

G9020

Signal Processing

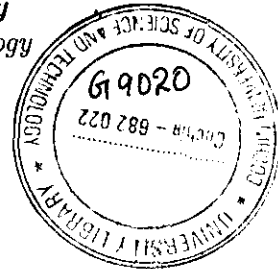
Ph.D Thesis

**PSWT BASED LINEAR PREDICTIVE CODING
AND
DEVELOPMENT OF A PARALLEL MULTIPLE SUBSEQUENCE
STRUCTURE FOR DWT COMPUTATION**

Submitted to the
Cochin University of Science And Technology
in partial fulfillment of the requirements for the award of the degree of
Doctor of Philosophy
in the Faculty of Technology

by

V.P. DEVASSIA



Under the guidance of

Dr. Tessamma Thomas

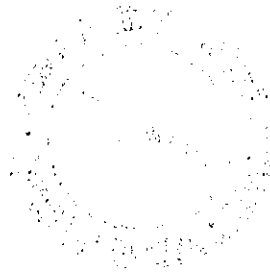
DEPARTMENT OF ELECTRONICS
COCHIN UNIVERSITY OF SCIENCE AND TECHNOLOGY
COCHIN, KERALA, INDIA 682 022

DECEMBER 2003

G9020



T
621.391
D&V



Dedicated to..

My PARENTS

DEPARTMENT OF ELECTRONICS
COCHIN UNIVERSITY OF SCIENCE AND TECHNOLOGY
COCHIN -22

C E R T I F I C A T E

This is to certify that this Thesis entitled PSWT Based Linear Predictive Coding and Development of a Parallel Multiple Subsequence Structure for DWT Computation is a bonafide record of the research work carried out by Mr. V.P. Devassia under my supervision in the Department of Electronics, Cochin University of Science And Technology. The results presented in this thesis or parts of it have not been presented for any other degree.

Tessamma Thomas

Dr. Tessamma Thomas
(Supervising Guide)

Reader

Department of Electronics


Cochin University of Science And Technology

Kochi
29-12-2003

D E C L A R A T I O N

*I hereby declare that this Thesis entitled **PSWT Based Linear Predictive Coding and Development of a Parallel Multiple Subsequence Structure for DWT Computation** is based on the original research work carried out by me under the supervision of **Dr. Tessamma Thomas** in the Department of Electronics, Cochin University of Science And Technology. The results presented in this thesis or parts of it have not been presented for the award of any other degree.*

Kochi
29-12-2003


V.P. Devassia

Acknowledgement

May I bow my head before **God Almighty** for showering His choicest blessings on me all through the years.

It is with deep sense of gratitude that I acknowledge the guidance of my research guide, **Dr. Tessamma Thomas**, Reader, Department of Electronics, Cochin University of Science And Technology, who has been the source of inspiration in accomplishing this research work.

The eminent personality who lead me to the world of Wavelet Transform is none other than **Prof. C.S. Sreedhar**, former head of the Department of Electronics, Cochin University of Science And Technology. I remember with much gratitude, the fruitful discussions I had with him, during various stages of the research work.

A good number of scholars and well-wishers have helped me in this venture. For fear of omission I am not mentioning them all here by name. Let me express my sincere thanks and deep sense of gratitude to all of them, especially to,

- the Director IHRD for giving me an opportunity to carry out this research work by sparing me from my responsibilities in the institution.
- Prof. K.G. Balakrishnan, Head of the Department of Electronics, Cochin University of Science And Technology and Prof. P.R.S. Pillai, former head of the Department, for extending the facilities in the department for my research work.
- Prof. K.T. Mathew, Dept. of Eelctronics, CUSAT, Prof. C.P. Kuriakose, former PVC of Cochin University of Science And Technology and Prof. K.S.M Panicker, Principal, N.S.S. College of Engineering, Palakkad for their continuous motivation and encouragement throughout the period of this work.
- Dr. V. Deepurajan, Dr. D. Rajaveerappa, and Mr. James Kurian of the department of Electronics, CUSAT and Mr. Jacob Thomas V. Asst. Professor, Model Engineering College, for helping me by providing with valuable resource materials.
- all the faculty members of the department of electronics, especially Prof. K. Vasudevan and Prof. P. Mohanan, for their timely help and support.
- Prof. M.N.N. Namboodiri, Department of Mathematics, Cochin University of Science And Technology who helped me in clarifying the mathematical concepts.

- Dr. Joe Jacob, Newman College Thodupuzha and Dr. Thomaskutty Mathew, M.G. University Regional Centre Edappally, for their valuable suggestions and selfless assistance.
- Mrs. M.G. Mini, the co-researcher, for her encouraging association.
- all the research scholars of the department, especially Mr. T.K. Mani, Dr. Bijukumar S., Dr. Jacob George, Dr. Binoy G.S., and Mr. Dineshkumar V.P., for the timely services and friendly association.
- the library and administrative staff of the department for their cooperation and services.
- Dr. V.P. Joseph, Christ College Irinjalakkuda, for his selfless help, suggestions and association.
- all my family members for their help and sacrifice, and
- M/s. Beeta Transcription Services Thrippunithura, for their help in editing and printing the thesis.

V.P. Devassia

Contents

1	Introduction	1
1.1	Signal Processing	4
1.2	The Wavelet Analysis	5
1.2.1	The Continuous Wavelet Transform	7
1.2.2	The Discrete Wavelet Transform	7
1.3	Wavelet Transform Computation	8
1.3.1	The Pyramid Structure	9
1.3.2	Alternate Structures for WT Computation	9
1.4	Wavelet Transform for Pseudo-Periodic signals	9
1.5	Linear Predictive Coding and Signal Compression	10
1.5.1	Linear Prediction	10
1.5.2	LPC Technique for Data Compression	11
1.6	Motivation for the Work	12
1.7	Layout of the Thesis	12
2	Literature Review	15
2.1	Linear Predictive Coding and Signal Compression	18
2.2	WT based Signal Processing	22
2.2.1	WT based Signal Compression	24
2.2.2	PSWT based Signal Processing	28
2.3	WT Computation	34
2.3.1	1D DWT Computation	34
2.3.2	2D DWT Computation	38
3	Review of Basic Theory	39
3.1	The Wavelets	41
3.2	The Continuous Wavelet Transform	42
3.3	The Discrete Wavelet Transform	44
3.4	Wavelets and Time-Frequency Analysis -Concept of MRA	46

3.4.1	Discrete multiresolution analysis	48
3.5	DWT computation	49
3.5.1	The Basic Multirate Operations	49
3.5.2	The Mallat's Pyramid Structure	52
3.6	The Wavelet Packet Transform	54
3.7	2D Discrete Wavelet Transform	57
3.7.1	Computation of 2D DWT	58
3.8	The Pitch Synchronous Wavelet Transform	61
3.9	Period Estimation	64
3.9.1	Autocorrelation-based Period Estimation	65
3.9.2	UDWT based Period Estimation	66
3.10	The Linear Predictive Coding	68
4	PSWT Based Linear Predictive Coding	73
4.1	Introduction	75
4.2	PSWT based LPC	75
4.2.1	Period Estimation	76
4.2.2	Pitch-Synchronous Representation	82
4.2.3	Normalization	82
4.2.4	Computation of PSWT Coefficients	85
4.2.5	Computation of Predictor Coefficients	85
4.2.6	Signal Encoding and Decoding Scheme	87
4.2.7	Performance Evaluation	88
4.3	Case Studies	90
4.3.1	Case Study 1: ECG signal	91
4.3.2	Case Study 2: Vocal sound	99
4.3.3	Case study 3: Instrumental Music	105
4.4	Feature Enhancement considerations in PSWT domain	109
4.4.1	Method of Noise suppression and Feature enhancement	111
4.4.2	Effect of Feature Enhancement in PSWT based Predictive Coding	114
4.5	Conclusions	115
5	The PMS Computational Structure	117
5.1	Introduction	119
5.1.1	Issues in Wavelet Transform Computation	119
5.2	Development of the PMS Structure for 1D Signals	121
5.2.1	Eliminating Irrelevant Computations	122
5.2.2	The Proposed Structure	123
5.2.3	Analysis of Computational Complexity	127
5.2.4	Comparison of PMS Structure with Pyramidal Structure	130

5.2.5	Case Study: PSWT Computation	134
5.3	WT Computation for 2D data	137
5.3.1	Development of PMS Structure for 2D DWT Computation	137
5.3.2	Comparison of 2D PMS Structure with 2D Pyramidal Structure	140
5.3.3	Case Study: Detection of Micro-Calcification in Mammograms	145
5.4	Conclusions	145
6	Summary and Conclusions	149
6.1	Summary of the Work and the important Conclusions	151
6.2	Scope for Further Investigations	153
 Appendix		
A	WT Based Signal Compression	155
A.1	Introduction	155
A.2	Implementation	156
A.3	Results and Discussion	157
A.4	Conclusion	160
B	WT based Signal Segmentation	161
B.1	Introduction	161
B.2	The Classification Algorithm	163
B.2.1	Short-Time Energy	163
B.2.2	Short-Time Zero Crossing Rate	163
B.2.3	Short-Time Zero-Crossing Energy Product	164
B.2.4	Pitch Correlation Factor	164
B.3	Results and Discussions	165
B.4	Conclusions	165
 Bibliography		
169		
 List of Publications		
189		
 Index		
193		

List of Figures

3.1	Schematic of MRA decomposition	48
3.2	Fractional sampling rate conversion by multirate techniques	52
3.3	Pyramid structure for 2 level DWT Computation	53
3.4	Frequency bands for the analysis tree of the pyramid structure	55
3.5	Time-Frequency tiling in wavelet decomposition	55
3.6	Tree structure for 2-level WP analysis (a) decomposition (b) reconstruction	56
3.7	Decomposition using Wavelet Packet	57
3.8	Coefficient layout of a 3-level DWT of an image.	60
3.9	Pyramid Structure for 2-level 2D DWT computation	60
3.10	Implementation structure for the computation of PSWT and its inverse	64
3.11	Linear Predictor of order p	69
4.1	A distorted signal creating difficulty in period estimation	77
4.2	A music note produced by an electric guitar and its autocorrelation plot.	78
4.3	Signal showing multiple dominant peaks in the Autocorrelation plot. . .	79
4.4	Distortion in the Autocorrelation plot due to dc shift in the signal. . . .	80
4.5	Signals showing wide variation in UDWT levels for period estimation. . .	81
4.6	Surface plot of a signal represented in PS form (Tambura note).	83
4.7	Surface plot of the signal shown in figure 4.6 after Normalization	84
4.8	Block schematic of the predictor estimator.	86
4.9	Block schematic of the PSWT based predictive coder.	88
4.10	Original ECG data segment at lead 1 of the arrythmia data record no.101	92
4.11	UDWT coefficients of the ECG data at different levels	93
4.12	Surface plot of the ECG data shown in fig. 4.10	93
4.13	Surface plot of the amplitude normalized ECG shown in fig. 4.10	94
4.14	Surface plot of the PSWT coefficients of the ECG in fig. 4.10	95
4.15	Reconstructed ECG and the error of prediction	96
4.16	Performance of PSWT based LPC on a segment of male voice sampled at 16kHz-variation with Normalizing factor	100

4.17	Performance of PSWT based LPC on a segment of male voice sampled at 16kHz-variation with Order of Predictor	101
4.18	Performance of PSWT based LPC on a segment of male voice sampled at 16kHz-variation with Level of Decomposition	101
4.19	Performance of PSWT based LPC on a segment of male voice sampled at 16kHz-variation with Wavelet Type	103
4.20	Original data segment from a female voice sampled at 16kHz	103
4.21	Frame-wise autocorrelation plot of the segment shown in fig. 4.20	104
4.22	PSWT based LPC of a typical voiced data segment taken from a female music.	104
4.23	PSWT based LPC of a typical segment taken from a Violin tone	107
4.24	PSWT based LPC of a typical segment taken from a Guitar tone	108
4.25	PSWT based LPC of a typical segment taken from a Harmonium tone	108
4.26	PSWT based LPC of a typical segment taken from a Flute tone	109
4.27	3-Level PSWT of a Violin signal mutilated by Bow noise	112
4.28	3-level PSWT of a Guitar signal mutilated by eccentric striking noise	113
4.29	3-level PSWT of a Harmonium signal mutilated by bellows blow noise	113
4.30	3-Level PSWT of a noisy Flute signal	114
5.1	Simple structure for the parallel computation of DWT & IDWT.	122
5.2	A PMS Computational block for decomposition	125
5.3	PMS Computational block for j^{th} level reconstruction.	127
5.4	The Complete PMS structure for 2-level DWT and IDWT Computation.	128
5.5	Comparison of arithmetic operations between the proposed PMS Structure and the Filter Bank Structure	131
5.6	Relative performance of the PMS and Filter Bank Structures for variable signal and wavelet lengths for a 4-level decomposition.	132
5.7	Relative performance of the PMS and Filter Bank Structures for variable signal and wavelet lengths.	133
5.8	Comparison of PMS structure with filter bank for computation of coefficients at selected levels	135
5.9	Multichannel Filter bank structure for a 3-level 2D DWT Decomposition	138
5.10	PMS structure for M-level 2D DWT decomposition	139
5.11	PMS structure for M-level 2D DWT reconstruction	140
5.12	Comparison of the number of multiplications between the PMS and pyramidal structures for 2D DWT decomposition	144
5.13	Comparison between the 2D pyramidal and PMS structures in terms of the number of multiplications needed for directly computing the details at any desired level	146

5.14	Comparison of computational efficiency between the 2D PMS and Pyramidal Structures for Variable wavelet lengths and data sizes.	146
5.15	Detection of micro-calcification in Mammogram using the PMS structure.	147
B.1	Flow chart showing segmentation and classification of vocal music using WT techniques	166
B.2	Classification of a piece of Classical music sung by a female artist	167

List of Tables

4.1	ECG Records used for compression study	91
4.2	Wavelet Optimization for ECG data compression.	97
4.3	Optimization of level of decomposition for ECG data compression.	97
4.4	Optimization of the Period normalizing factor for ECG data compression.	97
4.5	Optimization of the number of beats needed for computing the predictor for ECG data compression.	98
4.6	Optimization of the predictor order for ECG data compression.	98
4.7	Optimized parameters for ECG data compression	98
4.8	Summary of compression study on different ECG records	100
4.9	Optimized parameters for compression of human voice at 16kHz sampling rate	102
4.10	PSWT based LPC performance on voiced segments of human voice at constant pitch sampled at different frequencies.	105
4.11	Optimized parameters for the PSWT based LPC of few musical signals	107
4.12	Summary of compression study on different musical instrument tones	110
4.13	Details of mutilated music signals considered for illustration.	111
5.1	Comparison of computational efficiency on ECG data of length 10000 in PSWT computation on a small ECG segment.	136
5.2	Comparison of computational efficiency in PSWT computation on longer ECG segment.	136
A.1	Objective performance of wavelets on audio signal processing.	158
A.2	Effect of simple thresholding on audio signal compression.	159
A.3	Effect of change of wavelet on speech compression using simple thresholding.	159

Abbreviations

AASF	-	Average Amplitude Scale Factor
ADPCM	-	Adaptive Differential Pulse Code Modulation
AHA	-	American Hearts Association
ASEC	-	Analysis-by-Synthesis ECG Compression
ATH	-	Auditory Threshold of Hearing
AZTEC	-	Amplitude Zone Time Epoch Coding
BDWT	-	Block Discrete Wavelet Transform
BIH	-	Beth Israel Hospital
CCITT	-	International Telegraph and Telephone Consultative Committee
CELP	-	Code Excited Linear Prediction
CPBC	-	Cycle Pool Based Compression
CW	-	Characteristic Waveform
CWT	-	Continuous Wavelet Transform
DCT	-	Discrete Cosine Transform
DFT	-	Discrete Fourier Transform
DMRA	-	Discrete Multi Resolution Analysis
DSP	-	Digital Signal Processing
DTFT	-	Discrete Time Fourier Transform
DTWT	-	Discrete-Time Wavelet Transform
DWT	-	Discrete Wavelet Transform
DyWT	-	Dyadic Wavelet Transform
ERP	-	Evoked Response Potentials
EZW	-	Embedded Zerotree Wavelet
FHT	-	Fast Hartley Transform
FOS	-	Fast Orthogonal Search
FS	-	Fourier Series
FT	-	Fourier Transform
GBSD	-	Generalized Bark Spectral Distortion.

GCI	-	Glottal Closure Instant
IDWT	-	Inverse Discrete Wavelet Transform
IPSWT	-	Inverse Pitch-Synchronous Wavelet Transform
IWT	-	Integral Wavelet Transform
KLT	-	Karhunen-Loeve Transform
LP	-	Linear Prediction
LPC	-	Linear Predictive Coding
LTP	-	Long-Term Prediction
MBE	-	Multiband Excitation Coding
MBP	-	Mean Beat Period
MIAS	-	Mammographic Image Analysis Society
MIT	-	Massachusetts Institute of Technology
MOS	-	Mean Opinion Score
MPEG	-	Moving Picture Experts Group
MRA	-	Multi-Resolution Analysis
MRI	-	Magnetic Resonance Imaging
MUMS	-	McGill University Master Samples
MWT	-	Multiplexed Wavelet Transform
NMAE	-	Normalized Maximum Amplitude Error
NRMSE	-	Normalized Root Mean Square Error
OSC	-	Overlap-Save Convolution
OZWC	-	Optimal Zonal Wavelet Coding
PAN	-	Period and Amplitude Normalization
PCM	-	Pulse Code Modulation
PET	-	Positron Emission Tomography
PMS	-	Parallel Multiple Subsequence
PRD	-	Percentage Root Mean Square Difference measure
PS	-	Pitch-Synchronous
PSWT	-	Pitch-Synchronous Wavelet Transform

PWI	Prototype Waveform Interpolation
QMF	– Quadrature Mirror Filter
RDWT	– Redundant Discrete Wavelet Transform
REW	– Rapidly Evolving Waveform
RMS	– Root Mean Square
SC	– Sub-band Coding
Seg-SNR	– Segmental Signal-to-Noise Ratio
SEW	– Slowly Evolving Waveform
SIDWT	– Shift Invariant Discrete Wavelet Transform
SNR	– Signal-to-Noise Ratio
SPIHT	– Set Partitioning In Hierarchical Trees
STC	– Sinusoidal Transform Coding
STE	– Short-Time Energy
STFT	– Short-Time Fourier Transform
SVD	– Singular Value Decomposition
SWT	– Stationary Wavelet Transform
SZR	– Short-Time Zero Crossing Rate
TC	– Transform Coders
TF	– Time-Frequency
UDWT	– Undecimated Discrete Wavelet Transform
VLP	– Ventricular Late Potentials
WDD	– Weighted Diagnostic Distortion
WHOSC	– Wavelet Transform Higher Order Statistics-based Coding
WI	– Waveform Interpolation
WMWI	– waveform-matched waveform interpolation
WOSC	– Wavelet Overlap-Save Convolution
WP	– Wavelet Packet
WPT	– Wavelet Packet Transform
WT	– Wavelet Transform

-
- ZCR - Zero Crossing Rate
ZEP - Zero-Crossing-Energy Product

Chapter 1

Introduction

This chapter gives a brief introduction to the topic of research work undertaken. The importance of signal processing in this digital era is presented followed by a brief introduction to one of the modern non-stationary signal processing tools, viz. the wavelet analysis, and its computation. Other relevant areas such as the linear predictive coding and the signal compression technology are also introduced. Finally the motivation for the research work carried out is presented along with a brief layout of the thesis.

Over the past several decades, the field of Digital Signal Processing has been significantly contributing to the different areas of technology. Among the latest tools in DSP¹, the wavelet analysis has gained much attention due to its capability in non-stationary signal processing. The concept of wavelet analysis has been in place in one form or another since the beginning of 20th century. However, the wavelet theory attracted particular attention in the 1980s through the work of several researchers from various disciplines - Strömberg, Morlet, Grossmann, Meyer, Battle, Lemarié, Coifman, Daubechies, Mallat, Chui - to name a few. Because signal compression had been considered as a major application of wavelets, in many cases, the application of the WT² was regarded synonymous with data compression. Later, DSP has shown a fast growth phase, through its widening applications in various areas of technology.

Wavelet techniques enable us to divide a complex function into several simpler ones and study them separately. This property along with fast wavelet algorithms, makes these techniques attractive in analysis and synthesis problems. Wavelets have gained popularity in different areas like signal/image analysis, medical diagnostics, boundary-value problems, geophysical signal processing, statistical signal processing, pattern recognition, signal/image compression, and many others.

Since most of the man-made and natural signals are non-stationary in nature, unlike Fourier-based analysis, wavelet analysis offers much more compact and easier implementation. Since DWT³ is essentially a sub-band coding system and since sub-band coders have been successful in speech and image compression, it is clear that wavelets find immediate application in compression problems.

In this thesis, the problem of non-stationary signal processing and the techniques for compression of the pseudo-periodic regions in it have been investigated. The computational issues in wavelet analysis have been examined and the development of a novel

¹Digital Signal Processing

²Wavelet Transform

³Discrete Wavelet Transform

structure for DWT computation has been presented.

1.1 Signal Processing

A *signal* is defined as a physical quantity that varies with time, space, or any other independent variable. Signal processing is any operation, that changes the characteristics of a signal.

Conventional signal processing methods mainly include the Fourier techniques named after *Jean Baptiste Joseph Fourier* (1768-1830), a French mathematician and physicist. It transforms the signal in the time or spatial domain to the frequency domain in which many characteristics of the signal are revealed. Most of the signals encountered in the field of science and engineering are functions of a continuous variable such as time or space. Until World War II, *analog* methods played a dominant role in signal processing. The development of the theory of sampled data systems began in 1940's which lead to the development of *digital* signal processing. Eventually, due to the advances in integrated circuit technology, achievements in software engineering and improved algorithms in numerical analysis, the field of DSP experienced rapid expansion. There are several advantages in going for the digital processing of analog signals. These include flexibility, accuracy, storage considerations etc. A good majority of signals associated with physical phenomena are quasi-periodic in nature and has considerable processing advantages over totally non-periodic or random signals.

Depending on whether the signal is continuous, discrete, periodic or aperiodic, four categories of Fourier Transforms exist, *viz.* the FT⁴, the FS⁵, the DTFT⁶, and the DFT⁷. These apply respectively to aperiodic-continuous, periodic-continuous, aperiodic-discrete and periodic-discrete signals [1]. Practically the only type of FT that can be

⁴Fourier Transform

⁵Fourier Series

⁶Discrete Time Fourier Transform

⁷Discrete Fourier Transform

used in DSP is the DFT.

Although unquestionably the most versatile method, Fourier analysis becomes inadequate when the local frequency contents of the signal are of interest or when the signal is non-stationary in nature. To overcome this, a local analysis is needed combining the time and frequency domain techniques, by means of which one can extract the local frequency content of a signal. An elementary scheme in this line is referred to as STFT⁸. By this method, an approximate frequency content of a signal $f(t)$ in the neighborhood of some desired location in time, say $t = b$ can be obtained. This is achieved by first windowing the function using an appropriate window function $\phi(t)$ and then taking the FT of $f_b(t)$, where $f_b(t) = f(t)\phi(t - b)$ is the windowed function. This transform is also referred to as the *windowed FT* or *running window FT*.

Although STFT and its variations are widely used to resolve events in the frequency and time axis, the fixed time-frequency resolution of the STFT poses a serious constraint in many applications. The WT is an advanced time-frequency transform that overcomes the above constraint.

1.2 The Wavelet Analysis

Wavelet Analysis is a powerful concept that has highly influenced the field of applied mathematics and different areas of engineering research. It is the state-of-the art signal processing tool whenever a signal is dominated by transient behavior or discontinuities. Wavelets help in hierarchically decomposing functions. They allow a function to be described in terms of a coarse overall shape and details that range from broad to narrow. They offer an elegant technique for representing the various levels of details present in the signals. Signal characteristics can be efficiently located in the space and frequency domains. Thus, unlike the STFT, wavelets are adequate for the study of non-stationary

⁸Short-Time Fourier Transform

and unpredictable signals with both low frequency components and sharp transitions. The WT is a multiresolutional, multiscale analysis technique which has been shown to be well suited for music processing as well, due to its similarity to the processing of sound by the human ear.

In contrast to a Fourier Sinusoid, which oscillates forever, a wavelet is localized in time. They are functions which last for only a few cycles, and hence the name. They are basically oscillatory functions, which satisfy certain properties. Each wavelet is associated with a scaling function. They serve as excellent mathematical tools in the time-frequency analysis of both one-dimensional and two-dimensional signals. In practical wavelet analysis, this is achieved by representing the signals as a linear combination of scaled (durations) and shifted (positions) versions of the wavelet and scaling functions. Fine-scale wavelets are narrow and brief, and coarse-scale wavelets are wide and long-lasting.

There are various kinds of wavelets. Accordingly, to suit the application, one can choose from among smooth wavelets, compactly supported wavelets, symmetric and non-symmetric wavelets, orthogonal and biorthogonal wavelets etc. It is often a complex task, since there are so many properties like the smoothness, temporal/spatial localization, vanishing moments, frequency localization, symmetry, orthogonality etc., which are to be considered.

A Fourier Transform represents a signal in terms of superposition of sinusoids with different frequencies, the coefficients being a measure of the contributions of these sinusoids at these frequencies. Similarly the WT represents the signal as a sum of wavelets with different locations and scales. The wavelet analysis results in a set of wavelet coefficients, which indicate how close the signal is, to a particular basis function. The wavelet coefficients essentially quantify the strength of contribution of the wavelets at the corresponding locations and scales. The technique has been applied in such diverse fields as digital communication, remote sensing, audio signal processing, biomedical sig-

nal processing, medical imaging, astronomy, entertainment electronics, and numerical analysis.

1.2.1 The Continuous Wavelet Transform

The WT is computed as a correlation measure between the signal and a prototype wavelet function, also called the *mother wavelet*, at different scales and shifts. When the scale and shift parameters are continuous and the signal under consideration is a function of a continuous variable, the corresponding transform is called the *CWT*⁹. Specifically, if we choose the set of dilated and translated functions of $\psi(t)$ defined as $Q = \{(\psi_{a,b}(t), (a, b) \in (0, \infty) \times \mathbf{R})\}$, then the CWT $(W_\psi f)(a, b)$ of the signal $f(t) \in L_2(\mathbf{R})$ will be

$$(W_\psi f)(a, b) = |a|^{-1/2} \int_{-\infty}^{\infty} f(t) \overline{\psi_{a,b}(t)} dt \quad (1.1)$$

where, $\overline{\psi}$ is the complex conjugate of ψ , and $\psi(t) \in L_2(\mathbf{R})$ whose FT must satisfy [2]

$$C_\psi = 2\pi \int_{-\infty}^{\infty} |\omega|^{-1} |\widehat{\psi}(\omega)|^2 d\omega < \infty \quad (1.2)$$

Since the scale factor a is the inverse of the frequency ω , the value $(W_\psi f)(a_0, b_0)$ exhibits the frequency content of $f(t)$ in a frequency interval centered around $\omega_0 = a_0^{-1}$ at a time interval centered around b_0 . The CWT is highly redundant and in a strict sense, it is impossible to compute CWT using a digital computer

1.2.2 The Discrete Wavelet Transform

The CWT maps a signal of one independent variable t into a function of two independent continuous variables a, b . From a computational point of view, this transform is not efficient as it is highly redundant and we have to work on continuous variables. Although the discretized/sampled CWT enables the computation of the transform by computers, the CWT in true spirit is not achieved. The redundancy problem still exists in this

⁹Continuous Wavelet Transform. Sometimes referred as Integral Wavelet Transform (IWT)

sampled version of the CWT. One way to solve this problem of redundancy is to sample the CWT on a two dimensional grid $(a_j, b_{j,k})$, such that $a_j = 2^j$, and $b_{j,k} = k2^j$ to adapt to the scale factor a_j . For this choice of grid, the WT is called the DWT¹⁰ [2], [3], defined as:

$$(W_{\psi}f)(2^j, k2^j) = 2^{-j/2} \int_{-\infty}^{\infty} f(t)\psi(2^{-j}t - k)dt \quad (1.3)$$

DWT is still the transform of a continuous time signal, the discretization being only in the a and b variables. In this sense it is analogous to the Fourier Series and hence it has also been referred to as a continuous-time wavelet series [4], [5].

The DWT is often a tight and non-redundant representation of the signal. It is clear that this transform is not shift-invariant. If we select a redundant and shift-invariant transform, it has several advantages over its decimated counter parts. This type of WT is designated as the UDWT¹¹. The UDWT has been independently developed, for different purposes and under several names, viz. SIDWT¹², SWT¹³, RDWT¹⁴ etc [6], [7], [8]. The key point is that it is redundant, shift invariant, linear and is a better approximation to the CWT. The first and obvious way of computing a UDWT is by simply evaluating the DWT for all shifts. Therefore UDWT calls for increased storage space and computational complexity.

1.3 Wavelet Transform Computation

Processing in the discrete wavelet domain is generally carried out on the DWT coefficients, which are computed using multiband filtering operation. Both sequential and parallel computational structures are in use. The key points of consideration in imple-

¹⁰Discrete Wavelet Transform.

¹¹Undecimated/non-decimated Discrete Wavelet Transform

¹²Shift Invariant Discrete Wavelet Transform

¹³Stationary Wavelet Transform

¹⁴Redundant Discrete Wavelet Transform

menting computational algorithms are the structural complexity, storage requirements, number of arithmetic operations, and elimination of redundant computations.

1.3.1 The Pyramid Structure

For practical computation of DWT coefficients, a fast pyramid algorithm also called the *Mallat's algorithm*, which relates the wavelet function to a set of QMF¹⁵ bank [9], [3] is popularly used. This is a recursive algorithm. It is sometimes referred to as the two-channel sub-band coder as it involves filtering the input signal based on the wavelet function used. This algorithm basically follows a sequential structure and the computation of transform coefficients at any level is achieved by going through all the intermediate levels. Other algorithms in this category include the *àTrous* algorithm, *Vetterli* algorithm [4], etc.

1.3.2 Alternate Structures for WT Computation

There are certain applications in which we need not have to go for the WT coefficients at all levels. In such cases the sequential structure as above is not advisable and hence parallel implementations are often sought. This is a necessity in huge data processing applications as well. In parallel structures, the WT coefficients at each level is computed, as far as possible, directly from the original data itself. The specific structure of these algorithms greatly depend on the nature of the problem addressed and they vary in terms of the degree of parallelism, inter-processor communication, memory requirements etc.

1.4 Wavelet Transform for Pseudo-Periodic signals

A large class of signals that arise from physical systems are oscillatory in nature, though they are not periodic in a strict mathematical sense. Examples include voiced regions in speech/music signals, musical tones, ECG signals etc. Moreover, it has been estab-

¹⁵Quadrature Mirror Filter

lished that images can also be treated as pseudo-periodic signals [10]. This reveals the importance of pseudo-periodic signal processing.

In the case of pseudo-periodic signals, the WT can be applied in a modified way making use of the periodicity property. This transform called the Pitch-Synchronous Wavelet Transform was first introduced by G. Evangelista in 1993 [11]. Here the signal is first arranged in the PS¹⁶ form as a 2D¹⁷ data by making use of the local period information. The PSWT¹⁸ is computed using the PS data. It has got several advantages over the conventional DWT *viz.* rate-reduction coding, better feature extraction and noise suppression, bandwidth reduction, etc.

The usefulness of the PSWT technique always depends on the accuracy with which the local periods are evaluated. When no local period is detected, it turns out to be the ordinary DWT. When the local periods are same, the PSWT will turn out to be the MWT¹⁹.

1.5 Linear Predictive Coding and Signal Compression

1.5.1 Linear Prediction

Linear prediction is a particularly important topic in DSP, with application in a variety of areas such as speech signal processing, image processing, noise suppression in communication systems, biomedical signal processing, data rate reduction coding, etc. Predictive coding systems make use of the waveform redundancy to realize straightforward reductions in bit-rate for a specified quality of quantization. Any prediction system involves the evaluation of the optimal predictor, which reflects the signal/source charac-

¹⁶Pitch-Synchronous

¹⁷Two Dimensional

¹⁸Pitch-Synchronous Wavelet Transform

¹⁹Multiplexed Wavelet Transform

teristics, followed by a scheme to predict the signal in terms of the predictor itself. This results in a compact representation of signals for the purpose of efficient transmission and storage.

Predictors based on recent waveform history and time-invariant predictor coefficients leads to a class of coders which constitutes one example of *low-to-medium complexity* designs. On the contrary, *high complexity* predictors are characterized by the use of *adaptive predictors* matched to short-time input spectrum and/or *distant-sample-memory* for utilizing waveform periodicities.

1.5.2 LPC Technique for Data Compression

LPC²⁰ is a popular method employed in data compression applications. These coders usually belong to the analysis-synthesis type. Such coding systems utilize a compact set of parameters in the analysis stage, which are used to encode the original data efficiently. This results in considerable data size reduction for transmission and storage purposes. In the synthesis stage, these parameters are decoded and used in conjunction with appropriate inverse mechanism to reconstruct the original signal. Analysis can be open-loop or closed-loop. In closed-loop analysis, the parameters are extracted and encoded by explicitly minimizing a measure (usually the *mean square*) of the difference between the original and the reconstructed signals.

The above Parametric representations can be speech or non-speech specific. Non-speech specific coders or *waveform coders* are concerned with the faithful reconstruction of the time-domain waveform, whereas, speech specific coders or *voice coders* (vocoders) rely on speech models and are focussed on producing perceptually intelligible speech without necessarily matching the waveform. Hence vocoders are capable of operating at very-low rates whereas waveform coders generally operate at medium rates.

²⁰Linear Predictive Coding

1.6 Motivation for the Work

There has been outstanding contributions from various research groups towards transform-domain-based signal processing as reviewed in chapter 2. LPC based waveform coding is understood to be the best in waveform coding under medium-rate systems. The WT based systems are also found to give superior performance in sub-band coding systems. In the case of pseudo-periodic signals, once the signal is represented in PS form, data sequences across periods are noticed to be very much identical except for minor variations. Due to this, the PSWT, is proved to result in superior performance over conventional WT. This research work combines the relative advantages of LPC techniques in waveform coding and the WT techniques in sub-band coding along with the data reduction capability of PSWT in WT based sub-band coding.

In any transform domain system, computational issues deserve special attention. It is not an exception in WT based signal processing systems also, as state-of-the-art computational structures have the drawback of redundant computations. Hence in this work, reduction of computational burden and complexity in the evaluation of DWT and PSWT coefficients have also been taken up. Development of an efficient computational structure was aimed at to make the coding system attractive for implementation.

The motivating factor is that, this work could contribute considerably to a number of practical systems. Examples include speech processing systems, entertainment and multimedia systems, telemedicine and other modern biomedical systems, signal processing systems handling large volume of data like geological systems, weather forecasting systems etc.

1.7 Layout of the Thesis

This thesis is organized into two major sections. The first section describes a novel compression scheme in which the PSWT technique is uniquely combined with the pop-

ular Linear Predictive Coding technique. The second part deals with the computational issues of Discrete Wavelet Transform and the development of a new structure for its efficient implementation.

The remaining portion of the thesis is divided into 5 more chapters.

Chapter 2 gives the *literature review*. Here, an account of the previous research work that has been carried out in the related field by peer researchers is presented. The present work is related to various fields in signal processing including WT based signal analysis, period estimation methods, PSWT based signal processing, Linear Predictive Coding and compression, computational structures for 1D and 2D DWT, etc. Hence an elaborate account of the recent developments and state-of-the-art in such topics and related areas have been incorporated in this review chapter.

In Chapter 3, a brief description of the *introductory theory* is given. Topics included are Non-stationary signal processing, theory of MRA²¹ in discrete domain, basic theory of various types of WT, multi-rate techniques in DSP, theory of PSWT and LPC, computation of 1D and 2D wavelet transform etc. This chapter serves as a background for the work presented in chapters 4 and 5.

Chapter 4 describes the *PSWT based method developed for the Linear Predictive Coding and Compression* of pseudo-periodic signals, and its typical applications. Initially, the method has been presented in detail for *general* pseudo-periodic signal processing applications. Subsequently, the application of the compression scheme on typical signals have been presented as case studies. Some techniques for feature enhancement and source dependent noise suppression in PSWT domain also have been explained. The method has been evaluated in terms of standard performance measures. The chapter is concluded highlighting the results and discussing the important findings of the study.

The *development of the new computational structure* called the Parallel Multiple Subsequence (PMS) structure has been presented in chapter 5. It has been systemat-

²¹Multi Resolution Analysis

ically derived for 1D and 2D DWT computation. This is followed by the analysis of results and a description of its effectiveness in different signal processing applications. The performance of this structure for variations in data size, level of decomposition, and wavelet size are analyzed and compared with that of the popular pyramid structure. To highlight the computational efficiency of the PMS²² structure, Cases of complete wavelet decomposition and that of direct decomposition to arbitrary levels are separately presented. Application of the structure in practical situations like PSWT computation of ECG data, and edge detection in mammogram have been described as case studies. Important results and conclusions drawn thereof in using the PMS structure are given at the end of this chapter.

A *brief summary* of the research work conducted and the important *conclusions* thereon are highlighted in chapter 6. The scope for future work, as an extension of the present study, is mentioned towards the end of this chapter.

This thesis includes *two* appendices which describe some works of the author in the related field.

Appendix A deals with the results of a general study conducted on WT based data compression. The effect of wavelet length, type of wavelet, and level of decomposition in signal compression is presented and compared with that of Wavelet Packets for the same application. Audio data recorded at different sampling rates and storage resolution has been taken up for the study.

In *Appendix B* a study on WT based segmentation of 1D signals is presented. In order to apply the PSWT technique in general signal processing tasks, the signal has to be segmented into different regions and the pseudo-periodic ones are to be identified first. A preliminary study in this direction, employing the WT technique has been performed, taking vocal music as example. The signal could be automatically segmented into voiced, unvoiced, silent and transition regions as depicted in this appendix.

²²Parallel Multiple Subsequence

Chapter 2

Literature Review

A detailed account of the previous work in the field of WT based signal processing, particularly on the topic of PSWT based signal processing is given in this chapter. A brief description of the general work in this field is included. Some important published works in the allied areas such as Linear Predictive coding and WT based signal compression, period estimation techniques etc. are also briefly outlined. Developments in the field of PSWT based signal processing and techniques for period estimation of selected quasi-periodic signals are reviewed. Finally, significant contributions on algorithms and structures for WT computation are accounted.

The field of signal processing has always benefited from a close coupling between theory, applications, and technologies for implementing signal processing systems. The growing number of applications and demand for increasingly sophisticated algorithms goes hand-in-hand with the rapid pace of device technology for implementing digital processing systems. In practical systems it is quite natural that one may want to enhance some signal component or some parameter of a signal model. In communication systems, it is generally necessary to perform preprocessing operations such as modulation, signal conditioning and compression prior to transmission over a channel and then to carry out the corresponding post-processing at the receiver.

An important area in DSP is non-stationary signal processing. Almost all practical signals can be assumed non-stationary when considered for reasonably long duration. Hence conventional signal processing tools such as the Fourier techniques have been replaced by more efficient WT based methods. It is basically a time-frequency method employing multirate techniques. In the WT based processing of pseudo-periodic signals pitch-synchronous analysis, linear predictive analysis and computational issues play a significant role. Almost all solutions developed for 1D signals have a direct counterpart in multi-dimensional systems also, especially in 2D signal processing.

The first book devoted to digital signal processing was written by Gold and Rader [12]. A partial list of books on fundamental concepts and applications that followed include [13], [14], [15], [16], [17], [18], [19], [20] and [21]. More information concerning image analysis can be found in [22], [23], [24], [25] and [26], whereas [27] and [28] are devoted for geophysical and seismic applications. The important field of speech analysis and synthesis is the subject of [29], [30] and [31]. Feature enhancement is one among the common applications of signal processing [32], [33], [34], [35], [36].

2.1 Linear Predictive Coding and Signal Compression

The goal of any compression scheme is to achieve transparent compression, where the output signal is as perceptually similar to the original signal as possible. Many of the compression algorithms are similar, with variations existing primarily in the method of implementation. A fundamentally new approach to dynamic range compression based on a critical band multichannel structure, incorporating the attack and release rates, a level estimate mode control, and a normalization of the level estimates across the frequency bands is presented by Schmidt *et al* [37].

A CELP¹ based audio coding system which uses filter banks to decompose in the frequency domain, into constant width sub-bands is described in [38]. In order to obtain a high audio quality, they have used psycho-acoustic models. They could achieve excellent audio signal quality at bit rates of 50-60 kbit/s. Etemoğlu *et al.* has recently presented a sinusoidal speech model for low bit rate speech coding, where parameters of the model are extracted by a closed loop analysis based on matching pursuits [39]. The sinusoidal modelling of the speech LP residual is performed within the general framework of matching pursuits with a dictionary of sinusoids. The authors claim to have achieved a quality exceeding the 6.3kbps G.723.1 coder with a 4kbps matching pursuits sinusoidal speech coder.

The need for compressing biomedical signals is important due to the tremendous amount of data that need to be stored efficiently at low cost. As an example, consider an ambulatory or Holter ECG recording system [40]. The recent technological developments have made possible the recording of ambulatory ECG signal in digital form into solid state memory. However, memory requirement for say 24 hours of ECG monitoring are prohibitive (i.e., about 52 megabytes for two-channel, 12-bit resolution and 250 Hz sampling rate). For efficient storage of such large data records, effective data compression methods are of interest. The desired objective is to provide a high quality

¹Code Excited Linear Prediction

reconstruction of ECG signals at low bit rates and acceptable distortion levels. Solutions to the compression of ECGs resulted in different techniques, roughly two categories can be identified.

- *Dedicated techniques:* These are mainly time-domain techniques and were developed only for the compression of ECG signals. They include the following:
 - Heuristic algorithms, like the amplitude zone time epoch coding (AZTEC) [41], turning point (TP) [42], coordinate reduction time encoding system [43], FAN algorithm [44] and improvements to time-domain algorithms such as SLOPE [45] and AZTDIS [46].
 - Optimization algorithms, like long-term prediction (LTP) [47], [48], analysis by synthesis ECG compressor (ASEC) [49], and the cardinality constrained shortest path technique [50], [51].
- *General Techniques:* These can be used on a wide range of signals including speech, image, and video signals. They include differential pulse code modulation, sub-band coding (SC) [52], [53], [54], transform coding [49], [55], [56], and vector quantization [57].

A novel application of SVD² in data compression of ECGs is presented in [58]. Here, the quasi-periodic analysis of SVD was exploited to decompose an ECG sequence into a linear combination of a set of basic patterns with associated scaling factors. As done in [59], the beat information is obtained first, followed by period normalization. Consequently, the set of ECG segments are rearranged into a two-dimensional matrix and it is decomposed using SVD transformation. It has been shown that the information of the ECG signals will mostly be concentrated within a few dominant singular triplets, so that the balance can be suitably discarded resulting in compression. The performance

²Singular Value Decomposition

of this scheme is claimed to be superior in comparison with the existing schemes [55], [59], [60], [61], [62], [63].

A new type of ECG data compressor is presented in [64]. By differentiating the ECG signal and using proper thresholding, the ECG is first segmented into a sequence of straight lines, whose vertices are used to encode the signal. The decoder part applies the FOS³ method to reconstruct the signal from the partial data generated by the encoder. The compression ratio achieved using this method is very high compared to other methods, whereas the reconstructed signal quality is slightly inferior. The authors claim that this method preserves the diagnostic information contained in the original signal even after the transformation. A time domain algorithm based on the coding of line segments which are used to approximate the signal is presented by Nygaard *et al* [65]. Though applicable to any type of signals, the authors have illustrated the compression of ECG, and have compared the compression performance with other methods.

An analysis by synthesis ECG compression method has been developed by Zigel *et al* [66]. This scheme consists of a beat code book, long and short-term predictors, and an adaptive residual quantizer. A distortion measure has been defined in order to efficiently encode every heartbeat. The algorithm has been validated with the MIT-BIH⁴ Arrhythmia Database and it has been reported that they could achieve a mean compression rate of approximately 100bits/s (compression ratio of about 30:1), with a good reconstructed signal quality (WDD⁵ below 4% and PRD⁶ below 8%). They have conducted a MOS⁷ test also to validate the performance of the algorithm, in which the testers selected were expert cardiologists. It has been reported that for good diagnostic quality, each ECG lead should be sampled at a rate of 250-500 Hz with 12 bits resolution.

³Korenberg's Fast Orthogonal Search

⁴Massachusetts Institute of Technology- Beth Israel Hospital

⁵Weighted Diagnostic Distortion measure

⁶Percentage Root Mean Square Difference measure

⁷Mean Opinion Score

A family of CPBC⁸ has been described by Barlas *et al* [67]. This is found to be capable of encapsulating other template-matching algorithms and thus unify all similar compression approaches, to a certain extent.

The performance of a speech processing system is normally evaluated using MOS⁹, which is a formal subjective measure of received speech quality [68]. Generally, coding quality with MOS higher than 4 is considered as *toll-quality*, between 3.5 and 4 as *communication quality*, between 3 and 3.5 as *professional quality*, and below 3 as *synthetic quality* [69], [70]. Low rate speech coding and compression attempts to provide toll-quality speech at a minimum bit rate for digital transmission or storage. The trade-offs which depend on the particular coding technique and the application, are coding delay and distortion, and increased cost of equipments. The early years saw the standardization of 64 kbits/s PCM¹⁰, but research into more complex lower bit rate coding schemes was initially inhibited by practical implementation considerations imposed by the semiconductor technology of the day. As a consequence, research into sophisticated low bit rate algorithms were delayed and it was in 1984 that the first world-wide lower bit rate coding standard was achieved. This standard was the CCITT¹¹ G.721 recommendation for 32 kbit/s ADPCM¹². Since that time, the major advances made in microelectronics and DSP¹³ technology have spurred research into increasingly complex speech coding methods. Accordingly, low bit rate speech coders came into existence.

Recent research show that speech interpolation combined with noise masking speech coding techniques can achieve toll-quality performance at 4kbits/s [71]. When coding speech below 4 kbit/s, a representative waveform segment can be extracted regularly and efficiently coded; the signal between the coded segments can then be regenerated

⁸Cycle Pool Based Compression

⁹Mean Opinion Score

¹⁰Pulse Code Modulation

¹¹International Telegraph and Telephone Consultative Committee

¹²Adaptive PCM

¹³Digital Signal Processing

via interpolation. Among the techniques actively studied in recent years for very low bit rate speech coding are STC¹⁴ [72], [73], MBE¹⁵ [74], [75], and PWI¹⁶ [76], which use time-domain interpolation.

2.2 WT based Signal Processing

The beginning of the WT as a specialized field can be traced to the work of Grossman and Morlet [77]. The early works were related to the CWT. The term *wavelet* can also be found in the seismic signal processing literature in a context other than the WT [78]. A historical account of the WT is found in [79] and [80]. A tutorial presentation of theory and applications of WT is found in [5], [81], [9], [82], [83] and [84]. [85] and [2] are some of the good references for a rigorous mathematical analysis of wavelets. The contributions of WT in the field of electrical engineering is remarkable. Other fields also have equally benefitted from this field. Study of matter in universe [86], study of tropical convection [87], transient study in underwater acoustics [88], geo-acoustic data compression application [89], turbulence study in fluid mechanics [90] etc., are a few interesting general application areas. A number of works in the field of seismic signal processing is seen in the special issue of IEEE [91].

The field of Electrical engineering has already marked wide scope for WT based signal processing related to many of its allied application areas. One among them is the processing of pseudo-periodic signals like biomedical signals, speech, music etc. Due to the wide variety of signals and problems encountered in medicine and biology, the spectrum of applications of DWT has been extremely large. The main difficulty in dealing with biomedical objects is the extreme variability of the signals and the necessity to operate on a case by case basis. The application ranges from the analysis of the more traditional physiological signals such as the ECG, to the very recent imaging modalities

¹⁴Sinusoidal Transform Coding

¹⁵Multiband Excitation Coding

¹⁶Prototype Waveform Interpolation

including the PET¹⁷ and MRI¹⁸. The various uses of the WT and the corresponding research efforts in the biomedical area has been reviewed by Unser *et al* [92]. The wavelet properties that are most important for biomedical applications, and an analogy between the WT and some of the biological processing that occurs in the early components of the auditory and visual system have also been presented.

The most important biological signal showing pseudo-periodicity is the ECG. Earlier works of wavelet-based studies of ECG signals have focussed mainly on examinations of heart rate variability, the classification of ECG waveforms, or ECG data compression. But in present studies WTs are being utilized also for diagnostic purpose, feature enhancement etc. Detection of VLP¹⁹ [93], [94], [95], [96], wavelet-based feature extraction for discriminating between normal and abnormal cardiac patterns [97], study of heart rate variability in humans [98], [99], [100], [101] etc. are a few important works on ECG. A few promising applications of WT are described in the work [102] by Provaznik *et al*. It has been observed that cardiac dysfunction can be better discriminated based on study using the scale-dependent WT standard deviation and the corresponding spectral measures [100], [101], [103], [104]. It has been reported [105] as an exciting observation that previously unsuspected coordinated mechanical activity in the heart during ventricular fibrillation may be made visible in the surface ECG using the WT.

The use of WT in speech signal processing is manifold. Signal enhancement, period estimation, signal compression, signal classification, noise suppression etc. are important tasks in this field. It was in 1995 that the wavelet thresholding (shrinkage) was introduced as a powerful tool in denoising signals degraded by additive white noise [106]. Later on several works are reported for speech enhancement applying the principle of wavelet shrinkage [107], [108], of course with many problems to be resolved for its suc-

¹⁷Positron Emission Tomography

¹⁸Magnetic Resonance Imaging

¹⁹Ventricular Late Potentials

successful application in various stages. Removing the noise components by thresholding the wavelet coefficients is based on the observation that in many signals (like speech), energy is mostly concentrated in a small number of wavelet dimensions. The coefficients of these dimensions are relatively large compared to other dimensions or to any other signal (especially noise) that has its energy spread over a large number of coefficients. Due to the threshold adaptation, the modified method proposed in [109] is able to cope better with colored and non-stationary noise types.

The task of *music signal processing* demands more stringent requirements for TF²⁰ distributions when compared with the weaker requirements found in speech analysis. The major time and frequency methods that have been applied to musical signals are traced and the application areas are described in [110]. The techniques are examined in the context of Cohen's class²¹ [111]. In this paper, the impact of different analysis methods on pitch and timbre examination is shown which spans the Fourier Series and Transform, Pitch-Synchronous analysis, Heterodyne filter, STFT, Phase vocoder, constant-Q and Wavelet Transforms, the Wigner distribution, and the modal distribution. Some of the interesting works on WT based music signal processing are seen in [112], [113].

In general WT based signal processing applications, time-variance characteristics of non-redundant wavelet transforms causes serious limitations. Addressing this issue, potential applications based on shift-invariant wavelet transform is suggested in [89], [114], [8], [115], [116].

2.2.1 WT based Signal Compression

Signal compression is vital in massive data processing environment and WT has been extensively used in this application with different types of signals, due to its multireso-

²⁰Time-Frequency

²¹named after *L.Cohen*, one of the pioneers in this field, who has realized an infinite class of Time Frequency distributions, that could be obtained from the Wigner-ville distribution by linear transformation

lution capability. An elementary study on audio compression using orthogonal wavelets and wavelet packets using simple thresholding technique has been presented in [117]. A number of works are seen in this field making use of the perceptual characteristics of the auditory system [118], [119]. Painter *et al.* [119] reviews several different classes of audio coders and remark that all the current algorithms and the standards in particular are either high-resolution or low-resolution *sub-band* coders. They give a good collection of resources for research in this direction. An analysis-synthesis method of audio compression has been presented by Sathidevi *et al* [120]. The authors have made use of the perceptual characteristics of the human auditory system and its resemblance with WT, in developing a scheme for audio compression using WT and WPT. A comparison of WT based amplitude compression with other conventional audio compression schemes has been presented in [121].

A compression scheme employing vector quantization with WT has been developed by the Georgia Institute of Technology, Atlanta [122]. It is a hybrid subband-transform coding scheme which employs the perceptual masking properties to achieve an efficient reduction in bit rate. The coder has been applied with speech, music, and more diverse signals consisting of speech in the presence of eventful background sounds. Another wideband speech coder using WT in which each of the wavelet based subbands are separately coded using ADPCM²² technique and CELP with dynamic bit allocation has been presented by Li *et al.* [123] resulting in a bit rate of 12.8 kbits/s.

The WMWI²³ technique as suggested in [124], achieves an accurate representation of speech evolution by extracting consecutive pitch-periods of a time-warped, constant pitch residual. This pitch track optimization technique is found to ensure that the critically sampled pitch periods can be effectively decomposed into a slowly evolving and rapidly evolving waveforms, allowing efficient quantization. These type of WI²⁴

²² Adaptive Differential Pulse Code Modulation

²³ Waveform-Matched Waveform Interpolation

²⁴ Waveform Interpolation

coders are able to achieve high quality speech at low bit rates by using a decomposition motivated by human perception [125]. In this method, the warping operation removes the pitch variations of the linear prediction residual signal to enforce a constant pitch period. An unwrapping procedure is performed to reconstruct the residual, where a perfect reconstruction is said to be possible if the pitch track is accurately transmitted.

Bradie [60] has made a preliminary investigation of a wavelet-packet based algorithm for the compression of single lead ECG. This algorithm combines the flexibility and efficiency of WP expansions with the methodology of the KLT²⁵. For selected records from the MIT-BIH database, this algorithm could achieve a compression ratio of 21.4:1. Wavelet and waveletpacket-based compression based on the EZW²⁶ coding is presented by Hilton [63]. The author has analyzed the suitability of different types of wavelets for this application. An application oriented ECG data compression scheme for desired reconstruction quality has been developed by Miaou *et al* [126]. They make use of the SPIHT²⁷ compression technique proposed by Lu and team [127].

The compression performance and characteristics of two wavelet coding compression schemes of ECG signals suitable for real-time telemedical applications is presented in [128]. The two methods proposed are OZWC²⁸ method and the WHOSC²⁹ method. The WHOSC method employs higher order statistics and uses multirate processing with the autoregressive higher order statistics model technique to provide increasing robustness to the coding scheme. The OZWC algorithm used is based on the optimal wavelet-based zonal coding method developed for the class of discrete *Lipschitzian* signals. Both methodologies were seem to have evaluated using the NRMSE³⁰ and the average CR and the bits per sample criteria, applied on abnormal clinical ECG data samples selected

²⁵Karhunen-Loeve Transform

²⁶Embedded Zerotree Wavelet

²⁷Set Partitioning In Hierarchical Trees

²⁸Optimal Zonal Wavelet Coding

²⁹Wavelet Transform Higher Order Statistics-based Coding

³⁰Normalized Root Mean Square Error

from the MIT-BIH database and the Creighton University Cardiac Center database.

A work in the similar line, but employing 2D transforms is presented by Lee *et al* [129]. This method utilizes the fact that ECG signals generally show two types of redundancies - between adjacent heart beats and between adjacent samples. So heart-beat data sequences are cut and beat-aligned to form a 2D data array and 2D DCT³¹ method of transform coding is employed to achieve compression. This could exploit the sample-to-sample and beat-to-beat correlation. The study was based on the MIT-BIH arrhythmia and Medtronic databases. They claim to have achieved superior performance in using 2D transforms compared to 1D methods in general and the performance of 2D method is more superior for higher compression ratios. In comparison with earlier works reported in this direction [47], [130], in this work the beats are normalized to enhance the intercycle correlation. The type of preprocessing used in this method converts the ECG data into a near-cyclostationary sequence and enables the uniform choice of wavelet coefficients to be retained in each beat. The earlier wavelet based compression schemes proposed by Thakor and Bradie [60], [131], have not attempted this. They claim to have achieved a mean transmission rate of 180 b/s (for the tested data) with no compromise on the fidelity of reconstruction.

Rajoub has proposed a WT based ECG data compression algorithm [132] to claim a compression ratio of 24:1 for MIT-BIH record 117 with a root mean square difference as low as 1.08%. Another work combining the LPC techniques and DWT technique for coding high quality digital audio signals is seen reported in [133]. In this method, a linear predictor is first used to model each audio frame. Then the prediction error is analyzed using the DWT. The LPC coefficients and DWT coefficients are quantized using a novel bit allocation scheme which is claimed to be capable of minimizing the overall quantization error with respect to the masking threshold. They have claimed a near-transparent audio signal quality at encoding bit rates of 90-96 kb/s, which is

³¹Two Dimensional Discrete Cosine Transform

comparable to that of the MPEG³² layer II codec operating at 128 kb/s. The codec has been designated as LPC-DWT³³ codec. The objective measurements used for evaluating the performance were the Seg-SNR³⁴ and the GBS³⁵ measure [134].

2.2.2 PSWT based Signal Processing

WT has shown outstanding contributions in the field of pseudo-periodic signal processing. The signal can be represented in the *pitch-synchronous* form based on the local pitch-periods, which can be considered as a 2D representation of the 1D signal [135]. A new wavelet representation for this class of signals, called the PSWT has been explored in detail by *Evangelista* [11]. This is a noteworthy work on this topic. The transformation is applied on the pitch-synchronous vector representation of the signals. Compared to more classical representations, the author claims several potential advantages of PSWT. In this exploratory paper he has offered a panorama of prospective applications and reported the results of a few experiments performed on speech and music signals. As an extension of wavelet theory to pseudo-periodic signals, *Evangelista* has introduced yet another class of wavelet transform, called the MWT [136]. This transform has interesting extension in the field of image processing also. This transform is applicable for signals whose period is constant for the duration under consideration.

An efficient and accurate pitch-synchronized spectral analysis scheme to obtain the Fourier coefficients of a harmonic signal, sampled at an arbitrary rate above the Nyquist critical rate, has been outlined by *Medan et al* [137].

Speech coders based on the WI³⁶ paradigm allow efficient compression of signals by exploiting the perceptual importance of speech characteristics [138]. In state-of-

³²Moving Picture Experts Group

³³Linear Predictive Coding-Discrete Wavelet Transform

³⁴Segmental Signal-to-Noise Ratio

³⁵Generalized Bark Spectral Distortion. This is a perceptually motivated objective measure which compares the spectral difference between the original and coded signals using the Bark scale.

³⁶Waveform Interpolation

the-art WI coders, pitch-cycle waveforms [CW³⁷] are extracted from the LP³⁸ residual, aligned, and then filtered in the evolution domain to decompose the signal into a SEW³⁹ characterizing the voiced speech and a REW⁴⁰ representing noise-like unvoiced speech. Chong *et al* [139] has described a PSWT based CW decomposition method for the WI paradigm. This method has the feature of providing additional scalability in quantization than the other WI decomposition to meet desired quality requirements. The CW surface is decomposed into a series of reduced time resolution surfaces, which are quantized exploiting their perceptual importance and inherent transmission rate requirements.

Chong *et al.* has proposed a scheme in which the PSWT is used as an alternative characteristic waveform decomposition method for the WI paradigm. To meet desired quality requirements, this method provides additional scalability in quantization. Here, the PSWT is implemented as a quadrature mirror filter bank and decomposes the characteristic waveform surface into a series of reduced time-resolution surfaces. The quantization of these surfaces are made efficient by the perceptual importance of these surfaces and inherent transmission rate requirements. This method is established as one among the best for high quality speech storage applications. The exploitation of the perceptual importance of speech characteristics has been studied by Kleijn *et al* [138], [140]. The PSWT when applied to WI has offered significant advantage due to its perfect reconstruction properties and multi-scale decomposition of the evolving CW surface.

Another technique for the analysis and synthesis of pseudo-periodic signals based on a special kind of multiwavelet transform, the *harmonic-band wavelet transform*, has been introduced by Polotti *et al* [141]. The idea is inspired by the fact that pseudo-periodic signals from many physical and biological systems as well as man-made phenomena such

³⁷Characteristic Waveform

³⁸Linear Prediction

³⁹Slowly Evolving Waveform

⁴⁰Rapidly Evolving Waveform

as music signal, variations in traffic flow, economic data and fluctuation of pitch in music etc. show $1/f$ behavior [142], [143]. The technique presented in [141] allows one to control a highly complex stochastic process by means of relatively few parameters. The presence of non $1/f$ noise components due to external sources restricts the usefulness of this technique.

Even though the general signal compression studies have gone a long way, the combined potential of WT and LPC in signal compression is not much exploited. The only noteworthy work in this field is the work by Ramakrishnan *et al* [59]. In this the authors could achieve considerable gain in signal compression, than state-of-the-art compression schemes. They have applied LPC techniques on PS data after taking the WT for each period wise normalized data. To enhance the compression ratio, they have suggested a method to identify the significant coefficients which alone need to be transmitted to the decoder end. Information regarding the period and amplitude information of each beat also have been suitably transmitted. This is a reasonably complex method comprising of amplitude normalization, period normalization, DWT computation, LPC compression etc. Aliasing effect may occur due to improper design of interpolation filter and selection of sampling rate.

2.2.2.1 Local Period Estimation

The success of PSWT based methods highly depends on the accuracy with which the local periods are estimated. Unfortunately it is difficult, as testified by the hundreds of different pitch tracking algorithms that have been developed [144]. It plays an important roll in many speech and signal compression schemes [145]. Pitch detection algorithms can be classified in to two separate categories.

- *Spectral-domain pitch detectors*, which estimate the pitch period of a signal directly using windowed segments of speech, applying a *Fourier*-type analysis to determine a pitch average. They include Cepstrum [146], Maximum Likelihood [147], and

Autocorrelation methods [148].

- *Time-domain pitch detectors*, which estimate the pitch period by determining the GCI⁴¹ and measuring the time period between each *event* [149], [150].
- *Time-frequency based pitch detectors*, which is a combination of time-domain and frequency-domain methods. Wavelet based pitch detectors fall in this category.

One of the best known pitch tracking algorithms, and one against which other methods are often compared, is the Gold-Rabiner Scheme [151]. This algorithm was designed for speech applications, and performs over a range of input frequencies from 50Hz to 600Hz. McNab *et al.* [152] has used this scheme for implementing a pitch determination algorithm for Melody Transcription in music signal processing applications. They have modified the Gold-Rabiner scheme for faster implementation and also for higher frequency applications (up to 1000 Hz) like music signal processing. A comparison of seven popular pitch detection algorithms is given by Rabiner *et al* [153]. The general problems in pitch detection highlighting the importance of the subject also is given in this literature.

Kadambe *et al* [154], has proposed a wavelet based pitch determination algorithm, based on Mallat's work on images [155]. Mallat showed that when analyzing images, the use of wavelet functions with derivative characteristics produces maximums in the wavelet transform across many coincident scales along sharp edges. Kadambe *et al* used the assumption that when a GCI occurs in a speech waveform, maximums also occur in the adjacent scales of wavelet transform. In contrast with this, which chooses maximums if they occur in two adjacent wavelet coefficient scales, a DWT based pitch detection and speech segmentation method has been presented by Wendt *et al.* [156], in which a single derivative filtering function is defined to contain a specific bandwidth of voiced speech. They have shown that this wavelet function when convolved with a speech

⁴¹Glottal Closure Instant

signal will produce a filtered signal containing well defined local maxima where GCI occur in the speech signal. This method provides a dramatic simplification in processing, utilizing one convolution only to analyze and is claimed to be robust to noise. [157] also presents more or less a similar approach for pitch determination, giving more emphasis on musical signals, where there are some difficulties in estimating the pitch period as musical signals have broader range of frequencies [158]. This algorithm is based on DyWT⁴². A comparison with the autocorrelation method also has been given in this paper. Another pitch estimation work using the GCI detection based on WT has been presented by Limin *et al* [159].

Detecting the QRS complexes in the ECG is one of the most important tasks that need to be performed in any ECG processing environment. This information, which is similar to the pitch in speech signals, will help in performing accurate beat-by-beat processing. Li *et al* has presented a wavelet based approach for the detection of QRS complexes [160]. The interval measurement from digitized ECG usually contains an error due to the finite sampling frequency which may jeopardize the beat-to-beat analysis of the signal. Mario *et al* [161], has attacked this issue and have developed a model to describe and quantitate this error. The accuracy of the model has been tested by a simulation procedure. It has been shown that the statistics of the error are only function of the sampling frequency of the ECG. Closed form equations have been obtained for the variance, the autocorrelation function, and the power spectral density of the total error. They have proved that, for a good quality ECG which should be standard for any lab, the error due to the sampling frequency substantially contributes to the overall one, and it becomes dominant in case of a very low sampling frequency like 128 Hz which usually gives an acceptable SNR for various biomedical applications.

[162] and [163] describes some initial works on the subject of ECG beat detection. A new algorithm based on multirate DSP technique to detect beat periods in

⁴²Dyadic Wavelet Transform

ECG signal has been developed by Valtino *et al* [164]. This algorithm incorporates a subband filtering technique. It is claimed to have a sensitivity of 99.59% and a positive predictivity of 99.56% against the MIT/BIH database which is comparable with other algorithms reported in the literature. This can be categorized as a real-time algorithm since it has a minimal beat detection latency.

An algorithm based on WT has been developed by Cuiwei *et al.* [165] for detecting ECG characteristic points. Using this method the authors claim that the detection rate of QRS complexes is above 99.8% for the MIT/BIH database, even with serious baseline drift and noise.

A dyadic WT based QRS complex detector which is robust to time-varying QRS complex morphology has been developed by Kadambe *et al* [166]. The performance of the method has been illustrated with 70 hours ECG data from AHA⁴³ data base and has concluded that this method compared well with all standard techniques.

The application of the fuzzy neural network for ECG beat detection is dealt in [167]. The hybrid fuzzy neural network applied in the solution consists of a fuzzy self-organizing subnetwork connected in cascade with the multilayer perceptron, working as the final classifier. The method could be used for the classification of different types of heart beats.

Friesen *et al* [168] gives an exhaustive comparison of the noise sensitivities for nine different QRS detection algorithms. Synthesized ECG corrupted with five different types of synthesized noise including electromyographic interference, 60Hz powerline interference, baseline drift due to respiration, abrupt baseline shift, and a composite noise comprising all other type of noises, has been used for the study. Different algorithms performed in different style for different noise types. The results of this study is found to be helpful in the development of robust clinical instruments.

⁴³American Heart Association

A class of algorithm based on nonlinear transforms derived from multiplication of backward differences have been developed and evaluated with AHA ECG data base [169]. In comparison with other state-of-the-art methods, the authors present this algorithm to have desirable accuracy *vs* response time tradeoff for real-time microprocessor-based implementation.

The use of WT for ECG characterization in general has been presented in the tutorial paper by Sahambi *et al* [170]. It describes a real-time system that uses WT to overcome the limitations of other methods of detecting QRS and other important time domain features of ECG.

The design and test results of an ambulatory QRS detector is described in [171]. The QRS detection timing accuracy and reliability of this system was tested with an artificially generated ECG corrupted with various noise types and they could achieve a timing accuracy of less than 1ms.

2.3 WT Computation

Many of the WT based methods use complex procedures and needs stringent computations. To exploit the maximum benefit out of these methods, efficient computational structures and algorithms also have been developed. The CWT is often evaluated at dyadic scales. The *Shensa algorithm* [7] is a fast algorithm to compute CWT. This algorithm is defined by a pair of filters. An iterative and efficient method for the design of optimum filters has been presented by Chan *et al.* [172] to implement CWT computation using the Shensa algorithm.

2.3.1 1D DWT Computation

Several efficient algorithms have been derived for the implementation of DWT. Many researchers have addressed comparison of these algorithms based on computational complexity. The most popular algorithm for the computation of DWT is the filter bank tree

algorithm suggested by Mallat [3], [9]. This algorithm has its root in the Pyramid algorithm of Burt and Adelson [173]. Chapter 3 of [79] presents the pyramid algorithm from a historical perspective.

It is customary to view the DWT as a prediction error decomposition which has led to a new framework for DWT design under the name *lifting scheme* [174], [175]. Claypoole *et al.* [176] use lifting to design customized DWTs that adapt to match the signal under consideration. This has resulted in two adaptive lifting constructions which are invertible - *scale-adapted transforms* and *space-adapted transforms*. These schemes have proved their potential in various wavelet transform based applications including signal compression [177].

The pyramid scheme has been compared against the lifting scheme by Daubechies [178]. Rioul and Duhamel [4] discuss several other schemes to improve the standard algorithms. They suggest an FFT based algorithm known as Vetterli algorithm [179] for filters having large support and a "Fast Running FIR" algorithm [180] for short filters. Vetterli has suggested a method for mapping long running convolution into smaller ones by using multirate filter banks based on aperiodic convolution algorithm and STFT. This approach gives good tradeoffs among computational complexity, system architecture and input-output delay.

Viswanath *et al.* has given a VLSI⁴⁴ architecture for DWT implementation [181]. An efficient scalable parallel implementation of the DWT based on the pipeline processor methodology has been presented by Sava *et al.* [182], [183]. It is found attractive for large data size and long filters. According to Westenburg *et al.* [184], a Fourier domain implementation of the wavelet transform is more efficient than the direct computation, especially when the decomposition/reconstruction filters have large support. The single-rate VLSI architecture developed by Denk *et al.* [185] for the one dimensional orthonormal DWT makes two contributions. It is shown that the architectures that are

⁴⁴Very Large Scale Integrated Circuit

based on the QMF lattice structure require approximately half the number of multipliers and adders than corresponding direct form structures. Also, they have developed techniques for mapping the 1D orthonormal architecture to folded and digit-serial architectures which are based on the QMF lattice structure. Chandran *et al.* [186] present an FHT⁴⁵ based implementation of the filter bank structure for DWT and show that noticeable overall computational savings can be obtained over the conventional FFT-based implementation. This method claims other advantages like identical forward and inverse transforms, and also minimum storage requirement, as the outputs are all real-valued.

In real-time signal processing applications, the WT computation need to be performed on a frame-by-frame basis. In such cases, employing the above algorithms *per se* results in undesired artifacts at the ends of each frame. To resolve this difficulty, many researchers have come out with BDWT⁴⁶ algorithms [187], [188]. These algorithms are found to have difficulties related to large storage buffer requirement and frequent inter-processor communication, making it less optimized in real-time applications.

Parallel computation is particularly helpful for applications where wavelet coefficients at selected levels alone are required [189], as we need not perform unnecessary computations by way of evaluating some or all of the lower level coefficients. In a sequential structure, the delay associated with the implementation grows exponentially with the number of levels [182]. Parallel solutions have been proposed by Krishnaswamy *et al.* [190], Yang *et al.* [191], Fridman *et al.* [192] etc. for traditional message passing computing paradigms and most of them do not adapt effectively to the irregular computational structure of the WT.

The DWT, WPT, and M-band wavelet techniques are often implemented as a cascade of critically sampled filter banks. Many applications apply these techniques to finite frames of samples, leading to distortion in the filtered coefficients near the frame boundaries. In many-a-case, some assumptions are made about the signal characteris-

⁴⁵Fast Hartley Transform

⁴⁶Block DWT

tics beyond the frame boundaries. Signal extension techniques include zero and periodic extension [9], [193], and symmetric extension [9], [193], [194]. Muramatsu *et al* [195] formalized the OSC⁴⁷ technique for boundary elimination in single level multirate filter banks. Recently, Nealand *et al.* has reported [196] an application of WOSC⁴⁸ technique for cascaded filter banks, eliminating boundary distortion in the frame-based application of the DWT, WPT, and M-band wavelet techniques. WOSC eliminates boundary distortion with the exception of the first and last frames. A similar technique has already been developed for WPT [197]. A novel method to eliminate all edge artifacts in block-wise computation of DWT has been developed by Mini *et al* [198], [199]. This method effectively makes use of the overlap-add and -save method simultaneously in the DWT computation to achieve computational advantage as well as reduction in computational memory requirement.

In many of the DWT based data compression applications, compression is achieved making use of the sparsity of the quantized wavelet coefficients. Guo [200] proposes a mapped IDWT algorithm that takes advantage of this sparsity, and significantly lowers the complexity of IDWT to the level that is proportional to the number of non-zero coefficients.

It is noticed that many of the algorithms as presented in the literature do not provide sufficiently detailed exposition to expedite their use in practical applications by those not specializing in the theory. Moreover, these algorithms have been presented in general, for finite duration discrete-time signals restricted to a length equal to a power of 2. Also, most often, the issue of errors, periodization, redundancy, etc. are not well documented to relate algorithms to specific tasks and requirements. Taswell *et al.* [201] addresses these issues and relate to wavelet transform for finite-duration discrete-time signals of arbitrary length not restricted to a power of 2. They consider algorithms based on the pyramid structure.

⁴⁷Overlap-Save Convolution

⁴⁸Wavelet Overlap-Save Convolution

2.3.2 2D DWT Computation

The 2D DWT computation basically depends on whether separable or non-separable wavelets are used for its implementation. For computation with non-separable wavelets, the 2D data is directly convolved with the 2D wavelets and scaling function which are derived as the tensor products of the 1D wavelet and scaling function. When separable wavelets are used, the algorithms developed for 1D DWT can be adapted to 2D DWT computation as well. In such cases the computation is performed by separate row-wise and column-wise convolutions followed by multirate techniques.

The need for massive computation in 2D DWT, coupled with the demand for real-time operation in many image-processing tasks, has motivated the use of parallel processing to provide high performance at a reasonable cost. Parallelization of the 2D DWT has been described by Lu [202] based on the conventional pyramidal algorithm suggested by Mallat. Recently, multithreaded architectures have been promoted, as the semi-irregular computation structure of the WT and the embedded fine-grained parallelism is well suited for multithreading [203]. Marino *et al.* [204], has proposed a method for 2D DWT computation in standard form on parallel general-purpose computers in which the intermediate data transposition is reduced and the inter-processor communication is minimized. Since it is based on matrix-vector multiplication, there is no restriction on the size of the output data.

The VLSI architectures based on linear systolic arrays, proposed by Viswanath *et al.* [181] has an extension for computing the 2D DWT also. Here the separable wavelets are used to compute the DWT in real-time using the recursive pyramid algorithm and it is found to be suitable for single chip implementation. The authors also discuss the adaptation of the architecture for M -band DWT computation.

Chapter 3

Review of Basic Theory

The background theory closely related to the research work carried out, is briefly presented in this chapter. The topics include the basics of wavelets, wavelet transforms, the concept of multiresolution analysis, linear predictive coding, DWT computation etc. The theory of PSWT and LPC are slightly elaborated.

3.1 The Wavelets

Wavelet functions serve as excellent mathematical tool in the time-frequency analysis of both one-dimensional and two-dimensional signals. In practical wavelet analysis, this is achieved by representing the signals as a linear combination of scaled and shifted versions of a single wavelet function called the *mother wavelet*.

There are various kinds of wavelets. Accordingly, one can choose from among smooth wavelets, compactly supported wavelets, symmetric and non-symmetric wavelets, orthogonal and biorthogonal wavelets etc. It is often a complex task since there are so many properties to be considered, like the smoothness, temporal/spatial localization, vanishing moments, frequency localization, symmetry, orthogonality etc. In general, a wavelet function $\psi(t)$ will satisfy the following properties [82].

$$\int_{-\infty}^{\infty} \psi(t) dt = 0 \quad (3.1a)$$

$$\int_{-\infty}^{\infty} |\psi(t)|^2 dt < \infty \quad (3.1b)$$

$$0 < \int_{-\infty}^{\infty} \frac{|\Psi(\omega)|^2}{|\omega|} d\omega < \infty \quad (3.1c)$$

The condition in equation 3.1c is known as the *admissibility condition* on $\psi(t)$ [2]. $\Psi(\omega)$ is the Fourier Transform of $\psi(t)$.

One of the important classes of wavelet system is the **Orthogonal wavelets**. They satisfy the following orthogonality relationships [205].

$$\int \phi_{J,k}(t) \phi_{J,k'}(t) dt = \delta_{k,k'} \quad (3.2a)$$

$$\int \psi_{j,k}(t) \phi_{J,k'}(t) dt = 0 \quad (3.2b)$$

$$\int \psi_{j,k}(t) \psi_{j,k'}(t) dt = \delta_{j,j'} \delta_{k,k'} \quad (3.2c)$$

where $\phi_{i,j}(t)$ and $\psi_{i,j}(t)$ are the scaling and wavelet functions at the i^{th} scale and j^{th} shift respectively. The following are some of the orthogonal wavelets which have been considered in this study.

Haar: The haar wavelet, discovered by *Alfred Haar* [206], is a square wave. It is the only compact, orthogonal wavelet, which is symmetric [205]. However, unlike others the haar wavelet is not continuous.

Daubechies: The daubechies were the first type of continuous orthogonal wavelet with compact support. There are quite a large number of wavelets in this group *viz.* *db2*, *db4*, etc. all named after the discoverer *Ingrid Daubechies*, who is one of the pioneers in wavelets research [2]. The support and smoothness of these wavelets increase as the wavelet order number increases. These wavelets are all asymmetric and have the highest number of vanishing moments¹ among wavelets with similar properties, for a given support.

Symlets: The symlets also were constructed by Daubechies [2]. The wavelets in this family are nearly symmetric. The moment properties of these wavelets are similar to that of daubechies.

Coiflets: These wavelets are nearly symmetric and possess additional desirable properties. The scaling functions of this family also have vanishing moments similar to the wavelets and give rise to near-linear filters. The coiflets were constructed by Daubechies in association with *Ronald Coifman* [2].

3.2 The Continuous Wavelet Transform

The CWT is the wavelet transform of a continuous signal for continuous values of the dilation and translation parameters as given by equation 1.1. The transform is basically reversible and the signal $f(t)$ can be recovered from its transform using the inversion

¹The n^{th} order moment of a function $f(t)$ is defined as $M_n = \int_{-\infty}^{\infty} t^n f(t) dt$.

formula:

$$f(t) = \frac{1}{C_\psi} \int_0^\infty \int_{\mathbf{R}} a^{-2} (W_\psi f)(a, b) \psi\left(\frac{t-b}{a}\right) da db \quad (3.3)$$

where $\psi(t)$ is the mother wavelet function, and $(W_\psi f)(a, b)$ are the CWT coefficients for dilation a and translation b . Using the inner product notation in $L_2(\mathbf{R})$, the CWT can also be written as:

$$(W_\psi f)(a, b) = |a|^{-1/2} \langle f(t), \psi_{(a,b)}(t) \rangle \quad (3.4)$$

where

$$\psi_{(a,b)} = \psi\left(\frac{t-b}{a}\right), \quad (3.5)$$

and $\langle \cdot, \cdot \rangle$ is the inner product in $L_2(\mathbf{R})$. Since the scale factor a is the inverse of the frequency ω , the value $(W_\psi f)(a_0, b_0)$ exhibits the frequency content of $f(t)$ in a frequency interval centered around $\omega_0 = a_0^{-1}$ at a time interval centered around b_0 . The CWT is highly redundant and in a strict sense, it is impossible to compute CWT using a digital computer.

It is not always possible for the CWT to resolve practically all events in frequency and time [82]. Quantitative metrics for time and frequency resolution are based on the *duration* and *bandwidth* respectively of the mother wavelet². A measure of the duration and bandwidth indicating the spread of the wavelet in time and frequency domain is given by

$$\Delta t \equiv \sqrt{\frac{\int_{-\infty}^{\infty} (t - t_0)^2 |\psi(t)|^2 dt}{\int_{-\infty}^{\infty} |\psi(t)|^2 dt}} \quad (3.6a)$$

$$\Delta \omega \equiv \sqrt{\frac{\int_{-\infty}^{\infty} (\omega - \omega_0)^2 |\Psi(\omega)|^2 d\omega}{\int_{-\infty}^{\infty} |\Psi(\omega)|^2 d\omega}} \quad (3.6b)$$

where t_0 and ω_0 are respectively the first moment of $\psi(t)$ and $\Psi(\omega)$ given by the following equations. They define the location of the wavelet in the time-frequency plane.

$$t_0 \equiv \frac{\int_{-\infty}^{\infty} t |\psi(t)|^2 dt}{\int_{-\infty}^{\infty} |\psi(t)|^2 dt} \quad (3.7a)$$

²Also called the Root Mean Square (RMS) duration and RMS bandwidth

$$\omega_0 \equiv \frac{\int_{-\infty}^{\infty} \omega |\Psi(\omega)|^2 d\omega}{\int_{-\infty}^{\infty} |\Psi(\omega)|^2 d\omega} \quad (3.7b)$$

The product of the RMS duration and bandwidth is invariant to dilation and remains a constant which depends on the wavelet used. If its value is smaller, that wavelet provides better simultaneous localization in the time-frequency plane compared to the one with higher value. The translation parameter b affects only the location of the wavelet and not the duration and bandwidth, which are affected solely by the dilation parameter a .

3.3 The Discrete Wavelet Transform

DWT as defined in equation 1.3 is devoid of the problem of redundancy that exists in CWT. The time-scale parameters being discrete, it has been recognized as a natural wavelet transform for discrete time signals [207], [208]. The DWT of a discrete-time signal $x(n) \in l^2(\mathbf{Z})$ with respect to a set of analysis wavelets $(\xi_{n,m})_{n=1,2,\dots,m \text{ integer}}$ is defined as the set of coefficients

$$X_{n,m} = \sum_k x(k) \xi_{n,m}^*(k) \quad (3.8)$$

$\xi_{n,m}(k)$ is obtained by translating $\xi_{n,0}(k)$ by a scale dependent amount that monotonically increases with m at each fixed scale. Therefore the index m represents the time-shift. Mathematically,

$$\xi_{n,m}(k) = \xi_{n,0}(k - 2^n m) \quad (3.9)$$

Unlike the continuous-time case, in DWT, the scale-invariance may be only approximately satisfied. As a result, dilation and translation operations are defined in DWT on a time-scale grid. The most common grid is the regular dyadic grid given by the set of points $(2^n, 2^n m)_{n=1,2,\dots,m \text{ integer}}$. On the dyadic grid, the wavelet $\xi_{n+1,0}(k)$ is obtained by upsampling $\xi_{n,0}(k)$ by a factor of two.

The transform 3.8 is invertible [207], [209], [210] if, on the same time-scale grid, one can determine a dual set of synthesis wavelet sequences $(\psi_{n,m})_{n=1,2,\dots;m \text{ integer}}$ that satisfy the completeness relationship

$$\sum_n \sum_m \xi_{n,m}^*(s) \psi_{n,m}(p) = \delta_{s,p} \quad (3.10)$$

where $\delta_{s,p}$ is the Kronecker delta, that is $\delta_{s,p} = \begin{cases} 1 & s = p, \\ 0 & \text{otherwise.} \end{cases}$

In that case one can recover the signal $x(n)$, obtaining the following expansion

$$x(k) = \sum_n \sum_m X_{n,m} \psi_{n,m}(k) \quad (3.11)$$

The equations 3.8 and 3.11 define the wavelet decomposition of $x(k)$.

For bandlimited signals, a finite-level expansion will be quite sufficient. Hence such signals can be defined by the following finite sum.

$$x(k) = \sum_{n=1}^N \sum_m X_{n,m} \psi_{n,m}(k) + \gamma_N(k) \quad (3.12)$$

where

$$\gamma_N(k) = \sum_m \beta_{N,m} \phi_{N,m}(k) \quad (3.13)$$

and

$$\beta_{N,m} = \sum_k x(k) \lambda_{N,m}^*(k) \quad (3.14)$$

Here, $\lambda_{N,m}(k)$ is the analysis scaling sequence at N^{th} level, $\phi_{N,m}(k)$ being the corresponding synthesis dual. The sequence $\gamma_N(k)$ represents the residue of the expansion over the finite wavelet set. As $N \rightarrow \infty$, one must ensure that the residue converges in norm to zero for any arbitrary finite-energy input signal.

3.4 Wavelets and Time-Frequency Analysis -Concept of MRA

MRA³ forms the most important building block for the construction of scaling functions and wavelets, and for the development of algorithms for the time-frequency analysis of signals. MRA is something which helps us in viewing a function at various levels of approximation or resolution. Using this concept, a complicated function can be divided into several simpler ones and study them separately.

To achieve an MRA of a function $x(t)$, we must have a finite energy function $\phi(t) \in L^2(R)$, called a scaling function, that generates a nested sequence $\{\mathbf{A}_j\}$, namely $\{0\} \leftarrow \dots \subset \mathbf{A}_{-1} \subset \mathbf{A}_0 \subset \mathbf{A}_1 \subset \dots \rightarrow L^2$, and satisfies a dilation equation $\phi(t) = \sum_k g_0[k]\phi(at - k)$ for some $a > 0$ and coefficients $\{g_0[k] \in l^2\}$. In DWT we consider $a = 2$, which corresponds to octave scales. The function $\phi(t)$ is called the scaling function. The space \mathbf{A}_0 is generated by $\{\phi(\cdot - k) : k \in \mathbf{Z}\}$ and in general \mathbf{A}_s , by $\{\phi_{k,s} : k, s \in \mathbf{Z}\}$. Consequently, we have the following obvious results.

$$x(t) \in \mathbf{A}_s \Leftrightarrow x(2t) \in \mathbf{A}_{s+1} \quad (3.15a)$$

$$x(t) \in \mathbf{A}_s \Leftrightarrow x(t + 2^{-s}) \in \mathbf{A}_s \quad (3.15b)$$

These dilation equations are unique to MRA.

For each s , since \mathbf{A}_s is a proper subspace of \mathbf{A}_{s+1} , there is some space left in \mathbf{A}_{s+1} , called \mathbf{W}_s , which when combined with \mathbf{A}_s gives us \mathbf{A}_{s+1} . This space $\{\mathbf{W}_s\}$ is called the wavelet subspace and is complementary to \mathbf{A}_s in \mathbf{A}_{s+1} , meaning that

$$\mathbf{A}_s \cap \mathbf{W}_s = \{0\}, \quad s \in \mathbf{Z} \quad (3.16a)$$

$$\mathbf{A}_s \oplus \mathbf{W}_s = \mathbf{A}_{s+1} \quad (3.16b)$$

³Multi-Resolution Analysis

Subspaces $\{\mathbf{W}_s\}$ are generated by $\psi(t) \in \mathbf{L}^2$, called the wavelet, in the same way as $\{\mathbf{A}_s\}$ is generated by $\phi(t)$. In other words, any $x_s(t) \in \mathbf{A}_s$ can be written as

$$x_s(t) = \sum_k a_{k,s} \phi(2^s t - k), \quad (3.17)$$

and any function $y_s(t) \in \mathbf{W}_s$ can be written as

$$y_s(t) = \sum_k w_{k,s} \psi(2^s t - k) \quad (3.18)$$

for same coefficients $a_{k,s}, w_{k,s} \in l^2$.

Since we have

$$\begin{aligned} \mathbf{A}_{s+1} &= \mathbf{W}_s \oplus \mathbf{A}_s \\ &= \mathbf{W}_s \oplus \mathbf{W}_{s-1} \oplus \mathbf{A}_{s-1} \\ &= \mathbf{W}_s \oplus \mathbf{W}_{s-1} \oplus \mathbf{W}_{s-2} \oplus \dots, \end{aligned} \quad (3.19)$$

we have

$$\mathbf{A}_s = \bigoplus_{l=-\infty}^{s-1} \mathbf{W}_l \quad (3.20)$$

Observe that the $\{\mathbf{A}_s\}$ are nested while the $\{\mathbf{W}_s\}$ are mutually orthogonal. Consequently, we have

$$\mathbf{A}_l \cap \mathbf{A}_m = \mathbf{A}_l, \quad m > l \quad (3.21a)$$

$$\mathbf{W}_l \cap \mathbf{W}_m = \{0\}, \quad l \neq m \quad (3.21b)$$

$$\mathbf{A}_l \cap \mathbf{W}_m = \{0\}, \quad l \leq m \quad (3.21c)$$

The hierarchical nature of \mathbf{A}_s and \mathbf{W}_s can be shown by the schematic representation in figure 3.1. In the case of an *orthogonal decomposition*, in addition to the wavelet space \mathbf{W}_s being complimentary to \mathbf{A}_s , they are mutually orthogonal also, such that $\mathbf{W}_s \perp \mathbf{A}_s$.

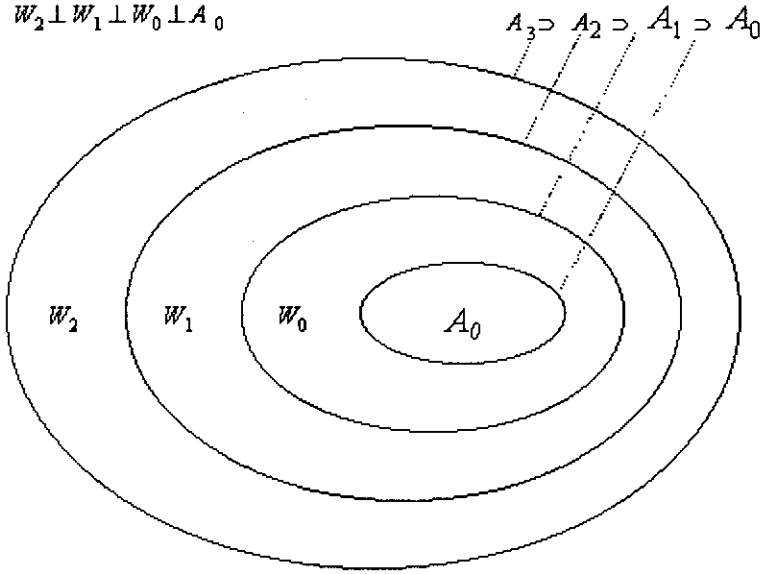


Figure 3.1: Schematic of MRA decomposition

3.4.1 Discrete multiresolution analysis

Similar to the MRA in $L_2(\mathbf{R})$, a DMRA⁴ for $l_2(\mathbf{Z})$ has been developed [5], [211], [212]. Accordingly any discrete-time sequence $f(n)$ of finite energy can be expressed in terms of the discrete-time basis functions $\psi_{j,k}(n)$ as:

$$f(n) = \sum_{j,k} d_j(k) \psi(2^j n - k) \quad (3.22)$$

This is in parallel to equation 1.3. If the expansion basis functions form a tight frame, the expansion coefficients from an inner product by

$$d_j(k) = \langle f(n), \psi(2^j n - k) \rangle = \sum_n f(n) \psi(2^j n - k) \quad (3.23)$$

⁴Discrete Multi Resolution Analysis

Now there are two ways of looking at signal processing using the WT at dyadic scales. The first is analogous to the use of Fourier Series where a continuous function is transformed into a discrete sequence of coefficients. The second is analogous to the DFT where a discrete function is transformed into another discrete function. The former is generally referred to as the DWT and the latter is designated as the DTWT⁵ [81]. The difference in these views comes partly from the background of various researchers (i.e., whether they are "wavelet people" or "filter bank people"), however, there are subtle difference between using the series expansion of the signal and using a multirate digital filter bank on samples of the signal.

3.5 DWT computation

For practical computation of DWT coefficients, a fast pyramid algorithm, also called the *Mallat's algorithm* is popularly used, relating the wavelet function to a set of QMF⁶ bank [3], [9]. This involves successive filtering and multirate operations.

3.5.1 The Basic Multirate Operations

The basic operations involved in sampling rate conversion of digital signals by digital means are the sampling rate reduction by an integer factor D called the *decimation*, and the sampling rate increase by an integer factor I , called the *interpolation*.

3.5.1.1 Decimation by a factor D

Decimation or downsampling a signal $x(n)$ by a factor of D simply means selecting every D^{th} sample of $x(n)$, thus resulting in an aliased version of $x(n)$, with a folding frequency of $F_s/2D$ where F_s is the sampling frequency of $x(n)$. If $x(n)$ was bandlimited to

⁵Discrete-Time Wavelet Transform

⁶Quadrature Mirror Filter

$F_{max} = F_s/2D$, there won't be any aliasing. Hence before the process of downsampling, usually a lowpass filtering is performed, which is characterized by the frequency response, $H_D(w)$, which ideally satisfies the condition

$$H_D(w) = \begin{cases} 1 & |w| \leq \pi/D \\ 0 & \text{otherwise} \end{cases} \quad (3.24)$$

The downsampled signal $x_d(m)$ can be expressed as

$$x_d(m) = x(mD) \quad (3.25)$$

The downsampling operation results in a time-variant system.

The frequency domain characteristics of the output sequence $x_d(m)$ can be shown to be [17]

$$X_d(z) = \frac{1}{D} \sum_{k=0}^{D-1} X(e^{-j2\pi k/D} z^{1/D}) \quad (3.26)$$

3.5.1.2 Interpolation by a factor I

An increase in the sampling rate by an integer factor of I can be accomplished by inserting $I - 1$ samples between successive values of the signal. If the sequence $x(m)$ is to be interpolated, it is first upsampled by putting $I - 1$ zeros between successive samples followed by a stage of filtering to remove the image frequencies. The upsampled sequence can be expressed as

$$x_u(m) = \begin{cases} x(m/I) & m = 0, \pm I, \pm 2I, \dots \\ 0 & \text{otherwise} \end{cases} \quad (3.27)$$

In frequency domain, it can be expressed as

$$X_u(z) = X(z^I) \quad (3.28)$$

whose spectrum is an I -fold periodic repetition of the input signal spectrum. The new sampling rate will be IF_s . Since only the frequency components of $x(n)$ in the range

$0 \leq w_s = 2\pi F_s \leq \pi/I$ are unique, the images of $X(w)$ in $X_u(w)$ beyond this range should be rejected by passing the sequence $x_u(m)$ through a filter of frequency response $H_U(w)$ that ideally has the characteristics

$$H_U(w) = \begin{cases} C & |w| \leq \pi/I \\ 0 & \text{otherwise} \end{cases} \quad (3.29)$$

where C is a scale factor required to normalize the output sequence.

3.5.1.3 Sampling rate conversion by a rational factor I/D

This type of sampling rate conversion is of frequent application in multirate systems as in the case of *PSWT based Linear Predictive Coding* described in chapter 4. Sampling rate conversion by a factor of I/D can be achieved by first performing interpolation by the factor I and then decimating the output of the interpolator by the factor D . The cascaded operation is illustrated in figure 3.2(a). The two cascaded filters in fig. 3.2(a) can be combined to a single filter as shown in fig. 3.2(b), which possess an ideal frequency response of

$$H(w) = \begin{cases} I & |w| \leq \min(\pi/D, \pi/I) \\ 0 & \text{otherwise} \end{cases} \quad (3.30)$$

' \downarrow ' indicates *down sampling* operation and ' \uparrow ' denotes *upsampling* operation.

In the time domain, the output of the upsampler is the sequence

$$v(l) = \begin{cases} x(l/I) & l = 0, \pm I, \pm 2I, \dots \\ 0 & \text{otherwise} \end{cases} \quad (3.31)$$

and the output of the linear time-invariant filter is

$$\begin{aligned} w(l) &= \sum_{k=-\infty}^{\infty} h(l-k)v(k) \\ &= \sum_{k=-\infty}^{\infty} h(l-kI)x(k) \end{aligned} \quad (3.32)$$

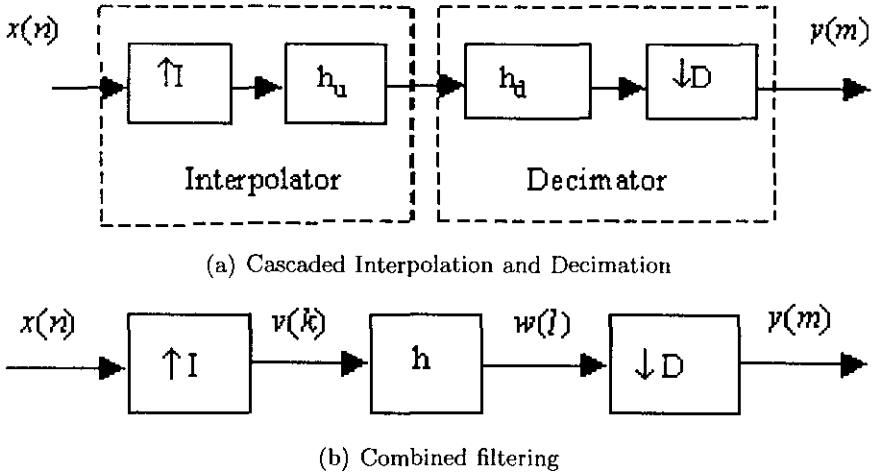


Figure 3.2: Fractional sampling rate conversion by multirate techniques

Finally the output of the sampling rate converter is the sequence $\{y(m)\}$, which is obtained by downsampling the sequence $\{w(l)\}$ by the factor D . Thus

$$y(m) = \sum_{k=-\infty}^{\infty} h(mD - kI)x(k) \quad (3.33)$$

3.5.2 The Mallat's Pyramid Structure

The DWT is generally computed using the *Pyramidal Tree* structure, as shown in figure 3.3 for a 2-level wavelet analysis on a discrete-time sequence $x(n)$. This algorithm is sometimes referred to as the two-channel sub-band coder as it involves the QMF filtering of the input signal based on the wavelet function used. The filters h, g, \tilde{h} , and \tilde{g} characterize the wavelet system. This is basically a sequential structure employing multirate digital filtering.

One can look at the Mallat's algorithm either as a way of calculating expansion coefficients of a continuous time signal at various scales or as a filter bank for processing discrete-time signals. The wavelet analysis coefficients are obtained at the output of each leaf of the tree structure. The sections are numbered from left to right in the

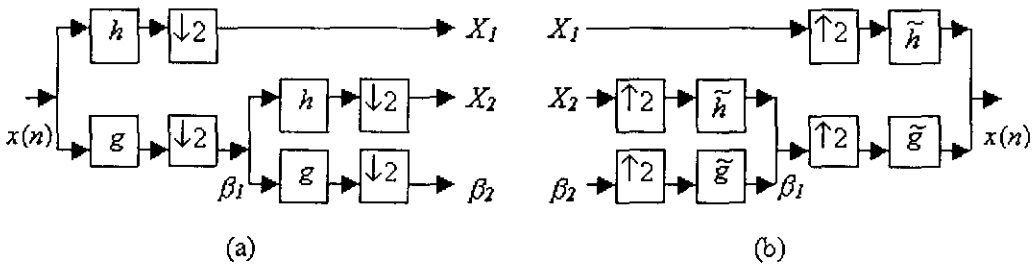


Figure 3.3: Pyramid structure for 2 level DWT Computation (a) Decomposition (b) Reconstruction

analysis structure and from right to left in the synthesis structure. The set of transform coefficients at j^{th} level are computed from the scaling output of the lower or previous stage of the analysis structure as

$$X_j(k) = (\beta_{j-1}(r) * h(r)) \downarrow 2 \quad (3.34a)$$

$$\beta_j(k) = (\beta_{j-1}(r) * g(r)) \downarrow 2 \quad j = 1, 2, \dots, J \quad (3.34b)$$

where $\beta_0(k) = x(k)$, $h(k) = \psi_{0,0}(k)$, and $g(k) = \phi_{0,0}(k)$.

In the frequency domain, the first stage divides the signal spectrum into two equal parts. The second stage divides the lower half into quarters and so on. This results in a logarithmic set of bandwidths as illustrated in figure 3.4. Correspondingly the tiling of the time-frequency plane will look as shown in figure 3.5. Each horizontal strip in the tiling corresponds to each channel, which in turn corresponds to a scale j . The span covered by each of the translations are marked along the horizontal axis. It is obvious that at lower resolutions (higher j), the translations are large and at higher resolutions the translations are small. The residue of the finite scale wavelet representation is then a low-pass signal that reproduces the trend of the input signal. The remaining terms in the expansion represent the fluctuations of the signal over this trend at several scales.

Similarly, employing the synthesis scaling and wavelet sequences of upper levels, the scaling coefficients at each of the immediate lower levels are computed using the

multirate filtering operations.

3.6 The Wavelet Packet Transform

Wavelet Packet analysis is an important generalization of wavelet analysis [213]. Wavelet packet basis functions comprise a rich family of building block functions which are still localized in time, but offer more flexibility than wavelets in representing different types of signals. The classical wavelet system results in a logarithmic frequency resolution. The low frequencies have narrow bandwidths and the high frequencies have wide bandwidths. The WP is an extension on this to allow a finer and adjustable resolution of frequencies at high frequencies. The cost of this richer structure is a computational complexity of $O(N \log(N))$, in contrast to the classical WT, which is $O(N)$. A wavelet orthonormal basis decomposes the frequency axis in dyadic intervals whereas the wavelet packets generalize this fixed dyadic construction by decomposing the frequency in intervals of varying bandwidth. Wavelet packet functions are generated by scaling and translating a family of basic functions denoted by, say $W_b(t)$. These functions are uniquely characterized by the oscillation index b . The fundamental scaling and wavelet functions correspond to $b = 0$ and $b = 1$ respectively. Larger values of b points to wavelet packets with more oscillations and higher frequency. The translated and scaled wavelet packet functions related to the oscillation index b , at any level j are generated from the following equation.

$$W_{j,b,m}(k) = W_{j,b,0}(k - 2^j m) \quad (3.35)$$

where m is an integer and

$$W_{j,b,0}(k) = 2^{-j/2} W_b(2^{-j} k) \quad (3.36)$$

The wavelet packet $W_{j,b,m}(k)$ has scale 2^j and location $2^j m$.

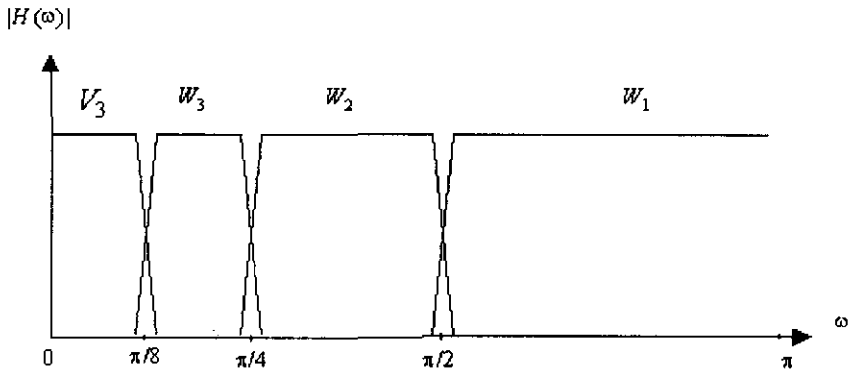


Figure 3.4: Frequency bands for the analysis tree of the pyramid structure

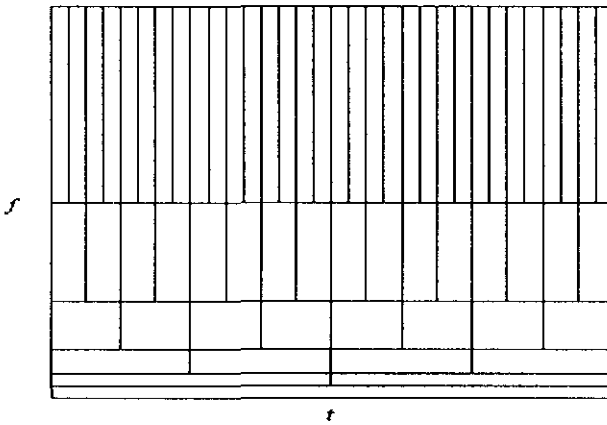


Figure 3.5: Time-Frequency tiling in wavelet decomposition

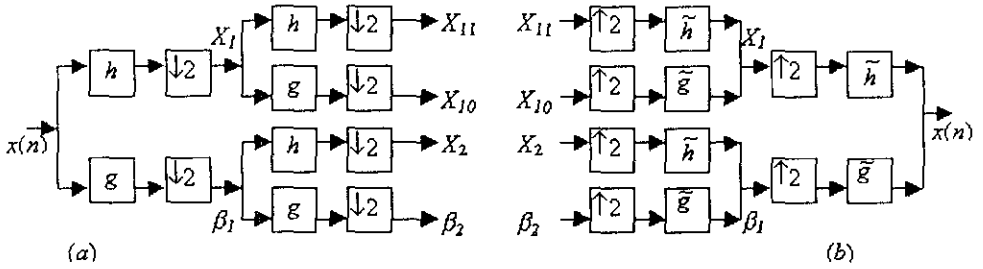


Figure 3.6: Tree structure for 2-level WP analysis (a) decomposition (b) reconstruction

The Mallat algorithm can be extended to WP⁷ analysis also. In WP decomposition, the high pass wavelet branch as well as the low pass approximation branch of the tree are successively decomposed. The resulting analysis tree and synthesis tree for a two-stage decomposition is shown in figure 3.6.

Figure 3.7(a) pictorially shows the vector-space decomposition for wavelet packet transform. Figure 3.7(b) shows the corresponding frequency response of the filter bank scheme.

A signal $x(n)$ can be expressed as a sum of orthogonal wavelet packet functions $W_{j,b,k}(n)$, at different scales, oscillations, and locations using the expression

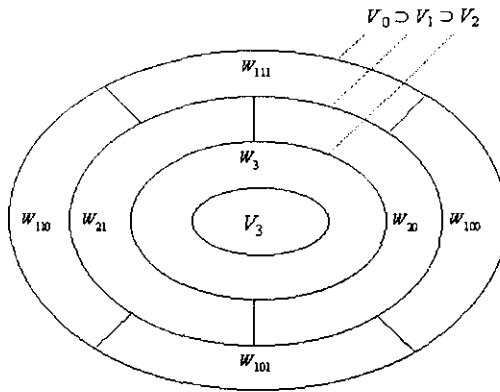
$$x(n) = \sum_j \sum_b \sum_k w_{j,b,m} W_{j,b,m}(n) \tag{3.37}$$

where

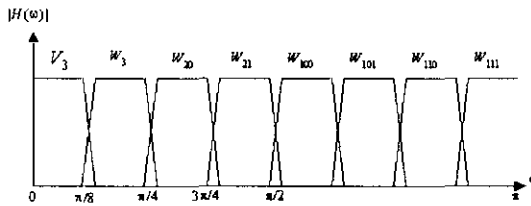
$$w_{j,b,m} = \sum_n W_{j,b,m}(n)x(n) \tag{3.38}$$

The range of the summation for the levels j and the oscillations b is chosen so that the wavelet packet functions are orthogonal.

⁷Wavelet Packet



(a) Vector-spaces of the Scaling and Wavelet functions



(b) Frequency response of WP filter bank

Figure 3.7: Decomposition using Wavelet Packet

3.7 2D Discrete Wavelet Transform

In order to apply wavelet decomposition to images, 2D extension of the wavelets are required. This can be achieved by the use of separable or non-separable wavelets. In this thesis, the case of separable wavelets only is considered. A 2D separable wavelet basis can be constructed from 1D using four basis functions viz. one scaling function $\phi_{j,k,l}^0(x, y)$ and three wavelet functions $\psi_{j,k,l}^\tau(x, y)$, $\tau \in 1, 2, 3$, $j, k, l \in \mathbf{Z}$, defined as follows [184].

$$\phi_{j,k,l}^0(x, y) = \phi_{j,k}(x)\phi_{j,l}(y) \tag{3.39a}$$

$$\psi_{j,k,l}^1(x, y) = \phi_{j,k}(x)\psi_{j,l}(y) \tag{3.39b}$$

$$\psi_{j,k,l}^2(x, y) = \psi_{j,k}(x)\phi_{j,l}(y) \quad (3.39c)$$

$$\psi_{j,k,l}^3(x, y) = \psi_{j,k}(x)\psi_{j,l}(y) \quad (3.39d)$$

where ϕ and ψ are the 1D wavelet basis defined earlier. These basis functions span the four j -level linear vector spaces rather than just two as in the 1D case. The first one spans the approximation subspace V_j , whereas the other three spans each of the details space W_{j1} , W_{j2} and W_{j3} respectively. It may be expressed making use of the MRA concept as

$$V_{j-1} = V_j \oplus W_{j1} \oplus W_{j2} \oplus W_{j3} \quad (3.40)$$

where V_{j-1} is the next high resolution approximation subspace. An analogous definition holds for the dual scaling function $\tilde{\phi}_{j,k,l}^0(x, y)$ and wavelet functions $\tilde{\psi}_{j,k,l}^r(x, y)$.

The M -level wavelet representation of a 2D-function $f(x, y)$ is given by

$$f(x, y) = \sum_{k,l} c_{k,l}^M \phi_{M,k,l}^0(x, y) + \sum_{j=1}^M \sum_{r \in \tau} \sum_{k,l} d_{k,l}^{j,r} \psi_{j,k,l}^r(x, y) \quad (3.41)$$

The approximation and the detail coefficients in the above expression are $c_{k,l}^M = \langle f, \tilde{\phi}_{M,k,l}^0 \rangle$ and $d_{k,l}^{j,r} = \langle f, \psi_{j,k,l}^r \rangle$ respectively, where $\langle \cdot, \cdot \rangle$ denotes the inner product in the $l^2(\mathbf{Z}^2)$ space.

3.7.1 Computation of 2D DWT

For fast DWT computation using the Mallat's pyramid algorithm, the sub band filtering scheme as in the case of 1D is used. The 2D wavelet basis may be represented by the four possible tensor products gg, gh, hg and hh of the 1D filters g and h . Let c^j and $d^{j,\tau}$ denote the 2D sequences $c_{k,l}^j$ and $d_{k,l}^{j,\tau}$; $k, l \in \mathbf{Z}$ and $\tau = 1, 2, 3$. The scaling and wavelet

transform coefficients at a coarser level j are computed from c^{j-1} by convolution followed by downsampling [184] as:

$$c^j = (\downarrow\downarrow 2)(gg * c^{j-1}) \quad (3.42a)$$

$$d^{j,1} = (\downarrow\downarrow 2)(gh * c^{j-1}) \quad (3.42b)$$

$$d^{j,2} = (\downarrow\downarrow 2)(hg * c^{j-1}) \quad (3.42c)$$

$$d^{j,3} = (\downarrow\downarrow 2)(hh * c^{j-1}) \quad (3.42d)$$

for $j = 0, 1, \dots, M - 1$. Here $*$ denotes 2D convolution and $(\downarrow\downarrow 2)$ denotes downsampling by a factor of 2 in both x and y directions. The given image is treated as c^0 .

To perform the 2D DWT computation as above, instead of using the 2D filters, one can employ a separable extension of the 1D decomposition algorithm. The rows of the data are convolved with the first 1D filter in equation 3.42 and alternate columns are retained. The resulting data is then convolved column-wise using the other 1D filter and every alternate row is dropped. Further stages of 2D decomposition are obtained by recursively applying the procedure to the low-pass filtered output of the previous stage. Following the terminology found in the literature, the coefficients given by equation 3.42 may be denoted as LL_j, LH_j, HL_j and HH_j respectively as shown in the layout of figure 3.8.

The wavelet reconstruction is performed recursively, starting at level M by upsampling in both directions (denoted by $\uparrow\uparrow 2$) followed by convolution using the 2D dual filters. The signal reconstructed at the j^{th} level from coefficients at the $(j + 1)^{th}$ level may be expressed as

$$c^j = \tilde{g}\tilde{g}*((\uparrow\uparrow 2)c^{j+1}) + \tilde{g}\tilde{h}*((\uparrow\uparrow 2)d^{j+1,1}) + \tilde{h}\tilde{g}*((\uparrow\uparrow 2)d^{j+1,2}) + \tilde{h}\tilde{h}*((\uparrow\uparrow 2)d^{j+1,3}) \quad (3.43)$$

Here also, separable wavelet filters can be employed as in the case of decomposition.

Figure 3.9 shows the block diagram for a 2-level decomposition and reconstruction of an image using the pyramidal algorithm employing separable filters.

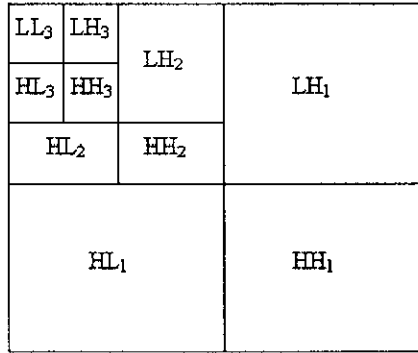


Figure 3.8: Coefficient layout of a 3-level DWT of an image.

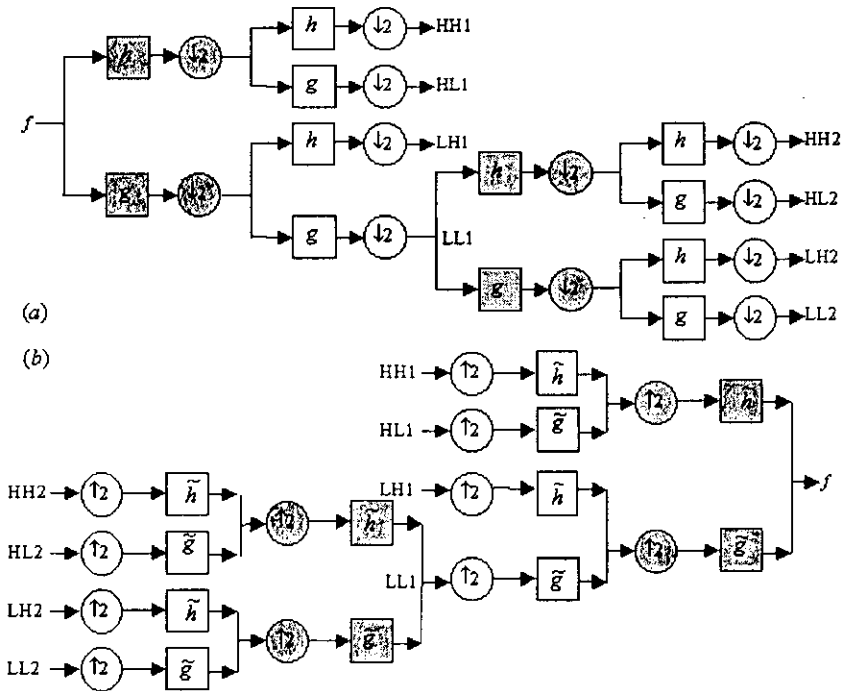


Figure 3.9: Pyramid Structure for 2-level 2D DWT computation (a) Decomposition (b) Reconstruction. (The shaded blocks represent row-wise operation, the rest being column-wise)

3.8 The Pitch Synchronous Wavelet Transform

By means of ordinary DWT the signal can be represented in terms of a low-pass trend plus fluctuations at several scales. However, pseudo-periodic signals may be better represented in terms of a periodic trend plus period-to-period fluctuations. The periodic trend may be obtained by averaging the *homologous samples*⁸, while period-to-period fluctuations are differences of the homologous samples. Looking from the frequency domain, the trend lies in narrow bands centered on the harmonics while the fluctuations consist of sidebands of the harmonics. In a perfectly periodic signal these fluctuations are null and the homologous samples are identical to the samples in any arbitrary period. Homologous samples in a pseudo-periodic signal are strongly correlated and may be slowly varying. By means of PSWT one can better analyze these changes at several scales.

The PSWT is based on a pitch-synchronous vector representation and it adapts to the oscillatory characteristics of pseudo-periodic signals. A lot of information is conveyed in period-to-period variations. These signals have a lot of redundancy between periods.

The PSWT rests upon the PS representation of the signal. It is nothing but the ordinary WT applied to each of the row vectors of the signal in PS form. For a quasi-periodic signal, the local pitch periods $P(r)$ for $r = 0, 1, 2, \dots, N$ may vary, which makes the column vectors in the PS matrix to be of different size. Hence the PS matrix is modified by inserting $p_m - P(r)$ zeros at the end of each column vector $v_q(r)$ for any fixed value of r . Here p_m stands for the maximum value of $P(r)$. Mathematically we define the new column vectors as sequences:

$$\hat{v}_q(r) = X_q(r)v_q(r) \quad (3.44)$$

where

$$X_q(r) = \begin{cases} 1 & q = 0, 1, \dots, P(r) - 1 \\ 0 & \text{otherwise} \end{cases} \quad (3.45)$$

⁸samples that are spaced one or more periods apart.

Obviously $\hat{v}_q(r)$ extends $v_q(r)$ as zeros, for q outside $P(r)$.

Now, if we try to represent each row in the PS signal matrix in terms of its DWT coefficients, the resulting matrix will constitute the PSWT of the signal represented by:

$$\hat{V}_{j,k,q} = \sum_n \hat{v}_q(n) \psi_{j,k}(n) \quad (3.46a)$$

where $\hat{V}_{j,k,q}$ are the PSWT coefficients of the q^{th} row of the signal $v_q(r)$ at j^{th} scale and k^{th} shift. Signal reconstruction is possible using the equation:

$$\hat{v}_q(n) = \sum_{j,k} \hat{V}_{j,k,q} \psi_{j,k}(n) \quad (3.46b)$$

The above PSWT coefficients can be expressed in terms of the original sequence $x(n)$ and the PS wavelet sequence as given below.

Considering a finite level decomposition of the signal, the PSWT may be expressed [11] by the following equations:

$$\hat{V}_{j,k,q} = \sum_r x(r) \theta_{j,k,q}(r); \quad j = 1, 2, \dots, J \quad (3.47a)$$

and

$$\hat{\sigma}_{J,k,q} = \sum_r x(r) \varphi_{J,k,q}(r); \quad k, q \in Z \quad (3.47b)$$

where $\hat{\sigma}_{J,k,q}$ are the PS scaling transform coefficients of the same section of the signal at J^{th} scale and k^{th} shift. This is the residue of the signal after J levels of PSWT decomposition. $\theta_{j,k,q}(r)$ are the PS wavelet sequences generated from $\psi_{j,k}$, and $\varphi_{J,k,q}(r)$ are the PS scaling function sequences generated from $\phi_{J,k}$ using the equations:

$$\theta_{j,k,q}(r) = \sum_n \delta(r - q - \sum_{u=0}^{n-1} P(u)) \psi_{j,k}(n) X_q(n) \quad (3.48a)$$

and

$$\varphi_{J,k,q}(r) = \sum_n \delta(r - q - \sum_{u=0}^{n-1} P(u)) \phi_{J,k}(n) X_q(n) \quad (3.48b)$$

where

$$\psi_{j,k}(n) = \psi_{j,0}(n - 2^j k) \quad (3.49a)$$

$$\phi_{J,k}(n) = \phi_{J,0}(n - 2^J k) \quad (3.49b)$$

Here $\psi_{j,0}$ and $\phi_{J,0}$ are the impulse responses h_j and g_J respectively, defined by

$$g_{j+1}(n) = \sum_r g_j(r)g(n - 2r) \quad (3.50a)$$

and

$$h_{j+1}(n) = \sum_r h_j(r)g(n - 2r) \quad (3.50b)$$

where, $g_1(n) = g(n)$ and $h_1(n) = h(n)$, are the impulse response of the analysis low pass and high pass filters respectively.

From the finite level PS wavelet and scaling coefficients, the original signal is reconstructed using the following relationship:

$$x(n) = \sum_{j=1}^J w_j(n) + \gamma_J(n) \quad (3.51)$$

where $w_j(n) = \sum_{k,q} \hat{V}_{j,k,q} \theta_{j,k,q}(n)$ is the PS wavelet partial at the j^{th} level

and $\gamma_J(n) = \sum_{k,q} \hat{\sigma}_{J,k,q} \varphi_{J,k,q}(n)$ is the PS scaling residue at the J^{th} level.

It may be noted that $w_j(n)$ corresponds to the extend of fluctuations at scale 2^j local pitch periods, whereas $\gamma_J(n)$ represents the asymptotic average behavior over several pitch periods.

After representing $x(n)$ in the PS form, the computation of the PSWT and its Inverse (IPSWT), can be performed by taking the DWT and IDWT of each of the q channels, one at a time using the PS wavelets and scaling sequences given by equation 3.48. In a parallel processing environment, the DWT and IDWT of all the channels may be computed concurrently, as they are independent of each other. The implementation structure for the above computation is given in figure 3.10. The DWT and IDWT blocks shown in the figure can be implemented using the same structure given in figure 3.3.

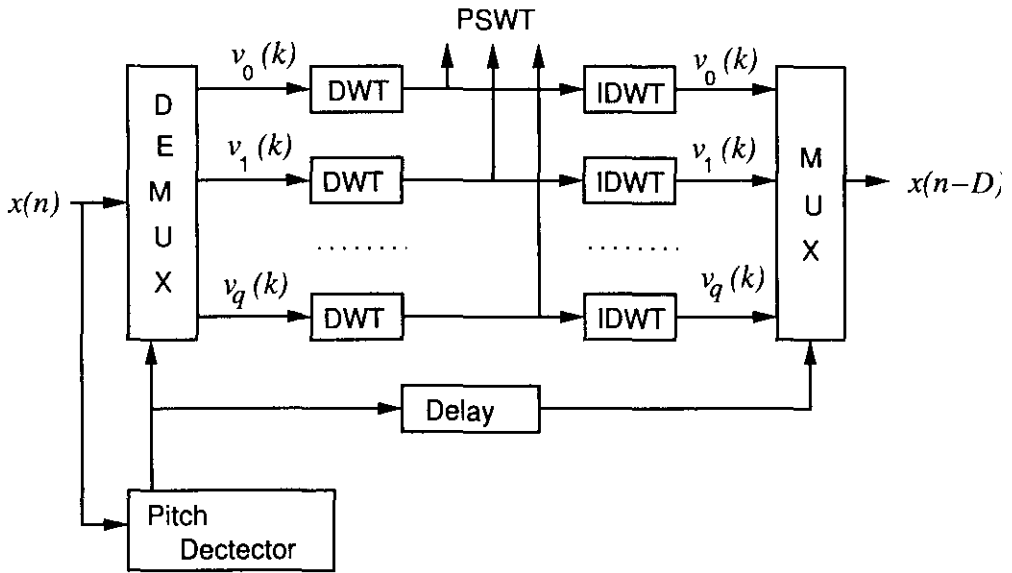


Figure 3.10: Implementation structure for the computation of PSWT and its inverse. (' D ' corresponds to the delay in the computation of the forward and inverse transform blocks).

3.9 Period Estimation

Period detection algorithms can be classified into two separate categories, *viz.* the spectral-domain based and the time-domain based period detection. Spectral period detectors such as Cepstrum, Maximum Likelihood, and Autocorrelation methods, estimate the fundamental period of a signal directly using windowed segments of the signal, applying a Fourier-type analysis to determine average period. A time based detector, however, estimates the period by determining the *events* such as zero crossing or GCI⁹ and then measuring the time period between each event.

⁹Glottal Closure Instant

3.9.1 Autocorrelation-based Period Estimation

This method employs the autocorrelation function [29] defined by

$$R(k) = \sum_{m=-\infty}^{\infty} x(m)x(m+k) \quad (3.52)$$

where $x(m)$ is any discrete-time deterministic signal. If $x(m)$ is periodic $R(k)$ also will be periodic with the same period. It is shown that, regardless of the time origin of the signal under consideration, the period of the signal can be estimated by finding the location of the maxima of the autocorrelation function and measuring the distance between the successive peaks. The autocorrelation function contains much more information about the detailed structure of the signal and hence we often consider the *short-time* autocorrelation function which is defined as

$$R_n(k) = \sum_{m=-\infty}^{\infty} x(m)w(n-m)x(m+k)w(n-k-m) \quad (3.53)$$

For finite duration sequences this expression can be modified as

$$R_n(k) = \sum_{m=0}^{N-1-k} [x(n+m)w'(m)][x(n+m+k)w'(k+m)] \quad (3.54)$$

where $w'(n) = w'(-n)$ and N is the block length selected. An important issue here is how N should be chosen to give a good indication of periodicity. Here the requirements are conflicting. Since the signal is basically non-stationary, N should be as small as possible. On the other hand, it should be clear that to get any indication of periodicity in the autocorrelation function, the window must have a duration of at least 2 periods of the waveform. It has been shown that [214], for best results, the block length has to be selected as $4M$ for voiced signals and as $2M$ for transient segments of signals, where M is the period.

Because of the finite length of the windowed signal segment involved in the computation of $R_n(k)$, there is less and less data involved in the computation as k increases. To

take care of the difficulty due to this roll off, sometimes samples from outside the interval from n to $n + N - 1$ are also involved in the computation. Strictly speaking, this leads to the cross-correlation computation rather than the autocorrelation and this is generally designated as the *modified* autocorrelation function. There are much more variants proposed [148], [215] on the basic scheme of period estimation using autocorrelation function to make it more adapted to different types of signals.

3.9.2 UDWT based Period Estimation

3.9.2.1 The Undecimated Discrete Wavelet Transform

One well-known drawback of wavelet transforms is their sensitivity to translations. This causes serious problems in some applications of wavelet methods. An immediate solution to this problem is the UDWT. This is a redundant type of shift-invariant WT which can be viewed as the DWT in which the multirate operations are eliminated. Hence as the number of levels of decomposition increases, the number of UDWT coefficients increases considerably, increasing the computational complexity as well. Although the UDWT is linear and shift-invariant, it has got the drawback of lack of orthogonality.

In discrete domain, the UDWT is implemented using the two QMF filters corresponding to the wavelet under consideration. The coefficients corresponding to the level j is computed using the following convolutions.

$$W_j(k) = h_j(k) * S_{j-1}(k) \quad (3.55a)$$

$$S_j(k) = g_j(k) * S_{j-1}(k) \quad (3.55b)$$

where $j = 1, 2, J, S_0(k) = x(k), h_0 = h, g_0 = g$. Here $h_j(k)$ and $g_j(k)$ are obtained by inserting $(2^j - 1)$ zeros between each of the two consecutive coefficients of the two filters $h(k)$ and $g(k)$, and J corresponds to the maximum level of decomposition.

3.9.2.2 Application of UDWT in period estimation

The UDWT can be used for period estimation, when special type of wavelets are used for analysis, as it exhibits local maxima at points of discontinuity or zero-crossing [155]. In order to detect a zero-crossing or GCI, a derivative function can be used as the wavelet. If the signal is filtered by a derivative function (such as Daubechies filters), a maximum will occur at each zero-crossing. Thus the time period between each maximum represents the pitch period of the signal at that moment. Let $\psi(t)$ be the mother wavelet with derivative properties. The functions

$$\psi(t) = 2^{k/2}\psi(2^k t) \quad (3.56a)$$

$$\varphi(t) = 2^{k/2}\varphi(2^k t) \quad (3.56b)$$

represent both the wavelet and scaling functions respectively at each scale. $\varphi(t)$ is a low pass function and is the conjugate mirror filter of $\psi(t)$, which is a high pass function. Since the range of period or fundamental frequency in any digital signal is between frequencies, say f_1 and f_2 Hz, the UDWT coefficients corresponding to levels contributing to this frequency band will show consistent peaks across successive levels. Therefore, the distance between these peaks can be measured as the period information. One can compute the scale parameters j_l & j_u corresponding to the frequency band of interest by selecting the nearest integer value of the parameters from the following expressions [154].

$$2^{j_l} = \frac{F_s}{f_1} \quad (3.57a)$$

$$2^{j_u} = \frac{F_s}{f_2} \quad (3.57b)$$

where F_s is the sampling rate of the signal.

3.10 The Linear Predictive Coding

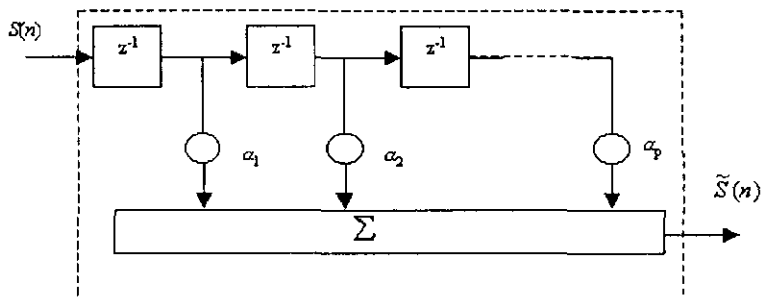
As applied to 1D signal processing, the term *Linear Prediction* refers to a variety of essentially equivalent formulations of the problem of modelling the signal. For 1D signal, the various formulations of linear predictive analysis have been:

- The Covariance method
- The Autocorrelation formulation
- The Lattice method
- The inverse Filter formulation
- The Spectral Estimation formulation
- The Maximum Likelihood formulation
- The Inner Product formulation

In each case the basic idea is that a signal sample can be approximated as a linear combination of past signal samples. By minimizing the sum of the squared differences (over a finite interval) between the actual signal sample and the linearly predicted ones, a unique set of predictor coefficients can be determined.

Because of the time-varying nature of signals, the predictor coefficients must be estimated from short segments of the signal, the segment (frame) size being dependent on the extent of non-stationarity of the signal and also its frequency characteristics. When the signal is transformed to another domain before the prediction is applied and transmitted, such schemes are called TC¹⁰. The signal in a TC is processed frame-by-frame and each frame is transformed using a discrete unitary transform. ie. $S = Ts$. The inverse transform is used for signal synthesis. ie. $s = T^{-1}S$.

¹⁰Transform Coders

Figure 3.11: Linear Predictor of order p

Predictive Coding systems utilize the data redundancy to realize straight forward reduction in bit rate for compression applications. The basic problem is to determine a set of predictor coefficients α_k directly from the signal in such a manner as to obtain a good estimate of the spectral properties of the signal being predicted. Because of the time-varying nature of non-stationary signals, the predictor coefficients must be estimated from short segments of the signal. A linear predictor with prediction coefficients α_k is defined as a system whose output is

$$\tilde{S}(n) = \sum_{k=1}^p \alpha_k S(n-k) \quad (3.58)$$

where p is the order of the predictor. The system function of the predictor is the polynomial $P(z)$, given by $\sum_{k=1}^p \alpha_k z^{-k}$. The structure of a p^{th} order linear predictor is shown in figure 3.11.

On using a linear predictor, the prediction error is defined as:

$$e(n) = S(n) - \tilde{S}(n) = S(n) - \sum_{k=1}^p \alpha_k S(n-k) \quad (3.59)$$

Similarly, the short-time average prediction error¹¹ is defined as:

$$\begin{aligned} E_n &= \sum_m e_n^2(m) = \sum_m (S_n(m) - \tilde{S}_n(m))^2 \\ &= \sum_m \left[S_n(m) - \sum_{k=1}^p \alpha_k S_n(m-k) \right]^2 \end{aligned} \quad (3.60)$$

where $S_n(m)$ is a segment of the signal that has been selected in the vicinity of sample n . To minimize the MSE, the values of α_k that minimize E_n by setting

$$\partial E_n / \partial \alpha_i = 0, \quad i = 1, 2, \dots, p,$$

thereby obtaining the equations

$$\sum_m S_n(m-1)S_n(m) = \sum_{k=1}^p \alpha_k \sum_m S_n(m-i)S_n(m-k); \quad 1 \leq i \leq p \quad (3.61)$$

Now, if we define $\phi_n(i, k)$ as $\sum_m S_n(m-i)S_n(m-k)$, equation 3.61 can be written more compactly as

$$\sum_{k=1}^p \alpha_k \phi_n(i, k) = \phi_n(i, 0); \quad i = 1, 2, \dots, p \quad (3.62)$$

This set of p equations in p unknowns can be solved in an efficient manner for the unknown predictor coefficients $\{\alpha_k\}$, that minimize the average squared prediction error for the segment $S_n(m)$.

There are two basic approaches [29] to proceed with the solution. One is the autocorrelation method and the other is the covariance method. In this thesis, only the former is discussed.

In the autocorrelation method it is assumed that the waveform segment $S_n(m)$, is identically zero outside the interval $0 \leq m \leq N-1$. i.e. $S_n(m) = S(m+n)w(m)$, where $w(m)$ is a finite word length window that is identically zero outside the interval $0 \leq m \leq N-1$. It is clear that, the error is large at the beginning of the interval $0 \leq m \leq p-1$ because we are trying to predict the signal from samples that have

¹¹To obtain the average, a division by the length of the data segment is necessary, but being irrelevant at this context, it is omitted.

arbitrarily been set to zero. Likewise the error can be large at the end of the interval $N \leq m \leq N+p-1$ because we are trying to predict zero from samples that are non-zero. For this reason, a window that tapers the segment $S_n(m)$ to zero is generally used for $w(m)$. It is shown [29] that

$$\sum_{k=1}^p \alpha_k R_n(|i-k|) = R_n(i); 1 \leq i \leq p \tag{3.63}$$

where $R_n(k) = \sum_{m=0}^{N-k-1} S_n(m)S_n(m+k)$. In matrix form the equation 3.63 can be expressed as

$$\begin{pmatrix} R_n(0) & R_n(1) & R_n(2) & \dots & R_n(p-1) \\ R_n(1) & R_n(0) & R_n(1) & \dots & R_n(p-2) \\ R_n(2) & R_n(1) & R_n(0) & \dots & R_n(p-3) \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ R_n(p-1) & R_n(p-2) & R_n(p-3) & \dots & R_n(0) \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \dots \\ \alpha_p \end{pmatrix} = \begin{pmatrix} R_n(1) \\ R_n(2) \\ R_n(3) \\ \dots \\ R_n(p) \end{pmatrix} \tag{3.64}$$

These equations are called the *normal equations*, *Yule-Walker Prediction equations*, or the *Wiener-Hopf equations*. The pxp matrix of autocorrelation values is a *Toeplitz* matrix and can be efficiently solved to derive the values of α_k . The most efficient of all methods for solving this system of equations is the *Durbin's recursive procedure* [216].

1984
1985
1986

Chapter 4

PSWT Based Linear Predictive Coding

This chapter describes the Pitch-Synchronous Wavelet Transform based method developed for the Linear Predictive Coding and Compression of pseudo-periodic signals, and its typical applications. Initially, the method has been presented in detail for general pseudo-periodic signal processing applications. Subsequently case studies have been presented on the application of the compression scheme on typical signals. Some techniques for feature enhancement and source dependent noise suppression in PSWT domain have also been briefly introduced. The method has been evaluated in terms of standard performance measures. The chapter is concluded by highlighting the results and discussing the important findings of the study.

4.1 Introduction

Conventional LPC methods and WT techniques have been in use for various signal processing applications. A noteworthy work combining the application of these techniques have been carried out by Ramakrishnan *et al* [59]. They have applied the LPC technique on the DWT of period-wise data for ECG data compression. In this chapter, a new method which is more beneficial for the analysis and compression of pseudo-periodic signals has been introduced. It combines the advantages of the LPC techniques and the WT methods with the inter-beat correlation of this class of signals.

The study has been presented for general pseudo-periodic signals, and case studies are presented on practical signals towards the end of the chapter. Practical signals may contain unwanted disturbances like random noise, harmonics, powerline interference etc. To reduce the effect of such disturbances, some feature enhancement techniques, which are exclusive to PSWT domain have also been developed and the details are briefly presented.

4.2 PSWT based LPC

The distinct steps involved in the PSWT based LPC system are listed below:

Step 1: Estimation of the local beat periods

Step 2: Organization of the data segment in the Pitch-Synchronous form

Step 4: Normalization of the PS data matrix

Step 5: Computation of the PSWT coefficients

Step 6: Estimation of the predictor coefficients

Step 7: Encoding the relevant information for transmission.

All the above operations are performed at the encoder end. At the decoder end all the inverse operations are performed in succession so that the signal is reconstructed.

4.2.1 Period Estimation

The effectiveness of PSWT based signal analysis techniques depend a lot on the accuracy with which the local periods are evaluated. For the PS representation, the local periods are to be evaluated for the whole signal continuously. In this work, two methods for period estimation have been adopted depending on the nature of signal under consideration, *viz.* the autocorrelation method and a WT based method. Suitable modifications have been made in the conventional algorithm to make the period estimation more adaptive for signals distorted by harmonics, noise, dc offset, etc.

4.2.1.1 The Autocorrelation based Period estimation

Autocorrelation based method is the most popular method for pitch determination of speech signals. Hence in this work, this method has been used to find the pitch period of signals like speech, vocal music etc. Modifications have been incorporated in the standard method so that even hard-to-detect periods could be estimated.

The signal was segmented into different frames, the size of which was selected to be roughly about 4 times the approximate local period. The autocorrelation function of each frame was computed using the simplified autocorrelation relation given by:

$$R(K) = \sum_{m=k}^N x(m)x(m-K); \quad 0 \leq K \leq N \quad (4.1)$$

When the signal is devoid of much disturbances, the period can be easily estimated as the distance between the successive peaks of the above autocorrelation function. But when the signal is contaminated by unwanted noise components, the period estimation becomes difficult and alternate techniques are to be employed. Some of the identified difficulties along with the proposed solutions are presented below.

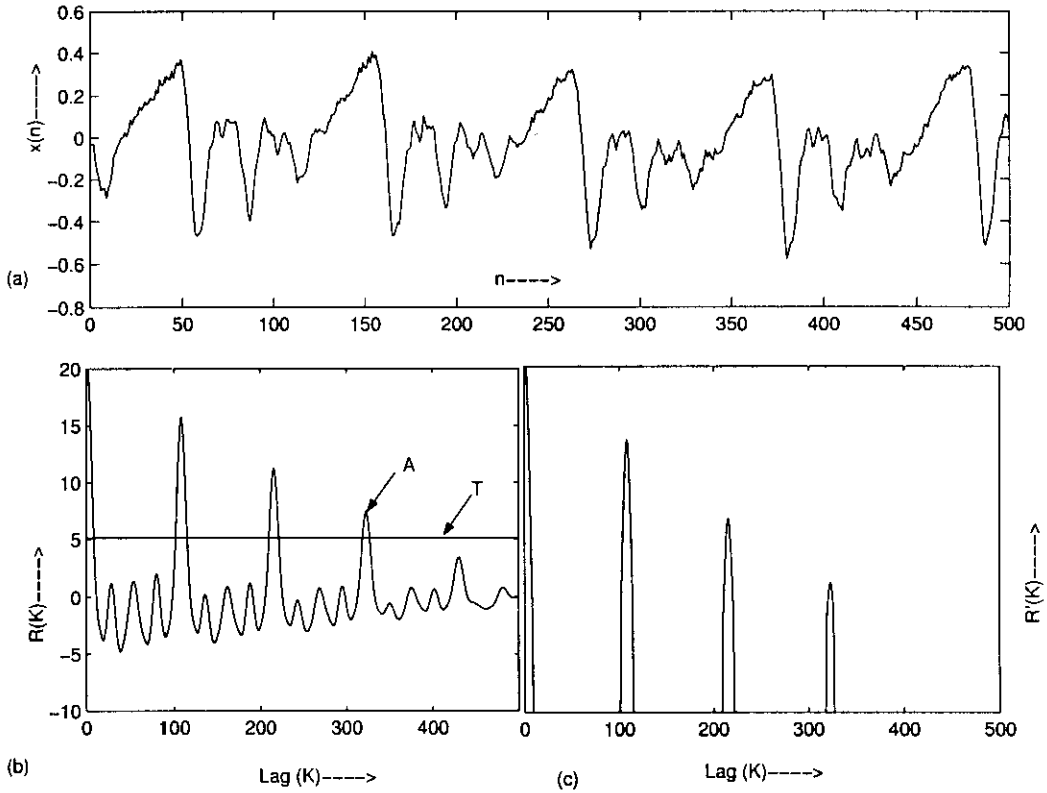


Figure 4.1: A distorted signal creating difficulty in period estimation (a)A segment of Original Signal (b)Autocorrelation Plot (c)Thresholded Autocorrelation plot

Consider one segment $x(n)$ of a signal as shown in figure 4.1(a). As seen from the autocorrelation plot (fig. 4.1(b)), many of the spurious spikes are smoothed out during the autocorrelation operation. But there are certain predominant intermediate peaks, which are still regular, making the period estimation difficult.

To get rid of this difficulty, a threshold $T = aR_{peak}$ was experimentally selected and $|R(K)| \leq T$ was made zero for all values of K , where a is the threshold factor and R_{peak} is the maximum value of the autocorrelation function. The resulting function $R'(K)$ is shown in figure 4.1(c). It may be noticed that, in this plot, the regions containing each local peak lies between two sets of zero-valued samples. The local peak from each of

these regions which are separated by sets of zero-valued samples are estimated and the distance between the consecutive local peaks are taken as the local pitch period. This method helps in partially eliminating the acceptance of the intermediate peaks which are not spaced by local periods.

Another difficulty in pitch estimation results from the unsymmetry between the positive and negative half cycles of the signal. One such case pertaining to a musical instrument tone produced by an electric guitar, is shown in figure 4.2. The corresponding autocorrelation plot is shown in (b). Here, if the negative peaks are considered for the period estimation, it gives almost double the actual period count. To take care of such situations, the peak detection is done for both positive and negative half cycles and the one with minimum peaks is further processed for detecting the local periods.

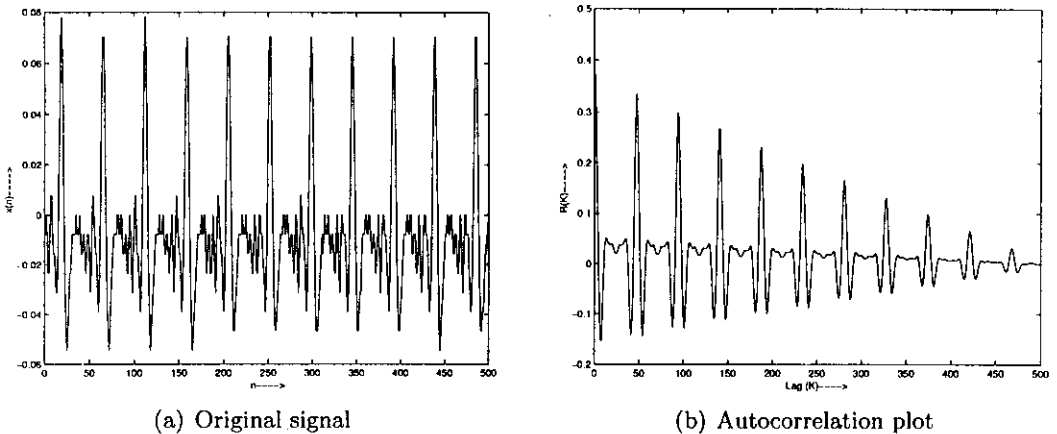


Figure 4.2: A music note produced by an electric guitar and its autocorrelation plot.

On estimating the period as the difference between the successive local peaks in the autocorrelation plot, a common difficulty encountered is the presence of multiple dominant peaks within the same period. So is the case with the female voice signal shown in figure 4.3. To avoid the contribution from the non-predominant peaks, the detected periods are compared with the local maximum. All detected peaks which are closer than by a preset percentage of the maximum period are discarded while selecting

the final periods. If the current one is less than a preset percentage of the previous, the current local peak is discarded and the succeeding peak is selected.

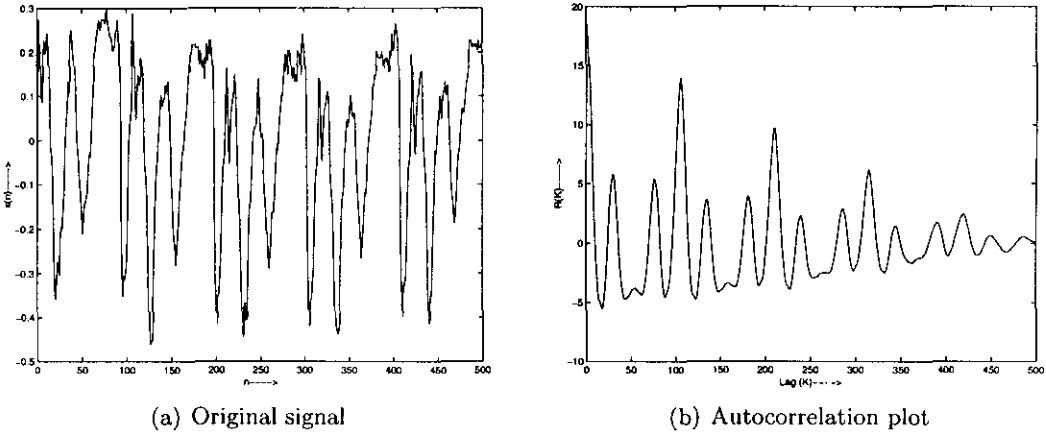


Figure 4.3: Signal showing multiple dominant peaks in the Autocorrelation plot.

Yet another difficulty in local peak detection is posed by the linear decrease of the magnitude of the successive peaks in the autocorrelation plot of a frame. As shown in figure 4.1(b), all local peaks after the point 'A' in the frame under study, are below the threshold and hence they could not be detected as such. For the PS representation of the signal, the local periods of the whole signal are to be determined. Therefore in the original signal, the sample corresponding to the peak 'A' is taken as the beginning of the next frame and the process is repeated. This procedure ensures that the local peaks of the entire input signal is detected irrespective of the selected frame size for the autocorrelation plot.

In rare occasions the signal may contain dc shift which also creates difficulty in period estimation. The presence of dc shift will distort the autocorrelation plot as depicted in figure 4.4. By subtracting the mean value of each segment before taking the autocorrelation plot, the undesirable effects of the dc shift could be eliminated.

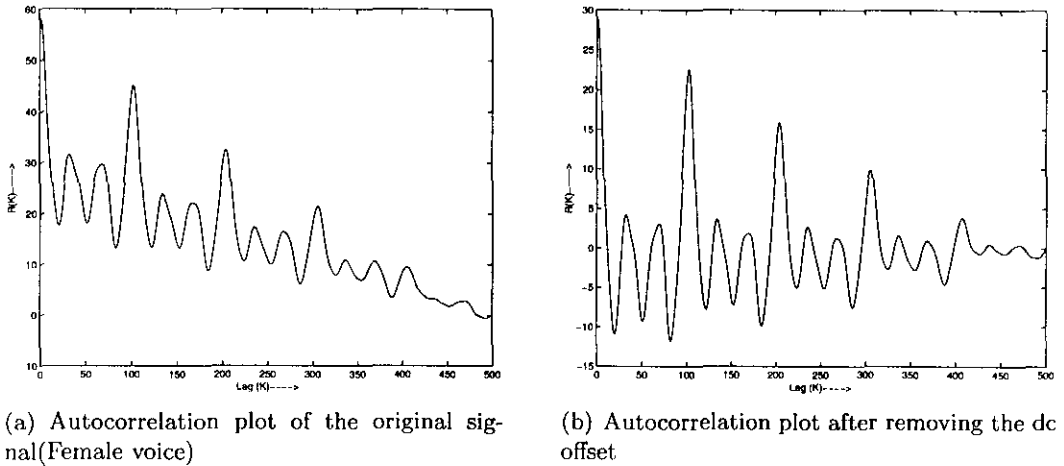


Figure 4.4: Distortion in the Autocorrelation plot due to dc shift in the signal.

4.2.1.2 The WT based Period estimation

For signals with large bandwidth, the autocorrelation based method of period estimation often results in inaccurate results. Musical sounds are one such category. They are harmonically rich and can contain inharmonic partials as well. For a minimum delay, the pitch has to be determined during the attack of a note, where the sound is noisiest and most harmonically complex. The pitch can also vary constantly without the onset of notes. Many ambiguously pitched sounds exist, such as multi-phonics, key clicks or unpitched sounds, which must be dealt with in a consistent manner. The wavelet based pitch tracker introduced in this thesis is able to estimate the pitch trajectory much more precisely and quickly than many classical techniques. This is mainly because, WT is a more natural way for windowing large bandwidth signals. Moreover, since the wavelet partials at different levels being dyadically related, the successive elimination of harmonics is possible without affecting the information content related to the fundamental period. The general transients that are present in the musical notes contain wide frequency components, making the smoothing operation difficult.

Another class of signals where the WT method is found superior is the ECG. The

frequency of ECG signals is small, whereas the QRS complex, being transient in nature, contains large bandwidth components.

In this method, the peaks corresponding to the fundamental period are made prominent by decomposing the signal in the UDWT domain using an appropriate wavelet filter. The level to which the decomposition is performed is dependent on the frequency content and sampling rate of the signal, making the technique to be signal dependent. When the successive levels show consistent peaks so that the position of these peaks can be uniquely determined, the decomposition is stopped and the period is computed as the distance between such peaks. For example, consider the signals shown in figure 4.5. The peaks corresponding to the fundamental period is not made uniquely detectable by even 8 levels of filtering in the case of the *tambura* note whereas it is achieved by 2 levels of filtering for the *flute* signal.

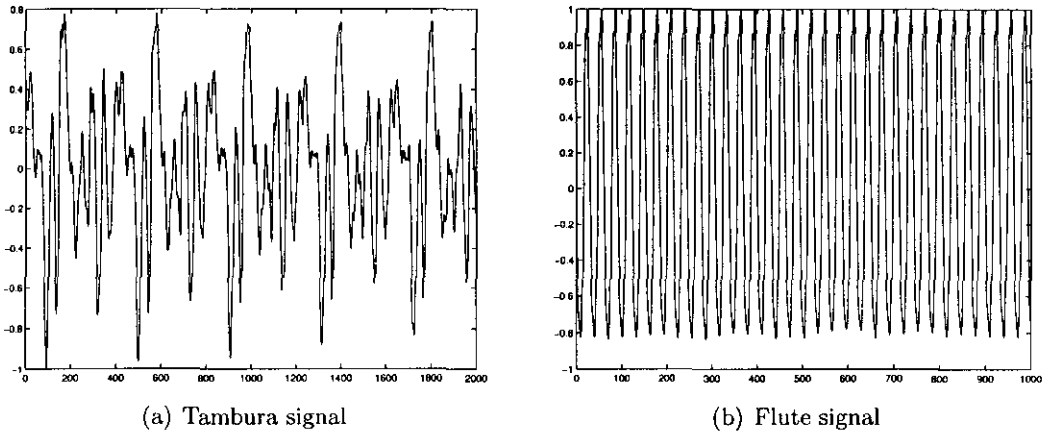


Figure 4.5: Signals showing wide variation in UDWT levels for period estimation.

Even after performing the filtering operation to the desired level, due to irregularities in the signal there can be other peaks in the near vicinity of the actual local peak. This can result in wrong period estimation. Hence the techniques employed along with autocorrelation method is adopted in this method also.

4.2.2 Pitch-Synchronous Representation

Once the local periods are estimated, any 1D data could be arranged in the PS form [135]. The details of this PS representation is given below.

Assuming $P(k)$ to be the sequence of integer local pitch periods extracted from a typical segment of the signal $x(n)$, the segment under consideration can be expressed in PS form as:

$$v_q(k) = \sum_i x(i)\delta(i - q - \sum_{r=0}^{k-1} P(r)), \quad i \in Z \quad (4.2)$$

where $q = 0, 1, 2, \dots, P(k) - 1$, and $k = 0, 1, \dots, N - 1$. Here the index q is the inter-period-count index and k is the period-count index. N is the total number of pitch periods that make up the segment under consideration. $v_q(k)$ gives a 2D representation of the 1D signal.

The signal $x(n)$ is now demultiplexed into q channels, the p^{th} channel being constituted by the p^{th} sample of each of the N pitch periods. The surface plot of a typical pseudo-periodic signal in PS form is given in figure 4.6.

The original signal can be reconstructed back from the PS form using the expression:

$$x(n) = \sum_i \sum_{q=0}^{P(i)-1} v_q(i)\delta(n - q - \sum_{r=0}^{i-1} P(r)) \quad (4.3)$$

4.2.3 Normalization

To achieve maximum interbeat correlation of the WT coefficients, beat normalization is carried out as done by Ramakrishnan et al. [59]. This has two stages, the period normalization and the amplitude normalization. Each beat in the PS data matrix is normalized separately. For period normalization, multirate techniques [217] are employed. This involves sampling rate change by different fractional factors for each period of the PS data matrix, employing equation 3.33.

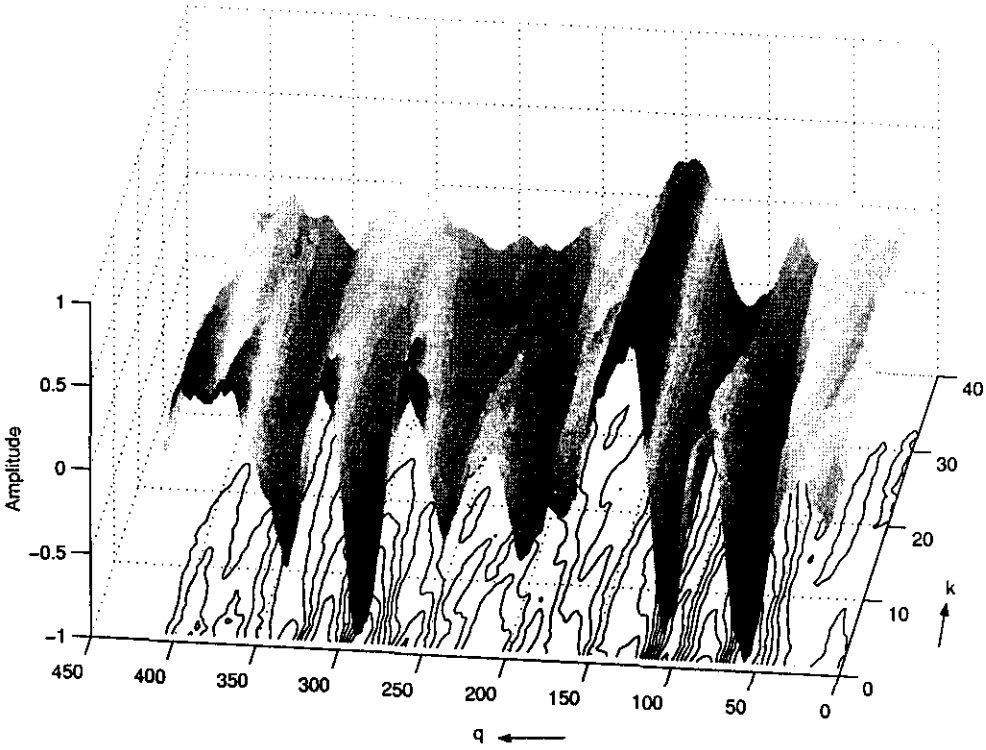


Figure 4.6: Surface plot of a signal represented in PS form (Tambura note).

The period normalization is achieved by first upsampling each period data by a factor L , which is an integral multiple of the maximum cycle period, p_m . This is followed by downsampling by the respective beat period.

After the period-normalization, equation 4.2 gets modified as

$$\hat{v}_q(k) = \sum_{r=0}^{P(r)-1} v_q(r) h_k(kP(k) - rL) \quad (4.4)$$

where $v_q(k)$ stands for the k^{th} column of the PS data and $\hat{v}_q(k)$ is the corresponding period normalized data vector. $h_k(z)$ is the impulse response of the smoothing filter. This converts the cycles of differing periods into cycles of a constant period, thus eliminating the cycle variability as shown in figure 4.7.

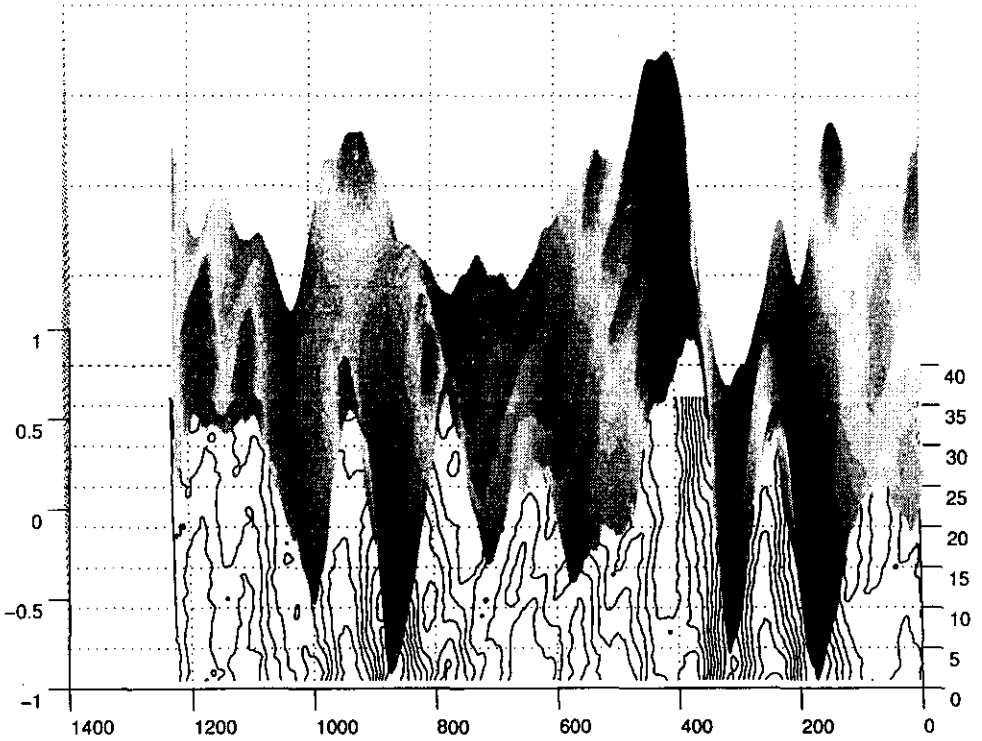


Figure 4.7: Surface plot of the signal shown in figure 4.6 after Normalization (period normalized to 3 times the maximum period)

The fixed period is selected based on the maximum local period of the signal under consideration. The modified sampling rate is selected still satisfying the Nyquist criterion. Since a sampling rate higher than the existing one is selected, there won't be any signal distortion provided the alias frequencies are eliminated by proper filtering using $h_k(z)$.

Amplitude normalization brings further similarity between the consecutive periods. Each sample of each period is divided by the magnitude of the largest sample of the corresponding period data. This makes the highest amplitude samples of each period equal to unity. Thus, the variations between the magnitudes of different cycles are

minimized. It can be seen that PAN¹ does not introduce any distortion in the signal whereas it will enhance the inter-period correlation.

4.2.4 Computation of PSWT Coefficients

For computing the PSWT, the method of DWT computation is utilized. Since after normalization, the PS data resembles a truly periodic signal, the PSWT computation reduces in effect to MWT computation. Here the DWT is computed row-wise on $\hat{v}_q(k)$ for each $q = 1, 2, \dots, L, q \in N$. The scheme for computation is given in figure 3.10. The PSWT coefficients are arranged in a matrix in the following format:

$$\begin{pmatrix} a_J^1 & d_J^1 & d_{J-1}^1 & \dots & d_2^1 & d_1^1 \\ a_J^2 & d_J^2 & d_{J-1}^2 & \dots & d_2^2 & d_1^2 \\ a_J^3 & d_J^3 & d_{J-1}^3 & \dots & d_2^3 & d_1^3 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ a_J^L & d_J^L & d_{J-1}^L & \dots & d_2^L & d_1^L \end{pmatrix}$$

where a_j^p and d_j^p are row vectors corresponding to the approximation and the details respectively of the WT of the p^{th} row of the data matrix at the j^{th} level.

In a parallel processing environment, the DWT and IDWT of all the channels may be computed concurrently, as they are independent of each other with absolutely no data dependency.

4.2.5 Computation of Predictor Coefficients

In this work, the predictor coefficients for the prediction of PSWT data are chosen according to the minimum error criterion. They are determined by minimizing the sum of squared differences (over a finite interval) between the actual PSWT coefficients and the linearly predicted ones using the autocorrelation method. The block schematic of the predictor estimation system is shown in figure 4.8.

The optimized set of predictor coefficients depend on both the order of predictor and the logical selection of the data input. The PSWT coefficients which form the data

¹Period and Amplitude Normalization

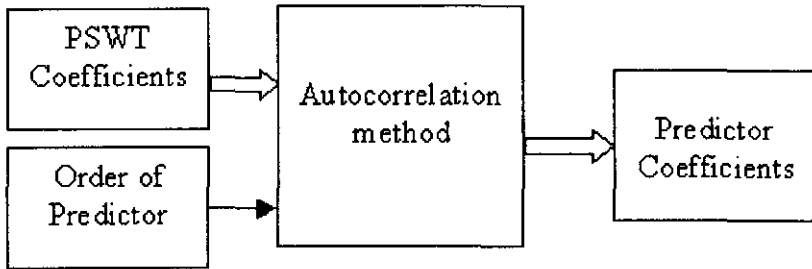


Figure 4.8: Block schematic of the predictor estimator.

input contain groups of varying spectral characteristics corresponding to different levels of decomposition. Hence different alternatives like separate predictors for prediction of each level-wise data, single predictor for complete data prediction etc. were tried out. Finally level-wise prediction, *i.e.*, different sets of predictors for predicting each level of transform coefficients, was selected.

Since the PSWT acts on homologous points of the normalized data, the average information contained over different beat periods get well localized in the scaling residue of the PSWT. Hence most of the energy contained in the signal appears in the scaling residue itself. This includes all the spectral components centered on the fundamental and its harmonics. On the contrary, the wavelet partials carry other informations, which lie near to the harmonics, indicating the degree of period-to-period fluctuations present in the signal. This is obviously different from the result of DWT on the same signal, as the wavelet coefficients in a DWT are indicative of intra-period variations, independent of the global changes in the signal. Taking advantage from the above facts, the prediction of the PSWT coefficients is performed with a minimum order linear predictor. Due to the strong correlation of successive beats, this method of prediction yields better results when compared with prediction of DWT coefficients of each beat of the signal as done in [59].

4.2.6 Signal Encoding and Decoding Scheme

After estimating the predictor coefficients, they are encoded for transmission to the decoder end. Along with these, the first few rows of the PSWT coefficients and other auxiliary information like mean beat period P_{mean} , the AASF², deviation of each beat period from P_{mean} and AASF, order of the predictor p , level of decomposition J , wavelet used for the study etc. are also transmitted to the decoder end.

Assuming the data transmission to be undistorted, at the decoder end, the PSWT coefficients of the data are extracted using the predictor coefficients and other auxiliary information, which were transmitted along with the predictor. The PS data is reconstructed using IPSWT methods. This is the normalized data which are denormalized employing the multi-rate techniques. The 1D signal is now obtained back from this denormalized PS data using the local period information. The block schematic for the PSWT based linear prediction along with the associated encoding and decoding scheme is shown in figure 4.9.

For different class of signals, the parameters for coding were optimized by changing the following variants.

1. order (support) of wavelet used for decomposition
2. level of decomposition
3. number of columns used for evaluating the predictor coefficients
4. without/with normalization
5. with and without transmitting the residue of prediction
6. order of predictor

²Average Amplitude Scale Factor

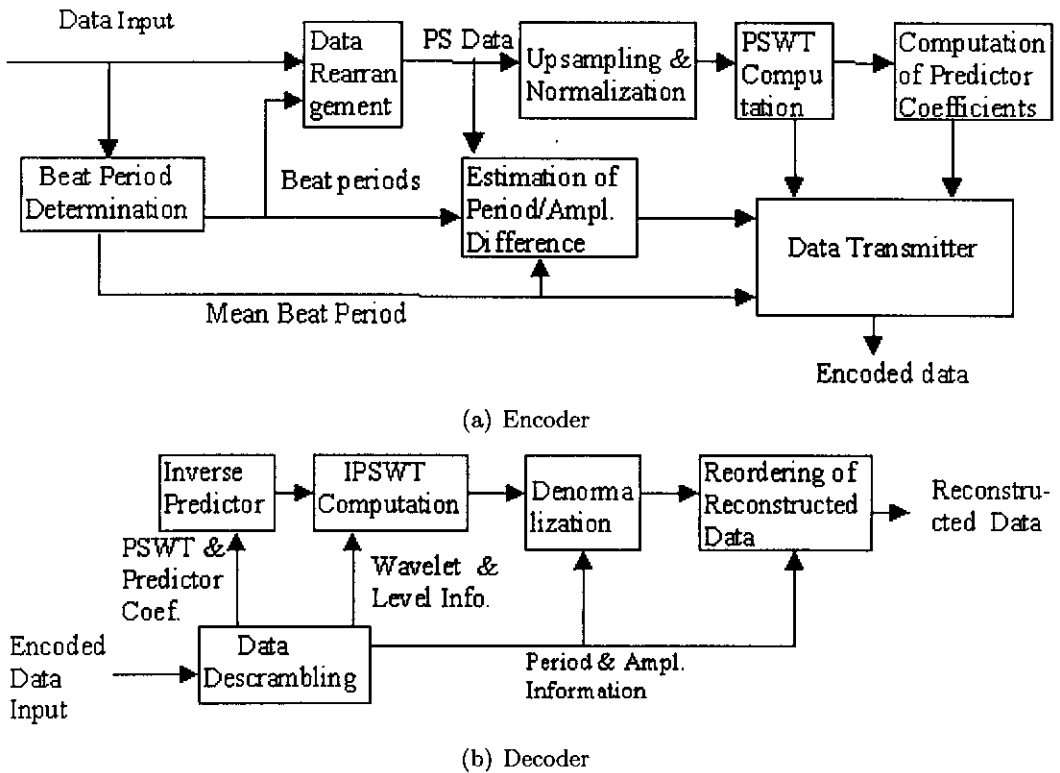


Figure 4.9: Block schematic of the PSWT based predictive coder.

4.2.7 Performance Evaluation

The performance of the coder was evaluated for all practical combinations of the above variants in terms of both objective and subjective measures depending on their applicability. The objective measures include, SNR, CR, NRMSE, NMAE [59]etc. whereas the important subjective measure used was the MOS [218].

SNR is one of the common objective measures for evaluating the performance of a compression algorithm. This is expressed as:

$$SNR = 10 \log_{10} \frac{\sum_{n=0}^{N_t-1} |S(n) - \tilde{S}|^2}{\sum_{n=0}^{N_t-1} |e(n) - \bar{e}|^2} \quad (4.5)$$

where $S(n)$ is the original data while $\tilde{S}(n)$ is the reconstructed data, $e(n) = (S(n) -$

$\tilde{S}(n)$), and N_l the length of the data segment. The SNR is long term measure for the accuracy of reconstruction and as such it tends to *hide* temporal reconstruction noise particularly for low level signals. Temporal variations of the performance can be better detected and evaluated using the *segSNR*, which is a short-time Signal-to-Noise ratio. i.e. by computing the *SNR* for each N -point segment of the data. A performance measure that better exposes a weak signal performance is the *segSNR*, which is given by:

$$segSNR = \frac{10}{L} \sum_{i=0}^{L-1} \log_{10} \left\{ \frac{\sum_{n=0}^{N-1} S^2(iN+n)}{\sum_{n=0}^{N-1} (S(iN+n) - \hat{S}(iN+n))^2} \right\} \quad (4.6)$$

In the above expression, since the averaging operation occurs after the logarithm, the *segSNR* more penalizes the coders whose performance is variant. L is the total number of data segments in the original data.

NRMSE, *NMAE* and the *CR* were evaluated using the following formulae.

$$NRMSE = \sqrt{\frac{\sum_{n=0}^{N_l-1} (S(n) - \tilde{S}(n))^2}{\sum_{n=0}^{N_l-1} S^2(n)}} \times 100 \quad (4.7a)$$

$$NMAE = \sum_{k=0}^{N_b-1} \frac{\max |v_q(k) - v_{qr}(k)|}{\max(v_q(k)) - \min(v_{qr}(k))} \times \frac{100}{N_b} \quad (4.7b)$$

$$CR = \frac{N_l n_s}{N_b n_s p + p(J+1)n_p + N_a} \quad (4.7c)$$

Here \bar{S} and \bar{e} represent the mean of the sequences $S(n)$ and $e(n)$ respectively. N_l , N_b , p , n_s , n_p and J stands for the number of samples in the original signal, number of beats, order of predictor, bit size of original samples, bit size for coding the predictor coefficients and the level of decomposition respectively. N_a is the auxiliary data to be sent to the decoder end which include the wavelet information, mean beat period, difference of each period from the mean, AASF, and the period-wise amplitude scale factor.

For signals having importance to perceptual characteristics, NMAE and NRMSE are not considered significant. Instead, it is customary to conduct subjective evaluation. For this listening tests [218] were conducted using ten subjects. Special care was taken to eliminate external interference, background noise, and echo-effects. Training sets were used to familiarize the subjects participated in the listening test. They were asked to rate the quality as *excellent*, *good*, *fair*, *poor*, or *bad*. These ratings were allotted grade numbers 5, 4, 3, 2, and 1, respectively. The MOS value was calculated by taking the arithmetic mean of the grades voted by them.

4.3 Case Studies

Three general classes of pseudo-periodic signals are considered for case study *viz*, ECG signal of both healthy and pathological subjects, vocal sound from both male and female speakers, tones produced by musical instruments.

As data samples for the study, both the research data available in the web for public use, and also local data from nearby institutions and recorded in our own laboratory were used. MUMS³ are one such standard set used for the study of instrument tones. For study on ECG data, other than samples arranged from local multi-speciality hospital, selected records from MIT-BIH data base also have been used. In all the cases it was observed that, the compression based on PSWT decomposition followed by prediction *along beats* gives better results compared to period-wise DWT followed by prediction *across beats*. The residue was totally discarded, as the SNR values obtained even without them are found good enough for the purpose. The application of separate set of predictors for the PSWT coefficients at each level is found to perform better compared

³McGill University Master Samples: This is the brain-child of Frank Opolko and Joel Wapnick. They released 11 CDs containing the sounds of most standard classical and popular musical instruments, which has a unique feature that every note of every instrument is included in this sampling CD set. All sounds are recorded at 44.1kHz. MUMS sounds are used for teaching and research in over 100 universities world wide as these samples are recorded in ideal environment with high precision recording set up to ensure that the special timbral personalities of each instrument are retained without any recognizable distortion.

to one set of predictor for the whole data.

4.3.1 Case Study 1: ECG signal

State-of-the-art medical technology in the field of modern medicine aims at a truly remote patient care by focussing on issues like telemedicine, tele-surgery etc. [219], [220]. This requires the simultaneous transmission and monitoring of huge data pertaining to various biomedical signals. Besides, signals like ECG, being very vital in patient monitoring, need to be continuously transmitted to the doctors' desk with a high degree of accuracy, retaining the critical morphological information in the signal. In due consideration of the importance, the developed method has been applied for ECG data compression and the results of study are presented in this section.

ECG data from the MIT-BIH data base -both the compression data base and the arrhythmia database- and a few data records collected from a local super specialty hospital pertaining to normal and pathological subjects have been used for this case study. The details of the signals are given in the table 4.1.

Sl. no.	Data base	Signal details
1	MIT-BIH compression data base	Sampling rate =250Hz Resolution =12 bit Number of samples =5120 each
2	MIT-BIH arrhythmia data base	Sampling rate =360Hz Resolution =11 bit Number of samples =10000 each
3	Local data base	Sampling rate =250Hz Resolution =16 bit Number of samples =7500 each

Table 4.1: ECG Records used for compression study

The method of signal coding on one of the data records is illustrated below. The ECG considered is of 11 – bit resolution sampled at 360Hz bearing record number 101 of the MIT-BIH arrhythmia data base, which is shown in figure 4.10.

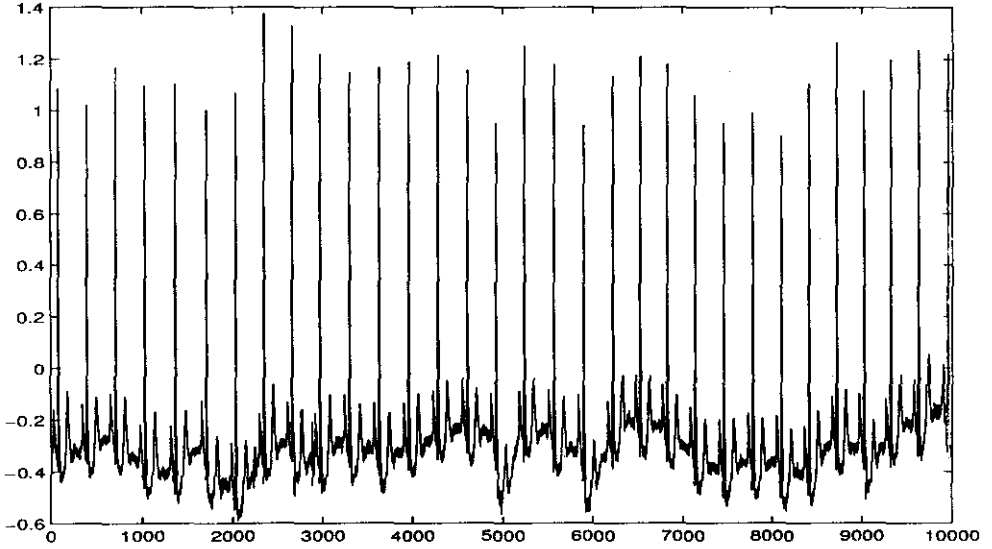


Figure 4.10: Original ECG data segment at lead 1 of the arrhythmia data record no.101

The local beat periods are evaluated using the UDWT based method. The UDWT partials are successively computed till the coefficients at consecutive levels have shown identical peaks. Here it is seen that the peaks in level 2 and 3 are identical as depicted in figure 4.11.

The local periods are estimated as the distance between the successive peaks in the 3rd level UDWT coefficients, which are given by the following row vector:

$$\begin{aligned}
 P(k) = & [314 \ 315 \ 321 \ 336 \ 344 \ 324 \ 313 \ 312 \ 311 \ 329 \ 330 \\
 & 329 \ 322 \ 327 \ 318 \ 313 \ 328 \ 330 \ 324 \ 305 \ 302 \ 313 \\
 & 320 \ 323 \ 320 \ 309 \ 307 \ 304 \ 295 \ 308 \ 324]
 \end{aligned} \tag{4.8}$$

The data segment is now arranged in the PS form using equation 4.2, the surface plot of which is given in figure 4.12.

The maximum local period of the whole segment under consideration is seen to be 344 samples. Each beat was period normalized to this maximum period. The MBP⁴

⁴Mean Beat Period

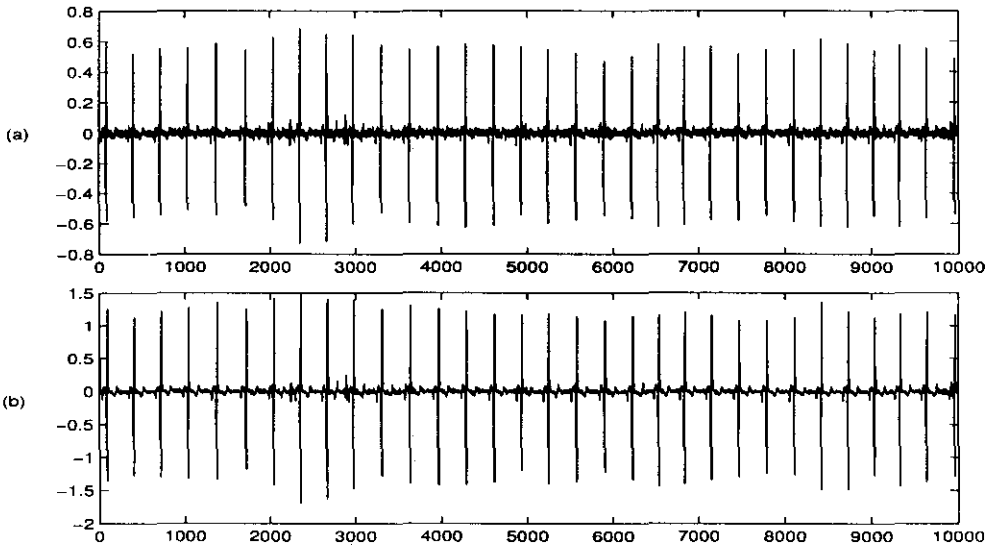


Figure 4.11: UDWT coefficients of the ECG data at different levels (a) 2nd level UDWT details (b) 3rd level UDWT details

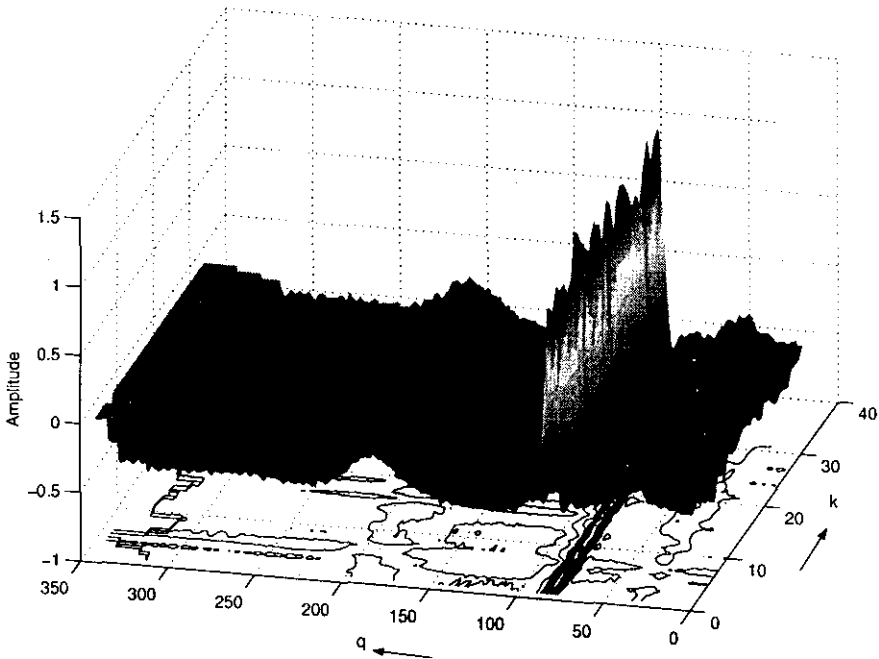


Figure 4.12: Surface plot of the ECG data shown in fig. 4.10

was found to be 318. The surface plot of a section of the normalized ECG segment is given in figure 4.13.

A 3-level PSWT decomposition was performed on the PS data matrix using the 'Haar' wavelet. The resulting PSWT coefficients are arranged in a matrix form, whose surface plot is given in figure 4.14.

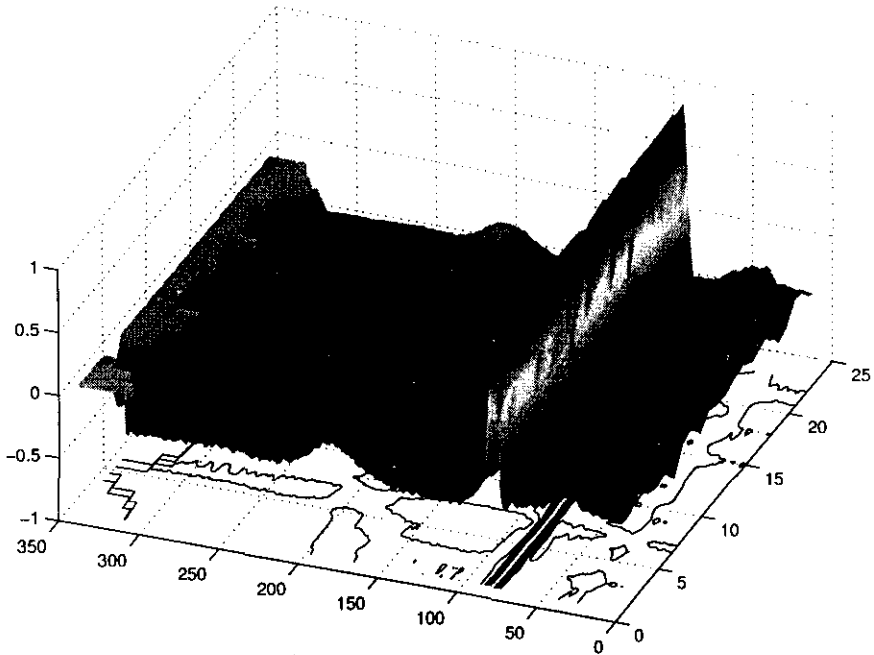


Figure 4.13: Surface plot of the amplitude normalized ECG shown in fig. 4.10

For prediction purpose, a 4th order predictor was chosen. The predictor coefficients were separately estimated for the approximation and details at each level of decomposition. They were estimated using the PSWT coefficients corresponding to *five* columns

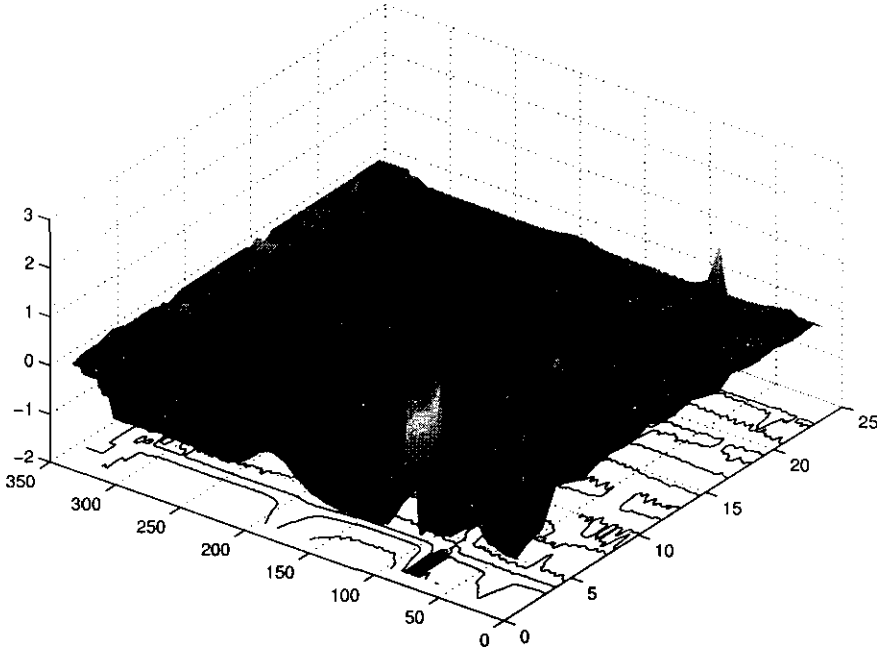


Figure 4.14: Surface plot of the PSWT coefficients of the ECG in fig. 4.10

at that level. The predictors in this case are given by the following row vectors.

$$\alpha_{d1} = [-1.8083 \quad 0.7042 \quad 0.3803 \quad -0.2643] \quad (4.9a)$$

$$\alpha_{d2} = [-1.2993 \quad 0.3309 \quad 0.0846 \quad -0.0747] \quad (4.9b)$$

$$\alpha_{d3} = [-1.2664 \quad 0.3352 \quad 0.0757 \quad -0.1029] \quad (4.9c)$$

$$\alpha_{a3} = [-0.9859 \quad 0.0651 \quad 0.0758 \quad -0.1121] \quad (4.9d)$$

where α_{dj} , $j = 1, 2, 3$ stands for the predictor related to the j^{th} level details and α_{a3} , that for the 3^{rd} level approximation.

The first 4 rows of PSWT coefficients are used at the decoder end along with the predictor and other auxiliary information to reconstruct the original signal.

The PSWT coefficients were reconstructed back using the inverse LPC techniques. The PS data matrix is recovered using the IPSWT computation as given in figure 3.10.

Employing the multirate techniques, the PS data were denormalized using the MBP and the period difference information. Equation 4.3 was then used to regenerate the original 1D signal. Figure 4.15 shows the reconstructed signal the error in reconstruction. It may be noted that the error of prediction is very insignificant.

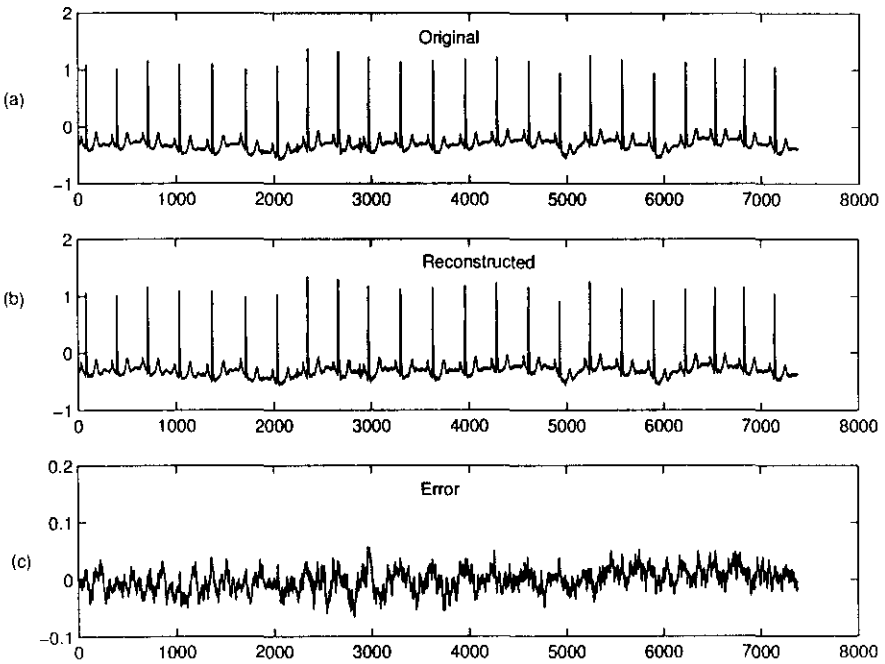


Figure 4.15: Reconstructed ECG and the error of prediction

The variants as discussed in section 4.2.6 were separately analyzed to optimize the parameters used for the encoding process. Tables 4.2 - 4.6 show the effect of change of these variants on the performance of the coder with respect to an ECG segment from the local data base.

DATA	WLT	J	PNF	NB	p	SNR	NRMSE	NMAE	CR
Local	db1					43	0.75	0.90	72.0
	db2					43	0.75	0.90	66.0
	db3					43	0.74	0.90	63.0
	db4	2	5	5	2	42	0.75	0.90	58.0
	db6					42	0.78	0.91	51.0
	db8					42	0.77	0.91	46.0
	db10					42	0.80	0.91	42.0
	db20					42	0.81	0.91	29.0

Table 4.2: Wavelet Optimization for ECG data compression.

DATA	WLT	J	PNF	NB	p	SNR	NRMSE	NMAE	CR
Local	db1	1				42	0.76	0.90	73.0
		2				43	0.75	0.90	72.0
		3	5	5	2	42	0.76	0.91	71.0
		4				42	0.77	0.91	70.0
		5				42	0.78	0.91	69.0

Table 4.3: Optimization of level of decomposition for ECG data compression.

DATA	WLT	J	PNF	NB	p	SNR	NRMSE	NMAE	CR
Local	db1	2	0			17	14.44	9.08	94.0
			1			20	10.42	6.37	82.0
			2			31	2.76	1.63	82.0
			3			38	1.32	0.85	82.0
			4			41	0.89	0.84	82.0
			5	5	2	43	0.75	0.90	82.0
			6			43	0.70	0.95	82.0
			7			43	0.69	0.99	82.0
			8			43	0.69	1.02	82.0
			9			43	0.70	1.04	82.0
10			43	0.70	1.05	82.0			

Table 4.4: Optimization of the Period normalizing factor for ECG data compression.

DATA	WLT	J	PNF	NB	p	SNR	NRMSE	NMAE	CR
Local	db1	2	5	1	2	42	0.77	0.90	72.0
				2		42	0.77	0.90	72.0
				3		42	0.76	0.90	72.0
				4		42	0.75	0.90	72.0
				5		43	0.75	0.90	72.0
				6		43	0.75	0.90	72.0

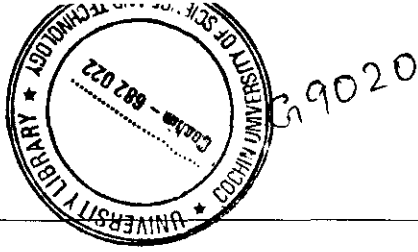
Table 4.5: Optimization of the number of beats needed for computing the predictor for ECG data compression.

DATA	WLT	J	PNF	NB	p	SNR	NRMSE	NMAE	CR
Local	db1	2	5	5	1	24	6.33	4.86	108.0
					2	43	0.75	0.90	72.0
					3	43	0.71	0.92	54.0
					4	43	0.70	0.93	43.0
					5	43	0.69	0.94	36.0
					6	43	0.68	0.94	31.0
					7	43	0.67	0.94	27.0
					8	44	0.67	0.94	24.0
					9	44	0.67	0.94	22.0
					10	44	0.66	0.94	20.0

Table 4.6: Optimization of the predictor order for ECG data compression.

Parameter	Abbreviation	Value
Wavelet	WLT	db1
Level of decomposition	J	2
Order of prediction	p	2
Period normalizing factor	PNF	5
Number of beats for evaluating the predictor coefficients	NB	5

Table 4.7: Optimized parameters for ECG data compression



Based on the observations on different signals as given in the table 4.1, the optimal parameters for the predictive coding of ECG data were selected which is given in Table 4.7. Wherever the improvement in the performance measures on variation of a parameter is not much appreciable in comparison with the computational burden, the one with minimum computation is selected. Table 4.8 summarizes the results obtained with the optimized parameters for the encoder, on a few ECG signals including that of pathological subjects.

The values of NMAE and NRMSE are high for pathological subjects due to large variation in period-to-period ECG. In such cases, the quality of reconstruction is found to have increased with the order of prediction. For normal subjects the CR went even up to 72 with SNR greater than 40 with NRMSE and NMAE less than unity, which is well acceptable reconstruction quality. Even higher rates of compression could be achieved by logically eliminating the less significant coefficients from the PSWT coefficients.

4.3.2 Case Study 2: Vocal sound

Speech and music signals comprise of voiced, unvoiced, silent and transition regions. The voiced region, which forms the major part in such signals is of pseudo-periodic in nature. Hence such regions were identified and segmented out to apply the compression technique. An automatic method that has been developed for this application is presented in *Appendix B*. Both male and female sounds were used for the study. The signals were sampled at different frequencies like $44.1kHz$, $22.05kHz$, $16kHz$, and $8kHz$.

The optimum parameters for the compression study has been experimentally obtained as in the case of the ECG compression. The effect of different variants on the performance of the compression scheme is shown in figures 4.16- 4.19.

It may be noted that, the SNR values in almost all cases are greater than 15 resulting in very good perceptual quality for speech. Hence in the case of speech, a MOS evaluation turns out to be irrelevant. For musical signals, slightly higher SNR is required for to

Wavelet used: Haar Order of Predictor: 2			Level of Decomposition: 2 Period Normalizing factor: 5			
Data base	Record No.	Signal details	SNR db	NRMSE (%)	NMAE (%)	CR
Local	Record 1	250 Hz, 16 bit	43	0.75	0.90	72.0
	Record 2		42	0.76	0.89	72.0
	Record 3		31	2.84	3.09	59.0
	Record 4		39	1.12	1.59	60.0
	Record 5		30	3.03	6.43	59.0
MIT-BIH -ADB	101	360 Hz, 11 bit	28	2.18	6.74	84.0
	103		30	2.51	5.85	84.0
	102		29	2.30	8.19	68.0
	104		32	2.01	5.54	76.0
	105		32	2.08	6.51	77.0
MIT-BIH -CDB	13420_06	250 Hz, 12 bit	30	3.13	3.67	70.0
	13420_10		28	3.29	6.46	65.0
	13420_02		30	2.53	5.74	89.0
	13420_03		33	1.75	2.88	87.0
	13420_07		38	1.15	1.95	85.0

Table 4.8: Summary of compression study on different ECG records

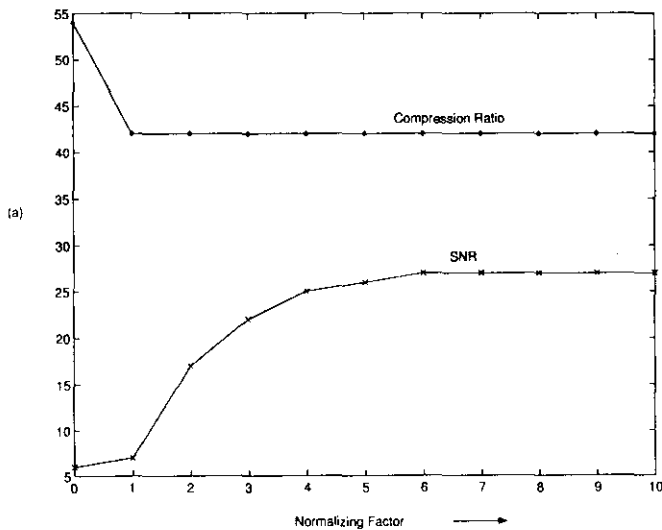


Figure 4.16: Performance of PSWT based LPC on a segment of male voice sampled at 16kHz-variation with Normalizing factor

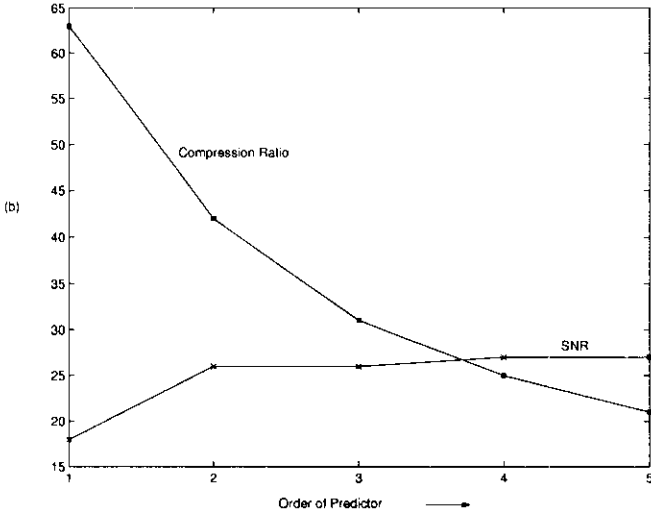


Figure 4.17: Performance of PSWT based LPC on a segment of male voice sampled at 16kHz-variation with Order of Predictor

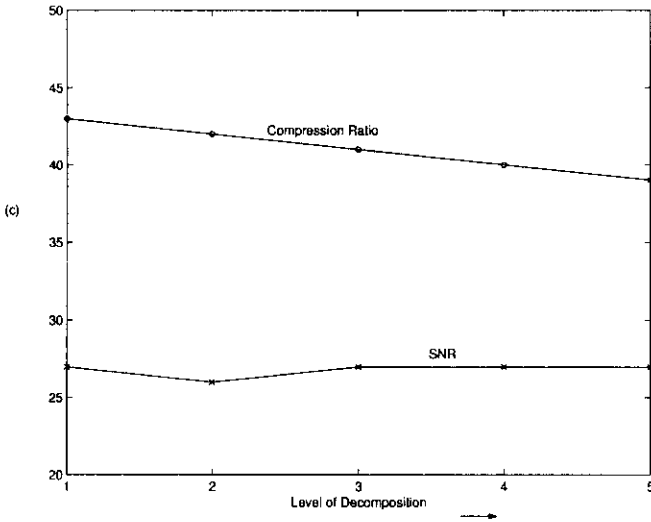


Figure 4.18: Performance of PSWT based LPC on a segment of male voice sampled at 16kHz-variation with Level of Decomposition

rank the MOS as *excellent*. The parameters optimized for voice compression study are summarized in table 4.9.

Parameter	Value
Wavelet	db1
Level of decomposition	1
Order of prediction	2
Period normalizing factor	5
Number of beats for evaluating the predictor coefficients	5

Table 4.9: Optimized parameters for compression of human voice at 16kHz sampling rate

This case study is presented on a typical voiced segment of 4000 samples at $16kHz$, a part of which is shown in figure 4.20, taken from a carnatic music sung by a female artist. For period estimation, the autocorrelation method was used. The frame size was selected as 220 samples. The frame-by-frame autocorrelation function for a few frames is given in figure 4.21.

The local period information was extracted as the distance between the successive local peaks of the function. The frames were repeatedly updated to get continuous information about the local periods. The extracted period information is given in the following row vector.

$$\begin{aligned}
 P(k) = & [60 \ 59 \ 59 \ 60 \ 60 \ 59 \ 59 \ 58 \ 58 \ 59 \ 60 \ 58 \ 58 \ 58 \ 58 \ 58 \ 58 \\
 & 59 \ 58 \ 57 \ 57 \ 58 \ 57 \ 58 \ 58 \ 57 \ 57 \ 58 \ 57 \ 57 \ 58 \ 57 \ 57 \ 58 \\
 & 58 \ 59 \ 58 \ 58 \ 59 \ 59 \ 59 \ 58 \ 59 \ 59 \ 58 \ 59 \ 60 \ 57 \ 59 \ 60 \ 57 \\
 & 59 \ 60 \ 57 \ 59 \ 58 \ 58 \ 59 \ 60 \ 58 \ 59 \ 60 \ 58 \ 59 \ 59 \ 58 \ 58 \ 58]
 \end{aligned}
 \tag{4.10}$$

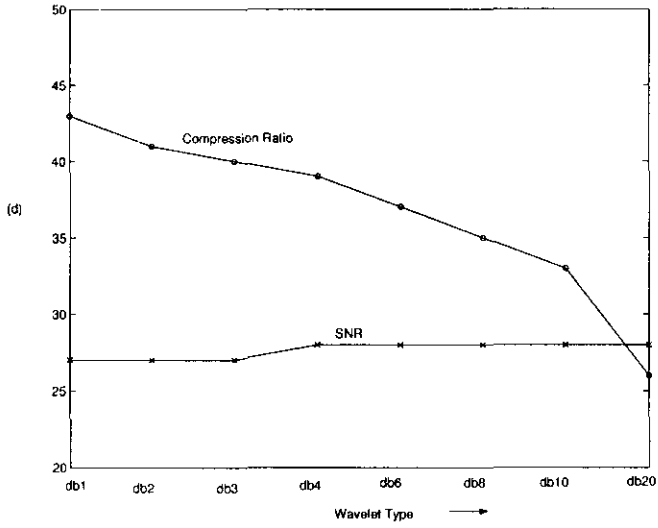


Figure 4.19: Performance of PSWT based LPC on a segment of male voice sampled at 16kHz-variation with Wavelet Type

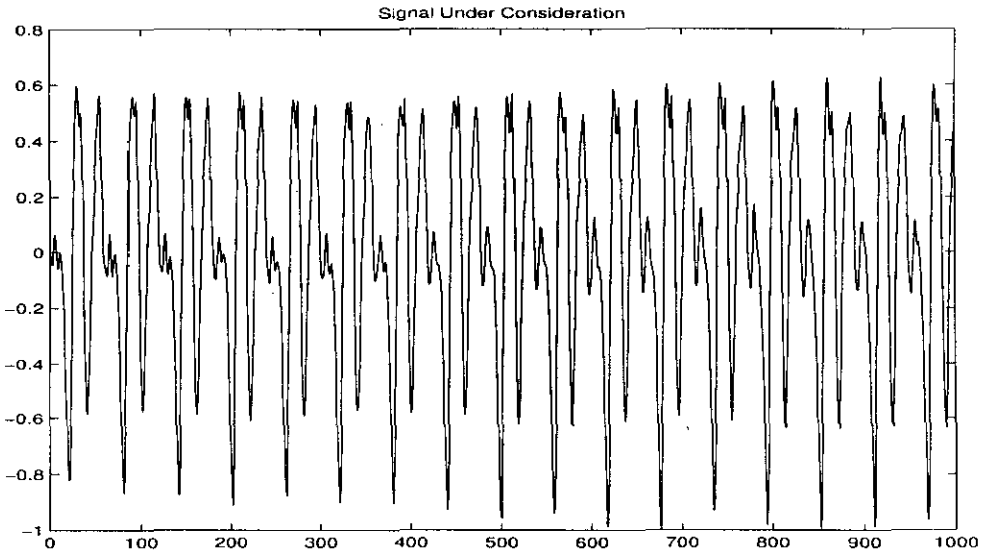


Figure 4.20: Original data segment from a female voice sampled at 16kHz

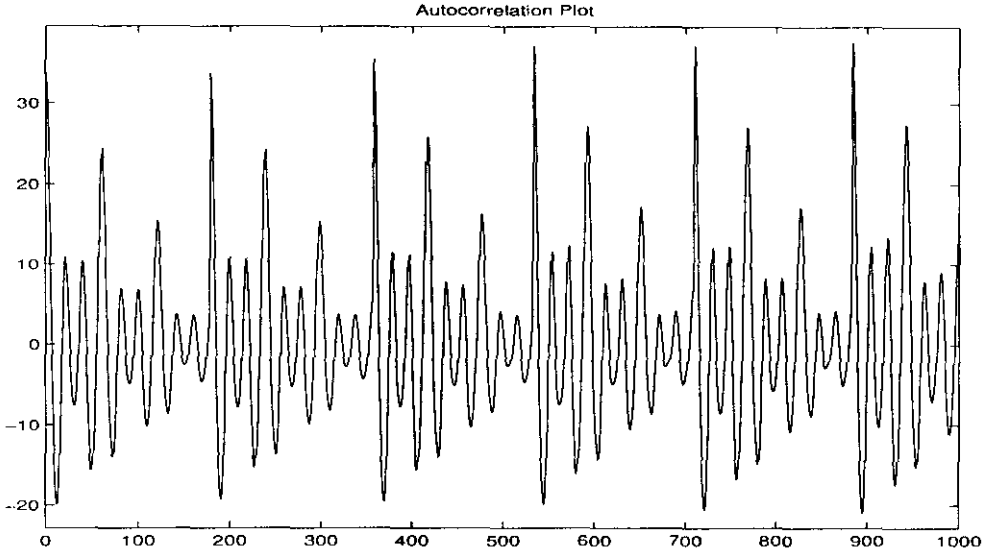


Figure 4.21: Frame-wise autocorrelation plot of the segment shown in fig. 4.20

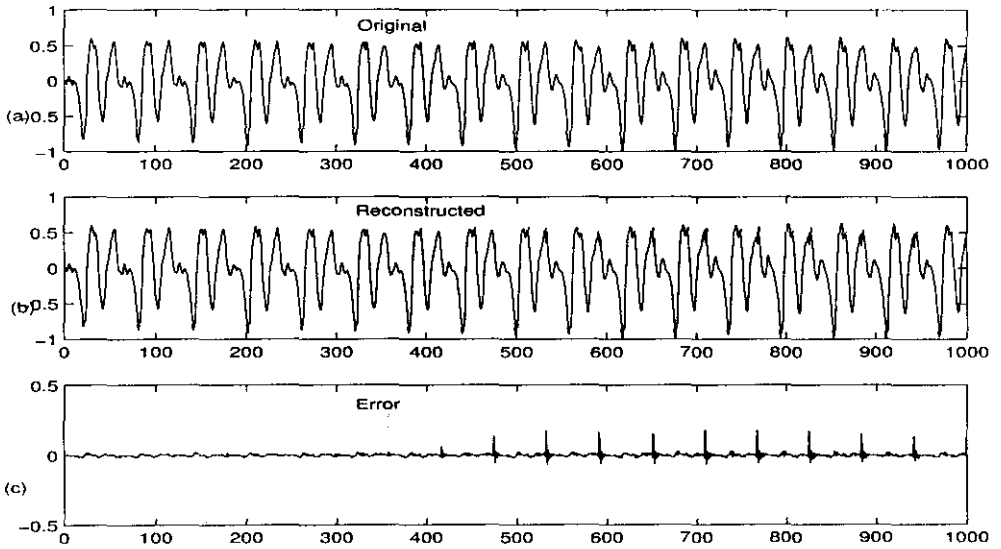


Figure 4.22: PSWT based LPC of a typical voiced data segment taken from a female music.

Type of Signal	Sampling rate & bit size	Performance	
		SNR	CR
Male Voice	8kHz, 16bit	24	24
	16kHz, 16bit	28	47
	22.05kHz, 16bit	34	65
	44.1kHz, 16bit	36	124
Female Voice	8kHz, 16bit	20	11
	16kHz, 16bit	26	23
	22.05kHz, 16bit	29	31
	44.1kHz, 16bit	30	59

Table 4.10: PSWT based LPC performance on voiced segments of human voice at constant pitch sampled at different frequencies.

Using this period data, the signal was arranged in the PS form. The rest of the procedure for this study was exactly the same as that followed for ECG data compression discussed in the previous section. Figures 4.22 shows the reconstructed signal and the error in reconstruction in comparison with the original.

The result obtained in applying the compression scheme on a few signals including male and female voice at a constant pitch period (Pitch corresponding to the D_4 note in the musical scale) sampled at different rates is given in table 4.10. From the table it is obvious that, both the compression ratio and the signal reconstruction quality are better for signals sampled at high rates. Moreover, the compression achieved for male voice is considerably higher than that for female voice at the same pitch frequency.

4.3.3 Case study 3: Instrumental Music

PSWT is quite appropriate for musical signal processing, as musical notes comprise of a number of frequency components, all harmonically related to the fundamental. Every musical instrument has its own characteristic tone qualities often called the *timbre* or *tone colour*. It is this attribute which enables the listener to identify the instrument producing the tone.

A musical signal is non-stationary in nature, and it is an ordered set of isolated

frequencies, $\{f_0, f_1, f_2, f_3, \dots, f_M\}$ where $f_i < f_{i+1}$, $i \in Z$, and $i > 0$ [110]. Here M is the highest perceivable harmonic present in the instrument tone and is different for different musical instruments. The constituent harmonic frequencies are called the instrument partials. In most of the cases, these partials are harmonically related to the fundamental f_0 as in the relationship, $f_i = i * f_0$. The partials in a harmonic set are integer multiples of the fundamental. For a monochromatic signal, the term pitch is identical to the frequency itself. But in the case of musical instruments producing polyphonic sound, the fundamental, rather than the wideband set of partials, determine the pitch.

Basically musical instruments fall into different categories like Reed-type woodwind, Blow-type woodwind, Plucked string, Bow-type string etc. Typical cases from each of the above classes of instruments as listed below have been considered in this study and the results are summarized.

Violin: Bow-type string instrument

Guitar: Plucked string instrument

Harmonium: Reed-type woodwind instrument

Flute Blow-type woodwind instrument

In all the cases, the UDWT based method has been used for the pitch determination. The parameters optimized for the PSWT based coding of the instrument tones are given in table 4.11. All the signals were sampled at 22.05kHz at 16 bit resolution.

Figures 4.23-4.26 illustrate the details of predictive coding and compression on typical violin, guitar, double reed harmonium, and flute tones. Based on the observations on a number of musical instrument tones, the values of SNR and compression achieved for some of the signals with the optimized parameters used for the compression study are tabulated in table 4.12. In all the cases the SNR values are high enough to rate the reconstructed signal quality as excellent.

Parameters	Value			
	Violin	Guitar	Harmonium	Flute
Wavelet	db1	db1	db1	db1
Level of decomposition	1	2	2	1
Order of prediction	2	2	2	2
Period normalizing factor	5	5	5	2
Number of beats for evaluating the predictor coefficients	5	5	2	2

Table 4.11: Optimized parameters for the PSWT based LPC of few musical signals

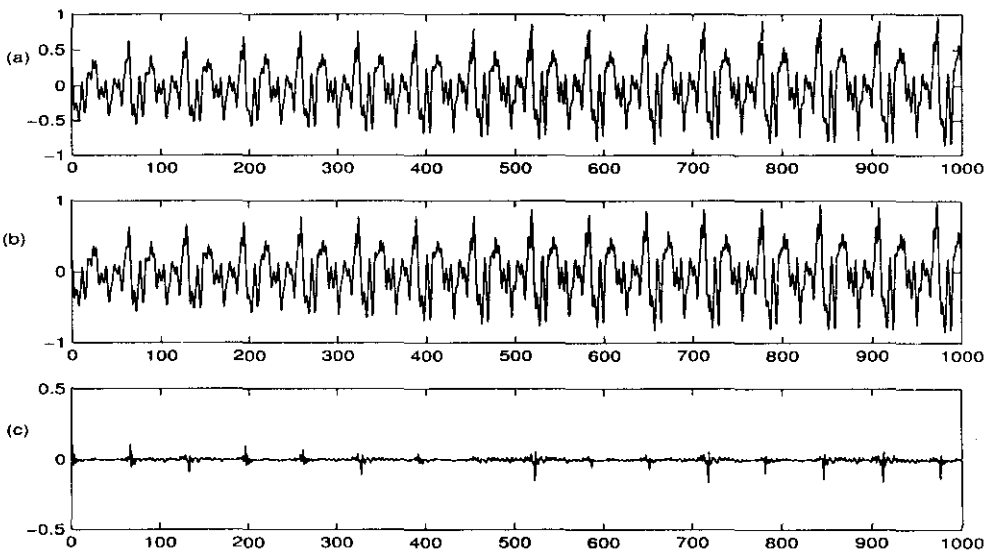


Figure 4.23: PSWT based LPC of a typical segment taken from a Violin tone (a) Original (b) Reconstructed (c) Error

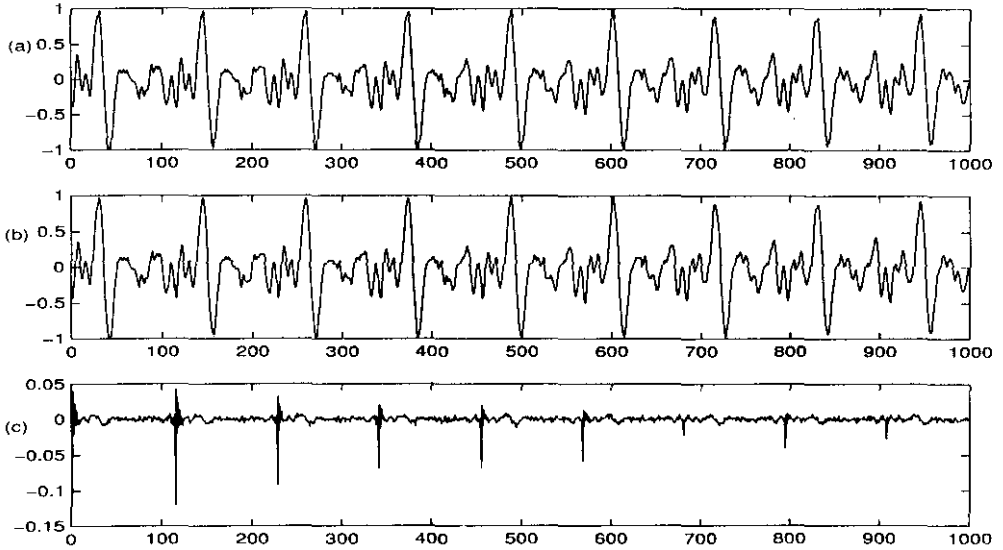


Figure 4.24: PSWT based LPC of a typical segment taken from a Guitar tone (a) Original (b) Reconstructed (c) Error

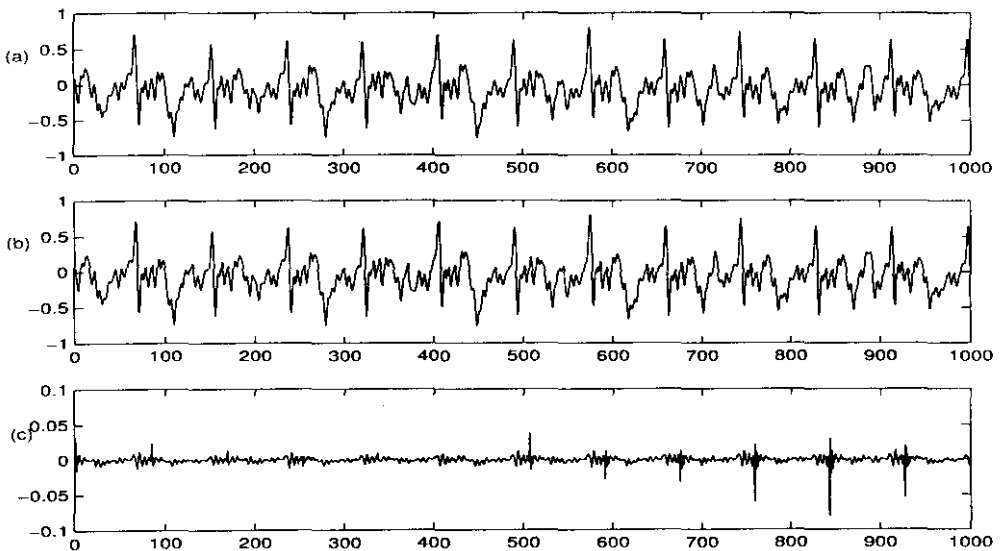


Figure 4.25: PSWT based LPC of a typical segment taken from a Harmonium tone (a) Original (b) Reconstructed (c) Error

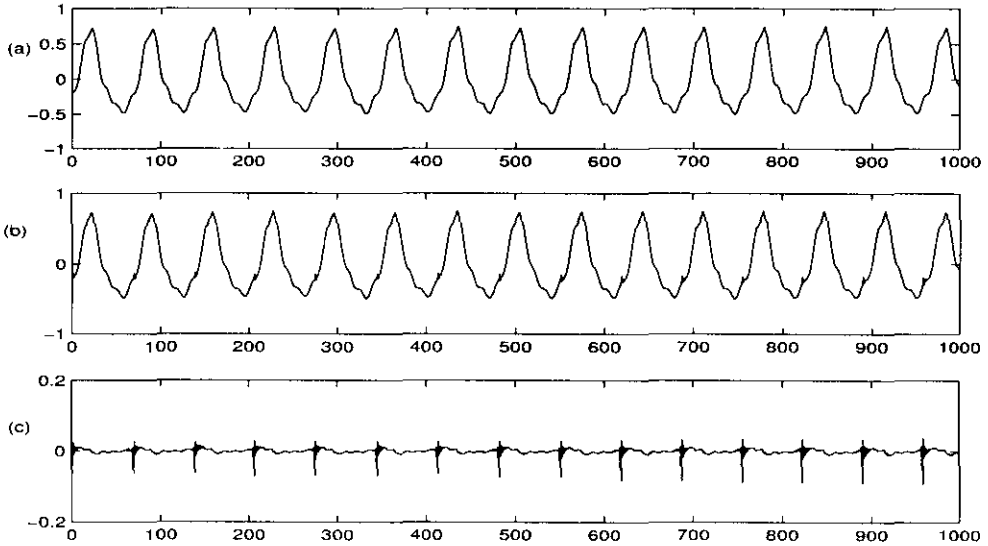


Figure 4.26: PSWT based LPC of a typical segment taken from a Flute tone (a) Original (b) Reconstructed (c) Error

4.4 Feature Enhancement considerations in PSWT domain

The PSWT based LPC technique presented in the previous section could be employed with any pseudo-periodic signal in general. The LPC technique is basically a lossy scheme and the error introduced in coding is inversely proportional to the order of the predictor. The irregularity in the signal, like the noise contamination, demands a higher order predictor to get acceptable reconstruction quality. This noise part has to be removed at an appropriate level, earlier if possible, the better. Since this noise part will get reflected more in the PSWT coefficients, the presence of noise will result in relatively high energy fluctuations at lower levels of PS wavelet decomposition. In view of this fact, some techniques are introduced here, which can be applied in the PSWT domain to enhance the signal quality.

Assuming $w_j(n)$ to be the signal component which gets decomposed into the j^{th} level of details, and $\gamma_I(n)$ be the residue signal getting transformed into the approximation

Instrument	Signal Characteristics	Mean period (samples)	SNR (db)	CR
Violin	22050 Hz, 16 bit	65	23	23
Guitar		114	32	38
Harmonium		83	30	28
Flute		68	28	23

Table 4.12: Summary of compression study on different musical instrument tones

space at a finite level J of the PSWT decomposition, the original signal $x(n)$ can be expressed in terms of the wavelet and scaling partials as:

$$x(n) = \sum_{j=1}^J w_j(n) + \gamma_J(n) \quad (4.11)$$

It may be noted that $w_j(n)$ corresponds to the extend of fluctuations at scale 2^j local pitch periods, whereas $\gamma_J(n)$ represents the asymptotic average behaviour including the melodious harmonics over several pitch periods.

Due to the inter-period similarity, most of the energy contained in the music signal appears in the scaling residue itself. This includes all the spectral components centered on the fundamental and its harmonics, which forms the vital melodious part of the music. On the contrary, the wavelet partials carry other information, which lies near to the harmonics. They also contain other non-melodious, unwanted side bands. Such information gets pushed more into the details spaces, indicating the degree of period-to-period fluctuations present in the signal. Hence the PSWT coefficients at different levels can be suitably thresholded or amplified to get rid of the noise components and enhance the melodious components. Being in the PSWT domain, in addition to enhancing the signal quality, it will improve the efficiency of the predictor as well, as we need to use lower order predictors only, for the same degree of reconstruction quality. Even though this method is sometimes not effective for low frequency noise components like power line noise etc., it is quite suited for instrument-dependent noise removal.

4.4.1 Method of Noise suppression and Feature enhancement

Musical instrument tones are often prone to noise contamination and peculiar characterization. This include the striking noise in the case of string instruments, deterioration of tonal quality due to immature playing, personal jesters used while playing, mechanical sound of any moving parts, etc. These are mainly instrument dependent and in many cases they contribute to unpleasant effects in the musical tones. Mutilated signals from four different instruments are considered in this study, where the proposed method is applied to separate the disturbance part from the melodious instrument tones. The signals were of 16 bit resolution, sampled at 22050Hz, the details of which are given in table 4.13.

Sl. no.	Instrument under study	Disturbance present
1	Violin: Bow type String	Unpleasant noise due to bow movement
2	Acoustic Guitar: Plucked string	Spurious oscillations due to eccentric striking
3	Harmonium: Reed type woodwind	Blow noise produced by the bellows
4	Flute: Blow type woodwind	Shrill oscillations due to change of blowing air pressure

Table 4.13: Details of mutilated music signals considered for illustration.

The results shown are that of a study based on a 3 – level PS wavelet decomposition using the *db2* wavelet. After the decomposition, the PSWT coefficients are thresholded by a factor dependent on the energy content in each of the wavelet subspaces. As the noise components gets transformed mostly into these subspaces, this operation considerably suppresses the noise part in the signal. Since the melodious portions are mainly sieved into the approximation subspace, the scaling coefficients are amplified by a factor

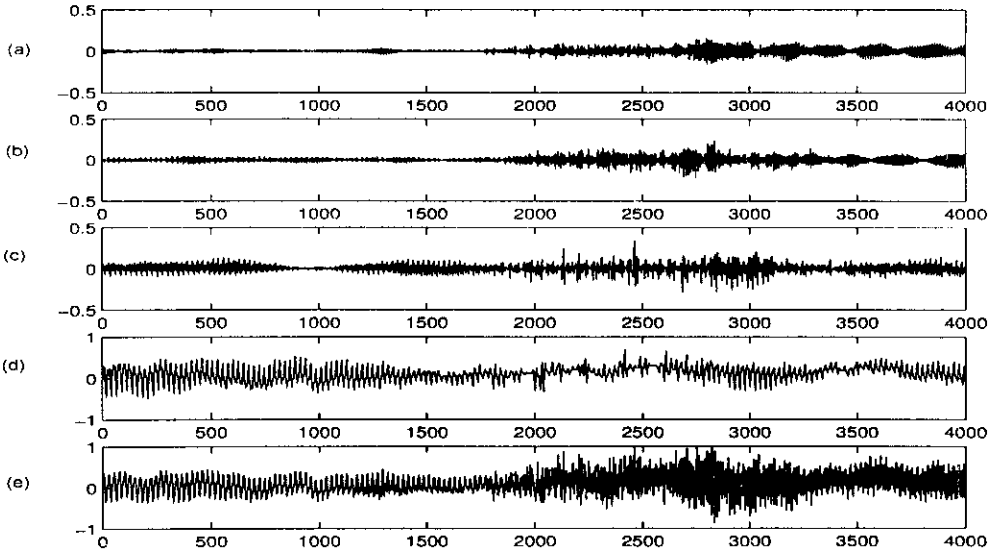


Figure 4.27: 3-Level PSWT of a Violin signal mutilated by Bow noise (a)-(c): Wavelet partials at level 1-3 respectively. (d): Scaling partial at level 3 (e): Original signal.

dependent on the total energy of the signal. This method is found to enhance the music quality.

Figures 4.27- 4.30 demonstrate the details of this study. In all the above cases, it is observed that the non-melodious instrument-dependent disturbances are mostly sieved out into the wavelet subspaces during the first and second levels of decomposition. The PSWT scaling partials at the third level is found to contain the melodious components present in the signals without appreciable loss of energy. This is clear from the figures 4.27*d-e*, 4.28*d-e*, 4.29*d-e* and 4.30*d-e*.

The trailing end of the violin signal is contaminated by heavy bow-noise and hence the wavelet partials corresponding to this region contain more energy which is unexpected in melodious signals. In the case of the guitar signal, the wavelet partials contain significant energy (fig 4.28*a-c*). It is attributed to the enormous distortion present in the signal. Similarly as the agitation in the flute signal is sustained throughout the note, the lower level wavelet partials contain a near-to-uniform distribution of the non-

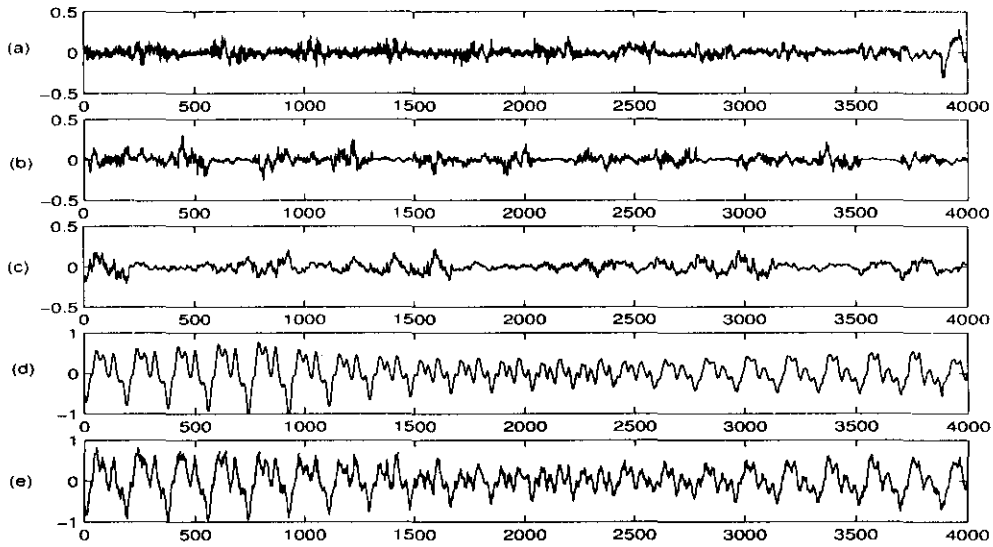


Figure 4.28: 3-level PSWT of a Guitar signal mutilated by eccentric striking noise (a)-(c) : Wavelet partials at levels 1-3 respectively. (d): Scaling partial at level 3 (e): Original signal.

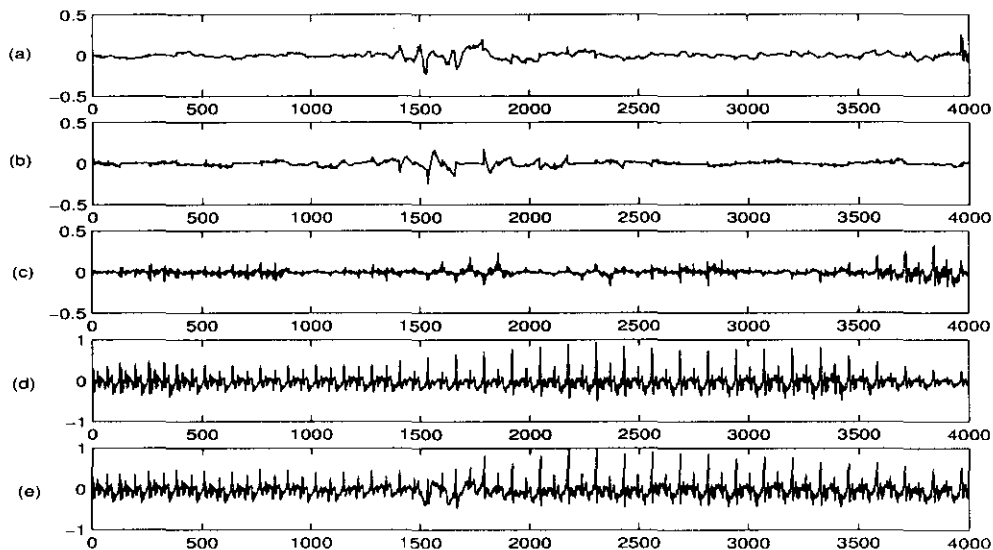


Figure 4.29: 3-level PSWT of a Harmonium signal mutilated by bellows blow noise (a)-(c): Wavelet partials at levels 1-3 respectively. (d): Scaling partial at level 3 (e): Original signal.

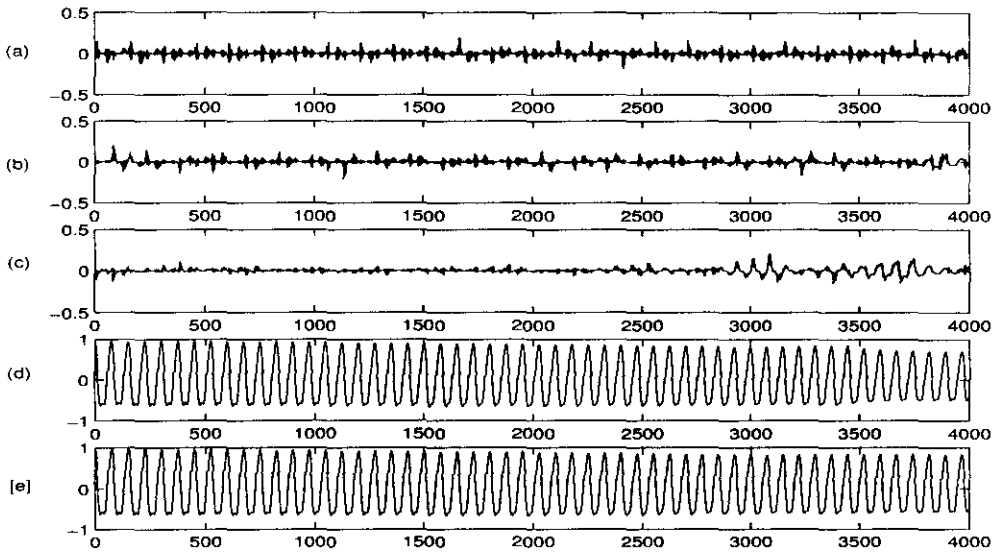


Figure 4.30: 3-Level PSWT of a noisy Flute signal (a)-(c): Wavelet partials at level 1-3 respectively. (d): Scaling partial at level 3 (e): Original signal.

harmonic signal components. The distortion in the harmonium signal is well localized and is almost transformed in the highest frequency wavelet partial.

4.4.2 Effect of Feature Enhancement in PSWT based Predictive Coding

The effectiveness of the PSWT based coding scheme, to a large extent, depends on the quality of the signal and the level of periodicity present in the signal. Almost all PSWT coefficients of a polluted signal are found significant for a true reproduction of the signal at the reconstruction end, whereas for a pure signal, quite a good number of them are insignificant or can be coded with less number of bits, enhancing the compression efficiency. As a result, the application of PSWT based feature enhancement techniques before the compression, improves the compression as well.

4.5 Conclusions

A new method for the analysis and compression of general pseudo-periodic signals has been introduced. In addition to making use of the periodicity property of the signals this method uniquely combines the advantages of LPC technique with the popular WT methods. The scheme includes various tasks like identification of pseudo-periodic regions in signals, continuous estimation of the local periods, sampling rate conversion, PSWT computation, predictive coding of the PSWT coefficients, and the corresponding inverse operations. It has been observed that predictive coding of PSWT coefficients has greater potential in signal compression compared to the ordinary DWT based methods. The proposed method has been validated with different classes of practical signals like ECG, human voice and musical instrument tones. The various parameters for the compression such as the type of wavelet, order of prediction, level of decomposition, etc. were optimized for each of these classes. The UDWT based methods are found superior in period estimation of signals like ECG and musical instrument tones, compared to the autocorrelation based methods. The period normalization is found to have improved the beat-to-beat correlation, contributing to better compression. The effectiveness of the method is largely dependent on the accuracy with which the local periods are estimated. The noise contamination in the signals is found to reduce the efficiency of compression. The PSWT-based feature enhancement techniques proposed in this chapter has resulted in better compression and reduction in bit size for coding the coefficients.

Chapter 5

The PMS Computational Structure

In the present chapter the PMS computational structure developed for DWT computation is presented. A brief introduction to the computational issues in WT implementation is given followed by the development of the PMS structure. The computational efficiency of this structure is compared against the popular pyramid structure. The case of PSWT based ECG data compression is taken as a case study to highlight the efficiency of the new scheme. Subsequently a description on the 2D extension of the PMS structure and a case study on edge detection is given. The chapter is concluded with the important results and observations regarding DWT computation employing the PMS structure.

5.1 Introduction

Signal decomposition studies using WT is an active research topic which finds extensive applications in science, engineering and technology. Development of new algorithms and the issues related to the computation of different types of WT are among the areas that gained primary importance. The competency of computational algorithms depends on various factors such as the number of multiplications, number of additions, number of data interchange operations, data dependency, computational delay, parallelism in computation, simplicity in algorithm, possibility of direct computation of intermediate results etc. As a result, all these became topics of interest to the research community. Several efficient algorithms have been derived for the implementation of wavelet transform. Many researchers have made comparison of these algorithms based on computational complexity as mentioned in chapter 2.

In the present work, the computational complexity of the pyramid structure¹ has been analyzed and the PMS structure is proposed as an efficient alternate scheme for real-time computation of DWT and IDWT. The underlying idea is to reduce the number of computations, incorporate parallelism computing coefficients at different levels of computation maintaining a regular computational structure, enable direct computation of transform coefficients at any preferred level and perform block wise computations after segmenting the signal based on an optimal selection of the frame size for any specific application. The PMS² structure is shown to be advantageous for PSWT computation as well as for various 1D and 2D signal processing applications.

5.1.1 Issues in Wavelet Transform Computation

In most of the algorithms developed so far, the DWT computation is performed based on the filter bank tree structure. The main disadvantage of using this serial structure

¹Interchangeably used as filter bank structure and tree structure

²parallel multiple subsequence

lies in the fact that, determination of the wavelet coefficients at the j^{th} level requires $(j - 1)$ times of successive low pass filtering followed by one high pass filtering. The delay associated with the implementation grows exponentially with the number of levels [182]. Further, for applications where wavelet coefficients at chosen levels alone are required, one has to perform unnecessary computations by way of evaluating the lower level coefficients.

In real-time signal processing applications, the WT computation need to be performed on a frame-by-frame basis. In such cases, employing the above algorithms *per se* results in undesired artifacts at the ends of each frame. To resolve this difficulty, many researchers have come out with BDWT³ algorithms [187], [188]. These algorithms are found to have difficulties related to large storage buffer requirement and frequent inter-processor communication.

In many applications like the functional neuro-imaging, feature enhancement and noise removal, period estimation, computer-assisted mammography etc., there are situations in which the information contained in the WT coefficients at a few selected levels only are essential to deduce the results. Consequently, in such applications, the WT coefficients at these desired levels only need be computed [189], [221]. Here the conventional Mallat algorithm proves to be an inappropriate choice as it is a sequential implementation, which essentially goes through all intermediate levels, whether it is ultimately required or not. Hence a parallel computational structure would be an ideal choice for this type of applications. The efficiency of parallel algorithms highly depends on the extent of inter-processor communications that the process demands and the intermediate storage requirements. In massive computational environment, the parallel computation of the WT coefficients with minimum inter-processor communication is much appreciated.

³Block Discrete Wavelet Transform

5.2 Development of the PMS Structure for 1D Signals

The DWT of any signal $x(n)$ may be expressed [11] as the ensemble of projections $X_{i,k}$ and $\beta_{J,k}$ of $x(n)$ over the wavelet sequences $\psi_{i,k}(n)$ and the scaling sequences $\phi_{J,k}(n)$ respectively, where the first index $i = 1, 2, \dots, J$ represents scale and the other index $k = 0, 1, 2, \dots$ is associated to time shift. Using mathematical notation,

$$X_{i,k} = \sum_n x(n)\psi_{i,k}(n) \quad (5.1a)$$

and

$$\beta_{J,k} = \sum_n x(n)\phi_{J,k}(n) \quad (5.1b)$$

where $\psi_{i,k}(n) = \psi_{i,0}(n - 2^i k)$ and $\phi_{J,k}(n) = \phi_{J,0}(n - 2^J k)$. Corresponding equation for IDWT computation will be:

$$x(n) = \sum_{i=1}^J \sum_k X_{i,k}\psi_{i,k}(n) + \sum_k \beta_{J,k}\phi_{J,k}(n) \quad (5.2)$$

Here the wavelets and scaling function at each level are iteratively updated using the relations [4]:

$$\phi_{1,0}(n) = g(n); \quad \psi_{1,0}(n) = h(n) \quad (5.3a)$$

$$\phi_{i+1,0}(n) = \sum_q \phi_{i,0}(q)g(n - 2q) \quad (5.3b)$$

and

$$\psi_{i+1,0}(n) = \sum_q \psi_{i,0}(q)g(n - 2q) \quad (5.3c)$$

where h and g are the impulse responses of the high pass and low pass filters respectively of the QMF⁴ bank corresponding to the wavelet. The wavelets in equation 5.3 can easily be shown to be a set of band pass filters with impulse response h_i for $i \neq 1$, and a high pass filter of response h_1 for $i = 1$, where h_i stands for the analysis discrete wavelets and

⁴quadrature mirror filter

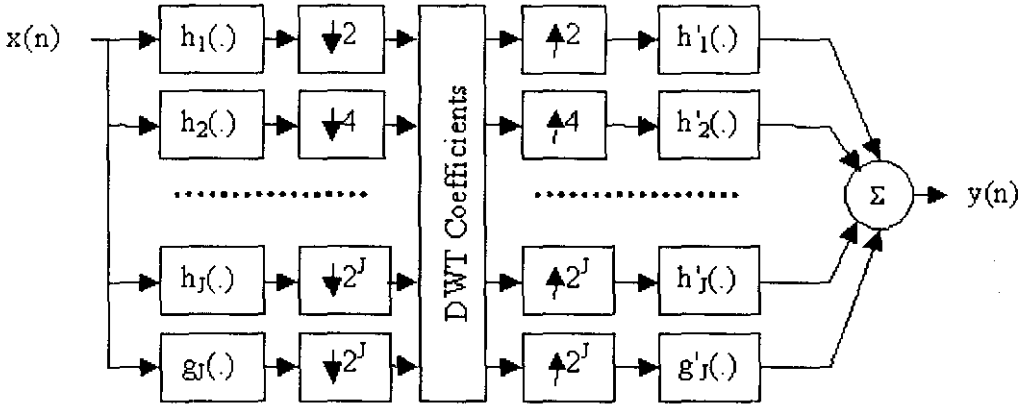


Figure 5.1: Simple structure for the parallel computation of DWT & IDWT.

g_j stands for the corresponding scaling sequence. The associated reconstruction filters h'_i and g'_j can be obtained by putting the associated reconstruction filters \tilde{h} and \tilde{g} in place of h and g in the above equation. The filters h, g, \tilde{h} and \tilde{g} are related one another which depends on the type of wavelet system selected, as derived in Burrus *et al* [81].

It can be easily understood from equations 5.1, 5.2 & 5.3 that, the process of decomposition and reconstruction are multi-band filtering operations as shown in figure 5.1, in agreement to the usual filter bank structure. In contrast, as all levels of decomposition and reconstruction can be performed independently, this is more suitable in parallel processing environment. From the figure it may be observed that a number of computations performed during each of the convolution operation is irrelevant as far as the final output of the stage is concerned.

5.2.1 Eliminating Irrelevant Computations

The $X_{i,k}$ in equation 5.1 can be viewed as a convolution operation between the input signal $x(n)$ and the wavelet filter $h_i(-n)$ followed by down sampling by a factor of 2^i .

It can be mathematically expressed as

$$X_{i,k} = (x(n) * h_i(-n)) \downarrow 2^i, \quad k = 2^{-i}n; n, k \in \mathbf{Z} \quad (5.4a)$$

Similarly,

$$\beta_{J,k} = (x(n) * g_J(-n)) \downarrow 2^J \quad (5.4b)$$

Here the ‘ \downarrow ’ represents the downsampling operation. The synthesis equation 5.2 accordingly gets modified as:

$$x(n) = \sum_{i=1}^J X_{i,n/2^i} * \tilde{h}_i(n) + \beta_{J,n/2^J} * \tilde{g}_J(n) \quad (5.5)$$

Here it is clear that, for each transform coefficient being added at the i^{th} level during decomposition, the downsampling operation following the convolution simply removes $2^i - 1$ samples from the output of convolution immediately after its computation, without being used anywhere else. Similarly, during reconstruction, the upsampling operation prior to the convolution increases the length of the sequence being convolved by $2^i - 1$ samples per final output sample without adding any new information. For an efficient implementation of the computation scheme, these unsolicited computations are to be eliminated. If the subsampling is performed before the convolution operation, the data length at the i^{th} level convolution is reduced by a factor of 2^i . Similarly, during reconstruction if the convolution is performed before the upsampling operation, considerable reduction is achieved in the length of the sequence being convolved. These modifications as done in the case of *lifting scheme* [178], along with the principle of *polyphase splitting* [217] could be made to advantage in the elimination of undesired and redundant computations.

5.2.2 The Proposed Structure

Consider the computation of the scaling and wavelet transform coefficients using the expression 5.4. The sequences involved in the convolution are $x(n)$, $h_i(n)$, and $g_J(n)$.

Taking the first level, to get the transform coefficients, the convolved result is downsampled by a factor of 2. Using the polyphase splitting, the same result can be achieved in a different way. The original sequences $x(n)$, $h_1(n)$ and $g_1(n)$ are first split into the even indexed samples

$$x_{1e} = x(2k), \quad h_{1e} = h_1(2k), \quad g_{1e} = g_1(2k)$$

and the odd indexed samples

$$x_{1o} = x(2k+1), \quad h_{1o} = h_1(2k+1), \quad g_{1o} = g_1(2k+1); \quad k \in \mathbf{Z}.$$

These sequences can be separately convolved as per the following formula to get the transform coefficients at the first level.

$$X_{1,k} = x_{1o}(k) * h_{1o}(-k) + x_{1e}(k) * h_{1e}(-k) \quad (5.6a)$$

$$\beta_{1,k} = x_{1o}(k) * g_{1o}(-k) + x_{1e}(k) * g_{1e}(-k) \quad (5.6b)$$

In the case of a multi-level decomposition, the above procedure can be extended to all the parallel branches in figure 5.1. Following a similar argument, each of the even and odd subsequences in equations 5.6a and 5.6b. is again split into their even and odd samples, resulting in 2^i subsequences at the i^{th} level. These multiple subsequences can be mathematically expressed as⁵.

$$x_{ip}(k) = x(2^i k + p - 1) \quad (5.7a)$$

$$h_{ip}(k) = h_i(2^i k + p - 1) \quad (5.7b)$$

$$g_{Jp}(k) = g_J(2^J k + p - 1) \quad (5.7c)$$

where $p = 1, 2, \dots, 2^i$. The signal subsequences are convolved with the relevant wavelet or scaling subsequences as the case may be, and the partial results are then added together to get the transform coefficients at the corresponding level. In view of this,

⁵Here it is assumed that the sequences are multiples of 2^i , else they may be zero padded to this length

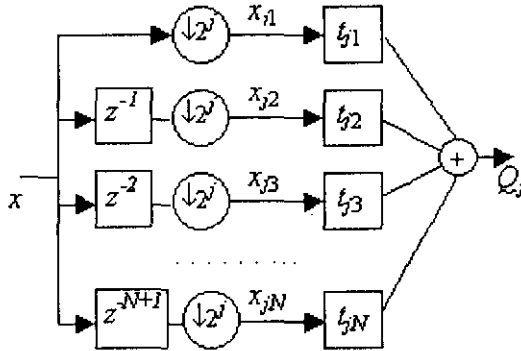


Figure 5.2: A PMS Computational block for decomposition ($N = 2^j$)

expressions 5.4 can be written as

$$X_{i,k} = \sum_{p=1}^{2^i} x_{ip}(k) * h_{ip}(-k) \quad (5.8a)$$

$$\beta_{j,k} = \sum_{p=1}^{2^j} x_{jp}(k) * g_{jp}(-k) \quad (5.8b)$$

These computations are all independent and hence can be performed in parallel environment without any inter-processor communication except for the final addition of the partial results. Figure 5.2 shows the configuration of one such block of the Parallel Multiple Subsequence computational structure applicable for the j^{th} level decomposition. Here t stands for the filters h or g as the case may be. Q_j represents the transform coefficients X_j or β_j .

A similar argument can be followed in the case of reconstruction to eliminate the unwanted computations. The upsampling operations shown in figure 5.1 could be performed after the convolution without affecting the end result. This also is achieved by a level-dependent subsequence decomposition of the wavelet and scaling sequences at each level as described below.

Considering the first level wavelet partial reconstruction, the filter sequence is divided into the even sampled sequence \tilde{h}_{1e} and the odd sampled sequence \tilde{h}_{1o} . Both of these

subsequences are convolved with the WT coefficients $X_1(k)$. The partial results are upsampled by 2 and added with one unit delay for the sequence $X_1(k) * \tilde{h}_{1e}(k)$. The first level wavelet partial be now expressed as

$$y_1 = (X_1(k) * \tilde{h}_{1o}(k)) \uparrow 2 + [(X_1(k) * \tilde{h}_{1e}(k)) \uparrow 2]z^{-1} \quad (5.9)$$

This procedure is extended to all the partial sums by incorporating a level dependent division of the wavelet/scaling coefficients as

$$\tilde{h}_{ip}(k) = \tilde{h}_i(2^i k + p - 1) \quad (5.10a)$$

$$\tilde{g}_{Jp}(k) = \tilde{g}_J(2^J k + p - 1) \quad (5.10b)$$

These subsequences are convolved with the transform coefficients at the appropriate level. These convolved partial results at each level i are upsampled by 2^i and then added resulting in the i^{th} level wavelet or scaling partial given by

$$y_i(n) = \sum_{p=1}^{2^i} y_{ip}(n - p + 1) \quad (5.11a)$$

$$\gamma_J(n) = \sum_{p=1}^{2^J} \gamma_{Jp}(n - p + 1) \quad (5.11b)$$

where,

$$y_{ip}(n) = (X_i(k) * \tilde{h}_{ip}(k)) \uparrow 2^i \quad (5.12a)$$

and

$$\gamma_{Jp}(n) = (\beta_J(k) * \tilde{g}_{Jp}(k)) \uparrow 2^J \quad (5.12b)$$

Here the \uparrow denotes upsampling operation. The PMS computational structure performing the j^{th} level reconstruction is shown in figure 5.3. Here q_j stands for the wavelet partial y_j or the scaling partial γ_j as the case may be.

The reconstructed signal can then be expressed as the sum of all the wavelet partials and the scaling partial corresponding to the J^{th} level as expressed by the final reconstruction equation,

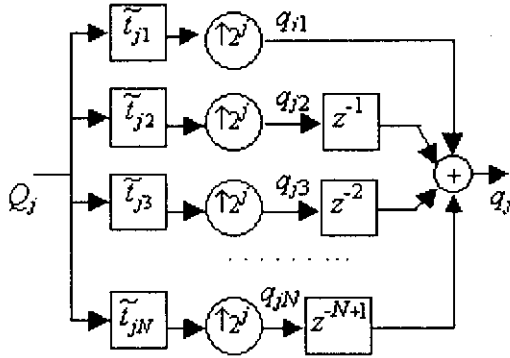


Figure 5.3: PMS Computational block for j^{th} level reconstruction.

$$y(n) = \sum_{i=1}^J y_i(n) + \gamma_J(n) \tag{5.13}$$

The remodelled computational structure for a two level DWT-IDWT implementation as per the new algorithm is shown in Figure 5.4. This being a regular structure, can be easily extended to any level. As evident from the figure, this structure has parallelism both within and between levels, making it suitable in parallel processing environment.

5.2.3 Analysis of Computational Complexity

In this section, the computational complexity of both the pyramid structure and PMS structure are estimated in terms of the real arithmetic operations involved in the computation of the transform coefficients. Expressions are derived for the number of multiplications and additions required to perform the decomposition and reconstruction.

5.2.3.1 Estimation of Computational burden for pyramidal structure

In of the DWT computation using the pyramidal structure, each level require two convolution operations followed by downsampling. The convolution between two sequences of length l_1 and l_2 requires $l_1 l_2$ multiplications and $(l_1 - 1)(l_2 - 1)$ additions. The resulting sequence will be of length $l_1 + l_2 - 1$. In using the pyramid structure, though

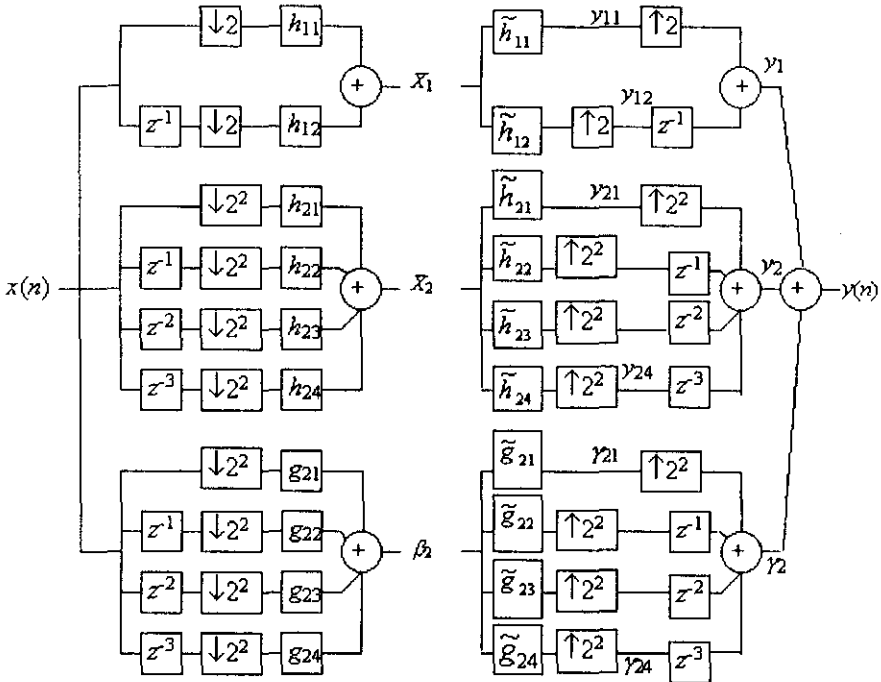


Figure 5.4: The Complete PMS structure for 2-level DWT and IDWT Computation.

the filter length remains the same for all levels, the input sequence length gets successively modified due to the *increase* in length resulting from convolution operation at the previous level, and the *decrease* in length resulting from downsampling operation. Assuming l_x is the length of the input sequence at level $i = 0$ and l_w , the length of the wavelet filter, input sequence to the i^{th} level will be of length⁶

$$[l_x + (2^{i-1} - 1)(l_w - 1)]/2^{i-1}.$$

For a J -level decomposition, the total computational burden is

$$\sum_{i=1}^J 2[l_x + (2^{i-1} - 1)(l_w - 1)]/2^{i-1} l_w \text{ multiplications}$$

and $\sum_{i=1}^J 2[(l_x + (2^{i-1} - 1)(l_w - 1)]/2^{i-1} - 1](l_w - 1)$ additions.

This is simplified to

⁶Here the length of wavelet and scaling filters are assumed to be the same. The fact that the length of the resulting sequence thus evaluated may not be an integer, is not of much concern at this point as it will not affect the discussions to follow

$2l_w\{J(l_w - 1) + 2(1 - 2^{-J})(l_x - l_w + 1)\}$ multiplications

and $2(l_w - 1)\{J(l_w - 2) + 2(1 - 2^{-J})(l_x - l_w + 1)\}$ additions respectively.

It is shown [4] that for IDWT, the number of arithmetic operations required is exactly the same as that of DWT.

5.2.3.2 Estimation of Computational burden for PMS structure

In the PMS structure, the length of both the input sequence and the filter sequences are varying depending on the level of computation. At level i , the input sequence and wavelet/scaling sequences are split into 2^i subsequences, each of length $l_x/2^i$ and $[(l_w - 1)(2^i - 1) + 1]/2^i$ respectively. For a J -level decomposition the total number of multiplications required is

$$\sum_{i=1}^J \{l_x[(l_w - 1)(2^i - 1) + 1]/2^i\} + l_x[(l_w - 1)(2^J - 1) + 1]/2^J.$$

This may be simplified to $l_x[1 + J(l_w - 1)]$.

The number of additions is

$$\begin{aligned} & \sum_{i=1}^J \{2^i(\frac{l_x}{2^i} - 1)[\frac{(l_w - 1)(2^i - 1) + 1}{2^i} - 1] + [\frac{l_x}{2^i} + \frac{(l_w - 1)(2^i - 1) + 1}{2^i} - 1](2^i - 1)\} \\ & + \{2^J(\frac{l_x}{2^J} - 1)[\frac{(l_w - 1)(2^J - 1) + 1}{2^J} - 1] + [\frac{l_x}{2^J} + \frac{(l_w - 1)(2^J - 1) + 1}{2^J} - 1](2^J - 1)\} \end{aligned}$$

This gets simplified to $[(l_w - 1)(l_x - 1) + 1]J$.

During the reconstruction phase, the wavelet/scaling filter coefficients are reorganized into subsequences as in the case of decomposition and then convolved with wavelet/scaling transform coefficients before upsampling. The resulting sequence is required to be finally truncated to a length of l_x . The operations that are needed for computing the terms being thus removed, are eliminated at each level during the convolution itself by modifying the convolution operation for $y_{ip}(n)$ and $\gamma_{Jp}(n)$ given in equation 5.12 as:

$$y_{ip}(n) = \left\{ \sum_{r=\frac{l_{w_i}}{2^i}-1}^{\frac{l_x}{2^i}+\frac{l_{w_i}}{2^i}-1} X_i(r)\tilde{h}_{ip}(k-r) \right\} \uparrow 2^i \quad (5.14a)$$

and

$$\gamma_{Jp}(n) = \left\{ \sum_{r=\frac{l_{w_J}}{2^J}-1}^{\frac{l_x}{2^J}+\frac{l_{w_J}}{2^J}-1} \beta_J(r)\tilde{g}_{Jp}(k-r) \right\} \uparrow 2^J \quad (5.14b)$$

where $l_{w_i} = (l_w - 1)(2^i - 1) + 1$. This implies that the number of multiplications required for reconstruction is the same as that of decomposition.

Finally, the number of additions needed for a J level IDWT computation can be expressed as

$$\sum_{i=1}^J \left\{ l_x \left[\frac{(l_w-1)(2^i-1)+1}{2^i} - 1 \right] + l_x \left[\frac{(l_w-1)(2^J-1)+1}{2^J} - 1 \right] + l_x J \right.$$

which gets simplified to $l_x J (l_w - 1)$. The operations required for adding the 2^i subsequences in each of the i^{th} level are ignored, as they are all null additions.

5.2.4 Comparison of PMS Structure with Pyramidal Structure

In order to visualize the computational advantage of the PMS algorithm over the pyramidal scheme, a comparison of the computational burden is made for both decomposition and reconstruction based on the number of real multiplications and additions.

Figure 5.5 shows the performance in two typical cases. The arithmetic operations for different data lengths (l_x) and wavelet lengths (l_w) are plotted as a function of the level of decomposition. Fig. 5.5(a)-(b) show that the PMS structure is best suited for DWT computation when the number of levels is less than 4. It is independent of the signal length and applicable for wavelets whose support is ≥ 4 . A good majority of practical applications of time-scale analysis need only few levels of wavelet decomposition [217]. For such applications the proposed computational structure is seen to be advantageous over existing schemes. Fig. 5.5(c)-(d) show the comparison for smaller wavelets like *Haar*. In such cases the number of multiplications required for the PMS implementation is

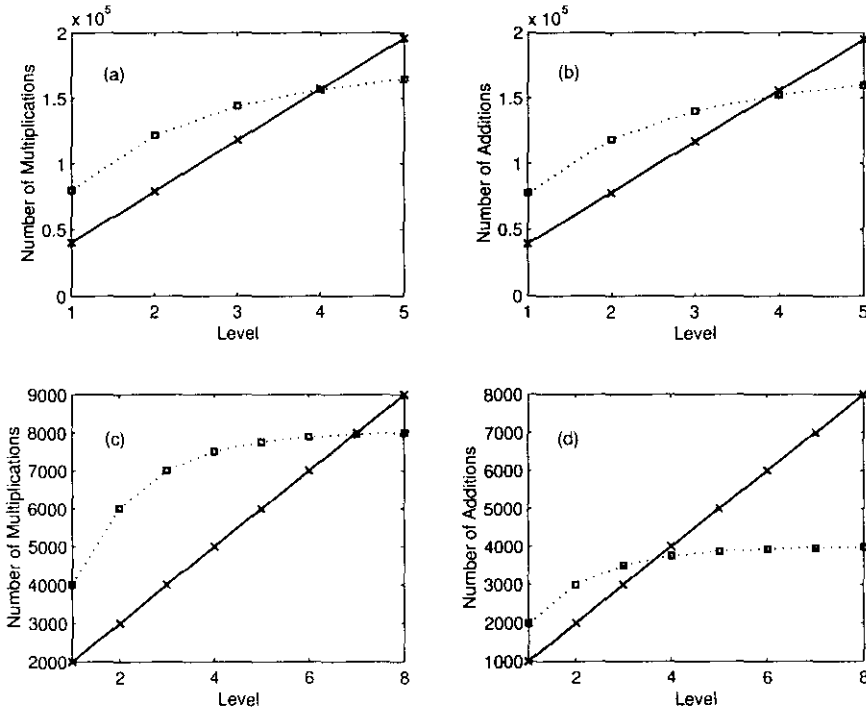


Figure 5.5: Comparison of arithmetic operations between the proposed PMS Structure (—) and the Filter Bank Structure (⋯), during decomposition (a)-(b) for $l_w = 40, l_x = 1000$ (Applicable for wavelets of support ≥ 4). (c)-(d) for $l_w = 2, l_x = 1000$. (Applicable for wavelets of support < 4)

considerably less than that for the corresponding filter bank method even up to 6 levels. The Haar wavelet has been identified as the best wavelet for a number of applications like speech processing, cosmological studies etc. [9], [156], [222].

Figures 5.6 & 5.7 are drawn to show the relative performance of the two algorithms on simultaneous variation of the signal length, wavelet length and level of decomposition. For each level of decomposition, the $l_x - l_w$ plane (fig. 5.6) can be split into two regions - one in which the PMS method performs better and the other in which the Filter Bank method is better, in terms of the total number of multiplications. For example the region to the left of the curve (marked $J=4$) indicates the area where the PMS method

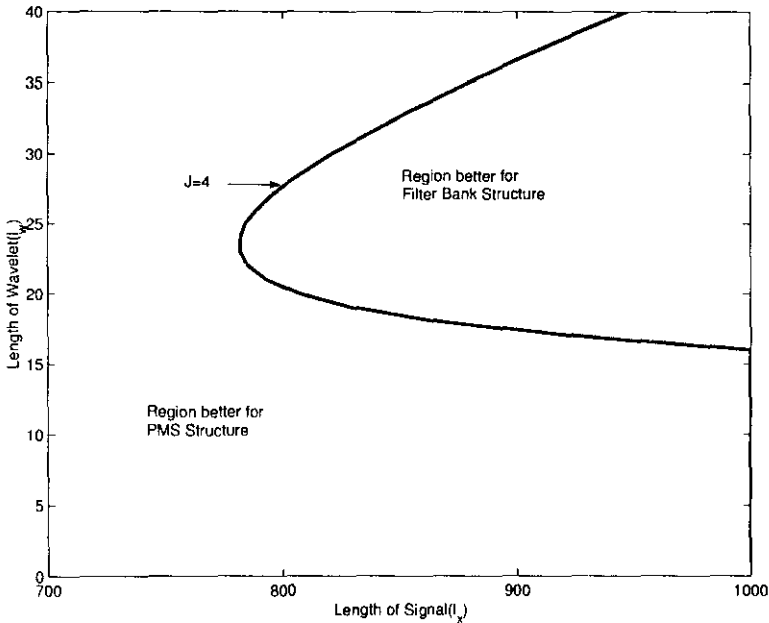
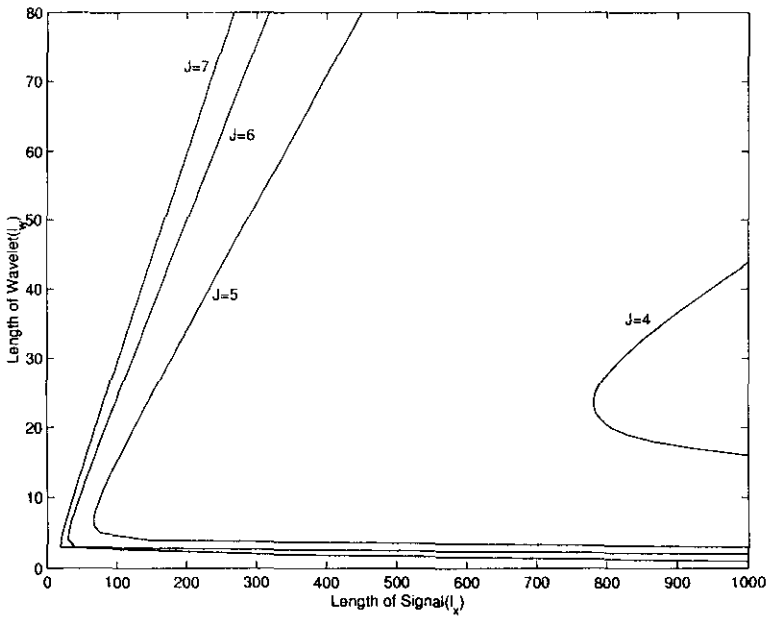


Figure 5.6: Relative performance of the PMS and Filter Bank Structures for variable signal and wavelet lengths for a 4-level decomposition.

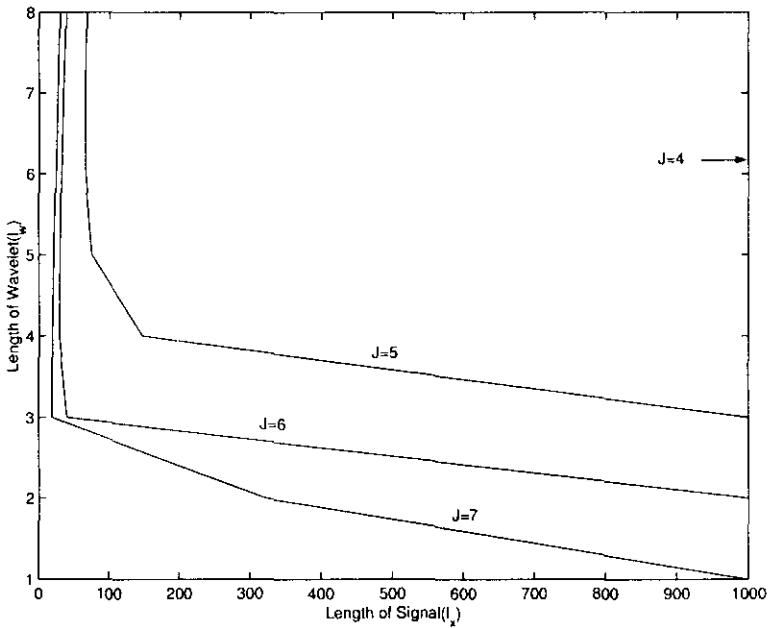
outperforms the filter bank scheme for different values of signal and wavelet lengths corresponding to a 4-level decomposition.

The curves in fig. 5.7(a) show the boundary between the two regions for different values of J . The portion of the plot related to wavelets of smaller support is redrawn in fig. 5.7(b). It is observed that the PMS structure is more advantageous for signal processing applications where smaller data sequences are involved. One typical example would be the PSWT based signal processing. The PSWT computation of a long sequence can be interpreted as a collection of DWT computations on smaller sequences. Hence the PMS structure helps in faster computation of PSWT.

Another remarkable observation that could be made from fig. 5.7 is that, once the wavelet system and the level of analysis are finalized, the input frame length can be suitably selected, so that the computations are reduced considerably. This is typically



(a) for different levels of decomposition



(b) enlarged view for wavelets of shorter support

Figure 5.7: Relative performance of the PMS and Filter Bank Structures for variable signal and wavelet lengths.

the requirement for real-time signal processing applications, where the input data is always taken in small frames, whose size depends on the extend of computations involved and also on the nature of application under consideration [223], [224]. In such cases, the DWT computation has to be performed blockwise [198]. The best value of block size for a PMS implementation can be selected using fig. 5.7 so that the total computational burden for long-time processing is minimized.

The computational advantage of the PMS scheme is further amplified for applications in which the transform coefficients corresponding to selected levels alone are needed. This is illustrated in figure 5.8. Here the PMS and filter bank structures are compared for the direct computation of transform coefficients at the selected levels. The PMS method is always superior for such applications. It is because that, for directly computing the transform coefficients at any arbitrary level using the filter bank structure, the computation of approximations at all the preceding levels also are to be performed.

5.2.5 Case Study: PSWT Computation

The PMS structure developed in the previous section can be adapted to any 1D signal processing tasks. Different applications employing DWT based MRA techniques like noise separation, transient detection, ECG signal analysis etc. were carried out to establish the suitability of the structure in DWT computation. In this section, the PSWT computation in the PSWT based Linear Predictive Coding and Compression of ECG signals discussed in subsection 4.3.1 is presented as a typical case study.

PSWT can be interpreted as the DWT of points on the 1D signal which are one pitch period apart. Once the signal is written in the PS form, the signal samples separated by one period will appear as the adjacent points in each row sequence. Hence, a row-wise DWT computation performed on all rows of the PS matrix results in the PSWT computation of the 1D data. Since these row sequences are small in length, the PSWT computation becomes much better if the PMS structure is used instead of the pyrami-

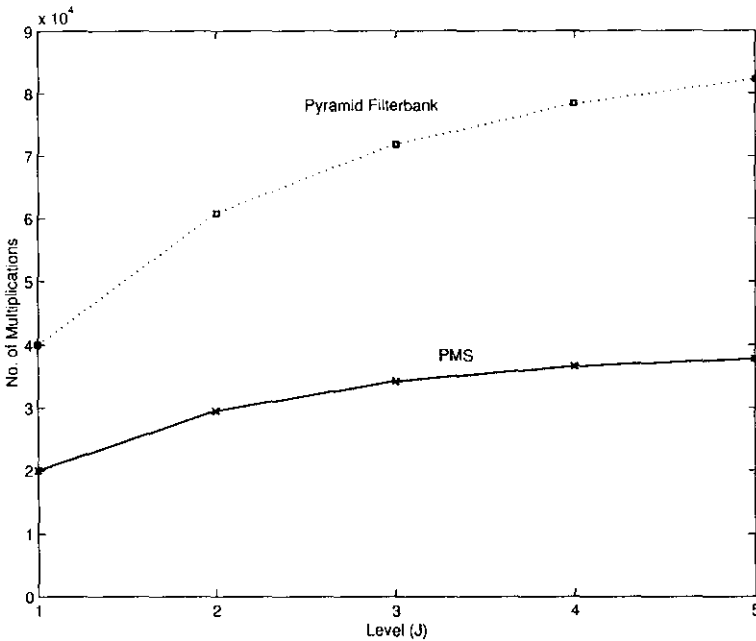


Figure 5.8: Comparison of PMS structure with filter bank for computation of coefficients at selected levels (for $l_w = 40, l_x = 1000$).

dal structure. To get a quantitative measure of the computational advantage, a 10000 sample ECG data taken from the MIT/BIH database as shown in fig. 4.10 is considered. The number of beats contained in the ECG segment was 31 and each beat was normalized to five times the maximum, making all the beats of period 1720. The normalized data matrix is now of size 1720×31 . To compare the computational advantage of the PMS structure, the PSWT computation of this data matrix is performed using different wavelets at different levels, using both pyramid and PMS structures. The computational burden in each case has been tabulated in table 5.1. Table 5.2 is the results of similar study on the signal with double the length. It can be noted that in the case of smaller input signals, the PMS structure is more efficient. Also, the computational efficiency is seen to be more in the case of higher levels of decomposition and for smaller wavelet lengths.

Signal:10000 samples of MIT-BIH ECG data(Record no.101)										
wavelet	No. of multiplications(x10000)									
	level=3		level=4		level=5		level=6		level=7	
	pyramid	PMS	pyramid	PMS	pyramid	PMS	pyramid	PMS	pyramid	PMS
db1	39	22	42	27	44	32	45	38	46	43
db2	80	54	89	70	96	86	101	102	106	118
db3	125	86	142	112	156	139	168	166	179	192
db10	226	150	266	198	302	246	335	294	367	342
db20	537	310	678	411	814	512	937	614	1079	715

Table 5.1: Comparison of computational efficiency on ECG data of length 10000 in PSWT computation on a small ECG segment.

Signal:20000 samples of MIT-BIH ECG data(Record no.101)										
wavelet	No. of multiplications(x10000)									
	level=3		level=4		level=5		level=6		level=7	
	pyramid	PMS	pyramid	PMS	pyramid	PMS	pyramid	PMS	pyramid	PMS
db1	76	43	82	54	85	64	87	75	89	86
db2	155	107	169	139	178	171	185	203	190	235
db3	237	171	262	224	280	278	294	331	306	384
db10	412	299	466	395	509	491	545	587	579	683
db20	910	619	1078	822	1227	1024	1367	1227	1503	1499

Table 5.2: Comparison of computational efficiency in PSWT computation on longer ECG segment.

5.3 WT Computation for 2D data

Computation of WT for 2D data finds manifold applications in various tasks including image compression, biomedical image processing for diagnostic purposes, edge detection, texture analysis, finger print analysis, pattern recognition, signal coding etc. Since these are prohibitively costly from a computational perspective, many researchers have addressed the computational complexity and implementation issues. The need for massive computation in 2D DWT, coupled with the demand for real-time operation in many image-processing tasks, has motivated the use of parallel processing to provide high performance at a reasonable cost. Hence, the PMS structure developed for 1D signal processing tasks has been extended for real-time computation of the DWT and IDWT of images also. The details are included in this section.

In addition to the general objectives as in 1D DWT computation, for efficient implementation of 2D DWT, the minimization of intermediate data transposition operation also play an important role.

5.3.1 Development of PMS Structure for 2D DWT Computation

2D DWT computation using the pyramidal structure has been illustrated in figure 3.9, where separable filters are employed. This scheme of computation requires a transposition of data after each stage of convolution, as horizontal and vertical convolutions are to be performed alternately. This sequential structure can be converted into an equivalent direct parallel structure by redefining the wavelet and scaling sequences at each level using the following equations [4].

$$g_{j+1}(n) = \sum_r g_j(r)g(n - 2r) \quad (5.15a)$$

$$h_{j+1}(n) = \sum_r h_j(r)g(n - 2r) \quad (5.15b)$$

where $g_1(n) = g(n)$ defines the scaling function, and $h_1(n) = h(n)$ defines the wavelet function. This is equivalent to a multi-channel filter bank decomposition with one low

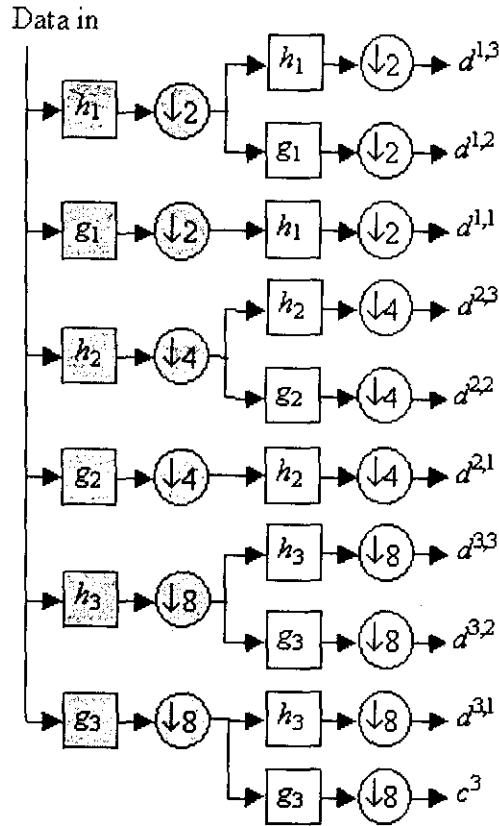


Figure 5.9: Multichannel Filter bank structure for a 3-level 2D DWT Decomposition. (Shaded blocks are row-wise operations)

pass filter, one high pass filter, the rest being band pass filters of appropriate bandwidth and center frequency, as illustrated in figure 5.9. It can be viewed as a 2D extension of the direct parallel structure shown in figure 5.1. A similar structure for reconstruction also has been developed by replacing the analysis filter sequences $h(n)$ and $g(n)$ with the corresponding synthesis filters $\tilde{h}(n)$ and $\tilde{g}(n)$.

An attraction of the above parallel multichannel structure is that the number of transpose operations has been reduced to a minimum of *two*, independent of the number of levels of decomposition. In the forthcoming paragraphs, the reconfiguration of the

above multi-channel structure into an efficient 2D PMS structure is described for both decomposition and reconstruction of images.

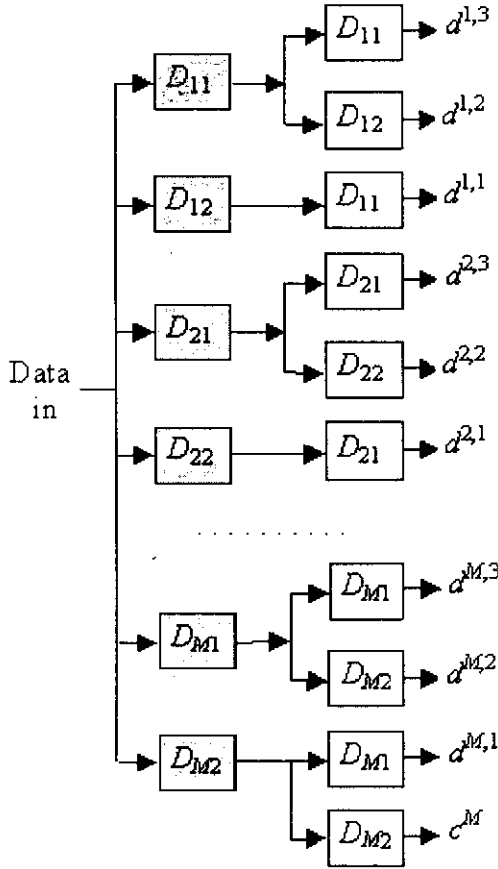


Figure 5.10: PMS structure for M-level 2D DWT decomposition. D_{js} denotes the structure shown in figure 5.2 for level j with filter h_j when $s = 1$ and g_j when $s = 2$. (Shaded blocks are row-wise operations, the rest being column-wise)

Each of the convolution cum downsampling block in fig. 5.9 can be replaced with the PMS block as done in the case of 1D (fig. 5.2). Similarly for reconstruction, each of the upsampling cum convolution block can be replaced with the single PMS block of figure 5.3. Introducing these fine-grained parallelism, the direct parallel structure of

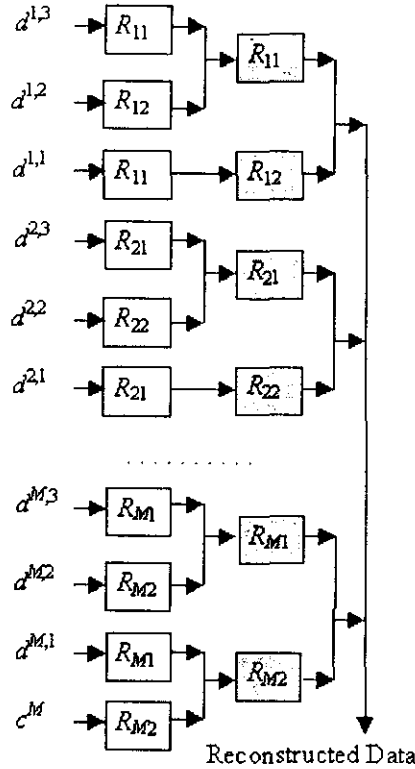


Figure 5.11: PMS structure for M-level 2D DWT reconstruction. R_{js} denotes the structure shown in figure 5.3 for level j with filter \tilde{h}_j when $s = 1$ and \tilde{g}_j when $s=2$. (shaded blocks are row-wise operations)

figure 5.9 gets modified as a 2D PMS structure as shown in figures 5.10 and 5.11 for decomposition, and reconstruction respectively. Here, c^M denotes the M^{th} level approximation and $d^{j,\tau}$ denote the details at j^{th} level. $\tau = 1, 2, 3$ for horizontal, vertical and diagonal details respectively. Similar to the 1D structure, this also has parallelism both within and between levels making it highly suitable in parallel computing environment.

5.3.2 Comparison of 2D PMS Structure with 2D Pyramidal Structure

In this section, both the conventional pyramidal structure and the PMS structure are separately analyzed in terms of the computational complexity.

5.3.2.1 2D Pyramidal structure

Consider an image of data size $X \times Y$. Each level of DWT computation involves two sets of convolutions, one being row-wise on all rows and the other being column-wise along all columns (Fig. 3.9). Recall that the convolution between two sequences of length l_1 and l_2 requires $l_1 l_2$ real multiplications and $(l_1 - 1)(l_2 - 1)$ real additions, and the resulting sequence will be of length $l_1 + l_2 - 1$. Here, eventhough the filter length remains the same throughout, the length of input sequence to each level gets successively modified due to the following reasons:

- increase in length resulting from convolution at the previous level
- decrease in length resulting from downsampling operation, and
- row-column transposition between each horizontal/vertical convolution.

Hence, at the j^{th} level, the input to each block in which convolution is carried out row-wise, will be of size $X_j^{row} \times Y_j^{row}$, where,

$$X_j^{row} = [X + (2^{j-1} - 1)(l_w - 1)]/2^{j-1} \quad (5.16a)$$

$$Y_j^{row} = [Y + (2^{j-1} - 1)(l_w - 1)]/2^{j-1} \quad (5.16b)$$

$j = 1, 2, \dots, M$; l_w is the length of the filter h or g . Similarly, for column-wise convolution, the data size at the input of the j^{th} level is $X_j^{col} \times Y_j^{col}$ where⁷,

$$X_j^{col} = [X + (2^{j-1} - 1)(l_w - 1)]/2^{j-1} \quad (5.17a)$$

$$Y_j^{col} = [Y + (2^j - 1)(l_w - 1)]/2^j \quad (5.17b)$$

Taking consideration of the change in data size during successive convolutions, the computational burden for an M-level pyramidal 2D-DWT decomposition can be derived

⁷The fact that the evaluated data size may not be an integer, will not affect the discussions to follow.

to be:

$$M_{d_{FB}} = \sum_{j=1}^M (2X_j^{row} Y_j^{row} + 4X_j^{col} Y_j^{col}) l_w \quad (5.18a)$$

and

$$A_{d_{FB}} = \sum_{j=1}^M (2X_j^{row} (Y_j^{row} - 1) + 4(X_j^{col} - 1) Y_j^{col}) (l_w - 1) \quad (5.18b)$$

where $M_{d_{FB}}$ is the number of real multiplications and $A_{d_{FB}}$ is the number of real additions during decomposition. Note that the decomposition at each level consists of two row-wise convolutions and four column-wise convolutions.

5.3.2.2 2D PMS structure

In the PMS structure, corresponding to the j^{th} level, each row of the input data and the wavelet/scaling sequence are split into 2^j subsequences (Fig. 5.2), each of length $X/2^j$ and $l_{w_j}/2^j$ respectively, where $l_{w_j} = (l_w - 1)(2^j - 1) + 1$. The number of multiplications required for the j^{th} level row-wise convolution is $2XY l_{w_j}/2^j$. Hence for an M -level decomposition, the total number of multiplications required for the row-wise convolution becomes $2XY \sum_{j=1}^M l_{w_j}/2^j$. After the row-wise convolution and downsampling by a factor of 2^j , the number of columns on which the column-wise operation is to be performed becomes Y_j^{col} . Considering all the three column-wise convolutions at each level, the number of multiplications required for the j^{th} level convolution would be $3XY_j^{col} l_{w_j}/2^j$. Therefore, the total number of multiplications required for an M -level decomposition will be $3X \sum_{j=1}^M Y_j^{col} l_{w_j}/2^j$. Hence the total number of multiplications required for the complete M -level 2D DWT decomposition becomes:

$$M_{d_{PMS}} = X(Y_M^{col} l_{w_M}/2^M + \sum_{j=1}^M (2Y + 3Y_j^{col}) l_{w_j}/2^j) \quad (5.19)$$

Proceeding similarly, the number of additions can be estimated to be:

$$\begin{aligned}
 A_{d_{PMS}} = & \sum_{j=1}^M \left\{ \left(\frac{l_{w_j}}{2^j} - 1 \right) [2X(Y - 2^j) + 3Y_j^{col}(X - 2^j)] \right. \\
 & + (2^j - 1) \left[2X \left(\frac{Y + l_{w_j}}{2^j} - 1 \right) + 3Y_j^{col} \left(\frac{X + l_{w_j}}{2^j} - 1 \right) \right] \left. \right\} \\
 & + Y_M^{col} [(X - 2^M) \left(\frac{l_{w_M}}{2^M} - 1 \right)] + \left(\frac{X + l_{w_M}}{2^M} - 1 \right) (2^M - 1)
 \end{aligned} \tag{5.20}$$

The computational burden during reconstruction has been estimated and is found to be more or less the same.

5.3.2.3 Comparison of the Computational complexity

A comparison of the complexity at different levels, between the PMS structure and the filter bank structure, has been performed for different data sizes and wavelet lengths. The results shown are based on the number of real multiplications needed for decomposition. Similar results have been obtained for addition operation also. During reconstruction, the number of multiplications remains unaltered, although the number of additions changes slightly. It is observed that the PMS structure is better for 2D DWT computation, independent of the data size for levels less than 4 and for wavelets of support less than 5. Figure 5.12 is an illustration of a typical case, which shows a comparison of the number of multiplications required for a 2-level decomposition with a wavelet of support 4. Here it is worth mentioning that, being the only known orthogonal symmetrical wavelet, *Haar* has been identified as one of the most suitable wavelets for a number of applications in image processing [9], [222]. Moreover, a good majority of practical applications of scale-space analysis require only few levels of wavelet decomposition [3], [9], [189], [221]. For such applications, the PMS computational structure is found to be advantageous over all existing schemes. The computational advantage of the PMS structure for directly estimating the transform coefficients at selected levels is more prominent in 2D case. Figure 5.13 gives a comparison of the same with the

pyramid filter bank scheme, against different levels of decomposition for an image of size 512×512 , using db_4 wavelet. Irrespective of the level of decomposition, it is found that the PMS structure shows superior performance.

Next a pictorial representation of the region based on the data size and wavelet length is given in figure 5.14, where the PMS method is superior to the filter bank method against variation in data size, wavelet support and level of decomposition. Corresponding to different values of M , the complete wavelet decomposition is per-

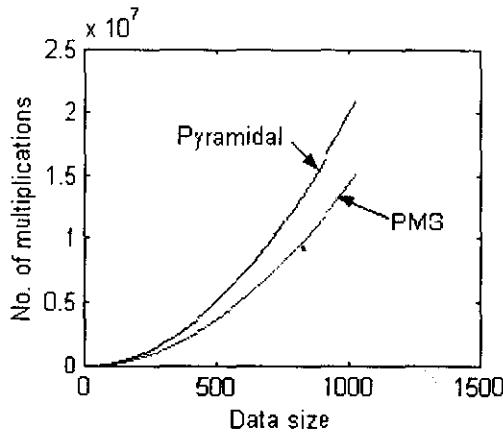


Figure 5.12: Comparison of the number of multiplications between the PMS and pyramidal structures for 2D DWT decomposition ($l_w = 4, M = 2$)

formed and two regions are identified - one in which the PMS method performs better and the other in which the Filter Bank method is superior. For example in fig. 5.14(a), the shaded region indicates the area where the PMS method outperforms the filter bank scheme for $M = 3$. Fig. 5.14(b) shows the boundary between the two regions for different values of M in the case of wavelets of larger support. Such wavelets are extensively used in various image-processing applications like texture analysis, finger print analysis, etc [225], [226]. A remarkable observation is that, once the wavelet system and the level of analysis are finalized, the input data size can be suitably selected for a frame-by-frame

processing, so that the computations are reduced considerably and the available parallel resources are fully utilized.

5.3.3 Case Study: Detection of Micro-Calcification in Mammograms

The performance of the algorithm has been studied for different applications especially with biomedical images. As a case study on edge detection, detection of micro calcification in mammograms is presented in comparison with the Mallat's algorithm. A typical mammogram from the MIAS [227] database⁸ bearing record number *mdb241* is used for the study. On analyzing the mammograms it is observed that, the resolution level 1 shows mainly the high frequency noise included in the mammograms, whereas levels 2 and 3 contained the data corresponding to microcalcifications. Higher levels contained a large correlation with the non-uniform background. Hence for detecting the microcalcifications, levels 2 and 3 were sufficient.

Figure 5.15 shows the the region of micro-calcification thus detected (without applying enhancement techniques) from the mammogram using the PMS structure employing *db4* wavelet. The locations of micro-calcifications are clearly identified (marked in circle) in the processed image given in fig. 5.15(b). It is found that the PMS structure could reduce the arithmetic operations considerably.

5.4 Conclusions

A novel parallel multiple subsequence computational structure has been presented in this chapter as an alternative for Pyramid scheme. This could eliminate all the unnecessary computations that are present in the pyramid structure. The PMS structure being parallel, it is found better for directly computing the transform coefficients at any

⁸Mammographic Image Analysis Society data base which is freely distributed digitized mammograms from the University of Essex, England. The images in the database are digitized at 200-micron pixel edge. The accompanied 'Ground Truth' with each image contains details regarding the character of the background tissue, class and severity of the abnormality and coordinates of the centre and radii of the cluster.

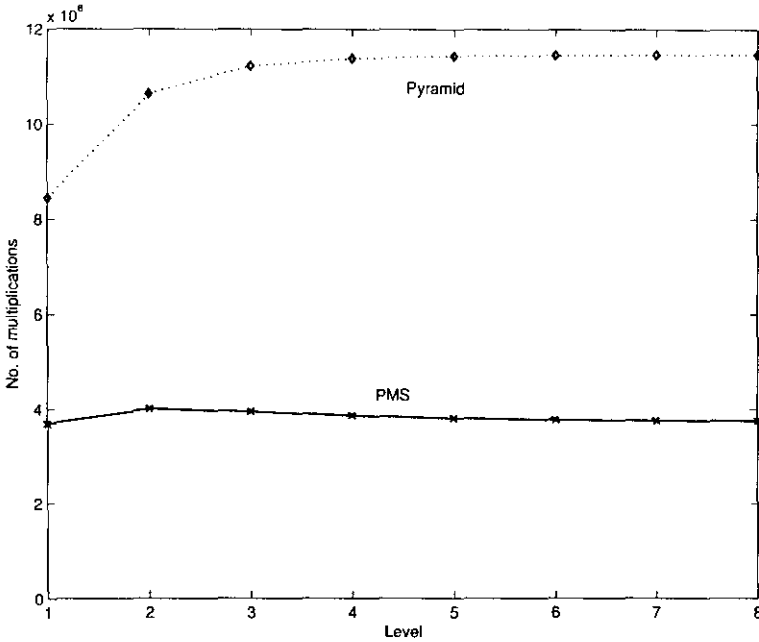


Figure 5.13: Comparison between the 2D pyramidal and PMS structures in terms of the number of multiplications needed for directly computing the details at any desired level

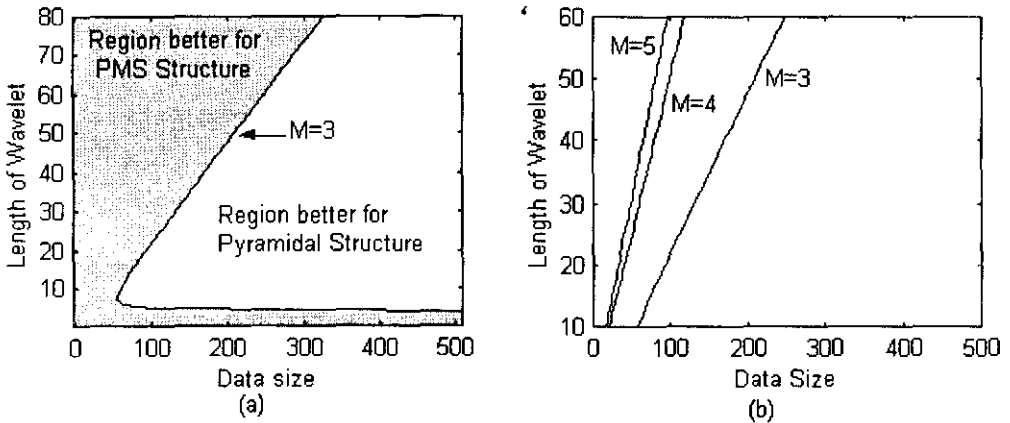


Figure 5.14: Comparison of computational efficiency between the 2D PMS and Pyramidal Structures for Variable wavelet lengths and data sizes. (a)for a typical level M=3 (b)for different levels and larger wavelets

arbitrary selected level of decomposition without going through the intermediate levels. The expressions for the computational complexity of the PMS structure has been derived and compared with the pyramid tree. The efficiency has been shown to be more in the case of lower levels of decomposition and is dependent on the length of wavelet also. The efficiency of the scheme is still better for computations using the *Haar* wavelet. The computational structure has been extended for the 2D wavelet system also. Cases of PSWT computation and edge detection in biomedical images have been presented as case studies to highlight the advantage of the structure in both 1D and 2D signal processing applications. For 2D applications it has the added advantage that the data transposition operation to be performed is made the minimum.

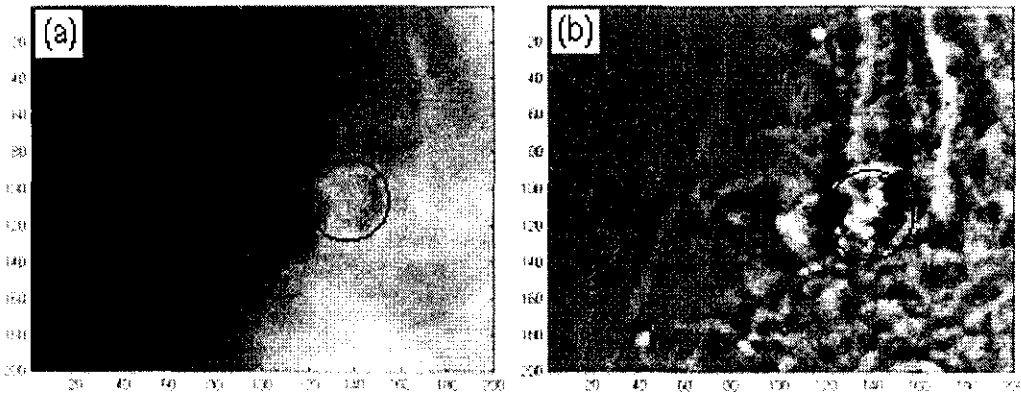


Figure 5.15: Detection of micro-calcification in Mammogram using the PMS structure.(a) Original image. (b) Processed image showing clear locations of micro-calcifications

Chapter 6

Summary and Conclusions

A brief summary of the research work conducted and the important conclusions thereon are highlighted in this chapter. The scope for further work in this field as an extension of the present study, has also been discussed.

6.1 Summary of the Work and the important Conclusions

The work presented in this thesis mainly concentrates on two important inter-related topics in time-frequency analysis.

The first topic of interest has been the methods for signal coding and compression. A new compression scheme has been developed for pseudo-periodic signals in which the Pitch-Synchronous Wavelet Transform technique is uniquely combined with the popular Linear Predictive Coding technique. The local periods of the signal under consideration are estimated and using the period information, the signal is represented in Pitch-Synchronous form. After normalizing the PS data, the PSWT is computed. The LPC parameters are estimated for the PSWT coefficients corresponding to each level of decomposition. The predictor coefficients are suitably encoded and packed with other relevant information for transmission to the remote location where it is decoded for reconstructing the original signal. Case studies on typical pseudo-periodic signals have been carried out and the performance of the scheme has been evaluated using standard measures. The PSWT based LPC technique results in better compression and higher reconstruction quality.

It has been observed that the predictive coding of PSWT coefficients has greater potential in signal compression compared to the ordinary DWT based methods. The method has been validated with different classes of practical signals like ECG, human voice and musical instrument tones. The various parameters for the compression such as the type of wavelet, order of prediction, level of decomposition, etc. were optimized for each of these classes. The UDWT based methods are found superior in period estimation of signals like ECG and musical instrument tones, compared to the autocorrelation based methods. The period normalization is found to have improved the beat-to-beat correlation, contributing to better compression.

The effectiveness of the method is largely dependent on the accuracy with which

the local periods are estimated. The noise contamination in the signals is found to reduce the efficiency of compression. Another shortcoming of this method is that the compression is computationally intensive.

The second topic of consideration has been the computational issues of Discrete Wavelet Transform. The most popular algorithm for DWT computation employs the 'pyramid structure' developed by S.G. Mallat. It basically follows a sequential structure. For a number of applications, this structure is found to be inefficient in terms of the number of computations. As a better choice in such cases, a novel Parallel Multiple Subsequence structure has been developed by uniquely combining the *noble identities* in multirate systems with the principle of *polyphase splitting*. The input data and wavelet/scaling filter sequences are divided into a level-dependent number of parallel subsequences resulting in a highly parallel environment especially at higher levels. The complexity involved while employing this PMS structure for DWT computation has been analyzed in detail and compared with that of the pyramid structure. The PMS structure could eliminate all the irrelevant computations that are to be carried out while using the pyramid structure. The PMS structure being parallel, it is found better for directly computing the transform coefficients at any arbitrary selected level of decomposition without going through the intermediate levels. The efficiency of the PMS structure has been shown to be more in the case of lower levels of decomposition and is dependent on the length of wavelet also. The scheme is still better for computations using the *Haar* wavelet.

Both 1D and 2D computational structures have been derived. Typical signal processing applications have been presented as case studies wherein the PMS structure is identified to be advantageous, both computationally and algorithmically. Case studies include the PSWT computation and the edge detection in biomedical images. This structure is shown to be better for PSWT and Block DWT computation. For 2D ap-

plications it has the added advantage that the data transposition operation has been reduced to the minimum.

6.2 Scope for Further Investigations

The PSWT based compression schemes are highly dependent on the accuracy of local period estimation. The methods for period estimation are found to be dependent on the signal characteristics and none of them are found universally acceptable. Eventhough lot of research has been carried out in this field, accuracy demands are still not met. A signal dependent adaptive period estimator is necessary to enhance the compression scheme.

The PSWT based compression scheme is applicable only for the pseudo-periodic regions present in the signal. Practical signals comprise of other regions also, which have to be separately dealt with in a real-time processing environment. This necessitates the use of a signal adaptive segmentation scheme at the preprocessing end.

The PSWT based feature enhancement technique proposed in this thesis is based on simple thresholding of PSWT coefficients. Application of energy dependent hard/soft thresholding can be attempted for signal enhancement by eliminating the instrument dependent noise part without any loss in the instrument tonal quality.

The compression scheme has concentrated on optimization of the variants used for signal coding only. Since majority of the PSWT coefficients of pseudo-periodic signals are insignificant, especially for noise free signals, dynamic bit allocation can be attempted thereby achieving higher levels of compression. The possibility of totally avoiding the insignificant coefficients also can be attempted. This will be of special importance in the case of signals where a good number of samples are insignificant. For example, in the case of ECG, the diagnostic information contained is in general localized to the QRS complex. Moreover by sending the residue part along with other information, the signal quality can be further enhanced, eventhough at the cost of compression.

The PMS structure developed in this thesis has been validated for 1D and 2D DWT computations. It is expected to give still better performance with Wavelet Packet Transform computation, which can also be taken up as an extension of this work.

Appendix A

WT Based Signal Compression

A.1 Introduction

Efficient coding and compression is vital in compact digital representation of signals. For high quality applications, signals are sampled at high frequencies and quantized at high resolution. This necessitates high storage space and increased transmission rate/bandwidth. For efficient data transmission and storage, the signals need to be represented with a minimum number of bits while achieving excellent signal reproduction, fully retaining all perceivable attributes in the signal. To accomplish this, one should eliminate the redundancies present in the signal. This is particularly significant in the case of audio signals, where one can exploit the human auditory perceptual characteristics also. Studies on human sound perception show that sound pressure at a particular frequency and time instant masks the sound below a threshold at nearby frequencies and time instants, a phenomenon known as *auditory masking* [119], [228]. Making use of this perceptual property, considerable reduction of data rate could be achieved.

Being a highly flexible means of signal analysis, the WT and the WPT¹ are very effective in audio data compression, feature extraction, signal source modelling etc. WT and WPT have been well established as a mathematical tool for non-stationary signal

¹Wavelet Packet Transform

analysis [118], [11] [229]. It has been remarked [205] that, there are no hard and fast rules for selecting the best wavelet for various applications. The central measure in choosing a wavelet lies on its match with the signal itself, in terms of its statistical characteristics. The choice of a particular wavelet basis to suit a specific class of signal is a major topic of interest to research community.

In this appendix, a comparison of the efficacy of the WT and the WPT in audio signal compression is presented. A study on selection of the best wavelet basis for this application has also been considered. Compression using the simple thresholding technique only has been carried out for this comparative study.

A.2 Implementation

Wendt et al. [156] has proved that Haar wavelet is the best in segmentation and pitch determination of speech signals. The study in this direction has been further extended by analyzing the performance of different wavelets for general audio processing applications. A collection of speech data at 16-bit resolution, from both male and female speakers sampled at 8 to 44.1 kHz was used for the study. Vocal music and instrumental tones also have been considered. The presenting the results, the following signals have been considered.

1. F1: Female voice ('Your Complaint Number is'), 8kHz, 16 bit, 10832 samples.
2. F2: Female voice ('The Pipe Started Rusting, While New'), 22kHz, 16 bit, 79792 samples.
3. F3: Female voice ('The Pipe Started Rusting, While New'), 44.1kHz, 16 bit, 105248 samples.
4. F4: Female voice ('The Pipe Started Rusting, While New'), 8kHz, 16 bit, 21728 samples.

5. V1: Violin tone (Natural Scale), 44.1kHz, 16 bit, 213856 samples.
6. M1: Male voice (Music-Shankarabharana Raaga), 8kHz, 16 bit, 64000 samples.

These signals were decomposed to 4 levels and reconstructed back using the pyramid structure shown in fig. 3.3 and fig. 3.6. It is seen that majority of the transform coefficients carry negligible information and hence they can be discarded without much loss of intelligibility. Moreover, for certain class of audio signals like speech, the information content is mainly concentrated in a narrow band. Hence, by decomposing the sampled speech into different sub-bands, irrelevant components in the signal could be eliminated, thereby achieving compression. The study was conducted using WT and WPT techniques with and without compression. To achieve compression, the coefficients below the specified threshold with respect to the maximum value of the transform coefficients, were made zero before attempting reconstruction. The objective evaluation of the reconstructed sound was done by calculating the SNR. For subjective evaluation, listening tests [218] were conducted using ten subjects. Special care was taken to eliminate external interference, background noise, and echo-effects. Training sets were used to familiarize the subjects participated in the listening test. They were asked to rate the quality as *excellent*, *good*, *fair*, *poor* or *bad*. These ratings were allotted grade numbers 5, 4, 3, 2, and 1, respectively. The MOS value was calculated by taking the arithmetic mean of the grades voted by them.

A.3 Results and Discussion

Table A.1 gives of the results of the objective evaluation based on a 4-level wavelet and wavelet packet analysis using different wavelets. The signals were reconstructed from the transform coefficients without applying any compression. In each case, the SNR was computed using equation 4.5.

Though the SNR is different for different wavelets, the subjective quality of recon-

Signal	Method	SNR obtained (dB)													
		haar	db2	db4	db10	sym5	sym8	coif1	coif5	bior1.3	bior2.8	bior3.9	bior4.4	bior5.5	bior6.8
F1	WT	308	236	231	220	246	249	232	157	307	307	305	232	230	245
	WPT	305	233	228	217	243	249	228	155	305	304	302	229	228	242
V1	WT	305	245	234	228	258	250	243	167	305	304	303	238	234	253
	WPT	305	244	234	228	257	250	242	167	304	304	303	238	234	253
M1	WT	306	236	229	220	247	246	232	156	305	305	303	231	229	245
	WPT	305	234	228	219	246	245	231	156	304	304	303	230	228	244
F2	WT	305	235	229	220	246	247	232	156	305	304	303	231	229	245
	WPT	305	234	229	219	246	247	231	156	304	304	302	230	229	244

Table A.1: Objective performance of wavelets on audio signal processing.

structured sound was found excellent in all the cases. This is justified, since the high values of SNR make the error in reconstruction well below the ATH². The tabulation shows that, for both wavelet and wavelet packet transforms, *Haar* and *Bior1.× / 2.× / 3.×* wavelets give better performance in respect of speech, music, instrumental tones, male voice and female voice, irrespective of the sampling frequency. In all the cases Haar wavelet was found to be the best.

To probe in to the possibility of low complexity signal compression using wavelets, simple thresholding technique was attempted. The signals were analyzed using the Haar wavelet. The corresponding results are summarized in table A.2. It is observed that, for female voice sampled at 8 kHz, very good quality audio is possible for a CR of up to 5.5 and good quality is attainable for a value of even 10. Due to data redundancy, better compression could be achieved for signals sampled at higher rates. For the same CR, though the objective quality of the reconstructed male voice is better than the female voice, the subjective quality is less.

Table A.3 gives a comparison on the effectiveness of different wavelets for speech compression, based on simple thresholding. The signal under consideration is 'F4'. Though Haar wavelet was identified as the best for audio signal analysis, the above study suggested that 'Db4' and 'Bior5.5' wavelets are more suitable for speech compression.

²Auditory Threshold of Hearing

Signal and sampling rate	Threshold (%)	Wavelet method			WP method		
		Compression Ratio	SNR (dB)	MOS (1-5)	Compression Ratio	SNR (dB)	MOS (1-5)
F1 8kHz	0	1.0	308	5	1.0	305	5
	1	1.3	31	5	1.3	32	5
	3	2.0	20	5	2.1	20	5
	5	2.9	15	5	3.0	15	5
	7	3.8	13	4.7	4.0	13	4.7
	10	5.5	10	4	5.8	10	3.5
F3 44.1kHz	0	1.0	305	5	1.0	305	5
	1	4.9	23	5	5.0	24	5
	3	11.8	15	4.2	11.7	16	4.2
	5	19.0	12	3.2	18.8	12	3.2
	7	26.0	10	2.5	26.0	10	2.5
	10	37.0	8	2.2	36.7	8	2.2
M1 8kHz	0	1.0	305	5	1.0	305	5
	1	2.3	22	4.5	2.3	22	4.5
	3	4.9	13	3.5	5.1	13	3.5
	5	8.2	9	2.2	8.4	9	2.0
	7	12.3	7	1.0	12.8	7	1.0
	10	21.7	5	1.0	22.3	5	1.0

Table A.2: Effect of simple thresholding on audio signal compression.

Method	Threshold (%)	CR and MOS obtained (Signal used: F4)													
		haar		db4		db10		sym5		coif5		bior3.9		bior5.5	
		CR	MOS	CR	MOS	CR	MOS	CR	MOS	CR	MOS	CR	MOS	CR	MOS
WT	0	1.0	5	1.0	5	1.0	5	1.0	5	1.0	5	1.0	5	1.0	5
	1	2.7	5	2.7	5	2.7	5	2.8	5	2.8	5	3.4	5	2.6	5
	3	3.5	4.5	4.3	5	4.3	5	4.6	5	4.5	5	5.7	5	4.0	5
	5	4.9	3.7	5.9	4.5	5.8	4.5	6.2	4.4	6.1	4.4	7.7	4	5.4	4.6
	7	6.3	3.5	7.5	4.2	7.3	4.0	7.9	3.6	7.6	3.5	9.7	3.2	7.0	4.4
	10	8.6	3.0	10	3.0	9.6	2.7	11	2.5	10	2.5	13	2.5	9.6	3.0
WPT	0	1.0	5	1.0	5	1.0	5	1.0	5	1.0	5	1.0	5	1.0	5
	1	2.6	5	2.8	5	2.9	5	3.0	5	3.0	5	4.3	5	2.9	5
	3	3.4	4.5	4.4	5	4.5	5	4.8	5	4.7	4.7	7.7	3.5	4.6	5
	5	4.5	3.7	5.9	4.2	6.1	4.2	6.5	4.2	6.4	4.0	11	2.5	6.4	4.2
	7	5.7	3.6	7.5	3.7	7.7	3.5	8.4	3.2	8.2	3.0	14	1.8	8.4	4.0
	10	7.9	3.1	10	2.8	11	2.5	12	2.2	12	2.2	21	1.2	13	2.5
	15	13	2.5	18	1.2	18	1.0	22	1.0	20	1.0	44	1.0	22	1.0

Table A.3: Effect of change of wavelet on speech compression using simple thresholding.

A.4 Conclusion

The application of different wavelets for audio signal processing has been explored. It was found that the *Haar* wavelet is best suited for general time-frequency analysis of audio signals, irrespective of the sampling frequency. But for compression applications based on simple thresholding techniques, *Db4* and *Bior5.5* wavelets were found to be even better.

Simple thresholding strategy could be efficiently applied for audio compression employing wavelet-based decomposition. For speech signals sampled at 8kHz, good quality speech output was obtained at a compression ratio of the order of 10. The value went even above 50 for a sampling rate of 44.1kHz, still maintaining the same audio quality. Compression achieved for male voice is comparatively less.

Though wavelet packets decompose the signal in both high frequency and low frequency bands with better resolution, noticeable difference is not perceived in comparison with wavelet transform. However, since wavelet packets are computationally more intensive, for audio signal processing applications the WT method is preferred over WPT.

Appendix B

WT based Signal Segmentation

B.1 Introduction

Accurate segmentation of signals into different distinguishable regions like pseudo-periodic, random, transition etc. is very important in signal processing and compression applications in particular, as the processing methods and strategy is highly dependent on the signal characteristics. Most of the classification methods that exist today [230], [231], [232], [233] as applicable to 1D signals are pertaining to speech as it has tremendous application in entertainment electronics. Moreover, these methods classify speech signals into unvoiced /voiced, or unvoiced /voiced / silent regions only. The regions of *transition* between any of the above have distinct characteristics when compared with voiced, unvoiced and silent regions [29], [224]. The characteristics of the transition region depend on the nature of the preceding and succeeding segments. Work has been recently reported [214] about a novel method of classification of speech signals into the above four distinct regions, in which the autocorrelation method was employed for pitch identification. It has been proved that codecs based on such a classification has better efficiency compared to other state of the art codecs [234]. Even though the features of music signals are quite different from that of speech signals [235], a classification of music signals into Voiced, Unvoiced, Silent and Transition regions exploiting the exclusive

characteristic features of music is not yet seen attempted. Segmentation and classification of audio signals could be made using moderately simple parameters derived from the audio signal such as RMS energy or ZCR¹. But such a method can achieve only limited accuracy. The voiced/ unvoiced/ silent classification is traditionally tied to the determination of periodicity (pitch period) [236]. Audio signals being quasi-periodic, accurate determination of periodicity always raised problems resulting in wrong classification. Threshold based classifiers like the conventional Cepstrum and autocorrelation methods [153] are typically used for voicing decisions.

Although encouraging results have been obtained for speech, the autocorrelation-based method of pitch determination is not often satisfactory when applied to music signals [157], [110]. This is primarily because of the large range of fundamental frequency and the variety of spectra encountered in music signals. It may be noted that a musical signal is a logarithmic organization of pitch based on the octave, which is the periodic dilation between two pitches, when one is twice the frequency of the other. Hence wavelet based pitch estimation [154], [156] is found to be a more natural choice for musical applications.

In this appendix, a WT based method for audio signal classification and segmentation in which signals are classified into Transition regions also in addition to the conventional classification into Voiced, Unvoiced and Silent regions, is presented. Appropriate threshold values for the statistical features such as SZR², STE³, the ZEP⁴, and the pitch correlation factor are utilized in the classification process. The UDWT techniques are employed for period estimation. The proposed method is made computationally attractive by restricting the WT computation only to a few selected levels.

¹Zero Crossing Rate

²Short-Time Zero Crossing Rate

³Short-Time Energy

⁴Zero-Crossing-Energy Product

B.2 The Classification Algorithm

The first step in the classification process is the statistical feature extraction. The signals under study are normalized and segmented into blocks of size corresponding to 20 ms of data approximately. It is assumed that the pitch of vocal music has a dynamic range of five octaves. Following statistical parameters are estimated for each segment of the signal.

B.2.1 Short-Time Energy

A measure of the energy for each segment is a convenient parameter that reflects the variations of the amplitude of the signal and has been widely used in classification problems. The *STE* of the i^{th} block of the signal, $x_i(n)$, is defined as:

$$STE_i = \sum_{n=0}^{N-1} |x_i(n)|^2 \quad (\text{B.1})$$

where N is the block size.

B.2.2 Short-Time Zero Crossing Rate

A zero crossing occurs in a discrete time signal if successive samples have different algebraic signs. Although the procedure needs only a comparison of the signs of two successive samples, the signal has to be preprocessed to eliminate noise, offset, etc. to ensure accurate measurement. The sampling frequency of the signal also determines the time-resolution of the zero-crossing measurements. The *SZR* corresponding to the i^{th} segment is:

$$SZR_i = \sum_{n=1}^N |sgn[x_i(n)] - sgn[x_i(n-1)]| \quad (\text{B.2})$$

If the *SZR* exceeds a given threshold, the corresponding segment is likely to be unvoiced, and it is too low a value for silent regions. It is observed that the *median* of

the SZR is an appropriate value to be used as a signal-dependent threshold.

B.2.3 Short-Time Zero-Crossing Energy Product

Since different classes of music segments may have comparable values of STE or SZR , their product ZEP has been defined as yet another discriminating parameter in the classification process. The ZEP of the i^{th} block is computed as:

$$ZEP_i = STE_i \cdot SZR_i \quad (B.3)$$

The value of ZEP will be considerably high for Transition from/to Voiced segments. For other Transition regions, its value is comparatively less.

B.2.4 Pitch Correlation Factor

The Pitch Correlation Factor β will be of use in the detection of Transition from Voiced/Unvoiced regions giving a marked discrimination when the signal energy is reasonably high giving a wrong notion of the block to be Voiced/Unvoiced. The β parameter for the i^{th} block is computed using the equation:

$$\beta_i = \frac{\sum_{n=P_{i_1}+1}^N x_i(n)x_i(n-P_{i_1})}{\sum_{n=P_{i_1}+1}^N [x_i(n-P_{i_1})]} \quad (B.4)$$

where P_{i_1} is the value of the first pitch period of the i^{th} segment. For highly voiced segments, β approaches unity as evident from equation B.4. During voicing Transitions also, especially in the case of vocal music, the value of β will be reasonably high and hence ample care should be taken to fix up the threshold of β in the decision making process. Moreover, during Transition phase from Voiced to Unvoiced/Silent regions, the successive pitch periods will show a gradual change which is strongly dependent on the *Thala* (Rhythm) of the music. The pitch identification is performed using UDWT coefficients as described in section 4.2.1.2.

The overall flowchart used for the segmentation followed by classification is given in figure B.1.

B.3 Results and Discussions

The proposed classification scheme has been applied on a wide range of classical music signals sung by a group of artists including both male and female. The sampling rates for the test signals were 8 kHz and 22.05 kHz. Using the experimentally selected values of the statistical parameters, accurate classification of the signals into Voiced, Unvoiced, Silent and Transition regions could be achieved. The results were verified by manual classification of the signals. The validity of the classifier was also tested with different test signals mutilated by noise. Except in occasions where the transition region is insignificant, the algorithm resulted in the exact segmentation of the signals. One typical case is illustrated in figure B.2.

B.4 Conclusions

An efficient scheme for classification of audio signals into Voiced, Unvoiced, Silent and Transition regions after segmenting into blocks of fixed frame size, has been developed. The conventional classification method based on audio features such as Short-Time Energy, Zero-Crossing Rate, measure of periodicity etc. are combined with the state-of-the-art Wavelet Transform methods. The proposed method gives better recognition score for classical vocal music when compared to auto-correlation based classification methods. The statistical parameters used for the classification process is required to be adapted to the signal properties. The classifier works well with wavelets, which are the first derivative of smooth functions. All the drawbacks of classical methods in classification of vocal music due to discriminative characteristics of music, are well taken care of in this method.

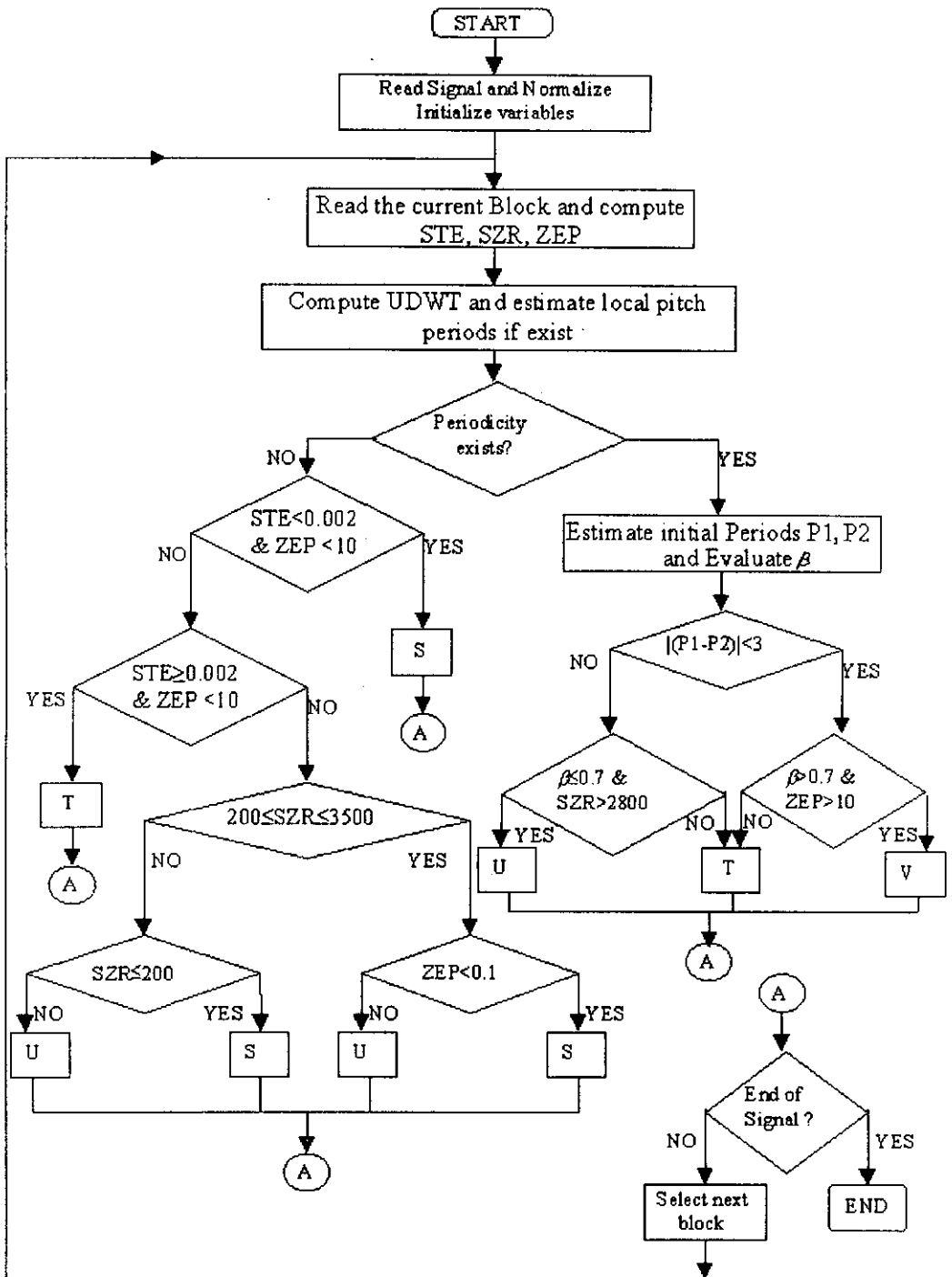


Figure B.1: Flow chart showing segmentation and classification of vocal music using WT techniques. V,U,S and T stands for Voiced, Unvoiced , Silent and Transition regions respectively

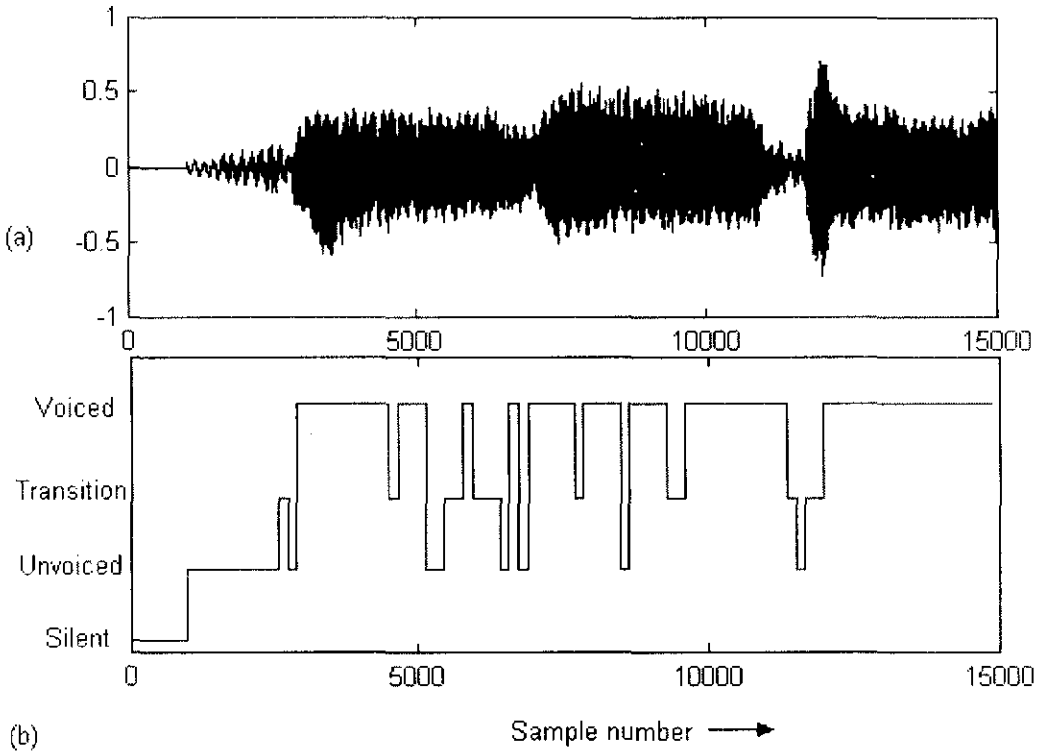


Figure B.2: Classification of a piece of Classical music sung by a female artist (a) Original Signal (b) Classifier Output

Bibliography

- [1] S. W. Smith, *The Scientists and Engineers Guide to Digital Signal Processing*. San Diego, California: California Technical Publishing, 2nd ed., 1999.
- [2] I. Daubechies, *Ten Lectures on Wavelets*. Society for Industrial and Applied Mathematics, 1992.
- [3] S. Mallat, "A theory for multiresolution signal decomposition: The wavelet rapresentation," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. PAMI-11, no. 7, pp. 674–693, 1989.
- [4] O. Rioul and P. Duhamel, "Fast algorithms for discrete and continuous wavelet transforms," *IEEE Trans. Inf. Theory*, vol. 38, pp. 569–586, March 1992.
- [5] M. Vetterli and J. Kovacevic, *Wavelets and Subband Coding*. Englewood Cliffs, NJ: Prentice-Hall, 1995.
- [6] G. Beylkin, "On the representation of operators in bases of compactly supported wavelets," *SIAM J. Numer. Anan.*, vol. 29, no. 6, pp. 1716–1740, 1992.
- [7] M. J. Shensa, "The discrete wavelet transform: wedding the átrous and mallat algorithms," *IEEE Trans. Signal Processing*, vol. 40, pp. 2464–2682, October 1992.
- [8] H. Guo and C. S. Burrus, "Convolution using the undecimated discrete wavelet transform," in *Proc. Int. Conf. Acoust. Speech, Signal Processing, ICASSP 96*, (Atlanta), May 1996.
- [9] G. Strang and T. Nguyen, *Wavelets and Filter banks*. MA: Wellesley-Cambridge Press, 1996.
- [10] M. G. Mini, V. P. Devassia, and T. Thomas, "One-dimensional processing for edge detection in biomedical images using pitch-synchronous wavelet transform." International Conference on Mathematical & Engineering Techniques in Medicine and Biological Sciences, METMBS'02, USA, June 24–27 2002.

-
- [11] G. Evangelista, "Pitch-synchronous wavelet representation of speech and music signals," *IEEE Trans. SP*, vol. 41, pp. 3313–3330, December 1993.
- [12] B. Gold and C. M. Rader, *Digital Processing of Signals*. NY: McGraw-Hill, 1969.
- [13] A. V. Oppenheim, ed., *Applications of Digital Signal Processing*. Englewood Cliffs, NJ: Prentice Hall, 1978.
- [14] A. V. Oppenheim and R. Schafer, *Digital Signal Processing*. Englewood Cliffs, NJ: Prentice Hall, 1975.
- [15] A. V. Oppenheim, R. W. Schafer, and J. R. Buck, *Discrete-Time Signal Processing*. India: Pearson Education Asia, 2001.
- [16] J. R. Johnson, *Introduction to Digital Signal Processing*. Englewood Cliffs, NJ: Prentice Hall, 1989.
- [17] J. G. Proakis and D. G. Manolakis, *Digital Signal Processing, Principles, Algorithms and Applications*. Prentice Hall of India, 3 ed., 1998.
- [18] S. K. Mitra, *Digital Signal Processing—A Computer Based Approach*. India: McGraw Hill Publishing Co., 2 ed., 2001.
- [19] P. Diniz S. R., E. da Silva A.B., and S. L. Netto, *Digital Signal Processing, System Analysis and Design*. Cambridge University Press, 2002.
- [20] E. C. Ifeachor and B. W. Jervis, *Digital Signal Processing—A Practical Approach*. India: Pearson Education Asia, 2 ed., 2002.
- [21] R. G. Lyons, *Understanding Digital Signal Processing*. India: Pearson Education Asia, 2001.
- [22] H. C. Andrews, "Digital image processing," *IEEE Spectrum*, vol. 16, pp. 38–49, April 1978.
- [23] H. C. Andrews, A. G. Tescher, and R. Kruger, "Image processing by digital computer," *IEEE Spectrum*, vol. 9, pp. 20–32, July 1972.
- [24] K. R. Castleman, *Digital Image Processing*. Englewood Cliffs, NJ: Prentice Hall, 1979.
- [25] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*. NY: Addison-Wesley, 1993.

- [26] B. R. Hunt, "Digital image processing," *Proc. IEEE*, vol. 63, pp. 693–708, April 1975.
- [27] E. A. Robinson and S. Treitel, *Geophysical Signal Analysis*. Englewood Cliffs, NJ: Prentice Hall, 1976.
- [28] L. C. Wood and S. Treitel, "Seismic signal processing," *Proc. IEEE*, vol. 63, pp. 649–661, April 1975.
- [29] L. R. Rabiner and R. W. Schafer, *Digital processing of speech signals*. Prentice Hall Inc., 1978.
- [30] W. A. Lea, ed., *Trends in Speech Recognition*. Englewood Cliffs, NJ: Prentice Hall, 1980.
- [31] B. A. Sherwood, "The computer speech," *IEEE Spectrum*, vol. 16, pp. 18–25, August 1979.
- [32] R. F. Lyon, "A computational model of binaural localization and separation," *Proc. IEEE Int. Conf. Acoust. Speech, Signal Processing, ICASSP 83, Boston*, pp. 1148–1151, 1983.
- [33] T. W. Parsons, "Separation of speech from interfering speech by means of harmonic selection," *J. Acoust. Soc. Amer.*, vol. 60, no. 4, pp. 911–918, 1976.
- [34] R. J. Stubbs and Q. Summerfield, "Algorithms for separating the speech of interfering talkers: Evaluations with voiced sentences and normal hearing and hearing-impaired listeners," *J. Acoust. Soc. Amer.*, vol. 87, no. 1, pp. 359–372, 1990.
- [35] A. Shamsoddini and P. N. Denbigh, "A system for speech separation," *Proc. 4th European Conference on Speech Communication and Technology, EUROSPEECH'95, Madrid*, pp. 1583–1586, 1995.
- [36] A. Shamsoddini and P. N. Denbigh, "Enhancement of speech by suppression of interference," *Proc. ICSP'96*, pp. 753–756, 1996.
- [37] J. Schmidt C. and J. Rutledge C., "Multichannel dynamic range compression for music signals," in *Proc. Int. Conf. Acoust. Speech, Signal Processing, ICASSP '96*, pp. 1013–1016, 1996.
- [38] O. van der Vrecken, L. Hubaut, and F. Coulon, "A new subband perceptual audio coder using celp," *Proc. Int. Conf. Acoust. Speech, Signal Processing, ICASSP '98*, vol. 6, pp. 3661–3664, 1998.

- [39] C. Ö. Etemoğlu and V. Cuperman, "Matching pursuits sinusoidal speech coding," *IEEE Trans. Speech and Audio Processing*, vol. 11(5), pp. 413-424, September 2003.
- [40] R. W. F. Campbell and A. Murray, *Dynamic Electrocardiography*. New York: Churchill Livingstone, 1985.
- [41] J. Cox, F. Noelle, H. Fozzard, and G. Oliver, "Aztec: A preprocessing program for real-time ECG rhythm analysis," *IEEE Trans. Biomed. Eng.*, vol. BME-15, pp. 128-129, 1968.
- [42] W. Mueller, "Arrhythmia detection program for an ambulatory ECG monitor," *Biomed. Sci. Instrum.*, vol. 14, pp. 81-85, 1978.
- [43] J. Abenstein and W. Tompkins, "New data-reduction algorithm for real-time ECG analysis," *IEEE Trans. Biomed. Eng.*, vol. BME-29, pp. 43-48, 1982.
- [44] D. A. Dipersio and R. C. Barr, "Evaluation of the fan method of adaptive sampling on human electrocardiograms," *Med. Biomed. Eng. Computing*, pp. 401-410, September 1985.
- [45] S. C. Tai, "Slope—a real-time ECG data compressor," *Med. Biomed. Eng. Computing*, vol. 29, pp. 175-179, March 1991.
- [46] S. C. Tai, "Aztdis - a two phase real-time ECG data compressor," *J. Biomed. Eng.*, vol. 15, pp. 510-515, November 1993.
- [47] G. Nave and A. Cohen, "ECG compression using long-term prediction," *IEEE Trans. Biomed. Eng.*, vol. 40, no. 9, pp. 877-885, 1993.
- [48] A. Cohen and Y. Zigel, "Compression of multichannel ECG through multichannel long-term prediction," *IEEE Engineering in Medicine and Biology*, pp. 109-115, 1998.
- [49] S. Jalaliddine, C. Hutchens, R. Strattan, and W. Coberly, "ECG data compression techniques—a unified approach," *IEEE Trans. Biomed. Eng.*, vol. 37, pp. 329-343, April 1990.
- [50] D. Haugland, J. Heber, and J. Hosoy, "Optimization algorithms for ECG data compression," *Med. Biomed. Eng. Computing*, vol. 35, pp. 420-424, July 1997.
- [51] R. Nygaard and D. Haugland, "Compressing ECG signals by piecewise polynomial approximation," *Proc. Int. Conf. Acoust. Speech, Signal Processing*, pp. 1809-1812, May 1998.

- [52] M. C. Aydin, A. E. Cetin, and H. Köymen, "ECG data compression by sub-band coding," *Electron. Lett.*, vol. 27, pp. 359–360, February 1991.
- [53] S. C. Tai, "Six-band sub-band coder on ECG waveforms," *Med. Biomed. Eng. Computing*, vol. 30, pp. 187–192, March 1992.
- [54] J. H. Husoy and T. Gjerde, "Computationally efficient subband coding of ECG signals," *Med. Eng. Phys.*, vol. 18, pp. 132–142, March 1996.
- [55] A. Djohan, T. Q. Nguyen, and W. J. Tompkins, *ECG compression using discrete symmetric wavelet transform*. Montreal, QC, Canada, 1995. presented at the 17th IEEE Int. Conf. Medicine and Biology.
- [56] M. Nakashizuka, H. Kikuchi, H. Makino, and I. Ishii, "Data compression by wavelet zero-crossing representation—application of ECG data," *Proc. ICICE, CAS 95-63*, pp. 57–64, September 1993.
- [57] C. P. Mammen and B. Ramamurthi, "Vector quantization for compression of multichannel ECG," *IEEE Trans. Biomed. Eng.*, vol. 37, pp. 821–825, September 1990.
- [58] J.-J. Wei, C.-J. Chang, N.-K. Chou, and G.-J. Jan, "ECG data compression using truncated singular value decomposition," *IEEE Trans. Info. Tech. Biomed.*, vol. 5, pp. 290–299, December 2001.
- [59] A. G. Ramakrishnan and S. Saha, "ECG coding by wavelet-based linear prediction," *IEEE Trans. Biomed. Eng.*, vol. 44, no. 12, pp. 1253–1261, 1997.
- [60] B. Bradie, "Wavelet-packet based compression of single lead ECG," *IEEE Trans. Biomed. Eng.*, vol. 43, no. 5, pp. 493–501, 1996.
- [61] I. Provaznik and J. Kozumplik, "Wavelet transform in electrocardiography-data compression," *Int. J. Med. Inf.*, vol. 45, pp. 111–128, 1997.
- [62] K. Nagarajan, E. Kresch, S. S. Rao, and Y. Kresh, "Constrained ECG compression using best adapted wavelet packet bases," *IEEE Signal Processing Lett.*, vol. 3, pp. 273–275, October 1996.
- [63] M. L. Hilton, "Wavelet and wavelet packet compression of electrocardiograms," *IEEE Trans. Biomed. Eng.*, vol. 44, pp. 394–402, May 1997.
- [64] A. M. Hazem, "Time series analysis for ECG data compression," *Proc. Int. Conf. Acoust. Speech, Signal Processing, ICASSP '99*, vol. 3, pp. 1537–1540, 1999.

- [65] R. Nygaard, G. Melnikov, and A. K. Katsaggelos, "A rate distortion optimal ECG coding algorithm," *IEEE Trans. Biomedical Engg.*, vol. 48, pp. 28–39, January 2001.
- [66] Y. Zigel, A. Cohen, and A. Katz, "ECG signal compression using analysis by synthesis coding," *IEEE Trans. Biomed. Eng.*, vol. 47, pp. 1308–1315, October 2000.
- [67] G. D. Barlas and E. S. Skordalakis, "A novel family of compression algorithms for ECG and other semiperiodical, one-dimensional, biomedical signals," *IEEE Trans. Biomedical Engg.*, vol. 43, pp. 820–828, August 1996.
- [68] R. D. Johnston, "Are speech recognisers still 98% accurate, or has the time come to repeal 'hyde's law'?", *BT Technol. J.*, vol. 14, pp. 165–176, January 1996.
- [69] I. Boyd, "Speech coding for telecommunications," in *Digital signal processing in telecommunications* (F. A. Westall and S. F. A. Ip, eds.), pp. 300–325, Chapman & Hall, 1993.
- [70] M. Delprat, A. Urie, and C. Evci, "Speech coding requirements from the perspective of the future mobile systems," *Proc. IEEE Workshop on Speech Coding for Telecommunications*, pp. 89–90, October 1993.
- [71] W. T. K. Wong, R. M. Mack, B. M. G. Cheetham, and X. Q. Sun, "Low rate speech coding for telecommunications," *BT Technol. J.*, vol. 14, January 1996.
- [72] R. J. McAulay and T. F. Quatieri, "Mid-rate coding based on a sinusoidal representation of speech," *Proc. Int. Conf. Acoust. Speech, Signal Processing, ICASSP '85*, pp. 945–948, 1995.
- [73] R. J. McAulay and T. F. Quatieri, "Low-rate speech coding based on the sinusoidal model," in *Advances in speech signal processing* (S. Furui and M. M. Sondhi, eds.), Marcel Dekker Inc., 1992.
- [74] D. W. Griffin and J. Lim, "A high quality 9.6 kbit/s speech coding system," *Proc. Int. Conf. Acoust. Speech, Signal Processing, ICASSP '86*, pp. 125–128, 1986.
- [75] J. C. Hardwick and J. Lim, "The application of the imbe speech coder to mobile communications," *Proc. Int. Conf. Acoust. Speech, Signal Processing, ICASSP '91*, pp. 249–252, 1991.
- [76] W. B. Kleijn, "Continuous representation in linear predictive coding," *Proc. Int. Conf. Acoust. Speech, Signal Processing, ICASSP '91*, vol. 1, pp. 201–204, 1991.

- [77] A. Grossman and J. Morlet, "Decompositions of hardy functions into square integrable wavelets of constant shape," *SIAM Journal of Mathematical Analysis*, vol. 15, pp. 723-736, July 1984.
- [78] D. J. Stone, "Wavelet estimation," *Proc. IEEE*, vol. 72, pp. 1394-1402, October 1984. Special issue on seismic signal processing.
- [79] Y. Meyer, "Wavelets: Algorithm and applications," *Philadelphia: SIAM*, 1993.
- [80] I. Daubechies, "Where do wavelets come from? a personal point of view," *Proc. IEEE*, vol. 84, pp. 510-513, April 1996. Special issue on wavelets.
- [81] C. S. Burrus, R. A. Gopinath, and H. Guo, *Introduction to Wavelets and Wavelet Transforms: A Primer*. NJ: Prentice-Hall, 1998.
- [82] R. M. Rao and A. S. Bopardikar, *Wavelet Transforms: Introduction to Theory and Applications*. India: Addison-Wesley, 2000.
- [83] J. C. Goswami and A. K. Chan, *Fundamentals of Wavelets: Theory, Algorithms, and Applications*. NY: John Wiley & Sons, Inc., 1999.
- [84] A. Aldroubi and M. Unser, *Wavelets in Medicine and Biology*. NY: CRC Press, Inc., 1996.
- [85] M. W. Frazier, *An Introduction to Wavelets Through Linear Algebra*. NY: Springer-Verlag, 1999.
- [86] B. C. Bromley, "The large-scale distribution of matter in the universe: A multiresolution analysis," *Proc. IEEE Int. Symp. Signal Processing*, 1992.
- [87] H. Weng and K. M. Lau, "Wavelets, period doubling, and time-frequency localization with application to organization of convection over the tropical western pacific," *J. The Atmospheric Sciences, American Meteorological Society*, vol. 51, pp. 2523-2541, September 1994.
- [88] R. E. Larned, W. C. Karl, and A. S. Willsky, "Wavelet packet based transient signal classification," in *Proc. IEEE-Signal Processing International Symposium*, pp. 109-112, 1992.
- [89] J. Liang and T. W. Parks, "A translation invariant wavelet representation algorithm with applications," *IEEE Trans. Signal Processing*, vol. 44, February 1996.
- [90] M. Farge, "Wavelet transforms and their applications to turbulence," *Ann. Rev. Fluid Mech.*, vol. 24, pp. 395-457, 1992.

- [91] *Proc. IEEE*, vol. 72(10), October 1984. Special issue on seismic signal processing.
- [92] M. Unser and A. Aldroubi, "A review of wavelets in biomedical applications," *Proc. IEEE*, vol. 84, pp. 626–638, April 1996.
- [93] F. Tuteur, "Wavelet transformations in signal detection," in *Wavelets: Time-Frequency Methods and Phase Space* (J. Combes, A. Grossman, and P. Tchamitchian, eds.), (New York), pp. 132–137, Springer-Verlag, 1989.
- [94] L. Khadra, H. Dickhaus, and A. Lipp, "Representations of ECG-late potentials in the time frequency plane," *J. Med. Eng. and Technol.*, vol. 17, no. 6, pp. 228–231, 1993.
- [95] H. Dickhaus, L. Khadra, and J. Brachmann, "Time-frequency analysis of ventricular late potentials," *Methods of Inform. in Med.*, vol. 33, no. 2, pp. 187–192, 1994.
- [96] O. Meste, H. Rix, P. Caminal, and N. Thakor, "Ventricular late potentials characterization in time-frequency domain by means of a wavelet transform," *IEEE Trans. Biomed. Eng.*, vol. 41, pp. 625–634, July 1994.
- [97] L. Senhadji, G. Carrault, J. J. Bellanger, and G. Passariello, "Comparing wavelet transforms for recognizing cardiac patterns," *IEEE Eng. in Med. and Biol. Mag.*, vol. 14, no. 2, pp. 167–173, 1995.
- [98] S. Thurner, M. C. Feurstein, and M. C. Teich, "Multiresolution wavelet analysis of heartbeat intervals discriminates healthy patients from those with cardiac pathology," *Phys. Rev. Lett.*, vol. 80, pp. 1544–1547, 1998.
- [99] M. C. Teich, "Multiresolution wavelet analysis of heart-rate variability for heart-failure and heart-transplant patients," *Proc. IEEE EMBS Meeting*, 1998.
- [100] M. Akay et.al., "Carotid-cardiac interaction: heart rate variability during the unlocking of the carotid artery," *Adv. Exp. Med. Biol.*, vol. 346, no. 72, pp. 365–372, 1993.
- [101] H. Tsuji and H. Mori, "New analysis of hrv through wavelet transform," *Int. J. Human-Computer Interaction*, vol. 6, no. 2, pp. 205–217, 1994.
- [102] I. Provaznik, J. Kozumplik, J. Bardonova, M. Novakova, and Z. Novakova, "Wavelet transform in ECG signal processing." Presented in the EuroConference BIOSIGNAL 2000, Brno, Czech Republic, 2000.

- [103] C. Heneghan, S. B. Lowen, and M. C. Teich, "Analysis of spectral and wavelet-based measures used to assess cardiac pathology," *Proc. Int. Conf. Acoust. Speech, Signal Processing, ICASSP '99*, vol. 3, pp. 1393-1396, 1999.
- [104] A. H. Tewfik and M. Kim, "Correlation structure of the discrete wavelet coefficients of fractional brownian motion," *IEEE Trans. Inform. Theory*, vol. 38, pp. 904-909, 1992.
- [105] P. S. Addison, J. N. Watson, G. R. Clegg, P. A. Steen, and C. E. Robertson, "Finding coordinated atrial activity during ventricular fibrillation using wavelet decomposition," *IEEE Eng. in Med. and Biol. Mag.*, pp. 58-61 & 65, January/February 2002.
- [106] D. L. Donoho, "Denoising by soft thresholding," *IEEE Trans. Info. Theory*, vol. 41, pp. 251-266, July 1995.
- [107] J. W. Seok and K. S. Bae, "Speech enhancement with reduction of noise components in the wavelet domain," *Proc. Int. Conf. Acoust. Speech, Signal Processing, ICASSP 97*, vol. 2, pp. 1323-1326, 1997.
- [108] E. Ambikairajah, G. Tattersal, and A. Davis, "Wavelet transform-based speech enhancement," *Proc. ICSLP*, 1998.
- [109] H. Sheikhzadeh and H. Abutalebi Reza, "An improved wavelet-based speech enhancement system," in *Proc. 7th European Conf. on speech communication and technology, EUROSPEECH2001*, (Scandinavia, Aalborg, Denmark), September, 3-7 2001.
- [110] W. J. Pielemeier, G. H. Wakefield, and M. H. Simoni, "Time-frequency analysis of musical signals," *Proc. IEEE*, vol. 84, pp. 1216-1230, September 1996.
- [111] L. Cohen, "Time-frequency distributions—a review," *Proc. IEEE*, vol. 77, pp. 941-981, July 1989.
- [112] S. Mariuz, "Separation of the noisy component in the musical sounds in the time-frequency domain," in *Proc. Int. Conf. Signal Processing Applications and Technology*, vol. I, (Boston, USA), pp. 360-364, 7-12 October 1996.
- [113] M. Goodwin and M. Vetterli, "Time-frequency signal models for music analysis, transformation, and synthesis," 1996.
- [114] M. Lang, H. Guo, J. E. Odegard, C. S. Burrus, and R. O. Wells Jr., "Nonlinear processing of a shift invariant DWT for noise reduction," in *Proc. SPIE Conf.*

- Wavelet Applications II* (S. H. H., ed.), vol. 2491, (Orlando, FL, Bellingham, Washington), pp. 640–651, April 1995.
- [115] S. D. Marco and J. Weiss, “Improved transient signal detection using a wavepacket-based detector with an extended translation-invariant wavelet transform,” *IEEE Trans. Signal Processing*, vol. 45, pp. 841–850, April 1997.
- [116] J.-C. Pesquet, H. Krim, and H. Carfantan, “Time-invariant orthonormal wavelet representations,” *IEEE Trans. Signal Processing*, vol. 44, pp. 1964–1970, August 1996.
- [117] V. P. Devassia, G. Sandeep, N. S. Antony, and T. Thomas, “Audio compression using wavelets and wavelet packets,” *AMSE Journal on Signal Processing and Pattern Recognition*, vol. 46, no. 3, pp. 61–72, 2003.
- [118] D. Sinha and A. H. Tewfik, “Low bit rate transparent audio compression using adapted wavelets,” *IEEE Trans. Signal Processing*, vol. 41, pp. 3463–3479, December 1993.
- [119] T. Painter and A. Spanias, “Perceptual coding of digital audio,” *Proc. IEEE*, vol. 88, pp. 449–513, April 2000.
- [120] P. S. Sathidevi and Y. Venkataramani, “An analysis-synthesis method for transparent speech and audio compression using wavelet transforms and wavelet packets,” in *Pro. Int. Conf. Knowledge-Based Computer Systems*, (Mumbai, India), KBCS2000, December, 17–19 2000.
- [121] A. Ramasubramanian, K. L. Payton, and A. H. Costa, “A comparison between conventional and wavelet based amplitude compression schemes,” in *Proc. IEEE-SP Int. Symp. on Time-Frequency and Time-Scale Analysis*, (Pittsburgh, PA), pp. 161–164, October 1998.
- [122] T. Trinkaas and M. A. Clements, “An algorithm for compression of wideband diverse speech and audio signals,” in *Proc. Int. Conf. Acoust. Speech, Signal Processing, ICASSP 99*, pp. 901–904, 1999.
- [123] Y. Li, R. wei Dai, J. jian Feng, and Y. gui Xie, “Wideband speech compression using celp and wavelet transform,” in *Proc. ICSP’96*, pp. 706–709, 1996.
- [124] N. R. Chong-White and I. S. Burnett, “Accurate, critically sampled characteristic waveform surface construction for waveform interpolation decomposition,” *Electronics Letters*, vol. 36, pp. 1245–1247, July 2000.

- [125] J. Haagen and W. B. Kleijn, "Waveform interpolation," in *Modern methods of speech processing* (R. Ramachandran and R. Mamm, eds.), Kluwer Academic Publishers, 1995.
- [126] S.-G. Miaou and C.-L. Lin, "A quality-on-demand algorithm for wavelet-based compression of electrocardiogram signals," *IEEE Trans. Biomedical Engg.*, vol. 49, pp. 233-239, March 2002.
- [127] Z. Lu, D. Y. Kim, and W. A. Pearlman, "Wavelet compression of ECG signals by the set partitioning in hierarchical trees (SPIHT) algorithm," *IEEE Trans. Biomedical Engg.*, 2000.
- [128] R. S. H. Istepanian, L. J. Hadjileontiadis, and S. M. Panas, "ECG data compression using wavelets and higher order statistics," *IEEE Trans. Inf. Tech. Biomed.*, vol. 5, pp. 108-115, June 2001.
- [129] H. Lee and K. M. Buckley, "ECG data compression using cut and align beats approach and 2-d transforms," *IEEE Trans. Biomed. Engg.*, vol. 46, no. 5, pp. 556-564, 1999.
- [130] P. S. Hamilton and W. J. Tompkins, "Compression of ambulatory ECG by average beat subtraction and residual differencing," *IEEE Trans. Biomed. Engg.*, vol. 38, pp. 253-259, 1991.
- [131] N. V. Thakor, Y. Sun, H. Rix, and P. Caminal, "Multiwave: A wavelet-based ECG data compression algorithm," *IEICE Trans. Inform. Syst.*, vol. E76-D, pp. 1462-1469, 1993.
- [132] B. A. Rajoub, "An efficient coding algorithm for the compression of ECG signals using the wavelet transform," *IEEE Trans. Biomed. Engg.*, vol. 49, no. 4, pp. 355-362, 2002.
- [133] S. D. Boland and M. Deriche, "Hybrid lpc and discrete wavelet transform audio coding with a novel bit allocation algorithm," *Proc. Int. Conf. Acoust. Speech, Signal Processing, ICASSP '98*, vol. 6, pp. 3657-3660, 1998.
- [134] W. Chang and C. Wang, "A masking-threshold adapted weighting filter for excitation search," *IEEE Trans. Signal and Audio Processing*, vol. 4, no. 2, pp. 124-132, 1996.
- [135] M. V. Mathews, J. E. Miller, and E. E. David Jr., "Pitch synchronous analysis of voiced sounds," *The Journal of the Acoustical Society of America*, vol. 33, February 1961.

- [136] G. Evangelista, "Comb and multiplexed wavelet transforms and their applications to signal processing," *IEEE Trans. Signal Processing*, vol. 42, February 1994.
- [137] Y. Medan and E. Yair, "Pitch synchronous spectral analysis scheme for voiced speech," *IEEE Trans. Acoustics, Speech, and Signal Processing*, vol. 37, September 1989.
- [138] W. B. Kleijn and J. Haagen, "Transformation and decomposition of the speech signal for coding," *IEEE Signal Processing Lett.*, vol. 1, pp. 136-138, September 1994.
- [139] N. R. Chong, I. S. Burnett, and J. F. Chicharo, "A new waveform interpolation coding scheme based on pitch synchronous wavelet transform decomposition," *IEEE Trans. Speech, Audio Processing*, vol. 8, pp. 345-348, May 2000.
- [140] W. B. Kleijn and J. Haagen, *Waveform interpolation for coding and synthesis. Speech Coding and Synthesis*, W. B. Kleijn and K. K. Paliwal Eds., Amsterdam, The Netherlands: Elsevier.
- [141] P. Polotti and G. Evangelista, "Analysis and synthesis of pseudo-periodic $1/f$ -like noise by means of wavelets with applications to digital audio," *Eurasip J. Applied Signal Processing*, vol. 2001, pp. 1-14, March 2001.
- [142] M. S. Keshner, " $1/f$ noise," *Proc. IEEE*, vol. 70, no. 3, pp. 212-218, 1982.
- [143] D. T. Gillespie, "The mathematics of brownian motion and johnson noise," *Amer. J. Phys.*, vol. 64, pp. 225-240, 1996.
- [144] W. Hess, *Pitch detection of speech signals*. New York: Springer-Verlag, 1983.
- [145] J. R. Deller Jr., J. G. Proakis, and J. H. L. Hansen, *Discrete-Time Processing of Speech Signals*. New York: Macmillan, 1993.
- [146] A. M. Noll, "Cepstrum pitch determination," *J. Acoust. Soc. Amer.*, vol. 41, no. 2, pp. 293-309, 1970.
- [147] J. D. Wise, J. R. Caprio, and T. W. Parks, "Maximum likelyhood pitch estimation," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-24, pp. 418-423, 1976.
- [148] M. M. Sondhi, "New methods of pitch extraction," *IEEE Trans. Audio Electroacoust.*, vol. AU-16, pp. 262-266, June 1968.

- [149] H. W. Strube, "Determination of the instant of glottal closure from the speech wave," *J. Acoust. Soc. Amer.*, vol. 56, no. 5, pp. 1625–1629, 1974.
- [150] Y. M. Cheng and D. O'Shaughnessy, "Automatic and reliable estimation of glottal closure instant and period," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 37, no. 12, pp. 1805–1815, 1989.
- [151] B. Gold and L. Rabiner, "Parallel processing techniques for estimating pitch periods of speech in the time domain," *J. Acou. Soc. Amer.*, vol. 46, no. 2, pp. 442–448, 1969.
- [152] R. J. McNab, S. A. Lloyd, and I. H. Witten, "Signal processing for melody transcription," in *Proc. 19th Australian Computer Science Conf.* (Melbourne, Australia), January 31–February 2 1996.
- [153] L. R. Rabiner, M. J. Cheng, A. E. Rosenberg, and C. A. McGonegal, "A comprehensive performance study of several pitch detection algorithms," in *IEEE Trans. On Acoust., Speech, Signal Processing*, vol. ASSP - 24, pp. 399–417, October 1976.
- [154] S. Kadambe and G. F. Boudreaux-Bartels, "Application of the wavelet transform for pitch detection of speech signals," *IEEE Trans. Info. Theory*, vol. 38, pp. 917–924, March 1992.
- [155] S. G. Mallat and S. Zhong, "Characterization of signals from multiscale edges," *IEEE Trans. Patt. Analy., Mach. Intell.*, vol. 14, pp. 710–732, July 1992.
- [156] C. Wendt and A. P. Petropulu, "Pitch determination and speech segmentation using the discrete wavelet transform," in *Proc. IEEE Int. Symp. Circuits, Systems, ISCAS*, vol. 2, (Atlanta), pp. 45–48, May 1996.
- [157] J. Fitch and W. Shabana, "A wavelet-based pitch detector for musical signals," in *2nd COST-G6 Int. Workshop on Digital Audio Effects, DAFx99*, (Trondheim, Norway), December 9–11 1999.
- [158] C. Roads, *The computer music tutorial*. Cambridge, USA: The MIT Press, 1995.
- [159] D. Limin and H. Ziqiang, "Determination of the instants of glottal closure from speech wave using wavelet transform," in *Proc. Int. Conf. Signal Processing Applications and Technology*, (Boston:USA), pp. 336–340, 7–12 October 1996.
- [160] C. Li, C. Zheng, and C. Tai, "Detection of ECG characteristic points using wavelet transforms," *IEEE Trans. Biomed. Eng.*, vol. 42, no. 1, pp. 21–28, 1995.

- [161] M. Merri, D. C. Farden, J. G. Mottley, and E. L. Titlebaum, "Sampling frequency of the electrocardiogram for spectral analysis of the heart rate variability," *IEEE Trans. Biomed. Eng.*, vol. 37, no. 1, pp. 99–106, 1990.
- [162] J. Pan and W. J. Tompkins, "A real-time QRS detection algorithm," *IEEE Trans. Biomed. Eng.*, vol. BME-32, pp. 230–236, 1985.
- [163] P. S. Hamilton and W. J. Tompkins, "Quantitative investigation of QRS detection rules using the mit/bih arrhythmia database," *IEEE Trans. Biomed. Eng.*, vol. BME-33, pp. 1157–1165, 1986.
- [164] V. X. Afonso, W. J. Tompkins, T. Q. Nguyen, and S. Luo, "ECG beat detection using filter banks," *IEEE Trans. Biomed. Eng.*, vol. 46, no. 2, pp. 192–202, 1999.
- [165] C. Li, C. Zheng, and C. Tai, "Detection of ECG characteristic points using wavelet transforms," *IEEE Trans. Biomed. Eng.*, vol. 42, no. 91, pp. 21–28, 1995.
- [166] S. Kadambe, R. Murray, and G. F. Boudreaux-Bartels, "Wavelet transform-based QRS complex detector," *IEEE Trans. Biomed. Eng.*, vol. 46, pp. 838–848, July 1999.
- [167] S. Osowski and T. H. Linh, "ECG beat recognition using fuzzy hybrid neural network," *IEEE Trans. Biomed. Eng.*, vol. 48, no. 11, pp. 1265–1271, 2001.
- [168] G. M. Friesen, T. C. Jannett, M. A. Jadallah, S. L. Yates, S. R. Quint, and T. H. Nagle, "A comparison of the noise sensitivity of nine QRS detection algorithms," *IEEE Trans. Biomed. Eng.*, vol. 37, no. 1, pp. 85–98, 1990.
- [169] S. Suppappola and Y. Sun, "Nonlinear transforms of ECG signals for digital QRS detection: A quantitative analysis," *IEEE Trans. Biomed. Eng.*, vol. 41, no. 4, pp. 397–400, 1994.
- [170] J. S. Sahambi, S. N. Tandon, and R. K. P. Bhatt, "Using wavelet transforms for ECG characterization," *IEEE Eng. Med. Biol.*, pp. 77–83, January/February 1997.
- [171] A. Ruha, S. Sallinen, and S. Nissilä, "A real-time microprocessor QRS detector system with a 1-ms timing accuracy for the measurement of ambulatory hrv," *IEEE Trans. Biomed. Eng.*, vol. 44, no. 3, pp. 159–167, 1997.
- [172] Y. T. Chan and K. C. Ho, "Filter design for CWT computation using the shensa algorithm," in *Proc. Int. Conf. Acoust. Speech, Signal Processing, ICASSP 99*, vol. 3, pp. 1321–1324, 1999.

- [173] P. J. Burt and E. H. Adelson, "The laplacian pyramid transforms for image coding," *IEEE Trans. Communications*, vol. 31, pp. 532-540, 1983.
- [174] W. Sweldens, "The lifting scheme: A custom-design construction of biorthogonal wavelets," *J. Appl. Comp. Harm. Anal.*, vol. 3, no. 2, pp. 186-200, 1996.
- [175] I. Daubechies and W. Sweldens, "Factoring wavelet transforms into lifting steps," tech. rep., Bell Laboratories, 1996.
- [176] R. L. Clypoole Jr. and R. G. Baraniuk, "Adaptive wavelet transforms via lifting," *Int. Conf. Acoust. Speech, Signal Processing, ICASSP 98*, vol. 3, pp. 1513-1516, 1998.
- [177] R. L. Clypoole, G. Davis, W. Sweldens, and R. G. Baraniuk, "Nonlinear wavelet tranforms for image coding," in *Proc. 31st Asilomar Conf.*, (Pacific Grove, CA), 1997.
- [178] I. Daubechies and W. Sweldens, "Factoring wavelet transform into lifting steps," *Journal of Fourier Anal. Appl.*, vol. 4, no. 3, pp. 247-269, 1998.
- [179] M. Vetterli and C. Herley, "Wavelets and filter banks, theory and design," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 40, no. 9, pp. 2207-2232, 1992.
- [180] M. Vetterli, "Running fir and iir filtering using multirate filter banks," *IEEE Trans. on Acoustic, Speech, Signal Processing*, vol. ASSP-36, pp. 730-738, May 1988.
- [181] M. Viswanath, R. M. Ovens, and M. J. Irwin, "Vlsi architectures for the discrete wavelet transform," *IEEE Trans. Circuits Syst. II*, vol. 42, pp. 305-316, 1995.
- [182] H. Sava, M. Fleuy, A. C. Downton, and A. F. Clark, "Parallel pipeline implementation of wavelet transform," *IEE Proc. Part I (Vision, image, and Signal processing)*, vol. 144, pp. 355-359, December 1997.
- [183] H. Sava, M. Fleuy, A. C. Downton, and A. F. Clark, "Fast implementation of discrete wavelet transform based on pipeline processor farming," in *Proc. IPA '97, IEE 6th Int. Conf. on Image Processing and its Applications*, pp. 171-173, 1997.
- [184] M. A. Westenburg and J. B. T. M. Roerdink, "Frequency domain volume rendering by x-ray wavelet transform," *IEEE Trans. Image Processing*, vol. 9, pp. 1249-1261, July 2000.

- [185] T. C. Denk and K. K. Parhi, "Vlsi architectures for lattice structure based orthonormal discrete wavelet transforms," *IEEE Trans. Circuits and Systems-II, Analog and Digital Signal Processing*, vol. 44, February 1997.
- [186] S. K. Chandran and K. M. M. Prabhu, "Computation of DWT via FHT-based implementation," *Electronics Letters*, vol. 32, pp. 1437-1438, August 1996.
- [187] F. Kossentini, "Spatially segmented wavelet transform," *Tech Report, UBC*, 1998.
- [188] A. Ortega, W. Jiang, P. Fernandez, and C. Chrysafis, "Implementations of the discrete wavelet transform: Complexity, memory and parallelization issues," in *Proc. SPIE in Wavelet Applications in Signal and Image Processing VII*, Denver, Co., July 1999.
- [189] R. N. Strickland and H. I. Hahn, "Wavelet transforms for detecting micro calcifications in mammograms," *IEEE Trans. Medical Imaging*, vol. 15, pp. 218-229, April 1996.
- [190] D. Krishnaswamy and M. Orchard, "Parallel algorithms for the two-dimensional discrete wavelet transform," in *International Parallel Processing Symposium, IPPS94*, vol. 3, (IL), pp. 47-54, 1994.
- [191] L. Yang and M. Misra, "Course-grained parallel algorithms for multi-dimensional wavelet transforms," *The J. Super computing*, vol. 11, pp. 1-22, 1997.
- [192] J. Fridman and E. Manolakos, "On the scalability of the 2d discrete wavelet transform algorithms," in *Multidimensional Systems and Signal Processing*, vol. 8, pp. 185-217, Boston, MA: Kluwer, 1997.
- [193] A. Mertins, *Signal Analysis Wavelets, Filter banks, Time-Frequency Transforms and Applications*. New York: Wiley, 1999.
- [194] J. N. Bradley, C. M. Brislawn, and V. Faber, "Reflected boundary conditions for multirate filter banks," *Proc. IEEE Int. Symp. Time-Frequency and Time-Scale Analysis*, 1992.
- [195] S. Muramatsu and H. Kiya, "Extended overlap-add and-save method for multi-rate signal processing," *IEEE Trans. Signal Processing*, vol. 45, pp. 2376-2380, September 1997.
- [196] J. H. Nealand, A. B. Bradley, and M. Lech, "Overlap-save convolution applied to wavelet analysis," *IEEE Signal Processing Letters*, vol. 10, pp. 47-49, February 2003.

- [197] B. Leslie and M. Sandler, "A wavelet packet algorithm for 1-d data with no block end effects," *Proc. IEEE Symp. Circuits and Systems*, 1999.
- [198] M. G. Mini, V. P. Devassia, and T. Thomas, "Block discrete wavelet transform for real-time processing of one-dimensional signals," *Proc. Int. Conf. Business, Decision & Technology, BDT'02, Kuwait*, vol. 1, pp. 155–162, March, 18–20 2002.
- [199] M. G. Mini, V. P. Devassia, and T. Thomas, "In-place computation of block discrete wavelet transform," *IETE Journal of Education*, vol. 43, pp. 79–82, April–June 2002.
- [200] H. Guo, "Mapped inverse discrete wavelet transform for data compression," *Proc. Int. Conf. Acoust., Speech, Signal Processing, Int. Conf. Acoust. Speech, Signal Processing, ICASSP 98*, vol. 3, pp. 1385–1388, 1998.
- [201] C. Taswell and K. C. McGrill, "Wavelet transform algorithms for finite-duration discrete-time signals," *ACM Transactions on Mathematical software*, vol. 20, pp. 398–412, September 1994.
- [202] J. Lu, "Parallelizing mallat algorithm for 2d wavelet transforms," *Information Processing Letters*, vol. 45, pp. 255–259, April 1993.
- [203] A. Khokhar, G. Heber, P. Thulasiraman, and G. Gao, "Load adaptive algorithms and implementations for the 2d discrete wavelet transform on fine-grain multi-threaded architectures," in *13th International Parallel Processing Symposium & 10th Symposium on Parallel and Distributed Processing*, (San Juan, Puerto Rico), 1999.
- [204] F. Marino, V. Piuri, and E. E. Swartzlander Jr., "A parallel implementation of the 2-d discrete wavelet transform without interprocessor communications," *IEEE Trans. Signal Processing*, vol. 47, pp. 3179–3184, November 1999.
- [205] A. Bruce and H. Y. Gao, *Applied Wavelet Analysis with S-PLUS*. NY: Springer-Verlag, 1996.
- [206] A. Haar, "Zur theorie der orthogonalen funktionen-systeme," *Math. Ann.*, vol. 69, pp. 331–371, 1910.
- [207] G. Evangelista, "Orthogonal wavelet transforms and filter banks." presented at IEEE 23rd Asilomar Conf. Circuits, Syst., Computers, November 1989.
- [208] M. Vetterly and C. Herley, "Wavelets and filterbanks: Relationships and new results," in *Proc. 1990 IEEE Int. Conf. Acoust., Speech, Signal Processing*, (Albuquerque, NM), pp. 1723–1726, April 3–6 1990.

- [209] I. Daubechies, "Orthonormal bases of compactly supported wavelets," *Comm. Pure and Applied Math.*, vol. XLI, no. 7, pp. 909–996, 1988.
- [210] G. Evangelista and C. W. Barnes, "Discrete-time wavelet transforms and their generalizations," in *Proc. Int. Symp. Circuits Syst.*, (New Orleans, LA), May 1990.
- [211] O. Rioul, "Regular wavelets: a discrete-time approach," *IEEE Trans. Signal Processing*, vol. 41, pp. 3572–3579, December 1993.
- [212] O. Rioul, "A discrete-time multiresolution theory," *IEEE Trans. Signal Processing*, vol. 41, pp. 2591–2606, August 1993.
- [213] R. Coifman, Y. Meyer, S. Quake, and V. Wickerhauser, "Signal processing and compression with wavelet packets," technical report, Yale University, 1990.
- [214] T. Thomas, "A modified method for the classification and detection of voiced, unvoiced, silent and transition regions in speech signals," *AMSE Journal on Signal Processing and Pattern Recognition*, vol. 46, no. 3, pp. 1–8, 2003.
- [215] W. B. Kendall, "A new algorithm for computing autocorrelations," *IEEE Trans. Computers*, vol. C-23, pp. 90–93, January 1974.
- [216] J. Makhoul, "Linear prediction: A tutorial review," in *Proc. IEEE*, vol. 63, pp. 561–580, 1975.
- [217] P. P. Vaidyanathan, *Multirate Systems and Filter Banks*. Englewood Cliffs, NJ: Prentice Hall, 1993.
- [218] W. R. Daumer, "Subjective evaluation of several efficient speech coders," *IEEE Trans. Commun.*, vol. COM-30, pp. 655–662, April 1982.
- [219] M. Kay, "Wavelets in biomedical engineering," *Annals of Biomed. Engg.*, vol. 23, pp. 531–542, 1995.
- [220] D. A. Peredina and A. Allen, "Telemedicine technology and clinical applications," *Jour. Am. Med. Assoc.*, vol. 273, pp. 483–488, 1995.
- [221] H. Yoshida, W. Zhang, W. Cai, K. Doi, R. M. Nishikawa, and M. L. Giger, "Optimizing wavelet transform based supervised learning for detection of micro calcification in digitized mammograms," *IEEE Proc. Int. Conf. On Image Processing, ICIP-95*, pp. 152–155, 1995.

- [222] B. C. Branley, "The large scale distribution of matter in the universe: A multiresolution analysis," *Proc. of the 1992 IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis*, pp. 201–204, October 4–6 1992.
- [223] T. Thomas and C. S. Sridhar, "A new variable frame length codec (vflc) for speech processing," *J. Acoust. Soc. India*, vol. 23, no. 1, pp. 55–60, 1995.
- [224] N. S. Jayant and P. Noll, *Digital Coding of Waveform: Principles and Applications to Speech and Video*. New Jersey: Prentice Hall Inc., 1984.
- [225] T. Chang and C. J. Kuo, "Texture segmentation with tree-structured wavelet transforms," in *Proc. IEEE-SP Int. Symp.*, pp. 543–546, 1992.
- [226] B. G. Sherlock and D. M. Monro, "Optimized wavelets for finger print compression," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Processing*, vol. 3, pp. 1447–1450, May 1996.
- [227] J. Suckling et al., "The mammographic image analysis society digital mammogram database," *Excerpta Medica. International Congress Series*, vol. 1069, pp. 375–378, 1994.
- [228] W. A. Ainsworth, *Mechanisms of Speech Recognition*. Pergamon Press, 1976.
- [229] P. Philippe, F. M. de Saint-Martin, and M. Lever, "Wavelet packet filter banks for low time delay audio coding," *IEEE Trans. Speech and audio Processing*, vol. 7, pp. 310–322, May 1999.
- [230] B. S. Atal and L. R. Rabiner, "A pattern recognition approach to voiced-unvoiced-silence classification with applications to speech recognition," *IEEE Trans. On Acoust. Speech, Signal Processing*, vol. 24, pp. 201–212, 1976.
- [231] L. R. Rabiner and M. R. Sambur, "Application of lpc to the voiced-unvoiced-silence detection problem," *IEEE Trans. On Acoust. Speech, Signal Processing*, vol. 27, pp. 338–343, 1979.
- [232] Y. Qi and B. Hunt, "Voiced-unvoiced-silence classification of speech using hybrid features & a network classifier," *IEEE Trans. On Speech, Audio Processing*, vol. 1, pp. 250–255, April 1993.
- [233] V. Ramamoorthy, "A simple time-domain algorithm for voiced-unvoiced detection, signal processing theories & applications," *EURASIP Journal of Signal Processing*, pp. 737–742, 1980.

-
- [234] O. S. Shiju, T. Thomas, and C. S. Sridhar, "Comparison of variable frame length codec (vflc) with similar low bit rate codecs," in *Communications '98* (U. Mukherji and K. V. S. Hari, eds.), pp. 210-214, New Delhi: Tata McGraw-Hill Pub. Co. Ltd, 1998.
- [235] M. J. Carey, E. S. Parris, and H. Lloyd-Thomas, "A comparison of features for speech, music discrimination," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing, Int. Conf. Acoust. Speech, Signal Processing, ICASSP 99*, (Phoenix, Arizona), pp. 149-152, March 15-19 1999.
- [236] B. Atal and S. Hanuer, "Speech analysis and synthesis by linear prediction of speech wave," *J. Acoust. Soc. Amer.*, vol. 50, pp. 637-655, August 1971.

List of Publications

International Journal/Conference Papers

1. **V.P. Devassia**, G. Sandeep, N. S. Antony, T. Thomas, "Audio Compression Using Wavelets and Wavelet Packets," *AMSE Journal on Signal Processing and Pattern Recognition*, vol. 46, no. 3, pp. 61-72, 2003.
2. **V.P. Devassia**, M.G. Mini, and T. Thomas, "A Novel Parallel Structure for Computation of Discrete Wavelet Transform of Images," *AMSE Journal on Computer Science and Statistics*, vol.7, no.3, pp.25-40, 2002.
3. **V.P. Devassia**, M.G. Mini, and T. Thomas, "An Efficient Algorithm for Time-Domain Computation of Discrete Wavelet Transform," Accepted in the *International Conference on Acoustics, Speech, and Signal Processing, ICASSP2001, USA*, May 7-11, 2001.
4. **V.P. Devassia**, M.G. Mini, and T. Thomas, "Feature Enhancement Using PSWT Technique: A Study on Instrumental Music," *Proceedings of the International Workshop on Linear Algebra, Numerical Functional Analysis, and Wavelet Analysis*, Ed. S.H. Kulkarni, M.N.N. Namboodiri, India, pp. 65-77, Aug. 6-15, 2001.
5. **V.P. Devassia**, M.G. Mini, and T. Thomas, "An Efficient Algorithm for Time-Domain Computation of Discrete Wavelet Transform," *Knowledge-Based Intelligent Information Engineering & Allied Technologies*, Ed. N. Baba, L.C. Jain, and R.J. Hewlett, IOS Press, Netherlands, pp.892-895, 2001.
6. **V.P. Devassia**, M.G. Mini, and T. Thomas, "Efficient Parallel Structure for Wavelet Transform Based Processing of Biomedical Images," Accepted in the *International Conference on Mathematical & Engineering Techniques in Medicine and Biological Sciences, METMBS'02, USA*, Jun. 24-27, 2002.
7. **V.P. Devassia**, M.G. Mini, and T. Thomas, "Wavelet Transform Based Classification of Vocal Music into Voiced, Unvoiced, Silent and Transition Regions," *Proceedings of the International Conference on Modelling & Simulation in Technical and Social Sciences, MS'02, Spain*, Jun. 25-27, 2002(In press).

8. **V.P. Devassia**, M.G. Mini, and T. Thomas, "PSWT Based Predictive Coding and Compression of ECG Signals, Accepted in the *International Conference on Mathematical & Engineering Techniques in Medicine and Biological Sciences*, METMBS'02, USA, Jun. 24-27, 2002.
9. **V.P. Devassia**, M.G. Mini, and T. Thomas, "A Novel Computational Structure for Real-Time Wavelet Analysis," Accepted in the *SPIE International Conference on Modeling, Signal Processing and Control*, SS03, San Diego, USA, vol 5049. pp. 661-669, March 2-6, 2003.
10. M.G. Mini, **V.P. Devassia**, and T. Thomas, "One-Dimensional Processing for Edge Detection in Biomedical Images Using Pitch-Synchronous Wavelet Transform," Accepted in the *International Conference on Mathematical & Engineering Techniques in Medicine and Biological Sciences*, METMBS'02, USA, Jun. 24-27, 2002.
11. M.G. Mini, **V.P. Devassia**, and T. Thomas, "Block Discrete Wavelet Transform for Real-Time Processing of One-Dimensional Signals," *Proceedings of the International Conference on Business, Decision & Technology*, BDT'02, Kuwait, Vol.1, pp.155-162, Mar. 18-20, 2002.
12. M.G. Mini, **V.P. Devassia**, and T. Thomas, "Detection of Microcalcifications in Digitized Mammogram using Wavelet Transform Local Extrema," Accepted in the *International Symposium for Computer Applications in Radiology*, SCAR, Boston, USA, June 7-10, 2003.
13. M. G. Mini, **V. P. Devassia**, and T. Thomas, "Low-memory Block Discrete Wavelet Transform for Image Processing", AMSE, Journal on Signal Processing and Pattern Recognition (In Press).
14. M. G. Mini, **V. P. Devassia**, and T. Thomas, "Block Discrete Wavelet Transform for real-time processing of one dimensional signals," AMSE Journal on Signal Processing and Pattern Recognition, (In Press).

National Journal/Conference Papers

15. **V.P. Devassia**, M.G. Mini, and T. Thomas, "Application of Pitch-Synchronous Wavelet Transform in the Analysis of Guitar Music," *J. Acous. Soc. Ind.*, vol.29, no.2-4, pp. 296-302, 2001.
16. **V.P. Devassia**, M.G. Mini, and T. Thomas, "Instrument-Dependent Noise Suppression using PSWT," *J. Acous. Soc. Ind.*, vol.30, no.1-2, pp. 164-167, 2002.

17. M.G. Mini, **V.P. Devassia**, and T. Thomas, "Block Discrete Wavelet Transform for Real-Time Processing of Speech Signals," *Proceedings of the 8th National Conference on Communications*, NCC2002, Bombay, pp.374-378, Jan. 25-27, 2002.
18. M.G. Mini, **V.P. Devassia**, and T. Thomas, "In-Place Computation of Block Discrete Wavelet Transform," *IETE Journal of Education*, vol.43, no.2, pp.79-82, April-June 2002.
19. M.G. Mini, **V.P. Devassia**, and T. Thomas, "A Comparative Study on the Application of Wavelet Based Techniques for Microcalcification Detection in mammograms," Accepted in the 9th National Conference on Communications, NCC2003, Madras, pp.601-605, Jan.31-Feb.2, 2003.

Index

- Autocorrelation method, 65
- Case study
 - 2D edge detection, 145
 - Compression of ECG, 91
 - Compression of instrument tone, 105
 - Compression of vocal sound, 99
 - PSWT computation using PMS, 134
- Continuous wavelet transform, 42
- Discrete wavelet transform, 44
- DWT computation, 49, 58
 - Computational complexity, 127, 140
 - comparison, 130, 143
 - in PMS structure, 129, 142
 - in pyramidal structure, 127, 141
 - for 2D, 57
 - Issues in, 119
 - Mallat's algorithm, 9, 52
 - parallel, 9
 - PMS structure, 121
 - Pyramidal structure, 9, 52
- Linear prediction, 10, 85
- Linear predictive coding, 10, 68
- LPC, *see* Linear predictive coding
 - PSWT based, 75, 95
- MRA, *see* Multi-resolution analysis
- Multi-resolution analysis, 46
 - Discrete-, 48
- Multirate operation, 49
 - downsampling, 49
 - Sampling rate conversion, 51
 - upsampling, 50
- Normalization, 82
- Parallel multiple subsequence structure, 121
- Performance evaluation, 88
 - Compression ratio, 89
 - MOS, 88, 157
 - NMAE, 89
 - NRMSE, 89
 - seg SNR, 89
 - SNR, 88
- Period estimation, 64, 76
 - Autocorrelation method, 65, 76
 - UDWT method, 92, 115
- Pitch-synchronous wavelet transform, 61
- PMS structure, *see* Parallel multiple subsequence structure
 - for 1D, 121
 - for 2D, 137
- PS representation of signals, 82
- PSWT, *see* Pitch-synchronous wavelet transform
 - Computation of, 64, 85
 - Feature enhancement in, 109
- Signal compression, 10
 - LPC technique for, 11
- Signal processing, 4

G9020

- Signal segmentation, 161
- Undecimated discrete wavelet transform,
66
- Voice coder, 11
- Waveform coder, 11
- Wavelet, 41
 - Coiflet, 42
 - Daublet, 42
 - Haar, 42
 - Orthogonal-, 41
 - Symlet, 42
- Wavelet analysis, 5
- Wavelet packet transform, 54
- Wavelet transform
 - Continuous, 7, 42
 - Discrete, 7, 44
 - for Pseudo-Periodic signals, 9
 - Integral, 7
 - UDWT, 66
- Wavelet transform computation, 8

