subband coding, is not unique from the Wavelet Transform, it greatly helps its practical implementation.

2.1.1 Discrete Wavelet Transform

In several practical problems, the signal under analysis is of discrete nature. Thus, there is the need to discretize the Wavelet Transform in order to use it in sampled discrete signals. This can be efficiently done using recursive filtering in iterated filter-bank as show in Fig. 4. Indeed, it can be shown that both scaling and wavelet coeficcients at a given scale can be calculated from the coefficients at the previous scale using appropriate filters. Furthermore, half of the samples after applying the filter can be eliminated according to the Nyquist's rule, since the signal now has only half of the bandwith.

There is an important relationship between these low and high-pass used recursively in the iterated filter-bank. Indeed, these filters are not independent from each other and satisfy the

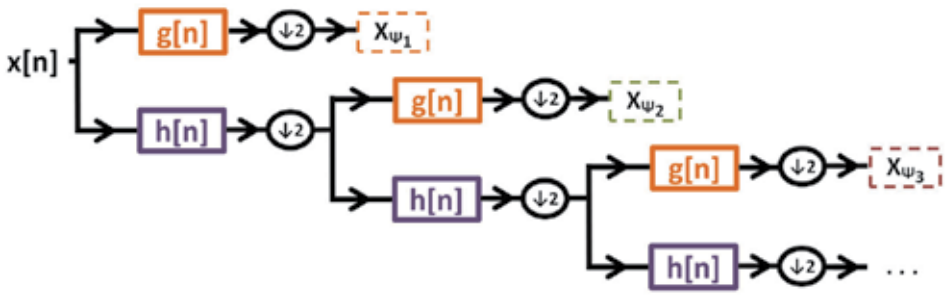


Fig. 4. Iterated filter-bank for practical implementation of the Discrete Wavelet Transform.

condition of Quadrature Mirror Filters (QMF):

$$g[L-1-n] = (-1)^n \cdot h[n], \tag{6}$$

where $g[n]$ is the high-pass filter, $h[n]$ is the low-pass filter and L is the filter lenght. The consecutive filter and subsampling operation at each scale can be expressed by:

$$y_{approx} = \sum_n x[k] \cdot h[-n+2k], \tag{7}$$

$$y_{detail} = \sum_n x[k] \cdot g[-n+2k], \tag{8}$$

Note that, given the finite time resolution of sampled signals, there is an upper limit in the scale at which the signal is analysed, which is of course dependent on the low-pass filter length. However, it is common that the DWT analysis is not done until this limit, being thus the output of the low-pass filtering corresponding to the scalling function coefficients. However, and for sake of readability, we shall globally denominate the coefficients arising from the application of these iterated filter-banks simply as DWT coefficients.
As for the continuous case, the original signal can be recovered from its DWT coefficients applying Inverse Discrete Wavelet Transform (IDWT). In order to reconstruct the signal, the iterated filter-banks has to be reversed. Thus, at each scale, the DWT coefficients are upsampled, by simply introducing a zero between every two samples, and then filter the signal with synthesis filters, closely related with the analysis filters used in the direct DWT. However, and in order to achieve perfect reconstruction, the analysis and synthesis filters must fulfill the following conditions (Daubechies & Sweldens, 1998):

$$\tilde{h}(z)h(z^{-1}) + \tilde{g}(z)g(z^{-1}) = 2, \tag{9}$$

$$\tilde{h}(z)h(-z^{-1}) + \tilde{g}(z)g(-z^{-1}) = 0, \tag{10}$$

where g and h are the analysis filters used in the direct DWT and $\tilde{h}$ and $\tilde{g}$ are the synthesis filters used in the inverse DWT.

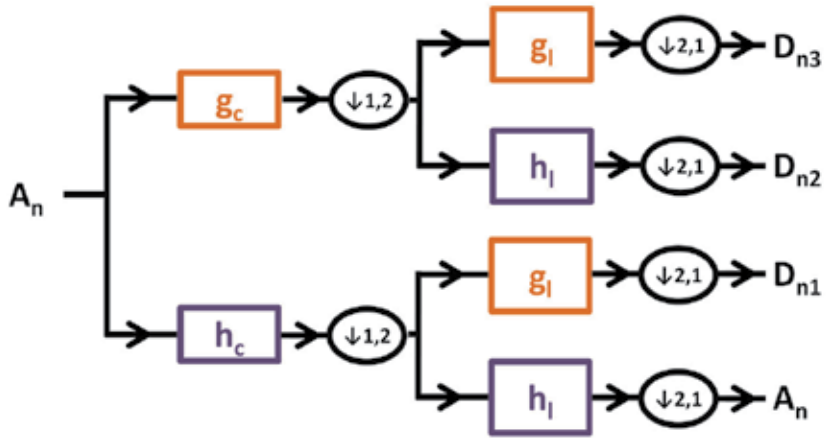


Fig. 5. Iterated filter-bank for 2D Discrete Wavelet Transform in images. Note that A will keep an approximation of the original image, whereas D will carry detail information.

2.1.2 2D Discrete Wavelet Transform

Given the one-dimensional nature of the signals under analysis in the Wavelet Transform framework, the direct application of this mathematical tool to higher dimensional signals is not directly possible. However, a N-D signal can still be analyzed along each of its N dimensions. Using this reasoning, Mallat also introduced in his seminal paper a very elegant extension of the concepts of multi-resolution decomposition in order to use the DWT in image processing problems. The key idea is to expand the application of 1D filterbanks to the 2D in a straightforward manner, by simply applying the designed QMF filters to the columns and rows separately. Thus, the DWT applied to an image can be implemented using an iterated filter-bank, illustrated in Fig. 5, which can be described as consecutive filtering and subsampling operations:

$$A_n = [h_l \star [h_c \star A_{n-1}]_{\downarrow 1,2}]_{\downarrow 2,1}, \tag{11}$$

$$D_{n1} = [g_l \star [h_c \star A_{n-1}]_{\downarrow 1,2}]_{\downarrow 2,1}, \tag{12}$$

$$D_{n2} = [h_l \star [g_c \star A_{n-1}]_{\downarrow 1,2}]_{\downarrow 2,1}, \tag{13}$$

$$D_{n3} = [g_l \star [g_c \star A_{n-1}]_{\downarrow 1,2}]_{\downarrow 2,1}, \tag{14}$$

where $(i, j) \in \mathbb{R}^2$, $\star$ denotes the convolution operator, $\downarrow 1,2$ ($\downarrow 2,1$) is the sub-sampling operation applied to the columns (lines) and A_0 is the original image I. The low and high-pass filter h and g are the QMF filters referred in the previous section. As shown nicely in Fig. 6, A_n arises from a low-pass filtering operation, being thus an approximated version of the image at scale n. On the other hand, D_{n1} and D_{n2} arise from high-pass filter in a specific direction, namely horizontal and vertical, while low-pass filtering is applied in the other direction. Thus, these coefficients held detail information along a specific direction in the image, at scale n. Lastly, D_{n3} arises from high-pass filtering in both horizontal and vertical directions, possessing thus information regarding details in the diagonal direction. This behaviour can be easily observed when applying the DWT transform to a square binary image, as shown in Fig. 6.

From the previous image, another important characteristic of DWT is evident: it provides a sparse representation of the information present in image. This has been a feature widely used in compression schemes, being perhaps the JPEG2000 the best know example.

Fig. 6. Example on the directional sensitivity of the 2D Discrete Wavelet Transform (right: original image, center: DWT coefficients up to j=4, left: correspondence between the DWT coefficients and the equations (11))-(14)).

Note that, as for the 1D case, the original image can still be recovered from its DWT coefficients, trough inverse transform. Therefore, the DWT can be manipulated in order to enhance features, which can be then synthesized to a new image using IDWT .

2.2 Curvelet transform

The multi-resolution capability of the DWT has been vastly explored in several fields of signal and image processing, as seen in the last section. The ability of dealing with singularities is another important advantage of the DWT, since wavelets provide an optimal representation for one-dimensional piecewise smooth signal (Do & Vetterli, 2005). However, and as seen in the previous section, the application of DWT to multidimensional data is done in a separate way along each dimension. As can be easily understood, natural images are not simply stacks of 1-D piecewise smooth scan-lines, and therefore singularities points are usually located along smooth curves rather than having an independent location. Indeed, these intermediate dimensional structures like discontinuities along curves often provided structures of interest and relevant information within the image. However, and being the DWT directional sensitivity limited to three directions, there is the need to develop mathematical tools to overcome this limitation. Thus, the limitations of the DWT triggered the quest for new concepts capable of overcome these limits. The alternatives proposed until now make use of a combination between concepts from traditional multi-scale analysis and ideas taken from geometry.

Given the focus of the present book being the DWT and its applications, the description of multi-directional and multi-scale transforms will be kept brief and mostly conceptual. The first attempts to extract multi-scale information at different orientations date from the beginning of the nineties, with the introduction of steerable pyramids. This was the first approach to this problem, being a practical, data-friendly strategy to extract information at different scales and angles. More recently, more advance techniques have been proposed, such as the Curvelet Transform (Candès & Donoho, 2000) and the Contourlet Transform (Do & Vetterli, 2005). These new and promising image analysis tools are already starting to prove its usefulness.

The Continuous Curvelet Transform has been introduced by Candès and Donoho and has as key concept the anisotropic notion of scale and high directional sensitivity in multiple directions (Candès & Donoho, 2000). Thus, and contrarily to DWT, the orientational sensitivity is not limited to the horizontal, vertical and diagonal directions but rather span

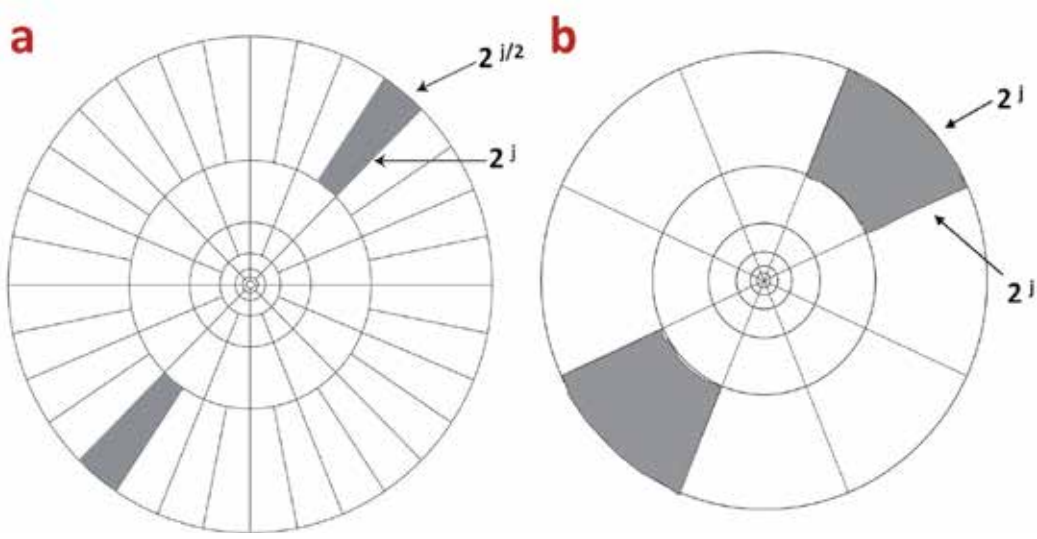


Fig. 7. Tilling of the 2D Fourier polar plane for the Curvelet Transform (a) and the 2D Discrete Wavelet Transform (b). Note that the radius stands for the frequency, where higher radius will encode higher frequency information, and the angle for the orientation.

a wider set of orientations, which stems from the anisotropic notion of scale (parabolic rescaling). In fact, the CCT is based in the tilling of the 2D Fourier space (which has both scale and orientation information) in different concentric coronae, one of each divided in a given number of angles, accordingly with a fixed relation shown in Fig. 7. As it can be observed, there is still a division of the scale information (the center of the 2D Fourier plane corresponds to low frequencies, while the outer regions to higher frequency components), being the directional sensitivity dependent on the scale.
Each of these different regions corresponds to a different curvelet. Indeed, these polar wedges can be defined by the superposition of a radial window and an angular window. To each of this polar wedges, a tight frequency window or coronae can be associated in the 2D Fourier space. This frequency window will then correspond to the Fourier transform of a curvelet function $\psi_{j,\theta}$ function. In fact, and for a single scale, all $\psi_{j,\theta}$ may be obtained by rotations and translations of a mother curvelet ψ_j. As for the wavelet case, the curvelets coefficients will then simply arise from the inner product between the image and the rotation/translation of the mother curvelet.
In the seminal work of Candès and Donoho a discretization scheme was also proposed. However, its complexity led to further research, from which arose the conceptually simpler, faster and less redundant second generation of curvelets, proposed in 2006 (Candès et al., 2006). However, it must be noted that the Curvelet Transform is continuous in its nature and will always have to be redesigned in order to be applied to discrete image data. Keeping this in mind, Do and Vetterli have introduced the Contourlet Transform, a discrete mathematical tool having similar multi-resolution and directional sensitivity characteristics (Do & Vetterli, 2005). However, the Contourlet Transform is formulated from a double filter bank approach, as shown in Fig. 8. The first stage corresponds to Laplacian pyramid which separates the image content into different detail levels. The second step is a a directional filter bank which links

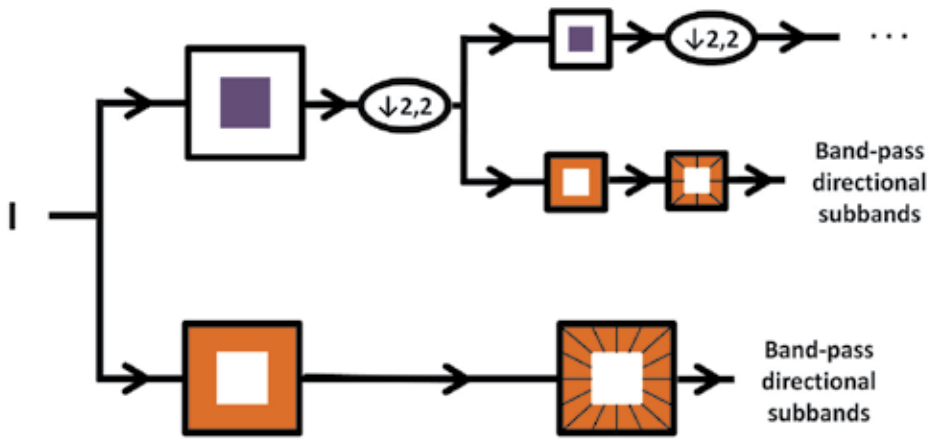


Fig. 8. Two-stage iterated filter-bank for implementation of Contourlet Transform.

point discontinuities into linear structures. Thus, there is multi-scale and multi-directional decomposition of the image content.
Although sharing similar fundamental concepts, Curvelet and Contourlet transforms are not equivalent. Indeed, the theoretical support of Curvelet Transform outranks the Contourlet Transform, in the sense that the curvelet elements have a sharply defined location in the frequency plane, materialized as a polar wedge (in contrast to other approaches with discrete nature). On the practical side, this means that contourlets lack smoothness along the ridge in the spatial domain and exhibit spurious oscillations which may be the source of numerous problems, especially if one wants to use these transforms for scientific computing. On the other hand, the Contourlet Transform is directly designed for discrete applications, whereas the discretization scheme of the curvelet transform faces some intrinsic challenges in the sampling of the Fourier plane in the outermost coronae. Furthermore, the Contourlet Transform presents less redundancy and has a lower computational burden.
As a final note, Curvelet coefficients are, as in the DWT, an accurate representations of the original image with different detail, given by the different frequency content in each scale, but also with different detail in multiple directions, overcoming the directional limitations of the DWT. This might be well suited for the analysis of complex spatio-frequency patterns as texture.

3. Feature extraction algorithm

3.1 Multiscale image pre-processing

In order to extract different scale information from capsule endoscopic frames, both Discrete Wavelet Transform and Discrete Curvelet Transform (DCT) were applied to CE data. A two level DWT and a three scales (including the coarsest) DCT were computed for each color channel of the CE video frames, leading to coarsest, medium and finest detail coefficients for both domains. Note that in the coarsest detail level there is no directional information for both transforms. On the other hand higher detail levels possess directional information. The DWT has 3 sub-bands of different directional information, while the DCT has a number of subbands dependent on the parameters used. In the present case, there were 8 and 16 sub-bands of different directional information in the medium and finest detail scales of the

Fig. 9. Example of a capsule endoscopic frame (center) and its DWT coefficients (left) and DCT coefficients (right). Note that the number of coefficients in the DCT image is not equivalent to the other two images.

DCT respectively. Nonetheless, a note should be made regarding the number of coefficients arising from these two transforms. While the DWT leads to a compact representation, with a number of coefficients equal to the number of pixels in the image under analysis, the DCT results in a higher number of coefficients. For illustrative purposes, the DWT and the DCT coefficients of a CE frame are show in Fig.9. Note that the DCT is far less compact.
Color transformations of the original image I result in three decomposed color channels:

$$I^i, \quad i = 1,2,3, \tag{15}$$

where i stands for the color channel.
A two level DWT is applied to each color channel, I^i, as shown in Fig. 9. This transformation results in a new representation of the original image by a low resolution image and the detail images, possessing higher frequency content in three orientations. Therefore the new representation is defined as:

$$\gamma^i_{DWT} = DWT(I^i) = \{A^i_n, D^i_{s,d}\}, \quad n=2\,,\ s=1,2\,,\ d=1,2,3 \tag{16}$$

where d stands for the wavelet directional sub-band, s for scale and n is the decomposition level.
A three scale DCT is applied to each color channel, I^i, as shown in Fig. 9. This transformation results in a new representation of the original image by a low resolution image and the detail images, possessing higher frequency content in multiple orientations. Therefore the new representation is defined as:

$$\gamma^i_{DCT} = DCT(I^i) = \{A^i_n, D^i_{s,d}\}, \quad n=3\,,\ s=1,2\,,\ d=1,\cdots,8 \ \vee \ d=1,\cdots,16 \tag{17}$$

where d stands for the curvelet directional sub-band, s for scale and n is the decomposition level. Note that, contrarily to DWT, the number of directional sub-bands at each scale depends on the scale itself. Thus, for the first scale there are 16 directions, while for the second there are 8 directions.
After the application of DWT and DCT transform to the image data, its content is divided in different multiscale and multi-directional information. It should be stressed that the

textural information is usually better presented in the middle scale coefficients channels. Thus, mid scale coefficients should be considered in theory. However, the relatively low image dimensions (256×256) limit the representation of the details, becoming the first level more adequate, than previously expected, for texture representation Barbosa et al. (2008). Nevertheless an image resolution normalization might be required for different image resolutions.

3.2 Statistical texture features

There are several statistical features that can be extracted from the wavelet and curvelet domains as texture descriptors, being the most common the mean, the standard deviation, the energy and the entropy of each DWT/DCT sub-band (μ, σ, E, Ent) (Dettori & Semler, 2007). Nonetheless, in authors' previous work it was observed that the introduction of energy and entropy in the feature set did not significantly contributed to an increase in the classification performance (Barbosa, Correia, Ramos & Lima, 2009; Barbosa, Ramos, Correia & Lima, 2009). Thus, only the mean and standard deviation in each DWT/DCT sub-band were computed. The proposed texture descriptors can be calculated as:

$$\mu = E\{P(i,j)\} = \frac{1}{N}\sum_i\sum_j P(i,j) \tag{18}$$

$$\sigma = \sqrt{E\{(P(i,j)-\mu)^2\}} = \sqrt{\frac{1}{N}\sum_i\sum_j (P(i,j)-\mu)^2} \tag{19}$$

where $P(i,j)$ corresponds to the pixel value at position (i,j). Note that capsule endoscopic video frames are usually square images of 256x256 but the information is restricted to a circular area in the middle of the image, as it is observable in Fig.9. Therefore, it is vital to only consider the pixels inside this area, since the information regarding to the CE exam is contained in this part. In order to cope with this constraint, the summation limits of the equations (18)-(19) were set in order to correspond to this area.

In the present work, it was decided to start the comparison between the different DWT and DCT detail levels only with the mean and variance as statistical descriptors, in order to better compare the two different multi-resolution domains. Since the low frequency components of the images do not contain major texture information, the most important scales in the DWT and DCT will be those in which are present medium and high frequency, texture encoding, information. Furthermore, the coarsest scale coefficients of the DCT and DWT are not directional, and consequently do not possess directional sensitivity. Therefore, each of the afore mentioned statistical features is computed at each sub-band of the DWT/DCT coefficients, for all the color channels, at medium and finest detail level.

The statistical dependence of textural descriptors taken in different color channels is useful to distinguish normal from abnormal texture patterns, as stated in authors' previous work. Furthermore, the same finding was previously reported in (Karkanis et al., 2003), for colonoscopy videos. Therefore, the covariance of textural descriptors in the different color channels will be used as a classification feature. Note that in the present framework, the high directional sensitivity of the DCT will be likely to lead to more robust descriptors than a similar scheme used in textural descriptors taken from DWT coefficients. This was recently suggested by the authors and shown in (Barbosa, Ramos, Correia & Lima, 2009). In the present chapter, this methodology is followed and the scope of the comparison of this recently proposed algorithm and other existing features is expanded. In the Color Curvelet Covariance

framework the covariance of textural descriptors in the different color channels is used as a classification feature. The Color Wavelet/Curvelet Covariance of a texture descriptor can be calculated as:

$$C{*}C(a,b,s,m) = \sum_{\alpha} ((F_m(a,s,\alpha) - E\{F_m(a,s,\alpha)\}) \times (F_m(b,s,\alpha) - E\{F_m(b,s,\alpha)\})) \quad (20)$$

where a and b represent the different color channels in the covariance calculation, F_m is the statistic textural descriptor, α is the considered angle of the DWT/DCT coefficients, s the considered detail scale and $E\{Fm(a,s,\alpha)\}$ the average of the statistical textural descriptor F_m over the different angles α, in the color channel a. As it is clear from Fig.9, the Color Wavelet Covariance features will arise from a covariance of only three different angles, while the Color Curvelet Covariance features will take more angles into account in this computation. One should note that if $a = b$, the $C{*}C$ features will enconde the angular variance of the textural descriptor F_m. On the other hand, for $a \neq b$, the $C{*}C$ features will give a measure of the similarity of the angular covariation of F_m between two color channels.
The proposed $C{*}C$ features allow to capture additional information regarding the texture patterns and can be regarded as second order statistical modelling of the texture descriptors F_m. However, the texture descriptors F_m already possess information regarding the existing textures in the image. Thus, it is proposed to add the average value $E\{F_m\}$ to the feature set. It should be referred that $E\{F_m\}$ corresponds to the mathematical expectancy for the value of F_m and therefore does not possess as much information as the sequence of the different F_m values taken from the DCT/DWT coefficients at different angles. Nonetheless, the inclusion of the entire F_m feature set helds dimensionality problems, which can compromise the training and consequently the performance of the classifier and the speed of the classification step. Furthermore, if the same texture pattern was rotated, the sequence of the different F_m would be affected, which could decrease the classification performance.

4. Implementation issues

4.1 General considerations

A 2.8 GHz Intel i7 dual core processor, with 4 GB of RAM, was used with MATLAB to run the proposed algorithm. The average processing time is between 0.2s (DWT) and 0.75s (DCT) for each CE frame. Note however that the implementation of the proposed algorithm was not optimized for speed, so the processing time can still be improved.
The DWT calculation was done using the MATLAB Toolbox Wavelets (made available by G. Peyré at http://www.ceremade.dauphine.fr/~peyre/matlab/wavelets/content.html). The basis used was the standard Daubechies wavelet with 4 vanishing moments.
The DCT calculation was done with the routines implemented in the toolbox CurveLab (available for research purposes at www.curvelet.org). The method chosen for the discretization of the Curvelet Transform was the wrapping algorithm, for the reasons referred in (Candès et al., 2006).
For both DWT and DCT, the reader is highly advised to explore the tutorial and demos offered in the used toolboxes in order to better understand the underlying principles of these mathematical tools.
The selected color space was the HSV, since it is more similar to the physiological perception of human eye (Li & Meng, 2009), and therefore more adequate than the standard RGB color space.

4.2 Feature sets

In order to assess the impact of the choice of texture descriptors, several feature sets were generated for the the analyzed data. The feature set *A* corresponds to a feature set comprising the mean and standard deviation of each DWT/DCT sub-band. A more advanced feature set is also formulated by using the covariance of the textural descriptors, as well as their average value for all the analyzed sub-bands. The covariance+average approach is applied to the texture descriptors mean and standard deviation and results in the feature set *B*. These different feature sets were computed for both the medium and finest detail scales of the DWT/DCT transforms, since it is known from before hand that low detail coefficients do not possess relevant texture information.

4.3 Classification scheme

The features were imported into the open source machine learning package WEKA (available at http://www.cs.waikato.ac.nz/ml/weka/). A stratified 10-fold cross-validation procedure was chosen to train a standard multilayer perceptron neural network. The default parameters were kept in the classifier options. The choice of a simple classification scheme, with default parameters, was done in order to make the results more representative of the choice of the features. The 10-fold cross-validation method is a standard procedure to validate machine learning classification outputs and has been found to provide an adequate and accurate estimate of the true error rate (Alpaydin, 2004). The 10-fold cross-validation algorithm splits the data into 10 partitions, where the proportion of both normal and abnormal frames in each partitions is similar to the entire dataset. The training and classification process is then repeated 10 times, where 9 partitions are used to train and 1 partition is used to assess the classification process. This way, each frame will be used exactly once as test data, allowing to efficiently use the available dataset. In order to have an accurate error estimate, the cross-validation process was repeated 10 times, being this a standard procedure (Witten & Frank, 2005). Note that the splitting of the training and testing datasets in the 10 repetitions of the 10-fold stratified cross-validation were similar for the different feature sets, allowing thus a paired comparison.

4.4 Evaluation of classification performance

The ultimate goal of the proposed method is to provide a binary classification of whether a given CE frame presents patterns of abnormal tissue. Thus, it is vital to understand how such classification performance is assessed. Confusion matrices offer a simple yet effective way of visualizing the classification errors:

		Outcome of classification test	
		Normal	Tumor
Ground Truth	Normal	TN	FP
	Tumor	FN	TP

A false negative, *FN* corresponds to a CE frame presenting tumor pathology which is misclassified as normal, while a false positive, *FP*, corresponds to a normal frame wrongly

considered as abnormal. While the overall accuracy can be defined as:

$$Accuracy = \frac{TP + TF}{TP + TF + FP + FN}, \tag{21}$$

it is more frequent to report the performance of a binary classification test using sensitivity and specificity:

$$Sensitivity = \frac{TP}{TP + FN}, \tag{22}$$

$$Specificity = \frac{TN}{TN + FP}. \tag{23}$$

There is an intrinsic trade-off between sensitivity and specificity of a classifier, which can be adapted to each situation in order to penalyze a specific type of error. This can be nicely visulized through Receiver Operating Characteristic (ROC) curves. Although no single number is able to capture this trade-off, being thus such 2D representation needed, the area under the ROC curve is usually used as a measure of the classifier performance (Witten & Frank, 2005).

Lastly, it was decided to include the Cohen's kappa coefficient, k, in the set of classification evaluation measures, since it is a statistical measure of agreement between the classifier output and theground truth. Furthermore, k takes into account the agreement that may occur simply by chance, being thus a better performance metric for the classification performance. It can be simply calculated as:

$$K = \frac{P(a) - P(e)}{1 - P(e)}, \tag{24}$$

where $P(a)$ is the relative observed agreement and $P(e)$ the probability of agreement due to random chance. Note that relative agreement is not more than simply the classification accuracy.

In order to assess the statistical significance of differences in the classification, a standard paired t-test was employed.

5. Experiments and results

5.1 Dataset

The experimental dataset was constructed with frames from capsule endoscopic video segments of different patients' exams, taken at the Hospital dos Capuchos in Lisbon by Doctor Jaime Ramos. The final dataset consisted in 400 normal frames and 196 abnormal frames. Examples of the dataset frames can be observed in the Fig.10. These frames have been manually classified by the expert physician.

5.2 Influence of the scale on the classification Performance

From authors' previous work, it is highly expectable that the most relevant information for classification purposes is encoded as high frequency content in the scale correspondent to the highest detail level. Table 1 shows that the most relevant information for classification purposes is encoded as high frequency content in both DWT and DCT finest detail coefficients. Note that M and F stands for medium and finest detail scales. For comparison purposes, the feature sets A was chosen. The values in Table 1 correspond to the mean and standard deviation for a ten-fold stratified cross-validation scheme repeated 10 times.

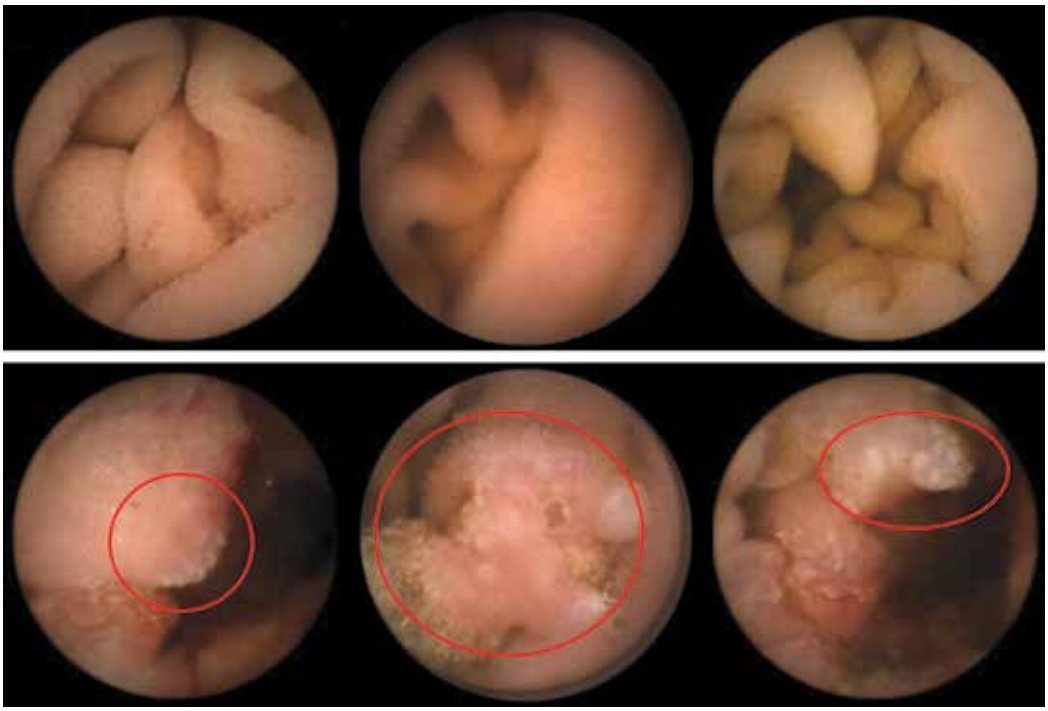


Fig. 10. Examples of Wireless Capsule Endoscopy video frames (top: normal frames, bottom: frames containing tumoral tissue).

Transform	**DWT**		**DCT**	
Feature Set	**A**		**A**	
Detail Level	**M**	**F**	**M**	**F**
Features (#)	18	18	48	96
Accuracy (%)	77.15 ± 1.56	96.34 ± 0.42	87.44 ± 0.96	95.54 ± 0.31
Specificity (%)	84.78 ± 3.23	97.34 ± 0.29	90.61 ± 2.29	96.76 ± 0.37
Sensitivity (%)	61.64 ± 4.60	94.29 ± 1.10	81.01 ± 3.39	93.13 ± 0.62
ROC Area	0.834 ± 0.012	0.990 ± 0.003	0.936 ± 0.008	0.989 ± 0.002
Cohen's K	0.473 ± 0.032	0.917 ± 0.010	0.716 ± 0.020	0.899 ± 0.007

Table 1. Influence of the Detail Level in the Classification Performance

From the results in the previous table, there is no clear advantage arising from the use of DCT, although better results are observed for the medium detail coefficients. However, the large difference between the number of features extracted does not allow a strong conclusion in this comparison, since it is known that large feature vectors pose difficulties to the training algorithm. Thus, the key observation from the results presented in Table 1 is that for both DWT and DCT domains, the features extracted from the sub-band corresponding to the higher detail yield indeed a better classification performance.

Transform (Detail Level)	DWT (F)		DCT (F)	
Feature Set	**A**	**B**	**A**	**B**
Features (#)	18	18	96	18
Accuracy (%)	96.34 ± 0.42	96.30 ± 0.45	95.54 ± 0.31	97.65 ± 0.55
Specificity (%)	97.34 ± 0.29	97.45 ± 0.65	96.76 ± 0.37	98.47 ± 0.58
Sensitivity (%)	94.29 ± 1.10	94.04 ± 1.40	93.13 ± 0.62	96.0 ± 0.61
ROC Area	0.990 ± 0.003	0.990 ± 0.003	0.989 ± 0.002	0.994 ± 0.003
Cohen's K	0.917 ± 0.010	0.916 ± 0.010	0.899 ± 0.007	0.947 ± 0.012

Table 2. Color Covariance Texture Descriptors Classification Performance

5.3 Color covariance texture descriptors

In the present section, the added value for the color covariance texture descriptors is shown. Given the results presented in the previous section, this approach was applied only to the finest detail coefficients of both DWT and DCT transforms, as these possess the most relevant texture information for classification purposes. The results are shown in the Table 2. The values correspond to a ten-fold stratified cross-validation scheme repeated 10 times.

Using the Color Covariance approach, a visible increase in the classification performance of the features extracted from the DCT domain was observed. Nonetheless, in the performance of DWT features there were no significant changes. These results support the hypothesis that $C*C$ features can extract complex texture patterns between color channels, leading to good texture classificantion performance. The Color Curvelet Covariance approach is able to significantly reduce the feature set dimensionality, while improving its classification performance. These results also point that features extracted from the DCT domain possess more relevant information for classification purposes when compared to the DWT. This can be intrinsically linked with the higher directional sensitivity of the DCT, which helps to better handle complex two-dimensional patterns such as texture.

5.4 Statistical analysis of the classification performance

Given the results shown in subsection 5.2, the statistical analysis was limited to the features extracted from the highes detail coefficients of both DWT and DCT domains. The significance of the statistical differences found between the different feature sets extracted from both domains can be observed in the Table 3, for the metrics used to evaluate the classification performance. For sake of clarity, NS stands for not statistically significantly different, † for $p < 0.05$ and ‡ for $p < 0.001$. Note also that $DWT(F, A)$ stands for the feature set A extracted from the finest detail level of the DWT transform, while $DCT(F, B)$ corresponds to the feature set B extracted from the DCT coefficients corresponding to the finest detail level.

From these results, some important highlights appear, confirming statistically some previous observations. First and foremost, the proposed Color Curvelet Covariance approach presents significantly higher classification performance for all the metrics used to compare the different methods. Second, the $C*C$ approach in the DWT domain does not lead to a significant difference in the classification performance. Last, the feature vector A extracted from the DCT scale of finest details helds a much higher dimensionality than the remaining ones. Thus, no strong conclusions can be taken from the statistical comparison with the remaining feature sets, since the high dimensionality of this feature set may pose a problem to the performance of the classifier.

Accuracy

Transform(Detail Level,Feature Set)	DWT(F,A)	DWT(F,B)	DCT(F,A)	DCT(F,B)
DWT(F,A)		NS	‡	‡
DWT(F,B)			‡	‡
DCT(F,A)				‡
DCT(F,B)				

Specificity

Transform(Detail Level,Feature Set)	DWT(F,A)	DWT(F,B)	DCT(F,A)	DCT(F,B)
DWT(F,A)		NS	†	‡
DWT(F,B)			†	‡
DCT(F,A)				‡
DCT(F,B)				

Sensitivity

Transform(Detail Level,Feature Set)	DWT(F,A)	DWT(F,B)	DCT(F,A)	DCT(F,B)
DWT(F,A)		NS	†	‡
DWT(F,B)			†	‡
DCT(F,A)				‡
DCT(F,B)				

ROC Area

Transform(Detail Level,Feature Set)	DWT(F,A)	DWT(F,B)	DCT(F,A)	DCT(F,B)
DWT(F,A)		NS	†	‡
DWT(F,B)			†	‡
DCT(F,A)				‡
DCT(F,B)				

K

Transform(Detail Level,Feature Set)	DWT(F,A)	DWT(F,B)	DCT(F,A)	DCT(F,B)
DWT(F,A)		NS	†	‡
DWT(F,B)			†	‡
DCT(F,A)				‡
DCT(F,B)				

Table 3. Statistical Significance of the Differences for the Classification Accuracy

6. Conclusion and future work

The more significant information content for classification purposes is encoded as high frequency information. Indeed, texture corresponds to medium/high frequency image content. However, and contrarily to what has been found in (Karkanis et al., 2003) for colonoscopy videos, the most relevant sub-bands for texture classification correspond to the highest detail levels, for both DWT and DCT domains.

The proposed color covariance approach achieves an optimal balance between feature vector size and classification performance, presenting promising results that support the feasibility of the proposed method. The Color Curvelet Covariance method yielded the higher classification accuracy, achieving $97.65 \pm 0.55\%$ of correctly classified frames.

Future work will include the extraction of different texture descriptors from the DCT and the use of different classifiers. The enlargement of the available dataset will be equally pursued, in order to further validate the observed results over a wider range of images and pathologies.

7. References

Alpaydin, E. (2004). *Introduction to machine learning*, MIT Press.

Barbosa, D., Correia, J. H., Ramos, J. & Lima, C. S. (2009). A multi-scale comparison of texture descriptors extracted from the wavelet and curvelet domains for small bowel tumor detection in capsule endoscopy exams, *Proc. World Congress on Medical Physics and Biomedical Engineering 2009*, pp. 1546–1549.

Barbosa, D., Ramos, J., Correia, J. H. & Lima, C. (2009). Automatic detection of small bowel tumors in capsule endoscopy based in color curvelet covariance statistical texture descriptors, *Proc. 31th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC2009)*, pp. 6683 – 6686.

Barbosa, D., Ramos, J. & Lima, C. S. (2008). Detection of small bowel tumors in capsule endoscopy frames using texture analysis based on the discrete wavelet transform, *Proc. 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC2008)*, pp. 3012 –3015.

Candès, E., Demanet, L., Donoho, D. & Ying, L. (2006). Fast discrete curvelet transforms, *Multiscale Modeling and Simulation* 5(3): 861–899.

Candès, E. & Donoho, D. (2000). Curvelets, multi-resolution representation, and scaling laws, *Wavelet Applications in Signal and Image Processing VIII, SPIE vol. 4119* pp. 1–12.

Chen, X., Ran, Z. & Tong, J. (2007). A meta-analysis of the yield of capsule endoscopy compared to double-balloon enteroscopy in patients with small bowel diseases, *World Journal of Gastroenterology* 13(32): 4372–4378.

Daubechies, I. (1988). Orthogonal bases of compactly supported wavelets, *Communications on Pure and Applied Mathematics* 41: 909 – 996.

Daubechies, I. & Sweldens, W. (1998). Factoring wavelet transforms into lifting steps, *Journal of Fourier Analysis and Applications* 4(3): 247–269.

Delvaux, M. & Gay, G. (2006). Capsule endoscopy in 2005: Facts and perspectives, *Best Practice & Research Clinical Gastroenterology* 20(1): 23 – 39.

Dettori, L. & Semler, L. (2007). A comparison of wavelet, ridgelet, and curvelet-based texture classification algorithms in computed tomography, *Computers in Biology and Medicine*, 37(2): 486–498.

Do, M. N. & Vetterli, M. (2005). The contourlet transform: an efficient directional multiresolution image representation, *IEEE Transactions on Image Processing* 14(12): 2091–2106.

Gay, G., Pennazio, M., Delmotte, J. & Rossini, F. (1998). *Atlas of Enteroscopy*, Springer, chapter Intraoperative enteroscopy, p. 51 – 54.

Herrerías, J. & Mascarenhas-Saraiva, M. (2007). *Atlas of Capsule Endoscopy*, Sulime Diseño de Soluciones, Sevilla.

Iddan, G., Meron, G., Glukhovsky, A. & Swain, P. (2000). Wireless capsule endoscopy, *Nature* 405(6785): 417–417.

Kaffes, A. (2009). Achieving total enteroscopy with capsule endoscopy in all patients: are we stretching the limits of technology?, *Gastrointestinal Endoscopy* 69(1): 81 – 83.

Karargyris, A. & Bourbakis, N. (2010). Wireless capsule endoscopy and endoscopic imaging: A survey on various methodologies presented, *IEEE Engineering in Medicine and Biology Magazine* 29(1): 72 –83.

Karkanis, S., Iakovidis, D., Maroulis, D., Karras, D. & Tzivras, M. (2003). Computer-aided tumor detection in endoscopic video using color wavelet features, *IEEE Transactions on Information Technology in Biomedicine* 7(3): 141–152.

Kodogiannis, V., Boulougoura, M., Wadge, E. & Lygouras, J. (2007). The usage of soft-computing methodologies in interpreting capsule endoscopy, *Engineering Applications of Artificial Intelligence* 20(4): 539–553.

Lee, N. & Eisen, G. (2010). 10 years of capsule endoscopy: An update, *Expert Reviews Gastroenteroly and Hepatoly* 4(4): 503–512.

Li, B. & Meng, M. (2009). Computer-based detection of bleeding and ulcer in wireless capsule endoscopy images by chromaticity moments, *Computers in Biology and Medicine* 39(2): 141–147.

Mackiewicz, M., Berens, J. & Fisher, M. (2008). Wireless capsule endoscopy color video segmentation, *IEEE Transactions on Medical Imaging* 27(12): 1769 –1781.

Mallat, S. (1989). A theory for multiresolution signal decomposition: The wavelet representatio, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 11(7): 674–693.

Mergener, K., Ponchon, T., Gralnek, I., Pennazio, M., Gay, G., Selby, W., Seidman, E. G., Cellier, C., Murray, J., de Franchis, R., Rösch, T. & Lewis, B. S. (2007). Literature review and recommendations for clinical application of small-bowel capsule endoscopy, based on a panel discussion by international experts, *Endoscopy* 39(10): 895–909.

Munoz-Navas, M. (2009). Capsule endoscopy, *World Journal of Gastroenterology* 15: 1584–1586.

Pennazio, M. (2006). Capsule endoscopy: Where are we after 6 years of clinical use?, *Digestive and Liver Disease* 38(12): 867 – 878.

Pennazio, M., Arrigoni, A., Risio, M., Spandre, M. & Rossini, F. (1995). Clinical evaluation of push-type enteroscopy, *Endoscopy* 27(2): 164–70.

Swain, P. & Fritscher-Ravens, A. (2004). Role of video endoscopy in managing small bowel disease, *Gut* 53(12): 1866–1875.

Unser, M., Aldroubi, A. & Laine, A. (2003). Guest editorial: Wavelets in medical imaging, *IEEE Transactions On Medical Imaging* 22(3): 285–288.

Westerhof, J., Koornstra, J. & Weersma, R. (2009). Can we reduce capsule endoscopy reading times?, *Gastrointestinal Endoscopy* 69(3, Part 1): 497 – 502.

Witten, I. H. & Frank, E. (2005). *Data Mining: Practical machine learning tools and techniques*, Morgan Kaufmann.

Yamamoto, H. & Kita, H. (2006). Double-balloon endoscopy: From concept to reality, *Gastrointestinal Endoscopy Clinics of North America* 16(2): 347 – 361. Small Bowel Enteroscopy.

10

Wavelet Transform for Electronic Nose Signal Analysis

Cosimo Distante[1], Marco Leo[1] and Krishna C. Persaud[2]
[1]*Consiglio Nazionale delle Ricerche – CNR*
[2]*University of Manchester*
[1]*Italy*
[2]*United Kingdom*

1. Introduction

Semiconductor gas sensors have been studied for many years and they are now used in many fields of application. Despite this wide trade further research needs mainly to improve sensitivity, selectivity and stability. In fact, in this sense, improving both, the sensor selectivity toward specific gaseous substances and discrimination capability, has been the goal of a great deal of work over the last few years. One strategy consists of using non selective sensor arrays and an appropriate pattern recognition system capable of recognizing simple or complex vapors based on the conductance or current at the saturation point of each transient response. We believe that most useful information (high and low frequencies in the curve) can be taken from the transient response. One of the most important processing part of an intelligent system, is its ability to extract useful information less redundant than the original one to aid fast processing and pattern classification. In other words: any selected feature (1) must discriminate clearly between two or more classes of objects, (2) must not be correlated with another feature to any moderate strong extent, and (3) should have meaning for humans. The first step toward the pre-processing of e-nose data was based on methods for extracting information of the transient only from the steady-state and baseline response values. However, these methods takes into account only stationary information about the transient (i.e. steady-state and baseline) but all the information related to the kinetic of the rise and recovery time is lost. Each sensor has its own behavior in response to an odor presentation that is stored along the response. For example, starting from the fact that light molecules (i.e. alcohol) get more dispersed than heavier one (i.e. acid oils), it is possible to analyse the recovery time of the response due to the end of the exposure to recognize the odor being presented. In some cases of long time sensor response (of the order of several minutes), it is possible to analyse the rise time of the transient response that can be computed in a few seconds instead of considering the saturation parameters.

The common problem encountered with electronic nose applications is the non-reproducibility of sensor responses over time due to drift effects. Drift is a slow change in sensitivity that occur in time due to ageing effects, slow morphological changes in the sensor material, poising and other long- term effects.

The instability of sensor responses could be connected to morphological evolution of the sensing material, so that it is needed to age the sensors for several weeks in the laboratory

bench test before their use. However long term drift is more associated with the contamination of the sensor material due to presence of atmospheric pollution that irreversibly react on sensor surface leading to a reduction of sensitivity.
Pijolat et al. (2003) proved that the presence of a very low concentration of SO_2 in environment can be at the origin of the tin dioxide sensor drift. In this case the problem was overcome by pre-treated the sensors in SO_2 before their use. Such a gaseous treatment had been used in the past to improve the stability of sintered tin dioxide sensor, especially with respect to effect of humidity. This type of treatment can be used with other gases if there is a clear view of the mechanism of interaction between the sensing layer and the gases. Unfortunately all processes concerning drift are not yet well known and the sensor instability depends upon time-varying parameters such as the historical use of the sensors, interactions with several volatile compounds present in the environment at various concentrations, the exposure period to certain substances. Actually a great deal of work has been directed towards the development of methods to minimize the drift effects Sisk & Lewis (2005). An efficient approach could be to use a continuous re-calibration for the training phase of the pattern recognition model, but in the real applications the use of sets of samples for the calibration is very expensive and time-consuming, which makes it unlikely to re-calibrate the pattern recognition model very often. Attempts for drift minimizing have been made by using a reference gas as a reference value and then correcting all subsequent readings accordingly Haugen et al. (2000). Component correction is a method using one or more reference gases. This linear method is based on PCA and PLS algorithms and removes the drift direction calculated from measurements of a reference gas Artursson et al. (2000). The technique use a linear model and has been used at feature level. Multiple Self-Organizing Maps (MSOM) has been developed in Distante, Sicilian & Persaud (2002) because new recognized data that match the stored recognized odors can be continuously used to retrain the classifier. This technique has the advantage of having self-recalibration mechanisms without the intervention of a user. However, the drift must be gradual, as a discontinuity in response between consecutive exposures (regardless of the time interval between the exposures) would immediately invalidate the classification model and would prevent adaptation Sisk & Lewis (2005). Others drift counteracting methods focus on the application of signal pre-processing techniques to filter out portions of the signal containing drift contaminations. Drift is in general a time scale longer than the duration of a single measurement, thus a selection of the lowest frequency components of the sensor responses are needed to filter it out.
A moving median filter and Fourier band-pass filters are some examples of filters applied to removing either high-frequency fluctuations (such as noise, spikes) or low-frequency changes such as drift. In comparison to these filters, Discrete Wavelet Transform (DWT) technique provides a flexible analysis of the signal at different resolutions by applying iteratively high-pass and low-pass filters Bakshi (1999); **?**. Therefore the wavelet transform is a powerful tool to point out drift contamination in the low-frequency behavior of the sensor responses. As it is shown in this paper, this technique allows to remove the selected low-frequency components easily and in such a way that the signal is not distorted. The comparison of the DWT filter and others commonly used have been compared with the relative method feature extraction technique which is commonly used in this context. Principal component analysis shows than that DWT approach is superior in terms of cluster dispersion. The DWT is here also ebmedded into MSOM neural network Distante, Sicilian & Persaud (2002); Zuppa et al. (2007) as it is able to detect and adapts to clusters trend Zuppa et al. (2004). The use of

smoothing filters as the Savitzky-Golay and Discrete Wavelet Trasform on this function allows to cut off the frequency components which do not influence its behaviour.

2. Related works

There has been a lot of work in using discrete wavelet trasform for electronic nose data since our initial paper in Distante, Leo, Siciliano & Persaud (2002). Since the interesting information resides in low frequency components as will be shown below, all the referenced works of DWT for electronic noses make use of the approximation coefficients which are the results of the low-pass filter with the analysing signal. In this case, approximation coefficients allows for signal representation, where there should be made distinction when extracting features for classification or compression tasks. Both tasks demand for signal simplification Rubinstein et al. (2010), but in Leone et al. (2005) it has been shown that for classification purposes, the fewer the number of decriptors, the better the accuracy. In Phaisangittisagul & Nagle (2008) features are extracted from the transient with discrete wavelet transform (DWT) in order to search for an optimal sensor array to be implemented in the e-nose system. Two different odor datasets such as coffee and soda are collected and a genetic algorithm is adapted to tailor a gas sensor array. Two different mother wavelets were used: Haar and Daubechies of second order ($db2$). Several decomposition levels were investigated, Haar third level and $db2$ at third fourth and fifth. As expected, Haar provide lower classification performance with respect to $db2$. The fifth decomposition level (approximation coefficients) has been selected and given to the $k - Nearest - Neighbor$ classifier. In Phaisangittisagul & Nagle (2010) a signal decomposition/reconstruction based on the discrete wavelet transform is proposed, whose coefficients are given to a support vector regression to predict a sensor's response to mixtures of odors. Prediction is performed on different mixing ratios of Regular, Sumatra coffees and Grean tea components. The idea is to predict sensor response by starting from the mixing ratios. Wavelet coefficients (both approximationa and details) have been used to train support vector regression (SVR) machines to predict the respective coefficients at lower level going backward for inverse discrete wavelet trasform to reconstruct the signal. So far, for each decomposition level, two SVR's (one for approximation coefficients and one for details) are trained. SVR parameters have been found using an optimization technique based on genetic algorithm. The considered sensor for the response prediction study is a MOSFET 101A sensor of the NST 3320 e-nose. The highest decomposition level must be manually chosen, where the 5th level has shown to be the most appropriate for the probelm at hand. Infact, choosing lower scales brings to the overfitting problem by loosing generalization capabilities, while going to higher scales increase the predicted error since it accumulates with the number of decomposition levels along the reconstruction process with IDWT. A study of the number of extracted features for the classification of several volatiles have been investigated in Acevedo et al. (2007). They show the number of coefficients needed to reach optimal classification accuracy with techniques such as: DWT, DCT, PCA and linear discriminant analysis (LDA).
The ability of the DWT to recover sensor signals subjected to *drift* effects has been addressed in Huang & Leung (2009); Zuppa et al. (2007). The drift, such as the seasonal fluctuation, resides in low frequencies. Wavelet analysis is used, in order to decompose the drifting signal at the greatest scale value to reveal signal trend. The trend is the slowest part of the signal and as the scale increases a better estimate of the unknown trend is obtained. So far, DWT is an efficient tool for pre-processing drifting sensor responses as it allows to select and discard signal components, where drift contamination is present, without distorting the signal by

excessive cutting off low-frequency components. The results are compared with ones obtained by applying usual high-pass filters.
The use of DWT for tin dioxide gas sensor has also been investigated on a robotic platform Trincavelli et al. (2009) in a continuos sampling setup. The plume to be tracked can be generated by ethanol, acetone or isopropyl. A feature vector composed on a combination of DWT, curve fitting and discrete Fourier Transform coefficients have been considered and provided to an SVM classifier. Similarly, in Loutfi et al. (2009) a mobile platform is used to integrate classification of odours with gas distribution mapping. The resulting odour map is then correlated with the spatial information collected from a laser range scanner to form a combined map. Two electronic noses are present onboard based on TGS Figaro technology. Each e-nose consists of four TGS sensors (TGS 2600 (ÃŮ2), 2620, 2602). A temporal analysing window of 20-30 seconds of measurements are considered for the classification task. Signals goes DWT analysis whose coefficients are projected onto the first two proncipal components (PCA). Then SVM is used to classify between ethanol, acetone and air.

3. Wavelet analysis

Wavelet transform is an extension of Fourier Transform, generalized to any wideband transient. Let us think to our input as a time-varying signal. To analyze signal structure of very different sizes, it is necessary to use time-frequency atoms with different time support. The wavelet transform decomposes signals over dilated and translated wavelets Mallat (1999). The signal may be sampled at discrete wavelength values yielding a spectrum. In continuous wavelet transform the input signal is correlated with an analyzing continuous wavelet. The latter is a function of two parameters such as scale and position. The widely used Fourier transform (FT) maps the input data into a new space, the basis functions of which are sines and cosines. Such basis functions are defined in an infinite space and are periodic, this means that FT is best suited to signal with these same features. The Wavelet transform maps the input signal int a new space, the basis function that are quite localized in space. They are usually of compact support. The term *wavelet* comes from well localized wave-like functions. Infact, they are well localized in space and frequency i.e. their rate of variations is restricted. Fourier transform is not local in space but only in frequency. Furthermore, Fourier analysis is unique, but wavelet not, since there are many possible sets of wavelets which one can choose. Our trade-off between different wavelet sets is compactness versus smoothness.
Working with fixed windows as in the Short Term Fourier Transform (STFT) may bring to problems. Infact, if the signal details are much smaller than the width of the window they can be detected but the transform will not localize them. If the signal details are larger than the window size, then they will not be detected properly. The scale is defined by the width of a modulation function. To solve this problem we must define a transform independent from the scale. This means that the function should not have a fixed scale but should vary. To achieve this, we start from a function $\psi(t)$ as a candidate of a modulation function and we can obtain a family starting from it by varying the scale s as follows:

$$\psi_s(u) = |s|^{-p}\psi(\frac{u}{s}) = \frac{1}{|s|^p}\psi\left(\frac{u}{s}\right) \quad p \geq 0 \quad \forall s \in \Re \quad s \neq 0. \tag{1}$$

If ψ has width T then the width of ψ is sT. In term of frequencies, we can state that small scales s implies ψ_s has high frequencies and increasing s the frequency of ψ_s decreases.

3.1 Continuous wavelet transform

As it is well known, FT uses basis functions consisting of sines and cosines functions. These functions are time-independent. Hence, the description of a signal provided by Fourier analysis is purely in the frequency domain. The windowed Fourier transform and the wavelet transform aims at the analysis of time and frequency. For non-stationary analysis, a windowed Fourier transform or Short Time FT (STFT) is best suited. The smaller the window size the more the number of discrete frequencies that will be reduced, leading to a weak discrimination potential among frequencies. Given a signal $f(t)$, a window g around the time-point τ and frequency ω is

$$STFT(\tau,\omega) = \int_{-\infty}^{+\infty} f(t)g(t-\tau)e^{-j\omega t}dt. \tag{2}$$

Now considering

$$k_{\tau,\omega}(t) = g(t-\tau)e^{-j\omega t} \tag{3}$$

as a new basis and rewriting this with window s inversely proportional to the frequency ω and the position parameter b that replaces τ gives the following

$$k_{b,s}(t) = \frac{1}{\sqrt{s}}\psi^*\left(\frac{t-b}{s}\right) \tag{4}$$

that yields to the continuous wavelet transform (CWT).

While in STFT the basis functions are sinusoids, in the CWT they are scaled versions of the so called mother wavelet ψ (ψ^* represents the conjugate i.e. $\psi_s^*(u) = \frac{1}{\sqrt{s}}\psi(-\frac{u}{s})$. A wavelet mother function can be constructed in several ways, but subjected to admissibility constraints.

Definition The Morlet-Grossman definition of the continuous wavelet transform for a 1-dimensional signal $f(t) \in L^2(\Re)$,the space of all square integrable functions, is given as follows:

$$\tilde{f}(a,b) = \frac{1}{\sqrt{a}}\int_{-\infty}^{+\infty} f(t)\psi^*\left(\frac{t-b}{a}\right)dt \tag{5}$$

where:

- $Wf(a,b)$ is the wavelet coefficient of the function $f(t)$,
- $\psi(t)$ is the mother wavelet,
- $a > 0$ is the scale parameter,
- b is the position parameter.

We can also rewrite eq. 5 as a convolution product:

$$\tilde{f}(a,b) = f \star \psi_s(t). \tag{6}$$

The continuous wavelet transform is the result of the scalar product of the original signal f with the *shifted* and *scaled* version of a prototype analysing function $\psi(t)$ called *mother wavelet* which has the characteristic of a bandpass filter impulse response. The coefficients $\tilde{f}$ of the transformed signal f represent how closely correlated the mother wavelet is with the section of the signal being analyzed. The higher the coefficient is, the more the similarity[1].

The continuous wavelet transform has the following properties:

1. CWT is a linear transformation,
2. CWT is covariant under translation,
3. CWT is covariant under dilation.

[1] Note that the result will depend on the shape of the wavelet you choose.

3.2 Wavelet functions

As it is know, to analyse signals of very different sizes, it is necessary to use time-frequency atoms with different time support. A wavelet $\psi \in L^2(\Re)$ is a function with a zero average

$$\int_{-\infty}^{+\infty} \psi(t)dt = 0, \tag{7}$$

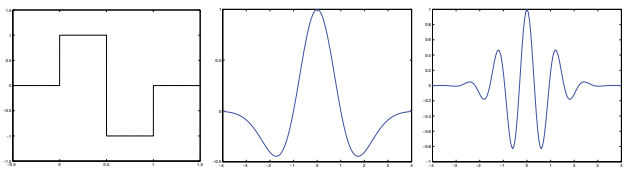
and is also normalized $\|\psi\| = 1$ and centered in the neighborhood of $t = 0$. The scaled and translated versions $\psi_s(u)$ remain normalized as well.

3.2.1 Haar wavelet

In 1910, Haar realized that one can construct a simple piece-wise constant function as shown in figure 1a and is defined as follows:

$$\psi(t) = \begin{cases} 1 \; if \;\; 0 \leq t < 1/2 \\ -1 \; if \;\; 1/2 \leq t < 1 \\ 0 \quad otherwise \end{cases} \tag{8}$$

whose dilation and translations generate an orthonormal basis of $L^2(\Re)$.
Application of this transform to data smoothing and periodicity detection



(a) (b) (c)

Fig. 1. Mother wavelet functions: (a) Haar, (b) Mexican hat and (c) Morlet.

3.2.2 Mexican hat

The Mexican hat is the second derivative of a Gaussian (see fig. 1b)

$$g(t) = (1 - t^2)e^{-t^2/2} \tag{9}$$

They were first used in computer vision to detect multiscale edges.

3.2.3 Morlet's wavelet

This wavelet is defined as follows

$$g(v) = e^{2\pi^2(v-v_0)^2} \tag{10}$$

which is shown in fig. 1c

3.3 Discrete wavelet transform

Calculating wavelet coefficients at every possible scale is a fair amount of work, and it generates an awful lot of data. What if we choose only a subset of scales and positions at which to make our calculations?
It turns out, that if we choose scales and positions based on power of two (called *dyadic* scales and positions) then our analysis will be much more efficient. This analysis is called the *discrete wavelet transform*.

In the discrete case, WT is sampled at discrete meshpoints and using smoother basis functions. But, how to discretize the time-scale domain in order to have a discrete wavelet transform (DWT)?
In the windowed Fourier transform the time frequency domain was discretized using a uniform lattice $\Delta(t_0, \omega_0) = \{(mt_0, n\omega_0)|m, n \in Z\}$ where t_0 and ω_0 are the time and position respectively.
It is known that scaling operation acts as a multiplicative way, that is, composing two consecutive scaling is attained by multiplying each of the scale factor. Thus, starting from an initial scale $s_0 > 1$ we consider all the discrete scales

$$s_m = s_0^m \quad m \in Z \tag{11}$$

Now, how to discretize the time? It is important to note that we must obtain a lattice in the time-scale domain in order to sample (with a minimum redundancy reconstruction of the original signal) the continuous wavelet transform we have seen before from the time-scale domain $\tilde{f}_{m,n}$.
Changing the scale, results in the increase of the width of the wavelet. Also, when the width of the wavelet reduces with a scale reduction operation, we must increase the frequency. One of the most important properties of the WT is the invariance under scale changes. Infact, if we change the scale in the function f and the scale of the underlying space by the same scaling factor, the WT does not change. If we take $f_{s_0}(t) = s_0^{-1/2} f(t/s_0)$ this implies $\tilde{f}_{s_0}(s_0 s, s_0 t) = \tilde{f}(s, t)$.
The invariance property of the wavelet transform is very important and should be preserved even when the WT is discretized. The preservation can be accomplished when we pass from one scale $s_m = s_0^m$ to the other $s_{m+1} = s_0^{m+1}$ by incrementing the time by the scaling factor s_0. We can choose an initial time t_0 and take the length of the sampling time intervals $\Delta t = t_0 s_0^m$. The time discretization lattice for each scale s_0^m is given by

$$t_{m,n} = n s_0^m t_0 \quad n \in Z \tag{12}$$

and the time-scale discretization domain in the lattice is

$$\Delta_{s_0,t_0} = \{(s_0^m, n s_0^m t_0)|m, n \in Z\} \tag{13}$$

the discretization of the WT $\tilde{f}(s, t) =< f, \psi_{s,t}(u) >$ in the time scale lattice is given by

$$\tilde{f}_{m,n} =< f, \psi_{m,n}(u) > \tag{14}$$

where

$$\psi_{m,n}(u) = s_0^{-m/2} \psi(s_0^{-m} u - n t_0). \tag{15}$$

The discrete WT has the following characteristics:

- the sequence $< f, \psi_{m,n} > \quad m, n \in Z$ is an exact representation of f,
- it is possible to reconstruct f from the family of wavelet time-scale atoms $\psi_{m,n}$,
- $\{\psi_{m,n}\}$ constitutes an orthonormal basis for $L^2(\Re)$.

3.4 Multiresolution and subband coding

In the previous sections we have seen the continuous wavelet transform and its discretization in the time-scale domain.

The idea of the scale is mainly related to the problem of point sampling of the signal. When we sample a signal, we have to fix the sampling frequency[2] and the sampling period[3]. If we are sampling the signal at a frequency 2^j this means that frequencies (*details*) outside the scale magnitude of the samples will be last in the sampling process. All of the details captured in a certain scale, will be present at higher scales $2^k \;\; k > m$.

The scaling process gives rise to a subspace generation. Infact, sampling with a frequency of 2^j can give rise to the formation f the subspace $V_j \in L^2(\Re)$ which is constituted by the functions in $L^2(\Re)$ whose details are well represented in the scale 2^j. Now let us define a representation operator that will represent the function $f \in L^2(\Re)$ in the scale 2^j. Let us suppose that there exist a function $\phi \in L^2(\Re)$ such that the family of functions

$$\phi_{j,k}(u) = 2^{-j/2}\phi(2^{-j}u - k) \quad j,k \in Z \tag{16}$$

that represent an orthonormal basis for the subspace V_J.

Defining different scales of ϕ we have

$$\phi_s(u) = \frac{1}{|s|^{1/2}}\phi(\frac{u}{s}) \tag{17}$$

where the width of ϕ is s-times the width of ϕ_s. Thus, as the scale increases or decreases the width of ϕ_s does the same. Now taking $s = 2^j \; j \in Z$ we have

$$\phi_{j,k} = \phi_{2^j}(u - k) = 2^{-j/2}\phi(2^{-j}u - k) \tag{18}$$

is an orthonormal basis for V_j.

For each space V_j with scale 2^j we can define the operator $R_j : L^2(\Re) \rightarrow V_j$. This represent the representation operator that orthogonally project a function $f \in L^2(\Re)$ in the space V_j as follows

$$R_j(f) = Proj_{V_j}(f) = \sum_k < f, \phi_{j,k} > \phi_{j,k}. \tag{19}$$

From the previous relationships emerges that we can represent a function f at several scales. It is important to change the representation from one scale to another without loosing information. The details of one scale at 2^j must appear at a smaller scale 2^{j-1}. Thus $V_j \in V_{j-1}$ that means: given a function $f \in L^2(\Re)$ then $f \in V_j$ iff $f(2u) \in V_{j-1}$ and recursively we can obtain the following

$$f \in V_j \quad iff \quad f(2^j u) \in V_0 \tag{20}$$

Let us now define the multiresolution representation as follows

Definition A multiresolution representation in $L^2(\Re)$ is defined as a sequence of closed subspaces $V_j \in L^2(\Re) \; j \in Z$ that satisfies the following properties:

1. $V_j \subset V_{j-1}$
2. $f \in V_j$ iff $f(2u) \in V_{j-1}$

[2] The number of samples in the time unit.

[3] The length of the sample interval.

3. $\bigcap_{j\in Z} V_j = \{0\}$
4. $\bigcup_{j\in Z} V_j = L^2(\Re)$
5. $\exists \phi \in V_0 \ni' \{\phi(u-k)|k \in Z\}$ is an orthonormal basis of V_0.

The function ϕ is called the *scaling function* of the multiresolution representation. Each of the spaces V_j is called *scale spaces*, or, more precisely, space of scale 2^j.
The orthogonal projection of $f \in L^2(\Re)$ in the space V_j is obtained by using a filtering process of f with the different kernels $\phi_{j,k}$ $k \in Z$ which define low-pass filters. Defining the Haar multiresolution representation as

$$\phi(t) = \begin{cases} 0 \; if \;\; t < 0 \;\; and \;\; t \geq 1 \\ 1 \qquad if \;\; t \in [0,1) \end{cases} \tag{21}$$

whose family represents basis function of the subspace

$$V_j = \{f \in L^2(\Re); f[2^j k, 2^j(k+1)] = constant, k \in Z\}. \tag{22}$$

That is the projection of f on the scale space V_j is given by a function that is constant in the interval $[2^j k, 2^j(k+1)]$. Thus the orthogonality projection of $f \in L^2(\Re)$ in the space V_j is obtained using a filtering process of f with the different kernels ϕ_{jk} $k \in Z$ which define low-pass filters.
Let us now interpret geometrically the sequence of nested scale spaces in a multiresolution representation. Indicating the cutting frequency α_j of this filters we can say that the space V_j is constituted of functions whose frequencies are contained in the interval $[-\alpha_j, \alpha_j]$, $\alpha_j > 0$. Going to a finer scale V_{j-1} we change to the interval $[-\alpha_{j-1}, \alpha_{j-1}]$, where the relation of the two subspaces V_j and V_{j-1} is given by

$$V_{j-1} = V_j \star W_j \tag{23}$$

where W_j is the *detail space* that comprises all the functions of $L^2(\Re)$ with frequencies in the band $[\alpha_j-, \alpha_{j-1}]$ of the spectrum. Thus W_j is orthogonal to V_j and the above states that a function represented on a finer scale space V_{j-1} is obtained from the representation on a coarser scale space V_j by adding details W_j. The details can be obtained by projecting a function f in each subspace W_j using band-pass filtering whose pass-band is exactly $[\alpha_j-, \alpha_{j-1}]$. In fact, this filtering process can be computed by projecting f on an orthogonal basis of wavelets. For each $j \in Z$ there exists an orthonormal basis of wavelets $\{\psi_{j,k}, k \in Z\}$ of the space W_j. Therefore, if R_j is the representation operator on the scale space V_j, we have, for all $f \in L^2(\Re)$

$$R_{j-1}(f) = R_j(f) + \sum_{k\in Z} < f, \psi_{j,k} > \psi_{j,k}. \tag{24}$$

The second term represent the orthogonal projection of function f on the space W_j and it will be denoted by $Proj_{W_j}(f)$. Now rewriting eq. 23 in terms of filters we have

$$R_{j-1}(f) = R_j(f) + Proj_{W_j}(f) R_{j-2}(f) = R_{j-1}(f) + Proj_{W_{j-1}}(f) ... \tag{25}$$

iterating this equation for $R_{j-2}, ..., R_{j-J_0}$, summing up both sides and performing the proper cancellations, we obtain

$$R_{j-J_0}(f) = R_j(f) + Proj_{W_{j-1}}(f) + ... + Proj_{W_{j-J_0}}(f). \tag{26}$$

The projection $R_j(f)$ represents a version of low resolution (*approximation*) of the signal obtained using successive low-pass filters $\phi_j, \phi_{j-1}, ..., \phi j - J_0$. The terms $Proj_{W_{j-1}}(f), ..., Proj_{W_{j-J_0}}(f)$ represent the *details* of the signal lost in each low-pass filtering. This details are obtained by filtering the signal using the wavelets $\psi_j, \psi_{j-1}, ..., \psi j - J_0$.

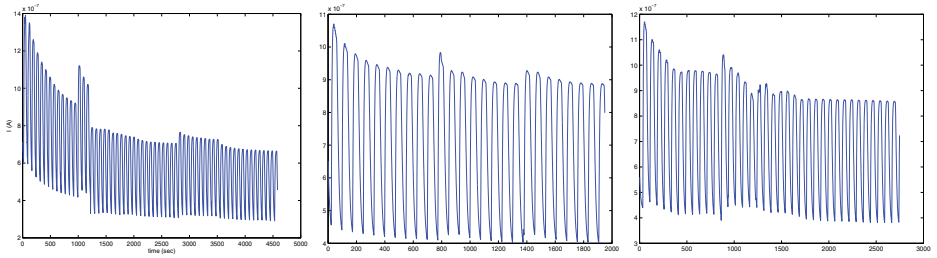


Fig. 2. Three responses relative to Pentanone, Acetone, and Exanal in 50%RH (from left to right).

4. Experiments

4.1 DWT analysis

The active layers of the array consist of pure and doped SnO_2 thin films prepared by means of sol-gel technology. Pd, Pt, Os, and Ni were chosen as doping elements starting from different precursors of the preparation of the modified films. The films, whose thickness was about $100nm$, were deposited on alumnia substrates supplied with interdigitated electrodes and platinum heater, by the spin coating technique at 3000 rpm, dried at 80 °C and heat treated in air at 600 °C. After deposition, the sensors were mounted onto a TO8 socket and inserted in the test chamber.

Samples of different compounds were introduced into a vial kept at room temperature by a thermostatic tank. Many subsequent measurements were performed, for each sample, by fixing the exposure time and the purging time at 20 minutes. The responses have been acquired with a sampling interval of 32 seconds then acquiring 75 points for each response.

Three gases have been taken under consideration (216 measurements in total): Acetone, Hexanal, and Pentanone in 50% relative humidity (RH) and dry air. For a preliminary analysis these gases in 50% RH.

Two different kinds of analysis have been carried out by using PCA and neural network. The first analysis is purely qualitative and for visualization purposes, project data obtained by the several feature extraction methods previously discussed onto the first two useful principal components.

Principal Component Analysis is usually carried out as a low pass filter, in order to reduce noise in the signal, keeping components corresponding to the first few eigenvectors that capture most of the variance contained in the data set. Usually the transformation is made into 2D or 3D spaces. Let $\mathbf{A}$ be the truncated transformation matrix constituted of the first useful principal components of the correlation matrix of the data set. In the experiments that follow in this paper, the first three principal components from the measurements y_{ij} have been extracted, so the matrix $\mathbf{A}$ is $3 \times n$ and then the observation becomes the 3D vector $x_j = \mathbf{A} \cdot y_j$.

The second analysis is more deep because is based on the classification results of a Radial Basis Function neural network, which is given a pattern composed of the coefficients extracted with a pre-processing method. The training and validation procedure is performed by using the

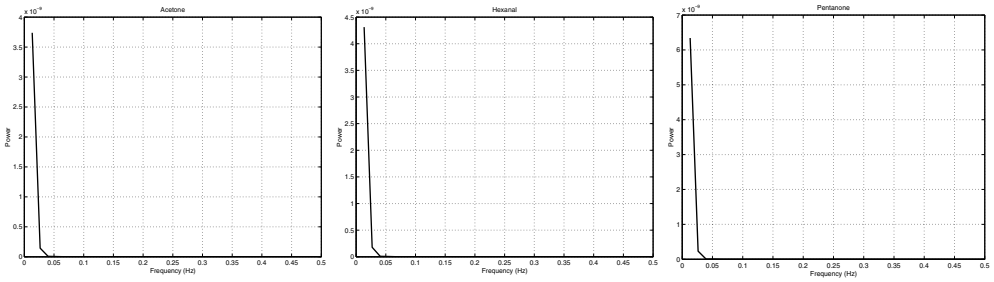


Fig. 3. Power spectra density computed on three different odour responses.

leave-one-out procedure which provides an estimate of the generalization performances of the final classifier.

The RBF network used, creates neurons one at a time. At each iteration the input vector which will result in lowering the network $sum - squared$ error, is used to create a new radial basis neuron. The error of the new network is checked, and if low enough the learning phase is finished. Otherwise the next neuron is added. This procedure is repeated until the error goal is met (0.001), or the maximum number of neurons is reached (i.e. the number of training vectors 215). A spread of 0.8 is used for radial basis functions in order to ensure that more than one neuron can respond to overlapping regions of the input space.

Analysis in the frequency domain has been carried out and figure 3 shows three curves relative to one of the three gases under consideration. The electronic nose responses possess the same frequency domain but different magnitudes for different odors. So an appropriate low pass filter is suitable for feature extraction since the responses have the useful information we need lying in low frequencies, as opposed to noise that is present in high frequencies. Unfortunately, drift also resides in low frequencies and this method carries also this problem in the transformation, but the reduction of the drift is outside the scope of this paper. As introduced in the previous sections, at each level (scale) of signal decomposition, wavelet coefficients are divided by approximation (low frequencies) and details (high frequencies). Also, the coefficients are obtained by convolving the wavelet function and the response curve, thus measuring their correlation degree. The mother wavelet used is the Daubechies family since they guarantee an orthogonal analysis, a necessary condition for feature extraction. After choosing the mother wavelet to use, the next investigation is the level to stop for getting the approximation coefficients of the DWT. The higher the scale the more low frequencies are amplified and high frequencies are cut off. But this is not an infinite process, infact the decomposition level has a lower limit that is the sampling period and a superior limit that is the signal support. The maximum decomposition level of the DWT is the 2-base logarithm of the length of the signal. Since each transient has a length of approximately 80 samples, then the 6th level is appropriate, where $2^6 = 64 < 80$ (the 7th level requires a signal greater than 80 samples).

The graph of the coefficients for one sensor and 3 exposures of the sensor array to the same odor (hexanal) is given in fig. 5 where coefficients up to the 6th decomposition level are shown. The first 13 coefficients on the x-axis shows approximation and the remainings show details. It is interesting to note that the first coefficients of the three curves fall all in the same position as shown by the leftmost patterns (they overlap each other). This may bring to the conclusion that the same order of coefficients related to the same odor are the same. So in this graph the first three order of coefficients can be taken as features to be successively classified by an opportune classifier. As a counter-example, fig. 4 shows coefficients related to

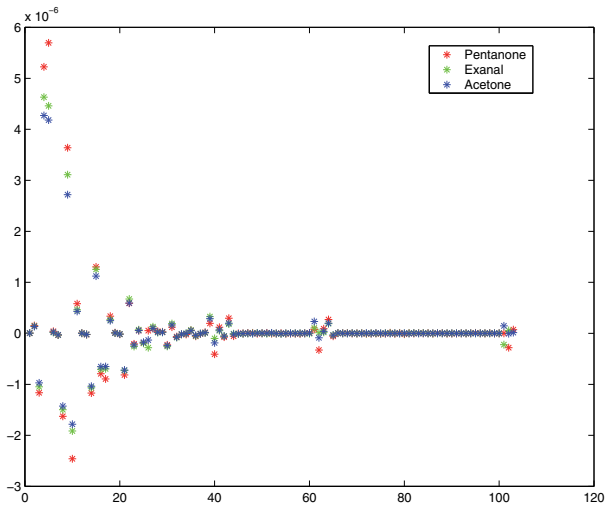


Fig. 4. Coefficients extracted up to the 6-th level of DWT decomposition for three responses of the three considered odour (pentanone, hexanal, and acetone).

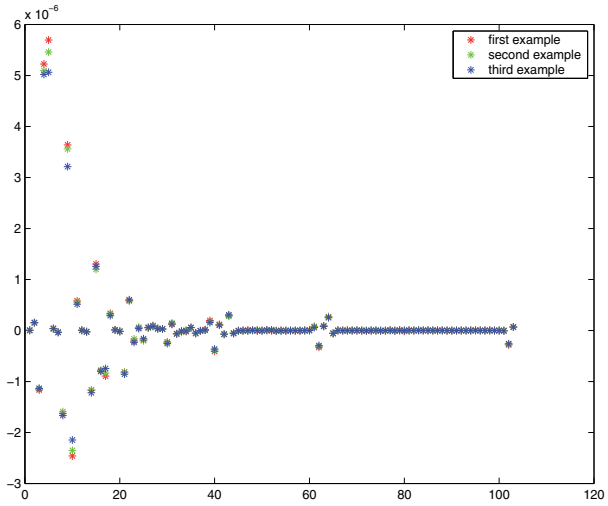


Fig. 5. Coefficients extracted up to the 6-th level of DWT decomposition for three response of the same odour (hexanal).

different odors. However, the first coefficients of the approximation of the three odor curves do not discriminate well since they overlap. The discrimination starts from the third order of coefficient until the sixth, and the scale is larger than the coefficients shown in the figure 5. This states that a number of 6 coefficients is suitable as features and that wavelet transform enlarge the range of the coefficients for different odors while maintaining small the variability of the range coefficients related to the same odor (compare the two figures for the 5th coefficient).
The transient has been analysed in three time parts response: rise time, recovery time and the complete curve (i.e. from the rise time to the next when a new odor is presented). Over all this experiments the extraction of the first 6 coefficients over the complete transient has shown

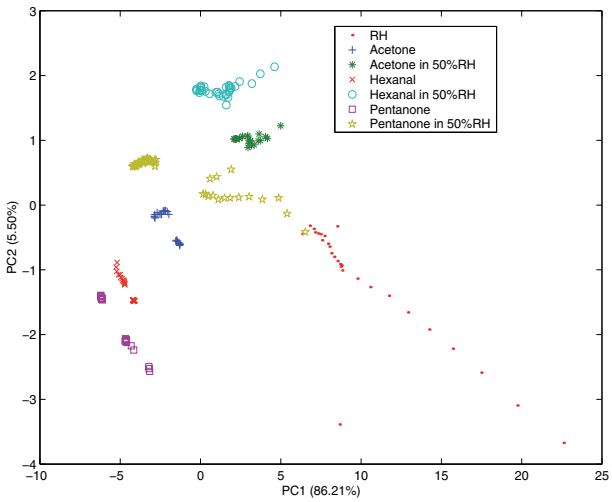


Fig. 6. Principal component analysis done in the wavelet space for the three gases measured in dry and humid air. Each observation has 30 coefficients (features) since for each response of the five sensors, the first six coefficients are extracted (compare the result with the traditional method shown in fig. 7)

better results. However it has been shown that the recovery time is more informative than the rise time, but in general better results are obtained with the complete curve. In fig. 6 is shown the score plot of the first two principal components computed in the wavelet space for visualization. Table 1 gives the classification results with RBF for input pattern of 30 wavelet descriptors.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
(1)	28	0	0	0	0	0	0
(2)	0	21	0	0	0	0	0
(3)	0	0	25	0	0	0	0
(4)	0	0	0	29	0	0	0
(5)	0	0	0	0	35	0	0
(6)	0	0	0	0	0	22	0
(7)	0	0	0	0	0	0	60

Table 1. Confusion matrix of the wavelet analysis. The total recognition percentage is 100%. (1) Humidity; (2) Acetone; (3) Acetone in 50% RH; (4) Hexanal, (5) Hexanal in 50% RH; (6) Pentanone; (7) Pentanone in 50% RH.

Traditional methods such as the relative, fractional, difference and log parameter, present almost the same behavior for the data under consideration in terms of classification results. Here we report the results of the relative method shown in fig. 7 and the corresponding confusion matrix in table 2 for input patterns to the RBF classifier of 5 coefficients (the number of sensors).

As a comparison, also Fourier descriptors have been tested against wavelet descriptors. Fourier descriptors provide a frequency measure of the curve under consideration without localizing them in the time domain. The Fast Fourier Transform (FFT) is used which allows us to compute very quickly the discrete transformation with a good approximation. The

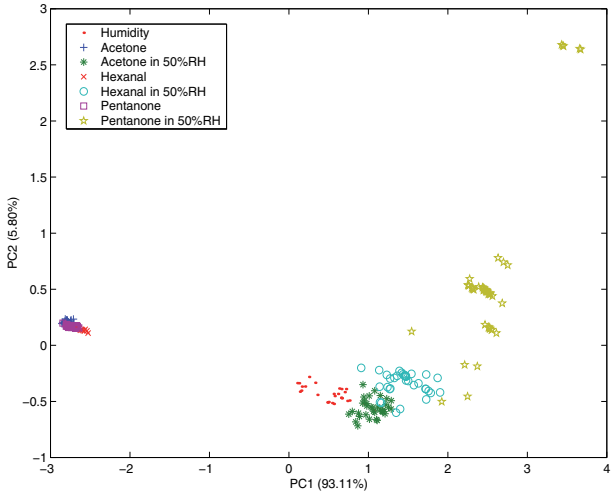


Fig. 7. Result using the traditional method of applying the relative method (I_s / I_0) for the extraction of the features from one response of the sensor and then performing the principal component analysis on the observation space of five elements (one observation from each sensor as opposed to the six extracted with the wavelet analysis).

	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**
(1)	26	0	2	0	0	0	0
(2)	0	21	0	0	0	0	0
(3)	4	0	16	0	5	0	0
(4)	0	0	0	28	0	1	0
(5)	1	0	1	0	31	0	2
(6)	0	0	0	22	0	0	0
(7)	0	0	0	0	3	0	57

Table 2. Confusion matrix of the relative method feature extraction. The total recognition percentage is 81.36%. (1) Humidity; (2) Acetone; (3) Acetone in 50% RH; (4) Hexanal, (5) Hexanal in 50% RH; (6) Pentanone; (7) Pentanone in 50% RH.

first 10 descriptors have been considered for each curve, since they possess the greatest magnitude and, low frequencies dominate in the response curve. The feature vector pattern is then composed of 50 descriptors, which have been: projected onto the first two principal components for visualization as shown in figure 8 and; given to RBF for discrimination capabilities (table 3).

Integrals and derivatives have also been investigated to benchmark them against the above discussed feature extraction methods.

The aim is the computation of the integral of the f function in the interval $[a, b]$, where a represents the time step of the concentration change in the rise time, and b the end of the recovery time (where $f(t) = 0$). The applied Newton-Cotes method Chapra & Canale (1988) is based on the substitution of the function that represents the transient response with a more simple approximation function:

$$I = \int_a^b f(x)dx \approx \int_a^b f_n(x)dx \tag{27}$$

where $f_n(x)$ is a polynomial function of order n defined as follows:

$$f_n(x) = a_0 + a_1 x + \ldots + a_{n-1} x^{n-1} + a_n x^n \tag{28}$$

The order of the polynomial function determines the accuracy of the method, in our experiments we used the "Simpson's 1/3 Rule" which uses a second order polynomial function.
Simpson's rule finds the area under the parabola which passes through 3 points (the end points and the midpoint, i.e. x_0, x_1, x_2) on a curve. In essence, the rule approximates the curve by a series of parabolic arcs and the area under the parabolas is approximately the area under the curve. The Simpson's rule as well as the other Newton-cotes methods can be applied only if the points are equally spaced. The transient response has been split into 25 intervals where the area is computed as follows:

$$I = (x_2 - x_0)\frac{f(x_0) + 4f(x_1) + f(x_2)}{6}. \tag{29}$$

The observation pattern is composed of five features each one of them describing the area of the corresponding sensor response. Figure 9 shows the PCA plot of the extracted features. The five features are given to the RBF for training and classification with leave-one-out procedure, and the confusion matrix is given in table 4. It is interesting to note that this method produces very informative features as compared with the results obtained with the wavelet descriptors. Another applied method is the study of the local gradient over the whole transient response. Even in this case the study of the local gradient has been carried out by approximation using the Taylor series. Starting from a first order Taylor series $f(x_{i+1}) = f(x_i) - f'(x_i)(x_{i+1} - x_i)$ we can approximate the derivative in the point x_i as follows:

$$f'(x_i) = \frac{f(x_i) - f(x_{i+1})}{x_{i+1} - x_i}. \tag{30}$$

The mean derivative has been computed over intervals of 10 point samples so then for the whole transient response a number of $(75/10) \approx 7$ features are obtained, leading to 35 features for each observation of the array five sensors. Figure 10 shows the result of mean derivative method by using PCA and in table 5 the confusion matrix of the 7 feature vectors classified by the RBF is given.

	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**
(1)	28	0	0	0	0	0	0
(2)	0	20	0	0	0	1	0
(3)	0	0	23	0	2	0	0
(4)	0	0	0	28	0	1	0
(5)	1	0	1	0	33	0	0
(6)	0	0	0	0	0	22	0
(7)	0	0	0	0	2	0	58

Table 3. Confusion matrix of the FFT method. The total recognition percentage is 96.36%. (1) Humidity; (2) Acetone; (3) Acetone in 50% RH; (4) Hexanal, (5) Hexanal in 50% RH; (6) Pentanone; (7) Pentanone in 50% RH.

All of the compounds in dry air have shown good results. Only the cluster related to humidity is well defined and separable from others in all the experiments, while hexanal and acetone

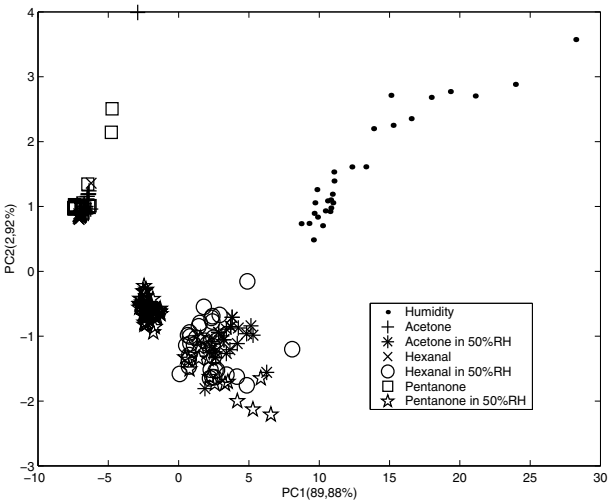


Fig. 8. Principal component analysis done in the Fourier space for the three gases measured in dry and humid air. Each observation has 50 features since for each response of the five sensors, the first ten coefficients are extracted.

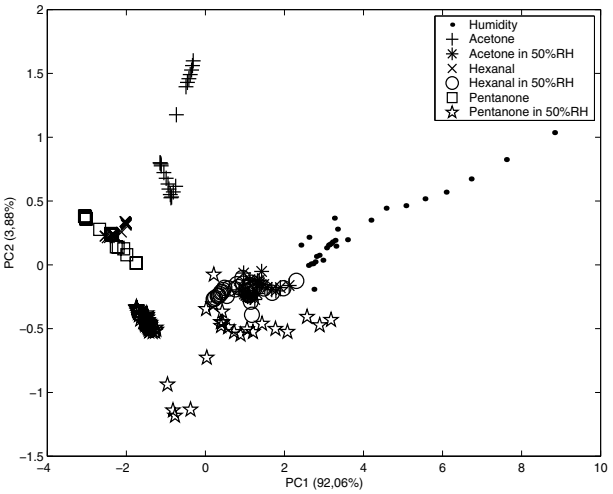


Fig. 9. Principal component analysis with the integration method for gases measured in dry and humid air. Each observation has 5 features corresponding to the areas computed on the responses starting from the release of the aroma in the test chamber.

are mixing up in dry air. Among all the steady-state methods, the difference method have shown better results.
However, when analysed in presence of humidity (50% RH) the performance of the classifier are decreased. It is interesting to note that after inspecting data with PCA, humid compounds and dry compounds are separate from each other. Traditional methods provides poor results in the first two principal dimensions, infact, in the two macro-clusters (dry and humid) the compounds have disordered positions. Hexanal from Pentanone in dry air have correlated

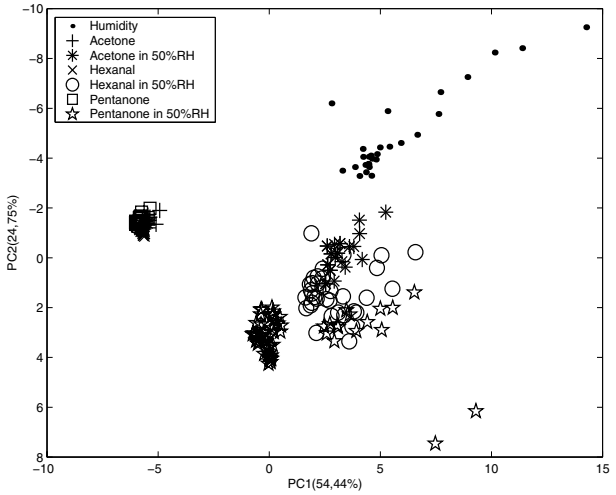


Fig. 10. PCA plot of the features extracted with the derivative method.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
(1)	28	0	0	0	0	0	0
(2)	0	21	0	0	0	0	0
(3)	0	0	25	0	0	0	0
(4)	0	0	0	29	0	0	0
(5)	0	0	0	0	34	0	1
(6)	0	0	0	0	0	22	0
(7)	0	0	0	0	0	0	60

Table 4. Confusion matrix of the Integration method. The total recognition percentage is 99.54%. (1) Humidity; (2) Acetone; (3) Acetone in 50% RH; (4) Hexanal, (5) Hexanal in 50% RH; (6) Pentanone; (7) Pentanone in 50% RH.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
(1)	28	0	0	0	0	0	0
(2)	0	20	0	1	0	0	0
(3)	0	0	25	0	0	0	0
(4)	0	0	0	29	0	0	0
(5)	0	0	4	0	29	0	2
(6)	0	1	0	0	0	21	0
(7)	0	0	0	0	2	0	58

Table 5. Confusion matrix of the gradient method. The total recognition percentage is 95.45%. (1) Humidity; (2) Acetone; (3) Acetone in 50% RH; (4) Hexanal, (5) Hexanal in 50% RH; (6) Pentanone; (7) Pentanone in 50% RH.

features, making RBF difficult to classify them correctly. It has been shown that features made only of steady-state information, weakly discriminate classes of odors. Features extracted from the transient response, are more informative but as opposed to traditional steady-state methods, they inherit sensor long term stabilities problems. These are shown with the humidity and Pentanone measurements. The derivative method presents better results with

Pre-processing Method	Recognition rate (%)
Relative	81
Log	81
Difference	83
Fractional	82
Derivative	95.45
Fourier coeff.	96
Integration	99.5
Wavelet coeff.	100

Table 6. Comparison between several feature extraction methods in terms of informative contents contained in each pattern. The results are based on the leave-one-out procedure performed with a RBF neural network.

95% classification rate considering the whole curve analysis. Some of the transient analysis methods have also been investigated for the rise and recovery time. A study of the derivative method restricted to the rise time and recovery time has produced 95% and 93% recognition rates respectively. The integration method have shown results better results both in the PCA space and with RBF classification leading to 99% over the whole response and 96% and 98% for the rise and recovery time respectively. The multiresolution approach with wavelet analysis conducted for the whole response have shown best performance overall methods here presented. In this case, both linear separation and better classification results have been obtained.
Rise time and recovery time parts of the curve gave classification rates of 99% and 100% accuracy respectively.

4.2 Drift counteraction

The micro-sensor array consisted of seven semiconductor metal oxide sensors whose sensing thin film material were pure and doped tin- dioxide SnO_2. Os, Ni, Pt, Pd elements were used as doping elements. The films had been prepared starting from tin tetrachloride, as precursor, with the aim to obtain stabilized SnO_2 solutions The films, whose thickness was 100nm, were deposited on alumina substrates, ($3 \times 3mm^2$) supplied with inter-digitated electrodes and platinum heater, by means of the spin coating technique at 3000 rpm, dried at 80 8C and heat treated in air at 600 8C. After deposition, the sensors had been mounted onto a TO8 socket and inserted in the test chamber. The conditioning of the sensors had been performed by using dry air (flow 100sccm), used also as reference gas for acquiring the baseline current values of the sensors. The device was exposed to three different gases: acetone, hexanal and pentanone in 50% relative humidity (RH). The measurements were performed by fixing the exposure time and the purging time at 20 min. The responses have been acquired with sampling interval of 32 sec. then acquiring 75 points for each response. Further long-term stability of the sensors was tested by performing measurements of baseline current values in dry air over a period of 76 days. At the same time a morphological study on sensor material was performed to reveal poising or ageing during their use by means of the Scanning Electron Microscopy (SEM) images and the X-ray Photoelectron Spectroscopy (XPS) study. As it can be seen in figure 11, it is evident a drifting behavior of all sensors. In this case a drift contamination is more evident for SnO_2-Os based sensor and the pair of identical sensors (SnO_2-Ni and SnO_2-Pd based sensors) show a similar trend in terms of baseline during the whole period

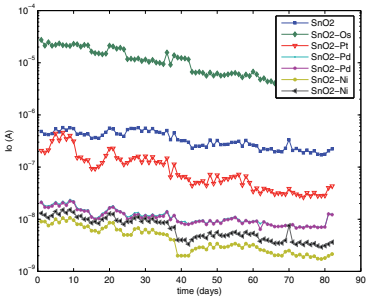


Fig. 11. Effect of drift on the baseline current I_o in dry air over a period of 76 days.

of measurements. The morphological analysis did not reveal effects of the ageing of sensing material while it showed a clear degradation of Ti/Au inter-digitated contacts due to high working temperature Capone et al. (2006). The signal pre-processing method based on

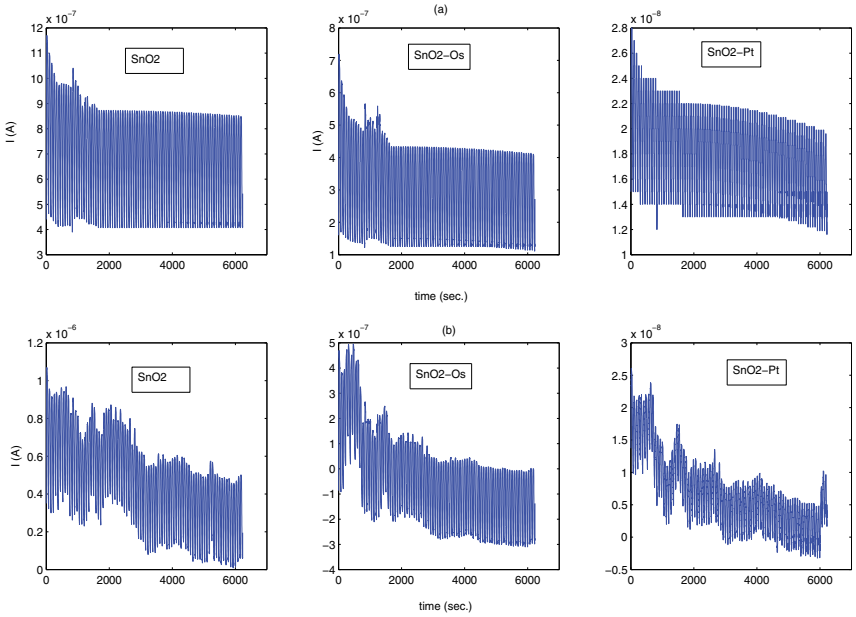


Fig. 12. Example of simulated metal oxide sensor responses to a same odor without drift effect (a) and subjected to drift(b).

DWT was tested on drifting sensor responses, which had been artificially generated in according to the drifting behavior of the sensors previously examined and showed in the fig 11. The drifting signal was obtained by adding the interpolation of sensor's baseline trend to a signal of n repeated cycles of sensor responses (fig. 12). The multilevel wavelet decomposition of these simulated signals was performed by applying iteratively high-pass and low-pass filters. The lowest frequencies influence the deepest levels (in this case the level fixed N), so that the associated approximation coefficients could have drift contamination. In this case it is needed to discard the N-level approximation coefficients containing drift

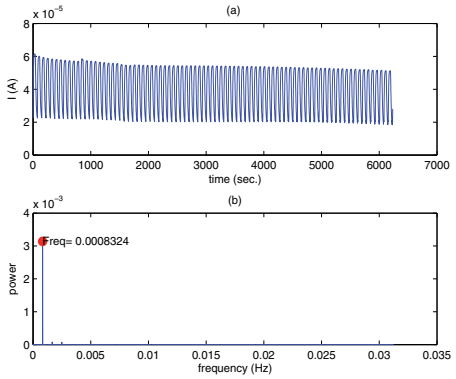


Fig. 13. SnO_2-based sensor response (a). Frequency content of signal via periodogram (b).

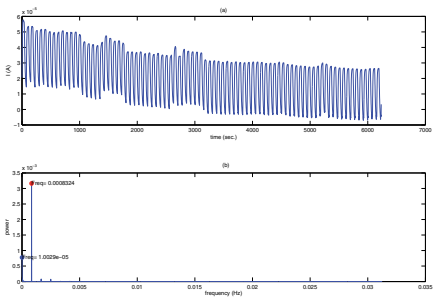


Fig. 14. SnO_2-based sensor response subjected to drift (a). Frequency content of signal via periodogram (b).

contamination. More precisely, each approximation coefficient was fixed to the value of the first approximation coefficient vector element to avoid shifting of the reconstructed signal with respect to the processed one. The wavelet reconstruction was computed by using these changed approximation coefficients and remaining wavelet coefficients from level 1 to N-1. The decomposition level was fixed once analysis in the frequency domain has been carried out to single out the frequency domain of the drifting trend. The figures 13 and 14 show two periodogramms of the SnO_2-based sensor response and the drifting one respectively. The drifting trend frequency component predominates at the scale a which has been gauged by means of the following relationship:

$$F_a = \frac{F_c}{a * \Delta}$$

where a is a scale with $a = 2^N$ (N is the number of decomposition levels), F_c is the center frequency associated to wavelet function, Δ is sampling period and F_a is the pseudo-frequency associated to scale a. The value of F_a must be close to the frequency whose component must be filtered out. The sixth order Daubechies ($db6$) was selected as analyzing wavelet, and the level of wavelet decomposition of the drifting SnO_2-based sensor response was fixed at $N = 6$. The pseudo-frequency F_a associated to scale 2^6 is $3.5511e-04Hz$, so it could filter out the frequency components less than F_a discarding the associated approximation coefficients.

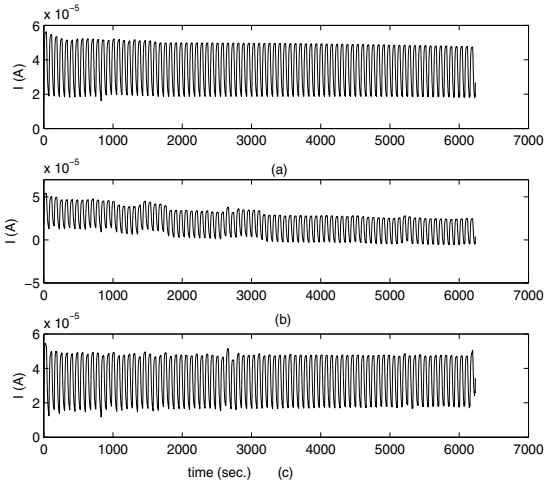


Fig. 15. The original signal (a), the drifting one (b) and DWT-filtered one (c)

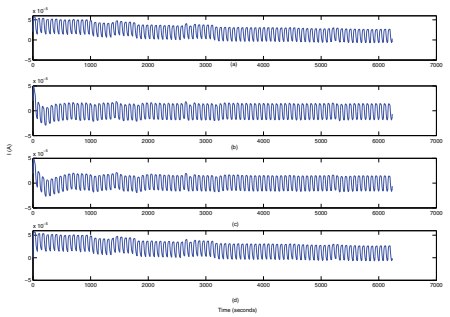


Fig. 16. The drifting signal (a), the Butterworth-filtered one (b), Elliptic-filtered one (c) and FIR-filtered one (d)

Figure 15 shows the filtered signal and tin comparison with the performance of the usual high-pass filters (fig. 16). After pre-processing sensor responses, the relative parameters I_s/I_0 (I_s represents the steady-state current value and I_0 the relative baseline value) were extracted from the drifting and filtered signals. These feature sets was subsequently normalized and analyzed by Principal Component Analysis to visualize the data set. The PCA score plot regarding to the feature set obtained from drifting signals (fig. 17) shows an overlap of the clusters relative to the three gases. A better cluster separation was obtained after the signal pre-processing based on DWT, and in the respective PCA score plot (fig. 18) the three clusters are distinct.

5. Conclusions

The chapter has shown the use of discete wavelet analysis to characterize electronic nose responses from array of gas sensors. The addressed problems were related to signal representation that finds its basis under the signal compression context in order to find compact ways for classification purposes and for denoising under the sensor drift problem.

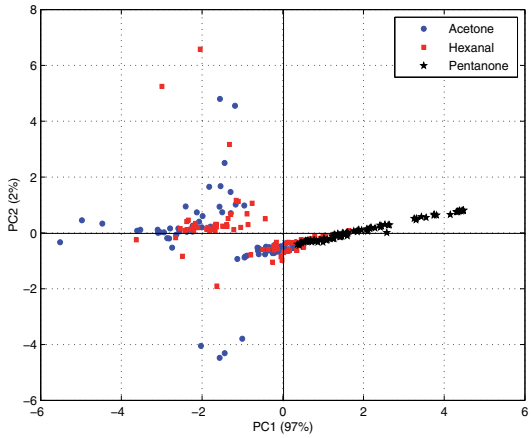


Fig. 17. The PCA score plot of feature sets extracted from drifting sensor responses.

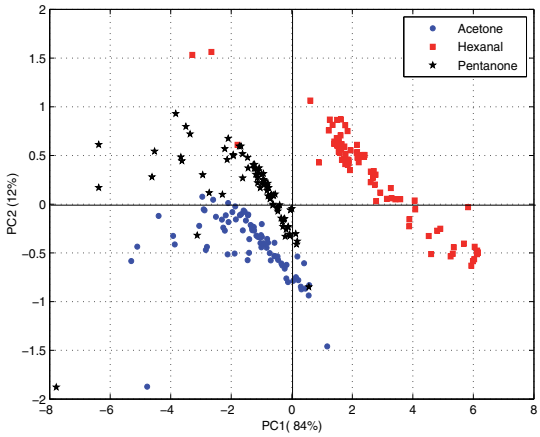


Fig. 18. The PCA score plot of feature sets extracted from DWT-filtered sensor responses.

In the first problem, the study of the transient response in terms of informative content stored in the curve (rise part, recovery part and both) have been investigated. It has been shown with this tool, that the recovery part is more informative than the rise one, but, the use of the whole curve carries better information in terms of discriminability. The feature vector is obtained from the approximation coefficients of multiresolution analysis with wavelet transform and is analysed with principal component analysis and Radial basis Function neural network. Although wavelet transform captures the frequency content of the signal at each decomposition level, it would have been reasonable to make a feature vector composed by maximums approximation coefficients in each level. This performed poorer than the method we used, since a good compression rate and the more discriminative features were obtained with the coeffifients of the highest decomposition level.
The method is compared with the most common techniques of signal processing in the context of electronic nose community. The performed measurements have shown that when we

use the wavelet descriptors the capability of the recognition system to classify three organic compounds improves notably. Table 6 gives the recognition results by using a RBF neural network trained with the features extracted with the considered methods.
This work showed a DWT-based method of signal filtering capable to remove drift effects from the sensor responses and recover the drifting signals to subsequently data analysis. The PCA analysis, employed on features extracted from pre-processed signals showed a clear improvement in the discrimination of the gases with an increase of the distance among the clusters in the principal components space. This signal pre-processing method is enough simple and not time-consuming and the filtered signals are not distorted in comparison with ones obtained by using usual high-pass filters, due to a a more flexible analysis of the signal that wavelet transform allows.

6. References

Acevedo, F., Maldonado, S., Domínguez, E., Narváez, A. & López, F. (2007). Probabilistic support vector machines for multi-class alcohol identification, *Sensors and Actuators B: Chemical* 122(1): 227–235.

Artursson, T., ov ov ov ov, T. E., om om om om, I. L., Mårtensson, P., ostr: om ostr: om, M. S. & Holmberg, M. (2000). Drift correction for gas sensors using multivariate methods, *Journal of Chemometrics* (14): 1–13.

Bakshi, B. R. (1999). Multiscale analysis and modeling using wavelets, *Journal of Chemometrics* 13(3-4): 415–434.

Capone, S., Epifani, M., Francioso, L., Kaciulis, S., Mezzi, A., Siciliano, P. & Taurino, A. M. (2006). Influence of electrodes ageing on the properties of the gas sensors based on sno2, *Sensors and Actuators B: Chemical* 115(1): 396–402.

Chapra, S. & Canale, R. (1988). *Numerical methods for engineers with personal computer applications*, 2nd edn, McGraw-Hill.

Distante, C., Leo, M., Siciliano, P. & Persaud, K. C. (2002). On the study of feature extraction methods for an electronic nose, *Sensors and Actuators B: Chemical* 87(2): 274–288.

Distante, C., Sicilian, P. & Persaud, K. C. (2002). Dynamic cluster recognition with multiple self-organising maps, *Pattern Analysis & Applications* 5: 306–315.

Haugen, J. E., Tomic, O. & Kvaal, K. (2000). A calibration method for handling the temporal drift of solid state gas-sensors, *Anal. Chim. Acta* (407): 23–39.

Huang, D. & Leung, H. (2009). Reconstruction of drifting sensor responses based on papoulis #x2013;gerchberg method, *Sensors Journal, IEEE* 9(5): 595–604.

Leone, A., Distante, C., Ancona, N., Persaud, K., Stella, E. & Siciliano, P. (2005). A powerful method for feature extraction and compression of electronic nose responses, *Sensors and Actuators B: Chemical* 105(2): 378–392.

Loutfi, A., Coradeschi, S., Lilienthal, A. J. & Gonzalez, J. (2009). Gas distribution mapping of multiple odour sources using a mobile robot, *Robotica* 27(02): 311–319.

Mallat, S. (1999). *A wavelet tour of signal processing*, 2nd edn, Academic Press.

Phaisangittisagul, E. & Nagle, H. (2008). Sensor selection for machine olfaction based on transient feature extraction, *Instrumentation and Measurement, IEEE Transactions on* 57(2): 369–378.

Phaisangittisagul, E. & Nagle, H. T. (2010). Predicting odor mixture's responses on machine olfaction sensors, *Sensors and Actuators B: Chemical* In Press, Corrected Proof: –.

Pijolat, C., Riviere, B., Kamionka, M., Viricelle, J. P. & Breuil, P. (2003). Tin dioxide gas sensor as a tool for atmospheric pollution monitoring: Problems and possibilities for improvements, *Journal of Materials Science* 38: 4333–4346.

Rubinstein, R., Bruckstein, A. & Elad, M. (2010). Dictionaries for sparse representation modeling, *Proceedings of the IEEE* 98(6): 1045–1057.

Sisk, B. C. & Lewis, N. S. (2005). Comparison of analytical methods and calibration methods for correction of detector response drift in arrays of carbon black-polymer composite vapor detectors, *Sensors and Actuators B: Chemical* 104(2): 249–268.

Trincavelli, M., Coradeschi, S. & Loutfi, A. (2009). Odour classification system for continuous monitoring applications, *Sensors and Actuators B: Chemical* 139(2): 265–273.

Zuppa, M., Distante, C., Persaud, K. C. & Siciliano, P. (2007). Recovery of drifting sensor responses by means of dwt analysis, *Sensors and Actuators B: Chemical* 120(2): 411–416.

Zuppa, M., Distante, C., Siciliano, P. & Persaud, K. C. (2004). Drift counteraction with multiple self-organising maps for an electronic nose, *Sensors and Actuators B: Chemical* 98(2-3): 305–317.

11

Wavelets in Electrochemical Noise Analysis

Peter Planinšič[1] and Aljana Petek[2]
[1]*University of Maribor, Faculty of Electrical Engineering and Computer Science*
[2]*University of Maribor, Faculty of Chemistry and Chemical Engineering*
Slovenia

1. Introduction

Electrochemical noise (EN) is the term used to describe the spontaneous fluctuations of current or potential, which are generated during the corrosion processes. It has been investigated extensively since the 1968, and data has shown the use of EN measurements offers valuable sources of information about complex electrochemical reactions such as those in corrosion systems (Gabrielli et al., 1985; Bertocci & Huet, 1995) .
Many methods can be used to analyze the data, such as the variance, standard deviation and root mean square in the time domain. Alternatively, the signal has been transformed from the time domain to the frequency domain using fast Fourier transform or the maximum entropy method, giving the power spectrum density (PSD). The technique of wavelet analysis may be used instead, where a set of wavelets of varying amplitude, duration, and location be constructed such that reproduces the signal of interest.
Wavelets based methods are modern mathematical tools for multiscale time frequency analysis and characterization of in general nonstationary EN signals. This work presents the short overview to usability and possibilities of wavelet transformation in comparison with classic analysis.

2. Corrosion processes and electrochemical noise

Corrosion can be defined as the deterioration of materials due to its interaction with its environment and is appearing in various forms: as localized corrosion and as general (uniform) corrosion. Localized corrosion results in the accelerated loss of material at discrete sites in a passive materials surface, leading to the perforation or other failure. Since the majority of the surface is unaffected the difficulty with localized corrosion is its detection and prediction. Uniform corrosion results from the sites that are distributed randomly over the surface regarding both the space and time. Uniform corrosion damage is manifested in dissolving and the progressive thinning of a metal. To prevent corrosion or to predict the outcome of a corrosion situation the knowledge of mechanism of various form of corrosion is fundamental. Thermodynamic principles can be applied to determine which processes can occur and how strong the tendency is for changes to take place. Kinetic laws then describe the rates of the reactions.
Corrosion of metals in aqueous environments is electrochemical in nature. It occurs when two or more electrochemical reactions take place on a metal surface, producing dissolved

species of metal or solid corrosion products and thus lowering the energy of the system. The corrosion process has been written as two separate reactions occurring at two distinct sites on the same surface: the anode (metal dissolution site) and the cathode (site of the accompanying reduction reaction). At corrosion of iron in an acid solution containing dissolved oxygen, iron is oxidized to ferrous ion which passes into solution in the anode region

$$Fe \rightarrow Fe^{2+} + 2e^- \tag{1}$$

and H^+ ion is reduced to hydrogen or O_2 is reduced to water in the cathode region

$$2\,H^+ + 2e^- \rightarrow H_2 \tag{2}$$

$$O_2 + 4\,H^+ + 4e^- \rightarrow 2\,H_2O \tag{3}$$

The overall corrosion reaction is the sum of the anodic and cathodic partial reactions: sum of the reactions (1) and (2)

$$Fe + 2\,H^+ \rightarrow Fe^{2+} + H_2 \tag{4}$$

and sum of the reactions (1) and (3)

$$2\,Fe + O_2 + 4\,H^+ + 4e^- \rightarrow 2\,Fe^{2+} + 2\,H_2O \tag{5}$$

Electrochemical corrosion processes can be investigated by observation of charge flows between the electrolyte and the corroding metal. The measure of the rate at which reducible or oxidizable species can gain or lose electrons is the current density, j i.e. the charge flux trough a metal/electrolyte interface. The potential dependence of charge transfer rate is known as the Butler-Volmer equation:

$$j = j_0 \left\{ e^{(1-\alpha) f \eta} - e^{-\alpha f \eta} \right\} \tag{6}$$

where j_0 is exchange current density, α is transfer coefficient, η is overpotential, i.e. measure of how far the reaction is from equilibrium, and f is defined by:

$$f = \frac{nF}{RT} \tag{7}$$

where F is Faraday constant, R is the gas constant and T is the absolute temperature. If the Butler-Volmer equation is used to express the current densities in corrosion processes, the anodic reaction is the metal dissolution and the cathodic reaction is the accompanying reduction of H^+ or O_2. The equilibrium is achieved (the overpotential is zero) at the corrosion potential where no measurable current j flows: the anodic current density of metal dissolution must be equal to the cathodic current density and is equal to the corrosion current density, j_0. The overpotential is said to be positive if it is such as to produce a positive current, i.e. if it drives the anodic oxidation reaction and suppresses the reduction reaction.

In measurements procedure a potentiostat is often used and the electrochemical cell with tree electrodes: working electrode (WE) represents the interface of interest, the reference

electrode (RE) acts as standard for potential measurements and the counter electrode acts as electron sink or source for reactions that occur on the surface of WE. The potential of the WE is controlled with respect to the RE at a constant value, and the current density j under those conditions is determined. If the mean current is compensated or subtracted from the data and only random fluctuations are remained then electrochemical current noise is obtained. Electrochemical current noise thus can be measured as the random fluctuation in current of WE that is held at fixed potential or as the galvanic coupling current between two nominally identical working electrodes. Electrochemical potential noise is measured as the random fluctuation in potential of a WE with respect to a RE or as the fluctuation in potential difference between two nominally identical working electrodes.

In our investigations the current noise was monitored by a low-noise battery-operated potentiostat Jaissle IMP88 PC-R at a sampling rate of 10 points per second. The cell assembly was put in a Faraday cage. An example of measured time series for two processes (pitting as signal I_0 and general corrosion as signal I_2) is presented in Fig. 1. As is seen, the signal has a relatively smooth appearance for general corrosion and occasional sharp increases and decreases in the amplitude of current noise data occur for localized corrosion.

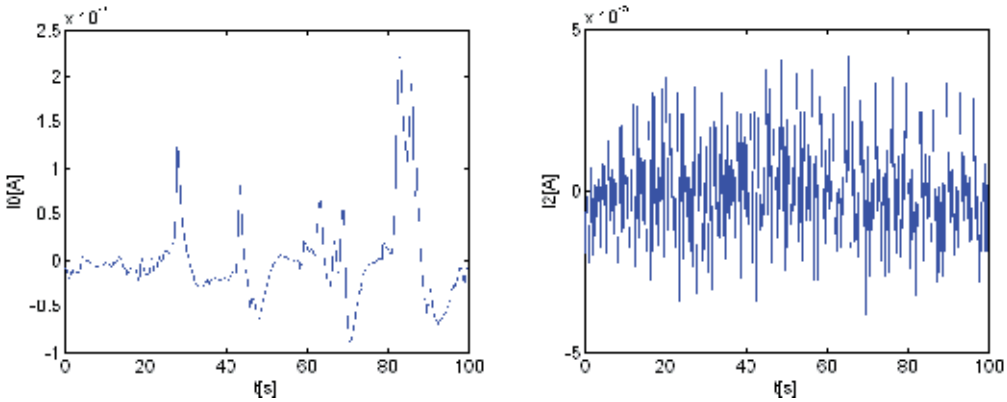


Fig. 1. Current noise signal corresponding to: a) X5CrNiMo17-13 stainless steel in borate buffer solution + 0.1 mol/dm^3 NaCl at the passive potential (I_0); b) Low Carbon Steel in 0.1mol/dm^3 H_2S0_4 at the corrosion potential (I_2) (Planinšič & Petek, 2008).

3. Stochastic processes and 1/*f* noise

EN-signals are generated from corrosion processes which are stochastic in their nature. A short theoretical overview of stochastic processes follows (Schroeder, 1991; Flandrin, 1992; Gao et al., 2007).

For understanding the stochastic processes it is essential to understand the concept of probability which is associated with random events. Often random events are presented by numbers, called a random variable. Let us denote a random variable by X, whose value depends on the outcome of random experiment ω. The probability $P(X < x)$ is denoted by $F_x(x)$ which is called the cumulative distribution function (CDF). When exist, it is usually more convenient to use its derivative $f_x(x)$, called the probability density function (PDF): $f_x(x) = dF_x(x) / dx$. Some commonly used distributions are normal or Gaussian

distributions, exponential and related distributions, binomial and related distributions and heavy tilled distributions. An important class of measures or parameters associated with CDF and PDF for a random variable is expected or mean value.
The mean or average of X can be obtained by relation:

$$E[X] = \overline{X} = \int_{-\infty}^{\infty} x dF_X(x) = \int_{-\infty}^{\infty} x f_x(x) \cdot dx \tag{8}$$

The p[th] order moment of X around zero is defined as:

$$E\left[X^p\right] = \int_{-\infty}^{\infty} x^p \cdot dF_X(x) = \int_{-\infty}^{\infty} x^p \cdot f_X(x) \cdot dx \tag{9}$$

The second moment represents the power of a random variable X.
The p[th] order central moment of X around mean value is defined as:

$$E\left[\left(x - \overline{X}\right)^p\right] = \int_{-\infty}^{\infty} \left(x - \overline{X}\right)^p \cdot f_X(x) \cdot dx \tag{10}$$

Variance of X is denoted as $Var(X)=\sigma_X^2$ and is the second order (p=2) central moment. The square rot of variance σ_X is called the standard deviation.
For a a given sample space S, a set of events E and a probability measure P, one can define a stochastic process as follows: For each sample point $\omega \in S$, we assign a time function $X(t,\omega)$. The stochastic process consists of the family of these functions. For each allowable parameter t, $X(t, \omega)$ is a random variable. For a fixed ω, $X(t, \omega)$ is a function of a time t; it is one realization of the stochastic process. There are many examples of stochastic processes. The well known examples are Markov processes and $1/f$ processes. We will focus our attention to $1/f$ processes. Stochastic process is also called a random process and for simplicity $X(t,\omega)$ is denoted as $X(t)$.
The activity of complex systems can usually be characterized by appearance of $1/f^{\beta}$ noise, a form of temporal fluctuations that has power-law power spectral density property over a wide range of frequencies. β is the power spectral exponent. A convenient framework for studying $1/f^{\beta}$ stochastic process is the self affine stochastic process $\mathbf{X} = \{X(t), t \geq 0\}$, which is defined by

$$X(\lambda \cdot t) \overset{d}{=} \lambda^H \cdot X(t), t \geq 0 \quad \lambda > 0 \quad , \quad 0 < H < 1 \tag{11}$$

where $\overset{d}{=}$ denotes equality in distribution, because of using the concept of statistical self similarity of time series. H is the Hurst parameter which is the measure of self-similarity. It can be derived, that a mean of such process is:

$$E[x(t)] = \frac{E[X(\lambda t)]}{\lambda^H} \tag{12}$$

the variance

$$Var\left[X(t)\right] = \frac{Var\left[X(\lambda t)\right]}{\lambda^{2H}} \tag{13}$$

and autocorrelation

$$R_{xx}(t,s) = \frac{R_{xx}(\lambda t, \lambda s)}{\lambda^{2H}} \tag{14}$$

By proving that $\lambda^{-H}X(\lambda t)$ and $X(t)$ have the same power spectral density, one can also prove that the power spectrum density of irregular, self similar processes has $1/f$ property:

$$S(\omega) \approx \frac{\sigma^2}{|\omega|^{\beta}} \qquad S(f) \approx \frac{\sigma^2}{|2\pi f|^{\beta}} = \frac{K \cdot \sigma^2}{f^{\beta}} \tag{15}$$

where ω is radial frequency, σ^2 is the variance, K constant value, and β the spectral exponent, which defines the slope of power spectral density over several decades.
An example of self-affine stochastic process is fractional Brownian motion process (fBm). It is nonstationary zero mean Gaussian process denoted as $B_H(t)$, characterized by scalar parameter H (Hurst parameter). The nonstrationary characteristic of fBm is evident from its covariance function structure:

$$E\left[B_H(t) \cdot B_H(s)\right] = \frac{\sigma^2}{2}\left(|t|^{2H} + |s|^{2H} + |t-s|^{2H}\right) \tag{16}$$

where E is the expectation operator. From this covariance function follows that the variance is of the type:

$$Var(B_H(t) = E\left[B_H(t)^2\right] = \sigma^2 |t|^{2H} \tag{17}$$

Although fBm process is nonstationary, it has stationary increments, which means that the probability properties of the difference process $B_H(t) - B_H(s)$ only depend on the lag t-s. It is this increment process which is self similar. The slope is in the range $1 \le \beta \le 3$. The slope β is 2 for the classic example Brownian motion.
The other example of $1/f^{\beta}$ processes is fractional Gaussian noise (fGn), with $-1 \le \beta \le 1$. fGn is stationary process. White Gaussian noise has the slope $\beta = 0$. It can be shown that Brownian motion ($\beta = 2$) is simply the integral of white noise.
It was reported, that β is related to the Hurst parameter H, which measures statistical self similar properties of signals:

$$\beta = 2 \cdot H + 1 \tag{18}$$

This is the reason for studying $1/f^{\beta}$ noise via self-affine stochastic processes. It was also shown, that for one dimensional signal, H is related to fractal dimension D by:

$$D = 2 - H \tag{19}$$

D is noninteger parameter in the range $1 < D < 2$ and H in the range $0 < H < 1$. The fractal dimension can also be used for characterizing the complexity of the stochastic signal.

Fractals are mathematical sets, which have a high degree of complex geometrical self similarity and can model many kinds of complex time series. The concept of statistical self-similarity and fractals was extended to time series to describe irregular characteristics of signals, from white noise to Brownian motion. The irregularity of a fractal curve or signal can be measured by capacity or fractal dimension D, a simplification of Hausdorf dimension, which is easier to calculate numerically. The roughness of such curve depends on D. The straight line have dimension 1. The more irregular the curve, the closer is its dimension to 2. There are many definitions and methods for calculation the fractal dimension. We will give a definition on which basis the popular box counting method.
Let S be a bounded set in $\Re^n$. The minimum number $N(s)$ of balls of radius s is needed to cover S; $N(s) \approx s^{-D}$. The fractal dimension is then defined as:

$$D = -\lim_{s \to 0} \inf \frac{\log N(s)}{\log(s)} \tag{20}$$

4. Classical statistical and Fourier analysis methods

In the past, the most common EN-analysis methods were statistical and Fourier methods. These methods assume the stationary or quasi stationary nature of processes and signals under consideration.
An early overview of different EN-data analysis methods was made in the work by R. A. Cottis (Cottis, 2001). Follows a little extended theoretical overview of classical methods (Orfanidis, 1996).

4.1 Background of statistical and Fourier methods

By analyzing random processes the statistical parameters as mean value and moments are defined by expectation operators, i.e. by statistical averaging of many realizations of stochastic process. In practice this is many times impossible and there is available only one block or array of N time signal samples. The statistical averaging is then replaced by the estimation obtained using sample or time average.
The pth moment of sample $x(n)$; n=0,..N-1 is defined as:

$$moment(p,x) = \frac{1}{N} \sum_{n=0}^{N-1} x^p(n) \tag{21}$$

The first moment is mean value $\bar{x}$. The square root of second moment gives the root mean square value x_{rsm}, which measures the amplitude of the signal. The square of x_{rsm} represents the signal power.
The pth central moment of sample is defined as:

$$central\_moment(p,x) = \frac{1}{N-1} \sum_{n=0}^{N-1} \left(x(n) - \bar{x} \right)^p \tag{22}$$

The second order central moment is the well known signal variance. The square root of variance is standard deviation, which is usually used for describing the amplitude of noise signals.

Classical spectral analysis bases on Fourier transform. The Fourier Transform of the deterministic continuous time signal $x(t)$ of duration T_1 is defined as :

$$X_{T_1}(j\omega) = \int_{T_1} x(t) \cdot e^{-j\omega \cdot t} dt \tag{23}$$

The Discrete Time Fourier Transform (DTFT) of the deterministic sampled signal with N samples is defined as:

$$X_N(\omega) = \sum_{n=0}^{N-1} x(n) \cdot e^{-j \cdot \omega \cdot n \cdot T} \tag{24}$$

where $x(n) = x(nT)$; $n=0,1,\ldots,N-1$, is according to sampling theorem sampled analog signal $x(t)$, n is time index, and T the sampling period. It can be efficient computed by the Discrete Fourier Transform (DFT) and its fast version the Fast Fourier Transform (FFT).

$$X(k) = DFT\{x(n)\} = \sum_{n=0}^{N-1} x(n) \cdot e^{-j \cdot k \cdot n \cdot 2 \cdot \pi / N}; k = 0,1...,N-1 \tag{25}$$

where n is time index and k is the frequency index . The corresponding frequency resolution is given by:

$$\Delta\omega = \omega_s / N \tag{26}$$

where ω_s is radial sampling frequency. The main shortcoming of classical Fourier transform is the averaging the features across the whole time domain.
EN signals are of stochastic nature; therefore sampled EN signals are random sequences. To obtain smooth spectra an ensemble averaging should be introduced and the spectrum calculated over autocorrelation function. The autocorrelation function of a zero mean random signal is defined as:

$$R_{xx}(k) = E[x(n+k) \cdot x(n)] \tag{27}$$

where E is the averaging or expectation operator. For stationary signals, R_{xx} do not depend on time n, but only on the relative time lag k between sequences $x(n)$ and $x(n+k)$. The power spectrum of the random signal $x(n)$ is defined as the Discrete Time Fourier Transform (DTFT) of its autocorrelation function $R_{xx}(k)$:

$$S_{xx}(\omega) = \sum_{k=-\infty}^{\infty} R_{xx}(k) \cdot e^{-j \cdot \omega \cdot n \cdot T} \tag{28}$$

where ω is the frequency in radians per sec. This power spectrum shows how the power is spread over frequencies and is also called PSD (Power Spectral Density).
EN measurements cannot often be repeated to obtain smoothed spectra by ensemble averaging. One can compute an estimate of expected or true value by so-called sample autocorrelation using time average:

$$\hat{R}_{xx}(k) = \frac{1}{N} \sum_{n=0}^{N-1-k} x(n+k) \cdot x(n) \tag{29}$$

for k=0,1,…N-1. It is known that $\hat{R}_{xx}(k)$ is an even function of the lag k. It is also well known that the results are statistical reliable only for small value of lag (5 to 10 percentages).
The DTFT of $\hat{R}_{xx}(k)$ is $\hat{S}_{xx}(\omega)$ and is referred to as periodogram spectrum and can be viewed as an estimate of power spectrum:

$$\hat{S}_{xx}(\omega) = \sum_{k=-\infty}^{\infty} \hat{R}_{xx}(k) \cdot e^{-j \cdot \omega \cdot n \cdot T} \tag{30}$$

Using the above equations we can express the periodogram also as:

$$\hat{S}_{xx}(\omega) = \frac{1}{N} \cdot \left| X_N(\omega) \right|^2 \tag{31}$$

where $X_N(\omega)$ is DFTF of N signal samples. It can be efficient computed using FFT. For wide sense stationary random signals the mean of periodogram converges to the true power spectrum $S_{xx}(\omega)$ in the limit for large N:

$$S_{xx}(\omega) = \lim_{N \to \infty} E\left[\hat{S}_{xx}(\omega) \right] = \lim_{N \to \infty} E\left[\frac{1}{N} \cdot \left| X_N(\omega) \right|^2 \right] \tag{32}$$

There are some problems with such classical Fourier spectral analysis method. To achieve high statistical reliability, very long signal sequences should be used. But long signal sequences can no longer be stationary. However, the main shortcoming is the averaging the futures over the whole time domain.
This have lead researchers to find and develop of an advanced signal analysis methods. Recently wavelet based methods for signal analysis found to be useful for nonstacionary signals. Therefore in this overview chapter we will consider wavelet-based methods for EN-signals analysis.

4.2 Overview of works using classical methods

In individual systems, the correlations between noise measurements and corrosion processes have been reported by many authors but only some can be mentioned here. The EN data for a passive system (SS 316L/Ringer's solution) and several active systems (mild steel/NaCl, brass/NaCl, Al 6061/NaCl and Al 2024/NaCl) have been analyzed in the frequency domain using power spectral density (PSD) and spectral plots, obtained from the ratio of PSD plots of the potential and current fluctuations. Comparisons of spectral noise spectra with traditional impedance spectra have been made and good agreement has been observed for all systems after trend removal (Lee & Mansfeld, 1998; Mansfeld et al., 2001). Current fluctuation during general corrosion was analyzed upon a simple model, derived on the assumption that elementary fluctuation sources are related to the fluxes of electrons that are transferred from the metal to electron-acceptor ions in solution. The number of successful electron transfers obeyed a Gaussian distribution, from which the corrosion current density and transfer coefficients were determined (Petek et al., 1997; Petek & Doleček, 2001). The time-series noise patterns of the steel in bicarbonate solution (the simulated geological environment) were transformed into frequency domain by fast Fourier transformations, and then their power spectral densities at a frequency were determined to be compared with the corrosion rate (Haruna et al., 2003). Two new indices (S_E and S_G) were

derived to evaluate pitting corrosion by dimensional analysis of three parameters of PSD, the slope of high frequency linear region, the critical frequency and the low-frequency plateau level. As shown, the value of S_E can be related to the fluctuation velocity, which can represent the pitting corrosion rate and S_G should contain some information about slow corrosion processes (Shi et al., 2008). PSD had been employed to analyze EN data associated with corrosion behavior of A291D magnesium alloy in alkaline chloride solution. Three corrosion stages, the anodic dissolution process companying with the growth, absorption and desorption of hydrogen bubbles, the development of pitting corrosion, and the inhibition process by protective MgH_2 film could be distinguished. However, the results obtained only from PSD was insufficient for better understanding the corrosion mechanism of alloy during the immersion and the wavelet transform was carried out (Zhang et al., 2007).

4.3 Our applications of classical methods

EN signal (Fig. 1) is represented as a time series, where one can easily distinguish the fluctuations but not the intensity and frequencies of fluctuations. In the paper (Planinšič & Petek, 2003) we analyzed EN corrosion signals also with some classical methods, which use correlation functions and histograms. Figure 2 shows estimated autocorrelation functions of two corrosion signals I_0 and I_2, respectively.

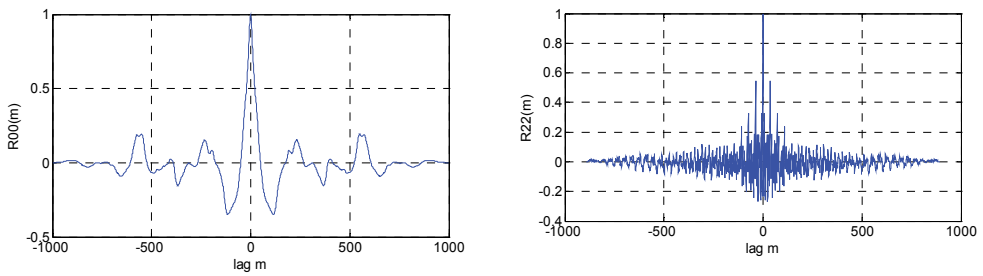


Fig. 2. Estimated autocorrelation functions of EN-signals: a) I_0; b) I_2

Noise data were transformed into frequency domain using FFT algorithm and presented as PSD in Figure 3. PSD of current noise data for pitting process exhibited two parts: a low-frequency plateau and high-frequency part, and the roll-off frequency, which is the frequency to separate the two parts of PSD. PSD plot of general corrosion can be characterized by "white noise" which is independent of frequency.

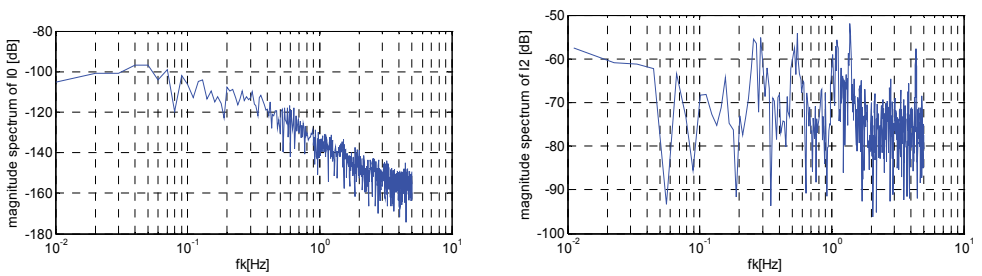


Fig. 3. Estimated power spectral density of signals I_0 (left) and I_2 (right)

Amplitude distribution was studied using normalized histograms. As demonstrated by Figure 4, a current noise amplitude distribution of general corrosion is Gaussian.

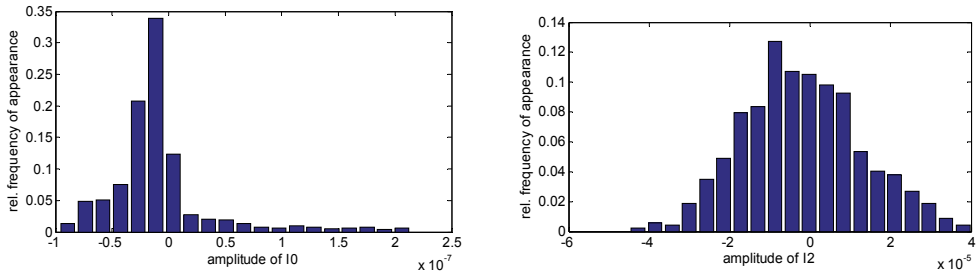


Fig. 4. Normalized histograms of signals I_0 (left) and I_2 (right)

5. Wavelet multiresolution analysis methods

The assumption of stationary behavior of corrosion processes and random signals is not always correct. Corrosion signals are a non stationary in general.

When we are interested on how signal frequency components vary with time, we should use joint time-frequency analysis. For this purpose we can use Window Fourier Transform (WTF), also called Short Time Fourier Transform (STFT) or spectrogram. It is known that the STFT can be considered as the filter bank, consists of Finite Impulse Response Filters (FIR) with equal bandwidth or equal frequency resolution. Therefore it is difficult to meet sharp localization in time and frequency simultaneously. For this reason, this technique is not always appropriate for analyzing natural signals or phenomena, where in the signal exist long duration low frequency components and short high frequency components at the same time. This problem can be elegantly solved using modern multiresolution time frequency analysis methods based on wavelets. It was shown that the Discrete Wavelet Transform can be viewed and realized as multirate filter bank with octave, also called constant Q frequency resolution.

A short theoretical overview of wavelet methods follows (Daubechies, 1992; Burrus, 1992; Fladrin,1992, 1993; Radolphe, 1994; Wornell, 1996; Mallat, 1998; Dai et al., 1994; Palawajjhalla et al., 1994).

5.1 Background of wavelet methods

Wavelets are waves which construct basis of signal decomposition in wavelet transforms, similar as trigonometric functions with different frequencies in Fourier Transform. Wavelets are scaled and shifted versions of the so called mother or primary wavelet function $\psi(t)$. Thus the family of functions is then defined as:

$$\left\{\psi_{a,b}(t)\right\}; \quad \psi_{a,b}(t) = |a|^{(-1/2)} \cdot \psi\left(\frac{t-b}{a}\right) \tag{33}$$

where parameters a and b ($a, b \in \Re; a \neq 0$) are called dilation (scaling) and translation (shifting) parameters, respectively. To be a good analyzing function, the mother wavelets must fulfill some conditions. The first is the so called ''admissibility'' condition:

$$\int_{-\infty}^{\infty} \frac{|\hat{\psi}(\omega)|^2}{|\omega|} \cdot d\omega < \infty \tag{34}$$

where $\hat{\psi}(\omega)$ is the Fourier transform of $\psi(t)$. Because the mother wavelet is absolutely inferable functions, we can show that:

$$\hat{\psi}(0) = 0 \Leftrightarrow \int_{-\infty}^{\infty} \psi(t) \cdot dt = 0 \tag{35}$$

Admissibility implies that a wavelet must be an oscillatory decaying function with zero mean. There are also additional other desirable properties for a function to be a useful wavelet, as smoothness, good time and frequency localization, number of vanishing moments. These properties suggest that wavelets are bandpass filters. $\psi(t)$ is the impulse response of filter $\hat{\psi}(\omega)$.

In the contrast with Fourier analysis where basis functions are trigonometric functions, by wavelet-based analysis different kind of mother wavelet function can be used, appropriate for particular application. There are many types of wavelet transforms according to input signals, time and scaling parameters, used wavelet functions, namely continuous, discrete, bi-orthogonal and semi-orthogonal and orthonormal bases version.

However wavelet transform can be broadly classified into Continuous Wavelet Transform (CWT) and Discrete Wavelet Transform (DWT). CWT of a function $f(t) \in L^2(\Re)$ involves the computation of scalar product. Wavelet coefficients are computed as:

$$C_{a,b} = C(a,b) = \int_{-\infty}^{\infty} f(t) \cdot \psi_{a,b}(t) \cdot dt \tag{36}$$

Discrete wavelet transform involves discretization of parameters, *a* and *b*, respectively:

$$a = a_0^{\ m} \qquad b = n \cdot b_0 \cdot a \tag{37}$$

$$\psi_{m.n}(t) = a_0^{\ -m/2} \psi(a_0^{\ -m} \cdot t - n \cdot b_0) \tag{38}$$

$$C_{m,n} = \left\langle f(t), \psi_{m,n}(t) \right\rangle = \int_{-\infty}^{\infty} f(t) \cdot \psi_{m,n}(t) \cdot dt \tag{39}$$

where $C_{m,n}$ are called discrete wavelet coefficients. Discrete wavelets $\psi_{m.n}(t)$ that satisfy the condition:

$$A \cdot \|f(t)\|^2 \le \sum_{m,n} \left| \left\langle f(t), \psi_{m,n}(t) \right\rangle \right| \le B \cdot \|f(t)\|^2 \tag{40}$$

are called frames (Daubechies,1992) and form Riesz basis. Discrete wavelets can be further classified into orthogonal, semi-orthogonal or non-orthogonal.

To obtain orthonormal basis, one can chose samples on dyadic grid (base 2):

$$a = 2^m \qquad b = n \cdot 2^m \qquad m \in \mathrm{Z}, n \in \mathrm{Z} \tag{41}$$

Orthonormal bases and orthonormal wavelet transform, play an important role in theory and practice of multiresolution analysis. The DWT can be further classified into Wavelet Series Transform (WST), when analyzed signal is continuous ($f(t)$), and into Discrete Time Wavelet Transform (DTWT), if the signal is time discrete ($f(n)$). One possibility of constructing wavelets is using a scaling function $\phi(t)$ and multiresolution analysis (Mallat, 1998). Namely, multiresolution algorithm is a natural way of constructing orthogonal wavelets. Multiresolution analysis is decomposition of square integrable functions $f(t) \in L^2(\Re)$ into closes subspaces V_j , where coarser subspace V_j is contained in finer subspace V_{j+1}: $V_j \subset V_{j+1}$. The subspaces also satisfy separation condition ($\underset{m \in Z}{\cap} V_m = \{0\}$) and condition for completeness ($\underset{m \in Z}{\cup} V_m \in L^2(\Re)$. Additionally, the functions $f(t)$ satisfy the scaling property ($f(t) \in V_m \Leftrightarrow f(2t) \in V_{m-1}$). And, there exist a scaling function in the coarsest space $\phi(t) \in V_0$, so that the family of functions $\{\phi_{m,n}\}$, $\phi_{m,n} = 2^{-m/2}\phi(2^{-m}t - n)$, form the so called Riesz basis of subspace V_m. Since $\phi(t) \in V_0 \subset V_1$ and the $\phi(2t)$ is a basis for the subspace V_1 , we can write scaling function as linear combination with the so called two scale difference equation:

$$\phi(t) = \sum_k h(k) \cdot \phi(2t - k) \tag{42}$$

where $h(k)$ is a finite sequence. It can be shown, that the frequency response of scaling function is a lowpass filter and $h(k)$ form the lowpass FIR-filter coefficients. Define W_{m-1} as the orthogonal complement of subspace V_{m-1} in V_m , than the direct sum of infinite subspaces V_j is the whole space $L^2(\Re)$. The subspace W_{m-1} contains the detail information needed to go from approximation of function at coarser to finer resolution level j. The multiresoluton analysis allows to approximate the given function $f(t)$ by $f_j(t)$ at each coarser subspace or resolution level. If $\psi(t)$ is a Riesz basis of space $W_0 \subset V_1$, we can also write:

$$\psi(t) = \sum_k g(k) \cdot \psi(2t - k) \tag{43}$$

where the finite sequence $g(k)$ form the highpass (bandpass) filter coefficients, as the frequency response of wavelet is like that of band-pass filter.
The multiresolution analysis form the theoretical basis for fast Discrete Wavelet Transform (fast DWT), using discrete signals $f(n)$, $n \in Z$, that is sampled version of $f(t)$. It was introduced by Mallat by the so called pyramidal multiresolution algorithm, where the signal $f(n)$ is decomposed into J decomposition levels. The idea of multiresolution analysis is to write a function as a limit of successive approximations, each of which is a smoothed version.
The sequences at scale j can be computed from sequences at scale j-1 by the following multirate filter or convolution, followed by subsampling by 2:

$$\begin{aligned} a_{2^j}(n) &= \left\langle f(t), \varphi_{m,n}(t) \right\rangle = \sum_k g(k - 2^j \cdot n) \cdot a_{2^{j-1}} \\ c_{2^j}(n) &= \left\langle f(t), \psi_{m,n}(t) \right\rangle = \sum_k h(k - 2^j \cdot n) \cdot a_{2^{j-1}} \end{aligned} \tag{44}$$

where $c_{2^j}(k) = c(j,k)$ are wavelet or detailed coefficients and $a_{2^j}(k) = a(j,k)$ are scaling or approximation coefficients. Data sequence $a_{2^1}(k) = a(1,k)$ at scale j=1 represents approximated or smoothed version of the original signal. The sequence $c_{2^1}(k) = c(1,k)$ at level j = 1 represents difference or detail information. The above equations together describe j^{th} level analysis filter bank. This calculation is repeated (iterated) up to scale J forming the multiresolution pyramidal algorithm; one stage is shown in Figure 5. As was mentioned, the lowpass filter is associated by scaling function and highpass filter by wavelet function. The filters for calculating the synthesis are the same by using orthogonal wavelet transform. Analysis and synthesis filters can have different length, as by using biorthogonal filter banks. However, this algorithm can be used for orthogonal and nonorthogonal wavelets.

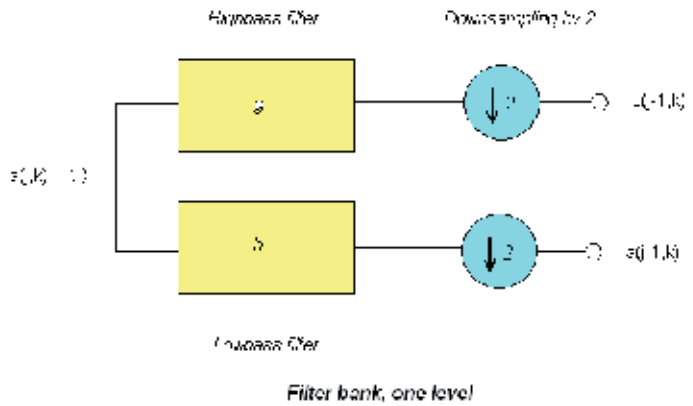


Fig. 5. One stage of analysis filter bank

For EN-signal originating from corrosion process, wavelet transform decompose it into approximation and detail signal components at different scales and locations. The wavelet transform is therefore convenient tool to analyze the self-similarity of 1/f time series.
The orthonormal wavelet transform based methods were used for estimating slope β, parameter H, and fractal dimension D (Akay,1998; Sekine, 2002; Planinšič & Petek, 2008). For orthonormal discrete wavelet decomposition the 1/f property can be replaced by the relation:

$$\sigma_j^2 = \frac{\sigma^2}{\left(2^j\right)^{\beta}} \tag{45}$$

where σ_j^2 is the variance of detail signal d_{2^j}. Then the slope β can be calculated from the plot $\log_2 \sigma_j^2$ versus level j, what can be obtained after short calculation:

$$\log_2 \sigma_j^2 = -\beta \cdot j + \log_2 \sigma^2 \tag{46}$$

5.2 Overview of works using wavelets

Wavelets have found many applications in different natural scientific disciplines, among them also in chemical engineering (Radolphe et al., 1994; Banjanin et al., 2001). The use of

wavelets to study electrochemical noise transients was reported by Aballe (Aballe et al., 1999; 2001). The wavelet analysis of electrochemical noise signals, where the signal was decomposed into wavelet-subbands was used for the characterization of pitting corrosion intensity (Smulko et al., 2002). Wharton et al. demonstrated how the wavelet variance exponent can be used to evaluate corrosion behavior for variety of stainless steels in chloride medium, i.e. be able to discriminate between various corrosion processes covering a wide range of EN signals (Wharton et al., 2003). Wavelet analysis based on the fractional energy contribution of smooth crystals and the lowest frequency detail crystal can provide information on the type and onset of corrosion (general corrosion, metastable pitting, stable pitting) in performed potentiostatic critical pitting temperature test for a superduplex stainless steel (Kim, 2007). In study of the copper anode passivation by electrochemical noise analysis using wavelet transforms it has been found that during active dissolution the electrode surface is dominated by long time scale process and the change of the position of the maximum relative energy from D7 to D8 could be an indication of future passivation (Lafront et al., 2010). It was shown, that electrochemical potential noise analysis of Cu-BTA system using wavelet transformation can be used to achieve the inhibition efficiency (Attarchi et al., 2009).

Some other authors also reported about the fractal nature of corrosion processes and corresponding electrochemical noise signals. The electrochemical potential and current noise originating from the corrosion of carbon steel in distilled water was analyzed using multifractal analysis. The multifractal spectra are found to be qualitatively different for different temporal stages of the corrosion process (Muniandy et al., 2011).

5.3 Our applications of wavelets methods

Our applications of wavelets transformation or combination with classical methods for the electrochemical current noise analysis were reported for different corrosion processes in several publications (Planinšič & Petek, 2003; 2004; 2007; 2008). For little more detailed insight the short overview of this research is as follows.

Daubechies wavelets ''db2'' were used to transform the EN signal from Fig. 1. The discrete wavelet transform (DWT) decomposition of signal into on joint time (position) and frequency (scale level) depended amplitudes are presented with color lightness in time-frequency plane in Fig. 6.

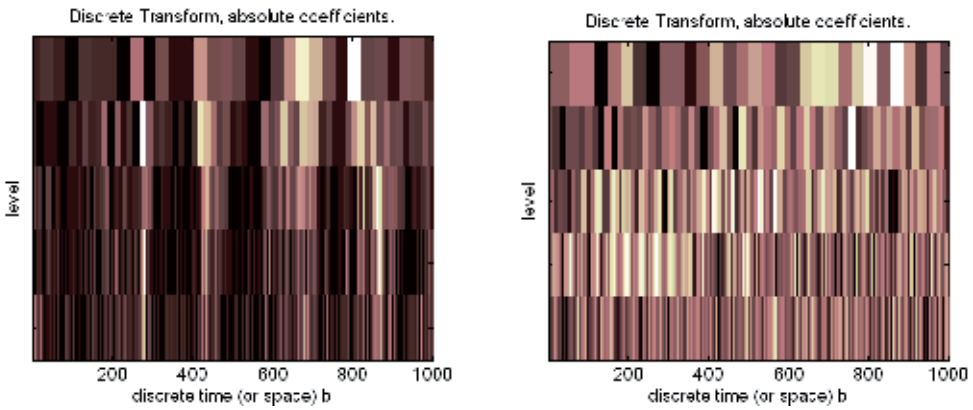


Fig. 6. Discrete wavelet transforms (DWT) of signals: a) I_0, b) I_2 (Planinšič & Petek, 2008)

Next, the DWT multiresolution decomposition of processes on 5 levels are shown in Fig. 7. The crystals from D_1 to D_5 are the details of the signal, and A_5 is the approximation of the signal. The frequency range which takes into account each series of detail coefficients, can be computed from relation $f_s / 2 / 2^j$ where f_s is the sampling frequency and j stands for the corresponding scale.

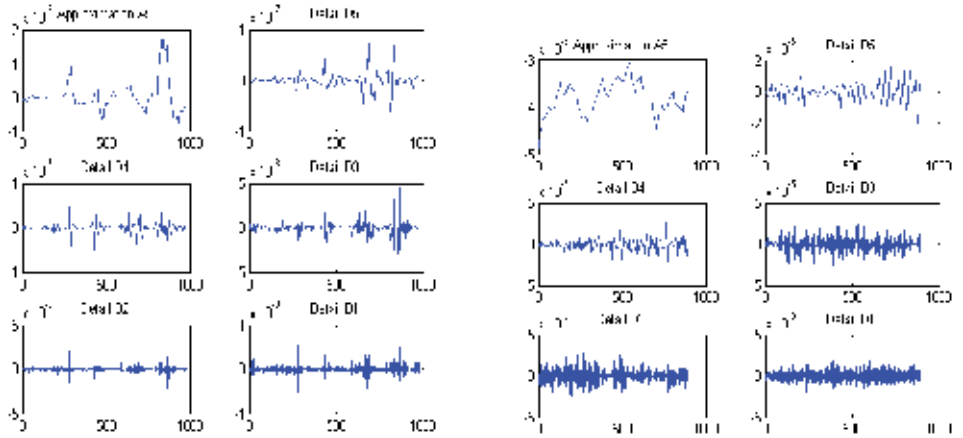


Fig. 7. Multiresolution decomposition of discrete signals: a) I_0, b) I_2 on approximation signal A_j and detail signals D_j on five levels (Planinšič & Petek, 2008).

Events with small time constants are taken into account by the fine scale coefficients, details D_1, D_2. The information dealing with larger time constant events is included in details D_4 and D_5. Therefore, these kinds of plot allow the signal to be viewed over the full time range and considering different scales, which contains information about corrosion events occurring at a determined time – scale.

Variances of details were calculated to detect the intensity of particular signal components on level j. Fig. 8 shows variances as a function of the decomposition level and also the logarithmic plot of details variances versus level j for the slope $ß$ determination.

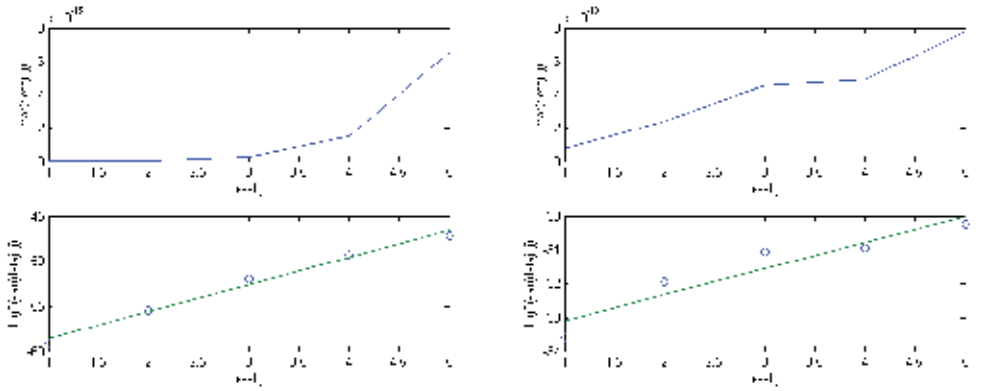


Fig. 8. Variances of details plotted as a function of the decomposition level j for two processes: a) I_0 (slope $ß$ = 3.0092) and b) I_2 (slope $ß$ = 0.7700) (Planinšič & Petek, 2008).

For the time series I_2 and I_0 were $ß$ = 0.7700 and $ß$ = 3.0092, respectively. As the value of $ß$ increases, the contributions of high-frequency components in time series are reduced. It is suggested that the events in the relatively higher frequency region may be associated with uniform corrosion. On the other hand, the events in the relatively lower frequency regions are responsible for pitting corrosion.
After short computation the value for Hurst parameter H, and fractal dimension D can be obtained:

$$D_i = 2,5 - 0,5 \cdot \beta_i \qquad H_i = 0,5 \cdot (\beta_i - 1); \quad i = 1,2. \tag{47}$$

The obtained slope $ß$ is the estimated power spectral exponent $ß$. It can be associated with the strength of persistence within a time series. The persistence defines the correlation between adjacent values within time series. If $0 \le \beta \le 1$, than persistence is weak. For the time series with β between 2 and 3, the persistence is strong. Fractional Gaussian noise with β between -1 and 1, and fractional Brownian motion, with β between 1 and 3, are considered as proper representatives of such processes. The first process is stationary and the second is non-stationary.
The obtained results indicating the presence of fractional Brownian motion in pitting corrosion with adjacent values in the time series being strongly correlated and fractional Gaussian noise in general corrosion, with adjacent values in the time series being weak correlated. The Hurst parameter in case of pitting is greater than ½, indicating also the persistence, i.e. a dependence of new values on old values. A summary of the wavelet – based fractal analysis is given in Table 1.

	β_i	H_i	D_i	process	persistence
I_0 – pitting corrosion	3.0092	1.0046	0.9954	non-stationary	strong
I_2 – general corrosion	0.7700	-0.1150	2.1150	stationary	weak

Table 1. The slope β, the Hurst parameter H, and the fractal dimension D, for two corrosion signals (Planinšič & Petek, 2008).

We proposed also a new way for determination of persistence nature of the electrochemical noise on the basis of correlation coefficients between original signal and details $R(I_i,D_j)$, Table 2. The pitting corrosion is positively correlated with long memory effect. Increasing correlation of signal I_2 to the D_2 detail and then decreasing to D_5, indicates weak persistence and short memory effect of general corrosion processes.

	$R(I_0,D_j)$	$R(I_2,D_j)$
D_1	0.0188	0.4196
D_2	0.0482	0.5303
D_3	0.1128	0.5097
D_4	0.2048	0.3648
D_5	0.3103	0.3138

Table 2. Correlation coefficients between original signal and details, $R(I_i,D_j)$, for two decomposed corrosion signals (Planinšič & Petek, 2008).

Correlation coefficients between successive details $R(D_j,D_{j+1})$ for two decomposed corrosion signals (Table 3) are all zero on the basis of analysis procedure.

	I_0	I_2
$R(D_1,D_2)$	0.0006	0.0006
$R(D_2,D_3)$	0.0000	0.0004
$R(D_2,D_3)$	0.0001	0.0004
$R(D_4,D_5)$	-0.0012	-0.0120

Table 3. Correlation coefficients between successive details, $R(D_j,D_{j+1})$, for two decomposed corrosion signals (Planinšič & Petek, 2008).

Additionally, DWT with 3-decomposition levels was made using different kinds of wavelet functions, from Daubechie's fractal-like wavelet "db2" (Massopust, 1994) to smoother wavelet functions, as Daubechie's wavelet function "db5" and symmetrical Coiflet wavelet function "coif5". After decomposition the coding gain (c_g) was calculated from variances of decomposed sub-bands:

$$c_g = \sum_{j=1}^{J}(\sigma_j^2 \cdot \alpha_j) / \prod_{j=1}^{J}(\sigma_j^2)^{\alpha_j} \tag{48}$$

where σ_j^2 are variances of sub-bands and α_j are the relative length of sub-band sequences and J the number of sub-bands. The coding gain is a measure of spectral flatness. For (uniform) white noise it has the value 1. Also the Shannon's entropy was calculated and can be viewed as a measure of signal complexity. The numerical experimental results are collected in Table 4.

	c_g, I_2	c_g, white noise	entropy, I_2	entropy, white noise
db2	1.3486	1.0090	4.5384	4.9304
db5	1.3411	1.0076	4.5384	4.9304
coif5	1.3411	1.0116	4.5384	4.9304

Table 4. Coding gain and Shannon's entropy obtained with different wavelet functions

The chose of different wavelets did not influence on the obtained c_g and entropy. We found also, that the coding gain and entropy can be used as an additional parameter to distinguish the corrosion processes.

To study the approximation properties of DWT's using different wavelet basis functions, a synthesis (inverse DWT) to different approximation levels was made. We expected better results with ''db2'' assuming the fractal–like shapes of EN-signals. However, no significant differences were found, a smaller approximation error was obtained even using smoother wavelets, what confirms the approximation theory.

6. Conclusion

The most attractive prospective benefit of EN measurement is the ability to obtain information about the type of corrosion that is occurring, but there is much less agreement

about the optimum analysis method for obtaining such information (Cottis, 2006). Wavelet transform has been developed over a number of years and only recently has been applied to electrochemical noise analysis. The main advantage of wavelet analysis of EN is the detection of transients which are localized in both the time and frequency domain and shows promise to be discriminatory for the intensity as well as the type of corrosion.

7. References

Aballe, A., Bethencount, M., Botana, F. J. & Markos, M. (1999). Using wavelets transform in the analysis of electrochemical noise data, *Electrochimica Acta,* Vol.44, No.26, (September 1999), pp. 4805-4816, ISSN: 0013-4686

Aballe, A., Bethencount, M., Botana, F. J., Markos, M. & Sanchez-Amaya, J. M. (2001). Use of wavelets to study electrochemical noise transients, *Electrochimica Acta,* Vol. 46, No. 15 (April 2001), pp. 2353- 2361, ISSN: 0013-4686

Akay, M., Mulder, E. J. H. (1998). Effects of Maternal Alcohol Intake on Fractal Properties in Human Fetal Breathing Dynamics, *IEEE Transaction on Biomedical Engineering*, Vol. 45, No. 9, (September 1998), pp.1097 – 1103, ISSN: 0018-9294

Attarchi, M., Roshan, M. S., Norouzi, S. & Sadrnezhad, S. K. (2009). Electrochemical potential noise analysis of Cu-BTA system using wavelet transformation, *Journal of Electroanalytical Chemistry,* Vol. 633, No. 1, (2009), pp. 240-245, ISSN: 1572-6657

Banjanin, B., Gergič, B., Planinšič,P. & Čučej, Ž. (2001). Entropy-threshold method for best basis selection, *Image and Vision Computing* Elsevier, Vol. 19, No. 7, (May 2001), pp. 477- 484, ISSN: 0262-8856

Burrus, C. S. (1992). *Introduction to wavelets and wavelet transforms*, *A primer*. Prentice Hall, ISBN: 0-13-489600-9, Upper Saddle River, New Jersey, USA

Cottis, R. A. (2001). Interpretation of Electrochemical Noise Data, *Corrosion*, Vol. 57, No.3, (2001), pp.265-285, ISSN: 0010-9312

Cottis, R.A. (2006). Sources of Electrochemical Noise in Corroding Systems, *Russian Journal of Electrochemistry*, Vol. 42, No. 5, 2007, pp. 497-505, ISSN: 1023-1935

Dai, X., Joseph, B. & Motard, R. L. (1994). Introduction to Wavelet Transform and Time-Frequency Analysis, In: *Wavelet Applications in Chemical Engineering*, Motard, R. L. &Joseph, B., (Eds.), 1-32, Cluwer Academic Publisher, ISBN: 0-7923-9461-5. Norwell, Massachusetts, USA.

Daubechies, I. (1992). *Ten Lectures on Wavelets*, Siam, ISBN: 0-89871-274-2, Philadelphia, Pennsylvania, USA

Farge, M. (1993). *Wavelets, Fractals, and Fourier Transforms*, J. C. Hunt, J. C. Vassilicos (Eds.), Clarendon Press, ISBN: 0-19-853647-X, Oxford, USA

Flandrin, P. (1992). Wavelet analysis and synthesis of fractional Brownian motion, *IEEE Trans. Information theory*, Vol. 38, No.2, Part 2, Mar. 1992, pp.904-909, ISSN: 0018-9448

Flandrin, P. (1993). Fractional Brownian Motion and Wavelets, In: *Wavelets, Fractals, and Fourier Transforms*, Farge, M., Hunt, J. C. R. & Vassilicos, J. C., (Eds.), 109-142, Clarendon Press, ISBN: 0-19-853647-X, Oxford, USA

Gao, Y., Cao, Y., Tung, W.-W. & Hu, J. (2007). *Multiscale analysis of complex time series (Integration of chaos and random fractal Theory, and Beyond)*, John Wiley & Sons, ISBN: 978-0-471-65470-4, New Jersey, USA

Haruna, T., Morikawa, Y., Fujimoto, S. & Shibata, T. (2003). Electrochemical noise analysis for estimation of corrosion rate of carbon steel in bicarbonate solution, *Corrosion Science*, Vol. 45, No. 9, pp. 2093-2104, ISSN: 0010-938X

Kim, J.J. (2007). Wavelet analysis of potentiostatic electrochemical noise, *Materials Letters*, Vol. 61, No. 18, 2007, pp. 4000-4002, ISSN: 0167-577X

Lafront, A.M., Safizadeh, F., Ghali, E.& Houlach, G.(2010). Study of the copper anode Passivation by electrochemical noise analysis using spectral and wavelet

transforms, *Electrochimica Acta* , Vol. 55, No. 22, 2010, pp. 2505-2512, ISSN: 0013-4686

Lee, C. C. & Mansfeld, F. (1998). Analysis of electrochemical noise data for a passive system in the frequency domain, *Corrosion Science*, Vol. 40, No. 6, pp. 959-962, ISSN: 0010-938X

Mallat, S. (1998). *A wavelet tour of signal processing*, Academic Press, 1998. ISBN :0-12-466606-X, San Diego, California, USA

Mansfeld, F., Sun, Z: & Hsu, C.H. (2001). Electrochemical noise analysis (ENA) for active and passive systems in chloride media, *Electrochimica Acta* , Vol. 46, No. 24-25, 2001, pp. 3651-3664, ISSN: 0013-4686

Massopust, P. R. (1994). *Fractal Functions, Fractal Surfaces, and Wavelets*, Academic Press, ISBN: 0-12-478840-8, San Diego, California, USA

Mathworks Inc.(2007), Matlab® 7.5.0 (R2007b), 2007

Muniandy, S.V., Chew, W.X. & Kan, C.S. (2011). Multifractal modeling of electrochemical noise of corrosion of carbon steel, *Corrosion Science*, Vol. 53, No.1, pp. 188-200, ISSN: 0010-938X

Orfanidis, S.J. (1996). *Introduction to Signal Processing*, Prentice Hall, ISBN 0-13-209172-0, Upper Saddle River, New Jersey, USA

Palavajjhalla, S., Motard, R. L & Joseph (1994), B. Computational Aspects of Wavelets and Wavelet Transforms, In: *Wavelet Applications in Chemical Engineering,* Motard, R. L. &Joseph, B., (Eds.), 33-83, Cluwer Academic Publisher, ISBN: 0-7923-9461-5. Norwell, Massachusetts, USA.

Petek, A., Doleček, V. & Vlachy, V. (1997). Stochastic Analysis of Current Fluctuations During General Corrosion of Stainless Steel in Sulfuric Acid, *Corrosion*, Vol. 53, No.12, (1997), pp.928-934, ISSN: 0010-9312

Petek, A: & Doleček, V. (2001). Interpretation of current noise generation by a simple model, *Materials and Corrosion*, Vol. 52, No.6, (2001), pp.426-429, ISSN: 1521-4176 Planinšič, P., Gergič, B., Gleich, D. & Čučej, Ž.(2001). Fuzzy control of subband codded image quality using standard and fuzzy quality measure, *Optical Engineering*, Vol.8, No.40, (August 2001), pp. 1529-1544, ISSN 0091-3286

Planinšič, P. & Petek, A. (2003). Analysis of electrochemical noise signals with classical methods and methods based on fractal-like wavelets, *Proceedings IEEE ICIT 2003*, 871-876, ISBN: 0-7803-7853-9, Maribor, Slovenia, December 10-12, 2003

Planinšič, P. & Petek, A. (2004). Analysis of electrochemical noise signals using fractal-like function wavelets, In: *Interdisciplinary applications of fractal and chaos theory,* R. Dobrescu, C. Vasilescu, (Ed.), 322-334, Editura Academica Romane, ISBN 937-27-1070-5, Bucaresti, Romania

Planinšič, P. & Petek, A. (2007). Electrochemical noise signals discrimination using wavelet-based fractal analysis, *Proceeding of 3rd international symposium on Interdisciplinary*

approaches in fractal analysis IAFA 2007, pp. 322-334, ISBN 842-6508-23-3, Printech, Bucaresti, Romania, May 23-25, 2007

Planinšič, P. & Petek, A. (2008). Characterization of corrosion processes by current noise - based fractal and correlation analysis, *Electrochimica Acta,* Vol. 53, No. 16, June 2008), pp. 5206-5214, ISSN: 0013-4686

Sekine, M., Tamura, T., Akay, M., Fujimuto, T., Togawa, T. & Fukuji, Y. (2002). Discrimination of Walking Paterns Using Wavelet-Based Fractal Analysis, *IEEE Transaction on Neural Sysems and rehabilitation Engineering,* Vol.6, No.3, (September 2002), pp. 188 -196, ISSN: 1534-4320

Shi, Y.Y., Zhang, Z., Cao, F.H. & Zhang J.Q. (2008). Dimensional analysis applied to pitting corrosion measurements, *Electrochimica Acta* , Vol. 53, No.6, 2008, pp. 2688-2698, ISSN: 0013-4686

Schroeder, M. (1991). *Fractals, Chaos, Power Laws (Minutes from an Infinite Paradise)*, W. H. Freeman and Company, ISBN: 0716721368, New York, USA

Smulko J., Darowicki, K. & Zielinski, A. (2002). Pitting Corrosion in Steel and Electrochemical Noise Intensity, *Electrochemistry Communication*, Vol. 4, No. 5 (May 2002), pp 388-391(4). ISSN: 1388-2481

Wharton, A., Woo, R. J. K. & Mellor, B. G. (2003). Wavelet analysis of electrochemical noise measurements during corrosion of austenitic and superduplex stainless steels in chloride media, *Corrosion Science*, Vol. 45, No. 1, (January 2003), pp 97-122. ISSN: 0010-938X

Wornell, G. W. (1996). *Signal Processing with Fractals, A Wavelet–Based Approach*, Prentice Hall PTR, ISBN 013120999X, Upper Saddle River, New Jersey, USA

Zhang, T., Shao, Y., Meng, G. & Wang, F. (2007). Electrochemical noise analysis of the corrosion of AZ91D magnesium alloy in alkaline chloride solution, *Electrochimica Acta* , Vol. 53, No.17, 2007, pp. 561-568, ISSN: 0013-4686

12

Applications of Discrete Wavelet Transform in Optical Fibre Sensing

Allan C. L. Wong and Gang-Ding Peng
School of Electrical Engineering and Telecommunications
University of New South Wales
Australia

1. Introduction

This chapter presents a comprehensive review of recent advances in the applications of discrete wavelet transform (DWT) in optical fibre sensing. DWT, like Fourier transform (FT), is a versatile and powerful mathematical tool to process, extract and analyse data. In fibre sensing, DWT is particularly useful in the demodulation, demultiplexing and denoising of sensor data, as well as in the detection, extraction and interpretation of measurand-induced change from an acquired sensor signal.

Both continuous and discrete has found a wide variety of applications in fibre sensing, and has been extensively used for fibre Bragg gratings (FBGs) and interferometric sensors. For example, Chan et al. (2007, 2010) used wavelet transform (both continuous and discrete) in reducing noise and increasing wavelength detection accuracy of FBGs; Jones (2000a, 2000b) used in the edge detection and crack detection of FBGs embedded in some structures; Staszewski et al. (1997) and Bang & Kim (2010) used for the detection of acoustic wave induced by impact and defect in composite plates; Gangopadhyay et al. (2005, 2006) used to extract and analyse fibre Fabry-Perot interferometer signals; Lamela-Rivera et al. (2003) used in the detection of partial discharges from high-power transformers; Tomic et al. (2010) in pressure sensing; and Wang et al. (2001) in the health monitoring of ship hull structure.

This chapter begins with a brief introduction on the applications of DWT in fibre sensing. This follows by the principles and approaches of using DWT. Several important and fundamental formulations and concepts, such as the use of DWT in signal demodulation, demultiplexing and noise reduction will be presented. Next, four exemplary application cases of using DWT in fibre sensing are presented. More specifically, representative topics with regard to sensor signal analysis, signal demodulation, noise reduction and demultiplexing of multiplexed sensor systems are described in details. Finally, conclusions to summarise the chapter is given.

2. Principles and approaches

In fibre sensing, there are two areas of signal analysis that DWT has been widely employed, namely sensor signal demodulation and noise reduction. The former is accomplished by the theory of Multiresolution analysis (MRA) and the latter by the wavelet denoising. This section will describe these two DWT techniques qualitatively, and all of the subsequent examples are essentially based on them.

2.1 Sensor signal demodulation

The DWT demodulation technique presented here is used to demodulate and demultiplex many types of multiplexed sensor signals, and it works best with periodic sinusoidal functions, which is often the case for most interferometric signals. Excellent references in the theory and properties of wavelets can be found in (Daubechies, 1992; Mallet, 1998; Sidney Burrus et al., 1998; Vetterli & Kovacevic, 1995). In this section, only the concepts and results that are essential to the understanding of this demodulation technique are included. The principle of operation is based on the theory of MRA, a representation of DWT from a digital signal processing perspective (Mallet, 1989). In the MRA, the sensor signal can be represented as a wavelet series,

$$I(\lambda) = \sum_m c_J(m)\phi_{J,m}(\lambda) + \sum_m \sum_{j=1}^{J} d_j(m)\psi_{j,m}(\lambda)\,, \tag{2.1}$$

where $j, m \in Z$, and the integer J sets the highest decomposition level.

$$c_j(m) = \left\langle I(\lambda), \phi_{j,m}(\lambda) \right\rangle = \int I(\lambda)\varphi^*_{j,m}(\lambda)d\lambda \tag{2.2}$$

are the jth-level DWT approximation coefficients, and

$$d_j(m) = \left\langle I(\lambda), \psi_{j,m}(\lambda) \right\rangle = \int I(\lambda)\psi^*_{j,m}(\lambda)d\lambda \tag{2.3}$$

are the jth-level DWT detail coefficients. * denotes the operation of complex conjugation. $\phi_{j,m}(\lambda) = 2^{j/2}\phi(2^j\lambda - m)$ is the scaling function, and $\psi_{j,m}(\lambda) = 2^{j/2}\psi(2^j\lambda - m)$ is the wavelet function. The DWT coefficients can be computed by a multistage two-channel quadrature mirror filter bank (QMFB). This is formed by the scaling function acting as a low-pass filter, i.e., $\phi(\lambda) = \sum_m h(m)\sqrt{2}\phi(2\lambda - m)$, where $h(m)$, $m \in Z$, are the low-pass filter coefficients. The complementary wavelet function acts as a high-pass filter, i.e., $\psi(\lambda) = \sum_m g(m)\sqrt{2}\phi(2\lambda - m)$, where $g(m)$ are the high-pass filter coefficients. The filter coefficients are related by $g(m) = (-1)^m h(1-m)$. In other words, in the QMFB the scaling and wavelet functions simultaneously perform the low-pass and high-pass filtering on the sensor signal. At the jth-stage of the QMFB, the DWT approximation coefficients are given by,

$$A_j = c_j(m) = \sum_k h(k-2m)c_{j-1}(k)\,, \tag{2.4}$$

and the DWT detail coefficients are given by,

$$D_j = d_j(m) = \sum_k g(k-2m)c_{j-1}(k)\,. \tag{2.5}$$

Then, Eq. (2.1) can be expressed in the form,

$$I(\lambda) = \sum_m A_J + \sum_m \sum_{j=1}^{J} D_j\,. \tag{2.6}$$

From this representation, the jth-level DWT coefficients can be computed by convolving the QMFB with the $(j-1)$th-level approximation coefficients, i.e., $A_{j-1} = A_j + D_j$. The schematic

diagram of the multistage decomposition operation using the QMFB is shown in Fig. 2.1. By repeating this decomposition in cascade using the most recent A_j as inputs, we can compute the DWT coefficients to the resolution level of interest. The usefulness of the MRA representation is that a signal can be decomposed into different levels of wavelet coefficients according to the frequency components that comprised the whole signal. The higher the wavelet levels, the lower the frequency components of the signal remain. Fig. 2.2 shows graphically the frequency representation of a signal being decomposed into various wavelet levels. f_c is the centre frequency of the whole spectral bandwidth of the signal. If the multiplexed sensor signal is tailored in such a way that each sensor has a different (signal) frequency, after this cascaded operation, their original sensor signals will appear on different wavelet levels.

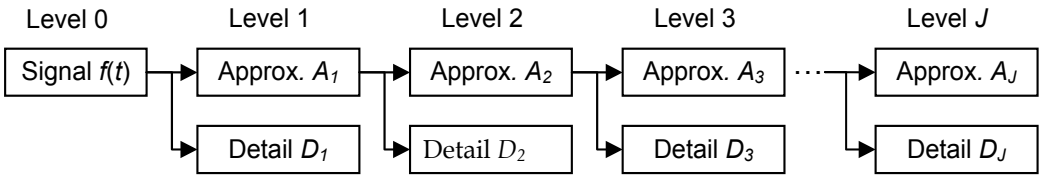


Fig. 2.1. Multistage decomposition of a signal using QMFB.

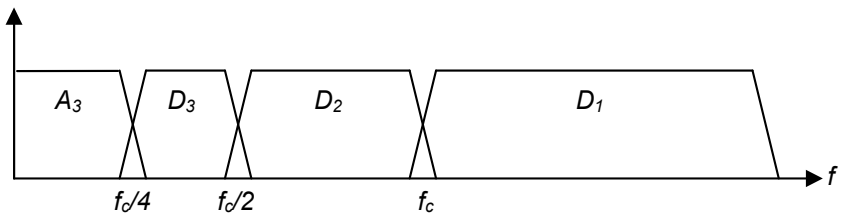


Fig. 2.2. Frequency representation of a signal decomposed into various wavelet levels.

2.2 Wavelet denoising

Wavelet denoising (Donoho, 1994) is a nonlinear noise reduction technique based on the DWT to remove/reduce the noise in the signal while preserving the overall signal features. The generic denoising algorithm is depicted in Fig. 2.3. In principle, wavelet denoising attempts to decompose a signal using the DWT to obtain the wavelet coefficients, and then apply a thresholding method or a shrinkage rule to each wavelet coefficient. The threshold can be obtained by using some risk estimators or by empirically finding a value. The method then either keeps or shrinks all wavelet coefficients that are above the threshold, and suppresses all those below the threshold. Then, the signal is reconstructed by taking the inverse DWT with the noisy part being removed. Here, two denoising techniques that are frequently employed in fibre sensing are discussed, namely hard thresholding (HT) and block-level thresholding (BLT). The former is a simple and effective method, while the latter is automatically incorporated with the DWT demodulation technique, making it a very attractive and convenient denoising method.

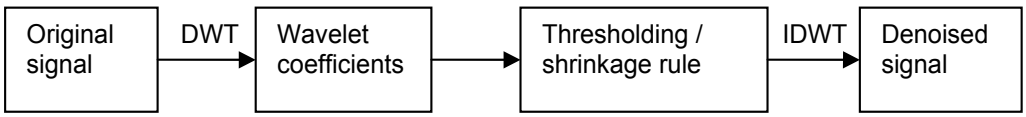


Fig. 2.3. Generic wavelet denoising algorithm.

2.2.1 Hard thresholding

When a sensor signal is acquired from a measurement system, the signal consists of broadband noise caused by the quality of the sensor, as well as the quantisation and finite resolution errors of the equipment. It is known that broadband noise is difficult to remove using conventional filtering without altering the signal. But this can be done by the HT wavelet denoising. Let the sensor signal intensity be represented as a wavelet series,

$$I(\lambda) = \sum_{m,n=-\infty}^{\infty} \alpha_{m,n} \psi_{m,n}(\lambda) , \tag{2.7}$$

where $\psi_{m,n}(\lambda) = 2^{m/2} \psi(2^m \lambda - n)$, $\psi \in L^2(\Re)$, $m, n \in Z$ are the orthonormal wavelet basis functions. By taking the DWT, the wavelet coefficients are,

$$\alpha_{m,n} = \left\langle \psi_{m,n}(\lambda), I(\lambda) \right\rangle = \int_{-\infty}^{\infty} \psi^*_{m,n}(\lambda) I(\lambda) d\lambda . \tag{2.8}$$

The HT then sets a threshold value, *h*, to the wavelet coefficients, such that any coefficients below the threshold are suppressed, whereas the coefficients above the threshold are retained. Since the noise components are usually of low magnitude, the threshold can be set to a value high enough to eliminate the noisy coefficients, and low enough to retain useful signal coefficients. The HT operation can be represented by (Wong et al., 2005),

$$\xi_h(\alpha_{m,n}) = \hat{\alpha}_{m,n} = \begin{cases} \alpha_{m,n}, & |\alpha_{m,n}| > h \\ 0, & |\alpha_{m,n}| \le h \end{cases} . \tag{2.9}$$

Since the HT is applied to the entire spectral range of the sensor signal, the broadband noise can be effectively removed. Once the noisy coefficients are removed, the signal can be reconstructed by taking the inverse DWT,

$$I'(\lambda) = \sum_{m,n=-\infty}^{\infty} \hat{\alpha}_{m,n} \psi_{m,n}(\lambda) . \tag{2.10}$$

From another perspective, suppose a sensor signal consists of an ideal noise-free part *f(λ)* and a noisy part *σ(λ)* with noise level *p*, i.e., $I(\lambda) = f(\lambda) + p\sigma(\lambda)$. If we set *h* = *p*, i.e., setting the threshold equal to the noise level, then the HT will suppress the noisy coefficients to zero while retaining the noise-free coefficients unaffected. The ideal case is to have $|I(\lambda) - f(\lambda)| \to 0$, i.e., the reconstructed signal is as close as practicable to the ideal, noise-free signal.

2.2.2 Block-level thresholding

The BLT can be considered as a modified version of the block-thresholding method (Wong et al., 2006b). In the block-thresholding method, instead of applying a threshold to each wavelet coefficient term-by-term (as is for HT), the threshold is applied to a block of wavelet coefficients, with the threshold value determined by calculating and minimising the risk using the block James-Stein estimator. On the other hand, the BLT method attempts to set an entire level of wavelet coefficients as a block, such that the entire level is either retained or discarded. There is no need to estimate the risk or to find a value empirically in order to obtain the threshold value. For example, if the sensor signal has Gaussian white noise associated with it, such that the noise components spanned the whole frequency range of the signal. From the theory of MRA, the 1st-level detail coefficients account for the upper-half of the whole frequency range of the original signal, the 2nd-level account for the upper-half frequency range of the 1st-level approximation coefficients, and so on (c.f. Fig. 2.2). Therefore, the first two levels of detail coefficients cover 75 % of the whole frequency range of the sensor signal, and the noise components within this range will be removed if we discard the detail coefficients at these levels. For the signal expressed in the form of Eq. (2.6) with levels $j = 1, 2, ..., J$, the BLT operation for a chosen level of approximation coefficients ($j = J_A$) used to demodulate the sensor signal is given by,

$$BLT(j = J_A) \Rightarrow I(\lambda) = \begin{cases} \sum_m A_{J_A}, & \text{for } j = J_A \\ 0, & \text{otherwise} \end{cases}, \tag{2.11}$$

Similarly, for a chosen level of detail coefficients ($j = J_D$) used to demodulate the sensor signal, the BLT operation is given by,

$$BLT(j = J_D) \Rightarrow I(\lambda) = \begin{cases} \sum_m D_{J_D}, & \text{for } j = J_D \\ 0, & \text{otherwise} \end{cases}. \tag{2.12}$$

In other words, the BLT discards all the levels of detail coefficients that are at (for approximation coefficients) or below (for detail coefficients) the desired level chosen to be used for sensor signal demodulation. Additionally, the BLT denoising technique can be applied as a standalone technique for any type of signal. In a more general form, the BLT for a chosen level j_{BLT}, where $1 \le j_{BLT} \le J$, is given by (Wong et al., 2007a),

$$BLT(j_{BLT}) \Rightarrow I(\lambda) = \begin{cases} \sum_m A_J + \sum_m \sum_{j=j_{BLT}+1}^{J} D_j, & \text{for } j_{BLT} < j \le J \\ \sum_m A_J, & \text{for } j = j_{BLT} \\ 0, & \text{for } 1 \le j < j_{BLT} \end{cases}. \tag{2.13}$$

An inverse DWT is then applied to the wavelet coefficients to reconstruct the signal with the noisy part removed. The choice of j_{BLT} requires the signal to preserve its characteristic features for demodulation purposes, while being able to maximally remove the broadband noise components associated with it.

In the subsequent sections, four exemplary applications, which are the research works conducted by the authors, will be presented to demonstrate the power and versatility of the DWT.

3. Application case 1 – simultaneous multi-sensor signal demodulation

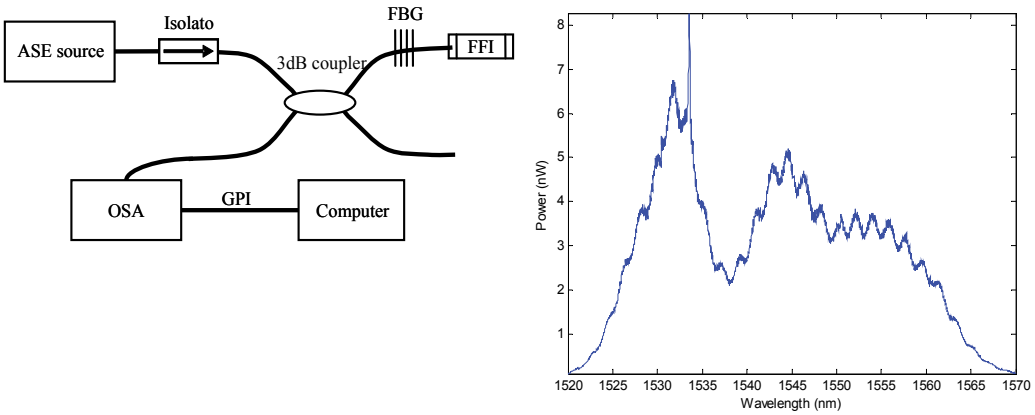
In the first example, a new simultaneous demodulation technique was proposed for a multiplexed fibre Fizeau interferometer (FFI) and FBG sensor system, based on the DWT technique described in the last section (Wong et al., 2006a, 2006b). In relation to this demodulation technique, the BLT denoising technique was applied automatically and simultaneously to reduce the noise associated with the sensor signal. It is known that FFIs and FBGs are two of the most widely studied types of fibre-optic sensors (Lee, 2003). They have distinct measurand-induced responses and dynamic ranges, and hence, when they are combined together, they can measure different measurands simultaneously. This demodulation technique outperforms currently reported techniques, and the key advantages are: (i) it overcomes the disadvantages associated with interferometric methods; (ii) it only requires a simple setup to interrogate and multiplex the sensors. All the data acquisition, signal processing, and calculations are carried out digitally by a computer program, and no complicated demodulating electronics are needed; (iii) it is a completely passive system that does not require any mechanical moving parts or active modulation, making it especially suited for continuous long-term quasi-static sensing; (iv) and it automatically reduces the signal noise through the BLT, without the need of any additional filtering techniques.

The FFI and FBG sensors are multiplexed using a hybrid of spatial-frequency-division multiplexing (SFDM) and wavelength-division multiplexing (WDM). Specifically, SFDM is used to multiplex the FFI sensors, in that each sensor produces a sinusoidal interference pattern with the spatial frequency specified by the cavity length. The multiplexed FFIs signal can then be demodulated using the FT peak detection method, which will be described later. For FBG sensors, WDM is employed in order to take the advantage of the wavelength-encoded nature. In the wavelength domain, each FBG has a narrowband Bragg peak in the reflection spectrum. As a result, the multiplexed signal is of two extremes: the FBG signal (i.e., the main peak) is localised with compact support in the original (wavelength) domain, but spans across the entire spectral bandwidth in the dual (spatial-frequency domain). The opposite case applies to the FFI signal. That is, in the original domain the signal is spanned along the spectral bandwidth of the light source. But in the dual domain, the signal is localised with compact support, and is shown as a sharp peak. In this case, the multiplexed signal can be easily separated by the DWT technique and individual sensor signals will appear on two distinct wavelet levels.

Fig. 3.1 (a) shows the experimental setup for the sensor system, An amplified spontaneous emission (ASE) light source was connected to a 3-dB fibre coupler to illuminate the FFI and FBG sensors that are serially multiplexed. This reflected signal was acquired by an optical spectrum analyser (OSA) and transferred to a computer. The multiplexed signal is the superposition of the two individual sensor signals. Assuming no insertion loss, for N FFIs and M FBGs, the measured spectrum is of the form (Wong et al., 2006b):

$$\begin{aligned} I(\lambda) &= A(\lambda)\Big(\sum I(\lambda)_{FFI} + \sum I(\lambda)_{FBG}\Big) \\ &= A(\lambda)\left\{\sum_{k=1}^{N} r_{1,k} r_{2,k}\left[1 + V_k S(\lambda)\cos(4\pi d_k / \lambda)\right] + \sum_{l=1}^{M} R_{B,l}\exp\left[-(\lambda - \lambda_{B,l})^2 / 2\delta\lambda_{B,l}^2\right]\right\}, \end{aligned} \tag{3.1}$$

where $A(\lambda)$ is the incident light source intensity, r_i are the reflection coefficients of the two fibre ends that form the FFI cavity, d is the cavity length, V is the fringe visibility, $S(\lambda)$ is the erbium-doped fibre (EDF) amplifier spectral profile, λ_B is the centre Bragg wavelength, R_B is the peak reflectivity of the FBG, and $\delta\lambda_B$ is a parameter related to the FBG bandwidth. A typical multiplexed signal for $N = M = 1$ is shown in Fig. 3.1 (b). From the figure, the FBG can be identified by the distinct narrowband peak (around 1533 nm), whereas the FFI produced a sinusoidal interference pattern that spanned the light source bandwidth, and superimposed on its spectral profile.



(a) (b)

Fig. 3.1. (a) Experimental setup for the multiplexed FFI and FBG sensor system; and (b) a typical multiplexed sensor signal.

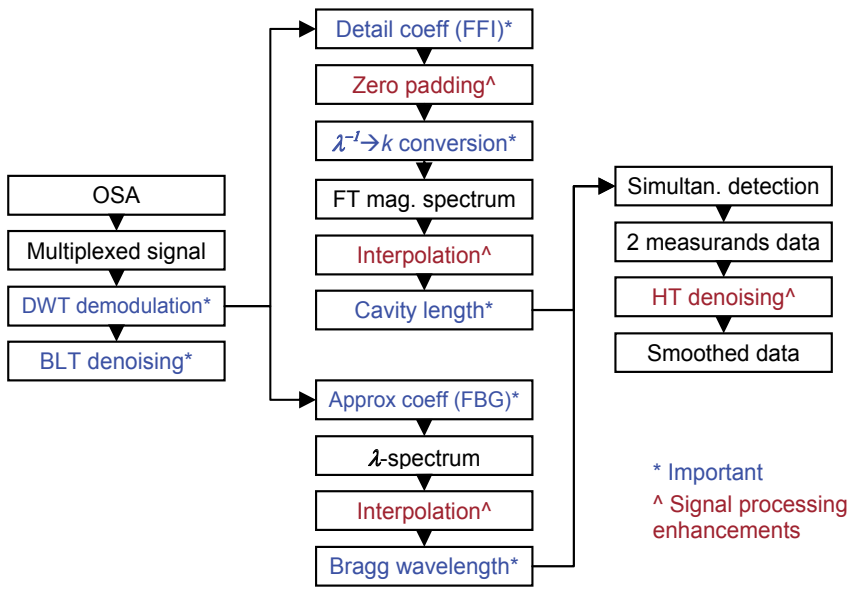


Fig. 3.2. Simultaneous demodulation algorithm with wavelet denoising.

The simultaneous demodulation algorithm is depicted in Fig. 3.2. For a multiplexed signal retrieved from the OSA, the DWT was applied to decompose it into multilevel wavelet coefficients. In parallel to the demodulation, the BLT was applied simultaneously such that the decomposed signal was denoised. The discrete Meyer wavelet was chosen to be the kernel function for DWT because the number of vanishing moments and regularity are suitable for analysing this type of multiplexed signal. For the FBG sensor, since the received signal from the OSA was a function of wavelength, the Bragg wavelength can be determined directly from a level of approximation coefficients, which effectively gave the denoised version of the original spectrum. The BLT technique automatically discarded the wavelet levels that were below the chosen level of approximation coefficients for the FBG signal, and thus reduced the signal noise without explicitly applying any additional filtering technique. Then, an interpolation utilising the piecewise-continuous cubic-spline function was applied to the Bragg peak to increase the wavelength resolution. Next, the cavity length of the FFI was determined. After taking the DWT, the detail coefficients were indeed the original interference pattern of the FFI, which was then demodulated using the FT peak detection method. First, the signal was Fourier transformed and each sensor had its own peak in the magnitude spectrum. The important step was to convert the variables of the detail coefficients from wavelength to wavenumber. In doing so, the FT dual-domain variable was related to the sensor cavity length. The cavity length can then be obtained directly from the location of the amplitude peak in the magnitude spectrum (Wong et al., 2005). Once the characteristic change of both sensors are known, i.e., Bragg wavelength of the FBG and cavity length of the FFI, simultaneous measurement of two measurands can be achieved by monitoring the changes of both sensors. In order to do so, the measurands-induced responses, i.e., the elements in the sensitivity matrix need to be determined. The HT denoising technique (described in previous section) can be used to denoise and smooth out the measured data. Thus, this demodulation technique provided a complete process for a practical sensor system from the acquisition of raw data to human understandable measurand outputs, and all steps were performed digitally by a computer program.

Having described in details the demodulation algorithm, for an acquired multiplexed signal [Fig. 3.1 (b)], after taking the DWT, the approximation coefficients that represent the FBG signal is shown in Fig. 3.3 (a). The approximation coefficients give a smoothed and noise-reduced version of the original noisy signal. Fig. 3.3 (b) shows the magnification around the Bragg peak, and the Bragg wavelength shift can be found directly by employing a wavelength detection method. A resolution of ~1.2 pm after interpolation was achieved. For the FFI sensor, after taking the DWT, Fig. 3.4 (a) shows the detail coefficients, which effectively represented the mean-removed and windowed version of the original interference pattern. Fig. 3.4 (b) shows the Fourier transformed magnitude spectrum of the detail coefficients around the amplitude peak. It can be seen that, after conversion of x-variable, the cavity length can be directly determined from its position in the magnitude spectrum. A resolution of ~0.04 μm was achieved after interpolation.

An application of this demodulation technique was demonstrated through the simultaneous measurement of strain and temperature of an aluminium (Al) plate. With the setup from Fig. 3.1 (a) the FBG was loosely adhered onto the Al-plate, such that it was not affected by the strain field. The FFI adhered firmly onto the Al-plate next to the FBG using some superglue, and so it was affected by both strain and temperature changes of the plate. Before proceeding to the actual experiment, the coefficients in the sensitivity matrix needed to be determined. To illustrate this, let the two sensors be represented as,

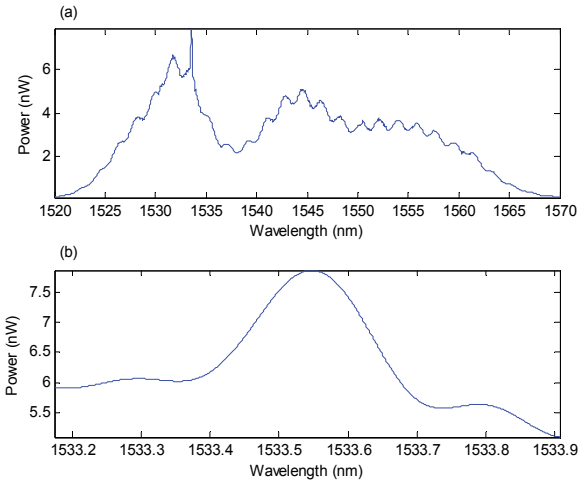


Fig. 3.3. (a) Approximation coefficients for the FBG; (b) around the Bragg peak.

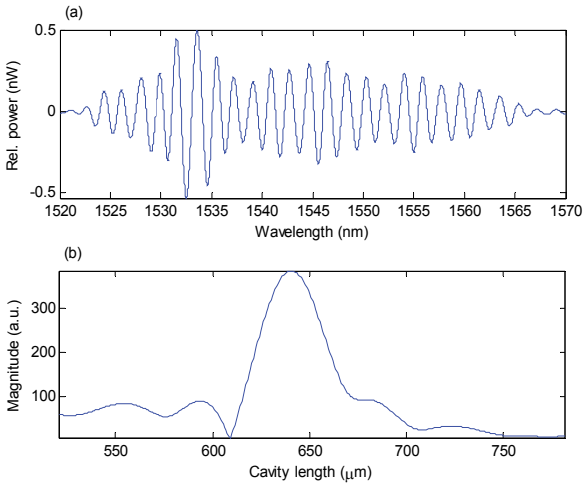


Fig. 3.4. (a) Detail coefficients for the FFI; (b) FT magnitude spectrum around the peak.

$$\begin{aligned} \Delta\zeta_1 &= K_{1\varepsilon}\Delta\varepsilon + K_{1T}\Delta T \\ \Delta\zeta_2 &= K_{2\varepsilon}\Delta\varepsilon + K_{2T}\Delta T \end{aligned}, \tag{3.2}$$

where $\Delta\zeta_i$ are the measurand-induced physical change of the sensors, and K_{ij} are the measurand-induced responses (i.e., sensitivity coefficients). $\Delta\varepsilon$ and ΔT are the strain and temperature changes, respectively. Eq. (3.2) is a set of two simultaneous linear equations, which can be expressed in matrix form as,

$$\begin{pmatrix} \Delta\zeta_1 \\ \Delta\zeta_2 \end{pmatrix} = \begin{pmatrix} K_{1\varepsilon} & K_{1T} \\ K_{2\varepsilon} & K_{2T} \end{pmatrix} \begin{pmatrix} \Delta\varepsilon \\ \Delta T \end{pmatrix}. \tag{3.3}$$

This equation can be solved for the measurand vector by inverting the *K*-matrix, i.e.,

$$\begin{pmatrix} \Delta\varepsilon \\ \Delta T \end{pmatrix} = \frac{1}{\det(K)} \begin{pmatrix} K_{2T} & -K_{1T} \\ -K_{2\varepsilon} & K_{1\varepsilon} \end{pmatrix} \begin{pmatrix} \Delta\zeta_1 \\ \Delta\zeta_2 \end{pmatrix}, \tag{3.4}$$

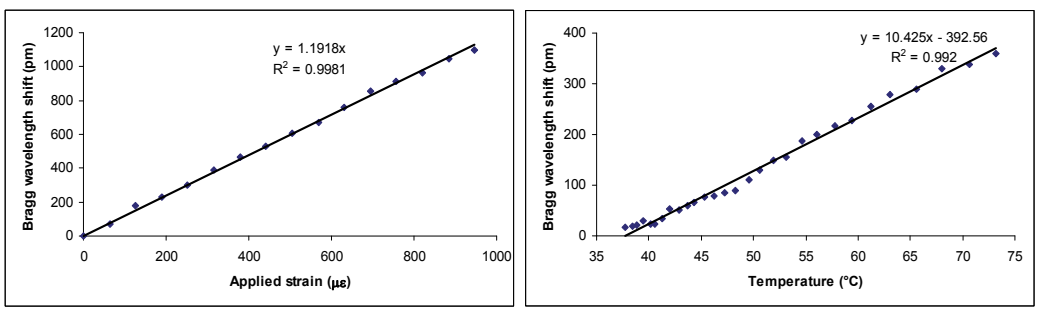
where $\det(K) = K_{1\varepsilon}K_{2T} - K_{1T}K_{2\varepsilon} \neq 0$. This shows the strain and temperature can be simultaneously separated and measured provided the elements in the square matrix are pre-determined. This can be done by measuring one measurand (while the other kept constant) at a time against the spectral response of each sensor. For this sensor arrangement, with the details described in (Wong et al., 2006b), the strain and temperature responses of the FBG and FFI sensors are shown in Figs. 3.5 and 3.6, respectively. From the linear regression fits, for the FBG sensor the strain and temperature sensitivity coefficients are 1.2 pm/με and 10.4 pm/°C, respectively; whereas for the FFI sensor, the respective sensitivity coefficients are 4.06×10^4 pm/με and 1.12×10^6 pm/°C. Since the FBG was not affected by strain change, the system of linear equations is given by,

$$\begin{pmatrix} \Delta d \\ \Delta\lambda_B \end{pmatrix} = \begin{pmatrix} 4.06\times10^4 & 1.12\times10^6 \\ 0 & 10.4 \end{pmatrix} \begin{pmatrix} \Delta\varepsilon \\ \Delta T \end{pmatrix}, \tag{3.5}$$

and

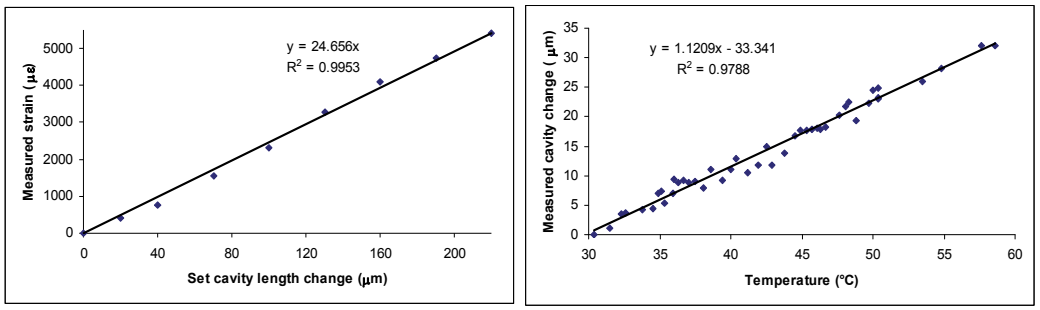
$$\begin{pmatrix} \Delta\varepsilon \\ \Delta T \end{pmatrix} = \begin{pmatrix} 2.5\times10^{-5} & -2.7 \\ 0 & 9.6\times10^{-2} \end{pmatrix} \begin{pmatrix} \Delta d \\ \Delta\lambda_B \end{pmatrix}, \tag{3.6}$$

where Δd and $\Delta\lambda_B$ are the FFI cavity length change and FBG Bragg wavelength shift, respectively. The sensors were left in the laboratory overnight for a period of 14 hours, with no external axial strain applied. A LabVIEW program was written to acquire the sensor signal in real-time at a rate of ~6 s per reading. The FFI subjected to both the thermal strain and temperature change of the Al-plate, while the FBG subjected to temperature change only. By using Eq. (3.6), the separated effects of thermal strain and temperature change over the measured period are shown in Fig. 3.7. Although both sensors experienced the same temperature change, the thermal strain of the Al-plate did not correlate very well with its temperature change. This is because the FBG was loosely adhered onto the Al-plate, it was influenced by the surrounding environment and temperature change more than the thermal response of the Al-plate.



(a) (b)

Fig. 3.5. (a) Strain and (b) temperature responses of the FBG sensor.



(a) (b)

Fig. 3.6. (a) Strain and (b) temperature responses of the FFI sensor.

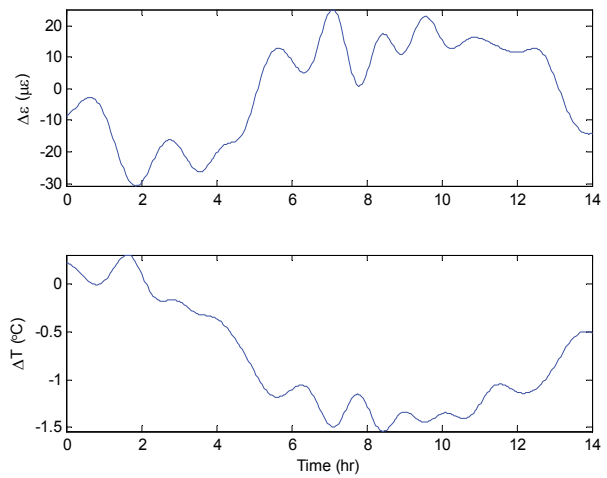


Fig. 3.7. Simultaneous measurement of strain and temperature of Al-plate.

4. Application case 2 – multiplexing and demultiplexing of multi-sensors

In the second example, a new type of FBG called amplitude-modulated chirped FBGs (AMCFBGs) was designed and fabricated, and based on that, a new multiplexing technique called spectral overlap multiplexing was proposed and experimentally demonstrated (Childs et al., 2010, Wong et al., 2007a, 2007b, 2010). We show that, the DWT played a central role in the demultiplexing, demodulation and analysis of these multiplexed novel sensor signals. The AMCFBGs are similar to ordinary chirped FBGs, i.e., they have a broad and flat reflection spectrum. The subtle difference is the addition of an amplitude-modulation function to the refractive index modulation of the fibre core, which is achieved by varying the induced DC refractive index during the writing process. With this index modulation, the reflection spectrum has a unique signature – a sinusoidal modulation on its flat-top region. The amplitude-modulation function for the refractive index modulation is given by $f_d(z) = w \cdot \sin^2\left(\pi dz / L_g\right)$, where w is an apodisation function such as a raised-cosine function, d is the number of periods (frequency) of modulation, and L_g is the length of the

grating. The reflection spectrum for one AMCFBG can be approximated by the expression (Wong et al., 2007a),

$$R(\lambda) \approx \chi_{\left(\lambda_{\min}, \lambda_{\min} + c_L L_g\right)}(\lambda)\left[R_{\min} + \left(R_{\max} - R_{\min}\right) f_d\left(\frac{\lambda - \lambda_{\min}}{c_L}\right)\right], \tag{4.1}$$

where λ_{min} is the initial wavelength of the grating, R_{min} and R_{max} are the reflectivities of the troughs and peaks of the grating spectrum, respectively, c_L is the linear chirp rate of the phase mask used for writing the gratings, and $\chi_I(x)$ is the characteristic function on the interval I which equals 1 when x is an element of I and zero otherwise. The fabrication procedure can be found in (Wong et al., 2007b). As an example of an AMCFBG that we fabricated, Fig. 4.1 (a) shows the amplitude-modulation function with a raised-cosine apodisation, d = 5 and L_g = 10 mm; and Fig. 4.1 (b) shows the corresponding measured reflection spectrum.

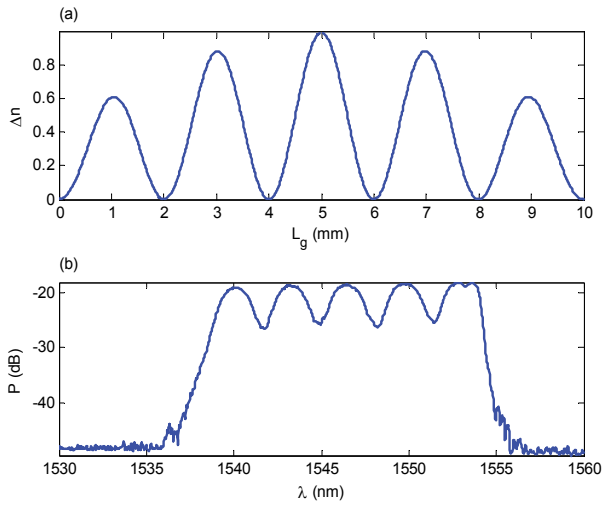


Fig. 4.1. (a) Index modulation function used in the fabrication; (b) measured reflection spectrum of the AMCFBG.

In the wavelength domain, ordinary FBGs cannot be easily distinguished if they are spectrally overlapped. Using the amplitude-modulation as a unique signature for each grating, the AMCFBGs are able to completely overlap one another having the same spectral characteristics, i.e., centre Bragg wavelength, bandwidth and reflectivity, yet still be uniquely distinguishable from each other. The uniqueness of each AMCFBG is defined by the number of periods (spatial-frequency) of its amplitude-modulation, and for a set of spectrally-overlapped gratings, no two gratings can have the same number of periods. With such unique signatures for the AMCFBGs, viable methods are needed to demultiplex and demodulate the multiplexed signals, and that was accomplished by using the DWT technique similar to that applied in the previous section. This is the basis of the new spectral overlap multiplexing technique. Since this multiplexing is fully compatible with the WDM, the sensor count can potentially be increased by several folds.

To demonstrate the spectral overlap multiplexing using AMCFBGs, we performed two experiments. In the first experiment, two spectrally overlapped gratings, S1 and S2, were employed, and the schematic diagram of the setup is depicted in Fig. 4.2. The individual spectra of S1 and S2, as well as the combined spectra are shown in Fig. 4.3 (a) – 4.3 (c), respectively. It is clear that, WDM scheme would not allow such overlapping, and without a suitable signal processing technique, it is unlikely the multiplexed signal be separated and analysed. However, by taking the DWT, the unique signatures, i.e., the

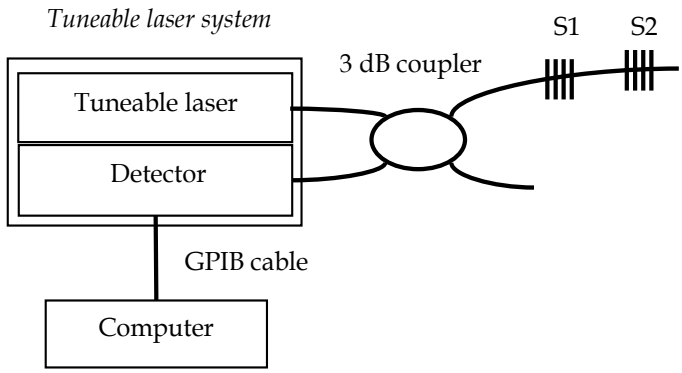


Fig. 4.2. Experimental setup of the multiplexed AMCFBG system.

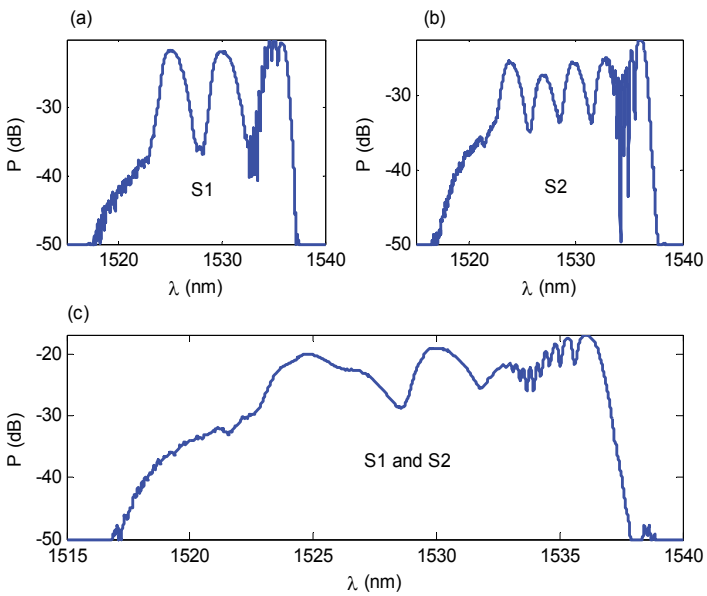


Fig. 4.3. Reflection spectrum of (a) S1, (b) S2, and (c) combined signal.

sinusoidal modulations of the gratings, can be recovered. Fig. 4.4 shows that the detail coefficients of the two gratings, and the unique modulated periods are unambiguously identified. That is, the multiplexed signal has been successfully demodulated whereby the

measurand-induced wavelength shifts can be measured individually by tracking the 'phase-shift' of the modulations. The strain response and crosstalk of the multiplexed AMCFBGs under this multiplexing scheme were conducted. A total of ten 20-cent coins (each weighs 11.3 g, which corresponds to a strain of 125 με) were applied to S1, while S2 was left unstrained. For each coin applied, ten readings were taken. Fig. 4.5 (a) shows the

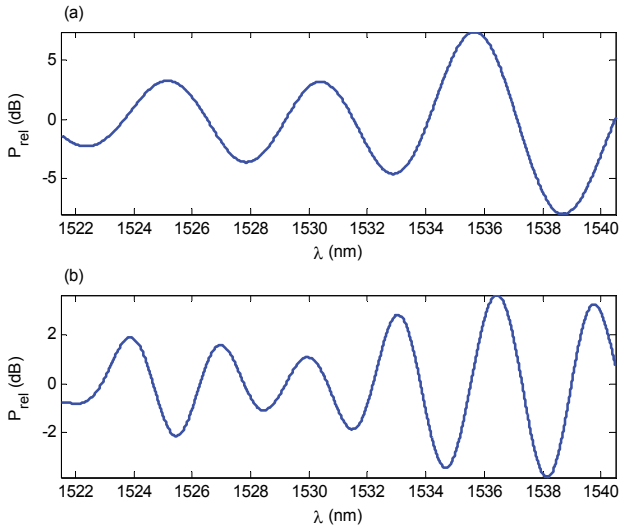


Fig. 4.4. Detail coefficients for (a) S1 and (b) S2 gratings.

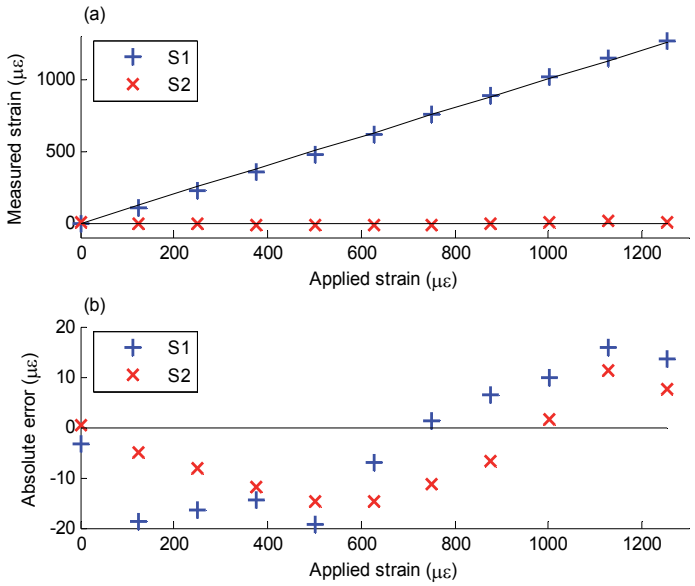


Fig. 4.5. (a) Strain measurements and (b) absolute errors of S1 (strained) and S2 (unstrained).

measured strain responses of S1 and S2 as a function of applied strain. The strain values were obtained by using the DWT technique mentioned before. The lines are the ideal

strain values. It can be seen that S1 measured the strain very accurately. For strains applied up to 1250 με, the absolute error is within ±20 με as shown in Fig. 4.5 (b). The corresponding strain response of S2 showed little crosstalk, with a maximum value of about 16 με as shown in Fig. 4.5 (b). This small amount of crosstalk implies that adverse effects, such as spectral shadowing, did not impose much of a problem on this multiplexing technique.

In the second experiment, simultaneous strain and temperature measurement of an Al-alloy plate was carried out. The experiment was similar to that in the previous section. That is, the AMCFBG sensors were first characterised to obtain the strain and temperature sensitivity coefficients, then simultaneous two-parameter measurement of the material was performed. Strain and temperature changes were obtained by measuring the wavelength (i.e., 'phase') shifts of the sensors and calculating their corresponding values. Two overlapped AMCFBGs were used, and to effectively utilise the advantage of having the same Bragg wavelength, the reference grating method (Wong et al., 2007b) was used. In such method, the strain sensor was firmly attached onto the structure, while the temperature sensor was placed under the same environmental conditions but unstrained. As such, the former experienced wavelength shifts due to both strain and temperature, while the latter only experienced the shift due to temperature change. With reference to Eq. (3.3), the set of simultaneous equations is given by,

$$\begin{pmatrix} \Delta\lambda_{B1}/\lambda_{B1} \\ \Delta\lambda_{B2}/\lambda_{B2} \end{pmatrix} = \begin{pmatrix} k_{1\varepsilon} & k_{1T} \\ 0 & k_{2T} \end{pmatrix} \begin{pmatrix} \Delta\varepsilon \\ \Delta T \end{pmatrix}, \text{ or } \begin{pmatrix} \Delta\varepsilon \\ \Delta T \end{pmatrix} = \frac{1}{k_{1\varepsilon}k_{2T}} \begin{pmatrix} k_{2T} & -k_{1T} \\ 0 & k_{1\varepsilon} \end{pmatrix} \begin{pmatrix} \Delta\lambda_{B1}/\lambda_{B1} \\ \Delta\lambda_{B2}/\lambda_{B2} \end{pmatrix}, \tag{4.2}$$

where the subscripts 1 and 2 represent strain and temperature sensors, respectively. Some approximations are made: (a) as both sensors experience the same temperature change, $k_T = k_{1T} = k_{2T}$; (b) the Bragg wavelengths are the same for both sensors and so $\lambda_B = \lambda_{B1} = \lambda_{B2}$. By setting $k_\varepsilon = k_{1\varepsilon}$ and modifying Eq. (4.2), the expressions for strain and temperature change are then given by,

$$\Delta\varepsilon = \frac{1}{k_\varepsilon}\frac{\left(\Delta\lambda_{B1} - \Delta\lambda_{B2}\right)}{\lambda_B}; \text{ and } \Delta T = \frac{1}{k_T}\frac{\Delta\lambda_{B2}}{\lambda_B}. \tag{4.3}$$

Eq. (4.3) shows that the strain of a structure under test can be simply obtained from the differential wavelength shift between the two sensors, and temperature from the temperature sensor alone. Both the strain and temperature coefficients can be determined experimentally by measuring their measurand-induced responses. Fig. 4.6 shows the strain and temperature responses of an AMCFBG. Based on the slopes of the linear regression fits, the strain and temperature sensitivity coefficients were 1.20 pm/με and 10.06 pm/°C, respectively. The temperature measurement was carried out when the sensor was free and unstrained. However, when it was used to measure the Al-alloy plate, due to the thermal expansion mismatch between the plate and the silica fibre, temperature response needed to be re-measured with the sensor adhered onto the plate. The measured temperature coefficient is found to be 38.12 pm/°C. Now, the strain and temperature can be expressed numerically as,

$$\Delta\varepsilon = 0.8295 \cdot \left(\Delta\lambda_{B1} - \Delta\lambda_{B2}\right); \text{ and } \Delta T = 0.0262 \cdot \Delta\lambda_{B2}, \tag{4.4}$$

where the Bragg wavelength shift, strain and temperature change are in units of pm, με and °C, respectively. To practically measure the strain and temperature change of the Al-alloy plate, one of the AMCFBG sensors was adhered firmly onto the plate with epoxy resin, whereas the other sensor was loosely attached such that it was free and unstrained from any strain field of the plate. The plate was placed inside a polyurethane foam box to minimise environmental perturbations. With the setup of Fig. 4.2, experiment was performed for a period of 18 h. By using Eq. (4.4), the strain and temperature changes are shown in Fig. 4.7. From the figure, both the strain and temperature curves changed in a much correlated manner, indicating that the strain was mainly thermally-induced. This is obvious as there was no external strain applied throughout the experiment, and unlike the previous results in Fig. 3.7, the experiment was isolated and so the thermal strain followed well with temperature change. The net change in strain and temperature during this 18 h period were about 40 με and 1.5 °C, respectively. Since the plate was placed inside a foam box that had a relatively high heat capacity, the change in temperature was a bit small. By plotting a graph of temperature vs. strain, the slope of the linear regression fit gives a thermal strain sensitivity of 33 με/°C.

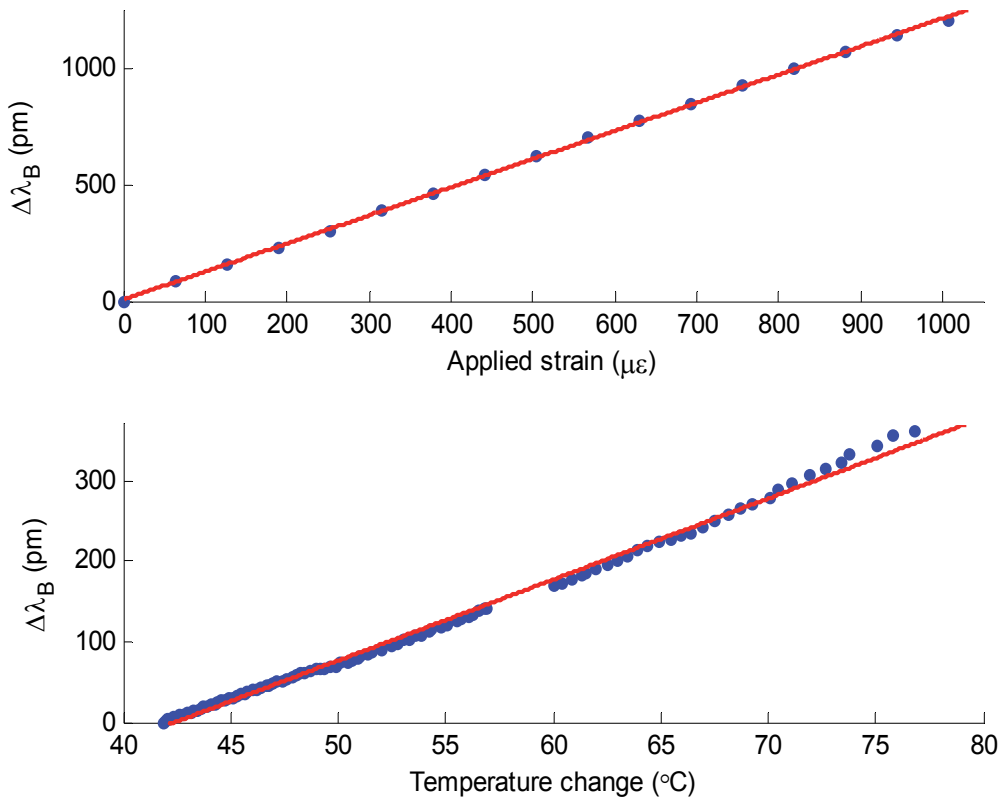


Fig. 4.6. Strain and temperature responses of the AMCFBG.

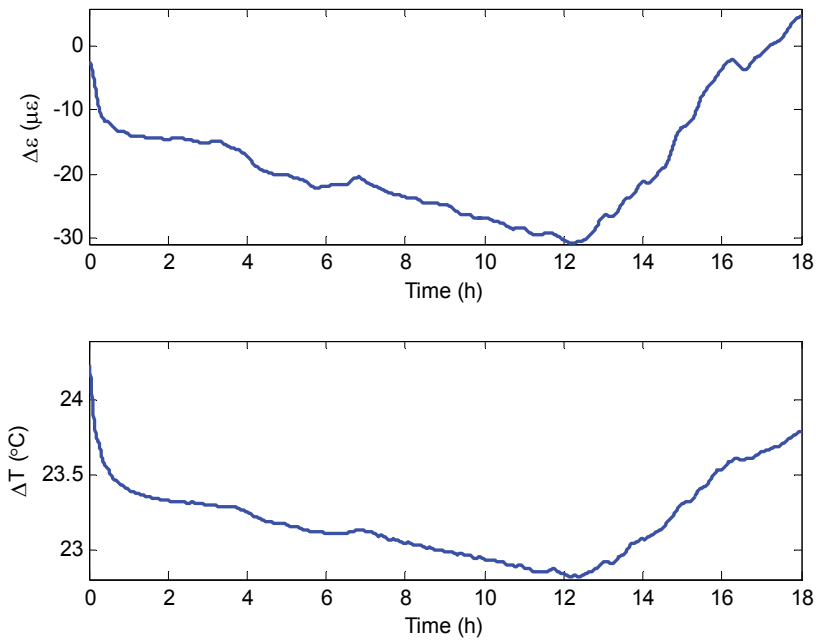


Fig. 4.7. Simultaneous measurement of strain and temperature of an Al-alloy.

5. Application case 3 – multiplexing of photonic crystal fibre sensors

In the third example, DWT was applied to the multiplexing and demultiplexing of a relatively new class of fibre – photonic crystal fibres (PCFs) (Fu et al., 2009). PCFs distinguish themselves from conventional fibers that they consist of microstructured air holes along the fibre, and a wide variety of cross-section air hole arrangements can be designed to suit different applications. The particular type of PCF used in the example was called polarisation-maintaining PCFs (PM-PCFs) that have the characteristics of high birefringence and low temperature sensitivity, and so is suitable for single parameter sensing where cross-sensitivity issue can be minimised. However, at present, all reported PCF sensors were operated as single sensors, and a main reason was due to the difficulty in demultiplexing and demodulating the multiplexed PCF sensor signals, even though the multiplexing schemes are simple and easy to implement. To overcome this, we demonstrate the use of DWT to separate the multiplexed sensor signal, such that the change from each individual sensor can be extracted and measured.

In our experimental setup, two PM-PCF sensors, PM-PCF1 and PM-PCF2, were multiplexed in series, as shown in Fig. 5.1. Each sensor unit was arranged in Sagnac interferoemter configuration with a section of PM-PCF as the birefrigent element, and the output signal can be represented by the transmission matrix as, $T = [1 - \cos(\delta)]/2$. The phase difference δ introduced by the PM-PCF with a length of L to the two light beams is wavelength dependent and is given by $\delta = 2\pi BL/\lambda$. The period of the output spectrum, i.e., the spacing between two adjacent minima, is $S=\lambda^2/(BL)$, where B is the birefringence of the PM-PCF. The birefringence change due to environmental parameters can then be detected by

measuring the 'phase shift' of minima. The output transmission spectrum of K sensor units multiplexed in series is given by (Fu et al., 2009),

$$\frac{P_{output}}{P_{input}} = 10Log_{10}\prod_{k=1}^{K}\left(\frac{1}{2}L_k\left[1-\cos(\frac{2\pi}{S_k}\lambda+\theta_k)\right]\right)[\mathrm{dB}], \tag{5.1}$$

where L_k, S_k, θ_k are the loss, the period of the output spectrum and the initial phase of the k-th sensor, respectively. Note that the output spectrum is indeed the multiplication of individual sensor signals. PM-PCF1 (length of 20 cm) was placed freely on a table, while PM-PCF2 (length of 60 cm) was placed inside a sealed pressure chamber. Pressure was

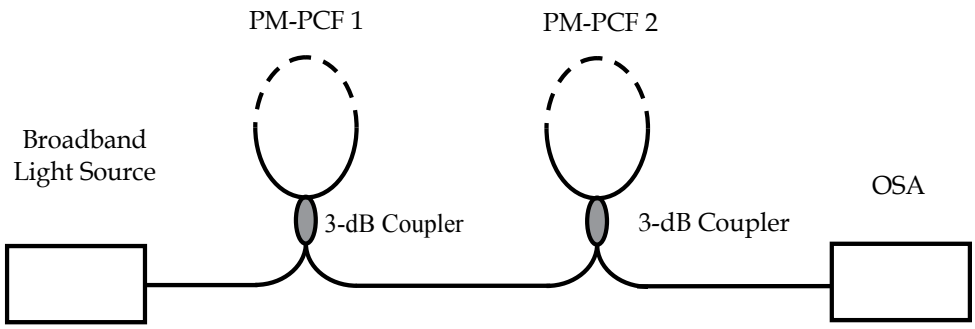


Fig. 5.1. Experimental setup of in series multiplexing technique for PM-PCF based Sagnac interferometric sensor.

applied to PM-PCF2 from 0 – 3 bars in steps of 0.5 bar, and was measured by a pressure gauge (COMARK C9557). Fig. 5.2 shows the output spectra of various pressure values measured by the OSA. In principle, to obtain the sensing information, the wavelength shift

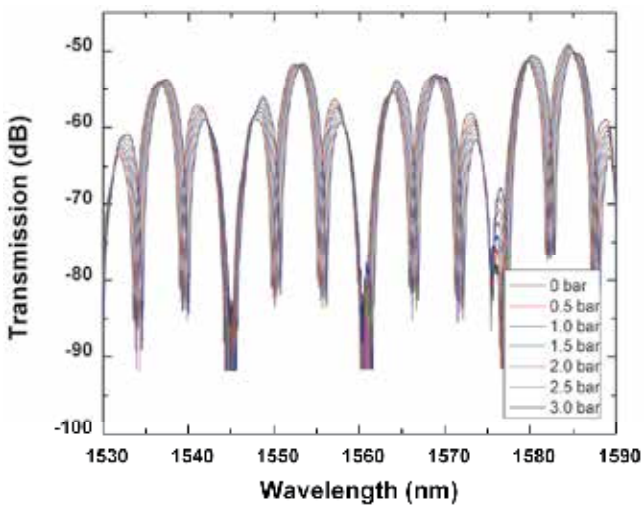
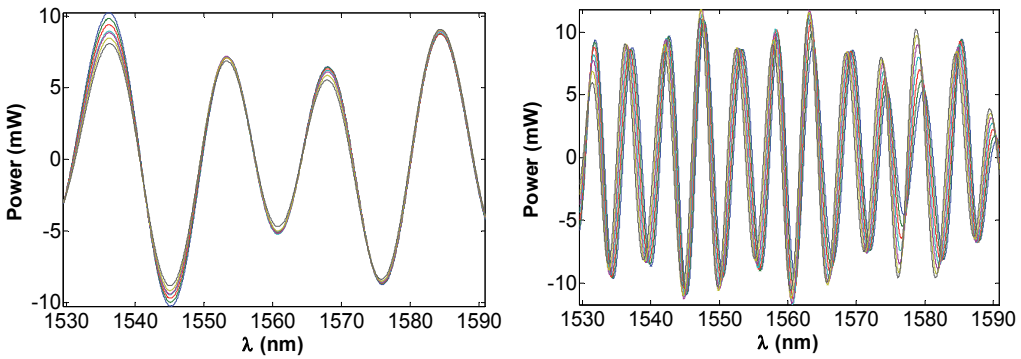


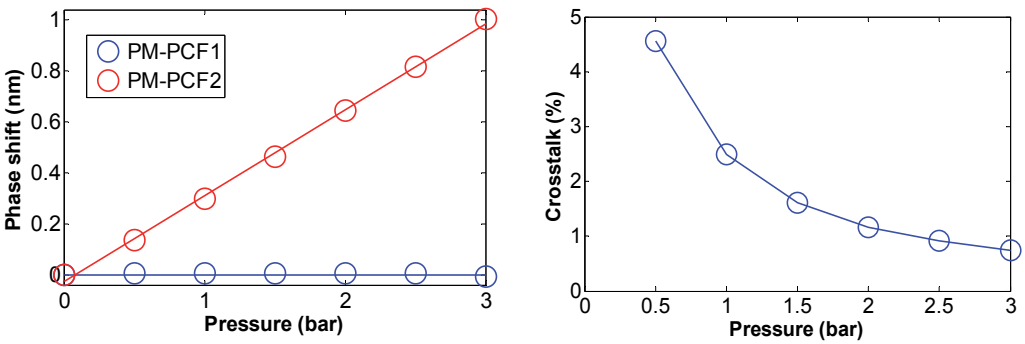
Fig. 5.2. Output transmission spectra of the two multiplexed Sagnac interferometric sensors in series with one sensor under applied pressure variations.

of the transmission minima of each sensor needs to be determined. However, as can be seen, the multiplexed sensor signal is more complex, and so simple tracing of the initial phase may not yield accurate results. We applied the DWT to the sensor signal, and the detail coefficients at two different levels are shown in Fig. 5.3. These two sets of coefficients are indeed the original sensor signal of the two sensors, and from the figure, the spectrum of PM-PCF2 shifted linearly with increasing pressure, while PM-PCF1 remained unchanged (at least the initial phase). Figs. 5.4(a) and 5.4(b) show the spectral shift of the two sensors as a function of applied pressure, and the crosstalk (includes other sources of errors, such as measurement error and ambient noise) between them, respectively. Thus, this example clearly demonstrated the capability of DWT in demultiplexing and demodulating multiplexed PCF sensor signals.



(a) (b)

Fig. 5.3. Detail coefficients of two PM-PCF sensors at two wavelet levels.



(a) (b)

Fig. 5.4. (a) Spectral phase shift and (b) crosstalk of the two PM-PCF sensor as a function of applied pressure.

6. Application case 4 – measurands analysis of novel fibre sensors

In the fourth example, DWT is used as a signal analysis tool to demodulate, analyse and interpret acquired signals from two novel fibre sensors, namely tilted moiré FBG and tilted Bragg reflector fibre laser (TBR-FL). These two sensors were proposed to perform simultaneous two-parameter sensing using only single sensing elements. The advantages are: (i) the capability of detecting more measurands using fewer sensors; (ii) the ability of mitigating the issue of cross-sensitivity (mostly temperature-induced) inherited from the sensor property (Rao, 1997); and (iii) sensor structure can be made more compact, which simplifies and eases the packaging and installation works for practical applications. It is known that, in order to measure two parameters simultaneously, either two distinctive sensor types or sensors of two different spectral characteristics are needed. The proposed sensors are unique in the sense that they are designed in such a way that, within a single sensor structure, their spectral characteristics response differently to different measurands, and so permitting them to distinguish individual measurand-induced changes. As a result, their spectral profiles are relatively more complex, and can be considered as having two parts from two different sensor types. Therefore, the spectra cannot be easily separated and analysed. However, with DWT, such complex sensor signals are readily separated without losing measurands information.

6.1 Tilted moiré fibre Bragg gratings

The tilted moiré FBG was originally proposed as a bandwidth controllable filter for telecommunications applications (Wong et al., 2010b). Here, we extended its application to fibre sensing, in particular, it was proposed as a single sensor for simultaneous two-parameter sensing (Wong et al., 2010c). The design criteria and fabrication procedure are detailed in (Wong et al., 2010c), and a typical spectrum is shown in Fig. 6.1. It can be seen that the sensor signal consists of two separate parts associated with the phase-shifted main Bragg mode (with a narrow resonance dip) and discrete cladding modes (including the ghost mode). The former is mostly confined in the core of the fibre, while the latter exist in the cladding region of the fibre. In that region, light is only loosely confined and its intensity decays exponentially along the radial direction and eventually radiates out of the fibre. As such, the cladding modes are capable of interacting with the surrounding environment at the fibre boundary, and sensing can be carried out via the principle of evanscent wave sensing.

To demonstrate the use of a single sensor for simultaneous two-parameter sensing, we performed simultaneous measurement of temperature and refractive index (RI) of an aqueous solution. The aim was to obtain the measurand-induced sensitivity coefficients independently, such that the wavelength shifts from each measurand can be unambigously determined. Thus, we carried out the characterisation experiments separately to obtain the temperature and RI sensitivity coefficients. First, temperature was varied by putting the sensor into a container filled with hot water, and the value was measured by a digital thermometer (Fluke 52II with K-type thermocouple) in the range of 43°C – 77°C, captured at every 1°C intervals. Next, the sensor was placed in a container filled with pre-mixed and saturated glucose solution (Dextrosol D-Glucose powder), and RI was varied by adding water to dilute the solution. A digital refractometer (Reichert

AR200) was used to obtain the absolute RI and temperature of the samples. The measured RI range was between 1.3331 and 1.4117, set by the RI of water and that of saturated glucose solution. The temperature (obtained from the refractormer) varied between 21.0°C – 21.6°C during the experiment. In both cases, wavelengths of the Bragg and cladding modes shifted according to measurands changes. More specifically, temperature changed the Bragg wavelength via thermo-optic effect (thermal expansion of glass fibre is very small that can be neglected), whereas the RI changed both the wavelength and transmission loss of the cladding modes. Thus, by applying the DWT, these two types of modes can be separated and their changes due to the two measurands can be extracted and analysed. Fig. 6.2 shows the detail wavelet coefficients for the (a) Bragg mode and (b) cladding modes at the 7th- and 6th-levels, respectively. Since the detail coefficients are mean removed, and so in the Bragg mode, the narrow resonance dip becomes a sharp peak, and temperature change can be measured by tracking this peak shift. As for the cladding modes, it is known that each individual cladding mode responds differently to the same measurand (e.g., RI) change, and taking an average of a number of cladding modes shifts will give more accurate readings than tracking just one particular mode. RI is thus measured by the averaged wavelength shift of about 10 cladding modes.

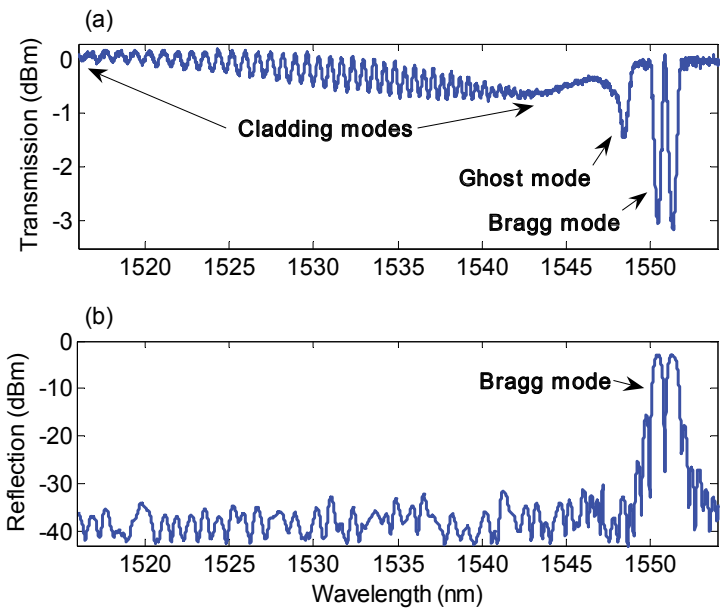


Fig. 6.1. (a) Transmission and (b) reflection spectra of the TMFBG sensor for simultaneous two-parameter sensing.

After taking the DWT to the measured spectra, Fig. 6.3 (a) shows the wavelength shifts of the wavelet coefficients for the Bragg and averaged cladding modes, respectively, as a function of temperature. The solid lines are the linear regression fits of the data points. It is clear that, in response to temperature change, the averaged cladding modes shifted by almost the same amount and in the same direction as the Bragg mode. From the regression lines, the temperature-induced sensitivity of the Bragg and averaged cladding modes are

10.53 pm/°C (with R^2 = 0.9979) and 10.48 pm/°C (with R^2 = 0.9925), respectively, which correlated very well with each other. Thus, it is sufficient to measure the temperature-only change by tracking the wavelength shift of the Bragg mode. Similarly, Fig. 6.3 (b) shows the wavelength shifts of the wavelet coefficients for the Bragg and averaged cladding modes, respectively, as a function of RI. From the figure, the Bragg mode remained unperturbed (within measurement errors), indicated that it is insensitive to RI change. The cladding modes, however, varied nonlinearly, with the curve best fitted by a polynomial function given empirically by,

$$P_{DWT}(n) = 169.58n^4 - 655.26n^3 + 843.59n^2 - 361.85n \tag{6.1}$$

(with R^2 = 0.9987) for the specified range $1.3331 \le n \le 1.4117$. Thus, for RI-only change, Eq. (6.1) can serve as a lookup table, such that by tracking the wavelength shift of the averaged cladding modes, the RI value can be obtained. By combining these two distinct measurands-induced responses, simultaneous measurement of temperature and RI is realised. Temperature change can be directly obtained from the wavelength shift of the Bragg mode, and the RI from the differential wavelength shifts between the Bragg and averaged cladding modes, i.e., the residual amount after subtracting the average shifted amount of the cladding modes from that of the Bragg mode. Fig. 6.3 (b) (and Eq. (6.1)) can then be used to find the RI value by looking-up the corresponding measured differential shift.

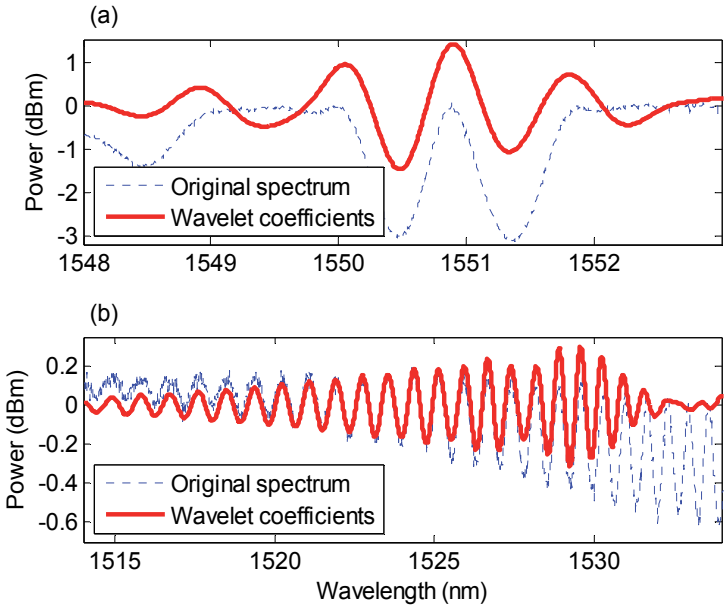


Fig. 6.2. Wavelet coefficients (solid lines) of a measured transmission spectrum: (a) 7th-level coefficients for the Bragg mode, and (b) 6th-level coefficients for cladding modes. Dotted lines are the original spectra.

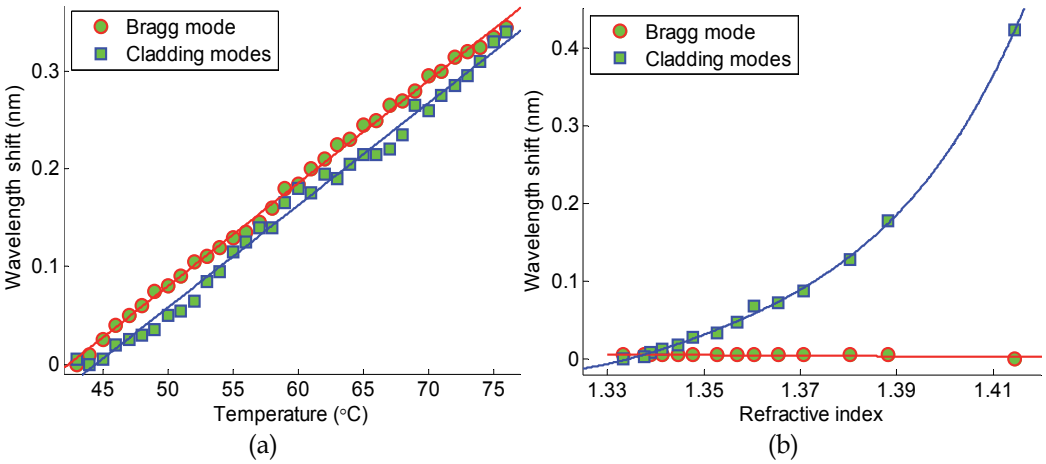


Fig. 6.3. Wavelength shifts of the wavelet coefficients for the Bragg mode and averaged cladding modes as a function of (a) temperature and (b) refractive index. Lines are the regression fits.

6.2 Tilted Bragg reflector fibre lasers

Up to present, there are mainly two types of grating-based FLs, namely distributed feedback and distributed Bragg reflector FLs. When applied in sensing, FLs have the advantage of high sensitivity, sensing output power and extinction ratio, and narrow linewidth/bandwidth. For simultaneous two-parameter sensing using single sensors, thus far, most of the proposed works are of passive type and only very few on using active sensors, e.g., fiber Raman lasers (Han et al., 2005, Tran et al., 2005) and distributed feedback FLs (Hadeler et al., 1999, 2001). TBR-FL is a new type of FL formed by a pair of wavelength and tilt-angle matched tilted FBGs (TFBGs) (Wong et al., 2011). In addition to the lasing peak, it possessed a grating tilt-induced cladding modes spectrum, which provided an extra sensing mechanism to detect the surrounding environment. We demonstrate that, with a simple experimental setup, the use of a single TBR-FL for simultaneous sensing of temperature and RI.

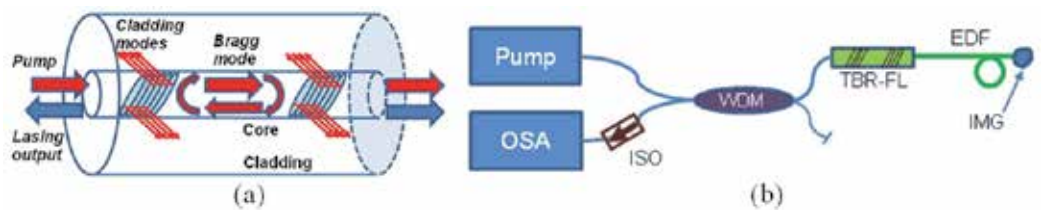


Fig. 6.4. (a) Structure of the TBR-FL, and (b) experimental setup. OSA = optical spectrum analyser, WDM = wavelength division multiplexer, EDF = erbium-doped fibre, ISO = isolator, IMG = index matching gel.

The structure of the TBR-FL comprises a pair of wavelength and tilt-angle matched TFBGs forming the resonant cavity and is depicted in Fig. 6.4 (a). The fabrication procedure is described in (Wong et al., 2011). The experimental setup depicted schematically in Fig. 6.4

(b). A laser diode was pumped to the TBR-FL via a wavelength-division-multiplexer (WDM), and the output signal was obtained by the OSA. At the other side of the FL was a continuous piece of EDF (~50 cm) looped in a diameter of ~3 cm, and so the sensor head comprised both the FL and the short coiled EDF section. Index matching gel (IMG) was applied to the far end of the EDF to minimise any reflections that may cause resonant feedback. This EDF section acted as an ASE source when excited by the excessive pump source. As such, the transmission spectrum of the constituent TFBG pair can be observed. With this setup, both the laser output and cladding modes spectra can be obtained simultaneously. A typical full spectrum is shown in Fig. 6.5 (a), which consisted of both the laser output [Fig. 6.5 (b)] and cladding modes [Fig. 6.5 (c)] spectra. Lasing occurred at the Bragg mode bound inside the core, whereas cladding modes were coupled out from the core and did not contribute to the laser operation. As mentioned, cladding modes can interact with the surrounding environment via evanescent wave, and therefore used to perform RI sensing. On the other hand, temperature change affected the FL as a whole and altered the entire spectrum (including lasing and cladding modes). Thus, simultaneous sensing was achieved by combining these two sensing properties, i.e., temperature was measured by tracking the lasing mode shift, and RI by the differential wavelength shift between the laser output and cladding modes. Similar to the previous case, the main objective was to empirically obtain the temperature and RI induced sensitivity coefficients. As such, characterisation experiments similar to that for tilted moiré FBGs were conducted [see also (Wong et al., 2011) for details], and the DWT was employed to demodulate and analyse the acquired sensor signals. Fig. 6.6 shows the wavelet coefficients of a typical measured spectrum. Detail coefficients are extracted for cladding modes [Fig. 6.6 (a)], as for the previous case; whereas approximation coefficients are extracted for the lasing mode [Fig. 6.6 (b)]. Since we are only interested in the lasing wavelength shift, approximation coefficients represent and resemble the original signal more accurately. After taking the DWT to the measured spectra, wavelength (i.e., wavelet coefficients) shifts of the laser output and averaged cladding modes (~20 modes) as a function of temperature are shown in Fig. 6.7 (a). Solid lines are the linear regression fits of the data. From the figure, both the laser output and averaged cladding modes yielded a very similar temperature-induced sensitivity, having values of 10.75 pm/°C (R^2 = 0.9991) and 10.88 pm/°C (R^2 = 0.9975), respectively. With such a high degree of correlation, it is sufficient to measure the temperature change by tracking the lasing wavelength shift alone. For RI, Fig. 6.7 (b) shows the wavelength shifts of the laser output and averaged cladding modes as a function of RI value. The laser output remained at the zero value (within measurement errors) and so have no direct relationship with RI. The averaged cladding modes varied nonlinearly with RI, and within the measured range the empirical relationship can be described by a polynomial function,

$$\Delta\lambda_{Cladding}(n) = 358.2n^4 - 1432.1n^3 + 1910.1n^2 - 849.9n \tag{6.2}$$

(R^2 = 0.9998). Thus, RI-only change can be obtained from the wavelength shift of the averaged cladding modes, as the laser output was not sensitive to it. Having established the temperature and RI sensitivity coefficients, simultaneous sensing of these two measurands using a single TBR-FL can be achieved. Temperature change can be obtained from the

wavelength shift of the laser output, whereas RI change from the differential wavelength shift between the averaged cladding modes and laser output. That is, the residual amount of the wavelength shift of the averaged cladding modes after subtracting from that of the laser output. As such, Fig. 6.7 (b) can be used as a look-up table to find the RI from the differential wavelength shift.

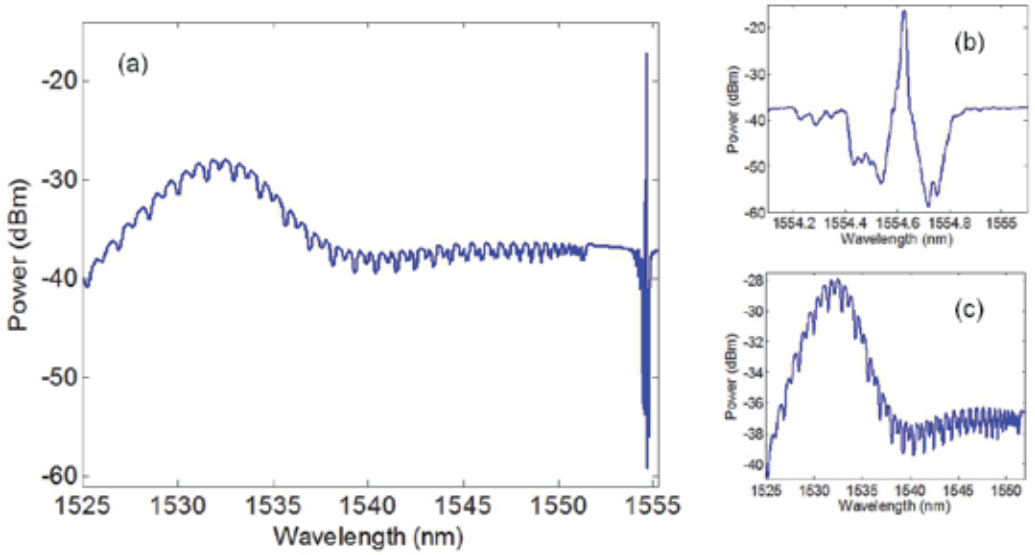


Fig. 6.5. (a) Full spectrum of the TBR-FL sensor; and the magnification around (b) the laser output and (c) cladding modes spectrum.

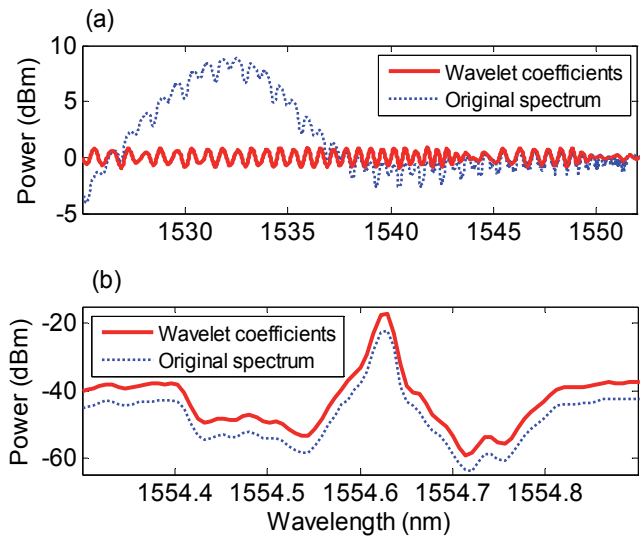
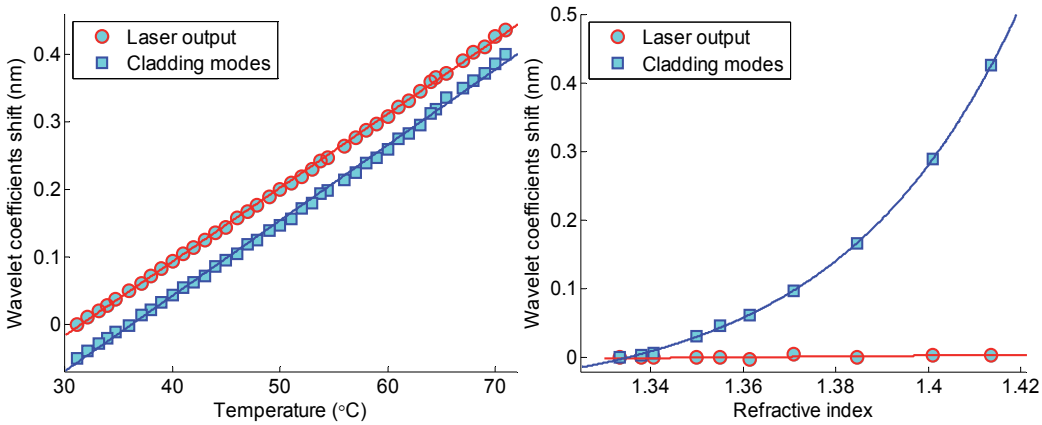


Fig. 6.6. Wavelet coefficients of a measured TBR-FL spectrum: (a) 6th-level detail coefficients for the cladding modes, and (b) 2nd-level approximation coefficients for the Bragg mode. Dotted lines are the original spectra (manually offset) for comparison.



(a) (b)

Fig. 6.7. Wavelet coefficients shift of the laser output and averaged cladding modes as a function of (a) temperature and (b) refractive index. Lines are the regression fits.

7. Conclusion

A review of the applications of DWT in optical fibre sensing is presented. Several representative application examples proposed by the authors have been discussed; and based on the implementation of DWT, novel signal processing techniques and fibre sensors have been designed and proposed. The concepts of DWT demodulation technique and the BLT wavelet denoising were introduced and applied to various application examples. First, we proposed and employed the DWT to demodulate and demultiplex a multiplexed FFI and FBG sensor system. Second, we designed a novel type of FBG (the AMCFBGs), and based on their unique overlapping properties, a new multiplexing technique called spectral overlap multiplexing was proposed and demonstrated. Third, DWT was employed in the multiplexing and demultiplexing of PCF-based sensors array, a relatively new class of fibre sensors that, up until now, has always been used as single sensors. Fourth, DWT was used as a signal analysis tool for two novel fibre sensors, namely the tilted moiré FBG (a passive sensor) and the TBR-FL (an active sensor) sensors. Complex sensor signals were separated and demodulated to obtain the individual measurands-induced changes, such that single sensing elements can perform simultaneous two-parameter sensing. In addition, BLT was applied in all the above cases to denoise sensor signals automatically during the use of the DWT demodulation technique.

8. Acknowledgment

Funding for an ARC Linkage Project (Project ID: LP0884100) from the Australian Research Council and the Roads and Traffic Authority, NSW, Australia, is gratefully acknowledged. Funding for an International Science Linkage Grant (Project ID: CG130013) from the Department of Industry, Innovation, Science and Research (DIISR), Australia, is also gratefully acknowledged.

9. References

Bang, H. J. & Kim, C. G. (2010). Study on the wavelet decomposed details of impact induced AE signals in composite laminates using fiber Bragg grating sensors, *Proc. SPIE*, Vol.7647, pp. 76473E

Chan, C. C.; Ni, N.; Sun, J.; Chu, Y. C.; Tang, Y.; Poh, C. L. (2007). Improving the detection accuracy in fiber Bragg grating sensors by using a wavelet filter, *J. Optoelectron. Adv. Mater.,* Vol.9, No.8, pp. 2376-2379

Chan, C. C.; Ni, N.; Sun, J.; Chu, Y. C.; Tang, Y.; Poh, C. L. (2010). Interferometric noise suppression in fiber Bragg grating sensors by using wavelet filter, *J. Optoelectron. Adv. Mater.,* Vol.12, No.6, pp. 1241-1246

Childs, P.; Wong, A. C. L.; Yan, B. B.; Li, M. & Peng, G. D. (2010). A review of spectrally coded multiplexing techniques for fibre grating sensor systems, *Meas. Sci. Technol.,* Vol.21, pp. 094007 (7pp)

Daubechies, I. (1992). *Ten Lectures on Wavelets,* Society for Industrial and Applied Mathematics, Philadelphia, PA, U. S. A.

Donoho, D. L. & Johnstone, I. M. (1994) Ideal spatial adaption by wavelet shrinkage *Biometrika,* Vol.81, pp. 425–55

Fu, H. Y.; Wong, A. C. L.; Childs, P. A.; Tam, H. Y.; Liao, Y. B.; Lu, C. & Wai, P. K. A. (2009). Multiplexing of polarization-maintaining photonic crystal fiber based Sagnac interferometric sensors, *Opt. Express,* Vol.17, pp. 18501-18512

Gangopadhyay, T. K.; Chakravorti, S.; Bhattacharya, K. & Chatterjee, S. (2005). Wavelet analysis of optical signal extracted from a non-contact fibre-optic vibration sensor using an extrinsic Fabry–Perot interferometer, *Meas. Sci. Technol.*, Vol.16, pp. 1075–1082

Gangopadhyay, T.K.; Chakravorti, S.; Chatterjee, S. & Bhattacharya, K. (2006). Time-frequency analysis of multiple fringe and nonsinusoidal signals obtained from a fiber-optic vibration sensor using an extrinsic Fabry-Perot interferometer, *J. Lightw. Technol.,* Vol.24, No.5, pp. 2122-2131

Hadeler, O.; Rønnekleiv, E.; Ibsen, M. & Laming, R. I. (1999). Polarimetric distributed feedback fiber laser sensor for simultaneous strain and temperature measurements, *Appl. Opt.*, Vol.38, pp. 1953–1958

Hadeler, O.; Ibsen, M. & Zervas, M. N. (2001). Distributed-feedback fiber laser sensor for simultaneous strain and temperature measurements operating in the radio-frequency domain, *Appl. Opt.*, Vol.40, pp. 3169–3175

Han, Y. G.; Tran, T. V. A.; Kim, S. H. & Lee, S. B. (2005). Multiwavelength Raman-fiber-laser-based long distance remote sensor for simultaneous measurement of strain and temperature, *Opt. Lett.*, Vol.30, pp. 1282–1284

Jones, K. J. (2000a). Wavelet edge detection applied to be used with composite embedded fiber optic sensors, *Proc. SPIE*, Vol.3986, pp. 397-406

Jones, K. J. (2000b). Wavelet crack detection algorithm for smart structures, *Proc. SPIE*, Vol.4328, pp. 306-313

Lamela-Rivera, H.; Macia-Sanahuja, C. & Garcia-Souto, J. A. (2003). Detection and wavelet analysis of partial discharges using an optical fibre interferometric sensor for high-power transformers, *J. Opt. A-Pure Appl. Opt.,* Vol. 5, No.1, pp. 66-72

Lee, B. (2003). Review of present status of optical fiber sensors, *Opt. Fiber Technol.*, Vol.9, pp. 57-79

Mallet, S. G. (1998). *A Wavelet Tour of Signal Processing,* Academic Press, Burlington, MA, U. S. A.

Mallet, S. G. (1989). A theory for multiresolution signal decomposition: the wavelet representation, *IEEE Trans. Pattern Anal. Mach. Intell.,* Vol.11, pp. 674–693

Rao, Y. J. (1997). In-fibre Bragg grating sensors, *Meas. Sci. Technol.*, Vol.8, pp. 355-375

Sidney Burrus, C.; Gopinath, R. A.; & Guo, H. (1998). *An Introduction to Wavelets and Wavelet Transforms,* Prentice Hall, Upper Saddle River, NJ, U. S. A.

Staszewski, W. J.; Gareth Pierce, S.; Worden, K.; Philp, W. R.; Tomlinson, G. R. & Culshaw, B. (1997). Wavelet signal processing for enhanced Lamb wave defect detection in composite plates using optical fiber detection, *Opt. Eng.*, Vol.36, No.7, pp. 1877–1888

Tomic, M.; Dinovic, Z. & Petricevic, S.; (2010). Fiber-optic pressure sensor based on Fizeau receiving interferometer, *Proceedings of the 33rd International Convention MIPRO*, pp. 100-104, Opatija, Croatia, May 24-28, 2010

Tran, T. V. A.; Han, Y. G.; Kim, S. H. & Lee, S. B. (2005). Long-distance simultaneous measurement of strain and temperature based on a fiber Raman laser with a single fiber Bragg grating embedded on a quartz plate *Opt. Lett.*, Vol.30, pp. 1632–1634

Vetterli, M. & Kovacevic, J. (1995). *Wavelets and Subband Coding,* Prentice Hall, Upper Saddle River, NJ, U. S. A.

Wang, G.; Pran, K.; Sagvolden, G.; Havsgard, G. B.; A E Jensen, A. E.; Johnson, G. A. & Vohra, S. T. (2001). Ship hull structure monitoring using fibre optic sensors, *Smart Mater. Struct.,* Vol.10, pp. 472-478

Wong, A. C. L.; Phillips, C. J. E. & Peng, G. D. (2005). Improved demodulation algorithm for spatial-frequency multiplexed fibre-optic Fizeau strain sensor system, Proceedings of 30th Australian Conference on Optical Fibre Technology, Sydney, Australia

Wong, A. C. L.; Childs, P. A. & Peng, G. D. (2006a). Simultaneous demodulation technique for a multiplexed fibre Fizeau interferometer and fibre Bragg grating sensor system, *Opt. Lett.*, Vol.31, pp. 23-25

Wong, A. C. L.; Childs, P. A. & Peng, G. D. (2006b). Multiplexed fibre Fizeau interferometer and fibre Bragg grating sensor system for simultaneous measurement of quasi-static strain and temperature using discrete wavelet transform, *Meas. Sci. Technol.,* Vol.17, pp. 384-392

Wong, A. C. L.; Childs, P. A. & Peng, G. D. (2007a). Multiplexing technique using amplitude-modulated chirped fibre Bragg gratings, *Opt. Lett.*, Vol.32, pp. 1887-1889

Wong, A. C. L.; Childs, P. A. & Peng, G. D. (2007b). Spectrally-overlapped chirped fibre Bragg grating sensor system for simultaneous two-parameter sensing, *Meas. Sci. Technol.,* Vol.18, pp. 3825-3832

Wong, A. C. L.; Childs, P. A. & Peng, G. D. (2010a). Spectrally coded multiplexing techniques in fibre-optic sensor systems, In: *Trends in Photonics*, J. Canning (ed.), Chapter 8, pp. 233-276, Transworld Research Network, ISBN: 978-81-7895-441-7, Kerala, India

Wong, A. C. L.; Giovinazzo, M.; Tam, H. Y.; Lu, C. & Peng, G. D. (2010b). Tilted moiré fiber Bragg grating optical filters with controllable passband and stopband, *J. Lightw. Technol.*, Vol.18, pp. 898-904

Wong, A. C. L.; Giovinazzo, M.; Tam, H. Y.; Lu, C. & Peng, G. D. (2010c). Simultaneous two-parameter sensing using a single tilted moiré fiber Bragg grating with discrete wavelet transform technique, *IEEE Photon. Technol. Lett.,* Vol.22, pp. 1574-1576

Wong, A. C. L.; Chung, W. H.; Lu, C. & Tam, H. Y. (2011). Single tilted Bragg reflector fiber laser for simultaneous sensing of refractive index and temperature, *Opt. Express,* Vol.19, pp. 409-414

Part 4

Identification and Diagnostics

13

Biometric Human Identification of Hand Geometry Features Using Discrete Wavelet Transform

Osslan Osiris Vergara Villegas, Humberto de Jesús Ochoa Domínguez,
Vianey Guadalupe Cruz Sánchez, Leticia Ortega Maynez
and Hiram Madero Orozco
Universidad Autónoma de Ciudad Juárez
Instituto de Ingeniería y Tecnología
Mexico

1. Introduction

Since the security factor became a basic need for civilization, a lot of systems have been developed. Those systems, try to ensure the safety in all the things that driving a certain degree of exclusivity. Historically, keys, cards and passwords were used as security systems; however, these methods are vulnerable to loss and theft. As a result biometric identification methods emerge in order to tackle the disadvantages of the non biometric classical methods.
Biometrics, is an emerging technology that addresses the automated identification of individuals, based on their physiological and behavioral traits. The main advantage of biometric methods is the ability to recognize, which is made by means of a physical feature or a unique pattern (Jain et al. (2008)). With these methods and individual can hardly be victim of plagiarism.
There exist several biometrics cues such as iris (Abhyankara & Schuckersa (2010)), face (Abatea et al. (2007)), fingerprint (Jimenez et al. (2010)), voice (Andicsa et al. (2010)), but one of the cheapest is the hand geometry. Hand geometry, as the name suggests, refers to the geometric structure of the hand (Singh et al. (2009)).
Hand geometry measurement is non intrusive and the verification involves a simple processing of the resulting features. Usually the hand geometry identification involves a digital picture acquisition and translation of the nodal points like: space between fingers, curvature, length and width of the hand into numerical representations used to cross reference with other hand prints stored in a database for a match.
The schemes which uses geometrical features of the hand, focused on characteristics as widths of fingers at articulations, finger and palm lengths, finger deviations and the angles of the inter-finger valleys with respect to the horizontal. The number of features obtained varied in the range of 20-30, and usually the acquisition stage need pegs to define the accurate finger position (Yoruk et al. (2006)).
The ability of associating an identity with an individual is called identification. Hand geometry measurements are easily collectible due to both the dexterity of the hand and due

to a relatively simple method of sensing which does not impose undue requirements on the imaging optics (Zunkel (1998)).
In this paper a method which uses 31 wavelet features for human hand geometry identification is presented. The paper is organized as follows: the related works about hand geometry identification are shown in section 2, the proposed methodology for identification is shown in section 3. Section 4 presents the tests and results obtained. Finally, the conclusions and further works are presented in section 5.

2. Related works

The biometric applied to hand geometry is an important issue which has been part of many investigations over the years, even when the theme is recent compared with other biometric models. There exists evidence to believe that since more than 3200 years ago the geometry of the hand was used to identify humans (Ratha & Govindaraju (1988)).
For example, in the Chauvet cave located in France, the walls were decorated with Palaeolithic art. Around the paintings of the cave there exist palms used to identify the creator of the painting. Additionally, reliable data show that in ancient Babylon and Egypt, the merchants had recognition techniques, by the impressions of the people fingerprint made in clay or by taking the morphology of the palm print.
The commercialization of hand geometry dates to the early 1970s with one of the first deployments at Georgia University in 1974. The US Army began testing hand geometry for use in banking in 1984. These deployments predate the concept of using the geometry of a hand for identification as patented by (Sidlauskas (1988)).
The first commercial device for hand geometry identification was made in 1994 by the Hungarian company Recowere Ltd. The hand identification systems are widely implemented for their ease of use, public acceptance and integration capabilities. In Fig. 1 an image of a commercial scanner is shown.
In the literature there are several works which tackle the problem of hand geometry identification, but never using wavelet features.
The paper of (Singh et al. (2009)) presents an overview of biometric hand geometry recognition. Five different methods were compared and the authors talks about the advantages and disadvantages of each method. An approach that uses the color of the skin of the hand as a feature for recognition is recommended. The best classifier proposed was Gaussian Mixture Models (GMM).
In (Sanchez et al. (2000)) a comparison of four different methods for hand geometry recognition is presented. 31 features were extracted and the classifiers compared were: Euclidean distance, Hamming distance, Gaussian Mixture Models (GMM) and Radial Basis Function Neural Networks. The authors reports a 97 percent of recognition success with the method of GMM. The False Rejection Rate (FRR) and False Acceptance Rate (FAR) always remain similar.
A work that uses the coefficients of the Fast Fourier Transform (FFT) and the coefficients of the Discrete Cosine Transform (DCT) as a features for hand recognition was presented by (Bolok et al. (2004)). Additionally, a database reduction algorithm is proposed. The classifier used was Euclidean distance and the percentage of recognition using FFT features was 99 and using DCT was 98.

Fig. 1. Example of hand image acquisition device.

The work in (Kumar & Zhang (2007)) presents an algorithm which exploit user-specific dependencies in the feature-level representation. The system employees a discrimination of hand geometry features using entropy-based heuristics. Four classifiers were used: K-Nearest Neighbor (K-NN), Bayes, Support Vector Machine (SVM) and Feedforward Neural Network (FFN). The results demonstrates the improvement of the classifier performance with feature discrimination, and the percent of recognition is equal to 95.
The work proposed by (Polat & Yildirim (2008)) presents a method for hand geometry identification, the system does not required the stages of image preprocessing and feature extraction before the identification. A general regression neural network is used for hand classification. The authors reports an FRR and FAR of 15.
In Table 1 a summary of the related works on hand geometry based is shown, the essential information of the table was collected at (Kanhangad et al. (2009)) and (Lai & Chaw (2009)). The first column, show the authors name and the year of publication, the second, column define the features used and in brackets the method used for classification. The third column, show the overall performance of the system in terms of FAR and FRR. The FAR is the probability of wrongfully accepting and imposter user. The FRR is the probability of wrongfully rejecting a genuine user. Finally, the fourth column show the population size.

3. Proposed methodology

The stages of the methodology for the solution of the problem statement are: a) an image of a hand is acquired, b) the image is preprocessed to enhance it, c) axes reallocation, d) wavelet transform, e) feature extraction in wavelet domain, f) nearest neighbor classification and g) the face of the individual of the hand recognized is displayed. In Fig. 2 an example of the methodology for identification using hand geometry is shown.

3.1 Image acquisition

120 hand images of different individuals were taken using a commercial scanner. The images were taken with a resolution of 300 dpi and the size of each image was 2550 x 3508 pixels as it is shown in Fig. 3.

Authors	Methodology	Performance	Database size
Golfarelli et al. (1997)	17 features including finger lengths (Mean of a multinomial pdf)	FAR = N/A FRR = N/A	100
Jain et al. (1999)	Measurements along 16 different axes (Euclidean distance)	FAR = 2.0% FRR = 15%	50
Jain & Duta (1999)	Alignment of finger shapes and shape distance measurement (match score)	FAR = 2.0% FRR = 3.5%	53
Sanchez et al. (2000)	several width, height and angle measurements (GMM)	FAR = 6.0% FRR = 6.0%	20
Lay (2000)	Distorted pattern of the back of hand (Quadtree)	FAR = 0.0% FRR = 3.9%	100
Kumar et al. (2003)	16 geometry measurements (Normalized correlation)	FAR = 5.3% FRR = 8.2%	100
Bolok et al. (2004)	Coefficients of FFT and DCT (Euclidean distance)	FAR = N/A FRR = N/A	40
Bulatov et al. (2004)	30 features (Training vector bounding box)	FAR = 1.0% FRR = 3.0%	70
Woodard & Flynn (2005)	Shape index (Normalized correlation coefficient)	FAR = 5.5% FRR = 5.5%	177
Xiong et al. (2005)	Elliptical model and finger tip/valley information	FAR = 2.4% FRR = 2.4%	108
Malassiotis et al. (2006)	96 curvature and 3D finger width measurements (L_1)	FAR = 3.6% FRR = 3.6%	73
Kumar & Zhang (2007)	Feature discretization (K-NN, Bayes, SVM, FFN)	FAR = N/A FRR = N/A	100
Polat & Yildirim (2008)	No features (Regression Neural Network)	FAR = 15.0% FRR = 15.0%	140
Adán et al. (2008)	14 Non-landmark features (Time averaged)	FAR = 0.45% FRR = 3.4%	470
Kanhangad et al. (2009)	Fusion of 3D and 2D hand geometry features	FAR = 2.6% FRR = 2.6%	177
This paper	31 Wavelet Features (Nearest Neighbor)	FAR = 11.4% FRR = 10.4%	120

Table 1. Summary of related works on hand geometry based identification

Five different pegs were placed on the scanner (since distances computed on the hand vary substantially with pose and finger configuration), then the user places the hand palm facing downwards, afterwards the hand image is acquired. Especially attention is placed to the localization of the center of the medium finger. The pegs are located symmetrically between the index and ring fingers. Another important point is the position of the thumb finger, because it defines the X axis inside the image, as it is shown in Fig. 4.

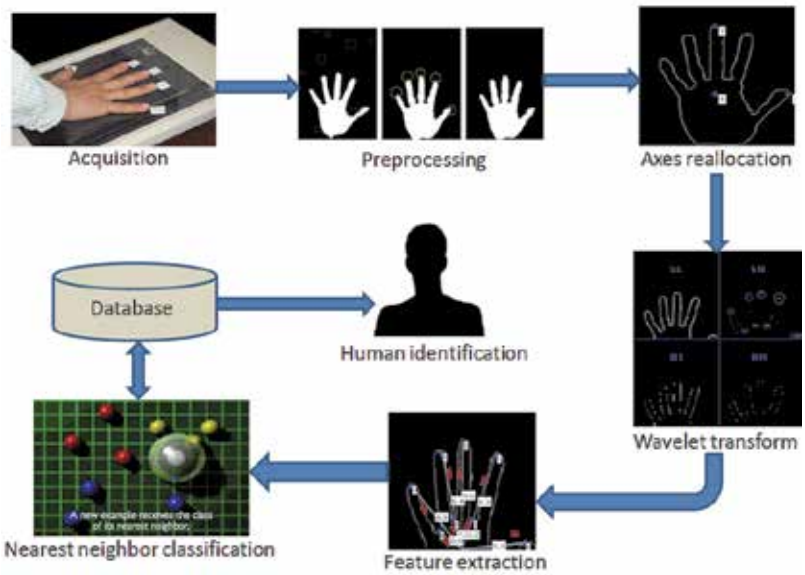


Fig. 2. Scheme of the proposed methodology.

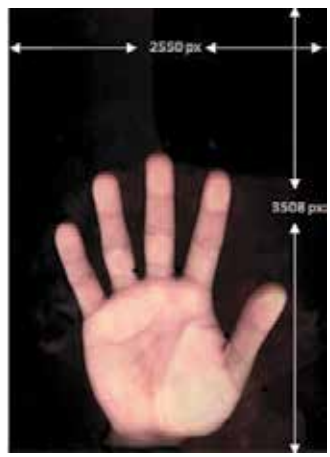


Fig. 3. Example of hand image acquired.

3.2 Preprocessing

The images were obtained in RGB plane, and then changed to grayscale. After that, the Otsu algorithm (Otsu (1979)) was used to compute a threshold in order to binarize the grayscale images. A bidimensional 4 x 4 median filter is applied to eliminate the salt noise. The next step, is to eliminate the shadows produced by the reflected scanner light against the fingernails. This is made by a morphological operation of erosion for selecting only the big objects inside the binary image (the hand). All the results obtained with image preprocessing are shown in Fig. 5.

Fig. 4. Pegs on the scanner for hand positioning.

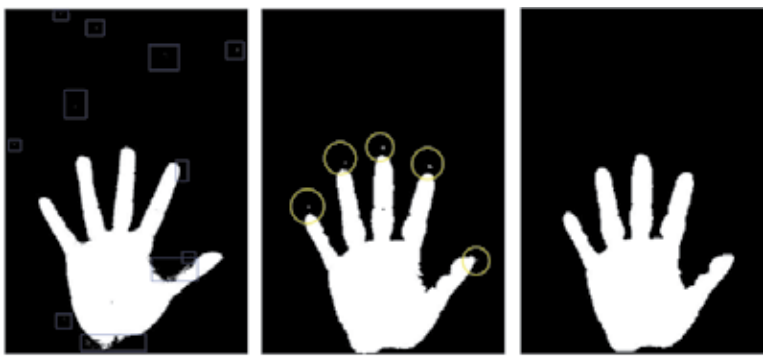


Fig. 5. Hand image preprocessing. a) Binary image, b) Denoised hand and c) Hand after shadow elimination.

3.3 Axes reallocation

After image preprocessing, the natural references of the hand are extracted in order to obtain the X and Y axes using the middle and thumb fingers. The axe Y corresponds to the skeleton of the middle finger, and the axe X corresponds to the straight skeleton of the thumb finger. With the axes located and by deduction a new origin O' is determined.

The translation of the points are computed by an scanning operation. The first pixel of the axes Y is located in order to detect the top of the middle finger. This pixel determine the new position of Y' axe. The same operations is computed to find the position of the thumb finger in X axe. This pixel determine the new position of X' axe.

After image axes reallocation, a Canny filter is implemented in order to obtain only the edges of the hand. The results obtained from axes reallocation and canny edge extraction are shown in Fig. 6.

The image of hand edges is cropped, only those sections defined by the new axes was preserved. The cut is made to avoid recognition mistakes if the people wear a bracelet. After that, all the pixels pertaining to hand contour are extracted and marked in order to obtain a matrix of xy points. The results obtained are shown in Fig. 7.

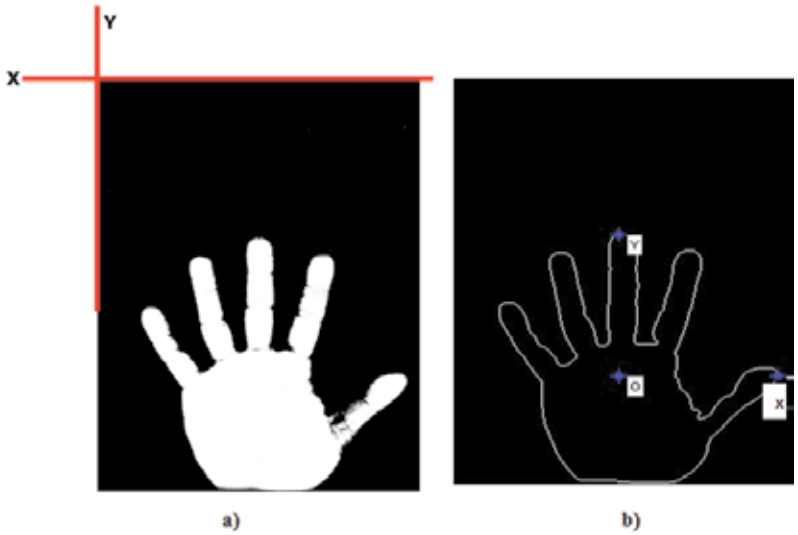


Fig. 6. Axes reallocation. a) Original axes and b) New axes on edge image.

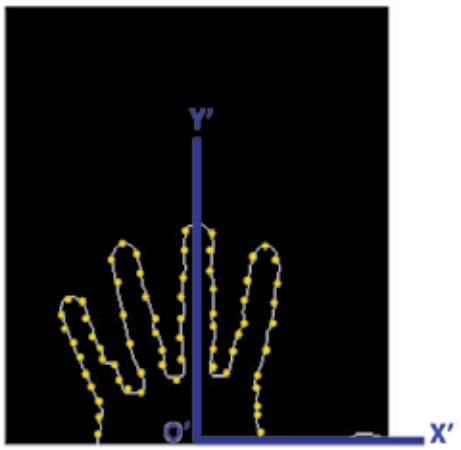


Fig. 7. Image cutting process and positioning of marks on the hand contour.

3.4 The Haar discrete wavelet transform

After axes translation the image is transformed to the wavelet domain. The Discrete Wavelet Transform (DWT) is a tool that can be applied on the discrete data to obtain a multiscale representation of the original data. From the digital point of view, the original information must be represented and delivered in efficient form. The representation efficiency, talks about the ability to capture significant information of an object of interest in a small description. From the practical point of view this representation is obtained by means of structured transformations and fast algorithms (Haar (1910)).

For this paper, the classical one level Haar decomposition is computed. The results obtained after wavelet transform are shown in Fig. 8.

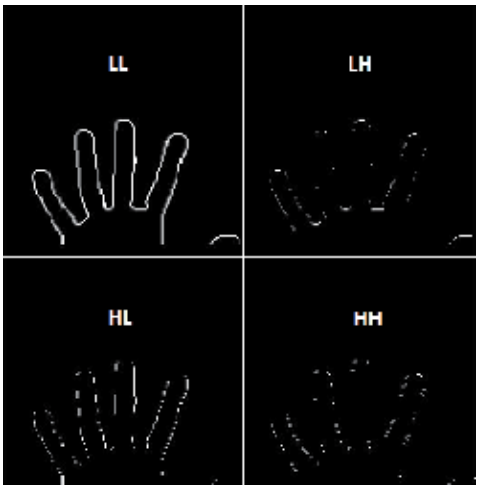


Fig. 8. The hand image transformed to Haar wavelet first level.

3.5 Feature extraction

The horizontal information obtained from LH subband is very important, because these horizontal details corresponds to the top points of the fingers and valleys of the hand. An scanning process is made on LH subband to detect the coordinates of all these points. An average is performed to select only eight control points defined as P_1 to P_8 as it is shown in Fig. 9.

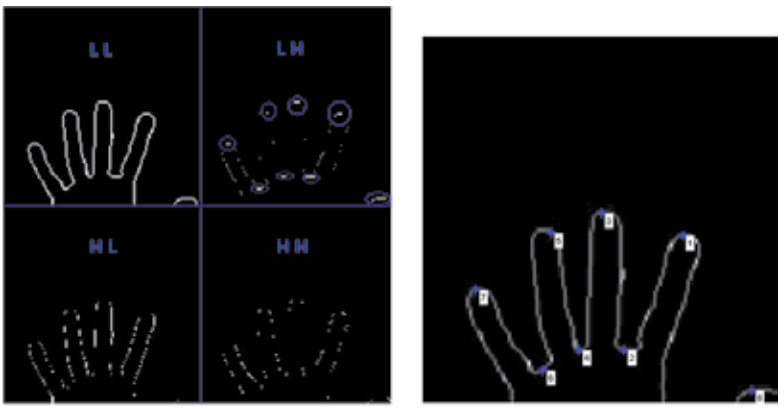


Fig. 9. Information on LH subband. a) Horizontal information and b) The eight control points.

The hand top points are defined by Equation 1 and the hand valleys points are defined by Equation 2.

$$Top\_points = \{P_1, P_3, P_5, P_7, P_8\} \tag{1}$$

$$Valley\_points = \{P_2, P_4, P_6\} \tag{2}$$

After the calculation of control points a vector with 31 W wavelet features is computed, the vector is called FV. The vector is divided into six sections as it is shown in Equation 3.

$$FV = \{f_1, f_2, f_3, f_4, f_5, f_6\} \tag{3}$$

The first group of eight features f_1 is obtained with the Euclidean distance measures between the eight control points and O' obtained in Fig. 9. The equation 4 show the computation of the first group of eight features.

$$f_1 = \{w_1 = \overline{p_1O'}, w_2 = \overline{p_2O'}, ..., w_7 = \overline{p_7O'}, w_8 = \overline{p_8O'}\} \tag{4}$$

The first group of features f_1 is shown in Fig. 10.

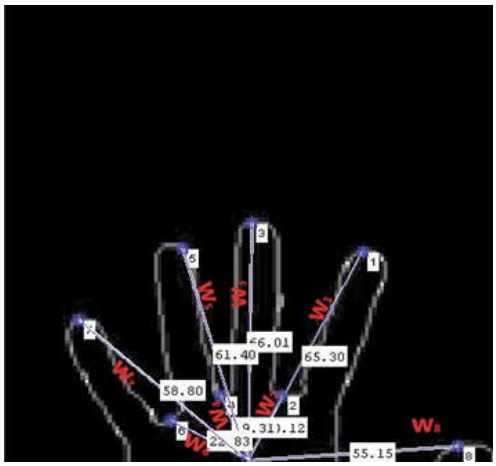


Fig. 10. The group of features f_1.

The second group of eight features f_2 is obtained with the computation of the angle between the eight control points and O', the angle is calculated with respect to the horizontal line. The equation 5 show the computation of the second group of eight features.

$$f_2 = \{w_9 = \angle p_1O', w_{10} = \angle p_2O', ..., w_{15} = \angle p_7O', w_{16} = \angle p_8O'\} \tag{5}$$

The second group of features f_2 is shown in Fig. 11.
The third group of features f_3 is obtained by means of a triangulation of the Euclidean distances between the points $p_1, p_2, p_6, p_7,$. The equation 6 show the computation of the third group of three features.

$$f_3 = \{w_{17} = \overline{p_2p_7}, w_{18} = \overline{p_7p_1}, w_{19} = \overline{p_1p_6}\} \tag{6}$$

The third group of features f_3 is shown in Fig. 12.
The fourth group of features f_4 is obtained by the computation of the Euclidean distances between the four fingers of the hand excluding the thumb. The equation 7 show the computation of the fourth group of six features.

$$f_4 = \{w_{20} = \overline{p_1p_2}, w_{21} = \overline{p_2p_3}, w_{22} = \overline{p_3p_4}, w_{23} = \overline{p_4p_5}, w_{24} = \overline{p_5p_6}, w_{25} = \overline{p_6p_7}\} \tag{7}$$

The fourth group of features f_4 is shown in Fig. 13.

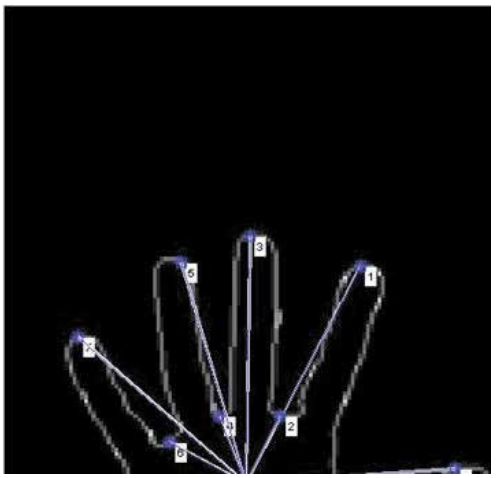


Fig. 11. The group of features f_2.

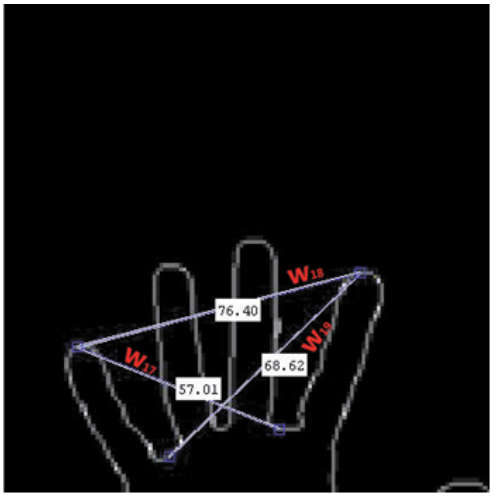


Fig. 12. The group of features f_3.

The fifth group of features f_5 is obtained by the computation of the Euclidean distances between the valleys of the hand which are distinctive features of each individual. The equation 8 show the computation of the fifth group of three features.

$$f_5 = \{w_{26} = \overline{p_2 p_4}, w_{27} = \overline{p_4 p_6}, w_{28} = \overline{p_6 p_4}\} \tag{8}$$

The fifth group of features f_5 is shown in Fig. 14.
The final group of features f_6 is obtained by the computation of the Euclidean distances between the thumb p_8, middle p_3 and pinky p_7 fingers with respect to O'. The equation 9 show the computation of the sixth group of three features.

$$f_6 = \{w_{29} = \overline{O' p_3}, w_{30} = \overline{p_3 p_7}, w_{31} = \overline{p_7 O'}\} \tag{9}$$

The sixth group of features f_6 is shown in Fig. 15.

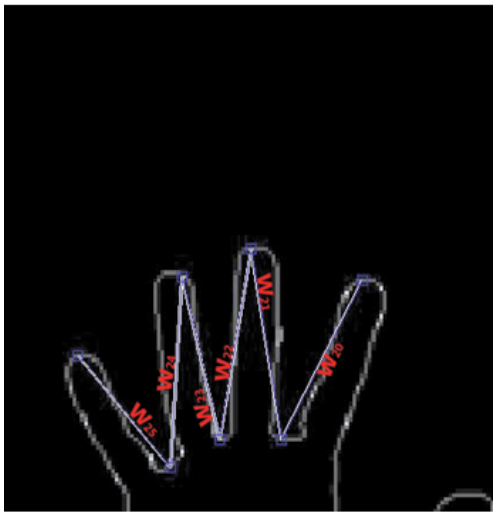


Fig. 13. The group of features f_4.

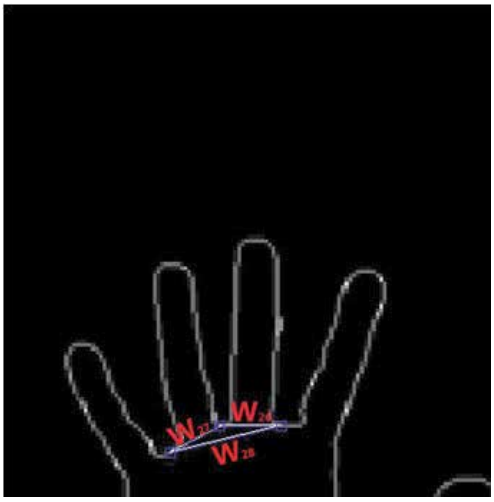


Fig. 14. The group of features f_5.

3.6 Nearest Neighbor classification

For the identification of hand geometry an algorithm of supervised learning called nearest neighbor was used. The examples are divided into training and test. A training example is an ordered pair $< x, y >$ where x is an instance and y is a label provided by a supervisor or expert. A test example is an instance x with unknown label. Then, the goal is to predict labels for test examples.

The supervised nearest neighbor algorithm was used for hand classification and works as follows (Samsudin & Bradleya (1988)): first, a database of sample hands images is created, the correct classification (label) for each image is already known, this is called the training phase. Then, when the system is given a query, i.e., a new hand to classify, the system simply computes its distance (Euclidean) to every training example and keep the k closest image in

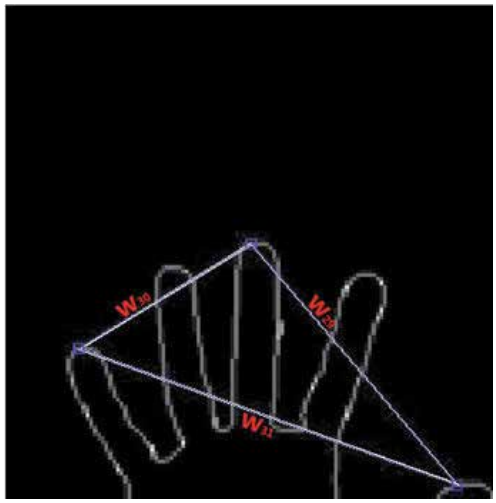


Fig. 15. The group of features f_6.

the database, in other words, finds the nearest neighbor of the query in the database. The system classifies the query as belonging to the same class as its nearest neighbor.
An example of the results obtained for classification with nearest neighbor algorithm with different wavelet features are shown in Fig. 16.

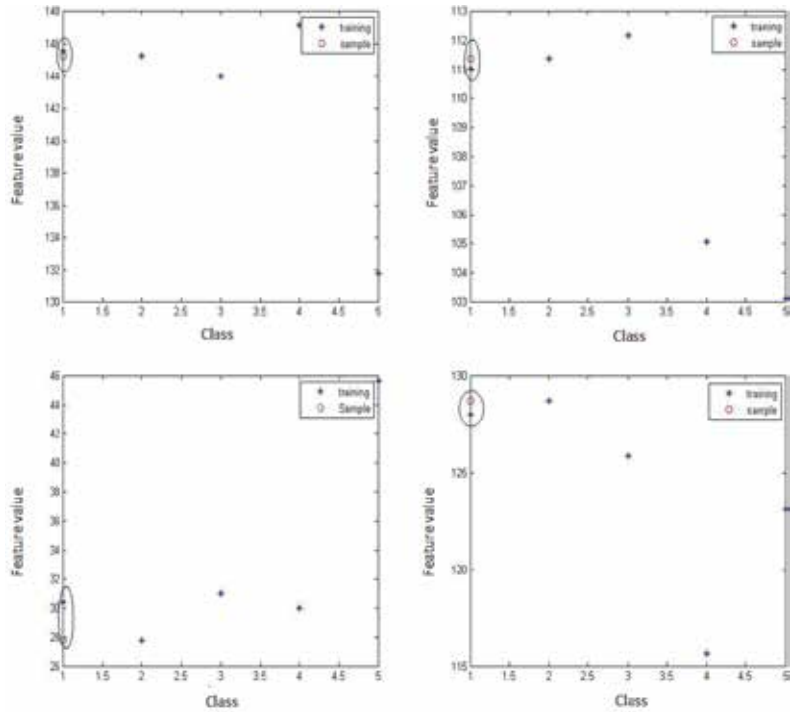


Fig. 16. Nearest neighbor results for features 5, 10, 15 and 25.

4. Experimental evaluation and results

The hand geometry identification system was trained and tested using a database of 120 hand images acquired using a commercial scanner. Additionally, images of the faces of the same individuals where acquired with a web cam and stored and relationes with the corresponding hand image. Then, when an input hand is recognized the system returns the face of the corresponding individual.
Four tests were made in order to observe the performance of the hand identification system proposed. 70 random individuals were selected for the recognition phase and a total of ten chances were executed for each person. Three different variables were measured for each of the four tests: false accept rate (FAR), false reject rate (FRR), and recognition rate (RR).
The first test, was made with natural conditions of hand images acquisition. At the second test the hands were exposed to noise by adding to the scanner pen litter. Finally, the third and four test was developed with extreme light conditions, such as high light and darkness. The results obtained for the tests are shown in Table 2.

Conditions	FRR	FAR	RR
Natural conditions	11.4%	10.4%	78.2 %
Participles of noise	30.56%	3.6%	65.84%
High light conditions	27.88%	7.0%	65.12%
Darkness	22.34%	10.0%	67.66%

Table 2. Tests results

The results showed in Table 2 demonstrated the average performance of the system, even when a simple classifier is used. Compared with the results showed in the literature our results are poor but represent a good start in a different new identifications method based on feature extraction on the wavelet domain.
Sometimes, the mistake in recognition is due to the bad position of the hand, even when the pegs are located in strategic positions. In Fig. 17 an example of a mistake in wavelet detection due to bad positioning is shown.

5. Conclusions and further works

A method to solve the old problem of human identification using a biometric cue known as hand geometry was presented. An input image of a hand was obtained using a scanner; the image is preprocessed and transformed to the wavelet domain. In the wavelet domain, 31 hand geometry features were obtained, after that, the input image is tested against 120 images of hands stored on a database. The stage of classification is performed using a simple nearest neighbor algorithm with Euclidean distances, Finally, a total recognition rate of 70.2 was obtained after experimental evaluation.
The proposed methodology can be applied in different environments such: parking lot, cash vault, interactive kiosk,anti-pass back, point of sale, time and attendance, etc.
In the future, and with the goal of obtaining better results, different more robust classifiers and different measurements distances will be tested.

Fig. 17. A mistake in horizontal information detection in wavelet domain due to bad hand positioning.

6. Acknowledgements

This work was partially supported by FOMIX CHIH-2009-C01-117569. Corresponding author: Osslan Vergara (overgara@uacj.mx)

7. References

Abatea, A.; Nappi, M.; Riccioa, D. & Sabatinoa, G. (2007). 2D and 3D face recognition: A survey. *Pattern Recognition Letters*, Vol. 28, No. 14, (October 2007) pp. 1885-1906.

Abhyankara, A. & Schuckersa, S. (2010). A novel biorthogonal wavelet network system for off-angle iris recognition. Pattern Recognition, Vol. 43, No. 3, (March 20010) pp. 987-1007.

Adán, M.; Adán, A.; Vázquez, A. & Torres, R. (2008). Biometric Verification Identification based on Hands Natural Layout, *Image and Vision Computing*, Vol. 26, No. 4, (April 2008) pp.451- 465.

Andicsa, A.; McQueena, J.; Petersson, K.; Gal, V.; Rudas, G. & Vidnyánszky, Z. (2010). Neural mechanisms for voice recognition, *Neuroimage*, Vol. 52, No. 4, (October 2010) pp. 1528-1540.

Bolok, E.; Adawy E.; Wafa A. & Mubarak, R. (2004). Person Identification through Hand Geometry Measurements. *Journal of Engineering Sciences*, Vol. 32, No. 4, (September 2004) pp.

Bulatov, Y.; Jambawalikar, S.; Kumar, P. & Sethia, S. (2004). Hand Recognition Using Geometric Classifiers. *Biometric Authentication*, Vol. 3072, Springer Berlin, 2004, pp.1-29.

Golfarelli, M.; Maio, D. & Maltoni, D. (1997) On the Error-Reject Trade-Off in Biometric Verification Systems. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.19, (July 1997) pp. 786-796.

Haar. A. (1910). Zur theorie der ortoghonalen funktionensysteme. *Mathematiche Annalen*, Vol. 69, pp.331Ű371.

Jain, A.; Ross A. & Pankanti, S. (1999). A Prototype hand geometry-based verification system, *2nd International Conference on Audio and Video Based Biometic Person Authentication (AVBPA)*, pp. 166-171, Washington, DC. USA, 22-24 March 1999.

Jain, A. & Duta N. (1999). Deformable matching of hand shapes for verification, *International Conference on Image Processing (ICIP)*, pp.857-861, Kobe, Japan, 24-28 October 1999.

Jain, A.; Flynn, P. & Ross A. (2008). Handbook of Biometrics, Springer

Jimenez, F.; Vergara, O.; Cruz, V. & Ochoa, H. (2010). Fingerprint Recognition Using Open Algorithms in Frequency and Spatial Domain, *IEEE Electronics, Robotics and Automotive Mechanics Conference (CERMA)*, pp.469-474, Cuernavaca, Morelos, Mexico, october 2010.

Kanhangad, V.; Kumar, A. & Zhang, D. (2009).Combining 2D and 3D Hand Geometry Features for Biometric Verification, *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 39-44, Miami, Florida, USA, June 2009.

Kumar, A.; Wong D.; Shen, H. & Jain, A. (2003). Personal verification using palmprint and hand geometry biometric, *2nd International Conference on Audio and Video Based Biometic Person Authentication (AVBPA)*, pp. 668-675, Guildford, United Kingdom, 2003.

Kumar, A. & zhang, D. (2007). Hand-Geometry Recognition Using Entropy-Based Discretization. *IEEE Transactions on Information Forensics and security*, Vol. 2, No. 2, (June 2007) pp. 181-187.

Lai, L. & Chaw, W. (2009). A Comparison Study on Hand Recognition Approaches, *International Conference of Soft Computing and Pattern Recognition (SOCPAR)*, pp. 364-368, Malacca, Malaysia, 4-7 December 2009.

Lay, Y. (2000). Hand shape recognition. Optics & Laser Technology, Vol. 32, No. 1, (February 2000) pp. 1-5.

Malassiotis, S.; Aifanti, N. & Strintzis, M. (2006). Personal Authentication using 3-D finger geometry", IEEE Transactions on Information Forensics and Security, vol. 1, No. 1, (March 2006) pp. 12-21.

Otsu, N. (1979). A Threshold Selection Method from Gray-Level Histograms. IEEE Transactions on Systems, Man, and Cybernetics, Vol. 9, No. 1, (January 1979) pp. 62-66.

Polat, O. & Yildirim T. (2008). Hand Geometry Identification without Feature Extraction by General Regression Neural Network. *Experts Systems with Applications*,Vol. 34, No. 2, (February 2008) pp. 845-849.

Ratha, N. & Govindaraju, V. (Eds.). (2008). Advances in Biometrics: Sensors, Algorithms and Systems, Springer, 1st edition.

Samsudin, N. & Bradleya, A. (2010). Nearest neighbour group-based classification, *Pattern Recognition*, Vol. 43, No. 10, (October 2010) pp. 3458-3467.

Sanchez, R.; Sanchez, C. & Gonzalez, A. (2000). Biometric Identification Through Hand Geometry Measurements. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 22, No. 10, (October 2000) pp. 1168-1161.

Sidlauskas, D. (1988). 3D hand profile identification apparatus, US Patent No. 4736203.

Singh, A.; Agrawal, A. & Pal, C. (2009). Hand geometry verification system: A review, *International Conference on Ultra Modern Telecommunications & Workshops (ICUMT)*, pp. 1-7, St. Petersburg, 12-14 october 2009.

Woodard D. & Flynn, P. (2005). Finger surface as a biometric identifier, Computer Vision and Image Understanding, Vol. 100, No. 3, (December 2005) pp. 357-384.

Xiong, W.; Toh, K.; Yau, W. & Jiang X. (2005). Model-guided deformable hand shape recognition without positioning aids. Pattern Recognition, Vol. 38, No. 10, (October 2005) pp. 1651-1664.

Yoruk, E.; Dutagaci, H. & Sankur, B. (2006). Hand Biometrics. Image and Vision computing, Vol. 24, No. 5, (May 2006) pp. 483-497.

Zunkel, R. (1998). Hand Geometry Based Authentication. in Biometrics: Personal Identification in Networked Society, Jain, K, (Eds), Kluwer Academic Publishers, Norwell, MA, USA.

14

Wavelet Signatures of Climate and Flowering: Identification of Species Groupings

Irene Lena Hudson, Marie R Keatley and In Kang
[1]*School of Mathematical and Physical Sciences, University of Newcastle, NSW,*
[2]*Dept of Forest and Ecosystem Science, University of Melbourne , Creswick, Vic.,*
[3]*Department of Mathematics and Statistics, University of Canterbury, Christchurch,*
[1,2]*Australia*
[3]*NZ*

1. Introduction

Phenology: the timing of biological events (life stages such as flowering, fruiting, bird arrival underpins or influences many different ecological processes (Dunlop and Brown 2008; Forrest and Miller-Rushing 2010). These processes also have a significant role in shaping society's values (e.g. on human health, biodiversity, forestry, agriculture and tourism (Beggs 2004; Fitter and Fitter 2002; van Vliet 2010)). Since the 1990s, primarily because of climate change (Keatley & Hudson 2010, Parmesan 2006, Root et al., 2008, Schwartz 2003, Sparks 1995, 2002) phenological time series have been used to determine and report the impacts of global warming in both natural and managed systems (Menzel et al. 2006; Rosenzweig et al. 2008; Sparks et al. 2005). Determining trends in relation to long-term climate, however, is not easy as trends can be confounded by short-term interannual trends. Hence not only are long term records required, but also needed is the development of novel statistical methods which can deal with confounding factors (Badeck et al. 2004; Hudson 2010; Hudson & Keatley 2010a).

Wavelet methods (Daubechies, 1992) have been extensively applied to many arenas (eg. to the study of change in European spring temperatures (Paluš et al. 2005) and rainfall (Koch and Marković 2007), changes in vegetation cover (Lu et al. 2007), and to brain imaging (Bullmore et al. 2003; Sendur et al. 2007). It is the ability of wavelets to cope with non-stationary data: to deconstruct a time series into its subcomponents and remove noise; to accommodate multi-scale information, and to minimize correlation and time-dependency in data (Cornish et al. 2006; Gencay et al. 2001; Percival and Walden 2000; Vidakovic (1999)) that have added to their popularity. As phenological time series are usually non-stationary and noisy, and as such wavelet methods present as a useful analytic method (Hudson et al., 2005, Hudson 2010, Hudson et al., 2010a,b) for examining phenological records and for the determination of possibly changing climatic impacts on flowering, at an annual and across years basis.

The utility of wavelets in investigating the relationship of flowering to climate (three temperature variants and rainfall) is shown in this chapter by examining the flowering intensity time series records of eight Australian eucalypts – namely, *E. camaldulensis*,

E. goniocalyx, *E. leucoxylon*, *E. macrorhyncha*, *E. melliodora*, *E. microcarpa*, *E. polyanthemos* and *E. tricarpa*. This work builds on an initial study by Kang et al. (2004) and on the early premise of Hudson et al. (2005) that wavelets *per se* could add integrity to the use of phenological records to detect climate change. This premise was recently confirmed by a study of 4 of the 8 Eucalypt species studied in this chapter by Hudson et al. (2010a, b) (see also Hudson (2010) and Keatley & Hudson, 2010).

The discrete wavelet transform (DWT), following the development of Percival and Walden (2000) and the maximal overlap DWT (MODWT) (Percival and Mofjeld 1997) is applied in this chapter. The rationale for this approach is that, given the resultant MODWT coefficients, the original (flowering) time series could be reconstructed as an additive decomposition - known as a multiresolution analysis (MRA) (Bratteli and Jorgensen 2002) and also that the individual detail (sub-component) series could be examined.

The aim of this research is to demonstrate the utility of wavelet analysis in phenology by extending the recent work of Hudson et al., (2010 a,b) from four to eight eucalypt species. This chapter contributes significantly to our understanding of the interplay between climate and the flowering of *Eucalyptus* flowering – a major southern hemisphere genus.

2. Methods

2.1 Maximal Overlap DWT (MODWT) and Multiresolution Analysis (MRA)

The discrete wavelet transform (DWT) as applied here, following Percival and Walden 2000, requires a discretization of the continuous time variable, Y_t. For most practical applications, the DWT, which analyzes signals over a discrete set of scales, that are usually sampled at dyadic sequences ($\lambda_j = 2^{j-1}$, j= 1, 2, 3, . . .), is sufficiently accurate and can recover signals perfectly (Mallat 1989). The maximal overlap DWT (MODWT) is one discretization choice which gives N wavelet coefficients for each scale (Percival and Guttorp 1994). The MODWT is a non-decimated variation of the DWT (Percival and Mofjeld 1997). It is also equivalent to the original time series, in the sense that, given the MODWT coefficients, Y, can be constructed as an additive decomposition, which is known as a multiresolution analysis (MRA) (Hernández and Weiss 1996, Gencay et al. 2001, Bratteli and Jorgensen 2002). The MRA decomposition is as follows:

$$Y = \sum_{j=1}^{J_0} \bar{d}_j + \bar{s}_{j_0} \tag{1}$$

where $\bar{d}_j$ are the "detail series", (j= 1, 2, . . ., J_0) for a pre-specified J_0, and are part of the MRA of Y that can be attributed to variations on a scale of λ_j (Gencay et al. 2001, Percival and Walden 2000). Recall that scale $\lambda_2 = 2$. In equation (1) $\bar{s}_{j_0}$ is a N dimensional vector, which depends on the scaling coefficients. The vector, $\bar{s}_{j_0}$, is called the "smooth series" since it is associated with averages over scales $2\lambda_{j_0}$ and longer. Note that $\bar{s}_{j_0}$ captures the slowly varying portion of Y, which is often considered to be the overall trend.

2.2 Wavelet correlation and cross-correlation

The above section considers decomposing the original time series into J_0+1 subcomponents. However, the scale λ_j MODWT coefficients may also be used to examine the wavelet

correlation and wavelet cross-correlation of bivariate time series X_t and Y_t (Serroukh and Walden 2000, Whitcher et al. 2000, Gencay et al. 2001), as described below.
The wavelet correlation (WCORR) of (X_t, Y_t) at scale $\lambda_j=2^{j-1}$ is defined as

$$\rho_{XY,\tau=0}(\lambda_j)=\frac{\mathrm{Cov}\left\{\overline{W}_{j,t}^{X}\,\overline{W}_{j,t}^{Y}\right\}}{\sigma_X(\lambda_j)\cdot\sigma_Y(\lambda_j)}=\frac{\gamma_{XY,\tau=0}(\lambda_j)}{\sigma_X(\lambda_j)\cdot\sigma_Y(\lambda_j)} \tag{2}$$

where $\sigma_X^2(\lambda_j)=\mathrm{var}\left\{\overline{W}_{j,t}^{X}\right\}$ is the wavelet variance with scale λ_j. $\overline{W}_{j,t}^{X}$ and $\overline{W}_{j,t}^{Y}$ are the scale λ_j MODWT coefficients for X_t and Y_t, respectively (Percival 1995). Note that Equation 2 is a theoretical quantity that is well-defined under an assumption of stationarity.
For a time lag τ, between the two series, X_t and Y_t, the wavelet cross-covariance and wavelet cross-correlation (denoted by WCCORR) (Gencay et al. 2001), for scale $\lambda_j=2^{j-1}$ and time lag τ, is

$$\rho_{XY,\tau}(\lambda_j)=\frac{Cov\left\{\overline{W}_{j,t}^{X},\overline{W}_{j,t}^{Y}\right\}}{\sigma_X(\lambda_j).\sigma_Y(\lambda_j)}=\frac{\gamma_{XY,\tau}(\lambda_j)}{\sigma_X(\lambda_j).\sigma_Y(\lambda_j)} \tag{3}$$

By setting τ = 0, $\gamma_{XY,0}(\lambda_j)$ reduces to the wavelet variance for X_t or Y_t denoted by $\sigma_X^2(\lambda_j)$ or $\sigma_Y^2(\lambda_j)$, respectively. The wavelet cross-correlation in equation 3 is thus able to provide the lead or lag relationship on a scale-by-scale basis (Gencay et al. 2001), just as is the case for conventional cross-correlations. These can determine lead or lag relationships between two time series (Percival et al. 2000).

2.3 Confidence interval determination

Construction of the 95% two-sided confidence interval (CI) of WCORR and WCCORR follows the development of Gencay et al. 2001. To produce CIs for the cross-correlation coefficient, the Fisher's nonlinear z-transformation $h(\rho)$ (Percival and Walden 2000), as follows, $h(\rho)=1/2\log(1+\rho/1-\rho)=\tanh^{-1}(\rho)$ is required. For $\hat{\rho}_X$ which is an unbiased estimator of WCORR based on the MODWT, the following asymptotic normal distribution holds

$$\sqrt{N-3}\left[\,h(\hat{\rho})-h(\rho)\,\right]\sim N(0,1). \tag{4}$$

Applying the transformation tanh maps the confidence interval back to [-1, 1] to produce an approximate 95% CI for $\rho_X(\lambda_j)$ as follows (Whitcher et al. 2000, Gencay et al. 2001),

$$\tanh\left\{\mathrm{h}\left[\hat{\rho}_X(\lambda_j)\right]\pm 1.96\left(\frac{1}{\overline{N}_j-3}\right)^{1/2}\right\}. \tag{5}$$

The quantity $\overline{N}_j$ in equation (5) is the number of DWT coefficients associated with scale λ_j.

2.4 Data analyses

DWT, MODWT and MODWT-MRA analyses were performed using code based on the wavelet methods developed by Percival and Walden (2000) and Gencay et al. (2001).

2.5 Phenological data

The phenological data was collected between 1940 to 1970 and comprises the flowering intensity profiles of eight eucalypts species growing in the region of Havelock, Victoria, Australia. The eight species examined in this study were: *Eucalyptus camaldulensis*, *E. melliodora, E. polyanthemos; E. goniocalyx*, *E. microcarpa, E. macrorhyncha, E. leucoxylon* and *E. tricarpa* (Table 1).

Observations on the timing, quantity and distribution of flowering of these species were collected on a monthly basis (Keatley et al. 1999) (Table 1). Flowering intensity (ranging from 0 to 5) of each species was quantified by assigning a rank value (Table 1) (Keatley and Hudson, 2007), thus producing a discrete time series (see Fig. 1 where only the period from Jan 1940 to Dec 1955 is shown for visual clarity).

A score of 0 indicates that no flowering occurred, in the given month, whilst a score of 5 indicates that flowering was heavy and distributed throughout the observation area. DWT methodology (Gencay et al. 2001; Whitcher et al. 2000) was applied to the afore-mentioned eight species' flowering intensity scores (0-5). This data comprises a rare long-term 30 year record of flowering (1940-1970) of over 350 monthly time points (Keatley et al., 2002).

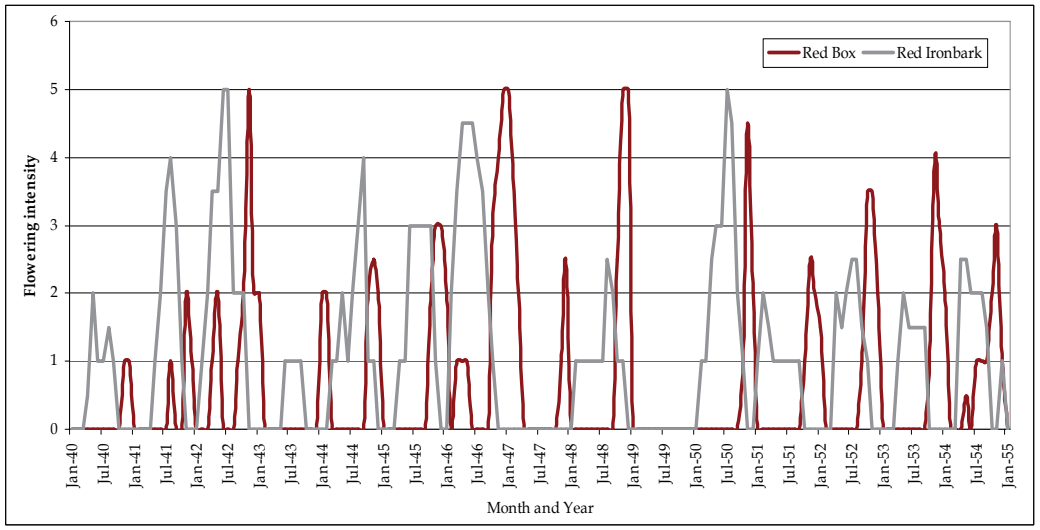


Fig. 1. Time series of E. *polyanthemos* (red box) and *E. tricarpa* (red ironbark) for the period between January 1945-January1955(wherein only the period January 1945 to January 1955 is shown for visual clarity).

Species	Common name & taxonomic authority	Median duration	Most probable month	Peak Flowering#	Flowering start month	Flowering cessation months	Flowering intensity score	Description	Assigned Value
E. camaldulensis	River red gum Dehnh	4.0	January	January	November /December	January	Quantity	No flowering	0
E. goniocalyx	Long-leaf box F. Muell	3.0	March/April	March	February	May		Very scattered or isolated	0.5
E. leucoxylon	Yellow gum F. Muell	9	October	September	May	December		Light Flowering	1
E. macrorhyncha	Red stringy bark F. Muell	4.0	March/April	January	January	May		Medium Flowering	2
E. melliodora	Yellow box Cunn	5.0	January	January	November	March		Heavy Flowering	3
E. microcarpa	Grey box Maiden	4.0	March	March	February	May	Distribution	Isolated	0.5
E. polyanthemos	Red box L. A. S. Johnson & K. D Hill	3.5	November	November	October	December		Scattered	1
E. tricarpa	Red ironbark L. A. S. Johnson & K.D Hill	6.0	July	July	April	September		Fairly General	1.5
								General	2

ϕ Based on actual likelihood/probabilities for Havelock (Keatley and Hudson, 2007) for all species, except for *E. camaldulensis, E. goniocalyx* and *E. macrorhyncha*.
The month with the highest mean intensity within a flowering year.

Table 1. Eight eucalypt species' median flowering duration and months in which the main phenophases occur. Terms describing flowering intensity and their assigned value (Keatley 1999) (Table sourced from Hudson et al., 2011a).

Daily minimum and maximum temperature along with daily rainfall were obtained from the Bureau of Meteorology, for the closest weather station (approximately 3.5 km away), in Maryborough, Victoria (37°03'22"S 143°43'55"E, 249.3 m elevation). The temperature dataset had some missing data. When only one day was missing, the average of the temperature either side of the missing date was used. If two or more days were not recorded, then the mean minimum or maximum temperature for that month was substituted. Daily mean temperature was calculated as the average of daily minimum and maximum temperature. Monthly means for minimum, maximum, diurnal temperature and rainfall were used in the analyses.

3. Results

3.1 MODWT-MRA

Maximal overlap discrete wavelet transform – multi-resolution analysis (MODWT-MRA) with $J_0 = 4$ provided the subcomponents of the flowering signal within each species (see Figure 2). The raw series is given by X, d1 to d4 are the different subcomponents (detail) of the original series X. Traditionally, d1 to d4 are associated with changes in averages over 1, 2, 4 and 8 months, respectively (Figure 2). Also s4, which is the smoothed series associated with averages over 8 months, is also provided by MODWT-MRA (see Figure 2 for *E. tricarpa*).

The first two sub-components reflect duration and the overall pattern of the original flowering data. d1 and d2 appear similar in their profiles (see Figure 3 for *E. macrorhyncha*), however, d1 contains subcycles per year (in a flowering year) but it is unclear as to what these cycles are related.

The annual cycle is delineated by d3, and d4 relates to the annual or biennial cycling of intensity of flowering years. S4 is the smoothed series and indicative of the overall trend.

d3 not only delineates the annual cycle of flowering, but reveals (by its peaks), in the majority of cases, the month of peak flowering in a flowering year (or within the month on either side) (Figure 4). Note that for *E. Goniocalyx,* all but one peak is selected (97%), however, this does include two years (1965 and 1966) when three months exhibited equal flowering intensity and for these years the middle month was nominated as the peak month. The selection of the middle month also occurred in *E. marcrorhyncha* (overall 68% peak selection) in five years (1945/46; 1952/52, 1957/58, 1960/61 and 1967/68). *Eucalyptus marcrorhyncha* flowering period usually includes December and January therefore two years are given. For the remaining species, *E. melliodora* 97%, *E. camaldulenis* 90%, *E. leucoxylon* 73%, *E. microcarpa* 72%, for *E. polyanthemos* 69% and for *E. tricarpa* 65%.

Additionally, d3 highlights years when flowering does not occur – this is seen in Figure 4 for *E. goniocalyx* when d3 is near zero for the years 1944, 1945, 1948, 1950, 1953, 1955 and 1956. This occurred in all species except *E. leucoxylon* which always flowered (Figure 4). For the other species, years of non-flowering are given in Table 2. As with *E. Macrorhyncha,* three other species, *E. camaldulensis*, *E. melliodora* and *E. polyanthemos* flowering periods also include December and January so two years are also given. The year 1958 or 1958/59 as a non-flowering year was common to five species: *E. camaldulensis, E. melliodora, E. microcarpa*, *E. polyanthemos* and *E. tricarpa.*

d4 outlines the flowering intensity for each species. For all species, except *E. leucoxylon* the cycle is associated with years of low flowering intensity (< 2) and also includes years of no flowering (i.e. flowering intensity = 0). The longest average period between peaks was

E .goniocalyx with 29.0 months with a range of 12 to 48 months. This is particularly influenced by the large number of non-flowering years (Table 2). The species with the smallest average number of months between peaks was *E. melliodora* (18.5 months with a range of 11 to 25 months). For the remaining species the average between peaks was *E. leucoxylon* 19.1 months, *E. marcrorhyncha* 19.5 months, *E. polyanthemos* 20.3 months, *E. camaldulensis* 20.5 months, *E. tricarpa* 21.3 months and for *E. microcarpa* 22.5 months (Figure 5). All species had at least one year between the period 1952 and 1958 when the number of months between peaks was between 11 to 14 months. For *E. camaldulensis* this occurred for five years between December 1953 and March 1958 (Figure 5).

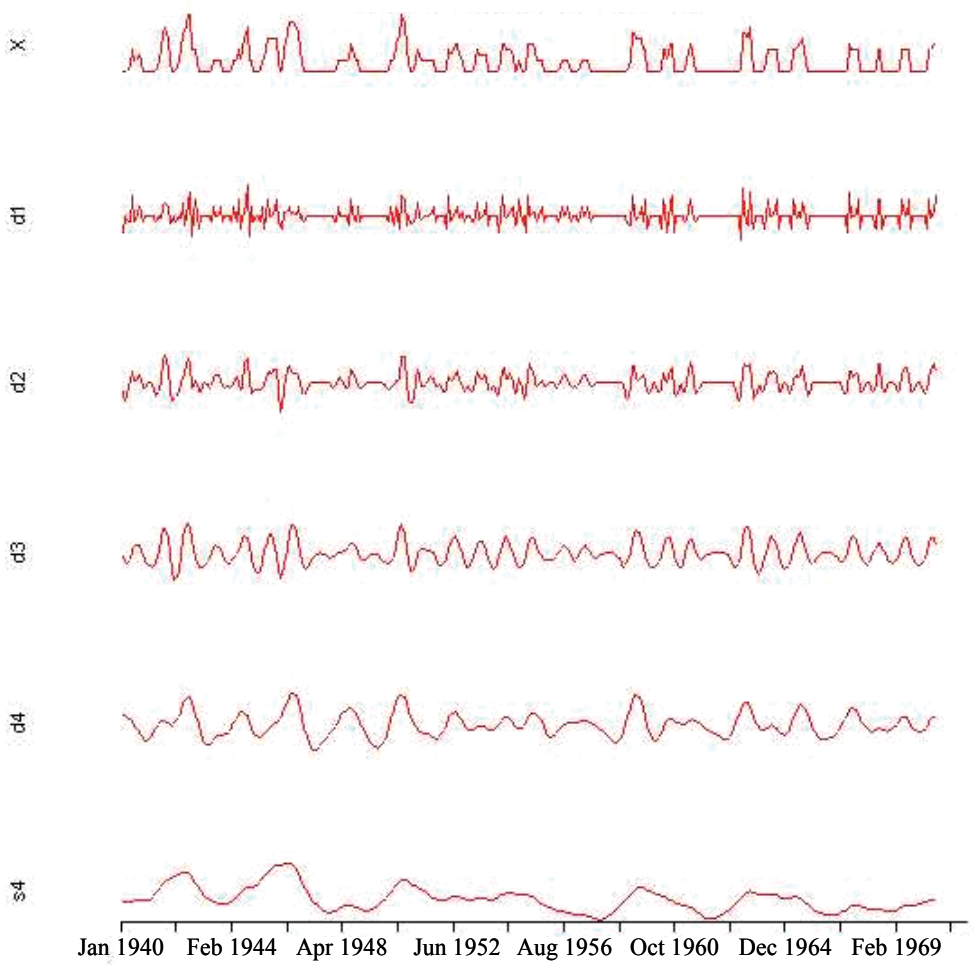


Fig. 2. Example of MODWT-MRA output for *E. tricarpa*. X = the raw data, d1 to d4 different subcomponents (details) of the original series X, and s4 = the smooth series

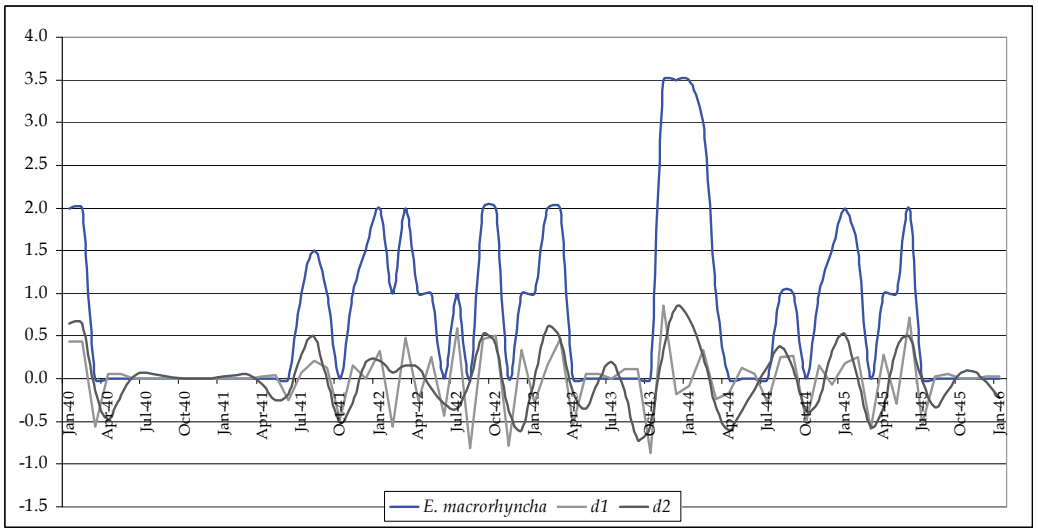


Fig. 3. Raw data, d_1 and d_2 of *E. macrorhyncha* between January 1940 and January 1946.

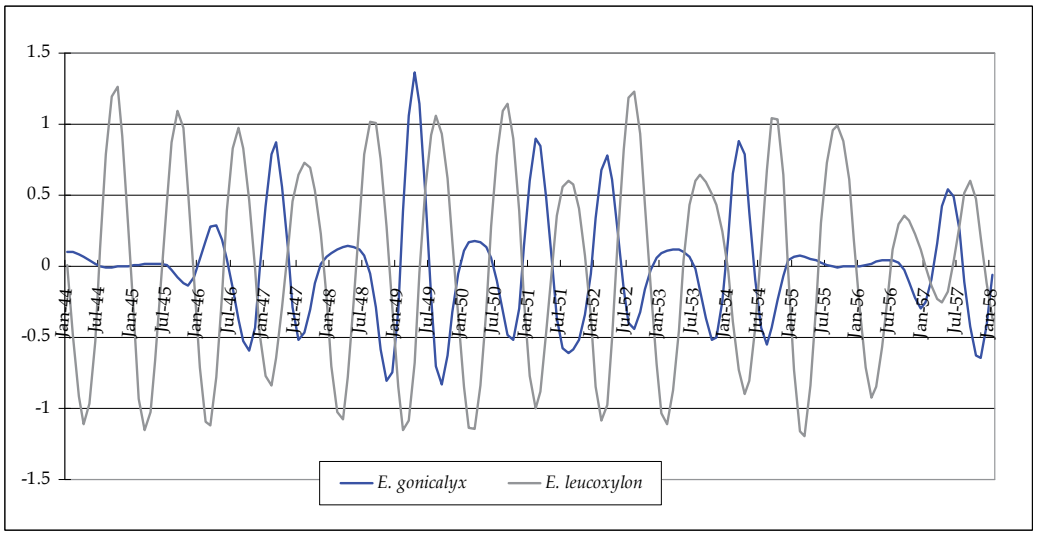


Fig. 4. d_3 subcomponent for *E. goniocalyx* and *E. Leucoxylon* between Jan. 1940 and Jan. 1946.

Although s4 is the smoothed series it is also reflective of the trend and variation in flowering intensity. s4 indicates that flowering has become less intense from 1940 to 1970 for all species except *E. camaldulensis* and *E. goniocalyx* (Figure 6A &B). These two species do not change their flowering intensity. *Eucalyptus leucoxylon* on average (mean intensity 1.63) flowers more intensely than the other 7 species except for two brief periods, between June 1945 and July 1946, when *E. tricarpa* flowered the most (Figure 6B) and January and August 1962 during which *E. melliodora* did. *Eucalyptus melliodora* is the next most intense flowering species (mean intensity = 0.97) (Figure 6A) and *E. goniocalyx* is the species which has flowered with the least intensity (0.37) (Figure 6B).

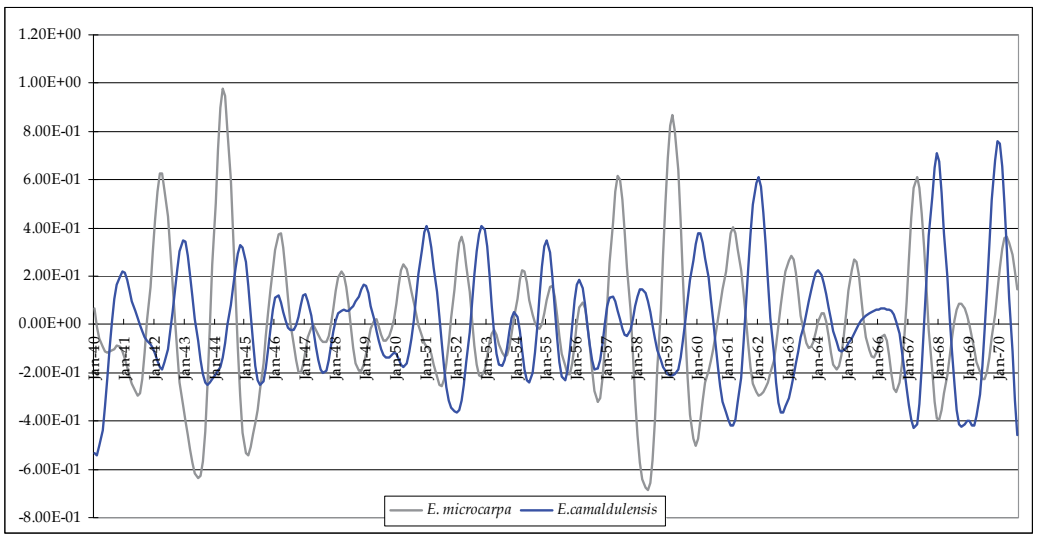


Fig. 5. d_4 subcomponent of E. microcarpa and E. Camaldulensis.

A trough indicating the non flowering year in all species except *E. leucoxylon* and *E. macrorhyncha,* where it is indicative of lower flowering intensity, is noticeable between March 1958 (*E. microcarpa*) to July 1959 (*E. macroryhncha*) (Figure 6A & B)). The most intense flowering period for all species (except *E. camaldulensis* and *E. Goniocalyx)* occurred prior to 1958. Hudson et al. 2005 noted wetter years than average in the mid 1940's to the end of the 1950's; and maximum temperature at Maryborough was increasing until late 1967/early 1968 (also reflected by the Southern Oscillation Index (Trenberth et al., 1996, Ghil et al., 2002).

Species	Non-flowering years
E. camaldulensis	1951/52, 1958/59, 1960/61, 1962/63 and 1966/67
E. goniocalyx	1942, 1944, 1945, 1948, 1950, 1953, 1955, 1956, 1959, 1961, 1963, 1964 and 1967
E. leucoxylon	Always flowered
E. macrorhyncha	1940/41, 1953/52, 1954/55, 1963/64 and 1969/70
E. melliodora	1958/59
E. microcarpa	1943, 1958, 1962, 1968 and 1969
E. polyanthemos	1949/50, 1956/57, 1958/59 and 1961/62
E. tricarpa	1947, 1949, 1958, 1962 and 1966

Table 2. Species and their years of non-flowering

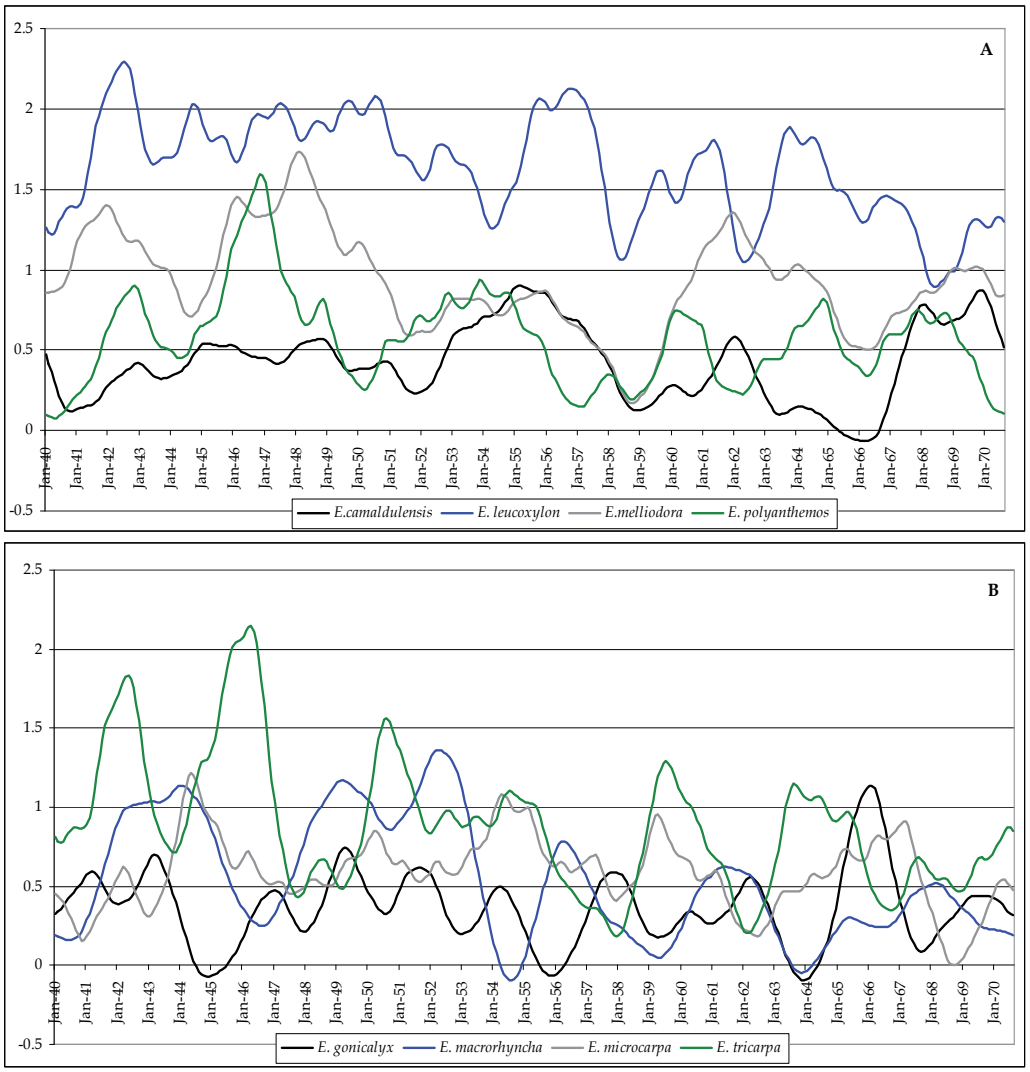


Fig. 6.A & B. s_4 subcomponents for each group of four species.

3.1.1 Wavelet correlation of flowering with temperature and rainfall

The wavelet correlation between minimum temperature and rainfall with peak flowering intensity at different wavelet levels or scales (1, 2, 4, 8, 16 and 32 months) were examined for all species. See Figure 7 and Figure 8, where significant ($P < 0.05$) wavelet correlations are those, where the upper (U) and lower (L) confidence limits are on the same side of the zero horizontal line. Correlations at 1, 8, 16 and 32 months with respect to minimum temperature (Figure 7) and rainfall (Figure 8) were not significant for any species. It can be seen that correlations with minimum temperature at either scale 2 (2^{2-1} = 2 months) or 3 (2^{3-1} = 4 months) are significant for all species, except for *E. goniocalyx* (Table 3, Figure 7).

For *E. tricarpa* minimum, maximum and mean temperatures have a significant negative relationship at these scales (Figure 7 and Table 3). *Eucalyptus tricarpa* was the only species

for which rainfall had a significant, positive relationship with flowering at the 4 month scale (Figure 8). This indicates that cooler, wetter periods lead to greater flowering intensity for *E. tricarpa*. The relationship for *E. leucoxylon* and rainfall was positive but was not significant, as in *E. tricarpa.* In contrast two species, *E. camaldulensis* and *E. melliodora* had a significant negative relationship with rainfall also at the 4 month scale (Table 3). Both of the latter species also had a significant positive relationship with all the temperature variables.

Species/ Period (months)	Tmean		Tmin		Tmax		Rainfall	
	2	4	2	4	2	4	2	4
E. leucoxylon	-0.49*	-0.59*	-0.52*	-0.62*	-0.46*	-0.56*	0.03	0.28
E. tricarpa	-0.51*	-0.75*	-0.49*	-0.72*	-0.51*	-0.76*	0.20	0.40*
E. polyanthemos	0.15	0.48*	0.14	0.43*	0.16	0.50*	-0.06	-0.17
E. camaldulensis	0.34*	0.75*	0.34*	0.74*	0.33*	0.75*	-0.01	-0.30*
E. melliodora	0.54*	0.86*	0.52*	0.84*	0.54*	0.87*	-0.09	-0.41*
E. gonicalyx	0.13*	0.23*	0.14	0.27	0.12	0.21	-0.06	-0.22
E. microcarpa	0.18	0.27	0.23*	0.31*	0.15	0.25	-0.01	-0.25
E. macrorhyncha	0.24	0.43	0.25*	0.44*	0.24*	0.42*	-0.05	-0.28

Table 3. Wavelet correlation of peak flowering intensity and climate variables (* P< 0.05)

For *E. camaldulensis* both mean temperature and maximum temperature at the 4 month scale had equal influence (Table 3), whereas for *E. melliodora* it was maximum temperature (and then mean temperature) at this scale that had the greatest impact. Therefore for these species warmer and dryer periods result in greater flowering intensity. For the remaining species a negative correlation (although not significant) with rainfall was indicated.

All temperature variables (mean, minimum and maximum) at 2 and 4 months were also significant for *E. leucoxylon*. At peak flowering intensity both *E. leucoxylon* and *E. tricarpa* have a negative relationship with each of the temperature variables, indicating increased intensity of flowering with decreasing temperature. However, there is a slight difference in the strength of the relationship of flowering across the temperature variants (Table 3). It is noteworthy that for *E. leucoxylon* the main relationship with flowering intensity is with minimum temperature and for *E. tricarpa* it is with maximum temperature (Table 3), both significant at the 4 month scale.

For each of the remaining six species an inverse relationship between flowering and the temperature variants was shown (Table 3 and Figure 7): that is an increase in flowering intensity with increasing temperature. However, for the species pair (*E. goniocalyx* and *E. microcarpa*) there was only one temperature variant which impacts significantly on flowering. For *E. goniocalyx* it is mean temperature and for *E. microcarpa* minimum temperature, indicating that for the individual species these climate predictors may be the primary climatic driver of flowering. In contrast, for *E. polyanthemos,* there was a significant positive relationship with all three temperature variables (at 4 months) with maximum

temperature being identified as having a slightly greater influence on flowering (Table 3). This is in agreement with a recent Generalised Additive Model for Location, Scale and Shape (GAMLSS) analysis of the data studied here (Hudson et al. 2009; 2011b). Flowering intensity in *Eucalyptus macrorhyncha* is influenced positively by both minimum and maximum temperatures with minimum temperatures at the 4 month scale being the most significant.

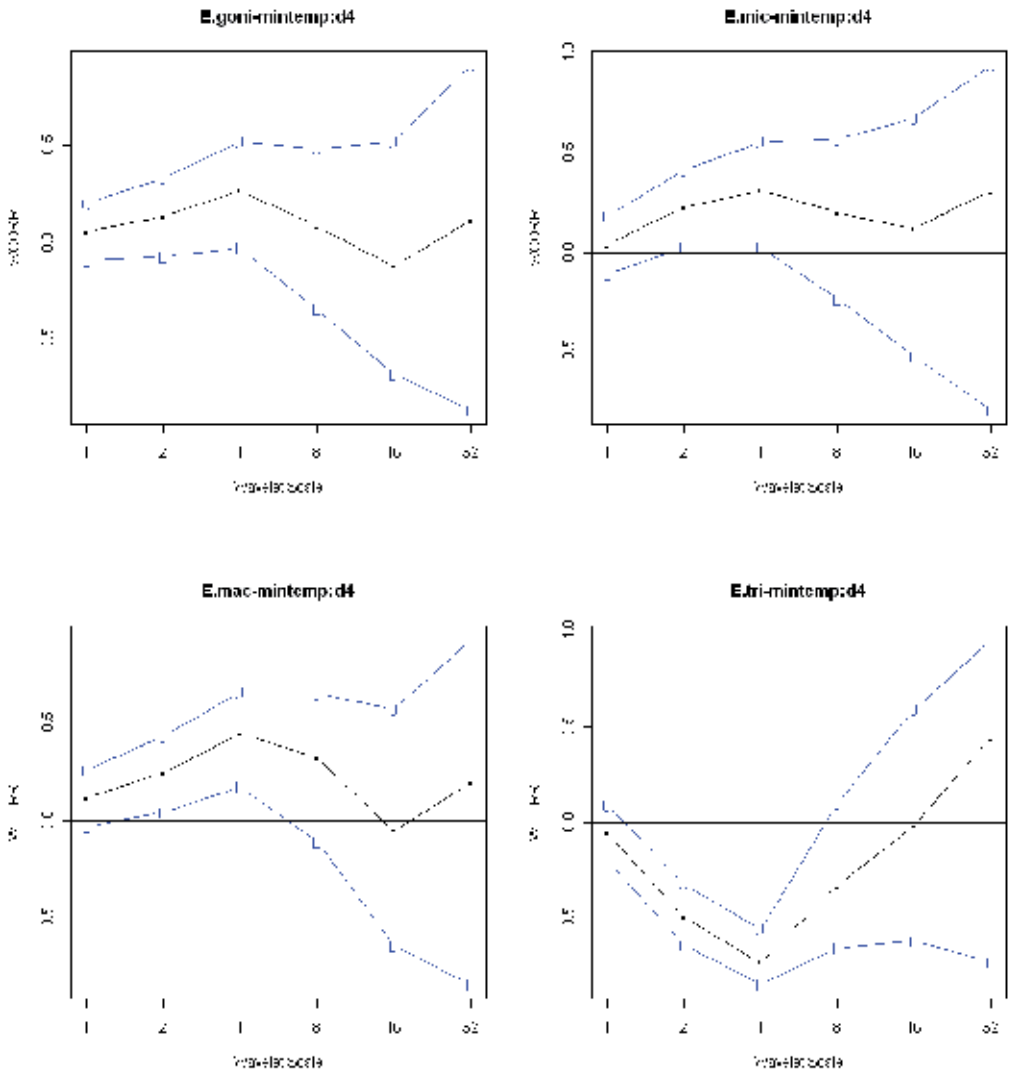


Fig. 7. Wavelet correlations (WCORR) of the flowering of *E. goniocalyx*, *E. microcarpa, E. macrorhyncha* and *E. tricarpa* with minimum temperature (95% confidence interval limits indicated by L and U).

Broadly speaking, *E. leucoxylon* and *E. tricarpa* can be grouped together with respect to their relationship with temperature and rainfall. The relationship of the remaining six species with climate is similar, however, when the significance of the relationship is taken into account two other species clearly grouped together: *E. camaldulensis* and *E. melliodora*. As already discussed they are the only two species which have a significant negative relationship with rainfall (similar WCORR signatures between their flowering and rainfall (Figure 8)) and a significant positive relationship with all three climate parameters at both

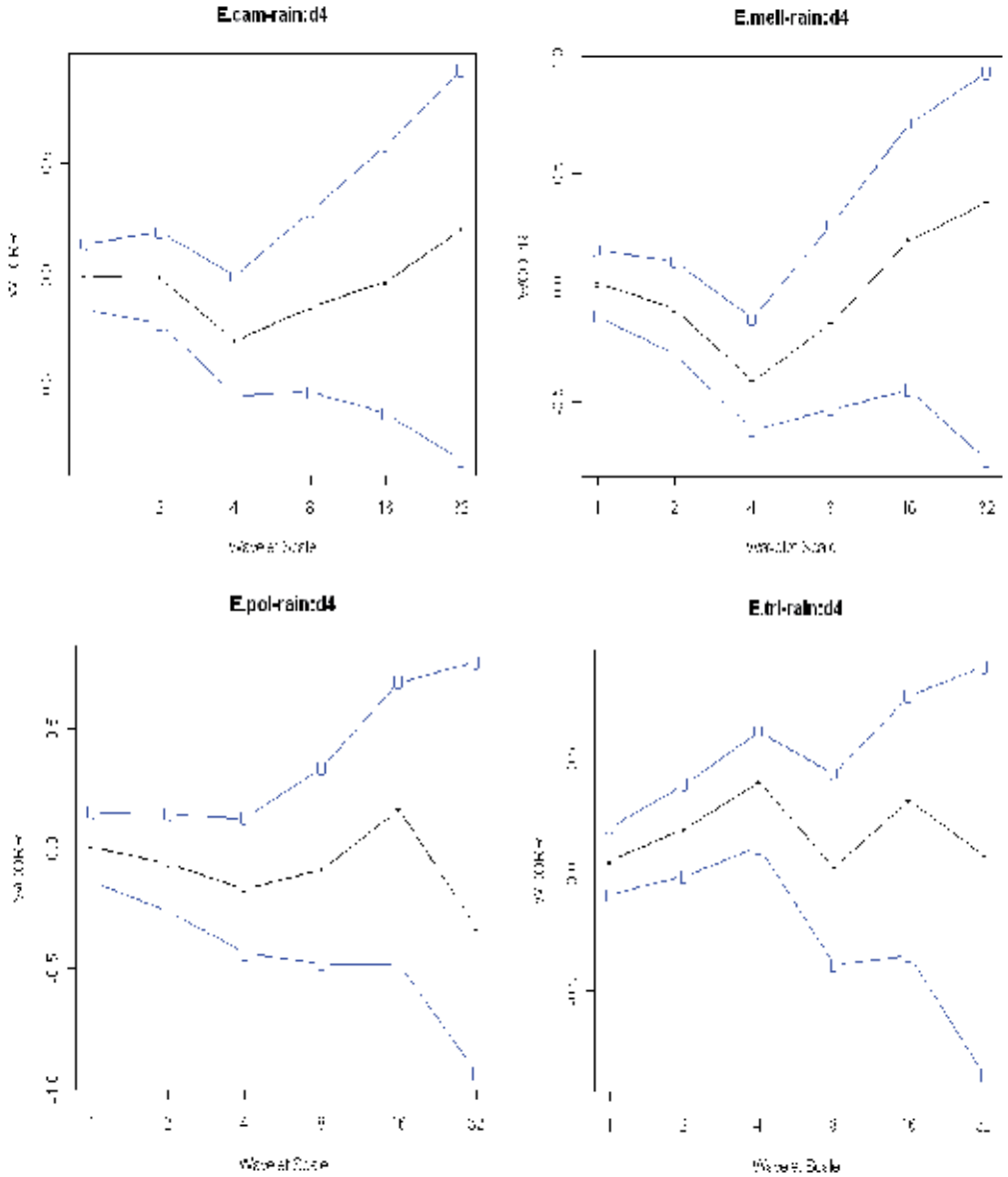


Fig. 8. Wavelet correlations (WCORR) of the flowering of *E. camaldulensis*, *E. melliodora*, *E. polyanthemos* and *E. tricarpa* with rainfall (95% confidence limits indicated by L and U).

scales. *Eucalyptus polyanthemos* aligns with this group as it also has a significant positive relationship with all three climate parameters, but only at the 4 month scale. Additionally, its main influence is maximum temperature as is the case with the other two species (albeit that *E. camaldulensis* is equally influenced by mean temperature).
Theother three species (*E. goniocalyx*, *E. microcarpa, E. macrorhyncha*) also broadly group, as evidenced by their similar climate signatures as shown in their WCORR profiles (Figure 7). Note that all members of this species triple, have a positive relationship between flowering and at least one temperature variant; and a negative (though not significant) relationship between flowering and rainfall.

3.1.2 Wavelet cross-correlation of flowering with temperature and rainfall variables

Wavelet cross-correlations were examined for all wavelet scales (1 to 32 months). Significant relationships ($P \leq 0.005$) were found only at 4 months (level 3) for rainfall; but at 2, 4 and 8 months (levels 2, 3 and 4) for the temperature variants (Figure 9 illustrates this for *E.leucoxylon*). The strongest (highest absolute value of the correlation profile) and most significant relationship for all species with temperature was at 4 months (level 3) (Figure 9). This mirrors the correlation results (Table 3), where level 3 correlations were significant. Recall that correlations are the equivalent of cross-correlations at lag zero.

Mean temperature	Positive to negative		Negative to positive	
Species	Month	Season	Month	Season
E. leucoxylon	August	Winter	February	Summer
E. tricarpa	April	Autumn	October	Spring
E. polyanthemos	December	Summer	July	Winter
E. camaldulensis	April	Autumn	October	Spring
E. melliodora	April	Autumn	October	Spring
E. microcarpa	September	Spring	March	Autumn
E. goniocalyx	October	Spring	April	Autumn
E. macrorhyncha	June	Autumn	December	Summer
Rainfall	Positive to negative		Negative to positive	
Species	Month	Season	Month	Season
E. leucoxylon	February	Summer	August	Winter
E. tricarpa	October	Spring	April	Autumn
E. polyanthemos	June	Winter	November	Summer
E. camaldulensis	October	Spring	April	Autumn
E. melliodora	October	Spring	April	Autumn
E. microcarpa	February	Summer	August	Winter
E. goniocalyx	August	Winter	February	Summer
E. macrorhyncha	December	Summer	May	Autumn

Table 4. Species specific change points (month, season) for mean temperature and rainfall.

3.1.3 Level 3 wavelet cross-correlations

Cross-correlations between peak flowering intensity and the temperature variants and rainfall were lagged from 1 and 12 months prior to the species specific peak flowering month. An illustration of these cross-correlation profiles between mean temperature and flowering, and between rainfall and flowering is given in Figure 10 and 11, respectively, for each species. These sinusoidal profiles show that there are positive and negative relationships of 6 months duration and as a consequence there are clear seasonal change points from positive to negative cross-correlation and vice-versa (see Table 4, where the season specific to each change point month is given per species).

The cross-correlation profiles also show there is a maximum positive and negative cross-correlation, and these occur systematically 6 months apart for each species (Figures 10 -11 and delta values (Δ) in Table 5). These cross-correlations can be viewed as the highest absolute value of the wavelet cross-correlations, for lags either in the short term (≤ 6 months to peak flowering intensity), or the long term (> 6 months). The number of months at which either lag, occurs prior to peak flowering intensity, differs for each species (Figures 10-11 and Table 5). Note that *E. polyanthemos's* maximum short term temperature lag is positive (0.73) and occurs in September, which is 2 months before peak flowering intensity, whereas *E. macrorhyncha's* short term lag is negative (-0.52) and occurs in August, this being 5 months prior to peak flowering intensity. The long term lag for *E. polyanthemos* is in March, which is 8 months before peak flowering intensity and negative (-0.71), compared to the February positive (0.52), long term lag for *E. macrorhyncha,* which occurs 11 months before its peak flowering intensity in July (Figure 10 and Table 5).

According to these wavelet cross-correlations the eight species form two groups based on their, similar cross-correlational relationships (or signatures) between flowering with temperature and between flowering and rainfall *E. leucoxylon, E. polyanthemos, E. camaldulensis and E. melliodora* are in one group and the remaining species in the other. This first group has cross -correlations with temperature which are positive in the short-term (≤ 6 months) and negative in the long-term (> 6 months) (Figure 10 and Table 5). With rainfall, the short-term cross-correlations are negative and positive in the long-term. The second group (comprising *E. goniocalyx*, *E. microcarpa, E. macrorhyncha* and *E. tricarpa*) has cross-correlations with temperature which are negative in the short-term (≤ 6 months) and positive in the long-term (> 6 months) – and with rainfall, the short-term cross-correlations are positive in the short-term and negative in the long-term (Figure 11 and Table 5).

It is noteworthy that, if one examines the whole 12 month pattern of the wavelet cross-correlation, so-called WCCORR *signatures* (sinusoids in Figures 10-11), in relation to flowering duration, flowering cessation and 12 months back (note the vertical line in Figures 10-11 delineate 12 months prior to peak flowering and the horizontal line depicts flowering duration) - rather than on focusing on primarily the signs of the WCCORR at months of maximum positive or negative cross-correlation (Table 5), there are three (3) clear species groupings, according to their similar WCCORR wavelet signatures, namely, *E. camaldulensis*, *E. melliodora, E. polyanthemos; E. goniocalyx*, *E. microcarpa, E. macrorhyncha;* and *E. leucoxylon* and *E. tricarpa*. Indeed the species within each of these groupings were recently shown to flower synchronously according to advanced methods for time series clustering of the patterns underlying the phenological series (Hudson et al., 2011a).

The cross-correlations in Figure 10 and Figure 11 also show, as did correlations, that for all species the temperature and rainfall variables do not act in concert. That is, within a given species, if the cross-correlation with temperature is positive, at a particular time of year, then

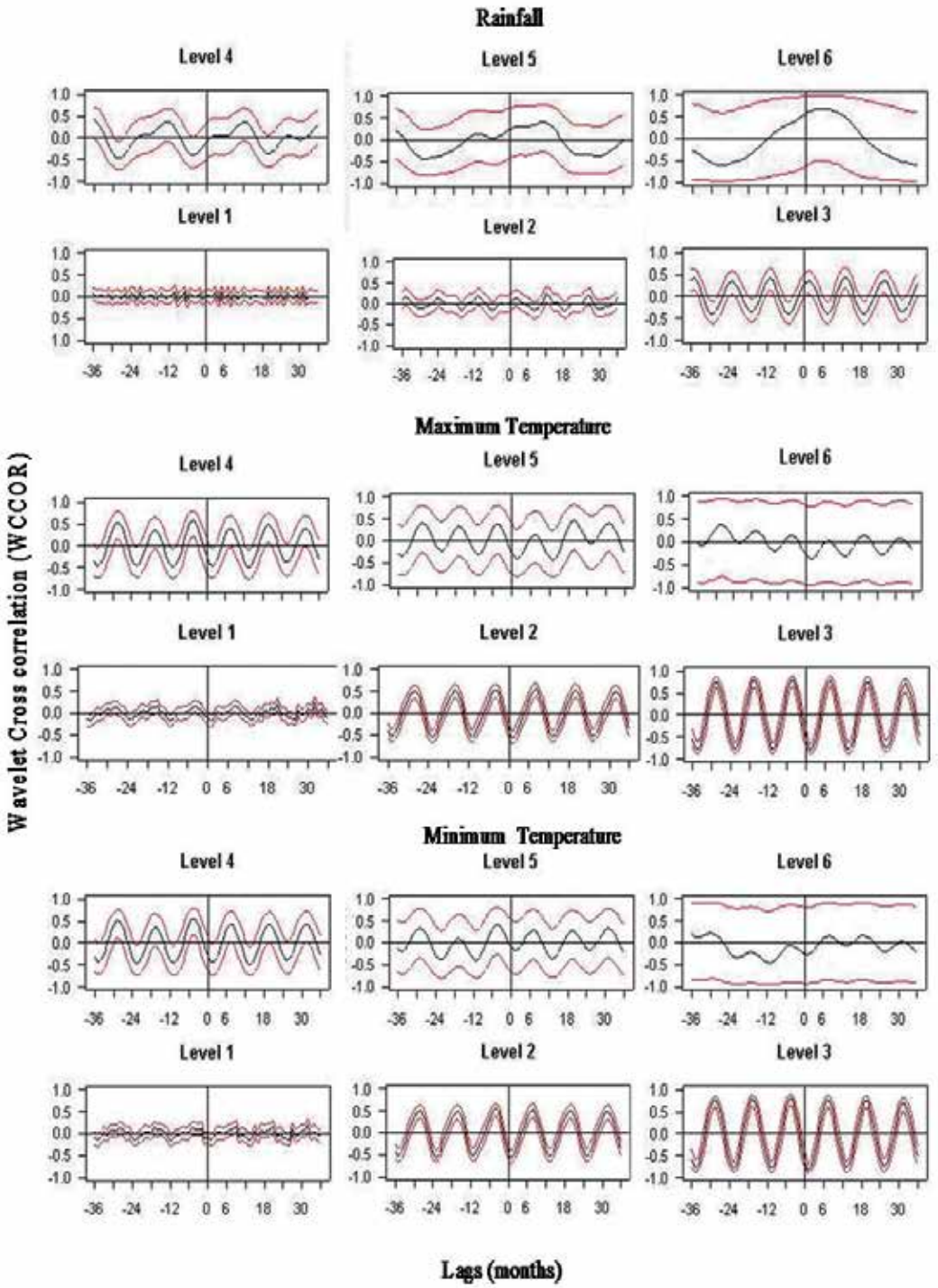


Fig. 9. Wavelet cross-correlation (WCCORR) plots of flowering with respect to rainfall, and maximum and minimum temperature for *E. leucoxylon*.

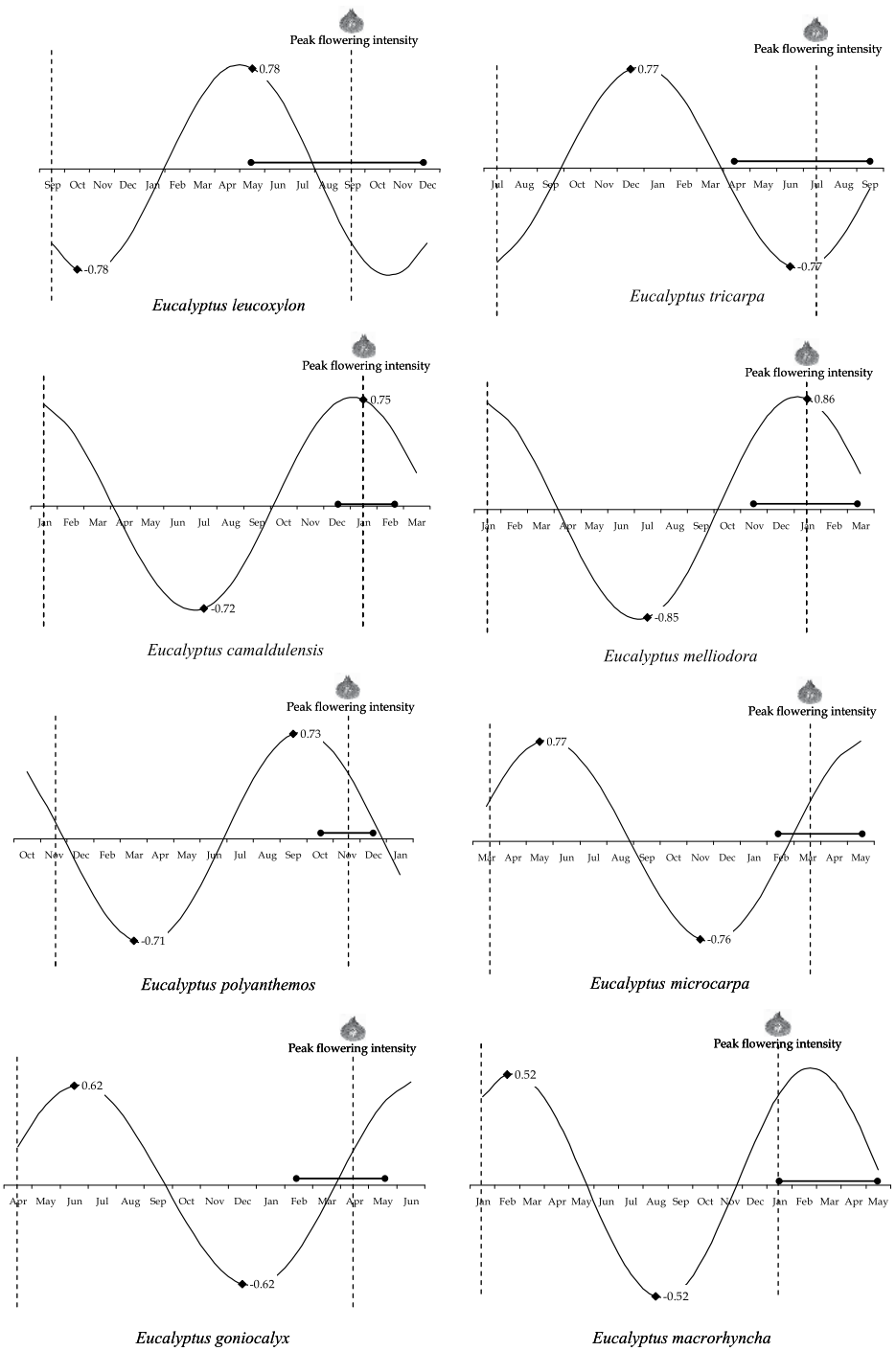


Fig. 10. Individual wavelet cross-correlation plots for daily mean temperature with flowering intensity for all species. Vertical lines encapsulate the 12 months prior to peak flowering. The small horizontal interval indicates the annual flowering period (duration).

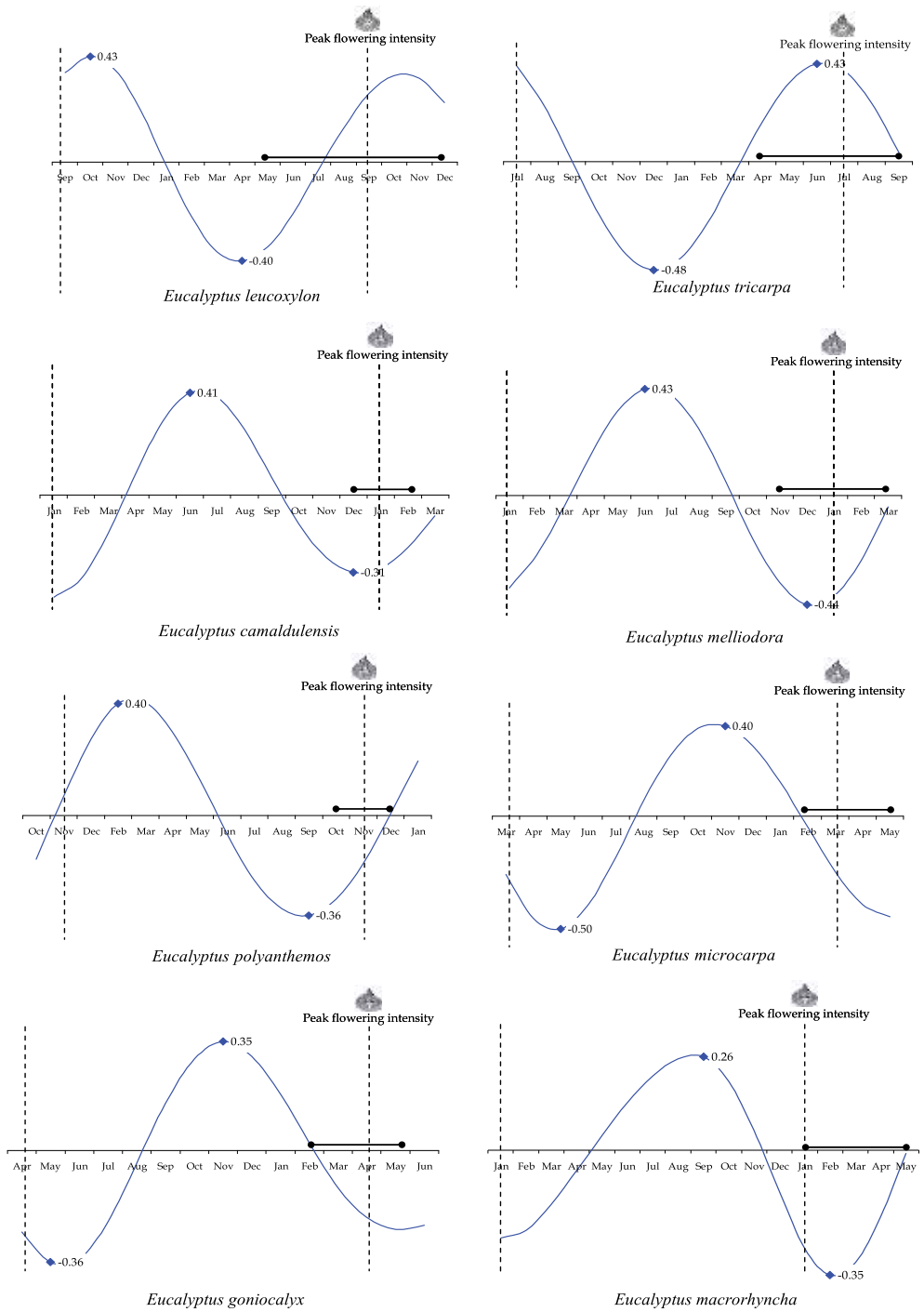


Fig. 11. Individual wavelet cross-correlation plots for daily rainfall with flowering intensity for all species. Vertical lines encapsulate the 12 months prior to peak flowering. The small horizontal interval indicates the annual flowering period (duration).

the relationship with rainfall is negative (Table 5) and vice versa (see opposite directionality between the wavelet WCCORR signatures in Figure 10 versus Figure 11). For example, each of *E. goniocalyx's* short term lags for temperature is negative and occurs in December. Its short term rainfall lag occurs in November, but is positive. Cross-correlations thus provide further insight into the dynamic, changing relationship between climate and peak flowering intensity for each eucalypt species.

Species	**Tmean**	**Tmax**	**Tmin**	**Rainfall**
E. leucoxylon	0.78 [-4]	0.79 [-4]	0.78 [-5]	-0.41 [-5]
Peak: September	May	May	April	April
Start: May	-0.80 [-11]	-0.80 [-10]	-0.81 [-11]	0.35 [-11]
End: December	October Δ = 7	Nov Δ = 6	October Δ = 6	October Δ = 6
E. tricarpa	-0.73 [-1]	-0.73 [-1]	-0.73 [-1]	0.46 [-1]
Peak: July	June	June	June	June
Start: April	0.75 [-7]	0.74 [-7]	0.75 [-7]	-0.45 [-7]
End: September	December Δ = 6	December Δ = 6	December Δ = 6	December Δ = 6
E. polyanthemos	0.73 [-2]	0.72 [-2]	0.72 [-2]	- 0.32 [-2]
Peak: November	September	September	September	September
Start: October	-0.71[-8]	- 0.72 [-8]	- 0.73[-8]	0.39 [-8]
End: December	March Δ = 6	March Δ = 6	March Δ = 6	March Δ = 6
E. camaldulensis	0.73 [0]	0.71 [-1]	0.75 [-1]	-0.31 [-1]
Peak: January	January	December	December	December
Start: Nov/Dec	-0.69 [-6]	0.69 [-7]	0.70 [-7]	0.41 [-7]
End: January	July Δ = 6	June Δ = 6	June Δ = 6	June Δ = 6
E. melliodora Peak:	0.86 [0]	0.82 [-1]	0.84 [-1]	-0.44 [-1]
January	January	January	December	December
Start: November	-0.80 [-6]	0.80 [-6]	-0.81 [-7]	0.43 [-7]
End: March	July Δ = 6	July Δ = 6	June Δ = 6	June Δ = 6
E. microcarpa	-0.77 [-4]	-0.77 [-4]	-0.76 [-4]	0.48 [-4]
Peak: March	November	November	November	November
Start: February	0.78 [-10]	0.78 [-10]	0.78 [-10]	-0.44 [-10]
End: May	May Δ = 6	May Δ = 6	May Δ = 6	May Δ = 6
E. gonicalyx	-0.62 [-4]	-0.61 [-4]	-0.63 [-4]	0.35 [-5]
Peak: March	December	December	December	November
Start: February	0.64 [-10]	0.62 [-10]	0.63 [-10]	-0.36 [-11]
End: May	June Δ = 6	June Δ = 6	June Δ = 6	May Δ = 6
E. macrorhyncha	-0.52 [-5]	-0.52	-0.52	0.26 [-5]
Peak: January	August	[-5] August	[-5] August	August
Start: January	0.52 [-11]	0.51 [-11]	0.52 [-11]	-0.22 [-11]
End: May	February Δ = 6	February Δ = 6	February Δ = 6	February Δ = 6

Table 5. Significant ($P < 0.0001$) wavelet cross correlations (4 months, λ_3 scale) between peak flowering and climate: shorter [≤ 6 months] and longer lags [> 6 months] prior to peak flowering intensity. (Δ signifies the difference in months between shorter and longer lags).

Even for species within a similar grouping, say *E. tricarpa* and *E. leucoxylon,* we observe that, although at zero lag *E. tricarpa* and *E. leucoxylon* have a similar correlational (WCORR)

relationship with climate (somewhat intuitive as their flowering overlaps, see Table 1 and Table 5), their cross-correlational relationship over the 12 months annual cycle differs, having opposite sinusoidal WCCORR profiles for peaks and troughs; with a positive and a negative short term lag for *E. leucoxylon* and *E. tricarpa,* respectively (Table 5) (see Figure 10 and Figure 11).

4. Discussion

4.1 Subcomponents

Wavelets Multiresolution analysis (MRA) has identified four subcomponents in each flowering series: these are a non-flowering phase, duration, annual cycle, flowering intensity, as well as the overall trend for each species. More specifically:

d1 is the non-flowering phase per species. These so-called "off phases" coincide with other reproductive phases (e.g. budding, seeding) in Eucalypts (Ashton 1975, Bassett 1995, Murray and Lutze 2004). This phenomenon has also been reported specifically in *E. tricarpa* (Keatley and Murray 2006).

The second sub-component, d2, reflects both the duration and the pattern of the original flowering record, for a given species; as shown also earlier for 4 of the 8 species studied in this chapter by Hudson et al. (2010a,b).

The annual cycle has been clearly delineated by subcomponent d3, with the mean month of peak intensity, so identified by peaks in d3 (e.g. *E. leucoxylon*: September, *E. tricarpa*: July, *E. melliodora*: January, *E. goniocalyx*: March). For five of the species (*E. leucoxylon*, *E. melliodora, E. microcarpa, E. tricarpa* and *E.polyanthemos*) this is in agreement with the trends identified by other analytic methods (e.g. the mean flowering intensity in a flowering year (Keatley & Hudson 2007) and/or by singular spectrum analysis (SSA) (Hudson et al. 2005; Hudson & Keatley, 2010b).

d4 relates to the annual or biennial cycling of the intensity of flowering years. In all eight species, except *E. leucoxylon,* the biennial cycle is related to years of low flowering intensity (< 2). Note that in these three species low flowering intensity also includes years of no flowering (i.e. flowering intensity = 0). In *E. leucoxylon* flowering generally alternates between a quasi-biennial (1950 to 1952, 1955 until 1963) and an annual cycle. It is the only species in this study, where the biennial cycle is associated solely with low (< 2) intensity, and/or late commencement with short duration (commencing in September or October compared to the most probable month of commencement in April (Keatley & Hudson, 2007).

Since Chambers (1893) and others, the variability in flowering intensity, in general, as well as its cyclic behavior has been long reported in eucalypts. Flowering intensity exhibiting two year cycles in these species has also been previously discussed, as based on the judgment and expertise of apiarists (Goodman 1973, Somerville and Campbell 1997, Paton et al. 2004); the collection of reproductive components (see Keatley and Murray 2006) and more recently confirmed by statistical autocorrelogram analysis (Wells 2000), and by singular spectrum analysis (SSA) (Hudson & Keatley, 2010b). A 4 year cycle, previously detected by SSA for *E. tricarpa* (Hudson & Keatley, 2010b), was however, not demonstrated by the wavelet analysis reported in this chapter. A four year cycle for this species has, on the other hand, been delineated earlier via autocorrelograms (Wells 2000). Wells's study and also the work of Somerville and Campbell (1997) have also identified a four year flowering cycle in *E. microcarpa.* Note that SSA is a global analysis (see Hudson et al. 2005; Hudson & Keatley,

2010b) in contrast to wavelets which afford a localized analysis. The reader is referred to Yiou et al. (2000) who discussed methods by which SSA can be made more wavelet-like specifically via the use of a windowing technique. This nonstandard version of SSA, however, was not applied in the analysis of the data of interest here, as reported in the SSA research applications of Hudson et al. (2005), and of Hudson & Keatley, 2010b).
s_4 is the smoothed series that illustrates the slowly varying portion of the flowering signal, and is taken to be an estimate of the trend. Again the results of wavelet analysis, namely a decreasing trend in flowering intensity over the period of study, agree with the SSA results (Hudson & Keatley, 2010b). Note that the so-called trend is usually the first primary reconstructed component in SSA (Elsner and Tsonis 1996) and mirrors the smoothed overall trend s_4 discussed in this chapter.

4.1.1 Temperature and rainfall wavelet correlations

Wavelet correlation analysis of the flowering and climate series found the same contemporaneous effects of temperature on the flowering of *E. tricarpa* and of *E. leucoxylon;* significantly increased flowering with decreased temperature; with the main drivers for each species being different, maximum temperature for *E. tricarpa* and minimum temperature for *E. leucoxylon* as the main species specific temperature influences on flowering. Maximum temperatures are often associated with day temperatures and minimum with night temperatures. Additionally, although both these species had a positive relationship with rainfall it was only significant in *E.tricarpa*, in opposition to *E. camaldulensis* and *E. melliodora* (which exhibited significantly increased flowering with increased minimum temperature and decreased rainfall). Hence for *E. leucoxylon,* it is cooler minimum temperatures, and in *E. tricarpa* cooler maximum temperatures which have a greater influence on peak flowering intensity. For *E. tricarpa* also wetter conditions tend to increase flowering intensity.
E. microcarpa, E. macrorhyncha and *E. polyanthemos* had a significant positive wavelet correlation with minimum temperature (at wavelet scale 4), with a trend (though not significant) for increased flowering with decreasing rain. *E. leucoxylon* and *E. tricarpa* were the only species of the eight species to exhibit a negative wavelet correlation with minimum temperature (at 2 and 4 months). *Eucalyptus tricarpa* was the only species for which rainfall had a significant, *positive* relationship with flowering (at the 4 month scale). This indicates that warmer, wetter periods lead to greater flowering intensity for *E. tricarpa*. The relationship for *E. leucoxylon* and rainfall was positive, but was not significant, as in *E. tricarpa. E. camaldulensis* and *E. melliodora,* which were the only species for which rainfall had a significant, *negative* relationship with flowering (at the 4 month scale); indicating that dryer periods lead to greater flowering intensity for these species. For *E. gonicalyx, E. microcarpa, E. macrorhyncha* and *E. polyanthemos* a negative correlation (but not significant) with rainfall was also indicated. Similar climatic drivers and relationships were found for 4 of the 8 Eucalypt species studied in this chapter by singular spectrum analysis (Hudson & Keatley (2010b)); by mixed transition distribution models (Kim et al. (2009); Hudson et al. (2010d)) and via Generalised Additive Models for Local, Scale and Shape (GAMLSS) modelling (Hudson et al., 2009, Hudson et al., 2010c;, Hudson et al., 2011b).
Eucalyptus microcarpa and *E. polyanthemos* were shown in this chapter to share a similar wavelet correlation relationship with temperature (positive) and rainfall (negative). Their flowering durations do not generally overlap (Keatley et al. 2004, Kim et al. 2008) and their

months of peak flowering intensity are separated by four months. However, mean diurnal temperature and mean monthly rainfall over their respective flowering duration is similar, 16.2ºC and 38.7 mm for *E. microcarpa* and 16.3ºC and 42.0 mm for *E. polyanthemos*. Rainfall did not have a significant influence in either of the latter species. In *E. microcarpa* flowering intensity is significantly correlated only with one of the three temperature variants, namely minimum temperature (at the 2 and 4 month scale), indicating that warmer minimum temperatures are associated with more intense flowering. Note that minimum temperature has also been previously nominated as the primary driver for *E. microcarpa* (Hudson et al. 2011b). Similar relationships between temperature and / or rainfall with flowering commencement (or flowering intensity), have also been shown by other studies of these species (Porter 1978, Keatley & Hudson 2000, Hudson et al. (2009, 2010a,b,c,d), and Kim et al., (2008, 2009)).

4.1.2 Temperature and rainfall wavelet cross- correlations

Four species (*E. leucoxylon, E. camaldulensis, E. melliodora and E. polyanthemos)* group together based on the 'sinusoidal' influence of the climate variables (evidenced by their wavelet WCCORR signatures), but separate themselves out by when that influence occurs - *E. leucoxylon* differentiates as such from the remaining 3 species (namely, *E. camaldulensis, E. melliodora and E. polyanthemos*).

Many of these species occur together. For example, *E. tricarpa* may be found with *E. leucoxylon*, *E. microcarpa*, *E. polyanthemos*, *E. goniocalyx* and *E. macrorhyncha*. Separation of flowering times is one mechanism to avoid competition for pollinators and reduce hybridization, being in a different sub-genus is another. For example, *E. tricarpa's* mean overlap with *E. polyanthemos* and *E. microcarpa* is low (0.13 and 0.36, respectively), however, in some years their flowering period can largely overlap (0.53 and 0.78, respectively) (Keatley and Hudson 2007). This higher overlap is probably due the similar temperature (influence) signatures (i.e. both species are negatively influenced by temperature at shorter lags and positively influenced in the longer term (prior to peak flowering), whereas *E. polyanthemos* has the opposite signature.

According to the wavelet cross-correlations, a similar relationship between flowering and temperature and between flowering and rainfall occurs in *E. leucoxylon* and *E. polyanthemos.* The reverse relationship between flowering and temperature and with rainfall was observed in *E. tricarpa* and *E. microcarpa.* In *E.camaldulensis,* and *E.melliodora* the rainfall 6 or 7 months prior to peak flowering intensity most strongly and positively influences subsequent flowering intensity, whilst the shorter term lags of *E. leucoxylon, E. polyanthemos*, *E. goniocalyx, E. microcarpa* and*E. macrorhyncha* were associated with negative cross-correlations with rainfall. Longer term lags of *E. leucoxylon, E. polyanthemos*, *E. goniocalyx, E. microcarpa* and *E. macrorhyncha* were associated with positive cross-correlations with rainfall.

Wavelet cross-correlation analysis identified the cyclical influence of temperature and rainfall on peak flowering intensity. For each species there are 6 months of the annual cycle in which any given climate variable positively influences flowering intensity and 6 months of negative influence. For all species, rainfall exerts a negative influence when temperature is positive. Wavelet cross-correlations between peak flowering intensity and the temperature variants and rainfall were lagged from 0 and 12 months prior to the species specific peak flowering month. These sinusoidal wavelet WCCORR signatures or profiles clearly demonstrated that there are positive and negative relationships of 6 months duration

and as a consequence there are seasonal change points from positive to negative cross-correlation and vice-versa.

Wavelets clearly demonstrate the cyclical influence of climate (temperature and rainfall) on peak flowering intensity for the eight species examined in this chapter. There are 6 months of an annual cycle in which any given climate variable positively influences flowering intensity and 6 months of negative influence, within a given species. This phenomenon results in distinctive seasonal change points. In very broad terms, for example, more intense peak flowering is likely to occur in *E. leucoxylon* when cool, wet conditions coincide with its peak flowering. However, peak flowering would be furthered enhanced if the preceding autumn and winter were warm and dry, and the previous spring and summer cool and wet. This supports the upper threshold mean temperature of 18.5ºC found for this species by Hudson et al. (2003) (see also Hudson et al., 2011b). *Eucalyptus tricarpa* requires the identical conditions at peak flowering intensity (i.e. cool and wet) but more or less the reverse conditions to *E. leucoxylon* in the months leading up to flowering: namely a cool, wet autumn and winter, with a warm, dry spring and summer. This supports the upper threshold temperature of 21.3ºC recently demonstrated for maximum temperature for this species by Hudson et al. (2011b). A warm, dry autumn and winter with a cooler, wet spring, and summer with warm, dry conditions at peak flowering, favours more intense flowering in *E. microcarpa*. Flowering is enhanced in *E. polyanthemos* with warm, dry conditions coinciding with peak flowering, preceded by a warm, dry winter and spring and cool, wet summer and autumn, in contrast to *E. microcarpa*. Correspondingly an upper threshold temperature of 16.1ºC was demonstrated for mean temperature for this species by Hudson et al. (2011b) via GAMLSS modelling.

There is evidence of this cycling in correlation between flowering and climate in previous research - specifically this 6 month cycling phenomenon can be observed in the reported tables and/or figures of the following studies; [1] in an examination of flowering commencement between 1954 - 1989 (by multiple linear regression) and the effect of mean monthly temperature by Fitter et al. (1995, Figure 4); [2] in an examination of flowering commencement, from 1978 to 2001, with respect to mean daily maximum temperature using P-splines by Roberts (2008, Figure 3) – there being an approximately 6 month period in which the sign of the smoothed regression coefficients of Roberts changed from negative to positive (see also Roberts 2010, 2011); [3] in Sparks and Carey (1995, see Table 2 of that study) there is evidence of this cycling in correlation between the flowering in wood anemone and turnip and monthly temperature in central England, for the months preceding mean observed date, over a 212 year period (1736 - 1947). Until now this phenomenon of 6 monthly cycling has not been commented on, apart from the recent studies of Hudson et al. (2010a,b) , nor formalised quantitatively as is achieved in this present chapter (via wavelets).

According to Hudson et al. (2010b) the relationships between phenophases and climatic variables, may be considered to be physiologically and/or statistically based (Yang et al. 1995, Spano et al. 1999). In this chapter the relationships between peak flowering intensity and the temperature and rainfall variables could partly be reflecting the time of year these species flower for each of the species examined. For example, *E. tricarpa's* peak flowering intensity occurs in winter, and this species has a negative relationship with temperature, indicating increased flowering with decreasing temperature. Previously, this scenario has been suggested for explaining, at least partially, the relationship between flowering commencement and climate in three (*E. leucoxylon, E. tricarpa* and *E. polyanthemos*) species

(Keatley and Hudson 2000). Additionally, it has been shown that when rainfall and temperature have been accounted for, there is no significant remaining trend in the flowering intensity of *E. leucoxylon* (Hudson et al. 2003), indicating that there is a physiological basis to this relationship. Previous work (Porter 1978, Keatley and Hudson 2000, Keatley et al., 2002, Hudson et al., 2003) examining the influence of temperature and rainfall on flowering intensity of these species have found similar results, but have used different analytical methods (singular spectrum analysis, generalized additive models (GAMs) and Bayesian hierarchical models (BHMs) and regression). As stated earlier comparable climatic drivers and relationships were found for 4 of the 8 Eucalypt species studied in this chapter by singular spectrum analysis (Hudson and Keatley, 2010b)); by mixed transition distribution models (Kim et al., 2009; Hudson et al., 2010d) and via Generalised Additive Models for Local, Scale and Shape (GAMLSS) modelling (Hudson et al., 2010c, Hudson et al., 2009, Hudson et al., 2011b).

In an earlier study House (1997) suggested that even though flowering time seems to be under the control of climatic conditions, it is during bud formation and development that there may be genetic control of the actual flowering dates in local (sub-) populations. Also noted was that the patchy distribution of time of flowering commencment and duration between groups of *E. tricarpa* and *E. regnans* was suggestive of local genetic control of these flowering traits (House, 1997). It was advocated that heavy and light flowering seasons appear to occur in cycles in many species (also evidenced in some species in this chapter) and are related to resource availability. House (1997) stated that there are few data available to demonstrate any systematic differences in reproductive traits that may help elucidate the co-existence of closely related species. Because hybridisation rarely occurs between different subgenera, flowering seasons should be more similar between subgenera than within them House (1997). In the research presented in this chapter all species are in the same subgenera except for *E. macrorhyncha.*

Wavelet cross-correlations add additional insight into and more detail on the changing relationship between climate and peak flowering intensity (over an annual cycle) for the eight eucalypt species (see also the study of Hudson et al., 2011c). They show that whilst the pairing of *E. leucoxylon* and *E. tricarpa,* and the pairing of *E. microcarpa* and *E. polyanthemos,* respectively, have similar relationships with the climate variables at zero lag; the relationship between peak flowering intensity and climate during the 12 months prior to each species' peak flowering is individual or species specific; and therefore is not just a mirroring of the time of year in which the species flower. Note that, four months prior to peak flowering intensity, *E. leucoxylon* is most strongly and positively influenced by temperature, whereas the influence on *E. tricarpa* is not significant. The influence of temperature on each of these species likewise switches from negative to positive and back again in different seasons pointing to a possible physiological basis for this interplay.

5. Conclusion

Broadly speaking, *E. leucoxylon* and *E. tricarpa* can be grouped together with respect to their relationship with temperature and rainfall. The relationship between flowering and climate of the remaining six species is similar, however, when the significance of the relationship is taken into account two other species clearly grouped together: *E. camaldulensis* and *E. melliodora*. As already discussed they are the only two species which have significant negative relationship with rainfall (similar WCORR signature between their flowering and

rainfall and a significant positive relationship with all three climate parameters at both scales. *Eucalyptus polyanthemos* aligns with this group as it has a significant positive relationship with all three climate parameters but only at the 4 month scale. Additionally, its main influence is maximum temperature as is the case with the other two species (albeit that *E. camaldulensis* is equally influenced by mean temperature). The other three species (namely, *E. goniocalyx*, *E. microcarpa* and *E. macrorhyncha*) also broadly group, as evidenced by the similar climate signature shown in their WCORR profiles. Note that that all members of this species triple, have a positive relationship between flowering and at least one temperature variant; and a negative (though not significant) relationship between flowering and rainfall.

According to wavelet cross-correlations the eight species form two groups based on their, similar cross-correlation relationship between flowering with temperature and between flowering and rainfall - *E. leucoxylon, E. polyanthemos, E. camaldulensis and E. melliodora* are in one group and the remaining species in the other. This first group has cross correlations with temperature which are positive in the short-term (≤ 6 months) and negative in the long-term (> 6 months). With rainfall, the short-term cross-correlations are negative in the short-term and positive in the long-term. The second group (comprising *E. goniocalyx*, *E. microcarpa, E. macrorhyncha* and *E. tricarpa*) has correlations with temperature which are negative in the short-term (≤ 6 months) and positive in the long-term (> 6 months) – and with rainfall the short-term cross-correlations are positive and negative in the long-term.

It is noteworthy that if one examines the whole 12 month pattern of the wavelet cross-correlation WCCORR signatures (sinusoids), in relation to flowering duration, cessation and 12 months back, rather than focusing on the signs of the WCCORR at months of maximum positive or negative cross-correlation, there are 3 groupings, namely, *E. camaldulensis*, *E. melliodora, E. polyanthemos; E. goniocalyx*, *E. microcarpa, E. macrorhyncha;* and *E. leucoxylon* and *E. tricarpa*. Indeed the members within each of these groupings were recently shown to flowering synchronously in a study using self organising map (SOM) methods for time series clustering (Hudson et al., 2011a). The resultant SOM correlations, which were based on clustering the *patterns* underlying the time series records in Hudson et al. (2011a), also showed similar trends to the correlations gleaned from Moran based correlations obtained from the raw flowering records (see Kim et al., 2008).

In agreement with the SOM clustering results, the wavelet based findings of this chapter (of the same eight species) showed via similarities in the wavelet correlation, wavelet cross-correlation signatures and in the identification of the main species specific climatic drivers of flowering: *Eucalyptus microcarpa* to be asynchronising with *E. leucoxylon* and *E. polyanthemos*. *Eucalyptus tricarpa* to be flowering asynchronously to the species triple (*E. camaldulensis, E. melliodora* and *E. polyanthemos). Eucalyptus leucoxylon* was demonstrated to be synchronous with both *E. polyanthemos* and *E. tricarpa,* and asynchronous to all remaining species. Hudson et al. (2011a) demonstrated that the main influence for grouping or clustering (synchronisation of flowering) was the season in which flowering commences, however, other flowering characteristics such as the timing of peak flowering and start and cessation of flowering were also contributing factors to synchronisation. In the wavelets based WCORR and WCCORR analyses in this chapter timing of peak intensity was also shown to be an important factor, as reflected by the common seasonal change points for synchronous species groupings (Table 4).

This study shows that the DWT and the MODWT, with multiresolution analysis (MRA) are ideally suitable for investigating the inter-relatedness between climate and phenological time series which may exhibit non-stationarity. As such this chapter adds further knowledge about the interplay between climate and the flowering of eight Eucalypt species - recognised as a major southern hemisphere genus. MODWT-MRA successfully identified the sub-components (annual cycle duration, non-flowering and trend) within each flowering series. Wavelets cross-correlation analysis: (i) determined the strength, directionality and lagged nature of the relationship between climate and flowering; (ii) identified the primary climatic drivers at peak flowering intensity, and (iii) confirmed the dynamic nature of the relationship between peak flowering and climate for each of the 8 species. Our results allude to a physiological basis for this interplay. Species specific wavelet correlation and cross-correlation signatures were established relating flowering intensity with climate; indicating climatic impacts and possibly signatures for synchronisation between specific species groups, in agreement to those recently identified by Hudson et al., 2010e; Hudson et al., 2011a). Wavelet methods are demonstrated to be valuable tools for the analysis of phenological time series, and in studies which aim to detect and understand local climatic impacts on phenological phases and possibly global climate change.

6. References

Ashton DH (1956) Studies on the autecology of *Eucalyptus regnans*. PhD thesis, The University of Melbourne, Parkville, 143 pp

Ashton DH (1975) Studies of flowering behaviour in *Eucalyptus regnans* F. Muell. Aust J Bot 23:399-411

Badeck F. W., Bondeau A., Böttcher K., Doktor D., Lucht W., Schaber J. & Sitch S. (2004) Responses of spring phenology to climate change. *New Phytologist* 162, 295-309

Bassett OD (1995) Development of seed crop in *Eucalyptus sieberi* L. Johnson and *E. globoidea* Blakely in a lowland sclerophyll forest of East Gippsland., pp. 38. Dept. of Conservation and Natural Resources, Victoria

Beggs P. J. (2004) Impacts of climate change on aeroallergens: past and future. *Clin. Exp. Allergy* 34, 1507-13

Bratteli O, Jorgensen P (2002) Wavelets through a looking glass. The world of the spectrum. Birkhäuser, Boston

Bullmore E, Fadili J, Breakspear M, Salvador R, Suckling J, Brammer M (2003) Wavelets and statistical analysis of functional magnetic resonance images of the human brain. Stat Methods Med Res 12:375-399

Chambers LT (1893) The Colonial Beekeeper. J.C. Stephens, Melbourne

Cornish CR, Bretherton CS, Percival DB (2006) Maximal overlap wavelet statistical analysis with application to atomspheric turbulence. Bound-lay Meteorol, 119:339-374

Daubechies I (1992) Ten lectures on wavelets. Society for Industrial and Applied Mathematics, Philadelphia

de Vries J (1980) Measuring the impact of climate on history: Searching for appropriate methodologies. J Interdiscipl Hist 10:599-630

Dose V, Menzel A (2004) Bayesian analysis of climate change impacts in phenology. Glob Change Biol 10:259-272

Dunlop M. & Brown P. R. (2008) Implications of climate change for Australia's National Reserve System: A preliminary assessment. Report to the Department of Climate Change, February 2008. Department of Climate Change,Canberra.

Elsner JB, Tsonis AA (1996) Singular spectrum analysis. A new tool in time series analysis. Plenum Press, New York

Fitter AH, Fitter RSR, Harris ITB, Williamson MH (1995) Relationships between first flowering date and temperature in the flora of a locality of central England. Func Ecol 9:55-60

Fitter A. H. & Fitter R. S. R. (2002) Rapid changes in flowering time in British plants. *Science* 296, 1689-91.Forrest J. & Miller-Rushing A. J. (2010) Toward a synthetic understanding of the role of phenology in ecology and evolution. *Proc. Roy. Soc. London B* 365, 3101–12.

Forrest AH& Miller-Rushing AJ (2010) Towards a synthetic understanding of the role of phenology in ecology and evolution. Proc. Roy. Soc. London B365, 3101-12

Gencay R, Selcuk F, Whitcher B (2001) An introduction to wavelets and other filtering methods in finance and economics. Academic press, San Diego

Ghil, M., M. R. Allen, M. D. Dettinger, K. Ide, D. Kondrashov, M. E. Mann, A. W. Robertson, A. Saunders, Y. Tian, F. Varadi, and P. Yiou (2002), Advanced spectral methods for climatic time series, Rev. Geophys., 40(1), 1003, doi:10.1029/2000RG000092

Goodman RD (1973) Honey Flora of Victoria. Department of Agriculture, Melbourne

Hernández E, Weiss G (1996) A first course on wavelets. CRC-Press, Boca Raton

House SM (1997) Reproductive biology of Eucalypts. In 'Eucalypt Ecology: Individuals to Ecosystems.' (Eds JE Williams and J Woinarski) pp. 30-55. (Cambridge University Press: Cambridge)

Hudson IL, Barnett A, Keatley MR, Ades PK (2003) Investigation into drivers for flowering in eucalypts: effects of climate on flowering. In: Verbeke G, Moelenberghs G, Aaerts M, Fieuws S (eds) 18th International Workshop on Statistical Modelling, Katholieke Universiteit Leuven, Belgium, pp 195-200

Hudson I.L., Keatley M.R. & Roberts A. M.I. (2005) Statistical Methods in Phenological Research. In: *Statistical Solutions to Modern Problems. Proceedings of the 20th International Workshop on Statistical Modelling* (eds A. R. Francis, K. M. Matawie, A. Oshlack and G. K. Smyth) 10-15 July, 2005, pp. 259-70, Sydney, Australia. ISBN 1 74108 101 7.

Hudson, I.L. Kim, S.W., & Keatley, M.R. (2009) Climatic influences on the flowering phenology of four Eucalypts: a GAMLSS approach. Proceedings of the 18th International Association for Mathematics and Computers in Simulation (IMACS) World Congress/MODSIM 2009, Cairns 13-17 July, 2009, ISBN: 978-0-9758400-7-8, pp 2611-2617

Hudson I (2010) Interdisciplinary approaches: towards new statistical methods for phenological studies. *Clim Change* 100(1):143-171.

Hudson, I.L., & Keatley, M. R. (2010a). *Phenological Research: Methods for Environmental and Climate Change Analysis*: Springer Dordrecht.

Hudson IL & Keatley MR (2010b) Singular Spectrum Analysis: Climatic Niche Identification. In: Hudson IL, Keatley MR (eds) *Phenological Research: Methods for Environmental and Climate Change Analysis*, Springer, Dordrecht, pp 393-424

Hudson IL, Kang I, & Keatley MR (2010a) Wavelet Analysis of Flowering and Climatic Niche Identification. In: Hudson IL, Keatley MR (eds) *Phenological Research: Methods for Environmental and Climate Change Analysis,* Springer, pp 361-391.

Hudson I. L., Keatley M. R. & Kang I. (2010b) Wavelet characterization of eucalypt flowering and the influence of climate. *Environmental and Ecological Statistics*. Published first online 27 June, 2010. DOI: 10.1007/s10651-010-0149-5

Hudson IL, Kim SW, & Keatley MR (2010c) Climatic Influences on the Flowering Phenology of Four Eucalypts: A GAMLSS Approach. In: Hudson IL, Keatley MR (eds) *Phenological Research: Methods for Environmental and Climate Change Analysis*, Springer, Dordrecht, pp 209-228.

Hudson IL, Kim SW, & Keatley MR (2010d) Modelling the Flowering of Four Eucalypt Species Using New Mixture Transition Distribution Models. In: Hudson IL, Keatley MR (eds) *Phenological Research: Methods for Environmental and Climate Change Analysis*, Springer, Dordrecht, pp 299-320.

Hudson, IL, Lee, SY, & Keatley, MR (2010e) Data visualization and mapping methods to assess synchrony. Phenology 2010, Trinity College, Dublin, UK. Climate change impacts and adaption. http://www.tcd.ie/Botany/phenology/2010/

Hudson, IL. Keatley, MR, Lee, SL (2011a) Using Self-Organising Maps (SOMs) to assess synchronies: an application to historical Eucalypt flowering records. *International Journal of Biometeorology* (DOI: 10.1007/s00484-011-0427-4) pp 00-27. Accepted 16 March, 2011

Hudson, IL, Kim, SW, Keatley, MR (2011b) Climate effects and thresholds for flowering of eight Eucalypts: a GAMLSS ZIP approach. Proceedings of International Congress on Modelling and Simulation (MODSIM). Sustaining our future: understanding and living with uncertainty, Dec., 12-16, 2011, Perth, Australia. (To appear)

Hudson IL, Keatley MR, Kang I (2011c) Wavelets and clustering: methods to assess synchronization. pp. 59-70. In: del Valle M, Muñoz R, Gutiérrez JM (eds) *Wavelets: Classification, Theory and Applications*. Nova Science Publishers (pp 00 -28) *(in press)*

IPCC (2001) Summary for policymakers. A report of Working Group 1 of the Inter-governmental Panel on Climate Change. IPCC, Shanghai

Kang, I., Hudson, I. L., & Keatley, M. R. (2004). Wavelet analysis in phenological research-the study of four flowering eucalypt species in relation to climate. Paper presented at the XXIInd *Intl Biometric Conference,* 11-16 July, Cairns, Australia.

Keatley MR, Hudson IL, Fletcher TD (1999) The use of long-term records for describing flowering behaviour: a case-study in Victorian Box-Ironbark Forests. In: Dargavel J, Wasser B (eds) Australia's Ever-changing Forests IV, Australian University Press, Canberra, pp 311-328

Keatley MR, Hudson IL (2000) Influences on the flowering phenology of three eucalypts. In: de Dear RJ, Kalma JD, Oke TR, Aucliems A (eds) Biometeorology and Urban Climatology at the Turn of the Century. Selected Papers from the Conference ICB-ICUC' 99 World Meteorological Organisation, Geneva, Switzerland, pp 191-196.

Keatley M.R, Fletcher T.F, Hudson I.L and Ades P.K. (2002). Phenological studies in Australia: Potential application in historical and future climate analysis. *International Journal of Climatology,* Vol 22 (14): 1769 - 1780.

Keatley MR, Hudson IL, Fletcher TD (2004) Long-term flowering synchrony of Box-Ironbark eucalypts. Aust J Bot 52:47-54

Keatley MR, Murray MD (2006) An examination of the reproductive phenology of *Eucalyptus tricarpa*. Forest Science Centre, Orbost, Victoria

Keatley MR, & Hudson IL (2007) A comparison of the long-term flowering patterns of Box-Ironbark species in Havelock and Rushworth forests. Environmental Modeling and Assessment, 12:279 - 292

Keatley M. R. & Hudson I. L. (2010) Introduction and overview. In: *Phenological Research: Methods for environmental and climate change analysis* (eds I. L. Hudson and M. R. Keatley) pp. 1-22. Springer, Dordrecht

Kim SW, Hudson IL, Agrawal MA, Keatley MR (2008) Modelling and synchronization of four Eucalypt species via MTD and EKF. *23rd International Workshop in Statistical Modelling,* Utrecht, Netherlands, 7-11 July: Ipskamp Partners, Enschede, pp. 287-292

Kim, S.W., Hudson I.L., & Keatley, M.R. (2009) Modelling flowering of four eucalypts species using MTDg analysis. Proc of the *18th Intl Assoc for Math and Computers in Simulation World Congress,* Cairns 13-17 Jul, ISBN: 978-09758400-78, pp 2625-2631

Koch M, Marković D (2007) Evidences for climate change in Germany over the 20th Century from the stochastic analysis of hydro-meteorological time-series. In: Oxley L, Kulasiri D (eds) *MODSIM 2007 International Congress on Modelling and Simulation.* Christchurch, New Zealand, pp 596-602

Lu X, Liu R, Liu J, Liang S (2007) Removal of noise by wavelet method to generate high quality temporal data of terrestrial MODIS Products. Photogramm Eng Rem S 73:1129-1140

Mallat S (1989) A theory for multiresolution signal decomposition: the wavelet representation. *IEEE T Pattern Anal*ysis 11:674-693

Menzel A., et al. (2006) European phenological response to climate change matches the warming pattern. *Global Change Biol.* 12, 1969 - 76

Murray M, Lutze M (2004) Seedcrop development in *Eucalyptus obliqua* and *Eucalyptus cypellocarpa* in High elevation Mixed Species forests of East Gippsland. Forest Science Centre, Orbost

Paluš M, Novotná D, Tichavský P (2005) Shifts of seasons at the European mid-latitudes: Natural Fluctuations correlated with the North Atlantic Oscillation. Geophys Res Lett 32 L12805, DOI:12810.11029/12005GL022838

Parmesan C (2006) Ecological and evolutionary responses to recent climate change. Ann Rev Ecol Syst 37:637–669

Paton DC, Crossfield EL, Hurrell B, Rogers DJ (2004) Floral resources used by the South Australian Apiary Industry. Rural Industries Research and Development Corporation, Barton, ACT

Percival D, Walden A (2000) Wavelet methods for time series analysis. Cambridge University Press, Cambridge

Percival DB, Guttorp P (1994) Long-memory processes, the Allan variance and wavelets. In: Foufoula-Georgiou E, Kumar P (eds) Wavelets in geophysics. Academic Press, New York, pp 325-344

Percival DB (1995) On estimation of the wavelet variance. Biometrika 82:619-31

Percival DB, Mofjeld O (1997) Analysis of subtidal coastal sea level fluctuations using wavelets. J Am Stat Assoc 92:868-880

Percival DB, Sardy S, Davison AC (2000) Wavestrapping time series: Adaptive wavelet-based bootstrapping. In: Fitzgerald WJ, Smith RL, Walden AT, Young PC (eds)

Nonlinear and nonstationary signal processing. Cambridge University Press, Cambridge, pp 442-471

Percival DB, Wang M, Overland JE (2004) An introduction to wavelet analysis with applications to vegetation monitoring. Community Ecol 5:19-30

Porter JW (1978) Relationships between flowering and honey production of Red Ironbark, *Eucalyptus sideroxylon* (A. Cunn.) Benth., and climate in the Bendigo district of Victoria. Aust J Agr Res 29:815-829

Roberts AMI (2008) Exploring relationships between phenological and weather data using smoothing. *Int J Biometeorol* 52:463-470

Roberts AMI (2010). Smoothing methods. In: *Phenological Research: methods for environmental and climate change analysis*. Eds. Keatley MR, Hudson IL. Springer Verlag., pp 255-270

Roberts AMI (2011) Comparison of regression methods for phenology. *International Journal of Biometeorology,* DOI: 10.1007/s00484-011-0472-z

Root TL, Price JT, Hall KR, Schneider SH, Rosenzweig C, Pounds JA (2003) Fingerprints of global warming on wild animals and plants. *Nature* 421:57-60

Rosenzweig C. et al. (2008) Attributing physical and biological impacts to anthropogenic climate change. *Nature* 453, 353-8

Schwartz MD (2003) Preface. In: Schwartz MD (ed) *Phenology: an integrative environmental science,* Kluwer Academic Publishers, Dordrecht, pp xvii-xix

Şendur, L., Suckling, J., Whitcher, B., & Bullmore, E (2007) Resampling methods for improved wavelet-based multiple hypothesis testing of parametric maps in functional MRI . *NeuroImage*, 37, Issue 4, 1 October 2007, Pages 118-1194

Serroukh A, Walden AT (2000) Wavelet scale analysis of bivariate time series I: Statistical Properties for Linear Processes. *J Nonparametr Stat* 13:1–36

Somerville D, Campbell S (1997) Beekeeping in the Narrandera State Forests. N.S.W Agriculture, Goulburn, N.S.W. Australia

Spano D, Cesaraccio C, Duce P, Synder RL (1999) Phenological stages of natural species and their use as climate indicators. *Int J Biometeorol* 42:124-133

Sparks TH, Carey PD (1995) The responses of species to climate over two centuries: an analysis of the Marshman phenological record, 1736-1947. J Ecol 83:321-329

Sparks TH, Menzel A (2002) Observed changes in seasons: an overview. *Int J Climatol* 22:1715-1725

Sparks T. H., Croxton P. J., Collinson N. & Taylor P. W. (2005) Examples of phenological change, past and present, in UK farming. *Ann. Appl. Biol.* 146, 531-7.

Sparks TH, Tryjanonwski P (2005) The detection of climate change impacts: some methodological considerations. *Int J Climatol* 25:271-277

Trenberth, Kevin E.; Hoar, Timothy J. (1996). "The 1990-1995 El Niño-Southern Oscillation event: Longest on record". *Geophysical Research Letters* 23 (1): 57-60

Vidakovic, B. (1999). *Statistical modelling by wavelets*. New York: John Wiley & Sons.

Wells K (2000) Long term cyclic and environmentally induced effects on flowering of four box-ironbark eucalypts. Fourth Year Project, University of Melbourne, Parkville

Whitcher BJ, Guttorp P, Percival DB (2000) *Wavelet Analysis of Covariance with Application to Atmospheric Time Series.* J Geophys Res 105:941–962

Yiou P, Sornette D, Ghil M (2000) Data-adaptive wavelets and multi-scale singular spectrum analysis. *Physica* D, 142:254-290

15

Multiple Moving Objects Detection and Tracking Using Discrete Wavelet Transform

Chih-Hsien Hsia, Jen-Shiun Chiang[1] and Jing-Ming Guo
Department of Electrical Engineering
National Taiwan University of Science and Technology, Taipei
[1]Department of Electrical Engineering
Tamkang University, Taipei
Taiwan

1. Introduction

In recent years, video surveillance systems for the purpose of security have been developed rapidly. More and more researches try to develop intelligent video surveillance systems to replace the traditional passive video surveillance systems (Hu et al., 2004) and (Jacobs & Pless, 2008). The intelligent video surveillance system can detect moving objects in the initial stage and subsequently process the functions such as object classification, object tracking, and object behaviors description. Detecting moving object is a very important aspect of computer vision and has a very wide range of surveillance applications. The accurate location of the moving object does not only provide a focus of attention for post-processing but also can reduce the redundant computation for the incorrect motion of the moving object. The successful moving object detection in a real surrounding environment is a difficult task, since there are many kinds of problems such as illumination changes, fake motion (Cheng & Chen, 2006), night detection (Huang, 2008), and Gaussian noise in the background (Gonzalez & Woods, 2001) that may lead to detect incorrect motion of the moving object. There are three typical approaches for motion detection (Hu et al., 2004), (Jacobs & Pless, 2008), and (Collins, 2000): background subtraction, temporal differencing, and optical flow. The background subtraction method detects moving regions between the current frame and the reference background frame. It provides the most complete motion mask data, but is susceptible to dynamic scene changes due to lighting and extraneous events. Therefore, it has to update the reference background frame frequently. The temporal differencing approach extracts the moving region by using consecutive frames of the image sequences. It is suitable for dynamic environment, but often extracts incomplete relevant motion object pixels. The optical flow method uses characteristics of flow vectors of moving objects over time to detect moving regions. However, most optical flow methods are with higher complex computation. Generally, the above three moving object detection methods are all sensitive to illumination changes, noises, and fake motion such as moving leaves of trees.

In order to solve the mentioned problems, several approaches for object detecting and tracking were proposed (Ahmed et al., 2005), (Alsaqre & Baozong, 2004), (Cheng & Chen, 2006), (Chen & Yang, 2007), (Collins, 2000), (Cvetkovic et al., 2006), (Hsieh & Hsu, 2007), (Hu et al., 2004), (Hu et al., 2009), (Huang et al., 2008), (Jacobs & Pless, 2008), (Liu et al., 2006), (Mckenna, 2000), (Sugandi, 2007), and (Tab, 2007). Video tracking systems have to deal with variously shaped and sized input objects, which often result in a massive computing cost of the input of images. Cheng *et al*. (Cheng & Chen, 2006) used discrete wavelet transform (DWT) to detect and track moving objects. The 2-D DWT can be used to decompose an image into four-subband images (LL, LH, HL, and HH). It only processes the part of LL-band image due to the consideration of low computing cost and noise reduction issues. Although this method provides low computing cost (low resolution) for post-processing and noise reduction based on the conventional DWT, the LL-band image produced by the original image size via two dimensions (row and column) calculation may cause high computing cost in the pre-processing. Especially they use the three-level low-low band image (LL_3) that does not only bring a great image size transfer computation, but also the slow motion of the real moving objects may disappear. After dealing with the background subtraction, Alsaqre *et al*. (Alsaqre & Baozong, 2004) used a local pre-process method to smooth the image with reducing noise and other small fluctuations. However, this approach is unable to reduce the post-processing computation. Sugandi *et al*. (Sugandi et al., 2007) proposed a method for detecting and tracking objects by using a low resolution image with the 2×2 average filter (2×2 AF), which is generated by replacing each pixel value of the original image with the average value of its neighbors and itself. They mentioned that the low resolution image is insensitive to illumination changes and can reduce the small movement like moving leaves of trees in the background. Although this method can deal with small movement, these low resolution images become more blurred than the LL-band image generated by using DWT.

To overcome the above-mentioned problems, we propose a method, direct LL-mask band scheme (DLLBS), for detecting and tracking moving objects by using SMDWT (Hsia et al., 2009). In DLLBS, we can select only the LL-mask band of SMDWT. Unlike the conventional DWT method to process row and column dimensions separately by low-pass filter and down-sampling, the LL-mask band of SMDWT can be used to directly calculate the LL-band image. Our proposed method can reduce the image transfer computing cost and remove fake motion that is not belonged to the real moving object. For objects occlusion, a new approach, characteristic point recognition (CPR), was proposed. Combined with DLLBS and CPR, it can have accurate object tracking for various types of occlusions. Furthermore, it can retain a better slow motion of objects than that of the low resolution method (Sugandi et al., 2007) and provide effective and complete moving object regions.

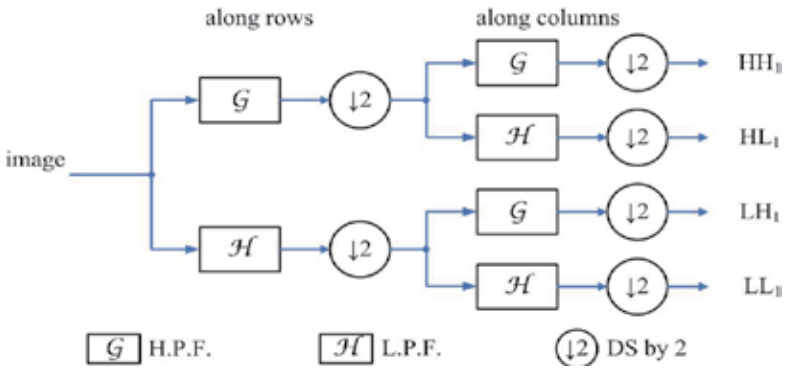
2. Discrete Wavelet Transform and low resolution technique

Due to the imperfection of video acquisition systems and transmission channels, images are often corrupted by noise. Therefor, this degradation leads to a significant reduction of image quality, especially for the task that performs high-level computer vision, such as object tracking, recognition, etc. Before dealing with motion object detection, there are several methods for removing noises or fake motion and reducing computing cost proposed in the past several years. DWT (Cheng & Chen, 2006) and low resolution technique (Andra et al., 2000) are two important approaches, and are briefly described in the following sub-sections.

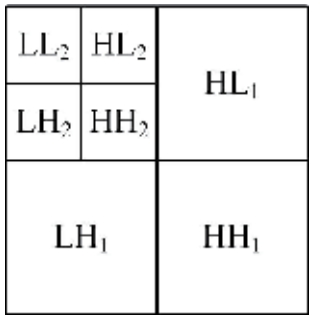
2.1 Discrete Wavelet Transform method

Wavelet transform (Mallat, 1989) was proposed in the mid-1980s, and it has been used in various fields such as signal processing, image processing, computer vision, image compression, biochemistry medicine, etc. For image processing, it provides an extremely flexible multi-resolution image and can decompose an original image into different subband images including low- and high-frequencies. Therefore people can choose the specific resolution data or subband images upon their own demands (Hsia et al., 2009), (Mallat, 1989) , (Ge et al., 2007), (Liu et al., 2006), (Ahmed et al., 2005), and (Tab et al., 2007).

A 2-D DWT of an image is illustrated in Fig. 1(a). When the original image is decomposed into four-subband images, it has to deal with row and column directions separately. First, the high-pass filter *G* and the low-pass filter *H* are exploited for each row data, and then are down-sampled by 2 to get high- and low-frequency components of the row. Next, the high- and the low-pass filters are applied again for each high- and low-frequency components of the column, and then are down-sampled by 2. By way of the above processing, the four-subband images are generated: HH, HL, LH, and LL. Each subband image has its own feature, such as the low-frequency information is preserved in the LL-band and the high-frequency information is almost preserved in the HH-, HL-, and LH-bands. The LL-subband image can be further decomposed in the same way for the second level subband image. By using 2-D DWT, an image can be decomposed into any level subband images, as shown in Fig. 1.



(a)



(b)

Fig. 1. Diagrams of DWT image decomposition: (a) the 1-L 2-D analysis DWT image decomposition process, (b) the 2-L 2-D analysis DWT subband.

Cheng *et al.* (Cheng & Chen, 2006) applied the 2-D DWT for detecting and tracking moving objects and only the LL_3-band image is used for detecting the moving object motion. Because noises are preserved in high frequency, it can reduce the computing cost for post-processing by using the LL_3-band image. This method can be used for coping with noise or fake motion effectively, however the conventional DWT scheme has the disadvantages of complicated calculation when an original image is decomposed into the LL-band image. Moreover if it uses an LL_3-band image to deal with the fake motion, it may cause incomplete moving object detecting regions.

2.2 Low resolution method

Sugandi *et al*. (Sugandi, 2007) proposed a simple method by using the low resolution concept to deal with the fake motion such as moving leaves of trees. The low resolution image is generated by replacing each pixel value of an original image with the average value of its four neighbor pixels and itself as shown in Fig. 2. It also provides a flexible multi-resolution image like the DWT. Nevertheless, the low resolution images generated by using the 2×2 average filter method are more blurred than that by using the DWT method, as shown in Fig. 3. The average filtering is a low pass filter which denoises the image and performs restoration by the noise reduction spatial domain. It may reduce the preciseness of post-processing operation (such as occlusion and object identification), because the post-processing depends on the correct location of the moving object detecting and accuracy moving object data.

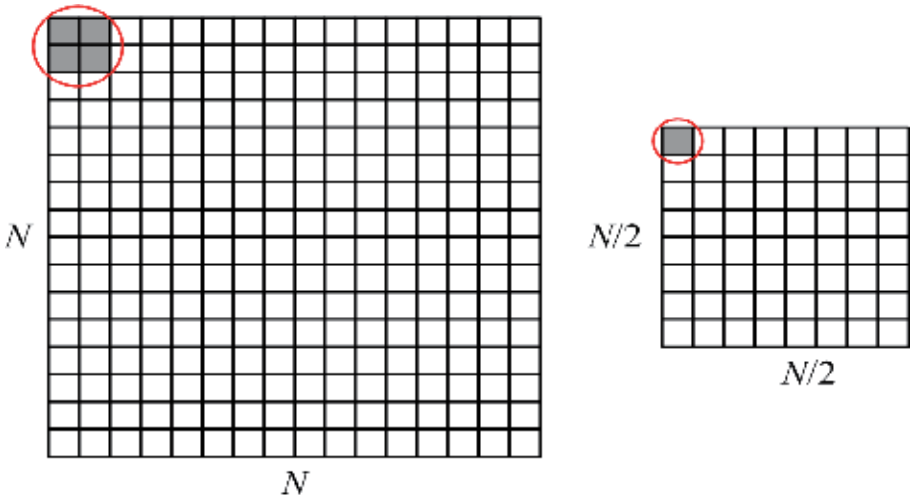


Fig. 2. Diagram of the 2×2 average filter method.

3. Direct LL-mask band scheme

In order to detect and track the moving object more accurately, we propose a new method called direct LL-mask band scheme (DLLBS) that is based on the 2-D integer symmetric mask-based discrete wavelet transform (SMDWT) (Hsia et al., 2009). It does not only retain the features of the flexibilities for multi-resolution, but also does not cause high computing cost when using it for finding different subband images. In addition, it preserves more

image quality of the low resolution image than that of the low resolution method (Sugandi, 2007).

(a)

(b)

(c)

Fig. 3. Comparisons of low resolution images: (a) the original image (320×240), (b) each subband image with DWT from left to right as 160×120, 80×60, and 40×30, respectively, (c) each resolution image with the 2×2 average filter method from left to right as 160×120, 80×60, and 40×30, respectively.

3.1 Symmetric Mask-based Discrete Wavelet Transform (SMDWT)

In 2-D DWT, the computation needs a large transpose memory requirement and has a long critical path. The SMDWT has many advanced features such as short critical path, high speed operation, regular signal coding, and independent subband processing (Hsia et al., 2009). The derivation coefficient of the 2-D SMDWT is based on the 2-D 5/3 integer LDWT. For computation speed and simplicity considerations, four-masks, 3×3, 5×3, 3×5, and 5×5, are used to perform spatial filtering tasks. Moreover, the four-subband processing can be further optimized to speed up and reduce the temporal memory of the DWT coefficients.

The four-matrix processors consist of four mask filters, and each filter is derived from one 2-D DWT of 5/3 integer lifting-based coefficients (Hsia et al., 2009). The coefficients of each subband mask are shown in Fig. 4, and the block diagram of the 2-D SMDWT is shown in Fig. 5.

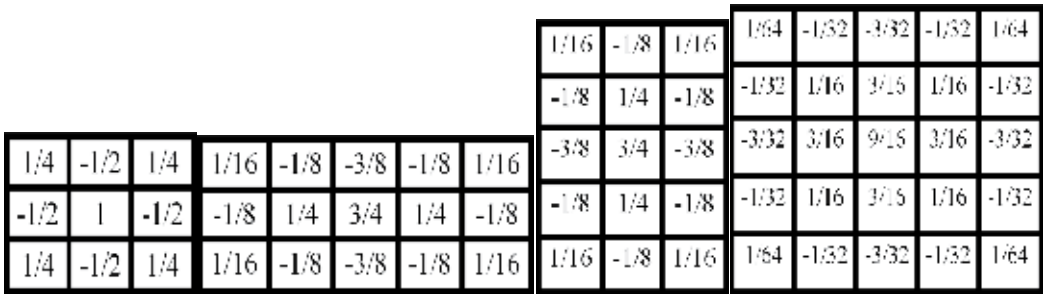


Fig. 4. The subband mask coefficients of (a) HH, (b) HL, (c) LH, and (d) LL.

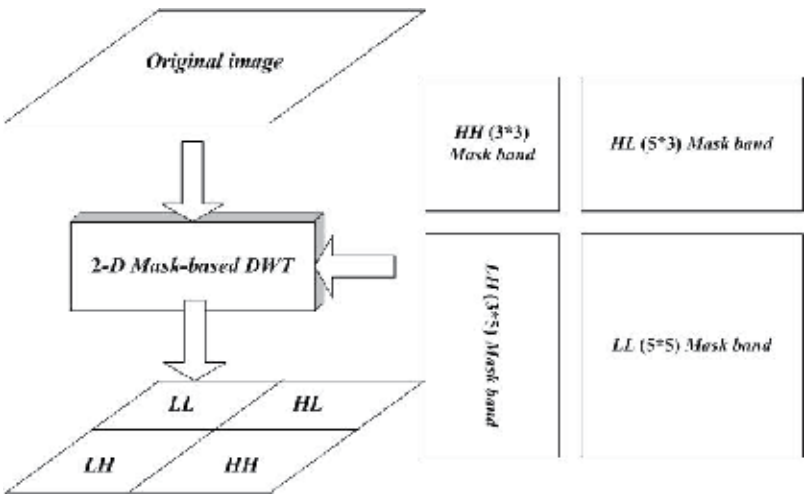


Fig. 5. The system block diagram of 2-D SMDWT.

3.2 Detection and tracking flow

The pre-processing flowchart of the proposed DLLBS moving object detection and tracking system is shown in Fig. 6. Frist, prior to color converting RGB data to YCbCr data (using Y data only). Basically we apply the double-change-detection method (Huang et al., 2004) to detect the moving objects. In order to decrease the holes left inside the moving entities, three continuous frames (F_{t-1}, F_t, and F_{t+1}) are used in this system for detecting moving object mask. These three continuous frames are decomposed into LL_2-band frames (LL_{2t-1}, LL_{2t}, and LL_{2t+1}) by using SMDWT. After most of the noises and fake motions are moved into the high-frequency subband as shown in Fig. 7, it can proceed with the post-processing by employing these three LL_2-band frames. Binary masks, B_{t-1} and B_t can be obtained by

computing the binary values of these three successive LL_2-band frames (in between LL_{2t-1}, LL_{2t}, and LL_{2t+1}) and a threshold value T in (1).

$$B_{t-1}(i,j) = \begin{cases} 1, & \text{if} \left| LL_{2t-1}(i,j) - LL_{2t}(i,j) \right| > T \\ 0, & \text{otherwise} \end{cases},$$
$$B_t(i,j) = \begin{cases} 1, & \text{if} \left| LL_{2t}(i,j) - LL_{2t+1}(i,j) \right| > T \\ 0, & \text{otherwise} \end{cases}. \quad (1)$$

The motion mask (MM_t) can be generated by using the union operation (logical OR) of B_{t-1} and B_t. The function is represented as follows:

$$MM_t = B_t \bigcup B_{t-1}. \quad (2)$$

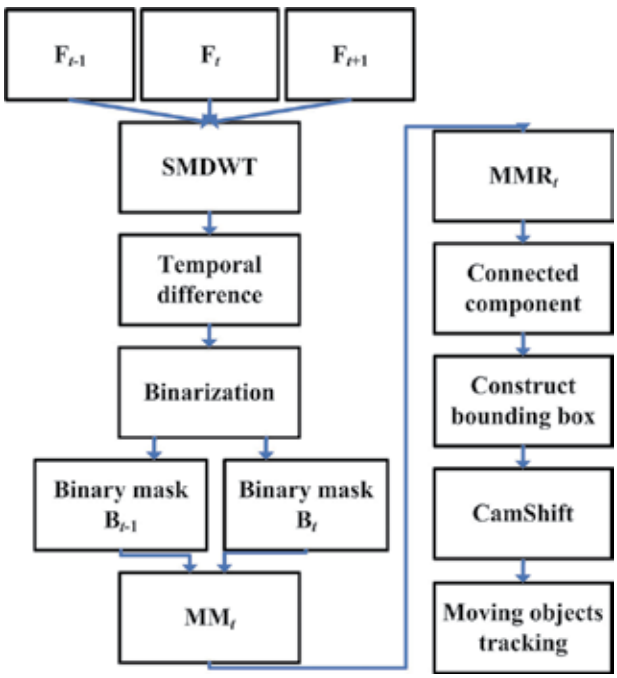


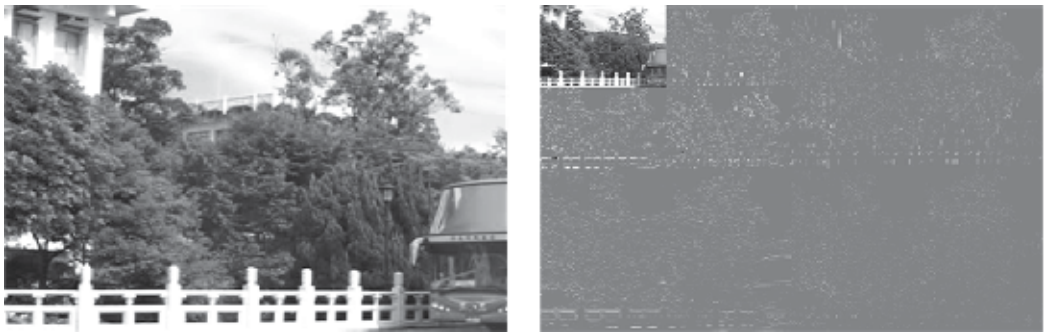
Fig. 6. The pre-processing flowchart of the moving object detection and tracking based on DLLBS.

The holes may still exist in the motion masks, because some motion pixels are too tiny such that it causes error judgments as non-motion ones. In order to increase the motion mask (MM_t) robustly, the morphological *closing* method (Hsieh & Hsu, 2007) is used to fill these holes. First, we apply the dilation operator for filling the middle of the isolated pixels that become related in the motion masks. It is defined as follows:

$$F_t(i,j) = \begin{cases} 1, & \text{if one or more pixels of the adjacent pixels of motion mask } MM_t(i,j) \text{ are 1,} \\ 0, & \text{otherwise.} \end{cases} \quad (3)$$

Then we apply the erosion operator for eliminating redundant pixels in the motion mask boundary as follows:

$$MMR_t(i,j) = \begin{cases} 0, & \text{if one or more pixels of the adjacent pixels of motion mask } F_t(i,j) \text{ are } 0, \\ 1, & \text{otherwise.} \end{cases} \quad (4)$$



(a) (b)

Fig. 7. After most of the noises and fake motions are removed using SMDWT (a) The original image, (b) LL_2-band image.

It scans eight neighbors of the motion mask MMR_t image pixel by pixel from top left to bottom right (raster scan). After extracting the connected component, it obtains several moving objects. In this work, we utilize the region-based tracking algorithm (Cheng &. Chen, 2006), (Mckenna, 2000), and (Chen & Yang, 2007) to track the moving object motion.
Labeling is useful when the moving objects in the scene are more than one (The connected component labeling is then employed to label each moving object and track each moving object individually). The labeling of the components based on pixel connectivity (intensity) (Gonzalez & Woods, 2001) is obtained by scanning an image and groups, pixel by pixel from top left to bottom right, in order to identify the connected pixel regions by comparing the eight neighbors that have already been encountered in the scan. If the pixel has at least one neighbor with the same label, we label this pixel as the neighbor. The labeled moving objects are thus found, and then we extract the boundary of the moving object using rectangle box to track the moving object. For this reason, the bounding box is found according to its motion mask from the foregoing work. The bounding box is made by finding the minimum and maximum values of row and column coordinates of the motion mask. In order to track moving objects in the original image size, we have to transform the coordinate from the LL_2 image size back to the original image size according to the spatial relationship of the DWT as follows:

$$O(i,j) = LL_n(i \times 2^n, j \times 2^n). \quad (5)$$

where $n = 0 \sim l$ and l is the number of level.
In the block-matching motion estimation, the motion vector is the displacement of a block with the minimum distortion from the reference block. The CamShift block-matching algorithm determines the motion vector by identifying a block with the minimum distortion

from fast search strategies of the diamond-arc-hexagon search patterns in the search area (Chiang et al., 2008).

3.3 Occlusion handling for multiple objects tracking

In the post-processing, occlusion handling is a major problem in a video surveillance system. The most popular color space is the RGB color space (Hu et al., 2009). If the multiple objects bounding boxes are occluded, the object bounding boxes are merged into the occlusion bounding box. Here we propose a new approach for occlusion in multiple objects tracking, called characteristic point recognition (CPR). Fig. 8 shows the operation flowchart of CPR. CPR uses bounding boxes during pre-processing of DLLBS. For each tracked individual, the system will detect whether it makes occlusion with other object or group. It can obtain the RGB information from the video capture device directly to calculate the color information of the moving pixels. Owing to the information of moving pixels the size of the inter-frame difference image (1/16 of the original image) is with the central pixels.

To recognize every object, it uses the bounding box to find the characteristic point (CP). CP represents the central point of the bounding box as shown in the following equation:

$$Cs^{q}[n] = B_{n}\{(x_1, y_1),(x_2, y_2),\ldots,(x_q, y_q)\}\ , \tag{6}$$

where $Cs^{q}[n]$ is an array to store the CP of every object, n the label of the object, q the amount of CP, B_n the bounding box of every object, and (x,y) the color information indexed by the position of CP. Therefore CP expresses the feature of the object. We would like to focus on each object bounding box in order to select one CP or more.

At first, the CP of every object is stored in the buffer when the first frame is input, and is regarded as the initial sample. In latter frames, the CP is matched with the sample. In other words, the CP of 1 to n matches with the CP of the sample as shown in the following equation:

$$Cd^{q}[n] = abs\{((Cs^{q}[n]- Cm^{q}[N])_{R},(Cs^{q}[n]- Cm^{q}[N])_{G},(Cs^{q}[n]- Cm^{q}[N])_{B})_{N}\}\ , \tag{7}$$

where $Cm^{q}[N]$ is a sample array to store the CP, N the label of the sample, and $Cd^{q}[n]$ the absolute values obtained from the difference between $Cs^{q}[n]$ and $Cm^{q}[N]$.

After the match step, $Cd^{q}[n]$ stores the sample N which is identical to the object n as shown in (8):

$$L[n] = N\ , \tag{8}$$

where $L[n]$ is the label N to label object n. Therefore the object is recognized and labeled as sample N.

However, the objects of a frame may disappear or be occluded in latter frames. In order to hold the information of the object, the CP of the object has to be retained. Hence, we must know the object which has ever occurred when the object appears again in some frames. Because the CP may be changed by the environmental factors, the buffer has to be updated whenever a new frame is input in order to obtain the latest CP. If a new object appears, the CP of the new object should be added into the buffer to update the CP information. The CPR flowchart is shown in Fig. 8.

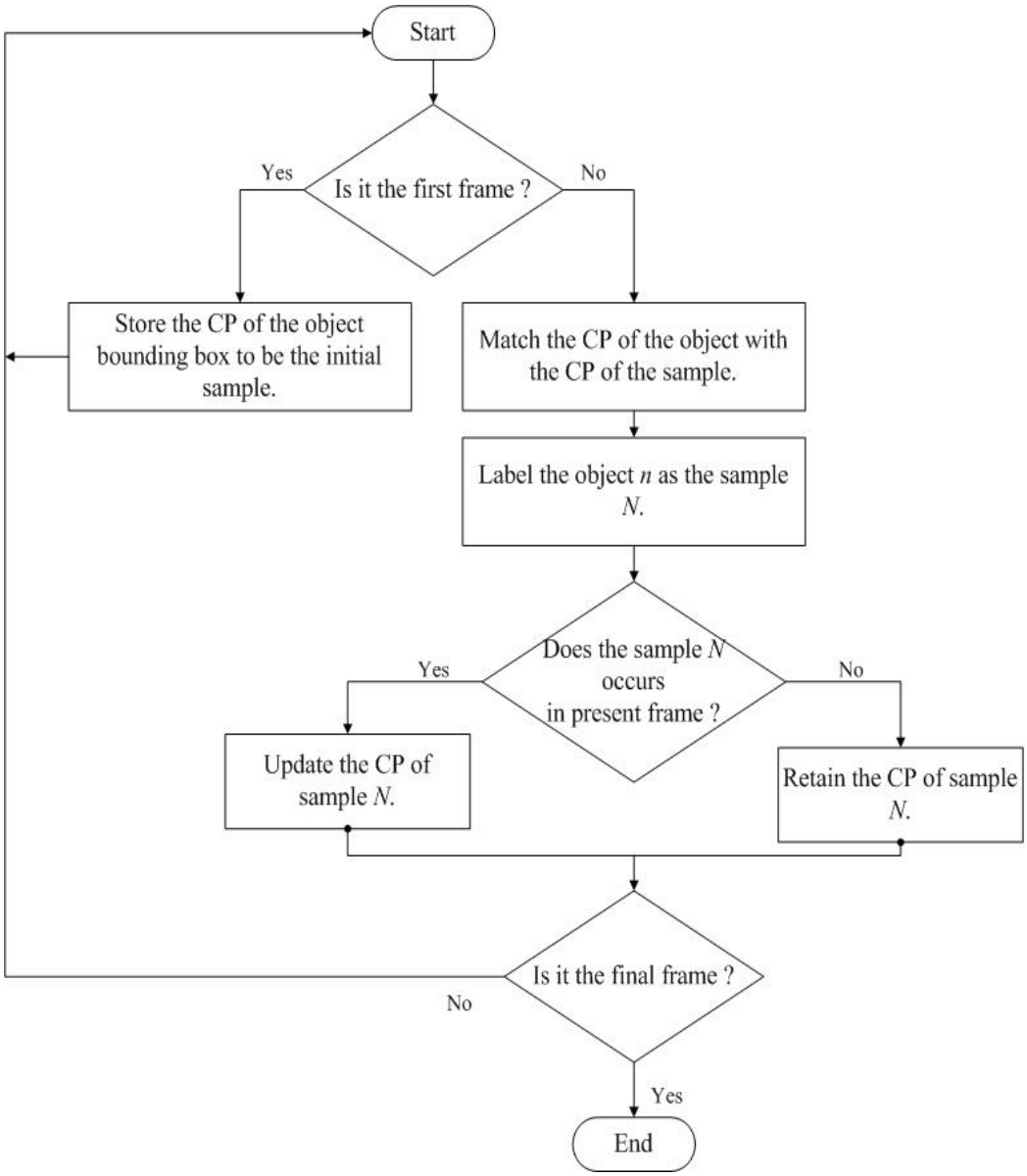


Fig. 8. CPR flowchart.

4. Experimental results

In this work, the experimental results of several different environments including indoor (all day) and outdoor (all day) environments with statistic video system are demonstrated. The original image frame sizes are 320×240, 640×480, and the format of color image frame is 24-bit in a RGB system. We use all gray level frames from transferring the RGB system to YCbCr system for detecting moving object motion and utilize the LL_2 (for 320×240) and the LL_3 (for 640×480) image size of 80×60 generated by using SMDWT from the original image for our proposed moving object detection and tracking system. The experimental environment is set using Intel 2.83 GHz Core 2 Quad CPU, 2 GB RAM, Microsoft Windows XP SP3, and Borland C++ Builder (BCB) 6.0. BCB is chosen as the software development platform. The software includes verifying for algorithms and image process for the moving objects detection.

4.1 Dealing with noise issues

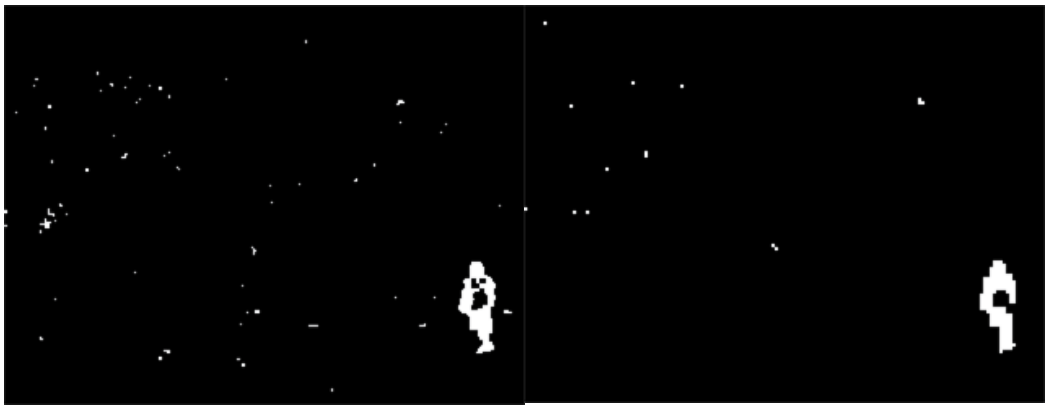
There are many kinds of difficulties such as illumination changes, fake motion, and Gaussian noise in the background. Different LL-band images including one-level, two-level, three-level, and multi-level LL-band images are used to deal with noises and compare their results. We suggest that a successful eliminating noise image has no other motion mask besides moving object motion masks, as shown in Figs. 9 and 10. Table 1 shows the average (Figs. 9 and 10) successful eliminating noise rate of each level LL-band image. The first row is in the indoor environment and the second row in the outdoor environment. Each level LL-band image has effective results when dealing with indoor noises like Gaussian noise produced by random noise and statistical noise. However, when dealing with the outdoor noise such as moving leaves of trees, the LL_1-band image has poor results because these outdoor noises sometimes are large that cannot be eliminated completely.

Resolution	[1]DLLBS	[2]LS	[3]2×2 AFS	[4]DSS
Level	Accuracy rate	Accuracy rate	Accuracy rate	Accuracy rate
LL_1 (160×120)	99.54 %	99.54 %	99.07 %	98.15 %
LL_2 (80×60)	99.07 %	99.07 %	93.07 %	81.94 %
LL_3 (40×30)	95.83 %	95.83 %	86.11 %	63.89 %

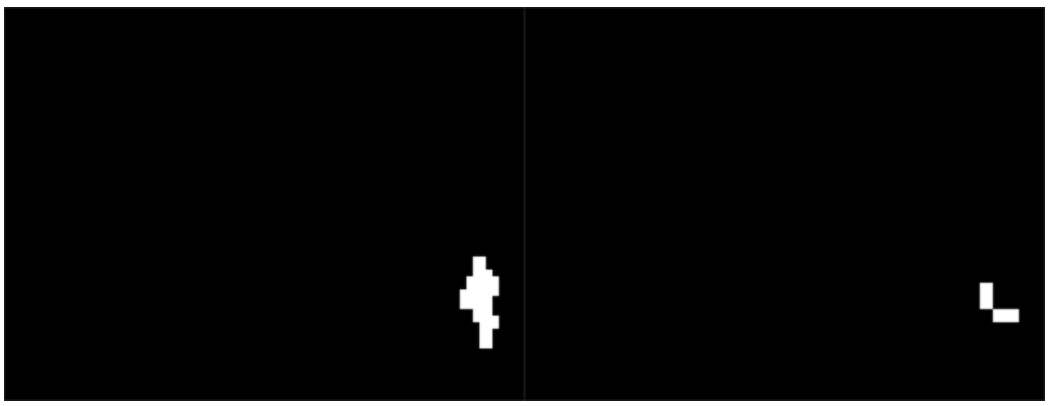
[1]DLLBS: Direct LL-mask Band Scheme; [2]LS: Lifting Scheme; [3]2×2 AFS: 2×2 Average Filter Scheme; [4]DSS: Down-Sampled Scheme; [5]Accuracy rate: Success Tracking/ Original Sequency.

Table 1. The moving objects detection and tracking results.

(a)



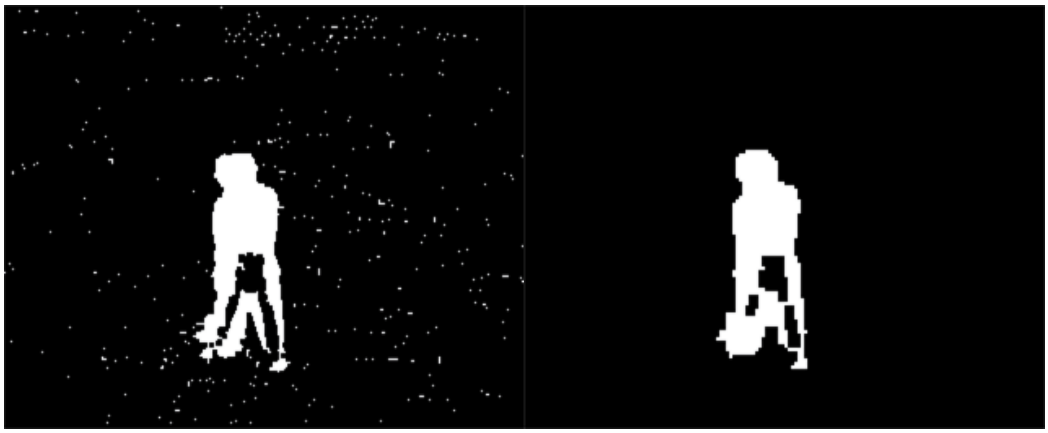
(b) (c)



(d) (e)

Fig. 9. Moving object detection in the outdoor environment with fake motion: (a) the original image of three consecutive frames, (b) the temporal differencing results of the original image, (c) the temporal differencing results of the LL_1-band image, (d) the temporal differencing results of the LL_2-band image, (e) the temporal differencing results of the LL_3-band image.

(a)



(b) (c)

(d) (e)

Fig. 10. Moving object detection in the indoor environment with Gaussian noise (Gonzalez & Woods, 2001): (a) the original image of three consecutive frames, (b) the temporal differencing results of the original image, (c) the temporal differencing results of the LL_1-band image, (d) the temporal differencing results of the LL_2-band image, (e) the temporal differencing results of the LL_3-band image.

4.2 Moving object tracking

We consider it to have a complete moving object region if it is a successful work, as shown in Fig. 11(a). In Fig. 11(b), the moving object regions have only a part of moving object, and that will be treated as a failure tracking. Fig. 12(a) expresses the original frame without detecting and tracking moving objects. Without DLLBS technique many noise masks are tracked. However, even if the moving objects are tracked, those moving regions are fragmented, as shown in Fig. 12(b). By using DLLBS, the noises can be filtered out, as shown in Fig. 12(c). It still generates incomplete moving object regions by using LL_1-band image, because the relevance of these pixels in the LL_1-band image is deleted. When using a three-level resolution image to detect the moving objects, it generates incompletely moving object regions, owing to the LL_3-band image causing too many slow motions belonged to the moving object disappeared, as shown in Fig. 12(e). Finally, let us look at the results of the LL_2-band image in Fig. 12(d). Using the two-level band image has a better tracking region and also can cope with noises and fake motion effectively, as shown in Table 1.

(a)

(b)

Fig. 11. Examples of (a) successful moving object tracking and (b) failure moving object tracking.

We use the 2×2 average filter scheme (AFS) in substituting the original DLLBS block system to demonstrate the moving object, however it is more blurred than the DLLBS technique. The accuracy rate of the successful object tracking with the 2×2 AFS are shown in Tables 1, 4, and 5. It is easy to perceive the contrasts between Tables 1 and 4 of any resolution image; the LL-band image generated by the DLLBS has a better successful ratio than the low resolution image generated by the 2×2 AFS.

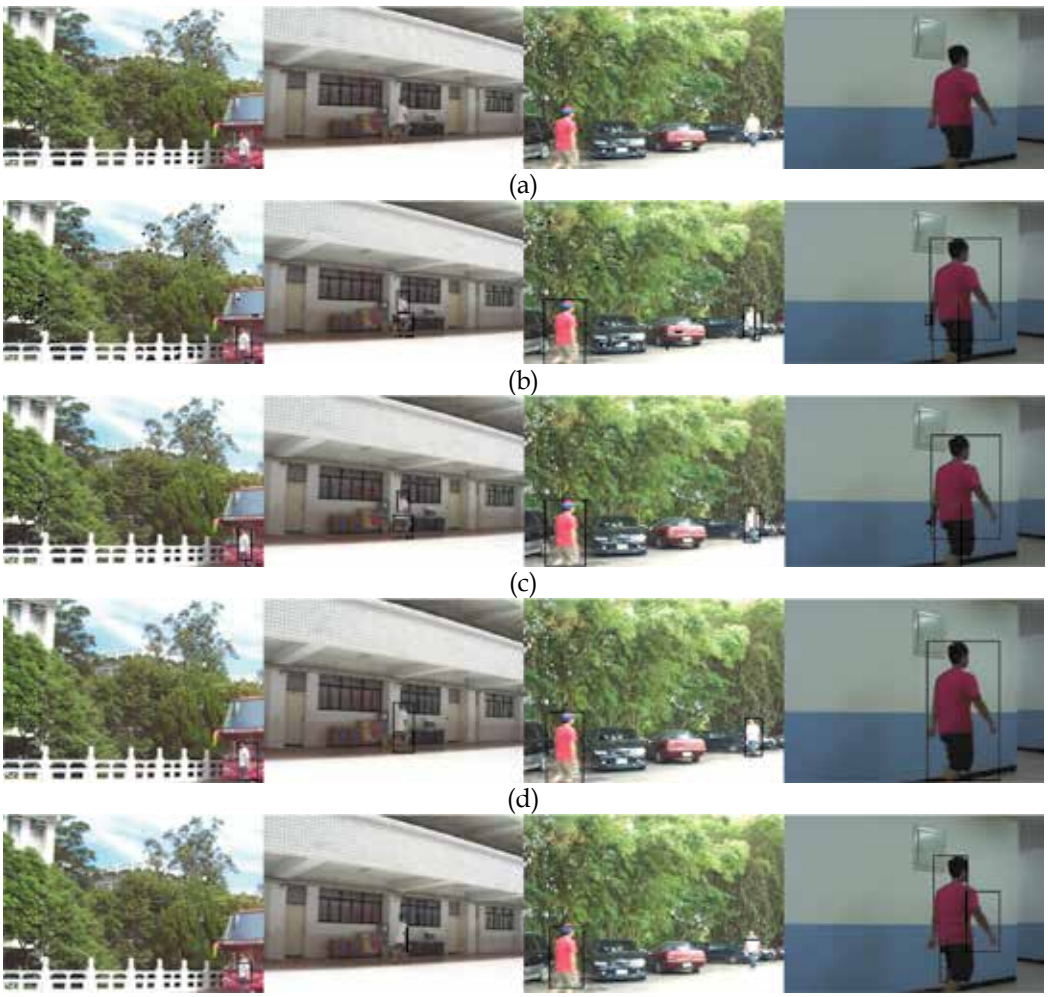


Fig. 12. Results of tracking moving objects in various environments: (a) original frames without region-based object tracking, (b) original frames with region-based object tracking, (c) LL_1-band frames with region-based object tracking, (d) LL_2-band frames with region-based object tracking, (e) LL_3-band frames with region-based object tracking.

Several experiments have been made to prove the feasibility of the proposed approach for moving object detection, tracking, and occlusion. We used an entry-level video camera and

capture card to capture the test sequences in our campus (Tamkang University), and simulated several cases of condition for moving objects, such as signal object in day time (indoor/outdoor), signal object at night (outdoor), and multiple objects in day time (outdoor) environments. All the test sequences are stored as the Microsoft AVI format with raw file of resolution 320×240, 640×480, and frame rate of 30 fps as shown in Fig. 13.
The choice of the threshold T is important. A too large value may lead to real targets missing; on the other hand a too small value may lead to noise binary images with pseudo features. The best threshold value also varies with the instantaneous illumination level. The threshold values of the best performance in different environments and DLLBS are listed in Table 2. According to the experiments, under day and midday time in outdoor for LL_2, the threshold values of 10, 14, and 16 were applied, some noisy pixels appeared. However, when the threshold value of 15 was applied, it outperformed all other threshold values.

Resolution	Night in the outdoor	Day and Night in the indoor	Day and Midday in the outdoor
LL_1(160×120)	T = 14	T = 20	T = 25
LL_2 (80×60)	T = 4	T = 10	T = 15
LL_3 (40×30)	T = 1	T = 6	T = 11

Table 2. The best threshold values, T, in different environments and DLLBS.

Environments	Fake motions	Low contrast	Reflection
DLLBS	Excellent	Excellent	Excellent
2×2 AFS	Excellent	Poor	Good
DSS	Poor	Excellent	Poor

Table 3. Features of various methods.

We established 16 test sequences at Tamkang University in different environments, such as day time, night time, rainy day, fast movement, slow movement, and occlusion, as shown in Fig. 13. Compared with other approaches (2×2 AFS and DSS), the DLLBS can obtain a good sparsity for spatially localized details, such as edges and singularities, as shown in Table 3. Because such details are typically abundant in natural images and convey a significant part of the information embedded therein, DWT has found a significant application for image denoising. From Tables 4, 5, and 6, we notice that some objects are not correctly identified in the test frame of the sequences. The wrong identification occurs in two reasons:

1. The moving object just enters or leaves the scene.

Because the moving object is detected and tracked at the border of the scene, the extracted features of the moving object in the case cannot represent the moving object very well.

2. The moving object is slowing down.

In this issue, the temporal difference image of the object becomes smaller and loses its situation.

Pattern	DLLBS		2×2 AFS		DSS	
	Accuracy rate	Detection + Tracking	Accuracy rate	Detection + Tracking	Accuracy rate	Detection + Tracking
Sequence1	98.61 %	53.8 FPS	98.61 %	58.5 FPS	71.76 %	52.1 FPS
Sequence2	95.90 %	56.5 FPS	96.31 %	57.1 FPS	56.15 %	49.0 FPS
Sequence3	97.40 %	54.1 FPS	92.36 %	60.7 FPS	96.59 %	63.1 FPS
Sequence4	93.55 %	54.3 FPS	82.61 %	62.0 FPS	92.65 %	63.2 FPS
Sequence5	82.97 %	56.7 FPS	80.44 %	60.2 FPS	77.29 %	65.6 FPS
Sequence6	82.57 %	53.9 FPS	78.90 %	61.5 FPS	46.79 %	63.9 FPS
Sequence7	90.28 %	54.7 FPS	37.50 %	61.4 FPS	40.28 %	63.5 FPS
Sequence8	83.33 %	55.1 FPS	73.46 %	62.4 FPS	78.40 %	61.2 FPS
Sequence9	90.16 %	53.5 FPS	75.13 %	59.0 FPS	83.94 %	58.5 FPS
Average	90.53 %	54.7 FPS	79.48 %	60.3 FPS	71.53 %	60.1 FPS

Table 4. Single moving object processing (without occlusion).

Pattern	DLLBS		2×2 AFS		DSS	
	Accuracy rate	Detection + Tracking + occlusion	Accuracy rate	Detection + Tracking + occlusion	Accuracy rate	Detection + Tracking + occlusion
Sequence10	92.94 %	56.7 FPS	88.61 %	60.6 FPS	82.92 %	57.6 FPS
Sequence11	89.98 %	54.8 FPS	83.67 %	60.9 FPS	92.60 %	58.7 FPS
Sequence12	90.43 %	53.9 FPS	79.79 %	60.8 FPS	82.98 %	55.9 FPS
Sequence13	90.00 %	52.6 FPS	86.67 %	59.7 FPS	75.56 %	63.7 FPS
Average	90.84 %	54.5 FPS	84.69 %	60.5 FPS	83.52 %	58.9 FPS

Table 5. Multiple moving objects processing (with occlusion).

Pattern	DLLBS		2×2 AFS		DSS	
	Accuracy rate	Detection + Tracking + occlusion	Accuracy rate	Detection + Tracking + occlusion	Accuracy rate	Detection + Tracking + occlusion
Sequence14	85.60 %	14.1 FPS	73.60 %	16.5 FPS	31.20 %	15.5 FPS
Sequence15	88.36 %	14.1 FPS	79.45 %	16.6 FPS	63.01 %	15.3 FPS
Sequence16	81.33 %	14.2 FPS	70.67 %	16.6 FPS	33.33 %	15.2 FPS
Average	85.10 %	14.1 FPS	74.57 %	16.6 FPS	42.51 %	15.3 FPS

Table 6. Multiple moving objects processing (with occlusion).

(a)

(b)

(c)

(d)

(e)

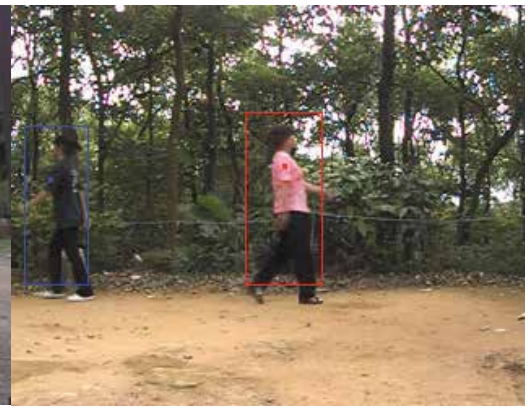
(f)

(g) (h)

(i) (j)

(k) (l)

(m)



(n)

(o)

(p)

Fig. 13. Test sequences at Tamkang University: (a)-(k) are sequences 1-13 (320×240) and 14-16 (640×480); (a)-(c) show the single moving object in the outdoor; (d) single moving object in the indoor; (e) single moving object in the outdoor (fast movement to slow movement); (f) single moving object in the outdoor (fast movement); (g) single moving object in the outdoor (slow movement); (h) single moving object in the indoor (zoom-out to zoom-in); (i) single moving object in the outdoor (rainy day); (j) multiple moving object in the outdoor (occlusion); (k) multiple moving object in the outdoor; (l) multiple moving object in the outdoor (occlusion); (m) multiple moving object in the outdoor (occlusion); (n) multiple moving object in the outdoor (occlusion); (o) multiple moving object in the outdoor (occlusion); (p) multiple moving object in the outdoor (occlusion).

5. Conclusions

The direct LL-mask band scheme (DLLBS) for moving object detection and tracking is proposed in this work. It is able to detect and track moving objects in indoor and outdoor environments with statistic video systems. The proposed DLLBS does not only overcome the drawbacks of high complex computation and slow speed for the conventional DWT, but also preserves the wavelet features of the flexible multi-resolution image and the capability for dealing with noises and fake motion such as moving leaves of trees. In the real-word application, the experimental results demonstrate that the 2-D LL_2-band (for 320×240) and the 2-D LL_3-band (for 640×480) can effectively track moving objects by region-based tracking under any environments (day and night), as well as it can cope with noise issues. For occlusion considerations, we propose a new approach, characteristic point recognition (CPR). Combined with DLLBS and CPR, it can accurately track various types of occlusions. The DLLBS system can be extended to the real-time video surveillance system applications, such as object classification and descriptive behaviors of objects.

6. Reference

Ahmed, J.; Jafri, M. N. & Ahmad, J. (2005). Target tracking in an image sequence using wavelet features and neural network, *IEEE TENCON*, (November 2005) pp.1-6.

Andra, K.; Chakrabarti, C. & Acharya, T. (2000). A VLSI architecture for lifting-based wavelet transform, *IEEE Workshop on Signal Processing Systems*, (October 2000) pp. 70-79.

Alsaqre, F. E. & Baozong, Y. (2004). Multiple moving objects tracking for video surveillance system, *IEEE International Conference on Signal Processing*, Vol. 2, (August 2004) pp.1301-1305.

Chen, D.-T. & Yang, J. (2007). Robust object tracking via online dynamic spatial bias appearance models, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 29, No. 12, (December 2007) pp. 2157-2169.

Cheng, F.-H. & Chen, Y.-L. (2006). Real time multiple objects tracking and identification based on discrete wavelet transform, *Pattern Recognition*, Vol. 39, No. 3, (June 2006) pp. 1126-1139.

Chiang, J.-S.; Lin, H.-T.& Hsia, C.-H. (2008). Novel fast block motion estimation using diamond-arc-hexagon search patterns, *Journal of the Chinese Institute of Engineers*, Vol. 31, No. 6, (September 2008) pp. 955-966.

Collins, R. T.; Lipton, A. J.; Kanade, T.; Fujiyoshi, H.; Duggins, D.; Tsin, Y.; Tolliver, D.; Enomoto, N.; Hasegawa, O.; Burt, P. & Wixson, L. (2000). A system for video surveillance and monitoring, *Carnegie Mellon University*, Technical Report, (2000) CMU-RI-TR-00-12.

Cvetkovic, S.; Bakker, P.; Schirris, J. & With, P. H. N. de. (2006). Background estimation and adaptation model with light-change removal for heavily down-sampled video surveillance signals, *IEEE International Conference on Image Processing*, (October 2006) pp. 1829-1832.

Daubechies, I. & Sweldens, W. (1998). Factoring wavelet transforms into lifting steps, *The Journal of Fourier Analysis and Applications*, Vol. 4, No.3, (1998) pp. 247-269,.

Ge, W.; Gao, L.-Q. & Sun, Q. (2007). A method of multi-scale edge detection based on lifting scheme and fusion rule, *International Conference on Wavelet Analysis and Pattern Recognition*, Vol. 2, (November 2007) pp. 952-955.

Gonzalez, R. C. & Woods, R. E. (2001). *Digital image processing*, Addison-Wesley Longman Publish Co., Inc., Boston.

Hsia, C.-H.; Guo, J.-M. & Chiang, J.-S. (2009). An improved low complexity algorithm for 2-D integer lifting-based discrete wavelet transform using symmetric mask-based scheme, *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 19, No 8, (August 2009) pp. 1201-1208.

Hsia, C.-H.; Guo, J.-M.; Chiang, J.-S. & Lin, C.-H. (2009). A novel fast algorithm based on SMDWT for image applications, *IEEE International Symposium on Circuits and Systems*, (May 2009) pp. 762-765.

Hsieh, C.-C. & Hsu, S.-S. (2007). A simple and fast surveillance system for human tracking and behavior analysis, *IEEE Conference on Signal-Image Technologies and Internet-Based System*, (December 2007) pp. 812-828.

Hu, W.-M.; Zhou, X.; Hu, M. & Maybank, S. (2009). Occlusion reasoning for tracking multiple people, *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 19, No. 1, (January 2009) pp. 114-121.

Hu, W.-M.; Tan, T.-N.; W, L. & Maybank, S. (2004). A survey on visual surveillance of object motion and behaviors, *IEEE Transactions on Systems, Man, and Cybernetics- Part C: Applications and Reviews*, Vol. 34, No. 3, (August 2004) pp. 334-352.

Huang, K.-Q.; Wang, L.-S.; Tan, T.-I. & Maybank, S. (2008). A real-time objects detecting and tracking system for outdoor night surveillance, *Pattern Recognition*, Vol. 41, No. 1, (January 2008) pp. 423-444.

Huang, J.-C.; Su, T.-S.; Wang, L.-J. & Hsieh, W.-S. (2004). Double-change-detection method for wavelet-based moving-object segmentation, *IET Electronics Letters*, Vol. 40, No. 13, (June 2004) pp. 798-799.

Jacobs, N. & Pless, R. (2008). Time scales in video surveillance, *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 18, No. 8, (August 2008) pp. 1106-1113.

Kharate, G. K.; Patil, V. H. & Bhale, N. L. (2007). Selection of mother wavelet for image compression on basis of nature of image, *Journal of Multimedia*, Vol. 2, No. 6, (November 2007) pp. 44-51.

Liu, H.-H.; Chen, X.-H.; Chen, Y.-G. & Xie, C.-S. (2006). Double change detection method for moving-object segmentation based on clustering, *IEEE International Symposium on Circuits and Systems*, (May 2006) pp. 5027-5030.

Mallat, S. G. (1989). A theory for multi-resolution signal decomposition: The wavelet representation, *IEEE Transaction on Pattern Analysis and Machine Intelligence*, vol. 11, no. 7, (July 1989) pp. 674-693.

Mckenna, S. J. (2000). Tracking groups of people, *Computer Vision and Image Understanding*, Vol. 80, No. 1, (October 2000) pp. 42-56.

Sugandi, B.; Kim, H.; Tan, J. K. & Ishikawa, S. (2007). Tracking of moving objects by using a low resolution image, *International Conference on Innovative Computing, Information and Control*, (September 2007) pp. 408-408.

Tab, F. A.; Naghdy, G. & Mertins, A. (2007). Multiresolution video object extraction fitted to scalable wavelet-based object coding, *IET Image Processing*, Vol. 1, No. 1, (March 2007) pp. 21-38.

16

Wavelet Signatures and Diagnostics for the Assessment of ICU Agitation-Sedation Protocols

In Kang, Irene Hudson, Andrew Rudge and J. Geoffrey Chase
[1]*Department of Mathematics and Statistics, University of Canterbury, Christchurch,*
[2]*School of Mathematical and Physical Sciences, University of Newcastle, NSW,*
[3]*Faculty of Health, Engineering and Science Victoria University, Melbourne,*
[4]*Department of Mechanical Engineering, University of Canterbury Christchurch,*
[1,4]*New Zealand*
[2,3]*Australia*

1. Introduction

The use of quantitative modelling to enhance understanding of the agitation-sedation (A-S) system and the provision of an A-S simulation platform are key tools in this area of patient critical care. A suite of wavelet techniques and metrics based on the discrete wavelet transform (DWT) are developed in this chapter which are shown to successfully establish the validity of deterministic agitation-sedation (A-S) models against empirical (recorded) dynamic A-S infusion profiles. The DWT approach is shown to provide robust performance metrics of A-S control and also yield excellent visual assessment tools. This approach is generalisable to any study which investigates the similarity or closeness of bivariate time series of, say, a large number of units (patients, households etc) and of disparate lengths and of possibly extremely long length. This work demonstrates the value of the DWT for assessing ICU agitation-sedation deterministic models, and suggests new wavelet based diagnostics by which to assess the A-S models.

Typically agitation-sedation cycling in critically ill patients involves oscillations between states of agitation and over-sedation, which is detrimental to patient health, and increases hospital length of stay (Rudge et al., 2006a; 2006b; Chase et al., 2004; Rudge et al 2005). Agitation management via effective sedation management is an important and fundamental activity in the intensive care unit (ICU), where in the hospitalized adult agitation is defined as excessive verbal behaviour that interferes with patient care, and the patient's medical therapies (Chase et al., 2004). The main goal of sedation is to control agitation, while also preventing over-sedation and over-use of drugs. In clinical practice, however, a lack of understanding of the underlying dynamics of A-S, combined with a lack of subjective assessment tools, makes effective and consistent clinical agitation management difficult (Chase et al., 2004; Rudge et al., 2005, 2006b). Early agitation management methods traditionally relied on subjective agitation assessment, and sedation assessment scales, combined with medical staff experience and intuition, to deliver appropriate sedation; and an appropriate sedation input response, from recorded at bedside agitation scales (Fraser &

Riker, 2001b; Jaarsma et al., 2001; Ramsay et al., 1974; Ricker et al., 1999; Sessler et al., 2002). The clinical staff at the bedside, usually nurses, then select an appropriate infusion rate based upon their evaluation of these scales, experience, and intuition (Kress et al., 2002). This approach usually leads to the administration of largely continuous infusions which lack a bolus-focused approach, and commonly result in either over sedation, or insufficient sedation (Rudge et al., 2006b). Several recent studies have emphasised the cost and healthcare advantages of drug delivery protocols based on assessment scales of agitation and sedation. A minimal differential equation (DE) model to predict or simulate the patients' agitation-sedation status over time (range [3,001-25,261] time points in minutes) was developed and validated statistically for the first time by Chase et al. (2004). This process is depicted in Figure 3 (see Chase et al., 2004). The goal of the research was to create a physiologically representative pharmacodynamic model that captured the fundamental dynamics of the A-S system. The resulting model can serve as a platform to develop and test semi-automated sedation management controllers that offer the potential of improved agitation management and thus, the clinically relevant outcomes of reduced length of stay in the ICU and reduced health care costs as a result. A-S models were later developed by Rudge et al. (2005, 2006a, 2006b). All these models used either kernel regression, tracking indices, kernel density estimation, a probability band or time within a band as metrics of similarity or closeness of the patient's simulated and recorded A-S profiles. Lee et al. (2005) also developed a nonparametric regression approach with an Epanechnikov kernel (Wand & Jones 1995) to assess the validity of the deterministic A-S models.

The work in this chapter develops novel wavelet signatures and wavelet based statistics and threshold criterion (to assess closeness between pairs of time series). These are applied to the recorded and the simulated infusion rates obtained from the DE models of Chase et al. (2004) to test for commonality across patients, in terms of wavelet correlations. A major aim of this study is to test the feasibility of wavelet statistics to help distinguish between patients whose simulated profiles were "close" to their mean profile a majority of the time profile versus those for whom this was not the case - so-called good versus poor trackers. This research builds on initial work by Kang et al. (2005), which was a preliminary study to assess wavelet signatures for modelling ICU A-S profiles to evaluate "closeness" or "discrimination" of simulated versus actual A-S profiles with respect to wavelet scales. Another earlier application of some of our methods was the study by Kang et al. (2004) on historical, flowering records of 4 Eucalypt species, where it was established that wavelets add credibility to the use of phenological records to detect climate change (see also Hudson, 2010, Hudson 2011, Hudson et al., 2010c and Hudson et al., 2005). This early phenological study was recently expanded from 4 to 8 Eucalypt species by Hudson et al. (2011a, 2011b) (see also Hudson et al., 2010a, 2010b and Hudson & Keatley, 2010).

2. Brief review of wavelets and associated mathematics

Section 2 gives a brief introduction of the basic ideas concerning wavelets. A wave is usually defined as an oscillating function that is localized in both time and frequency. A wavelet is a "small wave", which has its energy concentrated in time to give a tool for the analysis of transient, nonstationary, or time-varying phenomena (Goupillaud et al., 1984; Morlet, 1983). Wavelets have the ability to allow simultaneous time and frequency analysis via a flexible mathematical foundation. Wavelets are well suited to the analysis of transient signals in particular. The localizing property of wavelets allows a wavelet expansion of a transient

component on an orthogonal basis to be modelled using a small number of wavelet coefficients using a low pass filter (Barber et al., 2002). This allows application to a wide range of fields, such as signal processing, data compression, and image analysis (Mallat, 1998; Meyer, 2003; Kumar, 1993, 1994; Donoho, 1995; Chang et al., 2000a, 2000b). The wavelet decomposition of functions is analogous to Fourier decomposition methods (Ogden, 1997; Abramovich & Benjamini, 1995). The wavelet representation is presented first in terms of its simplest paradigm, the Haar wavelet (Haar, 1910). The Haar wavelet is used here to describe the concepts of multiresolution analysis (MRA). For more details about wavelets see, for example, Daubeches (1992), Chui (1992), Donoho & Johnstone (1994), Ogden (1997), Vidakovic (1999), Percival & Walden (2000), and Gencay et al. (2001).

2.1 The Discrete Wavelet Transform (DWT)

Wavelets may be formed from the mother wavelet function $\psi(t)$ via dyadic dilation and integer translation by the following,

$$\psi_{j,k}(t) = 2^{j/2}\psi(2^j t - k) \quad j,k \in Z, \tag{1}$$

where $\mathbb{Z}$ is the set of all integers and the factor $2^{j/2}$ maintains a constant norm independent of scale j. The entire set of wavelets $\psi_{j,k}(t)$ forms an orthonormal basis (Daubeches, 1992). The wavelet functions $\psi_{j,k}$ are ordered according to their dilation index j and translation index k. Higher j corresponds to lower frequency wavelets, and higher k corresponds to a rightward shift. The wavelet transform is usually considered to be a continuous wavelet transform (CWT) (Vidakovic, 1999; Percival & Walden, 2000; Gencay et al., 2001) since it is applied to a function $f(\cdot)$ defined over the entire real axis. However, we only have a finite number N of sampled values, as is usually the case for real data applications. This approach leads to the discrete wavelet transform (DWT). To some degree of approximation, we can regard the DWT as being formed by taking slices through a corresponding CWT (McCoy et al., 1995). Any wavelet in $L^2(R)$ then can then be written as a set of expansion functions,

$$f(t) = \sum_{j,k} a_{j,k} 2^{j/2} \psi\left(2^j t - k\right) \tag{2}$$

where the two-dimensional set of coefficients $a_{j,k}$ is called the discrete wavelet transform of $f(t)$. A more specific form indicating how the $a_{j,k}$'s are calculated, can be written using inner products (Swelden, 1996) as follows,

$$f(t) = \sum_{j,k} \left\langle f(t), \psi_{j,k}(t) \right\rangle \psi_{j,k}(t) . \tag{3}$$

Let X_t be a dyadic length column vector containing a sequence $X_1, X_2, \cdots, X_N$ of $N=2^J$ observations of a real-valued time series. The length N vector of discrete wavelet coefficients W is obtained via $W = \mathcal{W}X$, where $\mathcal{W}$ is an $N \times N$ orthonormal matrix defining the DWT. The vector of wavelet coefficients may be organised into $J + 1$ vectors,

$$\boldsymbol{W} = \left[W_1, W_2, \cdots, W_J, V_J \right]^T \tag{4}$$

where W_j is a length $N/2^j$ vector of wavelet coefficients associated with changes on a scale of length $\lambda_j = 2^{j-1}$ and V_J is a length $N/2^J$ vector scaling coefficients associated with averages on a scale of length $2^J = 2\lambda_J$. Wavelet coefficients thus tell us about variations in adjacent averages (Percival & Walden, 2000).

The structure of the $N \times N$ dimensional matrix $\mathcal{W}$ is visualised through the submatrices $\mathcal{W}_1, \ldots, \mathcal{W}_J$ and $\mathcal{V}_J$ (scaling coefficient matrix) via

$$\mathcal{W} = \left[\mathcal{W}_1, \mathcal{W}_2, \ldots, \mathcal{W}_J, \mathcal{V}_J \right]^T . \tag{5}$$

Let us now consider implementation of the DWT by using a pyramid algorithm (PA) (Mallat, 1989). Let $h = (h_0, \ldots, h_{L-1})$ be the vector of wavelet (high-pass) filter coefficients and $g = (g_0, \ldots, g_{L-1})$ be the vector of scaling (low-pass) filter coefficients (Daubechies, 1992). Graphical representation of the DWT as applied to a dyadic length vector X_t is given by Figure 1 and Figure 2. These depict the analysis of X_t into $\mathcal{W}_1$, $\mathcal{W}_2$ and $\mathcal{V}_2$ using the pyramid algorithm (PA). The synthesis of X_t from $\mathcal{W}_1$, $\mathcal{W}_2$ and $\mathcal{V}_2$ use the inverse of the PA (Figure 2).

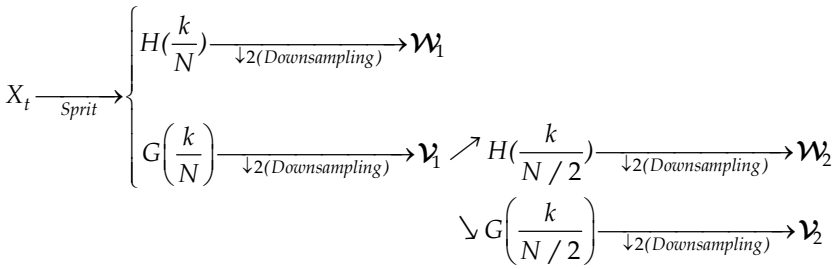


Fig. 1. Flow diagram illustrating the decomposition of X_t into first and second level wavelet coefficients $\mathcal{W}_{1,t}$ and $\mathcal{W}_{2,t}$ and their scaling coefficients $\mathcal{V}_{1,t}$ and $\mathcal{V}_{2,t}$ ($k=0,\ldots,N-1$).

The Inverse DWT (IDWT) is achieved through upsampling the final level of wavelet and scaling coefficients, convolving them with their respective filters and adding up the two filtered vectors. Figure 2 gives a flow diagram for the reconstruction of X_t from the second level wavelet and scaling coefficient vectors. Given a dyadic length time series, it was not necessary to implement the DWT down to level $J = log_2(N)$. A partial DWT (PDWT) may be performed instead that terminates at a level $J_p < J$. The resulting vector of wavelet coefficients will then contain $N - N/2^{J_p}$ wavelets (Percival & Walden, 2000; Gencay et al., 2001). PDWT's are more commonly used in practice than the full DWT because of the flexibility they offer in specifying a scale beyond which a wavelet analysis

into individual large scales is no longer of real interest. A PDWT of level J_0 allows us to relax the restriction that N satisfy, $N = 2^J$ for some J and replace this restriction with the condition that N can be an integer multiple of 2^{J_0} (Percival & Walden, 2000; Gencay et al., 2001).

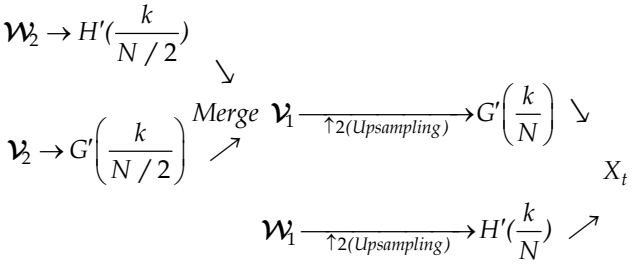


Fig. 2. Flow diagram illustrating the reconstruction of x_t^2 from first and second level wavelet coefficients $\mathcal{W}_{1,t}$ and $\mathcal{W}_{2,t}$ and their scaling coefficients $\mathcal{V}_{1,t}$ and $\mathcal{V}_{2,t}$ (k=0,…,N-1).

2.2 The Maximal Overlap Discrete Wavelet (MODWT)

The DWT is an alternative to the Fourier transform (FT) for time series analysis. The DWT provides wavelet coefficients that are local in both time and frequency. In this section the maximal overlap DWT (MODWT) which is a modified version of the discrete wavelet transform is discussed. Like the DWT, the MODWT is defined in terms of a computationally efficient pyramid algorithm (PA). The term MODWT comes from the relationship of the MODWT with estimators of the Allan variance (Allan, 1966). The MODWT gives up orthogonality in order to gain features the DWT does not possess. A consequence of this is that the wavelet and scaling coefficients must be rescaled in order to retain the variance preserving property of the DWT (Percival & Guttorp, 1994).

Property	DWT	MODWT
Data	N=2^J	Any sample size N
Detail and Smooth Coefficients of MRA	Downsampling	Associated with zero phase filters
Circularly shifting	Does not hold	Holds and Invariant
Wavelet Variance	Less efficient	Asymptotically Efficient

Table 1. Properties of the DWT and MODWT

The properties in Table 1 are important in distinguishing the MODWT from the DWT (Percival & Mofjeld, 1997; Percival & Walden, 2000; Gencay et al., 2001). The decomposition and reconstruction procedure and inverting of the MODWT is similar to the DWT. A key feature to an MRA using the MODWT is that the wavelet details and smooth are associated

with zero-phase filters. Thus, interesting features in the wavelet details and smooth may be perfectly aligned with the original time series. This attribute is not available through the DWT since it subsamples the output of its filtering operations (Percival & Walden, 2000; Gencay et al., 2001).

2.3 Wavelet-based estimators of correlation

The length N vector of discrete wavelet coefficients W is obtained via $\boldsymbol{W} = \mathcal{W}\boldsymbol{X}$ and $\boldsymbol{W} = \left[W_1, W_2, \cdots, W_J, V_J\right]^T$ where $\mathcal{W}$ is an $N \times N$ orthonormal matrix defining the DWT, $\boldsymbol{W}_j$ is a length $N/2^j$ vector of wavelet coefficients associated with changes on a scale of length $\lambda_j = 2^{j-1}$ and $\boldsymbol{V}_J$ is a length $N/2^J$ vector scaling coefficients associated with averages on a scale of length $2^J = 2\lambda_J$ in Section 2.1. Due to orthonormality, $X = \mathcal{W}^T\boldsymbol{W}$, and the squared wavelet coefficients of $\boldsymbol{W}$ the norm, $\|\boldsymbol{W}\|^2$, is the same as $\|X\|^2$ because of the equality $\|X\|^2 = X^T X = \left(\mathcal{W}^T\boldsymbol{W}\right)^T\left(\mathcal{W}^T\boldsymbol{W}\right) = \boldsymbol{W}^T\mathcal{W}\mathcal{W}^T\boldsymbol{W} = \|\boldsymbol{W}\|^2$. Then $\|X\|^2$ can be decomposed on a scale-by-scale basis via $\|X\|^2 = \|\boldsymbol{W}\|^2 = \sum_{j=1}^{J}\|\boldsymbol{W}_j\|^2 + \|V_J\|^2$. The wavelet correlation and cross-correlation between two time series can now be defined. The wavelet correlation (WCORR) is the correlation between the scale λ_j wavelet coefficients of bivariate time series. Likewise the wavelet covariance is the covariance between the scale λ_j wavelet coefficients from bivariate time series. We introduce a lag τ, between the two series, to obtain the wavelet cross-covariance (WCCOVA) and wavelet cross-correlation (WCCORR) (Percival & Walden, 2000; Gencay et al., 2001), as described below.

Let X_t and Y_t be two time series of interest. The wavelet cross-covariance of (X_t, Y_t) for scale $\lambda_j = 2^{j-1}$ and lag τ is defined (Percival, 1995; Percival & Guttorp, 1994; Whitcher et al., 2000; Gencay et al., 2001) as follows,

$$\gamma_{\tau,XY}\left(\lambda_j\right) \equiv Cov\left\{\overline{W}_{j,t}^{X}, \overline{W}_{j,t+\tau}^{Y}\right\} \tag{6}$$

where $\overline{W}_{j,t}^{X}$ and $\overline{W}_{j}^{Y}$ are the scale λ_j MODWT coefficients for X_t and Y_t. When the width of wavelet or scaling filter is equal or greater than two times of the number of differencing operations, the MODWT coefficients have mean zero and therefore the covariance reduces to an expectation of a product (Percival & Walden, 2000). By setting $\tau = 0$ and X_t to Y_t, $\gamma_{\tau,XY}\left(\lambda_j\right)$ reduces to the wavelet variance for X_t or Y_t denoted by $\sigma_X^2\left(\lambda_j\right)$ or $\sigma_Y^2\left(\lambda_j\right)$. The wavelet correlation of (X_t, Y_t) at scale λ_j=2^{j-1} is then defined as follows,

$$\rho_{XY}\left(\lambda_j\right) = \frac{Cov\left\{\overline{W}_{j,t}^{X}, \overline{W}_{j,t}^{Y}\right\}}{\sigma_X\left(\lambda_j\right).\sigma_Y\left(\lambda_j\right)} = \frac{\gamma_{XY}\left(\lambda_j\right)}{\sigma_X\left(\lambda_j\right).\sigma_Y\left(\lambda_j\right)}. \tag{7}$$

Rationale for wavelets

Wavelets allow time series data to be decomposed on a scale by scale basis, or to be discretized, into its so-called underlying subcomponents.

Conventional time frequency domain methods results may be difficult to interpret, whereas the wavelet-correlation is able to display how the association between two time series change with wavelet scale.

Transformation of the data (orthonormal) allows correlation, cross-correlational analyses of bivariate series to be performed - based on the derived wavelet coefficients.

DWT is often less computer intensive than other transformations (e.g. fast Fourier transform).

DWT offers easier analysis than the CWT as most time series are sampled as discrete values. DWT allows for the decorrelation of time series.

Qualitative Description of the DWT and MODWT

DWT transforms the original time series X into its DWT coefficients $W = wY$, where w is a $N \times N$ orthonormal matrix

− the components, W_j, of W are associated with coefficients for each scale $\lambda_j = 2^{j-1}$

Wavelet coefficients Wj inform on the variations in adjacent averages over λ_j.

There are $\frac{N}{2\lambda_j}$ wavelet coefficients for each scale $\lambda_j \equiv 2^{j-1}$, j = 1, 2 . . , J_0 where $\lambda_j = 2^{j-1}$ is a so-called standardized scale, whereas $\lambda_j\Delta t$ is a physical scale, where Δt is the sampling interval.

The MODWT is a non-decimated variation of the DWT, which defines the jth level MODWT detail subcomponent of the time series as $\overline{D}_j = \overline{W}_j^T \overline{W}_j$ and defines the jth level MODWT smoothed series, $\overline{S}_j = \overline{V}_j^T \overline{V}_j$, which is related to the average (over scale N), and is normally interpretable as the trend.

Cross- Correlation and correlation

The scale λ_j MODWT coefficients may be used to investigate the correlation and cross-correlation of two time series, X_t and Y_t.

The wavelet cross correlation of X_t, Y_t at scale $\lambda_j = 2^{j-1}$ for a time lag τ, is defined as

$$\rho_{XY,\tau}(\lambda_j) = \frac{Cov\left\{\overline{W}_{j,t}^X \overline{W}_{j,t}^Y\right\}}{\sigma_X(\lambda_j)\cdot\sigma_Y(\lambda_j)} = \frac{\gamma_{XY,\tau}(\lambda_j)}{\sigma_X(\lambda_j)\cdot\sigma_Y(\lambda_j)},$$

where $\overline{W}_{j,t}^X$ *and* $\overline{W}_{j,t}^Y$ *are*

scale λ_j MODWT coefficients. Correlation is the cross-correlation at lag $\tau = 0$.

MODWT-MRA

The orthonormal matrix, $N \times N$, leads to scale-based decompositions. Given the MODWT coefficients, Y can be constructed as an additive decomposition, known as a multiresolution analysis (MRA).

Specifically the level J_0 MODWT-based MRA is given by $Y = \sum_{j=1}^{J_0} \overline{d}_j + \overline{s}_{j0}$, $\overline{d}_j$ are the "detail series", (j= 1, 2, . . ., J_0)

for a pre- specified J_0, and are part of the MRA of Y that can be attributed to variations on a scale of λ_j.

Table 2. Summary of the wavelet mathematics and rationale.

where $\sigma_X^2(\lambda_j) = var\left\{\overline{W}_{j,t}^X\right\}$ is the wavelet variance with scale λ_j. $\overline{W}_{j,t}^X$ and $\overline{W}_{j,t}^Y$ are the scale λ_j MODWT coefficients for X_t and Y_t, respectively (Percival & Walden, 2000; Gencay et al., 2001). As with the usual correlation coefficient (between two random variables), the range of $\rho_{XY}(\lambda_j)$ is the interval –1 to 1 for all *j*. The typical cross-correlation statistic is purely a function of the cross-covariance and standard deviations. Thereby the MODWT estimator of the wavelet cross-correlation (WCCORR) of the two processes, which are at variance by an integer lag τ, is defined as

$$\rho_{XY,\tau}(\lambda_j) = \frac{Cov\left\{\overline{W}_{j,t}^X, \overline{W}_{j,t}^Y\right\}}{\sigma_X(\lambda_j).\sigma_Y(\lambda_j)} = \frac{\gamma_{XY,\tau}(\lambda_j)}{\sigma_X(\lambda_j).\sigma_Y(\lambda_j)}. \tag{8}$$

As the usual cross-correlation is used to determine lead or lag relationships between two series, the wavelet cross-correlation provides a lead or lag relationship between X_t and Y_t on a scale-by-scale basis. In terms of confidence intervals (CIs) for the WCORR and WCCORR parameters, a nonlinear transformation is required to produce reasonable CIs for the correlation coefficient (Gencay et al., 2001). We use the Fisher's z-transformation (Dépué 2003) which is defined as follows,

$$h(\rho) = \frac{1}{2} log\left(\frac{1+\rho}{1-\rho}\right) = tanh^{-1}(\rho). \tag{9}$$

An unbiased estimator of the WCORR based on the MODWT in Equation (7) is $\tilde{\rho}$. The given estimated correlation coefficient $\tilde{\rho}$, based on *N* independent Gaussian observations, has the following limiting distribution (Percival & Walden, 2000; Gencay et al., 2001).

$$\sqrt{N-3}\left[h(\tilde{\rho}) - h(\rho)\right] \sim N(0,1). \tag{10}$$

Applying the transformation tanh maps the confidence interval back to [−1, 1] to produce an approximate 95% CI for ρ as follows (Whitcher et al. 2000; Gencay et al. 2001; Hudson et al., 2010b)

$$tanh\left\{h\left[\tilde{\rho}_{XY}(\lambda_j)\right] \pm \zeta_{\frac{\alpha}{2}}\left(\frac{1}{\hat{N}_j - 3}\right)^{1/2}\right\}. \tag{11}$$

The quantity $\hat{N}_j$ in Equation (11) is the number of the DWT coefficients associated with scale λ_j. Table 2 gives a brief overview of the wavelet mathematics used in this chapter.

3. Application to Agitation Sedation (A-S) wavelet modelling

This section presents the application of a wavelets analysis of the agitation-sedation (A-S) data of 37 ICU patients' bivariate time series, sourced from the research of Chase et al. (2004). An extensive description of A-S modelling, as well as other references on A-S control;

along with details of the development and validation of the A-S model are given in Chase et al. (2004). The model by Chase et al. (2004) serves as a platform to develop and test semi-automated sedation management controllers that offer the potential of improved agitation management and reduced length of stay in the ICU. Figure 3 presents a diagram of the feedback loop employing nursing staff feedback of subjectively assessed patient agitation through the infusion controller (Chase et al., 2004). We refer the reader also to the later works of Lee et al. (2005) and of Rudge et al. (2005, 2006a, 2006b) who developed further A-S models and metrics. Table 3 summarises the equations used, mathematical methods employed and the aims of the given study, along with the performance indicators derived for each of Chase et al. (2004), Rudge et al. (2006a, 2006b, 2005), and Lee et al. (2005). As such Table 3 and subsequently Table 8 show how the research presented in this chapter adds knowledge and insight into A-S modelling in the context of these earlier works.

3.1 Using the DWT and MODWT

The DWT, the maximal overlap (MODWT) and multiresolution analysis (MRA) were applied to all pairs of patient specific infusion profiles (recorded (R) and simulated (S)) for the 37 ICU patients. The aim of the analysis reported in section 3.1 – 3.3 is to investigate

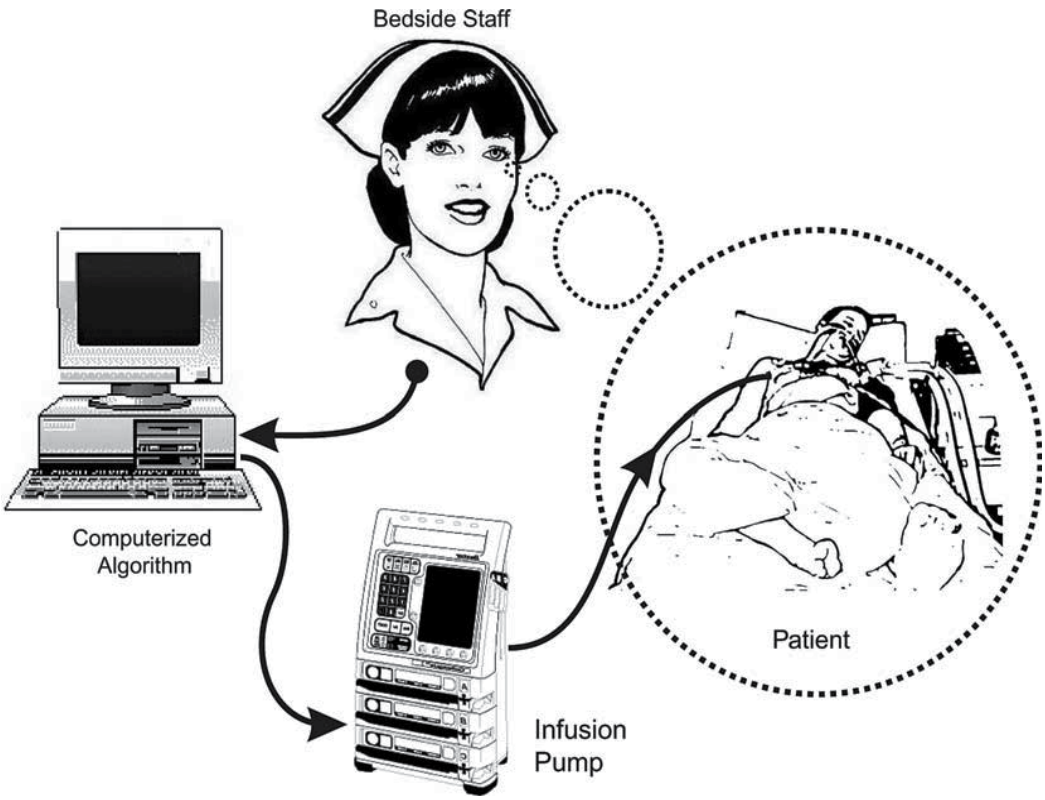


Fig. 3. Diagram of the feedback loop employing nursing staff feedback of subjectively assessed patient agitation through the infusion controller (diagram is sourced from Chase et al. (2004)).

whether wavelets based diagnostics can reliably assess how well the A-S model (simulation) captures the underlying dynamics of the true recorded infusion rates at different horizons via the DWT; and to compare these results with the diagnostics of Chase et al. (2004), Rudge et al. (2006a, 2006b, 2005) and Lee et al. (2005). For illustration of these concepts patient specific recorded (R) and simulated (S) profiles (as the thick line, according to the equation of Chase et al. (2004)) are shown in Figure 4. It is noteworthy that simulation of the A-S states using the model of Rudge et al. (2005) showed that a reduction in both the magnitude of agitation and the severity of agitation sedation cycling is possible. Mean and peak agitation levels were reduced by 68.4% and 52.9%, respectively, on average, with some patients exhibiting in excess of a 90% reduction in mean agitation level through increased control gains. Implementation of automated feedback infusion controllers based on such models could thus offer simple and effective drug delivery, without significant increases in drug consumption and expenses.

The lag/lead relationship between the S and R infusion series was investigated on a scale-by-scale basis via a MODWT-MRA (using the LaDaub (8) filter); thereby each patient's S and R series can be expressed as a new set of series, called details and smooth. Each of these series are associated with variations at a particular wavelet scale. The results of a MODWT-MRA (not detailed here due to space restrictions) reveal that thirteen patients (patients 3, 9, 11, 17, 20, 22, 26, 30, 33, 34, 35, 36, 37) (Figure 4) have recorded infusion series that lead their corresponding simulated infusion series. It is noteworthy that of these 13 patients, which exhibit such a lagged dependency, our DWT wavelet diagnostics (and those of Chase et al., (2004) and Rudge et al., (2006b)) identify the following as poor performers in common (Patients 9, 11, 17, 22, 33, 34 and 35). Overall it is thought that the simulated profile peaks later than the patient's recorded infusion possibly due to the delay in distribution time for the drug. This result implies that, while performing well most of the time, the simulated rate is lagging behind the patient's true infusion rate. These periods indicate times of the patient's hospital length of stay in ICU, where the DE model may not capture the subject's specific A-S dynamics (evidenced by the time lags). These periods may correspond to periods of marked distress or physiological alterations due to the patient's state. A common reason for the departure of the simulated profile is this apparent time-lag. Particularly small departures indicate rapid increases (or decreases) in the recorded infusion rate, where the simulated infusion rate appears to lag behind. These differences may be a result of the medical staff's over or under-assessment of the patient's agitation status, this hypothesis is as yet not proven.

3.2 Wavelet correlation and other diagnostics

In section 3.2 an estimate of wavelet correlation (WCORR), $\tilde{\rho}$ from Equation (7), is derived per patient. This WCORR between the scale λ_j wavelet coefficients of each patient's bivariate (S, R) time series is used to assess how the simulated (S) and recorded (R) infusion series correlate. Graphical assessment tools and derived wavelet-based metrics are then suggested and proven valid for ICU A-S management**. T**he wavelets based performance indices developed in this chapter pertain to the modulus of the wavelet correlation at wavelet scale (λ_j) (level 1) and also to summary statistics founded on measures of the wavelet correlations and wavelet cross-correlations over scales, as defined below (Table 6). For the ith patient (i=1, 2, …, 37) count the number of its wavelet scales (out of a maximum of 8) for which the WCORR value at scale λ_j is not significant (at the 5% level of significance). This variable is

denoted by "Count of NS" and given per patient, with specification of the significance (at the 5% level of significance) of WCORR at λ_j (j=1, 2, …, 8) as either significant (S) or not significant (NS) in Table 4. Patient specific AND (average normalized density), RAND (relative average normalized density), and TI (tracking index) values, all derived by Rudge et al (2006b) and Chase et al. (2004), are also shown in Table 4. Definitions for the AND, RAND and TI performance indicators or diagnostics are detailed in Table 3.

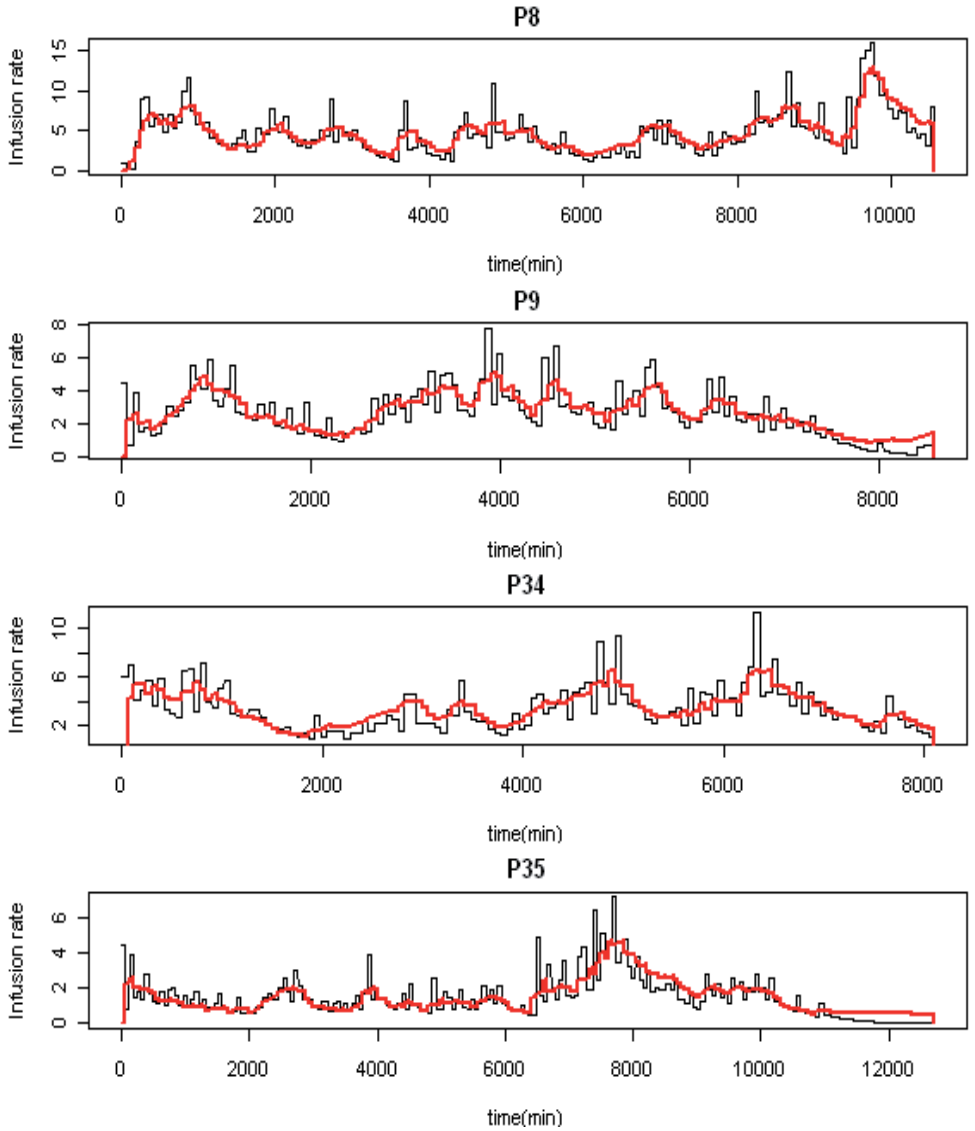


Fig. 4. Example of the delay between the recorded and simulated (thick line) infusion profiles for Patients 8, 9, 34 and 35 (denoted by P8, P9, P34 and P35). P8 is shown to be a good tracker, the remainder are poor trackers, according to the indices developed in this chapter.

Authors	Equations or Model used	Mathematical methods used	Aims of the study and the performance indicators derived	Patient numbers with poor fit to simulated profiles (poor trackers) and their summary profile
Rudge et al. (2006 a, b)	I. Pharmacokinetics of morphine : $V_c^o \frac{dC_c^o}{dt} = -\left(K_{CL}^o + K_{ce}^o + K_{cp}^o\right)C_c^o + P^o U + K_{ce}^o C_e^o + K_{pc}^o C_p^o$ $V_p^o \frac{dC_p^o}{dt} = -K_{pc}^o C_p^o + K_{cp}^o C_c^o, V_e^o \frac{dC_e^o}{dt} = -K_{ce}^o C_e^o + K_{ce}^o C_c^o$ II. Pharmacodynamics of midazolam : $V_c^s \frac{dC_c^s}{dt} = -\left(K_{CL}^s + K_{ce}^s\right)C_c^s + P^s U + K_{ce}^s C_e^s, V_e^s \frac{dC_e^s}{dt} = -K_{ce}^s C_e^s + K_{ce}^s C_c^s$ III. Pharmacodynamics of morphine and midazolam : $\frac{dA}{dt} = w_1 S - w_2 K_T \int_0^t E_{comb}(\zeta) e^{-K_T(t-\zeta)} d\zeta, E_{Comb} = E_0 + \left[E_{max}(\theta) - E_0\right] \frac{\left(\frac{C_o + C_S}{C_{50}(\theta)}\right)^{\gamma(\theta)}}{1 + \left(\frac{C_o + C_S}{C_{50}(\theta)}\right)^{\gamma(\theta)}}$	Kernel smoothing Chebychev's inequality for probability band (Chase et al., 2004) Relative average normalised density (RAND) Average normalised density (AND)	Develop a physiologically representative model that incorporates endogenous agitation reduction (EAR). Use performance measures as follows: 1. RTD: relative total dose (RTD) expresses the total dose administered in the simulation as a percentage of the actual total recorded dose. 2. Relative average normalised density (RAND) measures how probabilistically similar the model outputs are to the smoothed data, and hence the degree of comparability between the model and the empirical data. 3. Percentage time in band (TIB).	Exact patient identifiers of those who exhibit poor fit are not specified in Rudge et al. (2006b), but of 10 poorly modelled patients, 5 have RAND values from 0.45 to 0.5, and the remaining 5 have RAND values less than 0.45. These patients with poor fit are P7, P10, P11, P13, P22, P27, P28, P29, P32 and P35 The remaining 27 patients all have RAND values greater than 0.50.
Rudge et al. (2005)	I. Pharmacokinetics of morphine : $V_c^o \frac{dC_c^o}{dt} = -\left(K_{CL}^o + K_{ce}^o + K_{cp}^o\right)C_c^o + P^o U + K_{ce}^o C_e^o + K_{pc}^o C_p^o$ $V_p^o \frac{dC_p^o}{dt} = -K_{pc}^o C_p^o + K_{cp}^o C_c^o, V_e^o \frac{dC_e^o}{dt} = -K_{ce}^o C_e^o + K_{ce}^o C_c^o$ II. Pharmacodynamics of midazolam : $V_c^s \frac{dC_c^s}{dt} = -\left(K_{CL}^s + K_{ce}^s\right)C_c^s + P^s U + K_{ce}^s C_e^s, V_e^s \frac{dC_e^s}{dt} = -K_{ce}^s C_e^s + K_{ce}^s C_c^s$ III. Pharmacodynamics of morphine and midazolam : $\frac{dA}{dt} = w_1 S - w_2 K_T \int_0^t E_{comb}(\zeta) e^{-K_T(t-\zeta)} d\zeta, E_{Comb} = E_0 + \left[E_{max}(\theta) - E_0\right] \frac{\left(\frac{C_o + C_S}{C_{50}(\theta)}\right)^{\gamma(\theta)}}{1 + \left(\frac{C_o + C_S}{C_{50}(\theta)}\right)^{\gamma(\theta)}}$	Infinite Impluse Response (IIR) filter (Rorabaugh, 1998) Proportional-Derivative (PD) control with agitation for infusion rate (U) Moving blocks bootstrap (Efron & Tibshirani, 1993) Tracking Index (TI)	Develop a control model to capture the essential dynamics of the agitation-sedation system. Use performance measures as follows: 1. $U = K_p A + K_d A$ for the infusion rate. 2. Tracking Index (TI): A quantitative parameter to indicate how well the simulated infusion profile represents the average recorded infusion.	Exact patients who exhibit poor fit are not specified in Rudge et al. (2005). 8 patients have simulated infusion profiles that lie within the probability band less than 70% of the time. The median patient TI across all patients is 87.0 with a 95% CI of (85.9, 88.1).

Table 3. Overview of Agitation-Sedation studies of ICU patients.

Authors	Equations or Model used	Mathematical methods used	Aims of the study and the performance indicators derived	Patient numbers with poor fit to simulated profiles (poor trackers) and their summary profile
Lee et al. (2005)	I. Pharmacokinetics of morphine : $V_c^o \frac{dC_c^o}{dt} = -\left(K_{CL}^o + K_{co}^o + K_{cp}^o\right)C_c^o + P^o U + K_{pc}^o C_p^o$ $V_p^o \frac{dC_p^o}{dt} = -K_{pc}^o C_p^o + K_{cp}^o C_c^o, V_e^o \frac{dC_e^o}{dt} = -K_{eo}^o C_e^o + K_{ce}^o C_c^o$ II. Pharmacodynamics of midazolam : $V_c^s \frac{dC_c^s}{dt} = -\left(K_{CL}^s + K_{ce}^s\right)C_c^s + P^s U + K_{ec}^s C_e^s, V_e^s \frac{dC_e^s}{dt} = -K_{eo}^s C_e^s + K_{ce}^s C_c^s$ III. Pharmacodynamics of morphine and midazolam : $\frac{dA}{dt} = w_1 S - w_2 K_T \int_0^t E_{comb}(\zeta) e^{-K_T(t-\zeta)} d\zeta, E_{Comb} = E_0 + \left[E_{\max}(\theta) - E_0\right] \frac{\left(\frac{C_o + C_s}{C_{50}(\theta)}\right)^{\gamma(\theta)}}{1 + \left(\frac{C_o + C_s}{C_{50}(\theta)}\right)^{\gamma(\theta)}}$	Kernel regression (Wand & Jones, 1995) Kernel density estimation: marginal density function of the regression function estimate Nonparametric regression Chebychev's inequality for the probability band (Chase et al., 2004)	Develop a nonparametric approach for assessing the validity of deterministic dynamic models against empirical data. Use performance measures as follows: 1. Kernel regression and kernel density estimation to yield visual graphical assessment of the data. 2. Construct a probability band for the nonparametric regression curve and check whether the proposed model lies within the band, for a significant proportion of the time. 3. Average normalised density (AND) to measure how well the simulated values coincide with the maximum density at every time point and the relative average normalised density (RAND).	Exact patients who exhibit poor fit are not specified in Lee et al. (2005), however of the 37 patients, 27 have RAND values greater than 0.5. Of the remaining 10 patients, 7 have RAND values from 0.43 to 0.49, and 3 have low RAND values between 0.34-0.38.
Chase et al. (2004)	The agitation-sedation system model: Phamarcokinetic model (Wood & Wood, 1990) adding patient agitation as a third state variable $\frac{dC_c}{dt} = -K_1 C_c + \frac{U}{V_d}$ $\frac{dC_p}{dt} = -K_2 C_p + K_3 C_c$ $\frac{dA}{dt} = w_1 S - w_2 \int_0^t C_p(\tau) e^{-K_4(t-\tau)} d\tau$ Uniform kernel with bandwidwidth *h* (Wand & Jones, 1995) $K_h = \begin{cases} 0 & \text{if } t < -\frac{h}{2} \\ 1 & \text{if } -\frac{h}{2} < t \le \frac{h}{2} \\ 0 & \text{if } t > \frac{h}{2} \end{cases}$	Infinite Impluse Response (IIR) filter (Rorabaugh, 1998) Proportional-Derivative (PD) control with agitation for infusion rate (U) Tracking Index (TI) Chebychev's inequality for the probability band (Chase et al., 2004)	Develop a mathematical model to capture the essential dynamics of the agitation-sedation system and develop statistical validity tests using the recorded infusion data for 37 ICU patients. Use performance measures as follows 1. Kernel smoothing using the uniform kernel. 2. Tracking Index (TI). 3. Moving blocks bootstrap to gain an understanding of the reliability TI for a given patient's infusion profile 4. 90% Probability Band, by definition the range within at least 90% of the time, the estimated mean value of the recorded infusion rate lies within the band.	P6, P7, P9, P12, P17, P21, P27, P34 are the 8 patients which have a TI values less than 70%, i.e. do not have infusion profiles which lie within the probability band at least 70% of the time. The remaining 29 patients have TI values greater than 70%.

From Lee et al (2005) and Chase et al (2004): "Po and Ps are the concentrations of morphine and midazolam, respectively (mgmL−1), where terms with superscript 'o' relate to the opioid morphine, and terms with superscript 's' relate to the sedative midazolam. Time is represented by t (min), the variable of integration, and the terms w1 and w2 are the relative weights of stimulus and cumulative effect, representing the patient sensitivity. Finally, Ecomb is the combined pharmacodynamic effect of the individual effect site drug concentrations of morphine and midazolam determined using response surface modeling as defined in Minto et al. (Minto et al., 2000)."

Poor trackers are identified via wavelet diagnostics as follows: Patients with a "Count of NS" greater or equal to 2 and a non-significant WCORR value at scale λ_1 (level 1) are said to be a poor tracker; as are patients with a "Count of NS" less or equal to 3 and a *significant negative* WCORR value at scale λ_1 (level 1) and a *significant negative* WCORR value at scale λ_6. Table 4 indicates 15 such poor trackers (given in bold), as defined by the wavelet based diagnostics ("Count of NS" and WCORR at scales λ_1 and λ_6) derived in this chapter. Poor tracking implies that the patient's simulated A-S profiles do not mirror their actual (recorded) infusion profile over time (according to their profile of wavelet correlations).

Note that the threshold values delineating poor tracking for AND and RAND (according to Rudge et al. (2006b)) are not taken into account in classifying a patient as either a poor or good tracker (Table 4). In this chapter our criterion for tracking classification is based solely on the patient's WCORR values at scales λ_j (j=1, 2, …, 8), their significance or otherwise, and on their "Count of NS" wavelet correlations (Table 4). As mentioned above Table 4 also indicates patients which are considered to track poorly according to both Rudge et al's (2006b) and Chase et al.'s (2004) diagnostics, the latter are based on completely different mathematical methodologies as summarised in Table 3. It is noteworthy that 11 of our 15 wavelet based poor trackers are also considered to be poor trackers by either or both of Rudge et al.'s (2006b) and Chase et al.'s (2004) performance indices.

Figure 5 shows the estimated WCORR, $\tilde{\rho}$ and 95% CI for four patients, P2 and P4 (are poor trackers) and P8 and P14 (are good trackers). From Figure 5 WCORR is generally significant for wavelet levels 1, 2, 4, 8 and 16 for the good trackers (whether they are a positive or negative WCORR value). This is not the case for the poor trackers, who generally exhibit a non-significant WCORR at λ_j for j < 5. Figures 6 and 7 display each patient's multivariate profile of AND, RAND and "Count of NS" (divided by 10 for axis scaling purposes), for increasing $|\lambda_1|$, for the poor trackers and good trackers, respectively. Figures 6-7 clearly show that the profile of ("Count of NS"/10) is invariably higher for the poor trackers; and RAND, AND and $|\lambda_1|$ profiles are invariably higher for the good trackers.

By using the data per patient (from Table 4), we can perform a Kruskal Wallis test to statistically compare the medians of the performance indicators between the wavelet based good and poor trackers. These results are summarized in Table 5 and Table 6. Specifically Table 6 gives the results of the Kruskal Wallis (k-w) tests for our wavelet based poor versus good tracker groups for measures of WCORR at scale λ_j (j=1, 2, …, 8), "Count of NS", $|\lambda_1|$, in addition to AND, RAND and TI per patient. From Table 6 we observe that the median wavelet correlations for the first 5 wavelet scales are significantly lower for the poor trackers (15 of 37 patients), as are the median absolute value of the wavelet correlation at λ_1, and of AND, RAND and TI (P < 0.006). The median of the number of non significant wavelet correlations ("Count of NS") (an integer out of 8, at λ_j, j=1, 2,…, 8) is 5.0 for the poor trackers, significantly higher than the median of 2.0 for the good tracking group (P = 0.001) (Table 6). It is noteworthy also that the patient specific WCORR profiles are good visual "signatures" of the patient's tracking status (good or poor tracking) (see Figure 5).

Recall that 11 of the 15 DWT based poor trackers are also considered to be poor trackers by either or both Rudge et al.'s (2006b) and Chase et al.'s (2004) (non wavelet based) performance indicators. Indeed kappa tests of agreement show that our DWT WCORR

criterion for poor tracking, as developed in this chapter, agrees significantly with that of the performance thresholds of Chase et al (2004) (*kappa* = 0.2127, P=0.01) and with that of Rudge et al. (2006b) (*kappa* = 0.5856, P=0.001). It is noteworthy also that the threshold criterion for poor tracking for RAND by both Rudge et al.'s (2006b) and Chase et al.'s (2004) performance indicators is corroborated by our poor trackers, identified by the wavelet metrics derived in this chapter, shown to have a median RAND of 0.50 (Table 6), which is exactly the threshold used by Rudge et al. (2006b) and Chase et al. (2004).

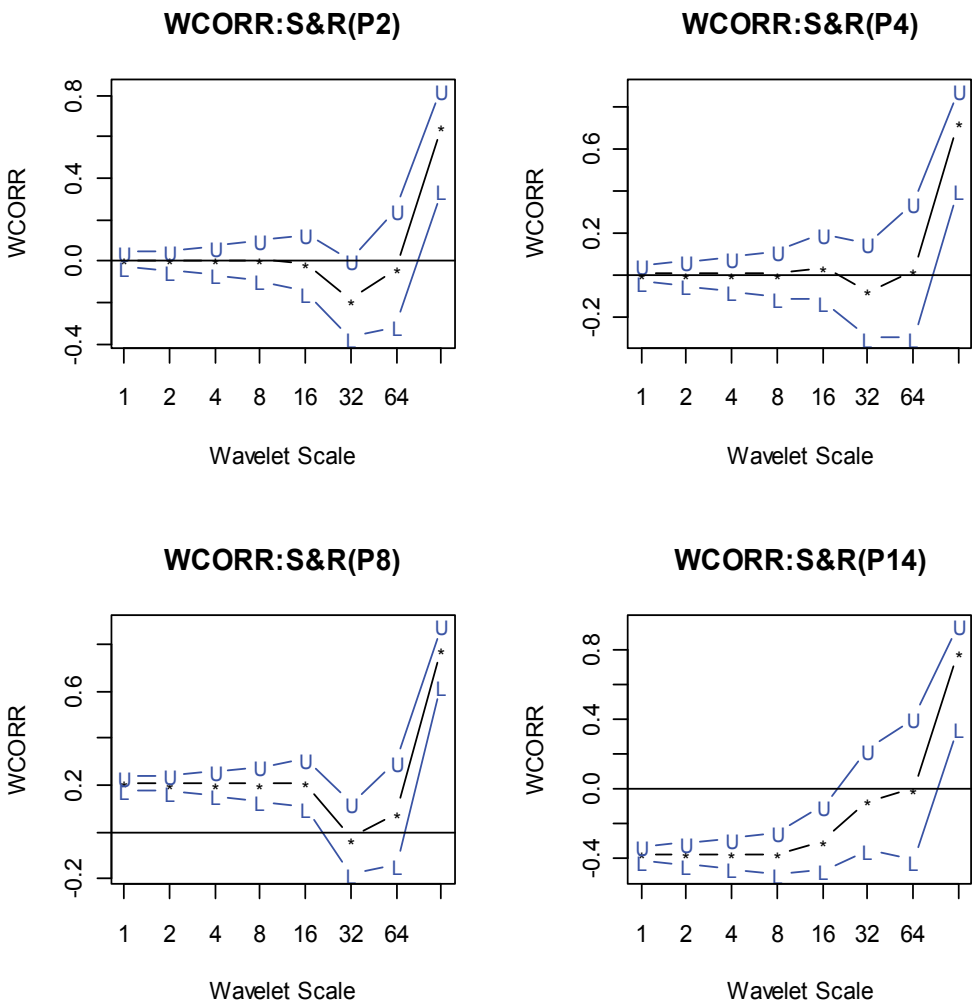


Fig. 5. Wavelet correlation for Patient 2 (Top LHS), Patient 4 (Top RHS), Patient 8 (bottom LHS) and Patient 14 (bottom RHS) with the approximate 95% confidence interval. Patients 2 and 4 are poor trackers in contrast to Patients 8 and 14 who are good trackers with >4 significant WCORR at wavelet scales (1, 2, 4, 8, 16).

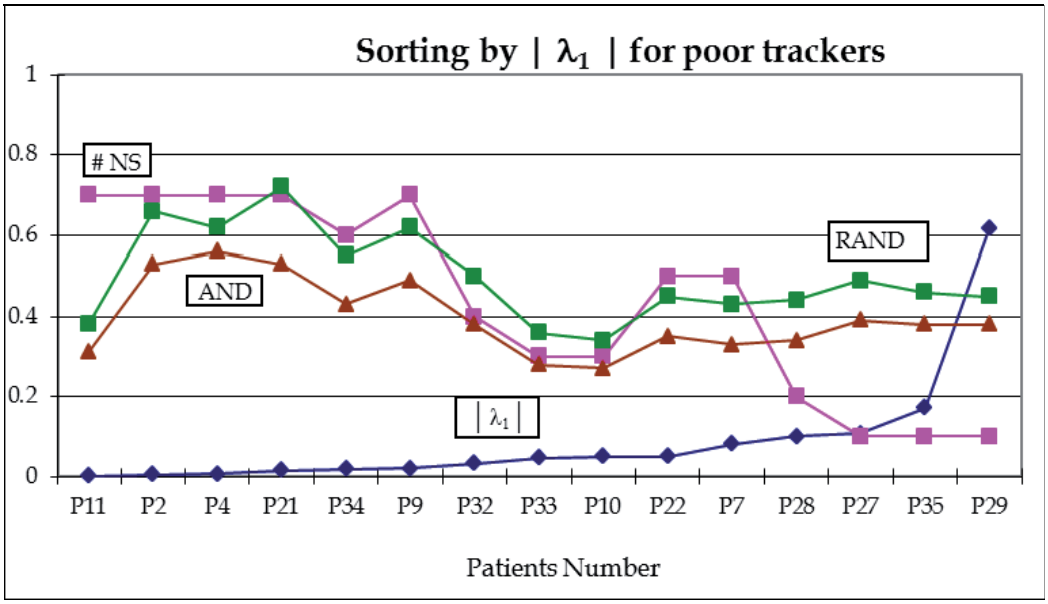


Fig. 6. A comparison of RAND (green square), AND (burgundy triangle), the number of non-significant λ_j (divided by 10) (pink square), and the modulus of λ_1 (blue diamond) for the poor trackers (P11, P2,…, P29) sorted by increasing $|\lambda_1|$.

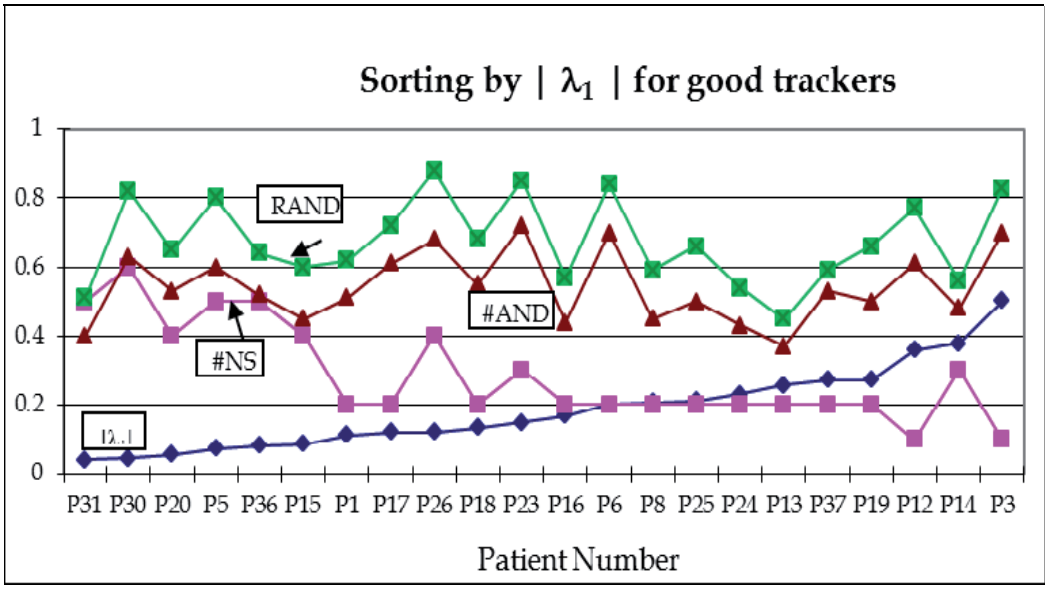


Fig. 7. A comparison of RAND (green square), AND (burgundy triangle), the number of non- significant λ_j (divided by 10) (pink square), and the modulus of λ_1 (blue diamond) for the good trackers (P31, P30,..., P3) sorted by increasing $|\lambda_1|$.

Patient ID	Wavelet scale λ_j (j= 1, 2, …, 8)								Count of NS*	Rudge et al. (2006b) diagnostics		Chase et al. (2004) diagnostics
	1	2	4	8	16	32	64	128		AND·	RAND·	TI (SE) ·
P1	S	S	S	S	S	NS	NS	S	2	0.51	0.62	87.0 (0.041)
P2	NS	NS	NS	NS	NS	NS	NS	S	7	0.53	0.66	86.2 (0.037)
P3	S	S	S	S	S	S	N.S	S	1	0.70	0.83	88.8 (0.015)
P4	NS	NS	NS	NS	NS	NS	NS	S	7	0.56	0.62	80.1 (0.095)
P5	S	S	NS	NS	NS	NS	NS	S	5	0.60	0.80	91.1 (0.016)
P6	S	S	S	S	S	NS	NS	S	2	0.70	0.84	87.0 (0.014)
P7	S	S	NS	NS	NS	NS	NS	S	5	0.33	0.43	84.5 (0.068)
P8	S	S	S	S	S	NS	NS	S	2	0.45	0.59	87.4 (0.027)
P9	NS	NS	NS	NS	NS	NS	NS	S	7	0.49	0.62	87.3 (0.024)
P10	S	S	S	NS	NS	S	NS	S	3	0.27	0.34	83.4 (0.041)
P11	NS	NS	NS	NS	NS	NS	NS	S	7	0.31	0.38	83.7 (0.080)
P12	S	S	S	S	S	S	NS	S	1	0.61	0.77	84.1 (0.033)
P13	S	S	S	S	S	NS	NS	S	2	0.37	0.45	86.1 (0.072)
P14	S	S	S	NS	S	NS	NS	S	3	0.48	0.56	93.1 (0.034)
P15	S	S	S	NS	NS	NS	NS	S	4	0.45	0.60	91.1 (0.011)
P16	S	S	S	S	S	NS	NS	S	2	0.44	0.57	87.9 (0.021)
P17	S	S	S	S	S	NS	NS	S	2	0.61	0.72	84.0 (0.037)
P18	S	S	S	S	NS	S	NS	S	2	0.55	0.68	94.6 (0.026)
P19	S	S	S	S	S	NS	NS	S	2	0.50	0.66	91.1 (0.014)
P20	S	S	S	NS	NS	NS	NS	S	4	0.53	0.65	87.3 (0.033)
P21	NS	NS	NS	NS	NS	NS	NS	S	7	0.53	0.72	78.5 (0.095)
P22	S	S	NS	NS	NS	NS	NS	S	5	0.35	0.45	85.2 (0.043)
P23	S	S	S	S	NS	NS	NS	S	3	0.72	0.85	84.8 (0.105)
P24	S	S	S	S	S	NS	NS	S	2	0.43	0.54	88.1 (0.023)
P25	S	S	S	S	S	NS	NS	S	2	0.50	0.66	92.4 (0.025)
P26	S	S	S	NS	NS	NS	NS	S	4	0.68	0.88	87.4 (0.031)
P27	S	S	S	S	NS	S	S	S	1	0.39	0.49	74.9 (0.074)

Patient ID	Wavelet scale λ_j (j= 1, 2, …, 8)								Count of NS*	Rudge et al. (2006b) diagnostics		Chase et al. (2004) diagnostics
	1	2	4	8	16	32	64	128		AND·	RAND·	TI (SE) ·
P28	S	S	S	S	NS	S	NS	S	2	0.34	0.44	89.2 (0.027)
P29	**S**	**S**	**S**	**S**	**S**	**S**	**NS**	**S**	**1**	**0.38**	**0.45**	**77.6 (0.083)**
P30	NS	NS	NS	NS	NS	NS	S	S	6	0.63	0.82	92.2 (0.021)
P31	S	S	NS	NS	NS	NS	NS	S	5	0.40	0.51	89.3 (0.030)
P32	S	S	NS	NS	NS	S	NS	S	4	0.38	0.50	89.3 (0.022)
P33	**S**	**S**	**S**	**NS**	**NS**	**S**	**NS**	**S**	**3**	**0.28**	**0.36**	**88.7 (0.020)**
P34	NS	NS	NS	NS	NS	S	NS	S	6	0.43	0.55	86.5 (0.034)
P35	**S**	**S**	**S**	**S**	**S**	**S**	**NS**	**S**	**1**	**0.38**	**0.46**	**85.9 (0.044)**
P36	S	S	NS	NS	NS	NS	NS	S	5	0.52	0.64	86.4 (0.095)
P37	S	S	S	S	S	NS	NS	S	2	0.53	0.59	79.9 (0.093)

Table 4. Wavelet correlation analysis of the 37 ICU patients. S-significant, NS-non-significant WCORR at given scale λ_j (j= 1, 2, …, 8). Bolded patients indicate poor trackers according to the WCORR and "Count NS" values. Patient 30 has NS=6, which is large, but their high AND=0.63, RAND=0.82, and TI=92.2% are indicative of good tracking.

Poor Patient ID	Wavelet Correlations (scale λ_j)								Rudge et al. (2006b) diagnostics		Chase et al. (2004) diagnostics
	λ_1	λ_2	λ_3	λ_4	λ_5	λ_6	λ_7	λ_8	AND	RAND	TI(SE)
P2	0.005	0.005	0.005	0.005	-0.008	-0.186	-0.034	0.649	0.53	0.66	86.2 (0.03)
P4	0.006	0.006	0.005	0.005	0.036	-0.077	0.023	0.724	0.56	0.62	80.1 (0.09)
P7	-0.081	-0.081	-0.081	-0.081	-0.037	-0.066	0.034	0.785	0.33	0.43	84.5 (0.06)
P9	-0.020	-0.021	-0.021	-0.021	0.012	-0.066	0.012	0.793	0.49	0.62	87.3 (0.02)
P10	0.051	0.051	0.051	0.051	0.052	-0.125	-0.040	0.591	0.27	0.34	83.4 (0.04)
P11	-0.001	-0.001	-0.001	-0.001	0.019	0.093	0.002	0.654	0.31	0.38	83.7 (0.08)
P21	0.016	0.017	0.017	0.017	0.026	-0.148	-0.031	0.779	0.53	0.72	78.5 (0.09)
P22	-0.051	-0.051	-0.051	-0.051	-0.027	-0.115	-0.052	0.739	0.35	0.45	85.2 (0.04)
P27	-0.108	-0.108	-0.108	-0.108	-0.088	-0.227	-0.201	0.660	0.39	0.49	74.9 (0.07)
P28	-0.100	-0.101	-0.101	-0.101	-0.090	-0.200	-0.157	0.628	0.34	0.44	89.2 (0.02)
P29	-0.616	-0.616	-0.616	-0.616	-0.582	-0.497	-0.294	0.627	0.38	0.45	77.6 (0.08)
P32	0.032	0.034	0.035	0.035	0.042	-0.139	-0.061	0.732	0.38	0.50	89.3 (0.02)
P33	0.046	0.046	0.046	0.046	0.049	-0.127	-0.012	0.691	0.28	0.36	88.7 (0.02)
P34	-0.019	-0.019	-0.019	-0.019	-0.029	-0.213	-0.131	0.676	0.43	0.55	86.5 (0.03)
P35	-0.172	-0.172	-0.172	-0.172	-0.139	-0.155	-0.166	0.576	0.38	0.46	85.9 (0.04)

Poor Patient ID	Wavelet Correlations (scale λ_j)								Rudge et al. (2006b) diagnostics		Chase et al. (2004) diagnostics
	λ_1	λ_2	λ_3	λ_4	λ_5	λ_6	λ_7	λ_8	AND	RAND	TI(SE)
Poor Median	**-0.019**	**-0.019**	**-0.019**	**-0.019**	**-0.008**	**-0.139**	**0.040**	**0.676**	**0.390**	**0.500**	**85.2**
(95%CI)	(-0.092 ,0.013)	(-0.094, 0.013)	(-0.094, 0.013)	(-094, 0.013)	(-0.069, 0.032)	(-0.195, -0.091)	(-0.147, -0.003)	(0.636, 0.736)	(0.343, 0.515)	(0.444, 0.620)	(81.333, 87.001)
Good Pt ID											
	λ_1	λ_2	λ_3	λ_4	λ_5	λ_6	λ_7	λ_8	AND	RAND	TI(SE)
P1	-0.112	-0.134	-0.134	-0.134	-0.107	-0.101	-0.115	0.662	0.51	0.62	87.0 (0.04)
P3	-0.503	-0.504	-0.504	-0.504	-0.439	-0.221	-0.242	0.675	0.70	0.83	88.8 (0.02)
P5	0.074	0.073	0.073	0.073	0.078	-0.118	-0.004	0.707	0.60	0.80	91.1 (0.02)
P6	-0.202	-0.202	-0.202	-0.202	-0.155	-0.090	-0.167	0.666	0.70	0.84	87.0 (0.01)
P8	0.208	0.206	0.206	0.206	0.211	-0.029	0.081	0.783	0.45	0.59	87.4 (0.03)
P12	-0.359	-0.359	-0.359	-0.359	-0.316	-0.264	-0.179	0.645	0.61	0.77	84.1 (0.03)
P13	0.258	0.258	0.258	0.258	0.257	0.007	0.107	0.768	0.37	0.45	86.1 (0.07)
P14	-0.378	-0.379	-0.380	-0.380	-0.301	-0.074	-0.003	0.785	0.48	0.56	93.1 (0.03)
P15	0.086	0.092	0.093	0.093	0.084	-0.149	-0.039	-0.697	0.45	0.60	91.1 (0.01)
P16	0.168	0.169	0.169	0.169	0.173	-0.067	0.035	0.758	0.44	0.57	87.9 (0.02)
P17	-0.122	-0.122	-0.122	-0.122	-0.069	0.131	-0.161	0.604	0.61	0.72	84.0 (0.04)
P18	-0.134	-0.134	-0.134	-0.134	-0.130	-0.243	-0.211	0.628	0.55	0.68	94.6 (0.03)
P19	0.272	0.273	0.273	0.273	0.277	0.066	0.195	0.726	0.50	0.66	91.1 (0.01)
P20	0.057	0.057	0.057	0.057	0.057	-0.106	-0.051	0.613	0.53	0.65	87.3 (0.03)
P23	0.149	0.149	0.149	0.149	0.164	-0.019	0.099	0.729	0.72	0.85	84.8 (0.11)
P24	0.231	0.232	0.232	0.232	0.232	-0.007	0.157	0.793	0.43	0.54	88.1 (0.02)
P25	-0.211	-0.214	-0.214	-0.214	-0.188	-0.218	-0.185	0.561	0.50	0.66	92.4 (0.03)
P26	-0.122	-0.123	-0.123	-0.123	-0.073	0.057	-0.011	0.579	0.68	0.88	87.4 (0.03)
P30	-0.045	-0.044	-0.044	-0.044	-0.049	-0.208	-0.150	0.647	0.63	0.82	92.2 (0.02)
P31	0.040	0.040	0.040	0.040	0.047	-0.108	0.024	0.701	0.40	0.51	89.3 (0.03)
P36	0.081	0.081	0.081	0.081	0.055	-0.172	-0.104	0.677	0.52	0.64	86.4 (0.10)
P37	0.272	0.273	0.273	0.273	0.273	0.066	0.195	0.726	0.53	0.59	79.9 (0.09)
Good Median	**0.049**	**0.049**	**0.0485**	**0.0485**	**0.051**	**-0.096**	**-0.025**	**0.676**	**0.525**	**0.655**	**87.7**
(95%CI)	(-0.122, 0.149)	(-0.134, 0.149)	(-0.134, 0.149)	(-0.134, 0.149)	(-0.108, 0.164)	(-0.149, -0.019)	(0.015, 0.036)	(0.065, -0.726)	(0.479, 0.610)	(0.590, 0.771)	(86.984, 91.100)
Kruskal Wallis P value (Poor vs Good)	P=0.32	0.32	0.32	0.32	0.30	0.16	0.40	0.84	**0.004**	**0.003**	**0.005**

Table 5. Kruskal Wallis test on the wavelet correlation, and on Rudge et al.'s (2006b) and Chase et al.'s (2004) diagnostics – testing for differences between the DWT based poor versus good tracker groups.

Variable	Poor group median	Good group median	**k-w#: P value**
Modulus WCORR at λ_1	0.046	0.159	**0.001**
"Count of NS"	**5**	**2**	**0.001**
AND	0.39	0.53	**0.004**
RAND	0.50	0.66	**0.003**
TI	85.20	87.7	**0.005**

Table 6. K-W tests of all wavelet and other diagnostics by wavelet based tracking group.

3.3 Using the Wavelet Cross-Correlation (WCCORR)

We can investigate possible lead or lag relationships between a given patient's modelled (simulated) versus observed (recorded) A-S profile by examining a plot of its MODWT based wavelet cross-correlation (WCCORR), according to Equation (8). Figure 8 shows this WCCORR plot for Patient 3 (P3: a good tracker) and Patient 4 (P4: a poor tracker). For Patient 3 the $\tilde{\rho}$ values are negative and statistically significant for all scales except 8 (Table 7), and there is also a large positive peak at a lag of 120 minutes for the first six wavelet scales λ_j (j=1, 2, …, 6) (Figure 8). At scale λ_7 ; a large positive peak occurs at 112 minutes for Patient 3; and at scale λ_8 at a lag of 33 minutes. We conclude that at scale λ_7 there is a period of 170 minutes (see Figure 15) for Patient 3. Likewise an examination of Figure 8 shows an inverse shaped profile of peaks to troughs for Patient 4 (a poor tracker), with generally non-significant positive $\tilde{\rho}$ values compared to Patient 3 (see also Table 7). It is noteworthy from Figure 8 that generally patients who are good trackers show a common type of WCCORR signature or pattern, with WCCORR being significant at zero lag (for all scales) and their 95% CI do not include zero (see Table 7), not so for the poor trackers.

Patient 3 (good tracker)

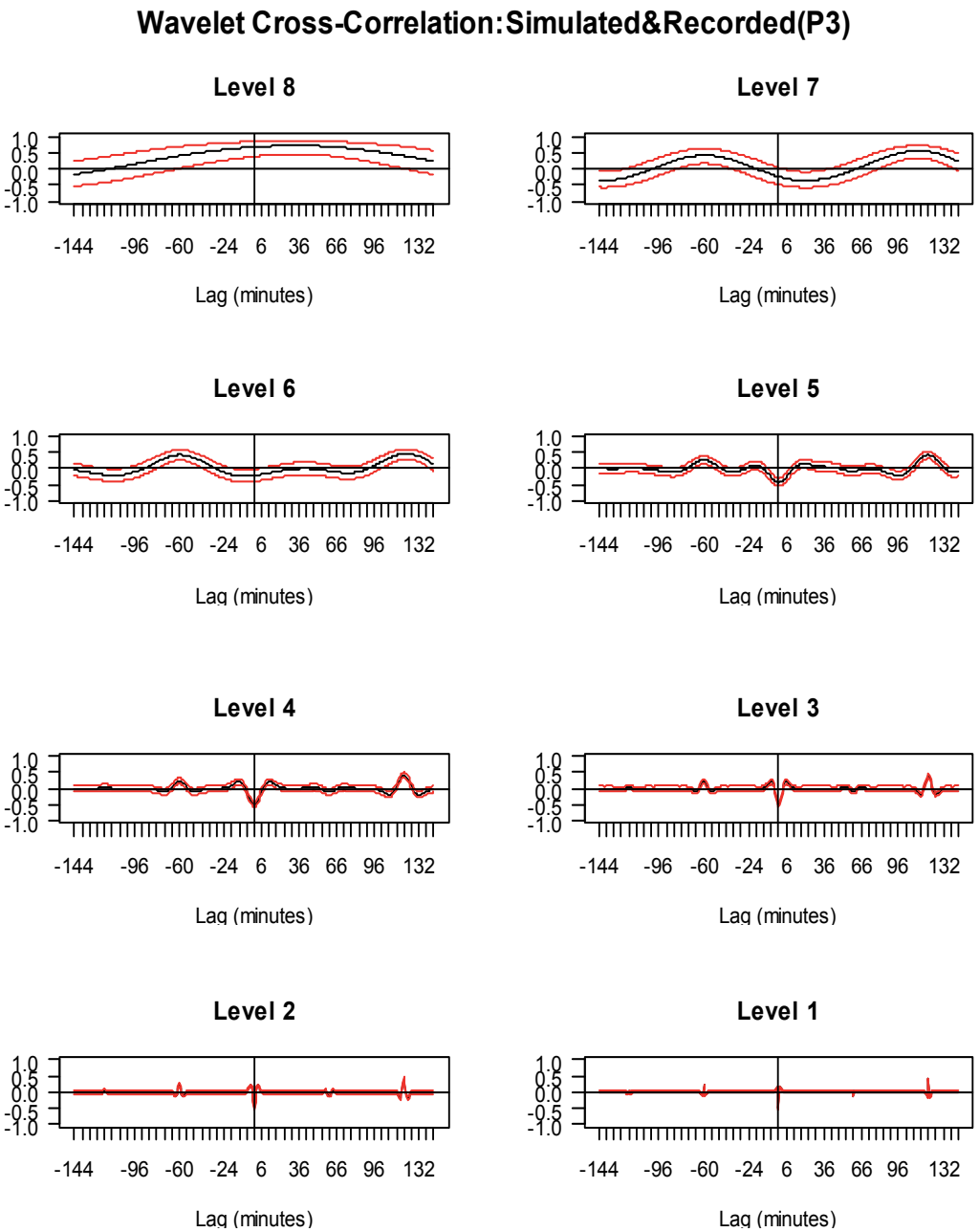
Scale	Level 1	Level 2	Level 3	Level 4	Level 5	Level 6	Level 7	Level 8
L. CI	-0.528	-0.539	-0.554	-0.573	-0.544	-0.398	-0.485	0.382
$\tilde{\rho}$	-0.503^{+}	-0.504^{+}	-0.504^{+}	-0.504^{+}	-0.439^{+}	-0.221^{+}	-0.242^{+}	0.675^{+}
U. CI^{++}	-0.477	-0.467	-0.451	-0.428	-0.321	-0.027	0.036	0.845

Patient 4 (poor tracker)

Scale	Level 1	Level 2	Level 3	Level 4	Level 5	Level 6	Level 7	Level 8
L. CI	-0.034	-0.050	-0.074	-0.107	-0.123	-0.297	-0.299	0.401
$\tilde{\rho}$	0.006	0.006	0.005	0.006	0.036	-0.077	0.023	0.724
U. CI^{++}	0.045	0.061	0.085	0.117	0.194	0.151	0.339	0.886

Table 7. The signature of WCORR for the Patient 3 and Patient 4 at zero lag based WCCORR. The positive subscript (+) indicates significant WCCORR $\tilde{\rho}$ values.

Recall that Table 7 gives the values (*signature* over wavelet scales) of $\tilde{\rho}$ and their associated 95% confidence limits (L.CI, U.CI) for Patient 3 and Patient 4 who are deemed, to be a good and poor tracker, respectively (as depicted by Figure 8). The main results of the work in this chapter are summarised in Table 8 and discussed in detail in the Conclusion (section 4).



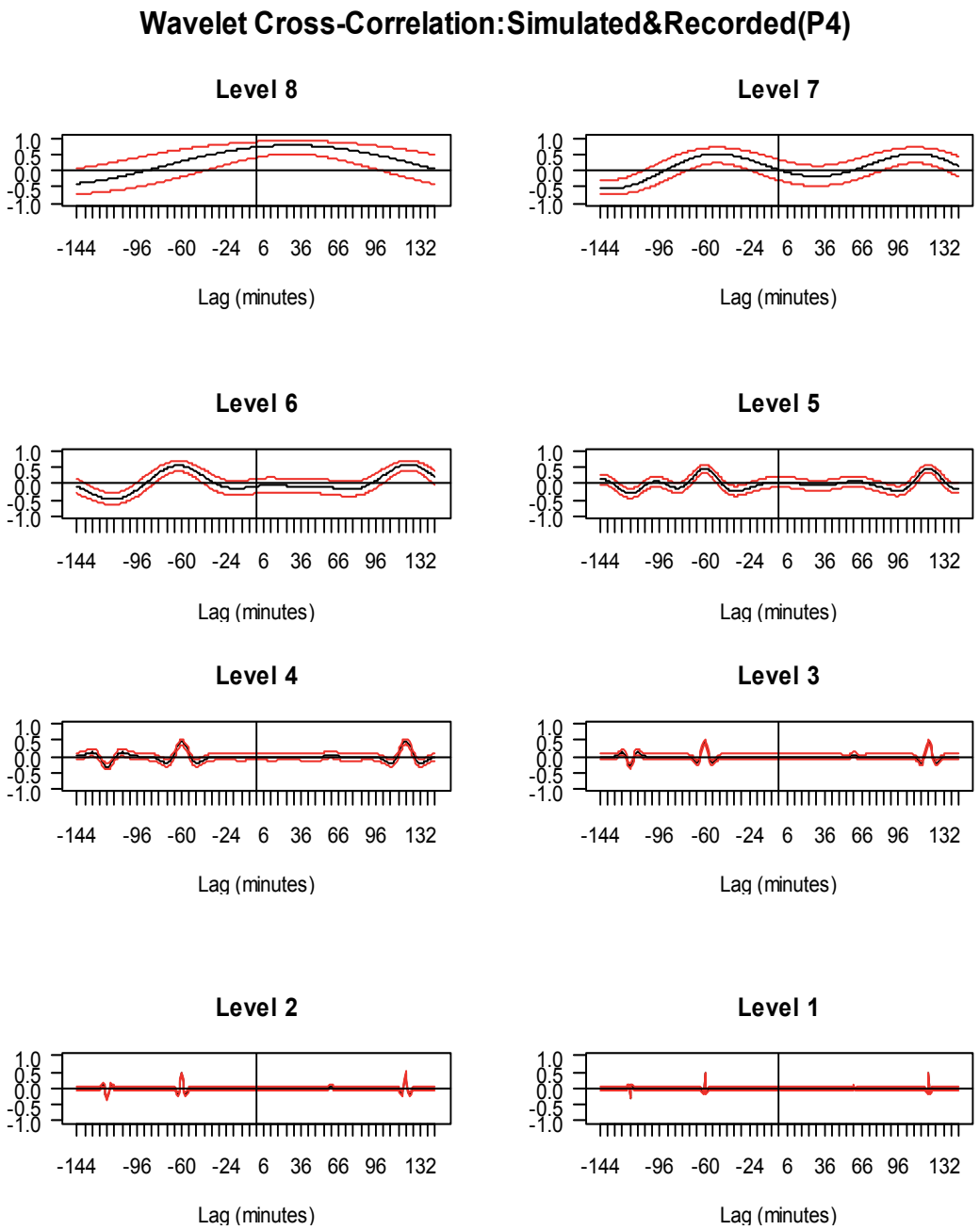
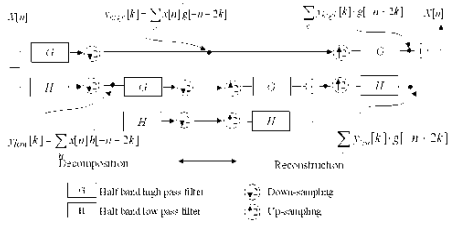


Fig. 8. MODWT estimated wavelet Cross-Correlation between the simulated and recorded infusion series for lags up to ±144 minutes for Patient 3 (a good tracker) and Patient 4 (a poor tracker) with approximate 95% CI (red broken lines).

Table 8. Overview of DWT and WCORR results.

Authors	Equations or Model used	Mathematical methods used	Aims of the study and the performance indicators derived	Patient numbers with poor fit to simulated profiles (poor trackers) and the summary profile of patients with poor versus good fit
Kang et al. (this chapter)	See the schema of our approach of Chapter 2 below: **DWT Procedure (second stage)**  **DWT analysis and synthesis equations** $\boldsymbol{X}=[\boldsymbol{X}_1, \boldsymbol{X}_2,\ldots, \boldsymbol{X}_N]$, $N=2^J$, DWT analysis equation $\boldsymbol{W}=\mathcal{W}\boldsymbol{X}$ $\boldsymbol{W}=$ discrete wavelet coefficients, $\mathcal{W}=N$ x N orthonormal matrix $\boldsymbol{W}=\mathcal{W}\boldsymbol{X}, \boldsymbol{W}=\left[W_1 W_2,\cdots,W_j,V_J\right]^T, \mathcal{W}=\left[\mathcal{W}_1^1,\mathcal{W}_2^1,\ldots,\mathcal{W}_J^1,\mathcal{V}_J^1\right]$ $\boldsymbol{X}=\mathcal{W}^T\boldsymbol{W}=\left[\mathcal{W}_1^1,\mathcal{W}_2^1,\ldots,\mathcal{W}_J^1,\mathcal{V}_J^1\right]^T\left[\boldsymbol{W}_1,\boldsymbol{W}_2,\ldots,\boldsymbol{W}_J,\boldsymbol{V}_J\right]^T$ $=\sum_{j=1}^{J}\mathcal{W}_j^T\boldsymbol{W}_j+\mathcal{V}_J^T\boldsymbol{V}_J \quad \Rightarrow$ DWT synthesis equation **DWT – MRA** $\boldsymbol{X}=\sum_{j=1}^{J}\mathcal{W}_j^T\boldsymbol{W}_j+\mathcal{V}_J^T\boldsymbol{V}_J=\sum_{j=1}^{J}D_j+S_J \Rightarrow$ Additive decomposition(=MRA) $D_j=\mathcal{W}_j^T\boldsymbol{W}_j$: Portion of synthesis due to scale λ_j, jth 'detail' $S_J=\mathcal{V}_J^T\boldsymbol{V}$:'smooth' of Jth order	Discrete Wavelet Transform (DWT) (Percival & Walden, 2000) Maximal Overlap Discrete Wavelet (MODWT) (Percival & Walden, 2000) Multiresolution analysis (MRA) (Percival & Walden, 2000) DWT-MRA, MODWT-MRA	Develop a wavelet correlation (WCORR) and wavelet cross-correlation (WCCORR) approach for assessing the validity of the deterministic dynamic models against the empirical agitation-sedation data of 37 ICU patients. Provide graphical assessment tools and wavelet based numerical metrics of comparability between the model and the data via the discrete wavelet transform (DWT), partial DWT (PDWT), maximal overlap DWT (MODWT) and via Multiresolution analysis (MRA). Investigate the lag/lead relationship between the simulated and recorded infusion series on a scale by scale basis via wavelet cross-correlation (WCCORR) methods. Develop performance measures as follows: 1. Modulus of the wavelet correlation at wavelet scale 1, λ_1. 2. Count number (out of 8) of the number of non-significant wavelet correlations at scales λ_j j=1,2,…,8). Test poor versus good fit or tracking groups via the: Kruskal Wallis test on the WCORR measures per patient and on the indicators in 1 and 2 above to test for group differences (for the poor versus good trackers group).	The set of 15 poor$^{\kappa}$ trackers are: **P2, P4, P7, P9, P10, P11, P21, P22, P27, P28, P29, P32, P33, P34, P35** With a median modulus of wavelet correlation at wavelet scale λ_1 of **0.046,** and 95% CI (0.017,0.093), significantly lower than the good trackers (P = 0.001 k-w test) and a median number of non-significant wavelet correlations at scales λ_j of **5**.0 (2.37, 7.0). The set of remaining good trackers have: a significantly lower median number of non-significant wavelet correlations at scales λ_j of **2.0,** and 95% CI (2.0, 4.0); a median of the modulus of wavelet correlation at wavelet scale λ_1 of **0.159**, and 95% CI (0.111, 0.232), significantly higher than the poor trackers. $^{\kappa}$ Poor trackers have a count of non-significant WCORR values ≥ 2 and a non-significant WCORR value at λ_1 (level 1) or a "Count of NS" ≤ 3 ; or have a significant *negative* WCORR value at scale λ_1 and a significant negative WCORR value at λ_6. Of these 15 poor performers, 11 are also considered to track poorly by either or both of the performance indicators of Rudge et al. (2006b) and of Chase et al. (2004).

4. Conclusion

DWT and MODWT-MRA decomposition and reconstruction are shown to provide clear and consistent, in regard to good or poor performance, "signatures" of, and values for the wavelet correlations and cross-correlations (at all dyadic scales) between an ICU patient's bivariate time series, namely their simulated and their recorded A-S infusion profiles over time under sedation. A suite of wavelet techniques are advocated, based on the DWT, and applied successfully to assess whether an ICU patient's mathematically simulated agitation-sedation (A-S) status reflects their true dynamic profile.

The wavelet correlation profiles of the good trackers are shown to be invariably significant at all scales (except at 32 and 64). Patients who exhibit poor tracking exhibit WCORR profiles which are invariably non significant at almost all wavelet scales, *particularly* at λ_1. Moreover, *cross*-correlation (WCCORR) signatures also show a common pattern for the good trackers, which is distinctly different to the pattern associated with poor tracking. It is also shown that the lead or lag relationship between a patient's simulated and recorded agitation-sedation infusion series can be investigated on a scale by scale basis via the implementation of the MODWT-MRA. The MRA is shown to successfully indicate local features of interest in the simulated and recorded series, with the smooth MRA series offering a good visual summary of the overall long-term trends in a patient's A-S status.

Fifteen poor trackers are identified by the DWT based diagnostics derived in this chapter. Specifically it is found that the modulus of WCORR at λ_1, $|\lambda_1|$, is invariably higher (and significantly so) for the good compared to the poor trackers. The profile of ("Count of NS") across all patients is higher for the poor trackers; and the values of RAND, AND and $|\lambda_j|$ invariably higher for the good trackers. Specifically the median absolute value of the wavelet correlation at λ_1, and median value of AND, RAND and TI are highly significantly lower for the poor trackers (15 of 37 patients). The median of the number of non significant wavelet correlations (the "Count of NS" variable) is 5.0 for the poor trackers, significantly higher than the median of 2.0 for the good tracking group (P = 0.001). It is noteworthy that 11 of the 15 DWT based poor trackers are also found to be poor trackers by either or both of Rudge et al.'s (2006b) and Chase et al.'s (2004) (non wavelet based) performance indicators, showing significant agreement between the wavelet DWT and the earlier kernel based methods - with a kappa test of agreement between the WCORR criterion for poor tracking and Chase et al (2004) of 0.2127 (P=0.01); and with Rudge et al. (2006b) of 0.5856 (P=0.001).

Other recent work by Kang et al. (*In Prep)* has used Bayesian densities and wavelet shrinkage methods to create a novel wavelet probability band (WPB). A 90% value for the WPB implies that for at least 90% of the time, the estimated mean value of the patient's recorded infusion rate lies within the band. A 90% WPB was constructed by Kang et al. *(In Prep*) for each of the 37 patient profiles, and the time and duration of any deviations from the wavelet probability band recorded for each patient. Likewise wavelet analogues of the AND and RAND diagnostics of Rudge et al. (2006a; 2006b), namely the average normalized wavelet density (ANWD) and the relative average normalized wavelet density (RANWD) have been derived by Kang et al. *(In Prep*); as has a Wavelet Time Coverage Index (WTCI) – all by using Bayesian wavelet thresholding. The resultant WTCI and 90% WPB provide very strong support for the DWT wavelet diagnostics derived in this chapter. Indeed of the 15 DWT based poor trackers identified in this chapter, 10 also

exhibit a low WPB (WPB 90% < 70%) and a low wavelet density based RANWD measure (RANWD ≤ 0.5) and are likewise deemed to be poor trackers according to these recent wavelet regression methods of Kang et al. (*In Prep)*. Statistically speaking the wavelet probability band and wavelet density diagnostics mirror the poor versus good classification of this chapter's WCORR DWT based diagnostics (kappa = 0.7701, P = 0.0001). The main reason for the reduced total time within the WPB (and for the non significant WCORRs) for this minority group of 10 (or 15) poor trackers (of a total of 37 patients), is the consistently poor performance of the DE model throughout their *total* length of the A-S simulation.

The work in this chapter provides a suite of new wavelet based diagnostics by which to achieve statistical model validation of the A-S models. The DWT, wavelet correlation and cross-correlation measures derived in this chapter are proved to be valid for assessing control, and mirror earlier validation measures; as do the more recent wavelet regression diagnostics (namely WTCI, ANWD, RANWD, and WPB 90%) of Kang et al. *(In Prep*). Wavelets are shown to visually and quantitatively discriminate patients for whom the A-S model captures their fundamental A-S dynamics, versus those, for whom this is not so. Wavelet WCORR and WCCORR signatures thus form a possibly alternative and appropriate feedback mechanism for comparison of improved sedation administration controllers and gain. The wavelets based visual tools and quantitative measures thus contribute to the task of improving and of refining A-S models. This chapter thus demonstrates that wavelets provide a new diagnostic tool by which to assess the agitation-sedation of ICU patients, and show that it is possible to evaluate A-S models via wavelet diagnostics for accurate evaluation of A-S management, where the latter represents a trade-off between the benefits of low patient agitation versus the cost of high infusion rates and increased total dose requirements (Rudge et al., 2003; 2005; 2006a; 2006b). Wavelets are thus suitable for clinical implementation in ICU agitation and sedation control.

Overall the various wavelet diagnostics strongly agree and confirm the value of A-S modelling in ICU. Wavelet DWT analysis also demonstrates that the models of the A-S studies of Chase et al., (2004), Rudge et al., (2005; 2006a; 2000b) and of Lee et al., (2005), are suitable for developing more advanced optimal infusion controllers. These offer significant clinical potential of improved agitation management and reduced length of stay in critical care. The use of quantitative modelling to enhance understanding of the A-S system and the provision of an A-S simulation platform are critical tools in this area of patient critical care.

The DWT approach gives robust performance metrics of A-S control and also yields excellent visual assessment tools - generalisable to any study which investigates the similarity or closeness between any bivariate time series of, say, a large number of units (patients, households etc) and of disparate lengths and possibly of extremely long length.

5. References

Abramovich, F., & Benjamini, Y. (1995). Thresholding of wavelet coefficients as multiple hypotheses testing procedure. *Lecture Notes in Statistics, 103*, 5-14.

Allan, D. W. (1966). Statistics of atomic frequency standards. *Proceedings of the IEEE, 54*(2), 221-230.

Barber, S., Nason, G. P., & Silverman, B. W. (2002). Posterior probability intervals for wavelet thresholding. *Journal of the Royal Statistical Society. Series B (Statistical Methodology), 64*(2), 189-205.

Chang, S. G., Yu, B., & Vetterli, M. (2000a). Spatially adaptive wavelet thresholding based on context modeling for image denoising. *IEEE Transactions on Image Processing, 9*(9), 1522- 1531.

Chang, S. G., Yu, B., & Vetterli, M. (2000b). Adaptive wavelet thresholding for image denoising and compression. *IEEE Transactions on Image Processing, 9*(9), 1532-1546.

Chase, J. G., Rudge, A. D., Shaw, G. M., Wake, G. C., Lee, D., Hudson, I. L., et al. (2004). Modeling and control of the agitation-sedation cycle for critical care patients. *Medical Engineering and Physics, 26*(6), 459-471.

Chui, C. K. (1992). *An introduction to wavelets*: Academic Press.

Daubechies, I. (1992). *Ten lectures on wavelets*. Philadelphia: Society for Industrial and Applied Mathematics.

Donoho, D. L. (1995). De-noising by soft-thresholding. *IEEE Transactions on Information Theory, 41*(3), 613-627.

Donoho, D. L., & Johnstone, I. M. (1994). Ideal spatial adaptation via wavelet shrinkage. *Biometrika, 81*, 425-455.

Dépué M (2003). Continuous variables. In: Jolliffe IT, Stephenson DB (eds) Forecast verification: a practitioner's guide in atmospheric science. John Wiley and Sons, Chichester, pp 97–120

Dose V, Menzel A (2004). Bayesian analysis of climate change impacts in phenology. Glob Fraser, G. L., & Riker, R. R. (2001b). Monitoring sedation, agitation, analgesia, and delirium in critically ill adult patients. *Crit Care Clin, 17*(4), 967-987.

Gencay, R., Selcuk, F., & Whitcher, B. (2001). *An introduction to wavelets and other filtering methods in finance and economics*: Academic Press.

Goupillaud, P., Grossmann, A., & Morlet, J. (1984). Cycle-octave and related transforms in seismic signal analysis. *Geoexploration, 23*(1), 85-102.

Haar, A. (1910). Zur Theorie der orthogonalen Funktionensysteme. *Mathematische Annalen, 69*(3), 331-371.

Hudson, I.L. (2011). Meta analysis, in *Encyclopedia of Climate and Weather*, S. H. Schneider, M. Mastrandrea and T. L. Root, (eds), pp 273-287, Oxford University Press.

Hudson, I.L. (2010). Interdisciplinary approaches: towards new statistical methods for phenological studies, *Climatic Change*, 100(1), 143-171.

Hudson, I. L., Keatley, M. R., & Roberts, A. M. I. (2005). Statistical Methods in Phenological Research. Proceedings of the Statistical Solutions to Modern Problems. Proceedings of the 20th International Workshop on Statistical Modelling, eds. AR Francis, KM Matawie, A Oshlack & GK Smyth, ISBN 1 74108 101 7, Sydney, Australia, 10-15 July 2005, pp. 259-270.

Hudson, I.L., & Keatley, M. R. (2010). *Phenological Research: Methods for Environmental and Climate Change Analysis*: Springer, Dordrecht, Netherlands.

Hudson, I.L., Kang, I., & Keatley, M. R. (2010a). Wavelet Analysis of Flowering and Climatic Niche Identification. In I. L. Hudson & M. R. Keatley (Eds.), *Phenological Research: Methods for Environmental and Climate Change Analysis*, pp 361-391, Springer, Dordrecht, Netherlands.

Hudson, I.L., Keatley, M. R., & Kang, I. (2010b). Wavelet characterisation of eucalypt flowering and the influence of climate. *Environmental and Ecological Statistics.* (Doi 10.1007/s10651-010-0149-5).

Hudson I.L., Keatley, M.R., & Kang, I. (2011a) Wavelets and clustering: methods to assess synchronization. pp. 59-70. In: del Valle M, Muñoz R, Gutiérrez JM (eds) *Wavelets: Classification, Theory and Applications*. Nova Science Publishers.

Hudson, I.L., Keatley, M.R., & Kang, I. (2011b). Wavelet Signatures of Climate and Flowering: Identification of Species Groupings In: Olkkonen H (ed) *Discrete Wavelet Transforms / Book 2*. InTech, Vienna, Austria, http://www.intechweb.org ISBN 979-953-307-241-7. (pp 00-30) *(in Press)*

Jaarsma, A. S., Knoester, H., van Rooyen, F., & Bos, A. P. (2001). Biphasic positive airway pressure ventilation (pev+) in children. *Crit Care Clin, 5*(3), 174-177.

Kang, I., Hudson, I. L., & Keatley, M. R. (2004). *Wavelet analysis in phenological research-the study of four flowering eucalypt species in relation to climate.* Paper presented at the International Biometrics Conference (IBC 04), July, Cairns, Australia.

Kang, I., Hudson, I. L., Rudge, A. D., & Chase, J. G. (2005). *Wavelet signature of Agitation-Sedation profiles of ICU patients.* In: Francis AR, Matawie KM, Oshlack A, Smyth GK (eds) Statistical Solutions to Modern Problems. 20th International Workshop on Statistical Modelling, Sydney, July 10–15, University of Western Sydney (Penrith), pp 293-296, ISBN 1 74108 101 7.

Kang, I., Hudson, I. L, Rudge, A, Chase, G. Density estimation and wavelet thresholding via Bayesian methods: application to ICU data, *Computer Methods and Programs in Biomedicine (In prep).*

Kress, J. P., Pohlman, A. S., & Hall, J. B. (2002). Sedation and analgesia in the intensive care unit. *Am J Respir Crit Care Med, 166*(8), 1024-1028.

Kumar, P. (1994). Role of coherent structures in the stochastic-dynamic variability of precipitation. *Journal of Geophysical Research-Atmospheres, 101(D21)*(26), 393-404.

Kumar, P., & Foufoula-Georgiou, E. (1993). A multicomponent decomposition of spatial rainfall fields. 1. Segregation of large- and small-scale features using wavelet transforms. *Water Resources Research, 29*(8), 2515-2532.

Lee, D. S., Rudge, A. D., Chase, J. G., & Shaw, G. M. (2005). A new model validation tool using kernel regression and density estimation. *Computer Methods and Programs in Biomedicine, 80*, 75-87.

Mallat, S. (1989). A theory for multiresolution signal decomposition: the wavelet representation. *IEEE Transactions on Pattern analysis and Machine Intelligence, 11*, 674-693.

McCoy, E. J., Percival, D. B., & Walden, A. T. (1995). *Spectrum estimation via wavelet thresholding of multitaper estimators*. Imperial College, London.

Meyer, F. G. (2003). Wavelet-based estimation of a semiparametric generalized linear model of fMRI time-series. *IEEE Transactions on Medical Imaging, 22*(3), 315-322.

Minto, C. F., Schnider, T. W., Short, T. G., Gregg, K. M., Gentilini, A., & Shafer, S. L. (2000). Response surface model for anesthetic drug interactions. *Anesthesiology, 92*(6), 1603-1616.

Morlet, J. (1983). Sampling theory and wave propagation. *Issues in Acoustic Signal/Image Processing and Recognition, 1*, 233-261.

Ogden, R. T. (1997). *Essential Wavelets for Statistical Applications and Data Analysis*. Boston: Birkhäuser.

Percival, D. B., & Guttorp, P. (1994). Long-memory processes, the Allan variance and wavelets, in *Wavelets and Geophysics*, edited by E. Foufoula-Georgiou and P. Kumar, Vol. 4, *Wavelet Analysis and Its Applications*, 325–344, Academic Press, San Diego, California.

Percival, D. B., & Mofjeld, H. O. (1997). Analysis of subtidal coastal sea level fluctuations using wavelets. *Journal of the American Statistical Association, 92*(439), 868-880.

Percival, D. B., & Walden, A. T. (2000). *Wavelet Methods for Time Series Analysis*. Cambridge, England: Cambridge University Press.

Ramsay, M. A., Savege, T. M., Simpson, B. R., & Goodwin, R. (1974). Controlled sedation with alphaxalone-alphadolone. *Br Med J, 2*(920), 656-659.

Riker, R. R., Picard, J. T., & Fraser, G. L. (1999). Prospective evaluation of the Sedation-Agitation Scale for adult critically ill patients. *Critical care medicine, 27*(7), 1325-1329.

Rudge, A. D., Chase, J. G., Shaw, G. M., & Lee, D. (2006a). Physiological modelling of agitation-sedation dynamics including endogenous agitation reduction. *Medical Engineering & Physics, 28*(7), 629-638.

Rudge, A. D., Chase, J. G., Shaw, G. M., & Lee, D. (2006b). Physiological modelling of agitation-sedation dynamics. *Medical Engineering and Physics, 28*(1), 49-59.

Rudge, A. D., Chase, J. G., Shaw, G. M., Lee, D., Wake, G. C., Hudson, I. L., et al. (2005). Impact of control on agitation-sedation dynamics. *Control Engineering Practice, 13*(9), 1139-1149.

Sessler, C. N., Gosnell, M. S., Grap, M. J., Brophy, G. M., O'neal, P. V., Keane, K. A., et al. (2002). The Richmond agitation-sedation scale: validity and reliability in adult intensive care unit patients. *Am J Respir Crit Care Med, 166*(10), 1338-1344.

Swelden, W. (1996). Wavelets: what next? *Proceedings of the IEEE, 84*(4), 680-685.

Vidakovic, B. (1999). *Statistical modelling by wavelets*. New York: John Wiley & Sons.

Wand, M. P., & Jones, M. C. (1995). *Kernel smoothing*: Chapman & Hall/CRC.

Whitcher, B., Guttorp, P., & Percival, D. B. (2000). Wavelet Analysis of Covariance with Application to Atmospheric Time Series. *Journal of Geophysical Research, 105*(D11), 14,941-962.

17

Application of Discrete Wavelet Transform for Differential Protection of Power Transformers

Mario Orlando Oliveira[1,2] and Arturo Suman Bretas[1]
[1]*Electrical Engineering Department, Federal University of Rio Grande do Sul (UFRGS)*
[2]*Energy Studies Center (CEED), National University of Misiones (UNaM)*
[1]*Brazil*
[2]*Argentina*

1. Introduction

Power transformers (PT) play an extremely important role on the reliability and energy supply continuity of Electric Power Systems (EPS). The inherent characteristic of power transformers introduce a number of unique problems that are not present in the protection of transmission lines, generators, motors or other power system apparatus (Horowitz & Phadke, 2008). When PT internal faults occur, immediate disconnection of the faulted transformer is necessary to avoid extensive damage and/or preserve power system stability and power quality (Harlow, 1999). Currently, percentage differential protection is a common practice for power transformer protection. However, nonlinearities in the transformer core, the current transformer (CT) core or in both, cause a substantial differential current to flow when there is no fault. Thus, these false differentials currents can cause a percentage differential relay miss trip. To mitigate some of these problems the differential relays are equipped with harmonic restraint, where the magnitudes of the second and fifth harmonic component are compared with the fundamental frequency component to discriminate internal faults from magnetizing inrush currents and transformer over-excitation, respectively (Anderson, 1999). However, performance limitations are still reported even for such phenomena. In order to overcome such limitation, a significant number of relaying formulations have been proposed (Abed & Mohammed, 2007; Eissa, 2005; Faiz & Lotfi-Fard, 2006; Mao & Aggarwal, 2000; Megahed et al., 2008; Morate & Nicoletti, 1999; Ngaopitakkul & Kunakorn, 2006; Saleh & Rahman, 2005. Thomas & Ozgönenel, 2007; Wang & Butler, 2001; Wiszniewski & Kasztenny, 1995; Zaman et al., 1996). These formulations are based on finite elements, artificial neural networks, fuzzy systems, dynamical principal components analysis, wavelet transforms (WTs) and hybrid systems. However, all mentioned relaying formulations have hard to design parameters, which make real life construction difficult.

In detecting faults in EPS and, specifically PT, frequency analysis is required so that the transient signal components can be isolated. This process helps to identify particular

phenomena that generated the transient signals. It should be noted that the waveforms associated with electromagnetic transients are typically non-periodic in nature, containing both high-frequency oscillations as short duration pulses superimposed on low frequency signals. Still, need to know the fault occurrence instant encourages the application of techniques with precise time and frequency resolution.
In this chapter, a novel percentage differential relaying algorithm for three-phase power transformers protection based on Discrete Wavelet Transforms (DWT) is presented. The proposed algorithm's formulation uses logical decision criteria based on wavelets coefficient spectral energy variation to identify and discriminate correctly external faults, inrush currents and incipient internal transformer faults. In order to analyze the proposed algorithms efficiency, it was built in MATLAB® platform (Matlab, 2010) and tested with simulated fault cases under BPA's ATP/EMTP software (ATP/EMTP, 2002).

2. Wavelet Transform (WT)

The Wavelet Transform (WT) theory is based on signal analysis using varying scales in the time and frequency domain. Formalization was carried out in the 80s, based on the generalization of familiar concepts. The wavelet term was introduced by French geophysicist Jean Morlet. The seismic data analyzed by Morlet exhibit frequency component that changed rapidly over time, for which the Fourier Transform was not appropriate as an analysis tool. Thus, with the help of theoretical physicist Croatian Alex Grossmann, Morlet introduced a new transform which allows the high-frequency events identification with a better temporal resolution (Polikar, 1999).
Faulted EPS signals are associated with fast electromagnetic transients, are typically non-periodic and with high-frequency oscillations. This characteristic present a problem for traditional Fourier analysis because its assumes a periodic signal and a wide-band signal requires more dense sampling and longer time periods to maintain good resolution in the low frequencies (Robertson et al., 1996). Thus WT is a powerful tool in the power system transient phenomena analysis. It has the ability to extract information from the transient signals simultaneously in both time and frequency domains and has replaced the Fourier analysis in many applications (Phadke & Thorp, 2009).

2.1 Continuous Wavelet Transform (CWT)

The Short-Time Fourier Transform (STFT) of the continuous signal $x(t)$, can be seen as the Fourier Transform (FT) of the signal with windowed $x(t).g(t - \tau)$ or also as a signal decomposition $x(t)$ into basis functions $g(t - \tau).e^{-jwt}$. The function based term refers to a complete set of functions that, when combined the sum with specific weight can be used to construct the signal (Bentley & McDonnell, 1994).
In the FT case the base function are complex sinusoid e^{-jwt} with a windows centred on τ time. The WT is described in terms of its basic functions, called wavelet or mother wavelet, and variable frequency w is replaced by an ever-escalating variable factor a (which represents the swelling) and generally to variable displacement in time τ is represented by b.
The main characteristic of the WT is that it uses a variable window to scan the frequency spectrum, increasing the temporal resolution of the analysis. The wavelets are represented by:

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

In the equation (1), the constant $1/\sqrt{a}$ is used to normalize the energy and ensure that the energy of $\psi_{a,b}(t)$ is independent of the dilation level (Simpson, 1993). The wavelet is derived from operations such as dilating and translating the mother wavelet, ψ, which must satisfy the admissibility criterion given by (Daubechies, 1990):

$$C_{\psi} = \int_{-\infty}^{+\infty} \frac{\left|\hat{\psi}(y)\right|^2}{|y|} dy \leq \infty \tag{2}$$

where $\hat{\psi}(y)$ is the FT of $\psi(t)$. This means that if $\hat{\psi}$ is a continuous function, then C_{ψ} is finite only if $\psi(0) = 0$, ie (Daubechies, 1990):

$$\int_{-\infty}^{+\infty} \psi(t)\, dt = 0 \tag{3}$$

Thus, it is evident that WT has a zero rating, property that increases the degrees of freedom, allowing the introduction of the dilation parameter of the window (Sarkar & Su, 1998).
The Continuous Wavelet Transform (CWT) of the continuous signal $x(t)$ is defined as:

$$(CWT)(a,b) = \int_{-\infty}^{+\infty} x(t) \cdot \psi_{a.b}(t)\, dt = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} x(t) \cdot \psi\left(\frac{t-b}{a}\right) dt \tag{4}$$

where the scale factor a, and the translation factor b are continuous variables.
The WT coefficient is an expansion and a particular shift represents how well the original signal $x(t)$ corresponds to the translated and dilated mother wavelet. Thus, the coefficient group of CWT(a,b) associated with a particular signal is the wavelet representation of the original signal $x(t)$ in relation to the mother wavelet (Aggarwal & Kim, 2000).

2.2 Discrete Wavelet Transform (DWT)

2.2.1 Why is DWT needed?

Although the discretized continuous wavelet transform enables the computation of the continuous wavelet transform by computers, it is not a true discrete transform. As a matter of fact, the wavelet series is simply a sampled version of the CWT, and the information it provides is highly redundant as far as the reconstruction of the signal is concerned. This redundancy, on the other hand, requires a significant amount of computation time and resources. The Discrete Wavelet Transform (DWT), on the other hand, provides sufficient information both for analysis and synthesis of the original signal, with a significant reduction in the computation time. The DWT is considerably easier to implement when compared to the CWT. The basic concepts of the DWT will be introduced in this section along with its properties and the algorithms used to compute it (Polikar, 1999).

2.2.2 DWT definition

The redundancy of information and the enormous computational effort to calculate all possible translations and scales of CWT restricts its use. An alternative to this analysis is the

discretization of the scale and translation factors, which leads to the DWT. There are several ways to introduce the concept of DWT, the main are the decomposition bands and the decomposition pyramid (or Multi-Resolution Analysis -MRA), developed in the late 70's (Rioul & Vetterli, 1991). The DWT of the continuous signal $x(t)$ is given by:

$$(DWT)(m,p) = \int_{-\infty}^{+\infty} x(t) \cdot \psi_{m,p}\, dt \tag{5}$$

where $\psi_{m,p}$ form wavelet function bases, created from a translated and dilated mother wavelet using the dilation m and translation p parameters, respectively.
Thus, $\psi_{m,p}$ is defined as:

$$\psi_{m,p} = \frac{1}{\sqrt{a_0^m}} \psi\left(\frac{t - p b_0 a_0^m}{a_0^m} \right) \tag{6}$$

The DWT of a discrete signal $x[n]$ is derived from CWT and defined as (Aggarwal & Kim, 2000):

$$(DWT)(m,k) = \frac{1}{\sqrt{a}} \sum_n x[n] \cdot g\left(\frac{k - n b_0 a_0^m}{a_0^m} \right) \tag{7}$$

where $g(*)$ is the mother wavelet and $x[n]$ is the discretized signal.
The mother wavelet may be dilated and translated discretely by selecting the scaling and translation parameters $a=a_0^m$ and $b=nb_0a_0^m$ respectively (with fixed constants $a_0>1$, $b_0>1$, m and n belonging the set of positive integers).

2.3 Multi-Resolution Analysis (MRA)

The problems of temporal and frequency resolution found in the analysis of signals with the STFT (best resolution in time at the expense of a lower resolution in frequency and vice-versa) can be reduced through a Multi-Resolution Analysis (MRA) provided by WT. The temporal resolutions, Δt, and frequency, Δf, indicate the precision time and frequency in the analysis of the signal. Both parameters vary in terms of time and frequency, respectively, in signal analysis using WT. Unlike the STFT, where a higher temporal resolution could be achieved at the expense of frequency resolution. Intuitively, when the analysis is done from the point of view of filters series, the temporal resolution should increase increasing the center frequency of the filters bank. Thus, Δf is proportional to f, ie:

$$\frac{\Delta f}{f} = c \tag{8}$$

where c is constant.
The main difference between DWT and STFT is the time-scaling parameter. The result is geometric scaling, i.e. 1, $1/a$, $1/a^2$, …; and translation by 0, n, $2n$, and so on. This scaling gives the DWT logarithmic frequency coverage in contrast to the uniform frequency coverage of the STFT, as compared in Fig. 1.
The CWT follows exactly these concepts and adds the simplification of the scale, where all the impulse responses of the filter bank are defined as dilated versions of a mother wavelet

(Rioul & Vetterli, 1991). The CWT is a correlation between a wavelet at different scales and the signal with the scale (or the frequency) being used as a measure of similarity. The CWT is computed by changing the scale of the analysis window, shifting the window in time, multiplying by the signal, and integrating over all times. In the discrete case, filters of different cut-off frequencies are used to analyze the signal at different scales. The signal is passed through a series of high pass filters to analyze the high frequencies, and it is passed through a series of low pass filters to analyze the low frequencies. Thus, the DWT can be implemented by multistage filter bank named MRA (Mallat, 1999), as illustrated on Fig. 2. The Mallat algorithm consists of series of high-pass and the low-pass filters that decompose the original signal $x[n]$, into approximation a(n) and detail d(n) coefficient, each one corresponding to a frequency bandwidth.

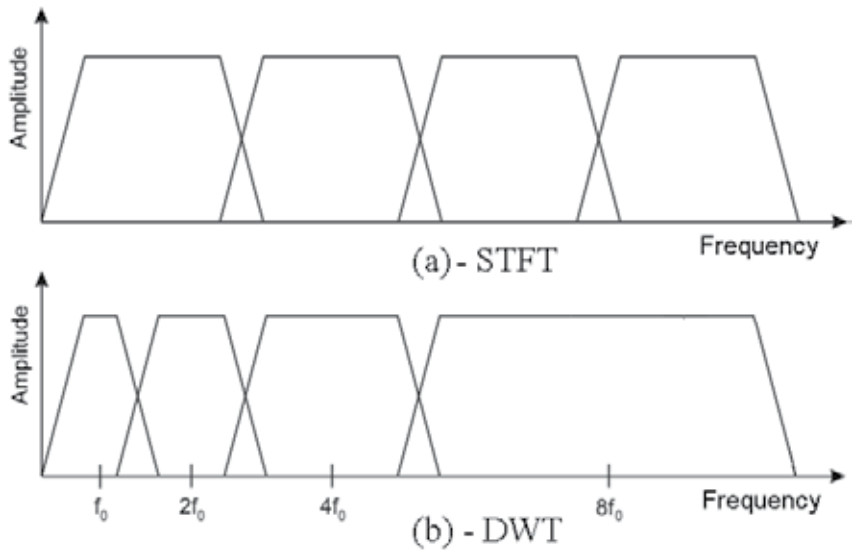


Fig. 1. Comparison of (a) the STFT uniform frequency coverage to (b) the logarithmic coverage of the DWT.

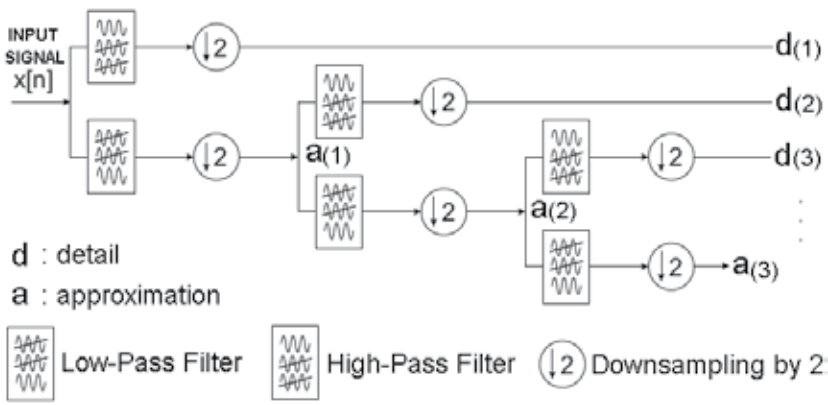


Fig. 2. DWT filter bank framework.

The resolution of the signal, which is a measure of the amount of detail information in the signal, is changed by the filtering operations, and the scale is changed by up-sampling and down-sampling (sub-sampling) operations. Sub-sampling a signal corresponds to reducing the sampling rate, or removing some of the samples of the signal. On the other hand, up-sampling a signal corresponds to increasing the sampling rate of a signal by adding new samples to the signal.
The procedure starts with passing this signal $x[n]$ through a half band digital low-pass filter with impulse response $h[n]$. The filtering process corresponds to the mathematical operation of signal convolution with the impulse response of the filter. The convolution operation in discrete time is defined as follows (Polikar, 1999):

$$x[n] * h[n] = \sum_{k=-\infty}^{\infty} x[k] \cdot h[n-k] \tag{9}$$

A half band low-pass filter removes all frequencies that are above half of the highest frequency in the signal. For example, if a signal has a maximum of 1000 Hz component, then half band low-pass filtering removes all the frequencies above 500 Hz. However, it should always be remembered that the frequency unit for discrete time signals is radians.
After passing the signal through a half band low-pass filter, half of the samples can be eliminated according to the Nyquist's rule. Simply discarding every other sample will subsample the signal by two, and the signal will then have half the number of points. The scale of the signal is now doubled. Note that the low-pass filtering removes the high frequency information, but leaves the scale unchanged. Only the sub-sampling process changes the scale. Resolution, on the other hand, is related to the amount of information in the signal, and therefore, it is affected by the filtering operations. Half band low-pass filtering removes half of the frequencies, which can be interpreted as losing half of the information. Therefore, the resolution is halved after the filtering operation. Note, however, the sub-sampling operation after filtering does not affect the resolution, since removing half of the spectral components from the signal makes half the number of samples redundant anyway. Half the samples can be discarded without any loss of information.
This procedure can mathematically be expressed as (Polikar, 1999):

$$y[n] = \sum_{k=-\infty}^{\infty} h[k] \cdot x[n-k] \tag{10}$$

The decomposition of the signal into different frequency bands is simply obtained by successive highpass and lowpass filtering of the time domain signal. The original signal $x[n]$ is first passed through a halfband highpass filter $g[n]$ and a lowpass filter $h[n]$. After the filtering, half of the samples can be eliminated according to the Nyquist's rule, since the signal now has a highest frequency of $p/2$ radians instead of p. The signal can therefore be sub-sampled by 2, simply by discarding every other sample. This constitutes one level of decomposition and can mathematically be expressed as follows (Polikar, 1999):

$$y_{high}[k] = \sum_{n} x[n] \cdot g[2k-n] \tag{11}$$

$$y_{low}[k] = \sum_{n} x[n] \cdot h[2k - n] \tag{12}$$

where $y_{high}[k]$ and $y_{low}[k]$ are the outputs of the high-pass and low-pass filters, respectively, after sub-sampling by 2.

2.4 Energy and power of discrete signal

The total energy of a discrete signal $x[n]$ is given for equation (Haykin & Veen, 2001):

$$E = \sum_{n=-\infty}^{+\infty} x^2[n] \tag{13}$$

and the average power is defined as:

$$P = \lim_{x \to \infty} \frac{1}{2N} \sum_{n=-N}^{N} x^2[n] \tag{14}$$

For a periodic signal of fundamental period N, the average power is given by:

$$P = \frac{1}{N} \sum_{n=0}^{N-1} x^2[n] \tag{15}$$

3. Differential protection of power transformers using DWT

3.1 Percentage differential protection

Differential protection schemes are widely used by electric companies to protect EPS equipments. This relaying technique is applied on power transformers protection, buses protection, and large motors and generators protection among others (Anderson, 1999). Considering power transformers rated above 10 MVA, the percentage differential relay with harmonic restraint is the most used protection scheme (Horowitz & Phadke, 2008). The percentage differential relay can be implemented with an over-current relay (R) and operation (o) and restriction coils (r), as illustrated on Fig. 3, connected between Current Transformer (CTs).

Under normal operating conditions or external faults, the CTs secondary currents, i_{2P} and i_{2S}, have close absolute values. The differential protection formulation compares the differential current to a fixed threshold value. To include CTs transformation errors, CTs mismatch and power transformer variable taps, the differential current (i_d) can be compared to a fixed percentage value of the restraint current (i_r). This percentage characteristic of the relay, named K, is given by:

$$K = \frac{i_{2P} - i_{2S}}{(i_{2P} + i_{2S})/2} = \frac{i_d}{i_r} \tag{16}$$

The relay identifies an internal fault when the differential current exceeds the percentage value K of the restraint current, where i_{op} is the operation current of relay:

$$i_d = i_{op} \geq K \cdot i_r = K \cdot \frac{(i_{2P} + i_{2S})}{2} \tag{17}$$

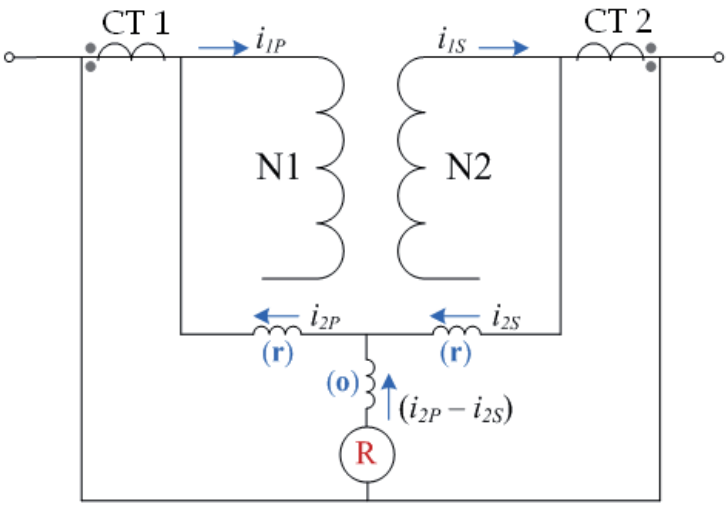


Fig. 3. Differential relay connections.

3.2 Proposed protection algorithm using DWT

A change in the spectral energy of the wavelets components of the current differential is noted when different electrical events (external faults, internal faults and/or inrush current) occur on the power transformers (Megahed et al., 2008). In this sense, the discrimination criterion of the proposed protection algorithm in this work is based in the spectral energy level generated by the electrical event type. The flow chart of the proposed algorithm is presented on Fig. 4.

In the disturbance detection (BLOCK 1) the activation current is calculated. The activation current is calculated for each phase through the percentage characteristic K and the restraint currents showed in equation (17). The activation current is given by:

$$\left|Id_{A,B,C}\right| \geq \left|I_a\right| = \left|k \cdot i_r\right| = \left|K \cdot \frac{(i_{2p} + i_{2s})^{A,B,C}}{2}\right| \tag{18}$$

where I_a is the activation current, $Id_{A,B,C}$ is the differential current on A, B and C phases, K is the percentage differential characteristic and i_r is the restraint current.

In the disturbance identification (BLOCK 2) the three-phase differential currents are initially processed through a DWT implemented as filter bank. After, a restraint index R_{ind}, is calculated. This index quantifies the relative magnitude characteristic of the differential signals in the 1st detail ($D1$) and is defined as the relation between the maximum detail coefficient from $D1$ and the detail-spectrum-energy (DSE) of the wavelet coefficient. Thus, R_{ind} is given by:

$$R_{ind} = \frac{d_{\max,D1}}{\sum_{c=1}^{M}\left|d_{(c)}\right|^2 \Delta t} \tag{19}$$

where $d_{max,D1}$ is the maximum detail coefficient from $D1$, M is the total number of wavelet coefficients from $D1$ and Δt is the sampling period.

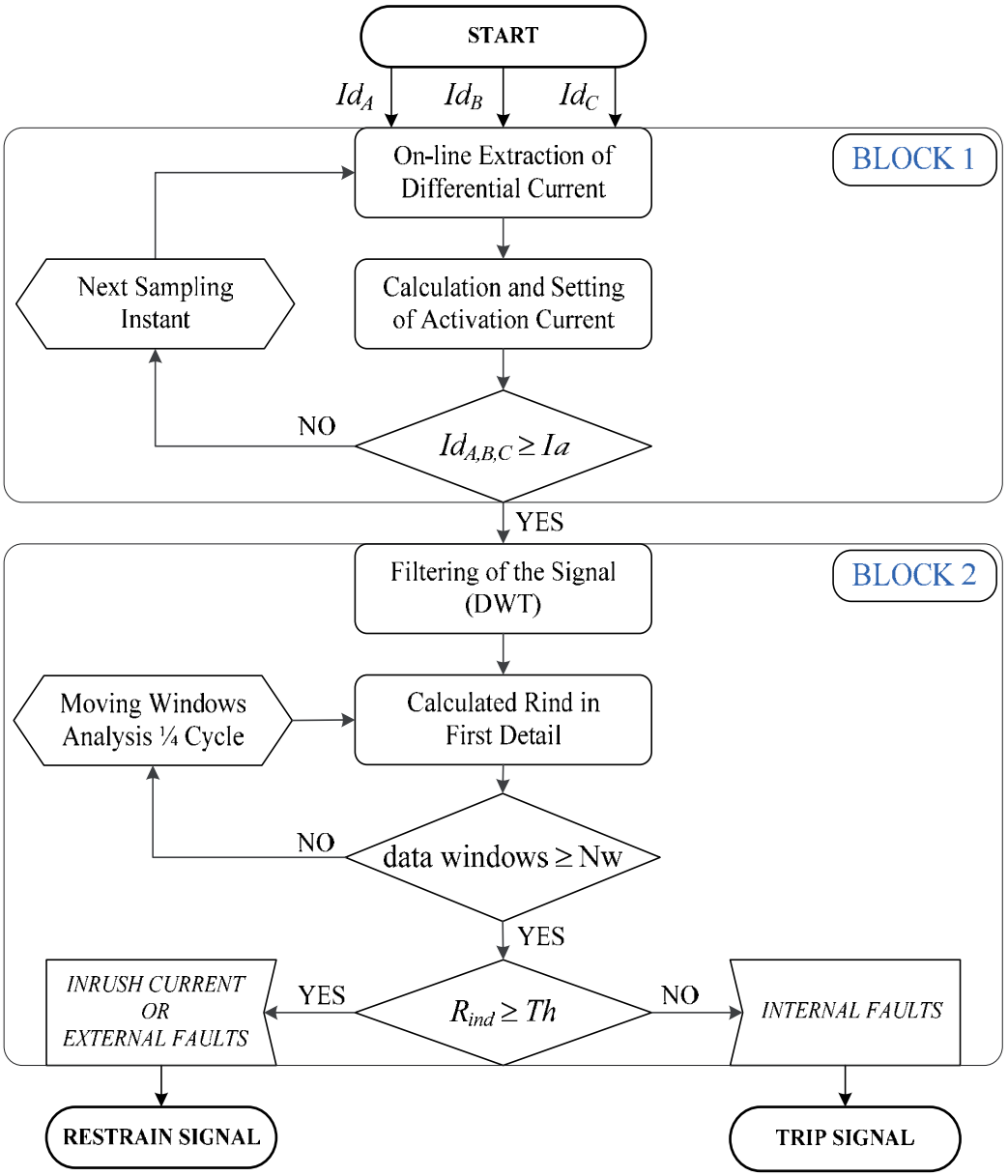


Fig. 4. Proposed Algorithm's Operation Scheme.

The proposed algorithm was implemented in MATLAB® platform (Matlab, 2010). Fig. 5 presents the graphical interface developed with three input block: 1) selecting the disturbance type; 2) selecting of the wavelet analysis characteristics; 3) analyzed results outputs.

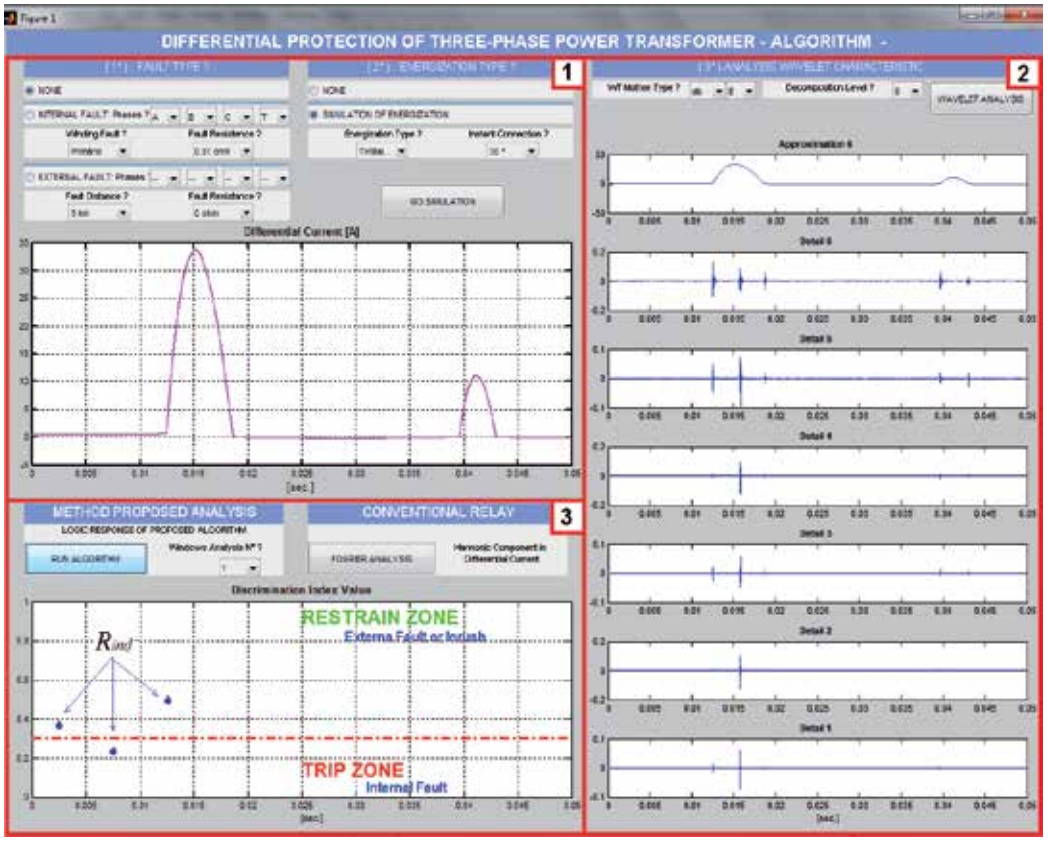


Fig. 5. Graphical implementation in MATLAB® environment.

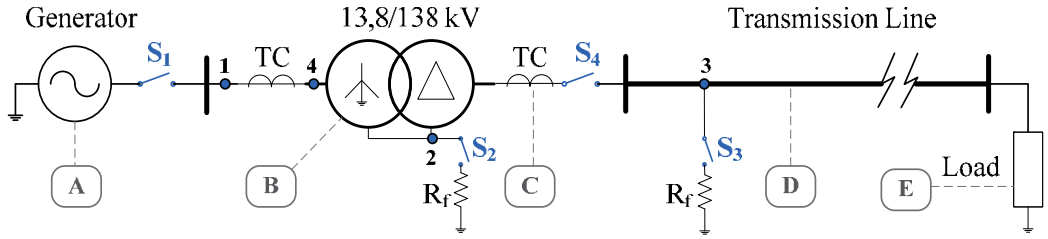


Fig. 6. Simulated electric power system.

4. Case study

Fig. 6. illustrates the studied electrical power system. The studied system consists of:

a. Generator: 13.8 kV, 30 MVA, 50 Hz;
b. Power Transformer (PT): 35 MVA, 13.8/138 kV, Yg–Δ;
c. Current Transformers (CT) with 1200/5 and 200/5 turns ratio;

d. Transmission line: with a length of 100 km;
e. Variable Load of 3, 10 or 25 MVA all with a 0.92 power factor.

The switches shown in Fig. 6, S_1 and S_4, are used to simulate the energization operation of the PT. In this phenomenon the transformer is connected without load. The switch S_3 simulates external faults through a fault resistor R_f. The closing of the switch S_2 simulates an internal faults to the PT in both the primary and secondary windings.

4.1 Types of analyzed events

The proposed algorithm operates through three-phase differential currents. The simulations performed are presented on Table 1:

N°	Type Event
1	Different energization cases, comprising different switching inception angles (0°, 30°, 60° and 90°) by closing the switch S_1 in the Low Voltage (LV) side.
2	Internal faults in both primary and secondary sides of the transformer. These faults were simulated with a fault resistance R_f values of 0 Ω, 0.01 Ω, 10 Ω, and 100 Ω.
3	Several cases of external faults with fault resistances R_f values: 0 Ω, 0.01 Ω, 10 Ω, and 100 Ω.
4	Faults applied between the PT and the CTs.
5	Energizing the PT with the presence of internal faults
6	Energizing the PT with the presence of external faults.

Table 1. Simulated Events.

5. Simulation and analysis result

In order to evaluate the proposed protection algorithm efficiency, internal faults and transient inrush currents have been simulated. For each simulation, the proposed algorithm used different mother wavelets to evaluate accuracy and speed. The mother wavelets tested in this study were: Daubechies (Db), Symlet (Sy), Haar (Hr), Coiflet (Coif) and Morlet (Mo).

5.1 Transient signal and fault current simulation

The transient signal (inrush current) and fault current simulated are concentrated in the following situations:

- Fig. 7 presents an energization case. Part (a) illustrates the voltages in the secondary side of the PT. Part (b) the differential current are presented.
- Fig. 8 illustrates a case of energization with internal fault (concurrent event). The internal fault was simulated in the A phase with fault resistance $R_f = 10$ Ω.
- Fig. 9 illustrates a case of external fault removal. The faults occurring at 3 km to the PT on the transmission line.

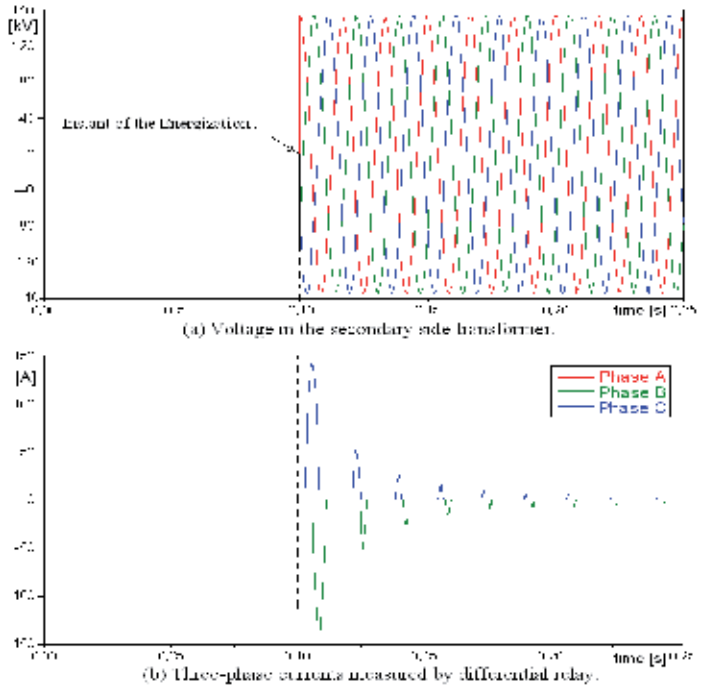


Fig. 7. Energization simulation on PT.

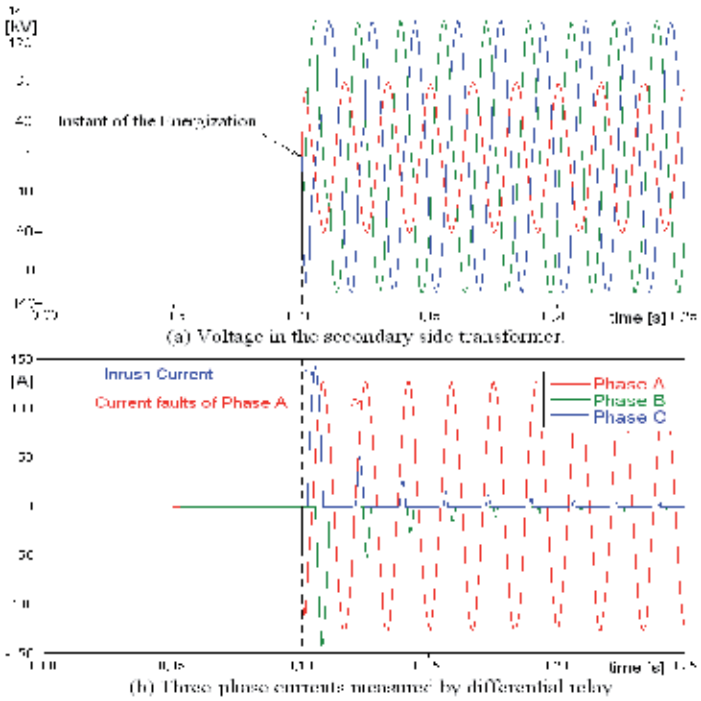


Fig. 8. Energization and internal faults simulation on PT.

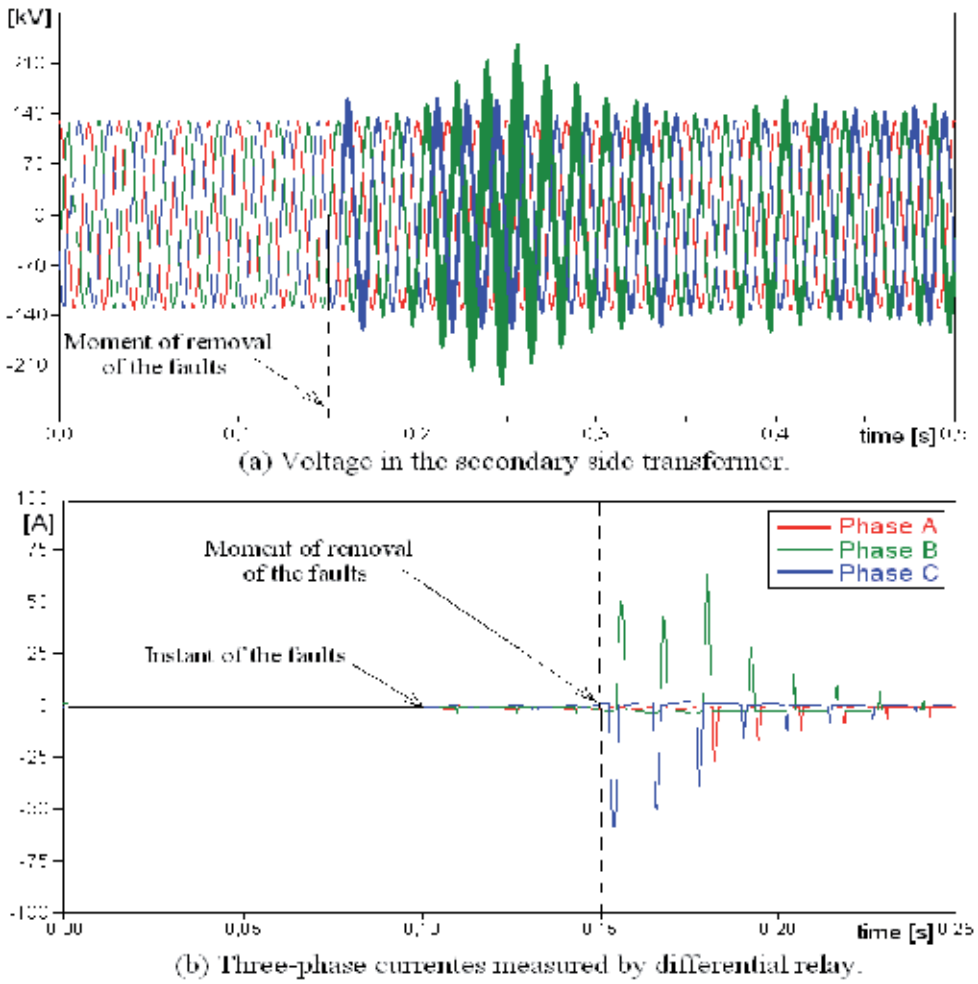


Fig. 9. External faults removal simulation.

5.2 Algorithm proposed analysis

Depending on the voltage angle in which the transformer is connected to the EPS, its residual flux can cause transient inrush currents which are correctly discriminated by the proposed protection algorithm.

Fig. 10 shows the algorithm response to a transient inrush current. Fig. 10(a) presents the inrush current in differential circuit of the power transformer. Fig. 10(b) shows the first detail of the DWT decomposition where a maximum number of three windows analyses are implemented on detail coefficient of the WT. Three windows analyses (N_w) are necessary to guarantee a correct decision by the methodology. The window analysis is moving 1/4 cycle for each restraint index (R_{ind}) calculated to avoid false operations of the protection algorithm. After calculating and analyzing the ratio index for event discrimination, the proposed algorithm sends a restrain signal to the protection relay. Note on Fig. 10(c) the adaptive threshold value is proportional to the differential current caused by the internal fault.

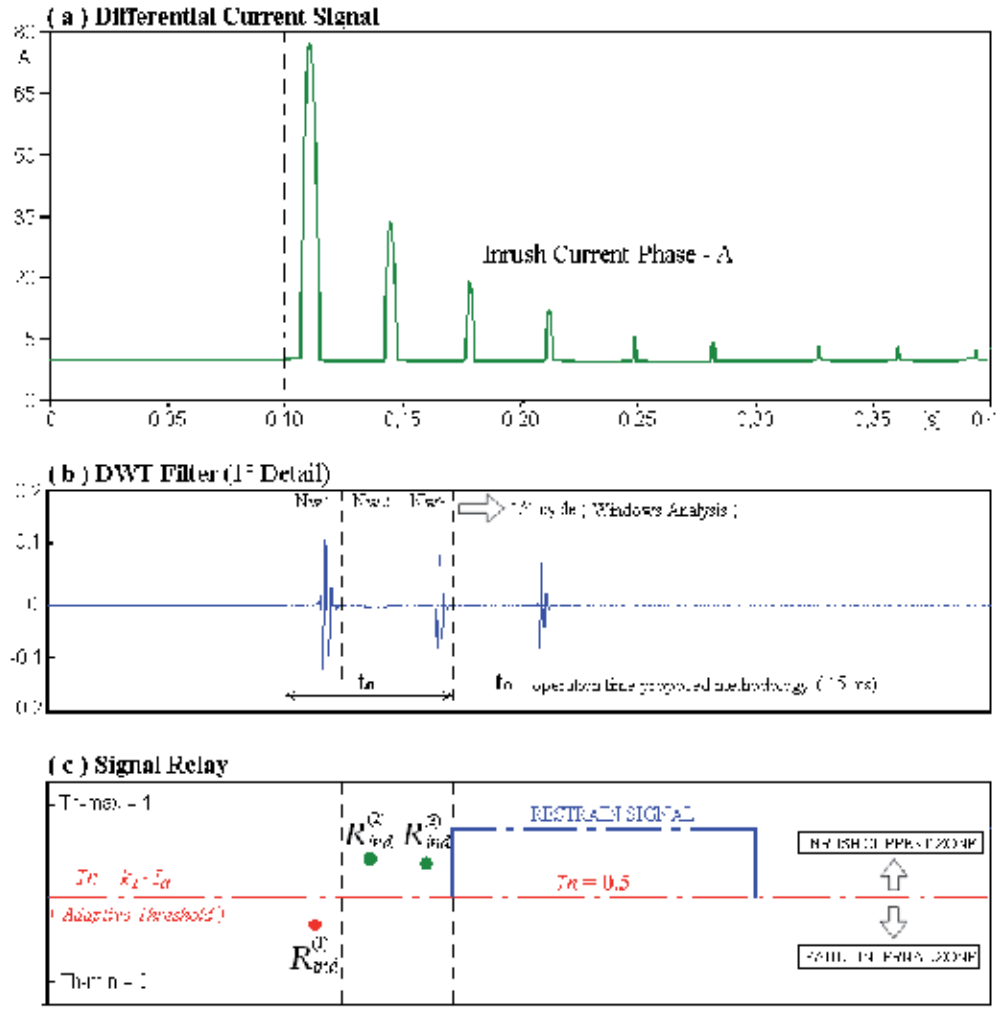


Fig. 10. Logical decision of the proposed algorithm to energization phenomenon.

5.3 Obtained results

The magnitude and shape of inrush current changes depending on several factors such as energization instant, core remnant flux, saturation of CTs and non-linearities of transformer core. However, in this work only the switching instant was evaluated. 12 energization cases were simulated for each switching angle and evaluated with the following mother wavelet: Daubechies (Db), Harr (Hr), Symlet (Sy), Coiflet (Coif) and Morlet (Mo).

Table 2 shows the proposed algorithm performance in correct operation number (OC[%]) for transformer energization. In test development, the Daubechies mother wavelets presented the best performance for all switching angles with 97.11% correct diagnosis. The Harr mother wavelet type appeared as the least efficient with 18.75% of correct diagnosis. Furthermore, at 90° switching angle presented the worse energization condition because it was the least correctly identified (56.66%). However, others switches angles tested did presented a significant effect on the inrush current identification.

Switch Angle	Mother Wavelet Type					OC [%]
	Db	Hr	Sy	Coif	Mo	
0°	12	7	12	12	12	91.66
30°	12	2	12	12	12	83.33
60°	12	0	10	12	10	73.33
90°	11	0	9	8	6	56.66
OC [%]	97.11	18.75	89.58	91.66	85.41	

Table 2. Performance of the proposed algorithm in percentage of correct operation (OC) [%] to different switching instants.

Table 3 summarizes the methodology efficiency in percentage of correct operation of the proposed algorithm for different internal faults types and different fault resistances (RF). The performance was evaluated considering a constant load of 10 MVA on the end of the transmission line. There was an important drop in accuracy of the protection algorithm to internal fault cases in faults type A-B and A-B-C. However, the discrimination of faults type A-G (phase-ground) and A-B-G showed little sensitivity to R_f variation.
It was noted that the mother wavelet Daubechies showed an excellent performance and high efficiency in discrimination of simulated disturbances. This is because the decomposition solutions using Daubechies wavelet function are orthogonal and no marginal overlaps will happen during the signal reconstruction. The mother wavelet Symlet and Coiflet presented a satisfactory performance with a greater efficiency than the Morlet type. On the other hand, the wavelet Haar type did not achieved a good performance, presenting many inaccuracies in the discrimination of all simulated disturbances.

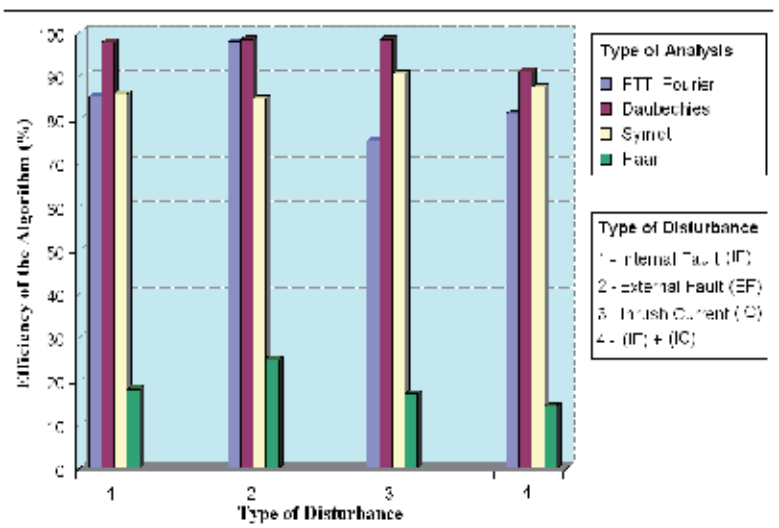


Fig. 11. Comparison between type wavelets functions and Fourier analysis (FTT).

To verify the wavelet function type effect on the proposed formulation, 3 wavelets function were compared with conventional protection methodology based in Fourier Analysis (FTT). The wavelet type used in the comparison study were: Daubechies, Haar and Symlet. The Fig. 11 shows the test results and the comparison between the proposed algorithm, a

conventional percentage differential protection relay. It can be observed that the conventional technique based on FTT obtained a lower efficiency than the proposed algorithm.

Mother Wavelet	R_f [Ω]	Internal Fault Type			
		A-G	A-B	A-B-G	A-B-C
Db	0.01	100.0	100.0	100.0	100.0
	10	100.0	100.0	100.0	100.0
	50	99.22	98.28	100.0	100.0
	100	98.90	97.66	98.44	100.0
Hr	0.01	82.36	81.65	83.15	84.15
	10	76.32	76.54	75.18	75.36
	50	72.65	71.54	73.21	73.26
	100	70.18	69.32	70.15	70.15
Sy	0.01	99.38	100.0	100.0	100.0
	10	98.75	98.75	99.68	100.0
	50	97.81	97.65	98.75	98.75
	100	97.18	97.03	98.12	95.75
Coif	0.01	100.0	100.0	100.0	100.0
	10	99.38	98.75	100.0	97.34
	50	98.75	91.25	97.50	92.81
	100	97.65	87.66	96.87	88.28
Mo	0.01	97.21	96.54	94.65	94.36
	10	96.24	95.64	95.63	94.62
	50	95.12	96.35	94.32	93.12
	100	90.15	84.71	85.63	89.34

Table 3. Performance of the proposed algorithm to internal fault cases.

6. Conclusions

In this chapter a novel formulation for differential protection of three-phase transformers, based on the differential current transient analysis is proposed. The algorithms performance is evaluated using fault simulations in a typical electrical system under BPA's ATP/EMTP software. The algorithm considers the different magnitudes assumed by the DWT coefficients, induced by internal faults and inrush currents. The wavelet decomposition allows good time and frequency precision to characterize the transient events.

The proposed algorithm is comprehensible and feasible for implementation showing a correct operation with the adaptive threshold value. The obtained results through various simulated fault cases and non-fault disturbances showed that the proposed algorithm is robust and accurate.

Based on these tests and after critical evaluation of the proposed protection algorithm important conclusions could be observed:

- The use of Wavelet Transforms to analyze differential signals produced by transient phenomenon proved to be an effective and robust tool.

- The variation of wavelets spectral energy coefficients proved to be an effective measure of discrimination.
- The proposed algorithm presents a perspective of practical application given the simplicity under which the methodology is based.
- The performance comparison made between the wavelet types: Daubechies (Db), Harr (Hr), Symlet (Sy), Coiflet (Coif) and Morlet (Mo), showed that the use of the Daubechies wavelet is the most appropriated.
- The comparative study with the conventional differential protection algorithm showed that the proposed formulation presents greater performance.

7. References

Abed, N. Y. ; Mohammed, O. A. (2007). Modeling and Characterization of Transformer Internal Faults Using Finite Element and Discrete Wavelet Transforms. *IEEE Transaction on Magnetics*, Vol. 43, No. 4, (April 2007), pp. 1425-1428, ISSN 0018-9464.

Aggarwal, R.; Kim, C. H. (2000). Wavelet Transform in Power System : part 1 –general introduction to the wavelet transform. *Power Engineering Journal*, Vol. 14, No. 2, (April 2007), pp. 81-87.

Anderson, P. M. (1999). *Power System Protection*, Wiley-Interscience: IEEE Press with McGraw-Hill. ISBN 0-7803-3427-2, New Jersey, USA.

Bentley, P. M.; McDonnell, J. T. E. (1994). Wavelets Transform: an introduction. *IEE Electronic & Communication Engineering Journal*, Vol. 6. No. 4 (August 1994), pp. 175-186, ISSN 0954-0695.

Bonneville Power Administration. (2002). *Alternative Transient Programs: ATP/EMTP*. Retrieved from: http://www.emtp.rog/.

Daubechies, I. (1990). The Wavelet Transform, Time-Frequency Localization and Signal Analysis. *IEEE Transactions on Information Theory*, Vol. 36, No. 5 (September 1990), pp. 961-1005, ISSN 0018-9448.

Eissa, M. M. (2005). A Novel Digital Directional Transformer Protection Technique Based on Wavelet Packet. *IEEE Transactions on Power Delivery*, Vol. 20, No. 3, (July 2005), pp. 1830-1836, ISSN 0885-8977.

Faiz, J.; Lotfi-Fard, S. (2006). A Novel Wavelet-Based Algorithm for Discrimination of Internal Faults From Magnetizing Inrush Currents in Power Transformer. *IEEE Transactions on Power Delivery*, Vol. 21, No. 4, (October 2006), pp. 1989-1996, ISNN 0885-8977.

Harlow, J. H. (2007). *Electric Power Transformer Engineering* (2nd Edition), CRC Press Taylor & Francis Group. ISBN 10-0-8493-9186-5, Boca Raton, USA.

Haykin, S.; Veen, B. V. (2001). *Signals and Systems* (2nd Edition), Jhon Wiley & Sons Inc. ISBN 8-57-307741-7, Porto Alegre, Brazil.

Horowitz, S. H.; Phadke, A. G. (2008). *Power System Relaying* (3nd Edition), Ed. Research Studies Press Ltd. ISBN 978-0-470-05712-4, Baldock, England.

Mallat, S. (1999). *A Wavelet Tour of Signal Processing* (2nd Edition), Academic Press. ISBN 0-12-466606-X, California, USA.

Mao, P . L. ; Aggarwal, R. K. (2000). A Wavelets Transform Based Decision Making Logic Method for Discrimination Between Internal Faults and Inrush Currents in Power Transformer. *International Journal of Electrical Power and Energy Systems*, Vol. 22. No. 6 (August 2000), pp. 389-395, ISSN 0142-0615.

Megahed, A. I.; Ramadan, A.; ElMahdy, W. (2008). Power Transformer Differential Relay Using Wavelet Transform Energies, *Proceedings in the Power and Energy Society General Meeting IEEE*, pp. 1-6, ISBN 978-1-4244-1905-0, Pittsburgh, USA, July 20-24, 2008.

Morate, M. G.; Nicoletti, D. W. (1999). A Wavelet-Based Differential Transformer Protection. *IEEE Transactions on Power Delivery*, Vol. 14, No. 4, (November 1999), pp. 1351-1358, ISSN 0885-8977.

Ngaopitakkul, A.; Kunakorn, A. (2006). Internal Faults Classification in Transformer Windings Using Combination of Discrete Wavelet Transform and Back-Propagation Neural Networks. *International Journal of Control, Automation, and Systems*, Vol. 4. No. 3, (June 2006), pp. 365-371, ISSN 1598-6446.

Phadke, A. G. ; Thorp, J. S. (2009). *Computer Relaying For Power System* (2nd Edition), Ed. Research Studies Press Ltd. ISBN 978-0-470-05713-1, Baldock, England.

Polikar, R. (1999). The Story of Wavelets, *Physics and Modern Topics in Mechanical and Electrical Engineering*, World Scientific and Eng. Society Press, pp. 192-197, USA, Retrieved from: http://www.public.iastate.edu/~rpolikar.

Rioul, O.; Vetterli, M. (1991). Wavelets and Signal Processing. *IEEE Signal Processing Magazine*, Vol. 8. No. 4 (October 1994), pp. 14-38, ISSN 1053-5888.

Robertson, D. C. ; Camps, O. I. ; Mayer, J. S. ; Gish, W. B. (1996). Wavelets and Electromagnetic Power System Transient. *IEEE Transactions on Power Delivery*, Vol. 11, No. 2, (April 1996), pp. 1050-1058, ISSN 0885-8977.

Saleh, S. A.; Rahman, M. A. (2005). Modeling and Protection of Three-Phase Power Transformer Using Wavelet Packet Transform. *IEEE Transactions on Power Delivery*, Vol. 20, No. 2 (April 2005), pp. 1273-1282, ISSN 0885-8977.

Sarkar, T. K.; Su, C. (1998). A Tutorial on Wavelets from an Electrical Engineering Perspective, Part 2: The Continuous Case. *IEEE Antennas and Propagation Magazine*, Vol. 40, No. 6 (December 1998), pp. 36-49, ISSN 1045-9243.

Simpson, D. M. (1993). An Introduction to the Discrete Orthogonal Wavelet Transform. *Revista Brasiliera de Engenharia*, Vol. 9, No. 1, (July 1993), pp. 57-81.

The Mathworks Inc. (2010). *Mathworks Matlab*. Retrieved from: http://www.mathworks.com/

Thomas, W.; Ozgönenel, O. (2007). Diagnostic of transformer internal faults through ANN based on radial basis functions and dynamical principal component analysis. *IET Generation,Transmission & Distribution,* pp. 1-11.

Vetterli, M. ; Herley, C. (1992). Wavelets and Filter Banks: Theory and Design. *IEEE Transactions on Signal Processing*, Vol. 40, No. 9, (September 1992), pp. 2207-1992, ISSN 1053-587X.

Wang, H.; Butler, K. L. (2001). Finite Element Analysis of Internal Winding Faults in Distribution Transformer. *IEEE Transaction on Power Delivery,* Vol. 16, No. 3 (July 2001), pp. 422-428, ISSN 0885-8977.

Wiszniewski, A.; Kasztenny, B. (1995). A Multi-Criteria Differential Transformer Relay Based on Fuzzy Logic. *IEEE Transaction on Power Delivery,* Vol. 10, No. 4 (October 1995), pp. 1786-1792, ISSN 0885-8977.

Zaman, M. A.; Hoque, M. A.; Rahman, M. A. (1996). On-line Implementation of the Artificial Neural Network Based Protection for Power Transformer, *Proceedings of NECEC*, pp. 5-11, NL, Canada, May 17-22, 1996.

Edited by Hannu Olkkonen

The discrete wavelet transform (DWT) algorithms have a firm position in processing of signals in several areas of research and industry. As DWT provides both octave-scale frequency and spatial timing of the analyzed signal, it is constantly used to solve and treat more and more advanced problems. The present book: Discrete Wavelet Transforms - Biomedical Applications reviews the recent progress in discrete wavelet transform algorithms and applications. The book reviews the recent progress in DWT algorithms for biomedical applications. The book covers a wide range of architectures (e.g. lifting, shift invariance, multi-scale analysis) for constructing DWTs. The book chapters are organized into four major parts. Part I describes the progress in implementations of the DWT algorithms in biomedical signal analysis. Applications include compression and filtering of biomedical signals, DWT based selection of salient EEG frequency band, shift invariant DWTs for multiscale analysis and DWT assisted heart sound analysis. Part II addresses speech analysis, modeling and understanding of speech and speaker recognition. Part III focuses biosensor applications such as calibration of enzymatic sensors, multiscale analysis of wireless capsule endoscopy recordings, DWT assisted electronic nose analysis and optical fibre sensor analyses. Finally, Part IV describes DWT algorithms for tools in identification and diagnostics: identification based on hand geometry, identification of species groupings, object detection and tracking, DWT signatures and diagnostics for assessment of ICU agitation-sedation controllers and DWT based diagnostics of power transformers.The chapters of the present book consist of both tutorial and highly advanced material. Therefore, the book is intended to be a reference text for graduate students and researchers to obtain state-of-the-art knowledge on specific applications.

ISBN 978-953-51-5599-7

Photo by LeeYiuTung / iStock

IntechOpen