

**PAPER MACHINE DATA ANALYSIS AND OPTIMIZATION USING  
WAVELETS**

By

Xuejun Jiao

B. E. (Electrical Engineering) Northeastern Heavy Mechanical Institute, P.R.China

M. E. (Electrical Engineering) Northeastern Heavy Mechanical Institute, P.R.China

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF APPLIED SCIENCE

in

THE FACULTY OF GRADUATE STUDIES

DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING

We accept this thesis as conforming  
to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA

January 1999

© Xuejun Jiao, 1999

In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the head of my department or by his or her representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Department of Electrical and Computer Engineering  
The University of British Columbia  
2356 Main Mall  
Vancouver, BC  
Canada V6T 1Z4

Date: January 1999

## **Abstract**

This thesis describes paper machine data analysis methods using wavelet and wavelet packets and their applications, aiming at improving paper machine efficiency.

First, the validity and accuracy of the wavelet transform are confirmed by applying discrete wavelet transform to paper samples using both the paper machine on-line scanner data and off-line analyzer data. Results show that the wavelet transform can represent paper machine process data economically without loss of detail, and that it can also provide excellent visualization to the operator.

Process monitoring and control performance assessment are then studied. By separating controllable and uncontrollable variations in the cross machine direction profile, the achieved performance and the best possible performance of the system are evaluated. A CD performance index can be calculated on-line, providing the operator with a quick assessment of the control system performance. Both wavelet and wavelet packets are used and the results are compared.

Finally, the processed paper machine profiles obtained through wavelet and wavelet packet analysis are used for trim-loss optimization taking into account paper quality. Three different optimization schemes are compared and the potential savings through trim-loss optimization before and after improving control are analyzed.

## **Acknowledgment**

Sincere appreciations to my supervisors, Prof. Michael S. Davies and Prof. Guy A. Dumont, for their kind guidance and support during the course of my thesis work.

I would also like to thank Zoran Nesic for his consistent help on my research project.

Thanks to all my friends at Pulp and Paper Centre and at the Department of Electrical and Computer Engineering for helping me and making my studies at UBC enjoyable.



## Table of Contents

<b>Abstract</b>	<b>ii</b>
<b>Acknowledgment</b>	<b>iii</b>
<b>List of Tables</b>	<b>vii</b>
<b>List of Figures</b>	<b>viii</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Overview of Estimation Theory . . . . .	3
1.2 Motivation . . . . .	3
1.3 Outline of Thesis . . . . .	4
<b>2 Wavelet and Wavelet Packet Theory</b>	<b>6</b>
2.1 Wavelet Analysis . . . . .	7
2.1.1 Wavelet Basis and Time-Frequency Resolution . . . . .	7
2.1.2 Wavelet Basis Properties . . . . .	9
2.1.3 Discrete and Inverse Discrete Wavelet Transform . . . . .	10
2.2 Wavelet Packet Analysis . . . . .	11
2.2.1 Wavelet Packets . . . . .	11
2.2.2 Discrete Wavelet Packet Transform(DWPT) . . . . .	12
2.2.3 Best Basis Algorithm . . . . .	15
2.3 Multi-Resolution Analysis . . . . .	16
2.4 De-Noising . . . . .	17

2.4.1	Thresholding . . . . .	18
2.4.2	Threshold Selection Methods . . . . .	20
2.5	Compression . . . . .	22
2.6	Two-Dimensional Signal Analysis . . . . .	23
<b>3</b>	<b>Comparison of On-machine and Off-machine Measurements of Paper Properties Using Wavelet Analysis</b>	<b>25</b>
3.1	Introduction . . . . .	25
3.2	On-line Scanner Data and Off-line Tapio Analyzer Data . . . . .	26
3.2.1	On-line Scanner Data . . . . .	26
3.2.2	Tapio Analyzer Data . . . . .	26
3.2.3	Transfer of Tapio Data . . . . .	27
3.3	Cross Machine Direction Analysis . . . . .	27
3.3.1	On-line Scanner Data Analysis . . . . .	28
3.3.2	Off-line Tapio Analyzer Data Analysis . . . . .	34
3.3.3	Comparison of On-line Data Analysis and Off-line Data Analysis .	35
3.4	Machine Direction Analysis . . . . .	36
<b>4</b>	<b>Wavelet and Wavelet Packet Analysis of Industrial Data</b>	<b>38</b>
4.1	CD Variation Separation Using Wavelet Analysis . . . . .	38
4.2	CD Variation Separation Using Wavelet Packets . . . . .	40
4.2.1	Frequency Order of Wavelet Packet Nodes . . . . .	41
4.2.2	Restricted Basis Algorithm for Tree Selection . . . . .	42
4.2.3	Profile Decomposition at Each Waveband . . . . .	45
4.2.4	Controllable and Uncontrollable CD Profiles . . . . .	47
4.3	Performance Assessment Using Wavelet and Wavelet Packet Analysis . .	51
4.4	Denoising . . . . .	55

4.5	Compression of Industrial Data . . . . .	55
<b>5</b>	<b>Trim-Loss Optimization</b>	<b>59</b>
5.1	Introduction . . . . .	59
5.2	Overview of Trim-Loss Optimization . . . . .	60
5.3	Mathematical Model for Trim-Loss Problem . . . . .	62
5.4	Lingo Optimization Solver . . . . .	64
5.5	Trim-Loss Optimization Using Visual Basic and Lingo . . . . .	66
5.6	Trim Optimization Schemes . . . . .	69
5.6.1	Scheme 1: Trim Optimization without Paper Quality Consideration	71
5.6.2	Scheme 2: Trim Optimization with Paper Quality Consideration .	73
5.6.3	Scheme 3: Trim Optimization after Improving Control . . . . .	74
5.6.4	Comparison and Conclusion . . . . .	76
<b>6</b>	<b>Conclusions</b>	<b>77</b>
6.1	Conclusions . . . . .	77
6.2	Further Work . . . . .	78
	<b>Bibliography</b>	<b>80</b>

## List of Tables

2.1	Minimax thresholds for various sample sizes . . . . .	20
4.2	Natural order and frequency order of wavelet packet nodes . . . . .	42
4.3	Summary of profile variation at each waveband . . . . .	45
4.4	Compression of industrial data . . . . .	58
5.5	Roll order information . . . . .	71
5.6	Trim optimization result: Scheme 1 . . . . .	71
5.7	Trim optimization result: Scheme 2 . . . . .	74
5.8	Trim optimization result: Scheme 3 . . . . .	75
5.9	Result comparison for three schemes . . . . .	76

## List of Figures

1.1	A simplified diagram of paper machine (J. Ghofraniha) . . . . .	1
1.2	Paper machine sensor path . . . . .	2
2.3	Time-frequency plane: (a)STFT (b)WT (c)WPT . . . . .	8
2.4	Block diagram of DWT and IDWT . . . . .	10
2.5	The Harr wavelet packets . . . . .	12
2.6	Filter bank implementation of DWPT . . . . .	13
2.7	Analysis of a chirp signal(a) using wavelets (b) and wavelet packets (c) .	14
2.8	Wavelet denoising . . . . .	18
2.9	Hard thresholding and soft thresholding . . . . .	19
2.10	Block diagram of compression . . . . .	22
2.11	Diagram of two-dimensional DWT . . . . .	23
2.12	Decomposition of signal using 2-dimensional DWT . . . . .	24
3.13	Original Tapio data . . . . .	28
3.14	Tapio data after transfer . . . . .	28
3.15	Raw profile . . . . .	29
3.16	On-line scanner data analysis using wavelet . . . . .	29
3.17	Multiresolution analysis and normalized wavelength . . . . .	30
3.18	MD scan average . . . . .	31
3.19	Raw profile . . . . .	32
3.20	CD approximation: level 2 . . . . .	32
3.21	CD residues: level 3 . . . . .	32

3.22	CD approximation: level 3 . . . . .	32
3.23	Raw profile . . . . .	33
3.24	CD approximation: level 2 . . . . .	33
3.25	CD residues: level 3 . . . . .	33
3.26	CD approximation: level 3 . . . . .	33
3.27	Wavelet approximation and detail: level 7, Tapio data . . . . .	34
3.28	Comparison of on-line scanner data and Tapio data . . . . .	35
3.29	Scanner data with CD profile removed . . . . .	36
3.30	Scan average and wavelet estimate of MD profile . . . . .	37
4.31	Wavelet decomposition tree and wavelength at each node . . . . .	39
4.32	Wavelet packet decomposition tree . . . . .	41
4.33	The db2 wavelet packets . . . . .	42
4.34	Wavelet packet decomposition tree and wavelength at each node . . . . .	44
4.35	Profile wavelet packet decomposition at each node . . . . .	46
4.36	Controllable profile estimates using wavelet and wavelet packets . . . . .	48
4.37	Uncontrollable profile estimates using wavelet and wavelet packets . . . . .	48
4.38	CD profile separation using wavelet and wavelet packets . . . . .	49
4.39	Controllable: Wavelet Packet . . . . .	50
4.40	Uncontrollable: Wavelet Packet . . . . .	50
4.41	Controllable: Wavelet . . . . .	50
4.42	Uncontrollable: Wavelet . . . . .	50
4.43	Performance index using wavelet and wavelet packets . . . . .	51
4.44	Performance index using wavelet and wavelet packets(Caliper) . . . . .	53
4.45	Wavelet decomposition tree and wavelength at each node(Caliper) . . . . .	54
4.46	Wavelet packet decomposition tree(Caliper) . . . . .	54

4.47	Wavelet coefficients before and after thresholding . . . . .	56
4.48	Scan 128: Energy distribution . . . . .	57
5.49	An example with 3 products ( $i=1,2,3$ ) and 3 cutting patterns ( $j=1,2,3$ ) .	63
5.50	Trim optimization diagram . . . . .	66
5.51	Trim optimization interface using Visual Basic . . . . .	67
5.52	Three trim optimization schemes . . . . .	70
5.53	Trim optimization: Scheme 1 . . . . .	72
5.54	Trim optimization: Scheme 2 . . . . .	73
5.55	Trim optimization: Scheme 3 . . . . .	75

## Chapter 1

### Introduction

The paper machine is the final stage of paper manufacturing after the fibre pulping and bleaching process. In the headbox, the fibres and white water are mixed and are delivered into the paper machine. Moisture is then removed through drainage, mechanical pressing and drying. Finally a sheet of paper is produced at the reel. A simplified diagram of paper machine is given in Figure 1.1.

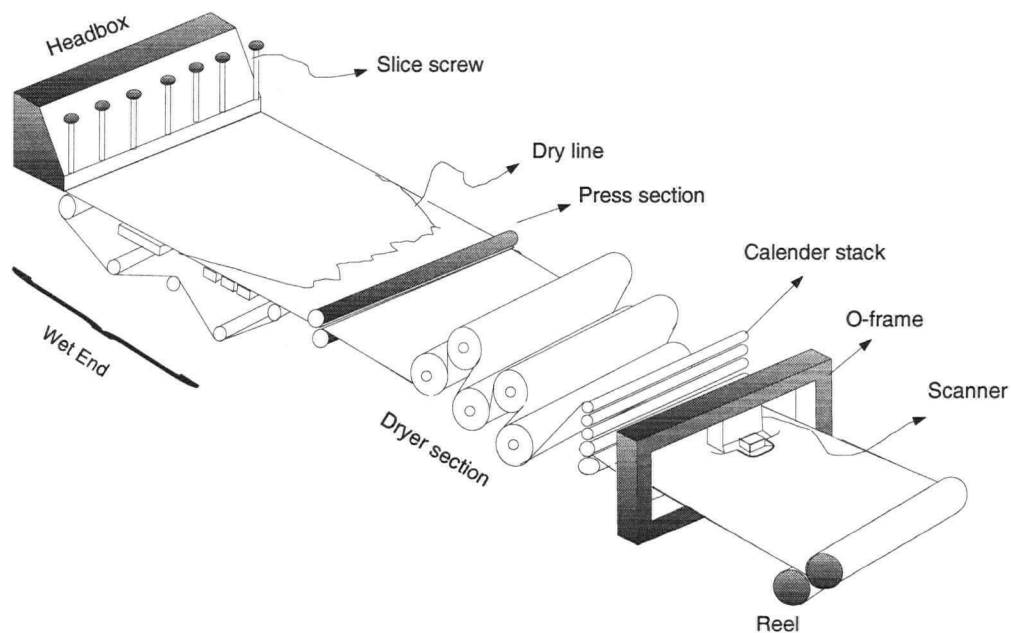


Figure 1.1: A simplified diagram of paper machine (J. Ghofraniha)

Hundreds of functional control loops are active to ensure the uniformity and quality of the paper produced. Among these, the basis weight and moisture control loops are



the most important.

A traversing sensor mounted on an O-frame at the dry end of the machine measures the sheet properties including basis weight, moisture etc. The sensor takes up to 3000 uniformly spaced measurements across the sheet during each traverse. As a result of sheet movement in machine direction (MD) and sensor moving in cross machine direction (CD), a zigzag pattern of measurements shown in Figure 1.2 is formed. Note that the machine direction speed is much greater than the sensor travel rate. The measured sequences of values thus contain both information about CD and MD variations. The MD variation is introduced by pressure and consistency variation and is considered to be fast and time-dependent. The CD variation is considered to be relatively time-invariant or slowly time-varying. For the purpose of CD control, the CD profile is extracted from the raw measurements and used to manipulate actuators distributed across the machine to achieve a uniform distribution of sheet properties.

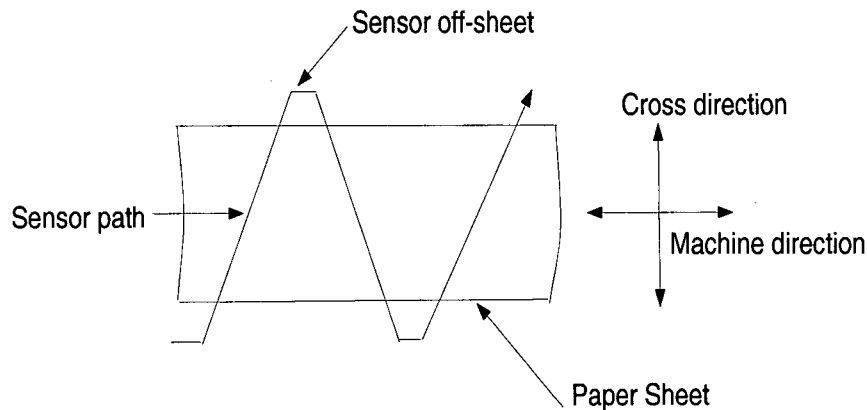


Figure 1.2: Paper machine sensor path

## 1.1 Overview of Estimation Theory

Effective estimation is important for the success of a proper control scheme in paper mills. Since automatic control was introduced on the paper machine, many estimation methods have been developed to process the scanner data. A commonly used technique is the exponential filtering (EXPO) technique [7]. The MD profile is extracted at the end of each scan and defined as the mean value of that scan. This estimator is usually slow and does not separate MD and CD variations optimally.

More advanced filtering techniques [26, 3, 14, 34, 23] use stochastic models for the profile variation. In one particular method [34, 25], the CD profile is estimated with a modified least-squares estimator, and the MD profile is estimated using a Kalman filter. Because the MD estimations are updated at each data point, this method results in an improved MD control bandwidth.

Recently, a new signal processing method – wavelet estimation has been used for paper machine data processing [27, 28, 4]. This method has been shown to be superior to previous estimation methods and has many advantages, such as small estimation error, fast computation, robustness in performance, better compression and better visualization.

One important application of wavelet multi-resolution analysis is the separation of the profile variations into different spatial components for performance assessment. Wavelet multi-resolution analysis of paper machine data has been used in [29]. In this thesis, results when using wavelet packets are compared to results of wavelet analysis.

## 1.2 Motivation

Advances in measurement and control systems have provided paper machines with high resolution profile data. The increased resolution potentially leads to overall improved control of paper machine. The operator benefits from a better presentation of process

data in order to detect and diagnose any change in the quality of the final product. Wavelet and wavelet packets filtering method can give a better and more clear picture of the profile data.

Quality control and production monitoring are increasingly important for paper companies. High speed sensors generate large amounts of data so that advanced data compression techniques become more and more important for data storage and data transfer. Profile variation separation and calculation of a performance index provide a quick and direct way of process assessment. As will be shown in this thesis, wavelet packet analysis provides a better separation of controllable profile variations and therefore a more accurate assessment of the control performance.

Process optimization is important to reduce the cost of raw fibre and energy in the modern paper industry. Trim optimization, which is studied in this thesis, minimizes the trim-loss when slitting a paper reel into individual rolls. Traditional trim optimization solving does not address the quality issue of the paper product. In this thesis, the profile data after wavelet filtering is used for trim optimization to maximize use of high quality product. Individual roll information following the slitter can also be displayed graphically.

### **1.3 Outline of Thesis**

A brief introduction of wavelet and wavelet packets theory and their application is given in Chapter 2. Chapter 3 focuses on assessing the validity and effectiveness of the wavelet transform by applying it to both paper machine on-line scanner data and off-line analyzer data. In Chapter 4, CD profile variation separation and performance assessment are compared by using wavelet and wavelet packet analysis respectively. In Chapter 5, the processed reel profile after denoising is further used for trim-loss optimization. Three different optimization schemes are discussed and tested. Finally conclusions and some

further remarks are given in Chapter 6.

## Chapter 2

### Wavelet and Wavelet Packet Theory

In recent years wavelet theory has found applications in different areas such as signal processing and statistical analysis. The basic idea of wavelet analysis dates back to work done by Littlewood-Paley in the 1930's or even earlier by A. Haar in the 1910's [18]. However it is only in 1982 that the wavelet was first proposed as a tool for signal analysis by Morlet. Later, the detailed mathematical theory of the continuous wavelet transform was carried out by Grossman and Morlet[17], followed by the detailed study on discrete wavelet transform by Daubechies, Grossmann and Meyer [9]. In 1988, Daubechies [8] gave a method for constructing compactly supported orthonormal wavelet basis functions from multi-resolution analysis, that attracted a lot of attention in many fields from theory to applications. Furthermore, wavelet packets were constructed by Coifman and Meyer in 1991 as a generalization of wavelets. More references on wavelet packets can be found in Coifman, Meyer and Wickenhausen [5, 30].

In this chapter, the basic theory of wavelet and wavelet packet theory, discrete wavelet transform and discrete wavelet packet transform are introduced followed by the description of denoising and compression schemes which are used in Chapter 3 and 4.

## 2.1 Wavelet Analysis

### 2.1.1 Wavelet Basis and Time-Frequency Resolution

The classical Fourier Transform is widely used for analyzing the frequency element of a signal by decomposing a signal into different frequencies of sine and cosine functions. However, the Fourier Transform is not suitable for dealing with non-stationary signals for which the frequency spectrum varies with time. To overcome this problem, Gabor proposed the Short-Time Fourier Transform (STFT) or the windowed Fourier Transform which is defined as following:

$$STFT(\tau, f) = \int x(t)g^*(t - \tau)e^{-2j\pi ft} dt \quad (2.1)$$

The STFT is popular in time-varying signal analysis because it maps a time-domain function to a time-frequency function  $F(t, \omega)$ , therefore allowing changes in spectrum with time to be traced.

Because the time window size is fixed for STFT, once a window is chosen, the resolution in both time and frequency are fixed. By comparison, the Wavelet Transform uses a shorter timing window at higher frequencies and a longer window at lower frequencies.

Wavelet basis functions are created by scaling and translating the same prototype  $\Psi(x)$  which is known as *mother wavelet*. Scaling corresponds to stretching or compressing the mother wavelet to generate new basis functions.

$$\Psi_{j,k}(x) = 2^{-j/2}\Psi(2^{-j}x - k) \quad (2.2)$$

where  $j$  is the scaling factor,  $k$  is the shift factor.

The time and frequency resolutions are defined as follows:

$$\Delta t = \sqrt{\frac{\int_{-\infty}^{\infty} t^2 |g^2(t)| dt}{\int_{-\infty}^{\infty} |g^2(t)| dt}} \quad (2.3)$$

$$\Delta\omega = \sqrt{\frac{\int_{-\infty}^{\infty} \omega^2 |G^2(\omega)| d\omega}{\int_{-\infty}^{\infty} |G^2(\omega)| d\omega}} \quad (2.4)$$

$g(t)$  and  $G(\omega)$  are the basis function and its Fourier Transform respectively.

According to Heisenberg inequality, the time and frequency resolution cannot be controlled independently. That is

$$\Delta\omega\Delta t \geq \frac{1}{2} \quad (2.5)$$

Figure 2.3 shows the time-frequency plane for STFT, wavelet transform(WT) and wavelet packet transform(WPT). From Figure 2.3, it can be seen that the resolution

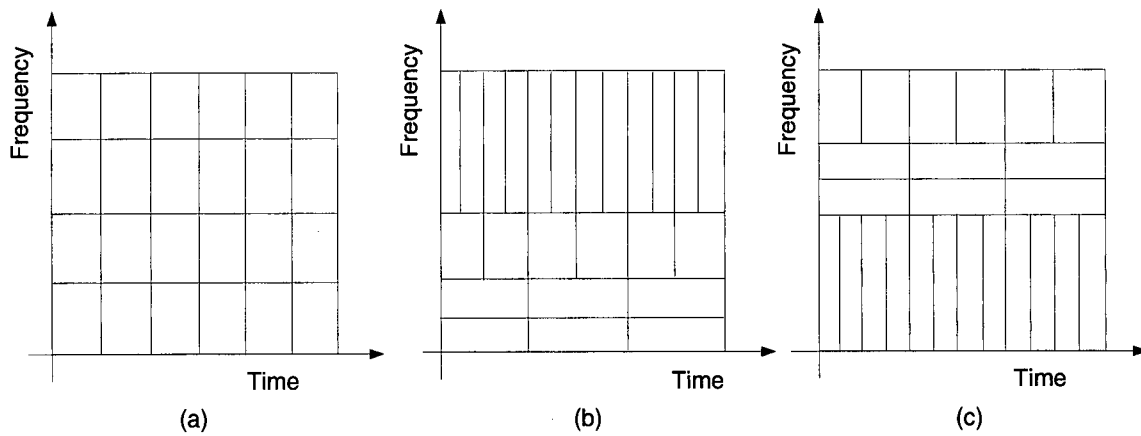


Figure 2.3: Time-frequency plane: (a)STFT (b)WT (c)WPT

$\Delta t$  and  $\Delta\omega$  are fixed for STFT, however they vary in the time-frequency plane for the wavelet transform, which leads to a coarser resolution at lower frequency and finer time resolution at higher frequency. Furthermore, wavelet packets (to be discussed later) offer more flexibility because they allow  $\Delta t$  and  $\Delta\omega$  to change in a signal decomposition.

### 2.1.2 Wavelet Basis Properties

Wavelet basis functions are created by scaling and translating a mother wavelet. The choice of mother wavelet depends on the wavelet properties and should be fit to the particular problem. Following are some important properties of wavelets, more details on wavelet properties can be found in [21].

*Compact Support* If the scaling function and wavelet are compactly supported, the wavelet filters are finite impulse response filters, as a result the fast wavelet transform is finite.

*Symmetry* Symmetry is necessary for wavelet filters to have linear phase and thus to avoid phase distortion.

*Smoothness* The  $M$  degree of smoothness means that the  $M^{th}$  derivative of a function is continuous at all points. Smoothness of wavelets plays an important role in compression. A higher degree of smoothness corresponds to better frequency localization of the filters.

*Number of vanishing moments* The number of vanishing moments  $N$  is defined as the number of moments of the wavelet that are zeros (see Equation (2.6)). It is related to the number of oscillations in wavelets and is important in singularity detection. A larger number of vanishing moments gives smoother wavelet.

$$\sum_k \Psi_k k^l = 0, \text{ for } 0 \leq l \leq N \quad (2.6)$$

*Orthogonality* The wavelets defined in 2.2 are orthogonal to each other:

$$\sum \Psi_{j,k}(x) \Psi_{j_0,k_0}(x) = \delta(j - j_0) \delta(k - k_0) \quad (2.7)$$

Orthogonality links the  $L^2$  norm of a function to the norm of its wavelet coefficients by (2.8).

$$\|f\| = \sqrt{\sum_{j,k} c_{j,k}^2} \quad (2.8)$$



where  $c_{j,k}$  is the wavelet coefficient at level  $j$  and position  $2^{-j}x - k$ .

Orthogonality is important for numerical calculation. If orthogonal wavelets are used, the fast wavelet transform is a unitary transformation and will not increase the error in the initial data.

### 2.1.3 Discrete and Inverse Discrete Wavelet Transform

The discrete wavelet transform(DWT) and inverse discrete wavelet transform(IDWT) are used in Chapter 3 to analyze the paper machine sheet property data. Decomposing a signal into wavelet components is referred to a Wavelet Decomposition or Wavelet Transform. The decomposition is done by correlating a signal with the scaled and translated versions of the mother wavelet. Figure 2.4 is the block diagram of the discrete wavelet transform(DWT) and inverse discrete wavelet transform (IDWT) of Mallat's pyramid algorithm which is used in this thesis.

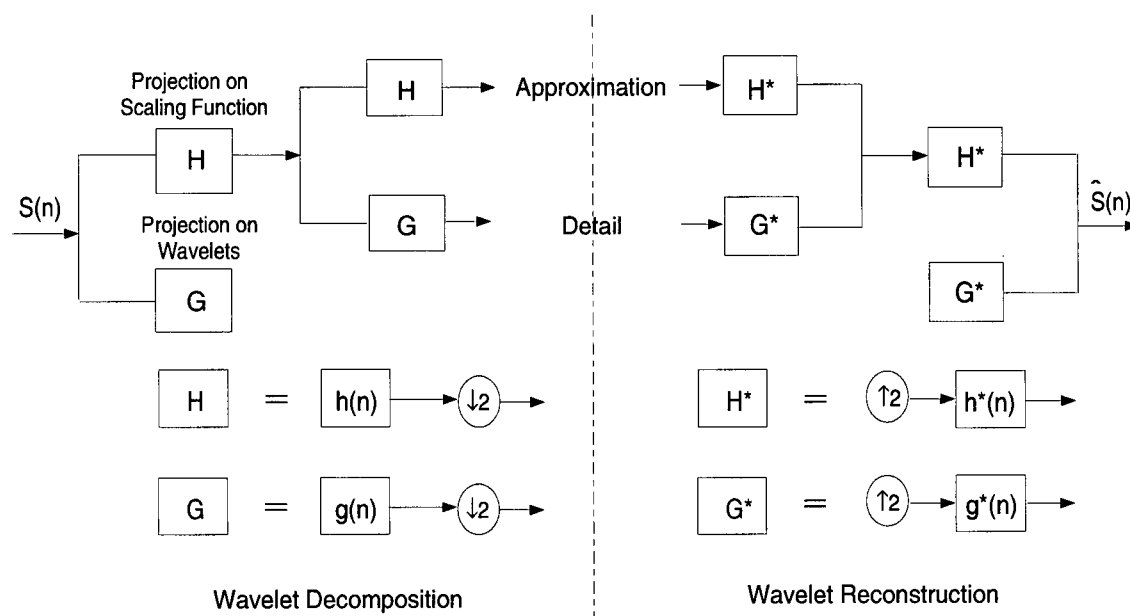


Figure 2.4: Block diagram of DWT and IDWT

Here  $h(n)$  is the lowpass filter,  $g(n)$  is the highpass filter. The discrete wavelet transform is implemented by filtering the signal and then downsampling. The filter  $h(n)$  and  $g(n)$  are not independent since:

$$g(n) = (-1)^{1-n} h_{1-n} \quad (2.9)$$

The filtering and downsampling are done iteratively until the filter length is 1, and the wavelet coefficients are obtained. The original data set has  $2^N$  points.

The inverse DWT is implemented by upsampling the lowpass components and highpass components and then convoluting the upsampled signal with the reconstruction lowpass and highpass filter respectively. When quadrature mirror filters are used in the wavelet decomposition and reconstruction, perfect reconstruction is guaranteed.

## 2.2 Wavelet Packet Analysis

The wavelet packet method is a generalization of wavelet decomposition that provides a richer signal analysis. As can be seen in Figure 2.3, the time-frequency resolution are predetermined for STFT and wavelet transform. However when using wavelet packets, it can be changed in a signal decomposition. An adaptive procedure for choosing the most suitable set of basis is also described in this section.

### 2.2.1 Wavelet Packets

Roughly speaking a wavelet packet  $W$  is a square integrable modulated waveform, well localized in both position and frequency. As shown in the following equation, it has three parameters: scale parameter  $j$  (resolution level), time-localization parameter  $k$  (translation level) and oscillation parameter  $p$ .

$$W_{j,p,k}(x) = 2^{-j/2} W_p(2^{-j}x - k) \quad (2.10)$$

The wavelet packet function is defined as Equation (2.11) and (2.12).

$$W_{2p}(x) = 2 \sum_{k=0}^{2N-1} h_k W_p(2x - k) \quad (2.11)$$

$$W_{2p+1}(x) = 2 \sum_{k=0}^{2N-1} g_k W_p(2x - k) \quad (2.12)$$

where  $W_0(x) = \Phi(x)$  is the scaling function and  $W_1(x) = \Psi(x)$  is the wavelet function.

Figure 2.5 gives the plots of Harr wavelet packets bases for  $p = 0$  to 7.

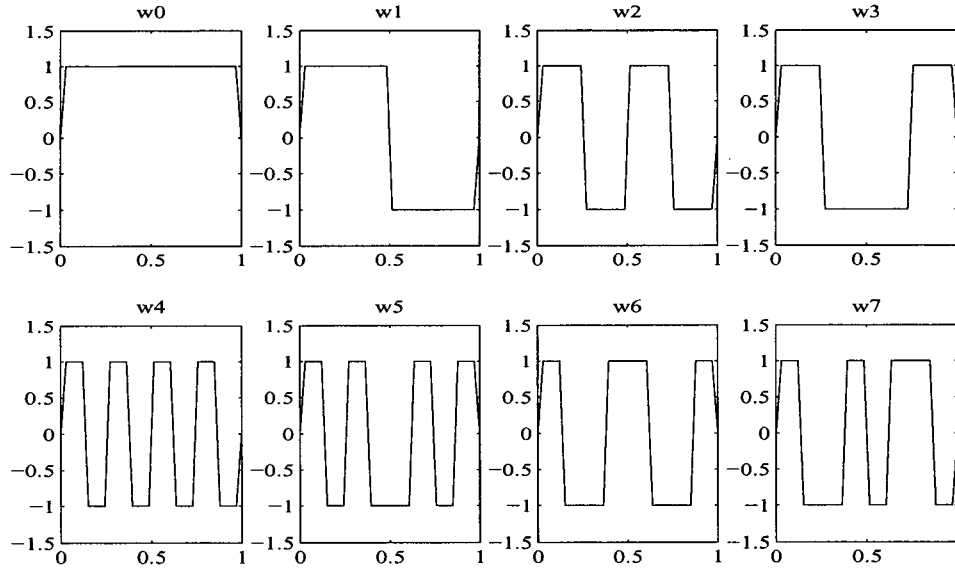


Figure 2.5: The Harr wavelet packets

### 2.2.2 Discrete Wavelet Packet Transform(DWPT)

In the wavelet packet framework, the projections on the scaling function and the wavelet are recursively decomposed to obtain a binary tree. Figure 2.6 shows the Filter bank implementation of Discrete Wavelet Packet Transform. It is noted that in the wavelet decomposition procedure, the generic step splits the approximation coefficients into two

parts. However, in the case of wavelet packet decomposition, both the approximation and the detail coefficients are decomposed into two parts.

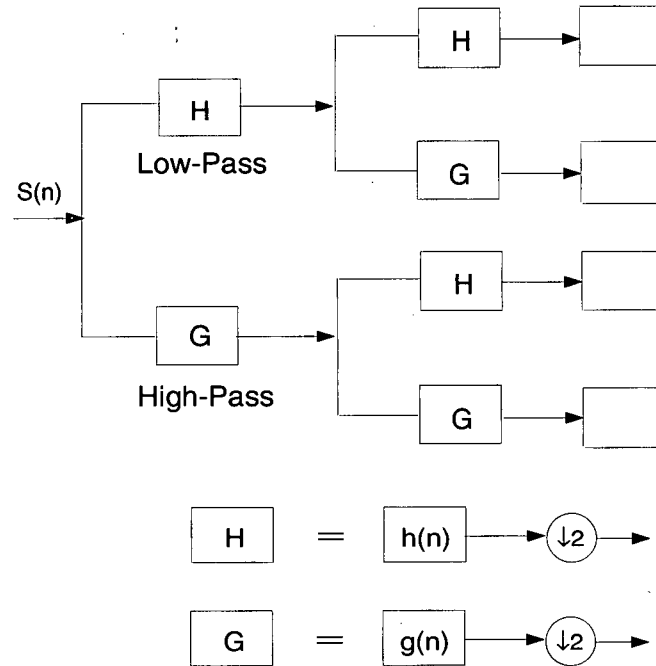


Figure 2.6: Filter bank implementation of DWPT

To illustrate the difference of wavelet and wavelet packets, a quadratic chirp signal in Figure 2.7(a) is analyzed. Figure 2.7 (b) and (c) plot the coefficients using wavelet and wavelet packets respectively.

The frequency of this chirp signal increases quadratically over time. In (b), a wavelet analysis does not easily detect the time-frequency property, however this property can be easily observed from the linear slope of the greatest wavelet packet coefficients in (c). As can be seen from Figure 2.7, wavelet packets can accurately single out the important frequency bands. This is used in Chapter 4 to achieve a better controllable and uncontrollable variation separation.

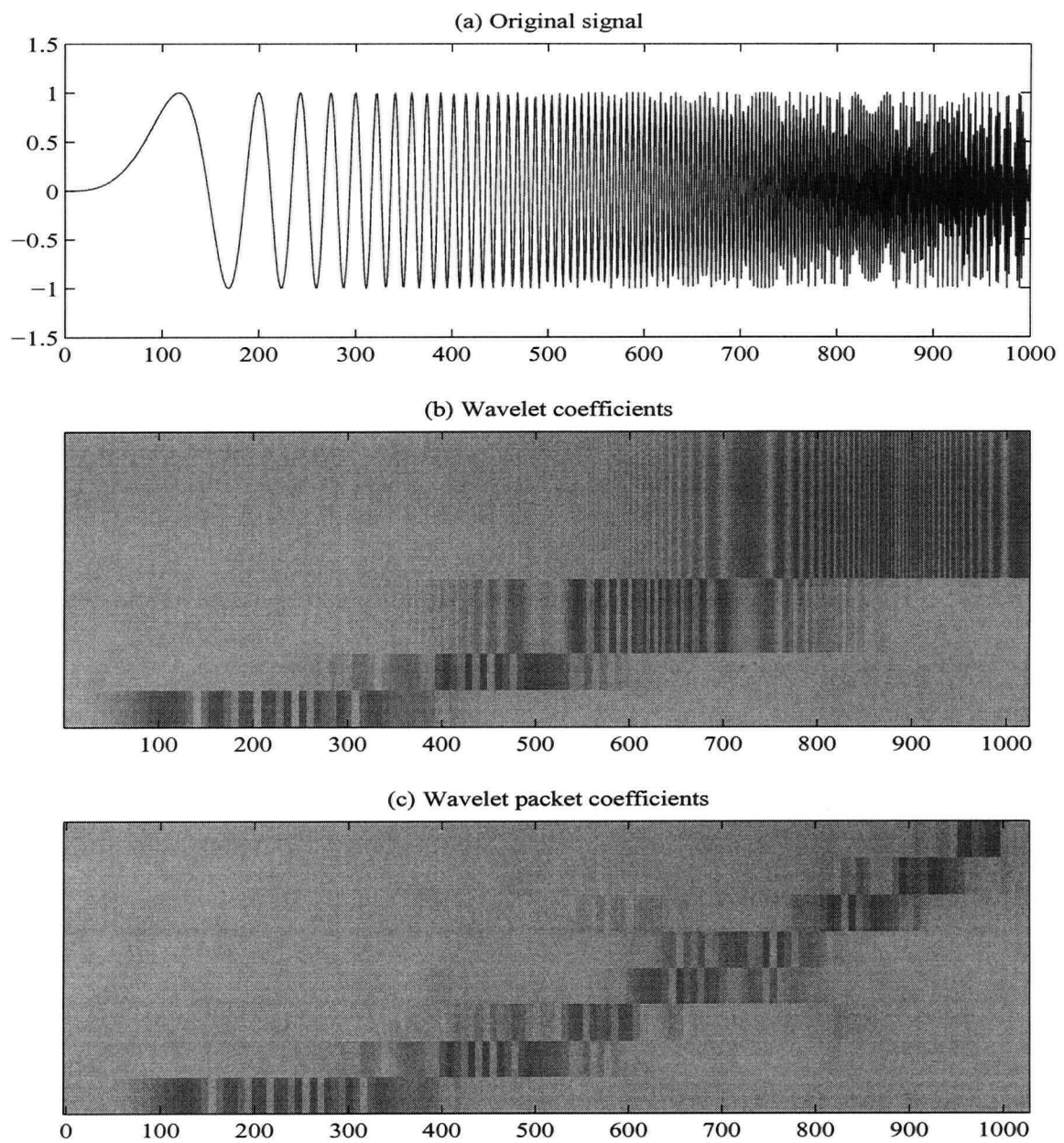


Figure 2.7: Analysis of a chirp signal(a) using wavelets(b) and wavelet packets(c)

### 2.2.3 Best Basis Algorithm

Wavelet packet decomposition provides an increased flexibility due to the large number of bases available for decomposition. The projection of the signal onto wavelet packet components produces a tree of  $N \log N$  coefficients. A signal of length  $N = 2^M$  can be decomposed in at most  $2^N$  different ways. To reduce the complexity and redundancy, a design objective can be used to choose a best representation of the original signal. The Best Basis Algorithm proposed by Coifman and Wickerhauser [38] is a widely used adaptive procedure. This algorithm finds the most efficient tree structure for a given signal in the sense that the signal will be economically represented. Since the number of non-zero coefficients is minimized, it is widely used for compression. Before each decomposition, this algorithm calculates the entropy (to be defined next) value of the child nodes and then compares this value to the parent node as a basis for deciding if further splitting is necessary in order to get a minimum-entropy decomposition tree.

Entropy is the measure of information concentration in a signal. If a parent node is of lower entropy, the decomposition is not carried out. If the children have lower entropy value, then further decomposition is carried out [38]. Some information cost functions that are normally used to calculate entropy are:

(1) *Shannon entropy*

The entropy involves the logarithm of the squared value of each signal sample which is defined as:

$$H_s = - \sum (p_i) \log(p_i) \quad (2.13)$$

where  $p_i = \frac{|s_i|^2}{\|s\|^2}$ , and  $x \log(x) = 0$  if  $x = 0$ .

(2) *The number of samples above a threshold level*

The number of samples in the signal whose absolute value exceeds the threshold value, the threshold is defined as:  $\sqrt{2 \log_e(n \log_2(n))}$

(3) *Concentration in  $l^p$  norm*

Concentration in  $l^p$  norm is defined as  $\|\{x\}\|_p$  for an arbitrary  $p < 2$ . The smaller is the  $l^p$  norm of a function, the more concentrated is its energy into a few coefficients.

(4) *Logarithm of energy*

Logarithm of energy is defined as the sum over all samples of the signal as following:

$$H(s) = \sum \log(s_i^2) \quad (2.14)$$

### 2.3 Multi-Resolution Analysis

In this thesis, multi-resolution analysis based on wavelet and wavelet packet will be applied to paper machine data. The idea of multi-resolution analysis is to write a function  $f$  as a sequence of approximations, each of which is a smoother version of  $f$ . The goal of multi-resolution analysis is to preserve the important features of the original signal using a smaller set of coefficients and thus get a parsimonious representation of the original signal. In multiresolution analysis, the space of square integrable function  $L^2(R)$  is decomposed into a closed sequence with the following properties:

1.  $V_j \subset V_{j-1}$  which means that the coarser subspace is contained in a finer subspace,
2.  $f(t) \in V_j \iff f(2t) \in V_{j-1}$ ,
3.  $f(t) \in V_0 \iff f(t+1) \in V_0$ ,
4.  $\bigcap_{j \in \mathbb{Z}} V_j = \{0\}$ , and  $\overline{\bigcup_{j \in \mathbb{Z}} V_j} = L^2(R)$ ,
5. There exists  $\phi \in V_0$ , such that  $\{\phi(t-n)\}_{n \in \mathbb{Z}}$  is an orthonormal basis for  $V_0$ .

By doing the scaling and translation for each level  $j$ , we have a collection of  $\phi_{j,k}$  (2.15) which form an orthonormal basis for  $V_j$ . And it is guaranteed that there is such an approximation shown in (2.15) which converges to the original function.

$$\phi_{j,k}(t) = 2^{-j/2} \phi(2^{-j}t - k) \quad (2.15)$$

## 2.4 De-Noising

One of the most important applications of wavelet and wavelet packet is de-noising. De-noising is the process of suppressing the unwanted noise part of the signal and recovering the signal. Wavelet and wavelet packets are superior to the traditional denoising methods, especially in estimating signals with jumps, spikes and non-smooth features [27, 4]. The denoising method in the wavelet packet framework is identical to that of wavelet framework which is shown as follows:

The noise in a signal is normally the part which lacks structure or coherence. A coherent part of the signal exhibits a concentration of energy in the representation domain and an incoherent part of the signal diffusely spread throughout in the representation domain.

Given the signal model in (2.16),

$$y_i = f_i + \delta\epsilon_i, i = 1, 2, \dots, n \quad (2.16)$$

where  $y_i$  is the noised measurement,  $f_i$  is the noise-free signal.  $\epsilon_i$  is a Gaussian white noise  $N(0,1)$ ,  $\delta$  is the noise level.

The decomposition of the signal normally concentrates into small number of coefficients, however the wavelet transform of white noise is still white noise which is evenly spread over all the coefficients [13]. Based on these facts, Donoho and Johnstone proposed the following three-step denoising procedure [10]:

1. Compute the wavelet decomposition of the original signal at given level  $N$ .
2. For each level 1 to  $N$ , apply soft thresholding to remove low magnitude wavelet coefficients.
3. Reconstruct the signal based on the original approximation and thresholded detail coefficients.



Figure 2.8 illustrates the result of wavelet de-noising using the above procedure. It can be seen that the sharp changes in the signal are preserved after denoising.

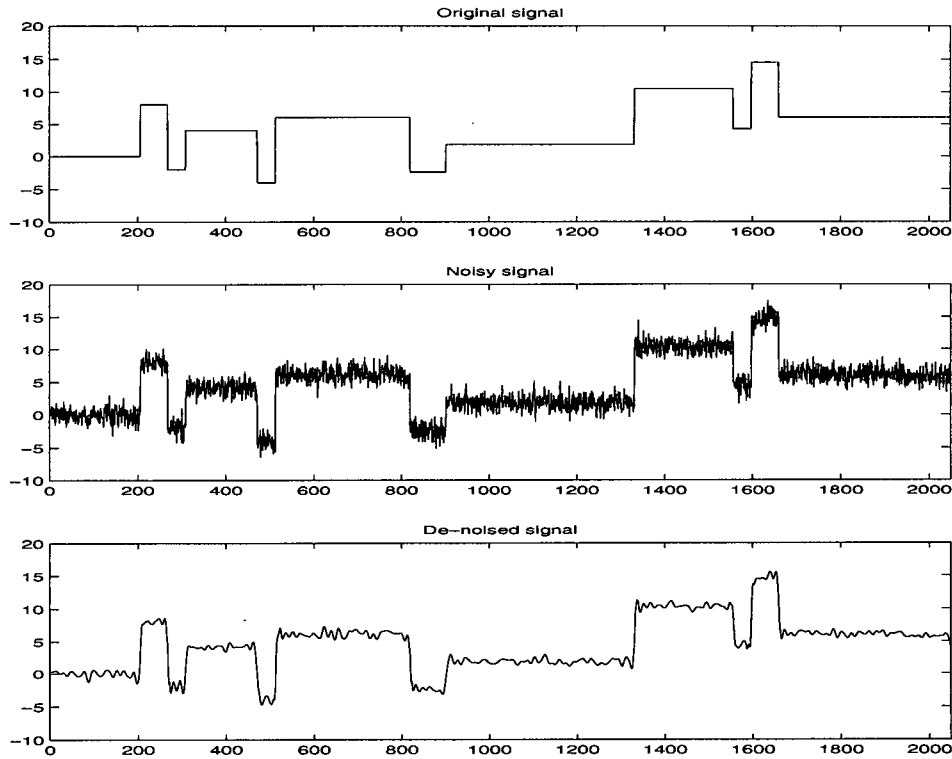


Figure 2.8: Wavelet denoising

In the de-noising procedure, the second step *thresholding* is the most important. Thresholding of the discrete representation is the key operation that may identify with the suppression of noise. Typically thresholding is only applied to the higher-resolution coefficient levels.

#### 2.4.1 Thresholding

There are two main types of thresholding: hard thresholding and soft thresholding. Hard thresholding sets all the coefficients with magnitude less than the threshold to zero while

leaving those coefficients greater than the threshold unchanged. The algorithm is given in (2.17).

$$c_{jk}^{hard} = \begin{cases} c_{jk} & \text{if } |c_{jk}| > \lambda \\ 0 & \text{otherwise} \end{cases} \quad (2.17)$$

Soft thresholding sets all the coefficients with magnitude smaller than the threshold to zero and shrinks the others by the threshold value. Its algorithm is given in (2.18).

$$c_{jk}^{soft} = \begin{cases} c_{jk} - \lambda & \text{if } c_{jk} > \lambda \\ 0 & \text{if } |c_{jk}| \leq \lambda \\ c_{jk} + \lambda & \text{if } c_{jk} < -\lambda \end{cases} \quad (2.18)$$

The graphical display of hard thresholding and soft thresholding is given in Figure 2.9

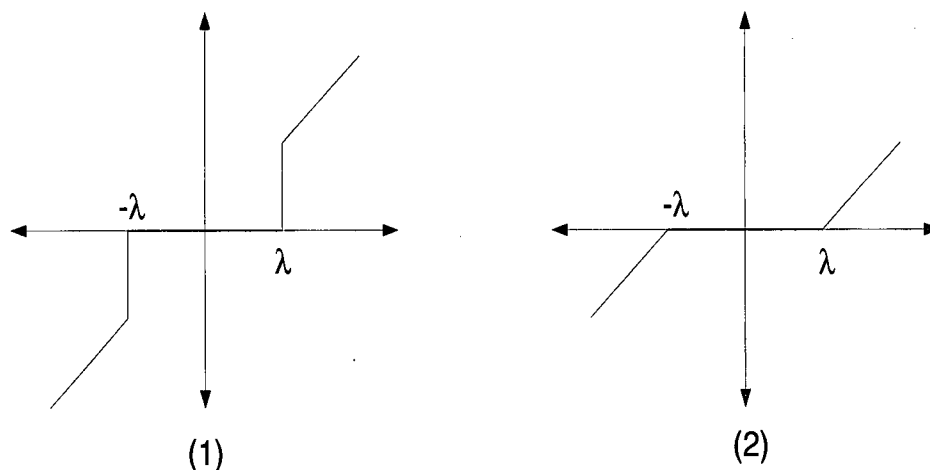


Figure 2.9: Hard thresholding and soft thresholding

Due to the discontinuity of the shrink function, hard thresholding tends to have larger variance but smaller bias [2]. Soft thresholding tends to have larger bias but smaller variance because of shrinking all large wavelet coefficients towards zero by  $\lambda$ .

V. Solo recently reformulated the soft thresholding method as an  $L_1$  regularised least squares problem. He proposed a new iterative algorithm to deal with the wavelet estimation in coloured noise and used it for transfer function estimation. Detailed reference can be found in [31].

### 2.4.2 Threshold Selection Methods

Donoho and Johnstone [10, 12, 22] have done extensive work on different threshold selection methods. The various thresholding values can be expressed by

$$t = \hat{\delta} \cdot \lambda \quad (2.19)$$

where  $\hat{\delta}$  is the noise level estimation and can be estimated using *Median Absolute Deviation(MAD)* method [11].

$$\hat{\delta} = \text{Median}(|c_{jk}|)/0.6745 \quad (2.20)$$

There are different criteria for choosing  $\lambda$ . The four most important threshold selection methods are described as follows:

(1) *Minimax threshold* applies the optimal threshold in terms of  $L_2$  risk. Minimax threshold depends on the sample size  $n$  and is derived to minimize the upper bound of the  $L_2$  risk in estimating a function. It does not have any closed form expression.

Table 2.1: Minimax thresholds for various sample sizes

n	$\lambda$	n	$\lambda$
64	1.474	2048	2.414
128	1.669	4096	2.594
256	1.860	8192	2.773
512	2.047	16384	2.952
1024	2.231	32768	3.131

Table 2.1 gives the approximate threshold value for different sample sizes. The minimax method does a better job at picking up abrupt jumps at the expense of smoothness.

(2) *Universal Thresholding* The thresholding value  $\lambda$  is given by

$$\lambda = \sqrt{2 \log(n)} \quad (2.21)$$

where  $n$  is the sample size. The universal threshold value is substantially larger than its minimax counterpart. So the universal method often gives smooth estimates but does not pick up jumps very well.

(3) *SURE Thresholding* is based on the principles of minimizing the *Stein Unbiased Risk Estimate (SURE)* for threshold estimates. It is smoothness adaptive and the advantage of this method is evident when the underlying function has jump discontinuities on a smooth background.

(4) *Hybrid SURE Thresholding* When the signal to noise ratio is very small, the SURE estimate may be very noisy and the Universal Thresholding is used. Otherwise SURE Thresholding is used.

The threshold selection methods described above apply to the noise model in (2.16). For correlated noise, thresholds must be rescaled by a level-dependent estimation of the level noise as in (2.22) [22].

$$t_j = \sigma_j \sqrt{2 \log(n)} \quad (2.22)$$

where  $\sigma_j$  is the standard deviation of wavelet coefficients at  $j_{th}$  level and  $n$  is the data sample size.

MultiMAD to be used in Chapter 3 and 4 is a resolution-dependent thresholding method using MAD noise estimator (2.20) to estimate the noise strength at each resolution level.

## 2.5 Compression

Data compression is another important application of wavelet and wavelet packet theory. Because wavelet and wavelet packet transforms and thresholding can compress the signal energy in a small number of coefficients, they can be used for data compression. In the paper industry, data compression is very useful for the storage of historical data for future use and the transfer of data between different paper mills.

There are two basic compression schemes: lossless compression and lossy compression. Here we only consider lossy compression, that is, we can accept some error as long as the reconstruction after compression is acceptable.

The wavelet lossy compression diagram is shown in Figure 2.10.  $s$  and  $\hat{s}$  are the input sequence and the recovered sequence respectively. The use of a quantizer is optional and can result in a high compression ratio at the expense of an additional error due to the quantization of wavelet coefficients. The compression procedure using wavelet packet is

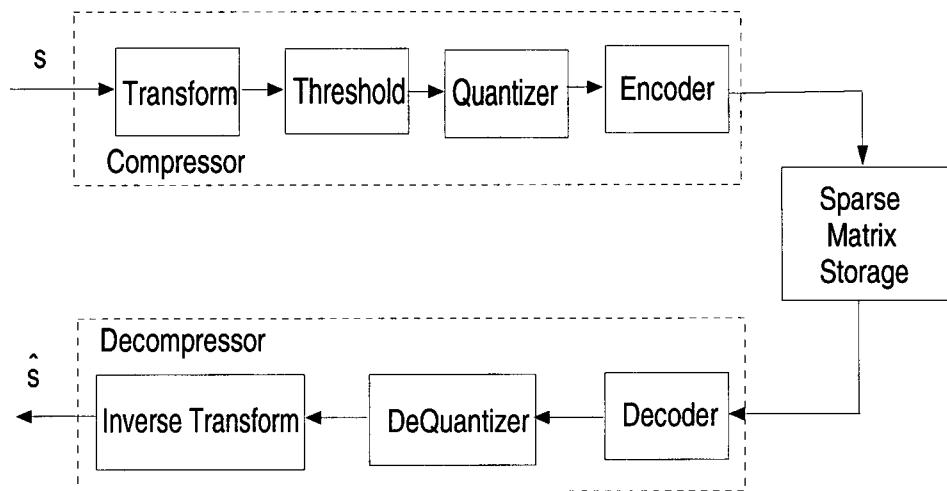


Figure 2.10: Block diagram of compression

identical to that of wavelet. The only new feature is the increased flexibility due to a

large number of bases from the signal decomposition. A design objective can be used to choose the best representation of the original signal.

Because wavelet packet coefficients can represent signals at least as efficiently as wavelet coefficients, the method normally achieves better compression. Smooth oscillatory signals such as speech or music can be compressed significantly better using wavelet packet bases, which can accurately single out the important frequency bands.

## 2.6 Two-Dimensional Signal Analysis

The measured paper machine process data used in this thesis is two-dimensional data, and the two-dimensional transform is used. The two-dimensional Discrete Wavelet Transform(DWT) and Discrete Wavelet Packet Transform(DPWT) can be achieved by two separate one-dimensional transforms. As shown in Figure 2.11, the rows of the two-dimensional signal are decomposed first and then the columns of the signal are decomposed.

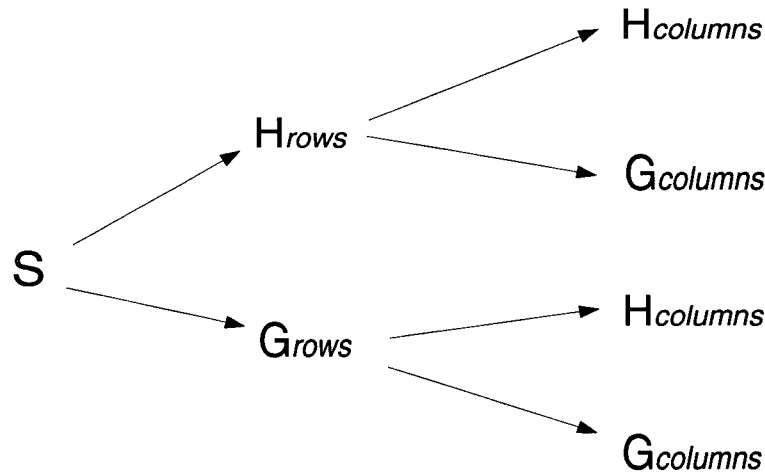


Figure 2.11: Diagram of two-dimensional DWT

Figure 2.12 illustrates the decomposition of a fingerprint image using two-dimensional discrete wavelet transform [24]. After the first level decomposition, there are four sets

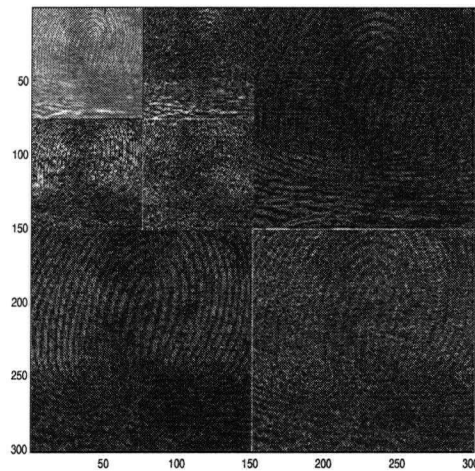


Figure 2.12: Decomposition of signal using 2-dimensional DWT

of coefficients, which correspond to the approximation, the vertical detail, horizontal detail and diagonal detail at level one. At the next level, the DWT only operates on the approximation coefficients while DWPT operates on both the approximation and detail coefficients.

## Chapter 3

### Comparison of On-machine and Off-machine Measurements of Paper Properties Using Wavelet Analysis

#### 3.1 Introduction

Improvements in measurement and control systems now provide paper machine profile data with high resolution. In recent years, wavelet filtering has been used in the paper machine data analysis and has been shown to give a superior performance in comparison to the traditional estimation methods. Wavelet processing can separate the CD and MD variations as well as remove high frequency measurement noise. Wavelets are effective for the detection of signals in the noisy data, leading to a better visualization and estimation. The sheet properties can be represented economically and without loss of detail. Detailed reference can be found in [27, 28, 4]. This chapter is concerned with the adaptation of wavelet techniques to the analysis of two-dimensional paper sheet properties. The accuracy of analyzing paper machine data using discrete wavelet transform is first examined by applying wavelet filtering to both the paper machine on-line scanner data and off-line analyzer data from the same paper sheet. The Cross Machine Direction(CD) analysis is used to show the controllable and uncontrollable variation that can and cannot be removed by the CD control system. The machine direction(MD) analysis is intended to obtain clean MD profile. These results are also of interest since it is unusual for both scanner data and laboratory test data to be available for the same paper sheet. Such



direct comparisons are of value to the paper manufacturers since they provide confirmation of the validity of the data used to control production. The logistics of identifying the sheet samples, connecting them and transporting the CD and MD samples to the laboratory from the mill site are formidable.

## **3.2 On-line Scanner Data and Off-line Tapio Analyzer Data**

### **3.2.1 On-line Scanner Data**

As described in chapter 1, paper machine on-line scanner data are collected from zigzag sampling of the paper by the on-line scanner located at the dry end. The scanner has a moving head with six sensors to measure paper properties including dry weight, caliper, moisture, top side gloss and wire side gloss. In this thesis, the basis weight variations are used for study, and caliper is also studied in Chapter 4.

The raw basis weight data consists of 130 scans and 685 data points per scan, this represents one entire jumbo reel, which is approximately 58 km of paper. The speed is 23 seconds per scan, and there are 130 scans so the speed is:  $58000/(130 * 23) = 19.4m/s$  or  $70km/h$ .

### **3.2.2 Tapio Analyzer Data**

The Tapio analyzer is an off-machine tool that measures the paper properties and calculates the variability at a high resolution. The Tapio software uses traditional signal processing analysis techniques to determine the variation and frequency content of the paper properties. The information from the analyzer is often used to determine how a paper machine is operating by measuring the spectral content of each paper property at a high resolution. The Tapio data consists of Cross Machine Direction(CD) samples and Machine Direction(MD) samples.

The CD samples are 150 CD strips taken over the entire width (7.8m) of the jumbo reel. This is approximately equivalent to two scans at the scan positions #127 and #128. Each CD strip is about 30cm wide, has about 9500 measurements and thus contains more high frequency information. During the data collection, the CD strips are taped together in order to run through the Tapio analyzer continuously, so the CD sample data has many spikes caused by the tapes following the CD strips (as shown in Figure 3.13).

Two MD butt-roll samples were taken from the winder. The rolls were 8 inch wide and had a diameter of 39 inches, and were taken 25 feet from the front side of the machine. The length of MD samples are 415743 and 385535 respectively.

### 3.2.3 Transfer of Tapio Data

Before the analysis starts, the spikes in the CD samples of Tapio data are removed. Figure 3.13 shows strips 26 to 75 of the Tapio CD sample data for basis weight property. As can be seen, the average value for basis weight is around  $45 \sim 55g/m^2$ . The spikes between the strips have much higher values (around  $120g/m^2$ ) and should be removed. This is achieved by first detecting the tapes by searching the unusually high or low values and removing those values. The data shown in Figure 3.14 is the data after transfer, i.e. the spikes in Figure 3.13 have been removed. Then all the 150 CD strips are aligned together and are averaged into two scans for use in later analysis.

## 3.3 Cross Machine Direction Analysis

The CD analysis is intended to separate the controllable and uncontrollable variations. Here in the absence of actuator response knowledge, the CD control bandwidth is assumed to be more than 2 actuator spacings. So the variations with a wavelength over this value should be removed by the CD control systems.

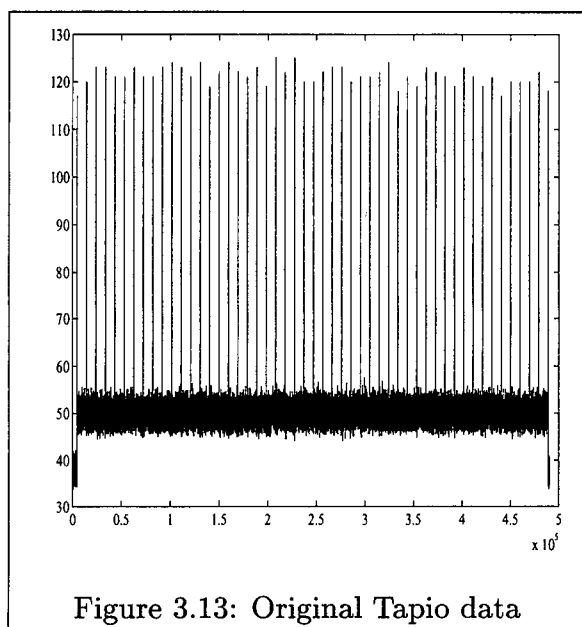


Figure 3.13: Original Tapio data

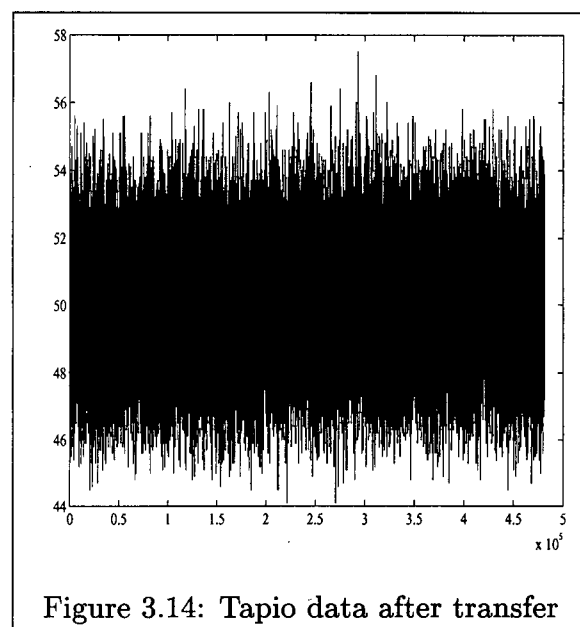


Figure 3.14: Tapio data after transfer

Wavelet analysis requires the input data to be dimension  $(2^{m_1}, 2^{m_2})$ . Paper machine measured data seldom gives this kind of size and the data should be padded to the proper size before the transformation. There are three ways of padding: *zero padding*, *periodic extension* and *symmetric replication*, detail reference can be found in [24, 32]. The symmetric replication is used here because it can reduce the boundary errors due to the periodic nature assumed by the DWT algorithm. After reconstruction the extra data sets are removed.

### 3.3.1 On-line Scanner Data Analysis

Figure 3.15 shows the original raw profile from the on-line scanner. Figure 3.16 gives the diagram of on-line scanner data analysis. First two-dimensional discrete wavelet transform is performed on the raw profile. It has been found that for the paper machine data analysis, the Symlet wavelet family and Daubechies family produce better results. The Symlet wavelet ( $Sym_4$ ) of filter length 8 is selected for this set of data, MultiMAD

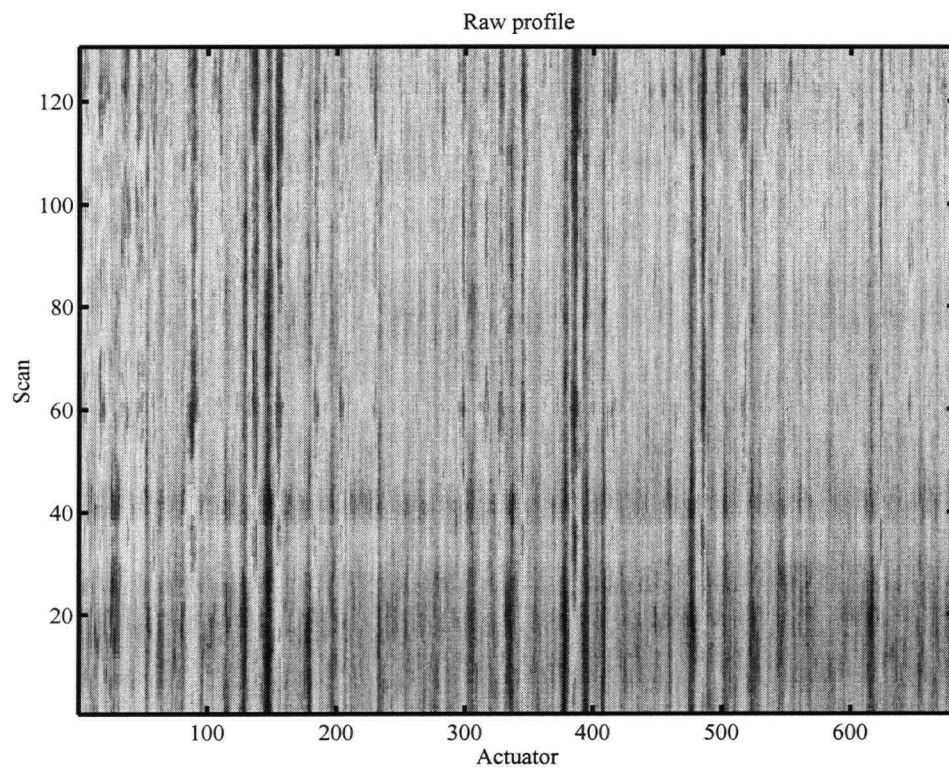


Figure 3.15: Raw profile

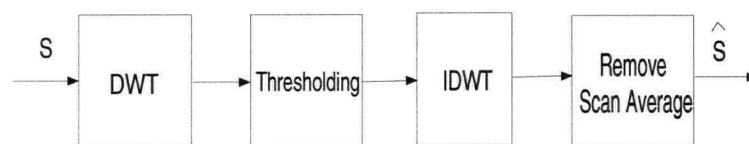


Figure 3.16: On-line scanner data analysis using wavelet

resolution-dependent thresholding is used here due to its ability to adjust the thresholds at different levels.

The multiresolution plot for scan 128 and the normalized wavelength is shown in Figure 3.17. The normalized wavelength is the corresponding spatial wavelength, normalized in terms of actuator spacing rather than absolute dimensions. The wavelength of the wavelet details from level 1 to level 3 is:  $0.21 \sim 0.43$ ,  $0.43 \sim 0.86$ ,  $0.86 \sim 1.72$  respectively. The decomposition level that can best separate the controllable and uncontrollable variations is the third level, because the wavelength at this level is 1.72 which is close to 2 and is the best dividing wavelength that can be reached using wavelet analysis.

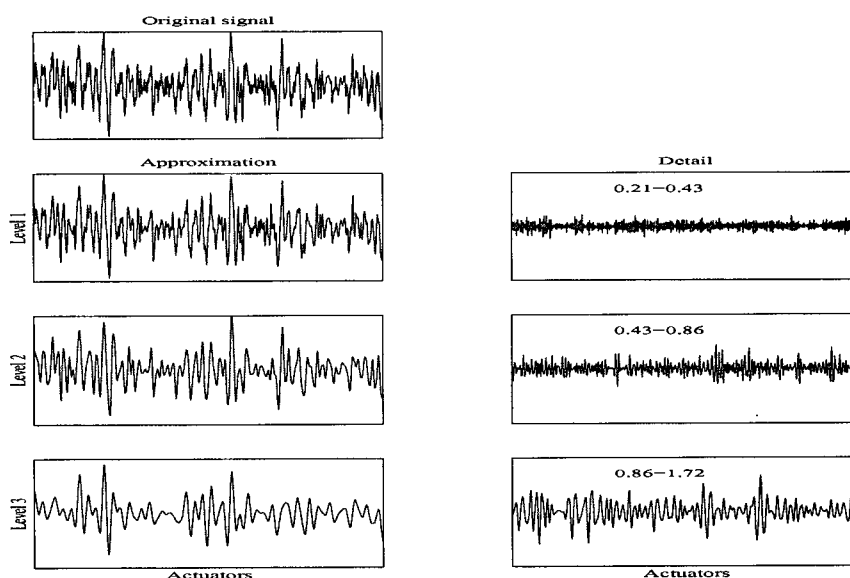


Figure 3.17: Multiresolution analysis and normalized wavelength

After wavelet filtering, the thresholded details and approximation are used to reconstruct the profile by Inverse Discrete Wavelet Transform(IDWT). Finally the scan average is removed and the clean CD profile is obtained. Because the zigzag sampling data contain information about both CD and MD variations, they should be separated

for different control and troubleshooting purposes. The separation is achieved in this case by subtracting the scan average after wavelet filtering to remove the MD trend.

Figure 3.18 is the scan average after wavelet filtering. Figure 3.19 to Figure 3.22 are

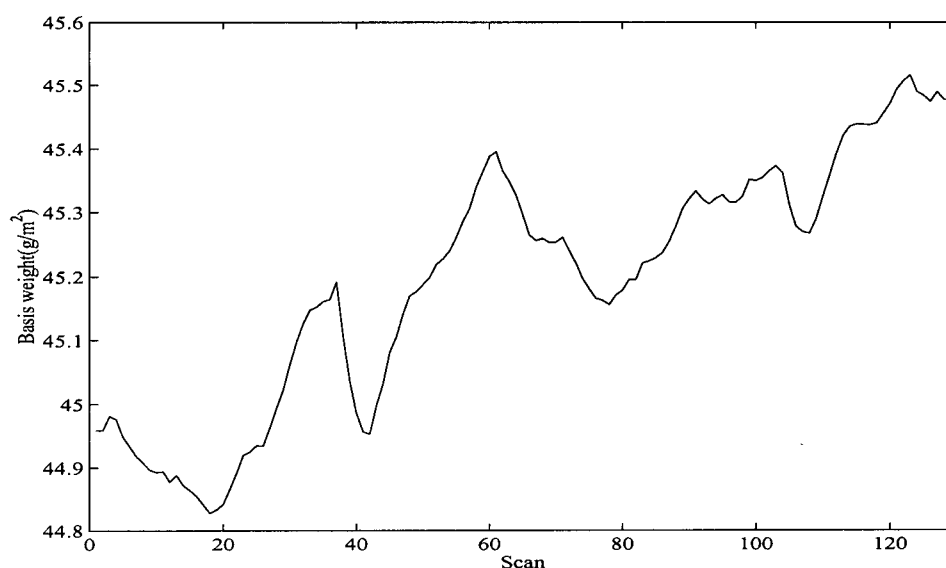
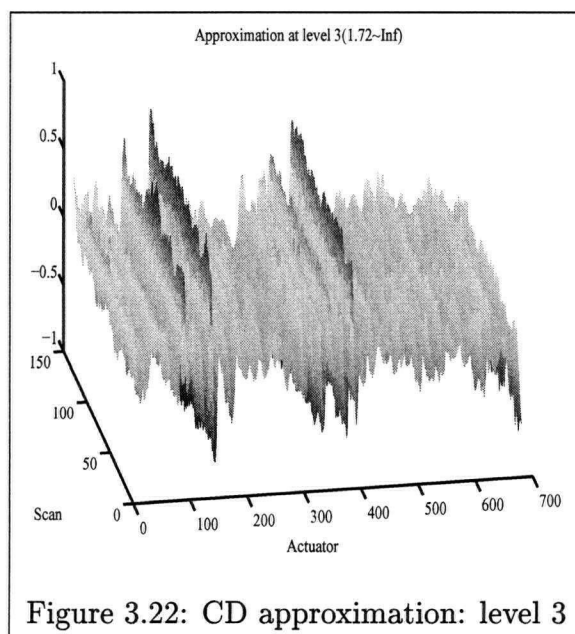
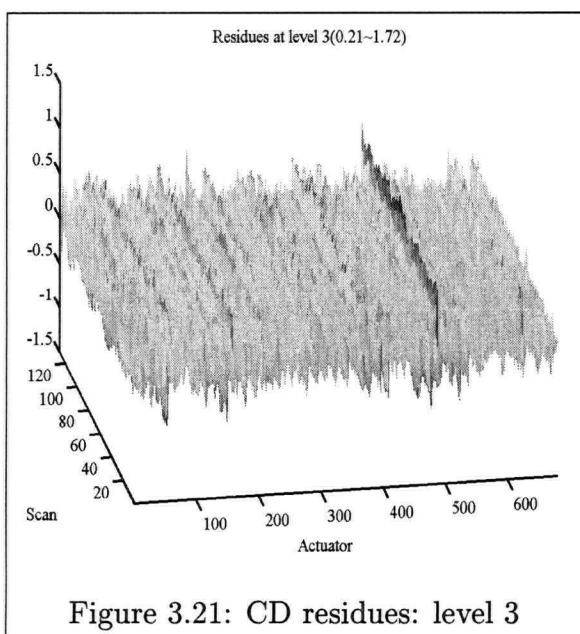
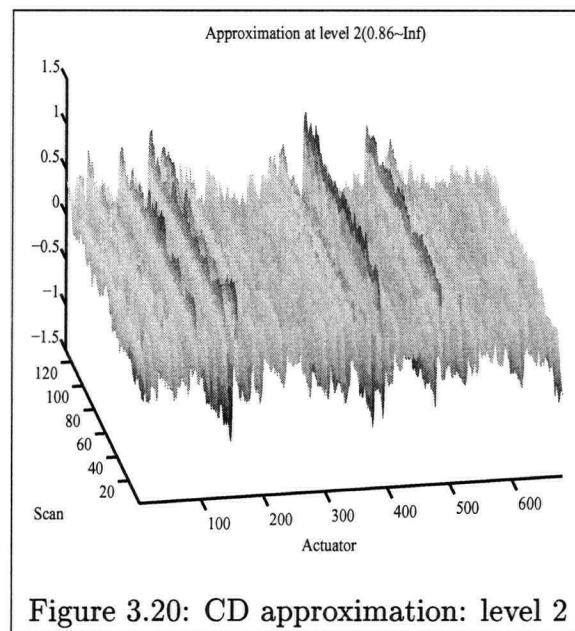
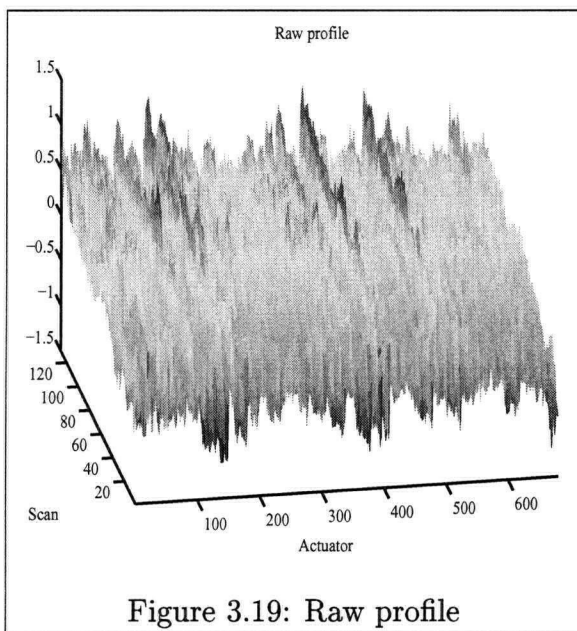


Figure 3.18: MD scan average

the raw profile, CD approximation at level 2, CD residues at level 3 and CD approximation at level 3 respectively. The image displays of profiles are plotted in Figure 3.23 to Figure 3.26 respectively.

The controllable variations that should be removed by the control system are the streaks shown in the level 3 approximation, while the streaks in the residues are the residual variations that are inherent in the system. The profile after wavelet filtering provides better visual reference for the operators, who can quickly see if the paper being produced is within specifications.





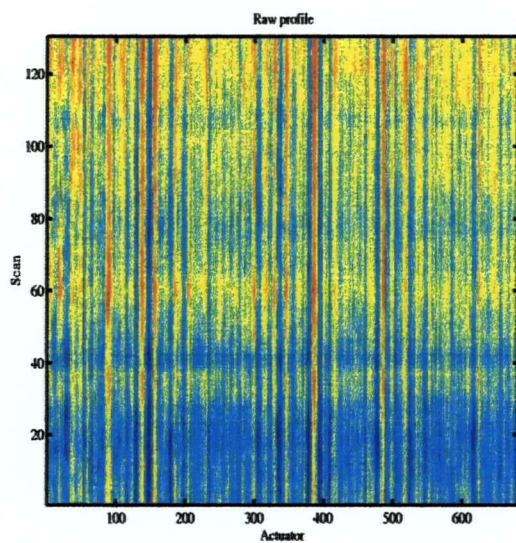


Figure 3.23: Raw profile

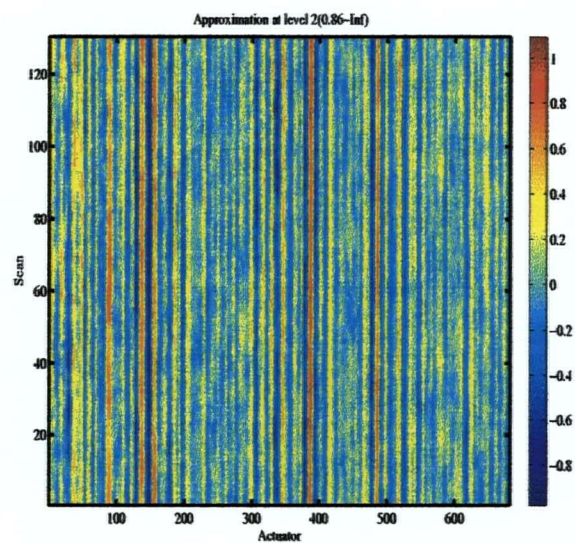


Figure 3.24: CD approximation: level 2

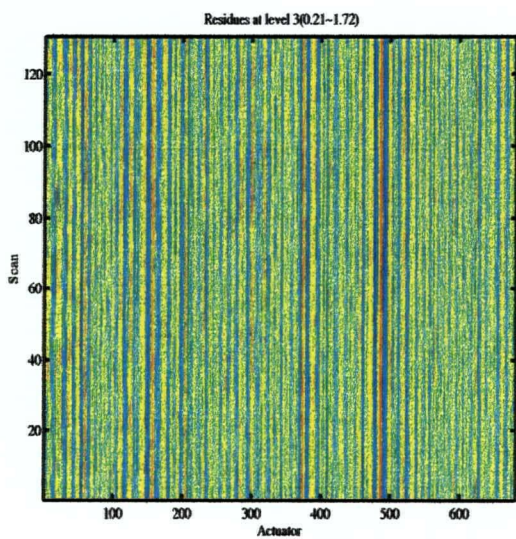


Figure 3.25: CD residues: level 3

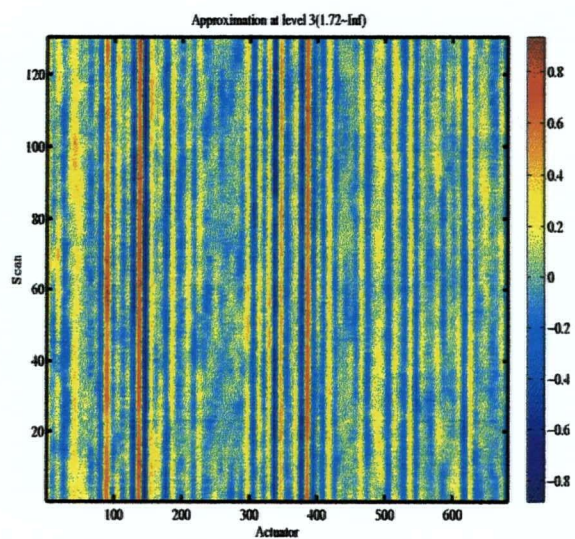


Figure 3.26: CD approximation: level 3



### 3.3.2 Off-line Tapio Analyzer Data Analysis

The Tapio data set was averaged into two scans, 75 plies were averaged together for each scan. Then one dimensional wavelet filtering was applied to each scan. Because there is a large resolution difference between Tapio data (150 strips, 9500 measurement for each strip) and Measurex data (130 scans, 685 data points for each scan), the Tapio data is decomposed to a deeper level to compare variations for the same wavelength.

The wavelength of variations for the Tapio data from level one to level eight are:

0.031, 0.062, 0.124, 0.248, 0.496, 0.992, 1.98, 3.96 respectively.

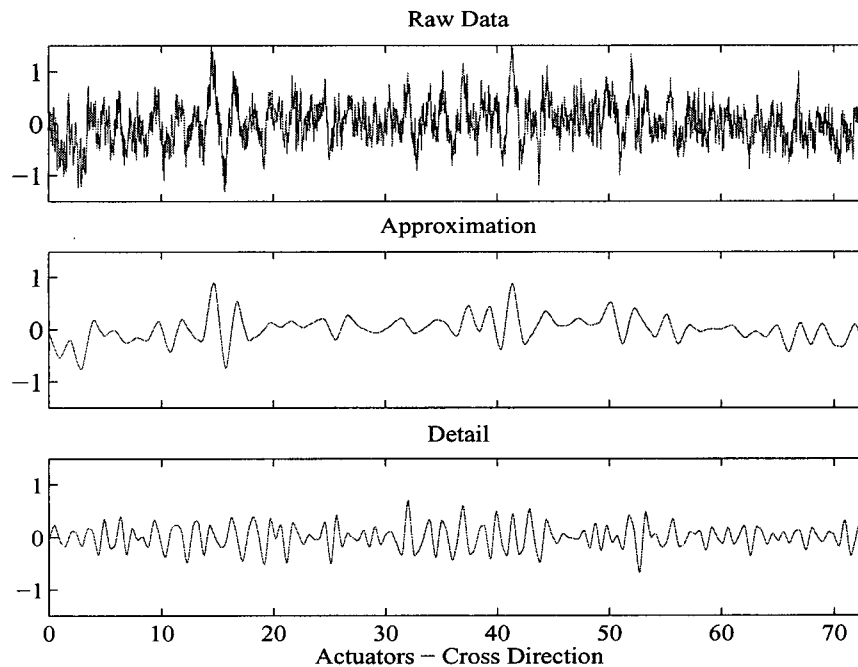


Figure 3.27: Wavelet approximation and detail: level 7, Tapio data

The decomposition level that can separate the controllable and uncontrollable variations is level 7 corresponding to a wavelength of 2. Figure 3.27 show the wavelet approximation and detail at level 7.

### 3.3.3 Comparison of On-line Data Analysis and Off-line Data Analysis

Figure 3.28 plots the third level wavelet approximation and detail of on-line data scan 128 and the seventh level wavelet approximation and detail for Tapio data at the same position(plies 76 to 150). Given the big difference of resolution between the two data sets, the match is excellent at the corresponding level.

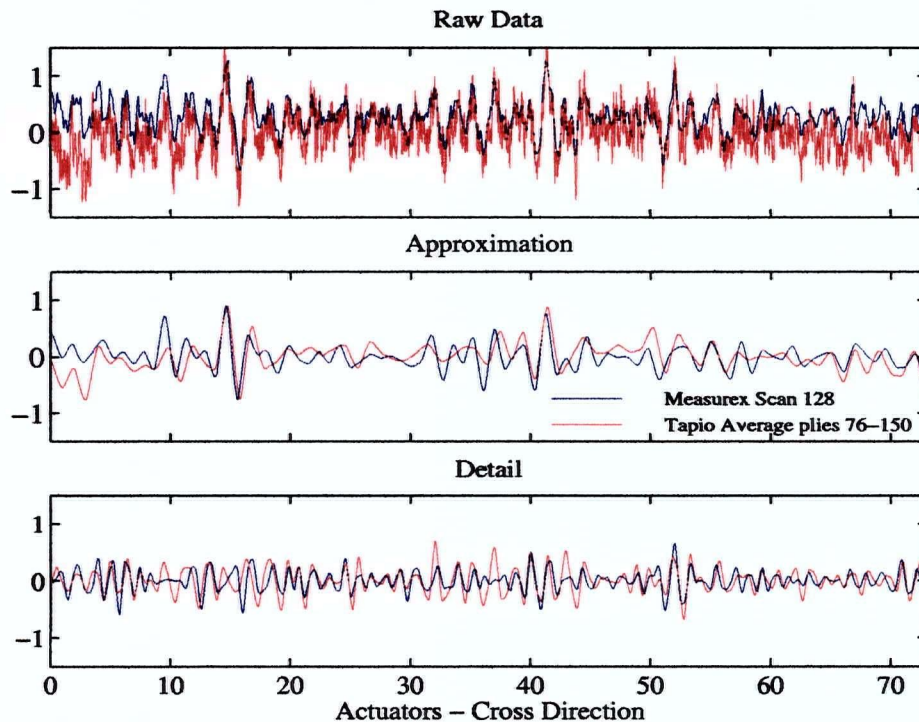


Figure 3.28: Comparison of on-line scanner data and Tapio data

The match between the on-line data and off-line data shows these two data sets are measuring the same properties and the wavelet transform does not produce any artifacts in the data. It also shows that wavelet filtering can be used to successfully separate the CD and MD variation of the on-line scanner data.

The differences between the two data sets are caused by the following reasons:

1. In order to remove the spikes between two consecutive plies, the Tapio data transfer is carried out before the data analysis. However during this process some human errors occur.
2. There exists difference between the dividing wavelength for the on-line scanner data(1.72) and the Tapio data(1.98).

### 3.4 Machine Direction Analysis

The MD direction analysis is intended to further remove noise and to obtain clean MD profile. Due to the data collection problem of the MD data set, only the on-line scanner data analysis is discussed here. The scanner data with CD profile removed is shown in Figure 3.29. After 2D wavelet filtering was performed on the raw data to remove the

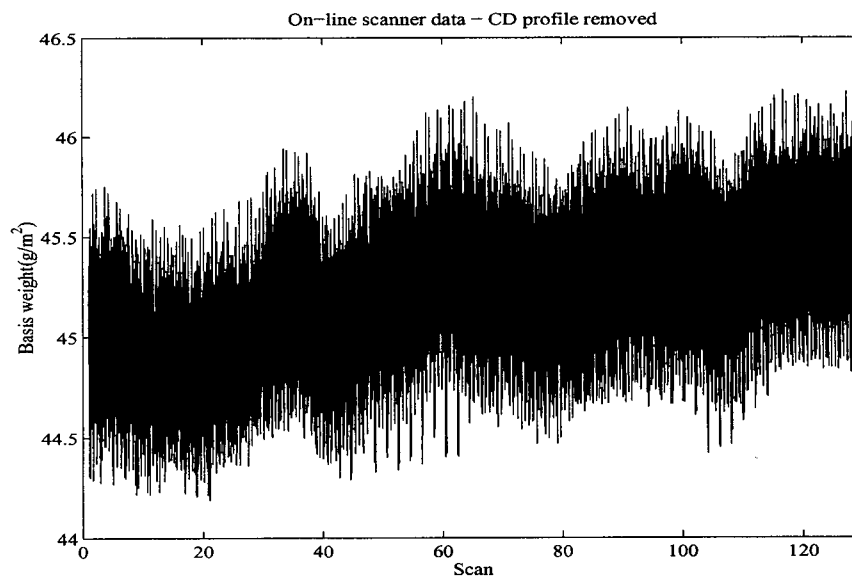


Figure 3.29: Scanner data with CD profile removed

high frequency MD and CD variations (as shown in the CD analysis), the resulting CD

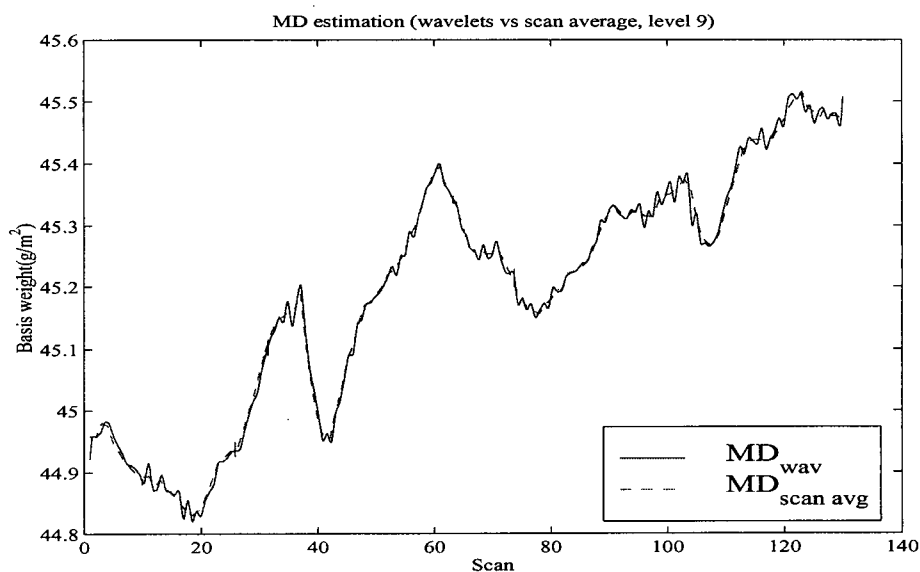


Figure 3.30: Scan average and wavelet estimate of MD profile

approximation profile was subtracted from the raw data to leave only lower frequency MD variations and the high frequency residual variations in the data. Next, the data was transformed into a vector representing the MD path that the scanner has traced on the paper. Finally this vector was further decomposed to level 9 using one-D wavelet transform. The wavelet MD estimate versus the scan average is plotted in Figure 3.30. The close match between the wavelet estimation and the scan average shows that wavelet can separate the CD and MD variation of paper machine on-line scanner data successfully.

## Chapter 4

### Wavelet and Wavelet Packet Analysis of Industrial Data

In this chapter, both wavelet and wavelet packet transforms are used for industrial data analysis. One important application of multi-resolution analysis is performance monitoring and performance assessment. Both wavelet and wavelet packet analysis can be used for CD profile separation and control performance assessment, but wavelet packet analysis provides more flexibility for decomposition and this characteristic is used to get a better separation of CD profile variations, and thus a more accurate assessment of the system. Denoising and compression of the basis weight profile are also achieved using both wavelet and wavelet packet transforms and the results are compared.

#### 4.1 CD Variation Separation Using Wavelet Analysis

The profile measurements of paper machine contain various frequency components. Relating the different frequency ranges to the process, the variation can be classified as [6]:

short term variation:  $< 1s$  period

medium term variation:  $1s - 200s$  period

long term variation: Longer than  $200s$  period

Short term variation starts where formation leaves off (wavelength of  $100mm$ ) and includes all wavelengths up to the length of paper made in one second. Short term

variation is primarily affected by hydraulic pulsation, hydraulic stability and equipment vibration. Medium term variation is primarily affected by blending, flow stability and fast control loops. Long term variation is primarily affected by system stability and slow control loops.

After wavelet decomposition to a suitable level, the signal is decomposed into different resolution levels which correspond to different wavelengths. Based on this, the process variations can then be divided into categories associated with the controllable and uncontrollable wavelengths and noise. If there exist controllable components in the profile, improved control actions are required to remove those variations.

Here the CD control bandwidth is assumed to be wavelengths above 2 actuator spacings. This is a realistic assumption that may be modified in some cases when the response to an individual actuator adjustment is known. The on-line scanner data consist of 130 scans, 685 data boxes. It is already known from chapter 3 that the wavelet decomposition level that separates the CD controllable and uncontrollable variations is level 3.

The wavelet decomposition tree is shown in Figure 4.31, each node is also labeled with its wavelength. According to this wavelet decomposition tree, the signal is divided

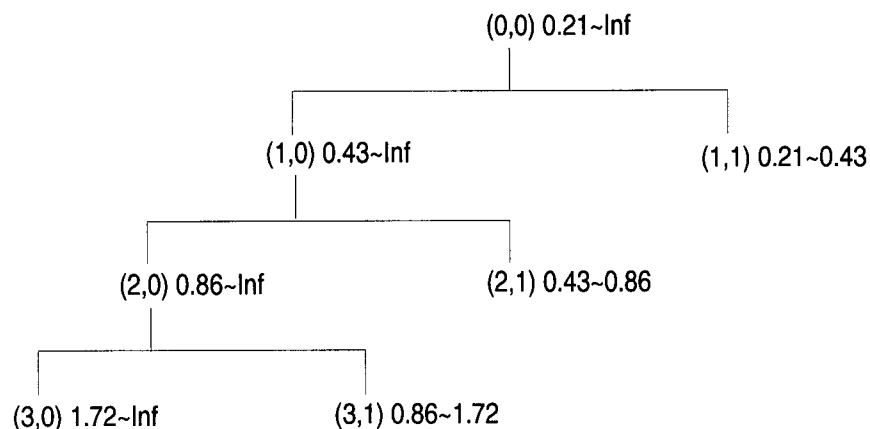


Figure 4.31: Wavelet decomposition tree and wavelength at each node

into the following consecutive wavelength divisions (units: actuator spacing):

node(1,1): 0.21 to 0.43

node(2,1): 0.43 to 0.86

node(3,1): 0.86 to 1.72

node(3,0): 1.72 to  $\infty$

Node (3,0) in the decomposition tree has a wavelength over than 1.72, and is regarded as controllable in the wavelet analysis. However, this would include part of the uncontrollable variation, because the controllable wavelength should be longer than 2 actuator spacings. Since in the wavelet decomposition tree, only the low frequency approximation component is further decomposed, the detail information is not decomposed any more. Thus 1.72 is the best dividing wavelength that can be reached using wavelet analysis. Once the wavelength  $\lambda$  of the original data is determined, with wavelet analysis, separation is only possible at wavelength corresponding to  $2^n \lambda$ .

## 4.2 CD Variation Separation Using Wavelet Packets

In order to get a more accurate separation of CD variation, wavelet packet analysis is used. Wavelet packets provide more flexible decomposition, both the approximation and the detail may be decomposed each time. A design objective can be used to select the required bases to obtain an accurate separation of controllable and uncontrollable variation. Furthermore arbitrary multiples of the initial wavelength can be used for separating the controllable and uncontrollable components. In this case the wavelength division can be any value  $n\lambda$ , where  $\lambda$  is the original wavelength.

### 4.2.1 Frequency Order of Wavelet Packet Nodes

In the process of wavelet packet decomposition, there are two ways to group the wavelet packet coefficients, frequency order and natural order [24].

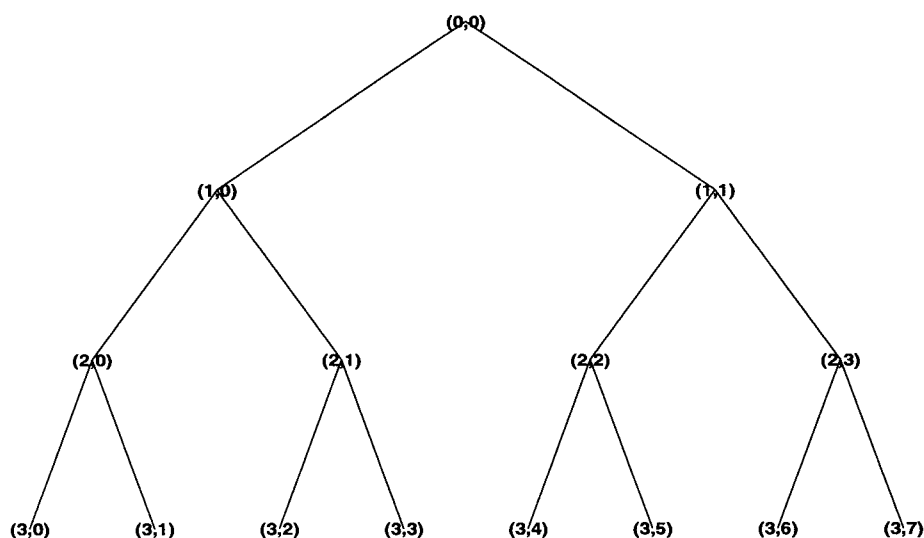


Figure 4.32: Wavelet packet decomposition tree

Consider the three-indexed family of wavelet packet functions  $W_{j,p,k}(x)$  which is defined in Chapter 2, the natural order of a node is the same as its position in the decomposition tree. As shown in Figure 4.32, the natural order of the nodes at level 3 is 0,1,2,3,4,5,6,7 which is same as its position in the tree. The frequency order corresponds to the oscillating property, as can be seen from the *db2* wavelet packets in Figure 4.33,  $W_n(x)$  oscillates approximately  $n$  times.

The frequency order of the wavelet packet function is not the same as the natural order. The frequency order of the nodes can be obtained from the natural order recursively as in Table 4.2. Here for the convenience of calculating the wavelength at each



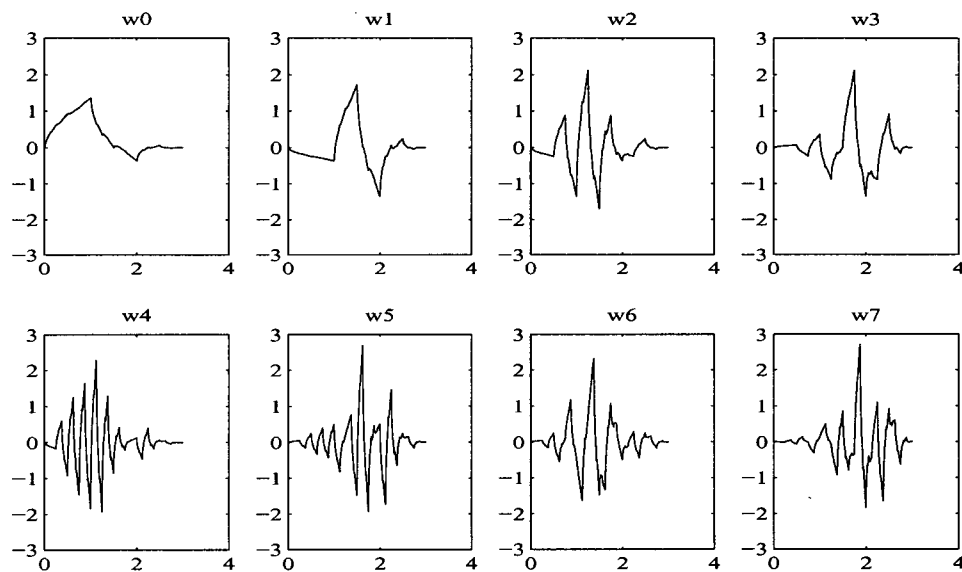


Figure 4.33: The db2 wavelet packets

node and reconstructing the profile at each node, the coefficients are grouped according to the frequency order.

Table 4.2: Natural order and frequency order of wavelet packet nodes

Natural order	0	1	2	3	4	5	6	7
Frequency order	0	1	3	2	6	7	5	4

#### 4.2.2 Restricted Basis Algorithm for Tree Selection

The restricted basis selection algorithm is used here to search among the family of wavelet packet bases. Given the complete rectangle of wavelet packet coefficients down to some level, certain coefficients are excluded for statistical or other reasons [5]. Here the purpose of the restricted basis algorithm is to select those nodes which constitute consecutive wavelength divisions of the original signal.

For a given level, full wavelet packet decomposition is first performed. The wavelet packet nodes are grouped in the frequency order and the wavelength for each node is calculated. Next the wavelet packet separation node whose wavelength includes  $\lambda = 2$  is identified. After the dividing node is found, a group of nodes is found whose wavelengths together with the wavelength of the separation node can constitute a set of consecutive wavelength divisions of the original signal. In this way, a tree structure which satisfies our purpose can be obtained.

The above procedures can also be summarized as: Given the initial wavelength  $\lambda_0$ , the decomposition find  $2/\lambda_0$  and truncates it to an integer, then level  $n$  is found (here  $n = 4$ ) which is the smallest integer that satisfies:

$$2^n \lambda_0 > 2 \quad (4.23)$$

Then the approximation node at level  $n$  and all detail nodes above  $n$  are selected. Next decompose the detail node at level  $n$  and repeat the research until the node whose wavelength is equal to 2 (with the resolution of  $\lambda_0$ ) is reached.

Figure 4.34 shows the resulting wavelet packet decomposition tree based on the restricted basis algorithm. The wavelength and profile variance corresponding to each node are also labeled in the tree diagram. By performing the wavelet packet decomposition according to the tree in Figure 4.34, the signal can be divided into the following consecutive wavelength divisions (units: actuator spacing):

node(1,1): 0.21 to 0.43

node(2,1): 0.43 to 0.86

node(3,1): 0.86 to 1.72

node (7,8): 1.72 to 1.93

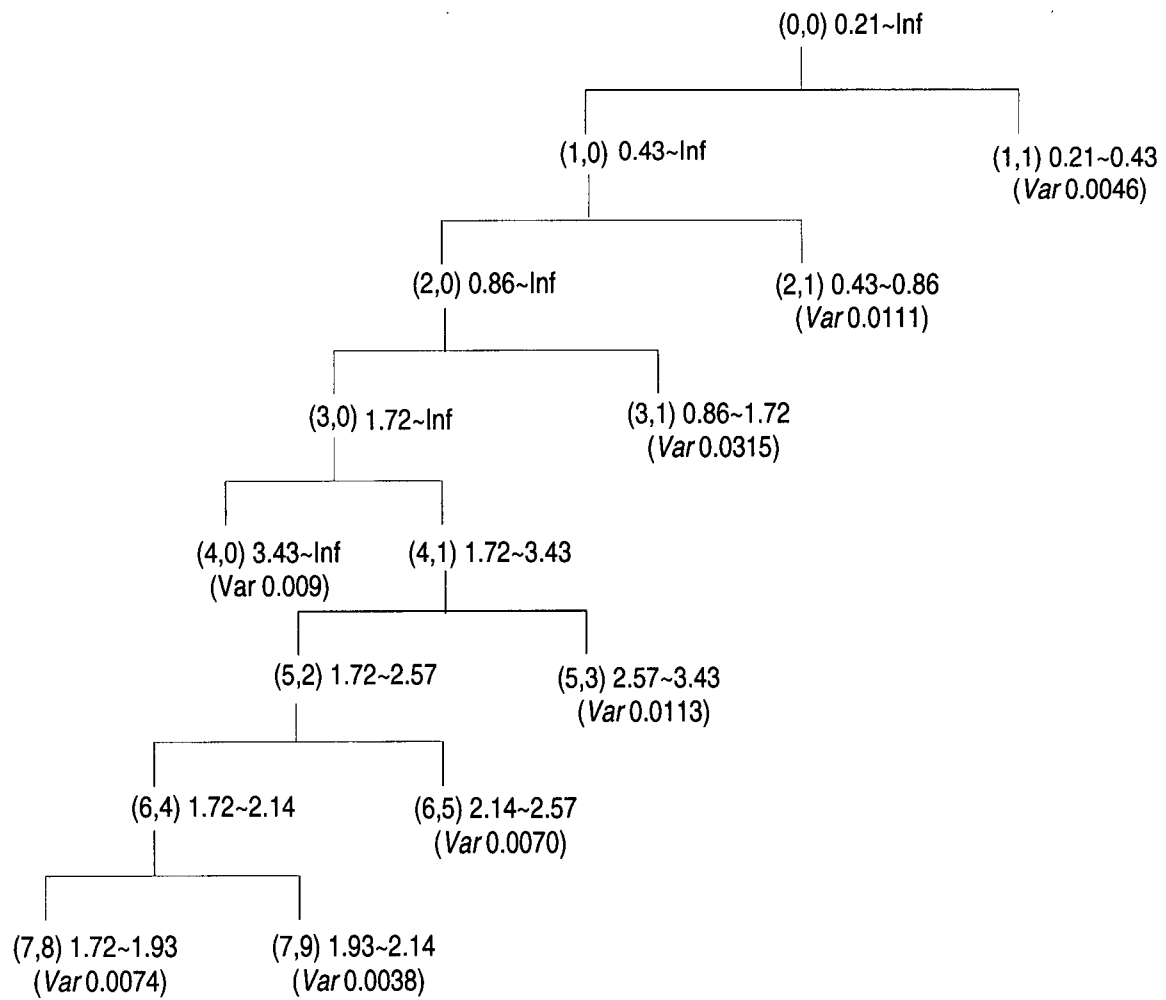


Figure 4.34: Wavelet packet decomposition tree and wavelength at each node

node (7,9): 1.93 to 2.14

node (6,5): 2.14 to 2.57

node (5,3): 2.57 to 3.43

node (4,0): 3.43 to  $\infty$

Here those nodes with wavelength shorter than 1.93 are considered to be uncontrollable, and those nodes with wavelength longer than 1.93 are considered to be controllable. By doing so, it can be seen that the variations of wavelength(node(7,8))  $1.72 \sim 1.93$  should be classified as uncontrollable variation, whereas it is improperly included in the controllable variation when using wavelet decomposition method.

#### 4.2.3 Profile Decomposition at Each Waveband

After the wavelet packet decomposition, the profile in each node is extracted using wavelet packet reconstruction and is shown in Figure 4.35. The variance of the profile in each waveband are summarized in Table 4.3.

Table 4.3: Summary of profile variation at each waveband

Node	Wavelength		STD	Var	Var Percentage(%)
15(4,0)	3.43	$\infty$	0.0949	0.009	10.3
34(5,3)	2.57	3.43	0.1063	0.0113	12.9
68(6,5)	2.14	2.57	0.0837	0.0070	8.0
136(7,9)	1.93	2.14	0.0616	0.0038	4.3
135(7,8)	1.72	1.93	0.0860	0.0074	8.4
8(3,1)	0.86	1.72	0.1775	0.0315	36.1
4(2,1)	0.43	0.86	0.1054	0.0111	12.7
2(1,1)	0.21	0.43	0.0678	0.0046	5.3

The first column shows the node number in the wavelet decomposition tree, the second column is the corresponding wavelength division. The third and fourth column are the

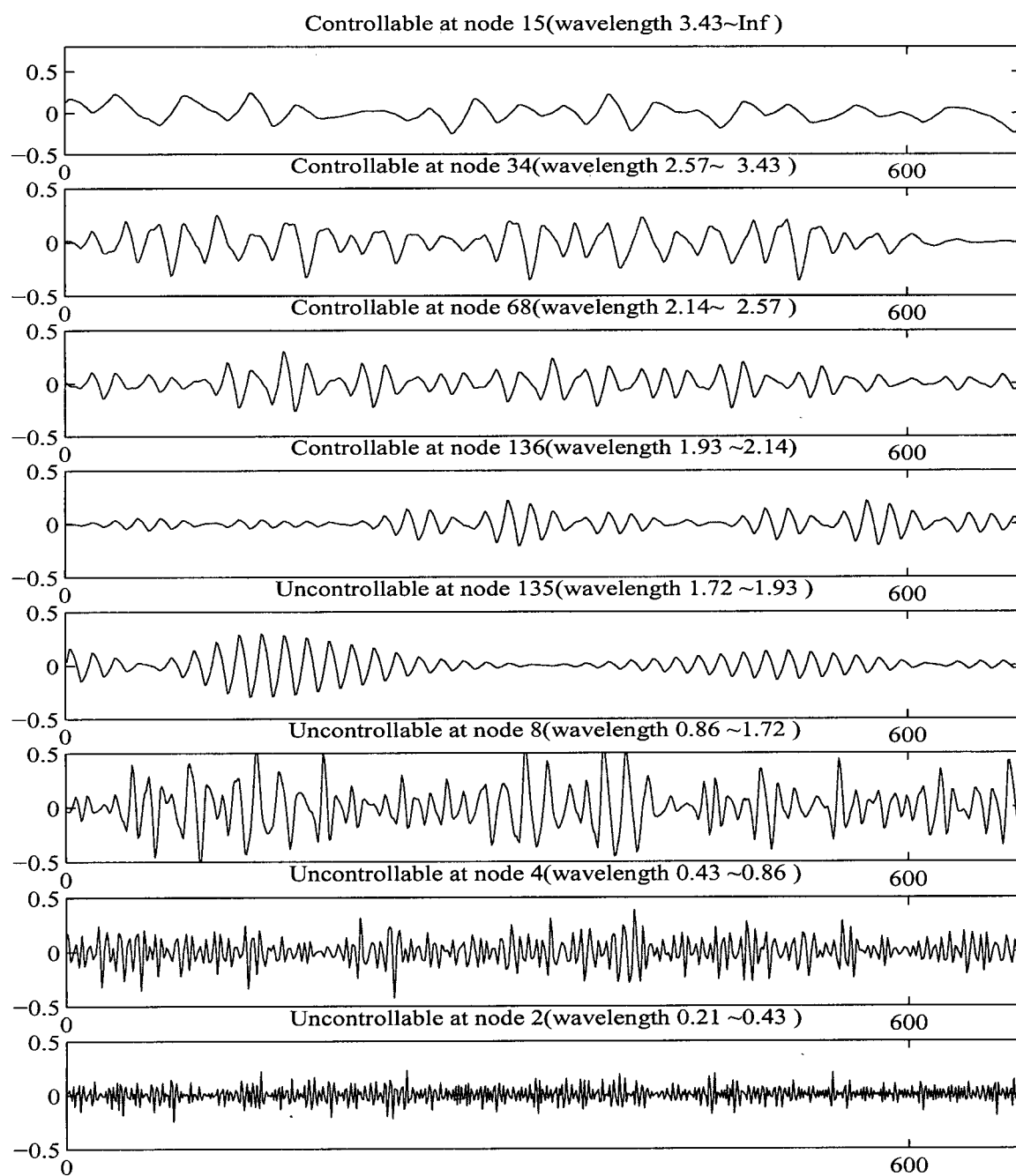


Figure 4.35: Profile wavelet packet decomposition at each node

standard deviation and the variance of the profile decomposition at each node. The last column shows the variance percentage at each waveband. Knowledge of this information is helpful for CD performance monitoring and diagnosis.

#### 4.2.4 Controllable and Uncontrollable CD Profiles

Combining all controllable nodes and uncontrollable nodes in Figure 4.35 respectively, gives the controllable and uncontrollable profiles.

Figure 4.36 and Figure 4.37 illustrate the controllable and uncontrollable profile after separation using wavelet analysis and wavelet packet analysis respectively.

Based on the analysis for each scan, the controllable and uncontrollable profiles for the entire jumbo reel are shown in Figure 4.38. (a) and (b) are the controllable and uncontrollable profiles using wavelet packet analysis for the entire reel, (c) and (d) are the controllable and uncontrollable profiles using wavelet analysis. These pictures (normally displayed in color) are very helpful for the operators, and useful for the process retrieving and monitoring. It can be seen that some streaks in (c) are actually uncontrollable, so they are included in the uncontrollable profile properly shown in (b) by using wavelet packet analysis.

Figure 4.39 to 4.42 display images of the profiles in Figure 4.38.

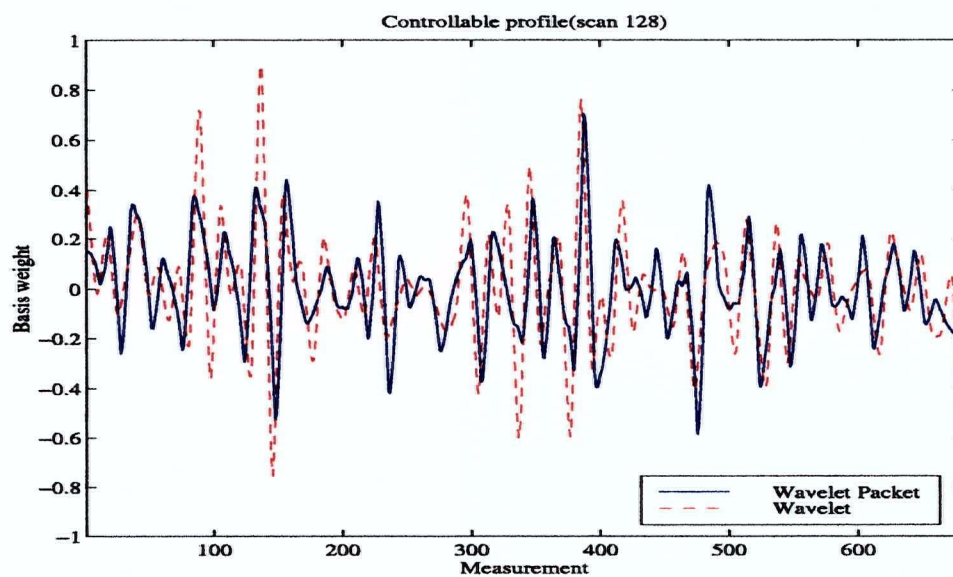


Figure 4.36: Controllable profile estimates using wavelet and wavelet packets

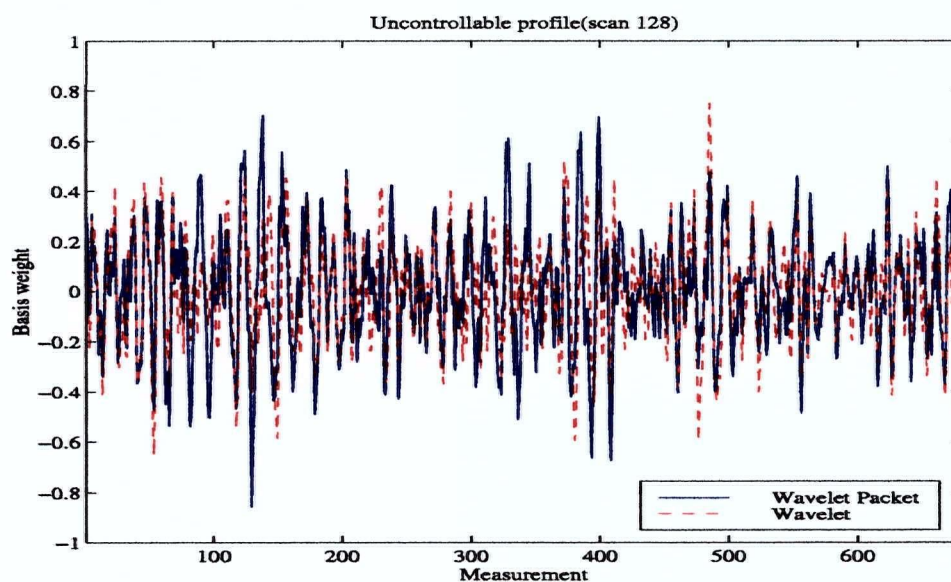


Figure 4.37: Uncontrollable profile estimates using wavelet and wavelet packets

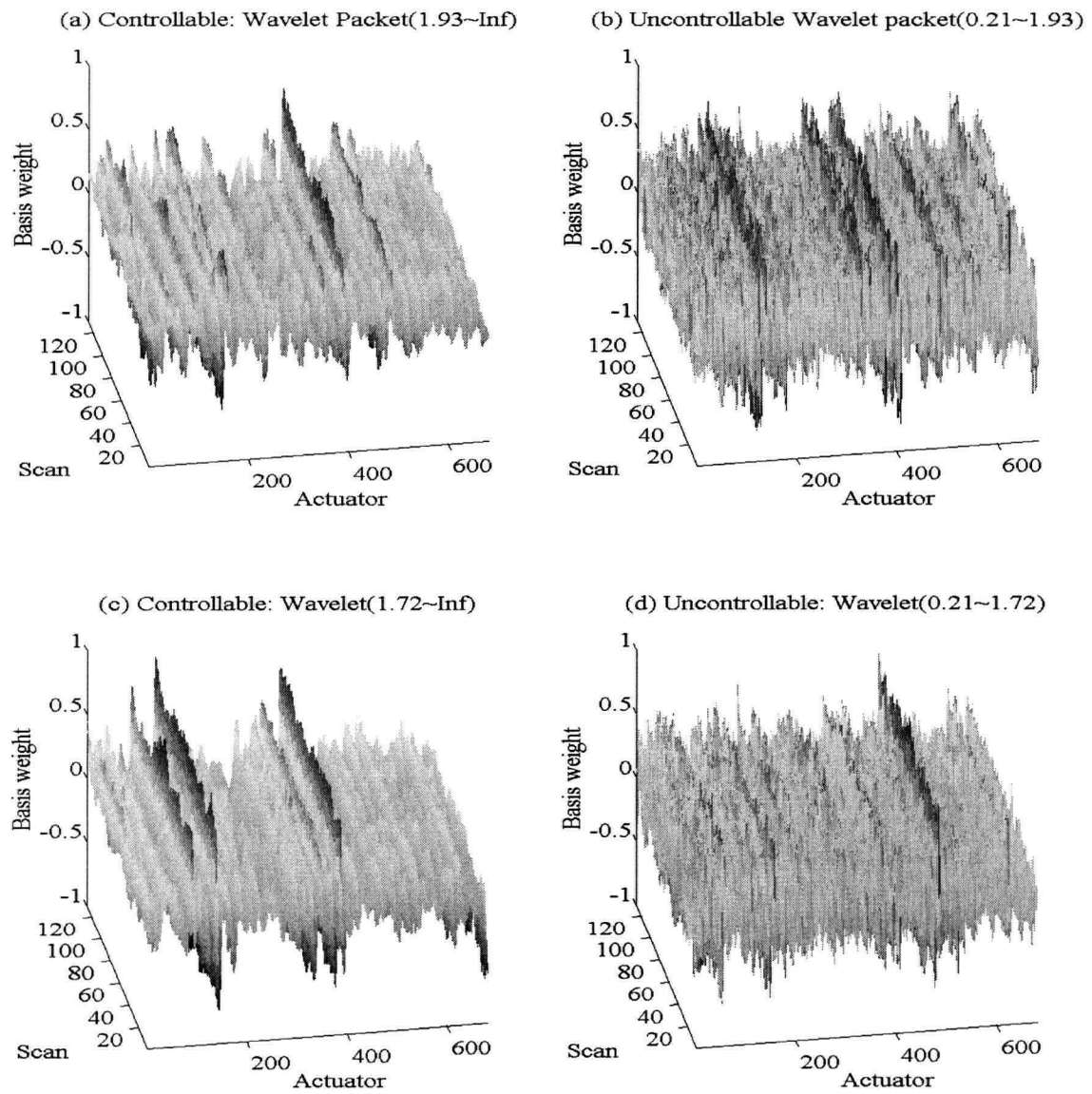


Figure 4.38: CD profile separation using wavelet and wavelet packets



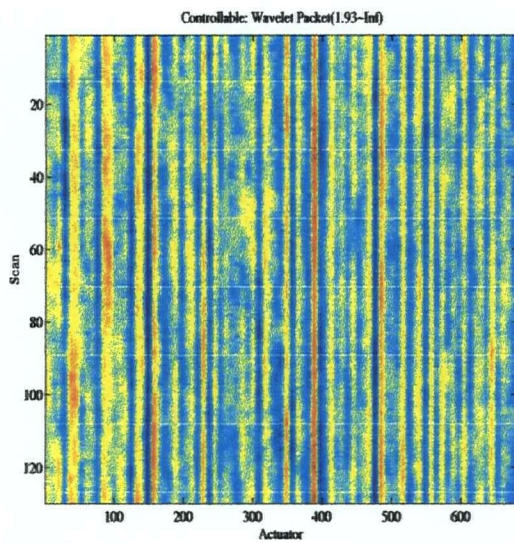


Figure 4.39: Controllable: Wavelet Packet

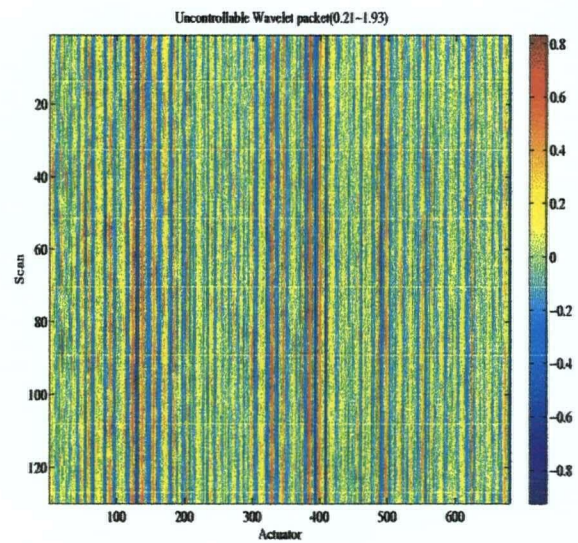


Figure 4.40: Uncontrollable: Wavelet Packet

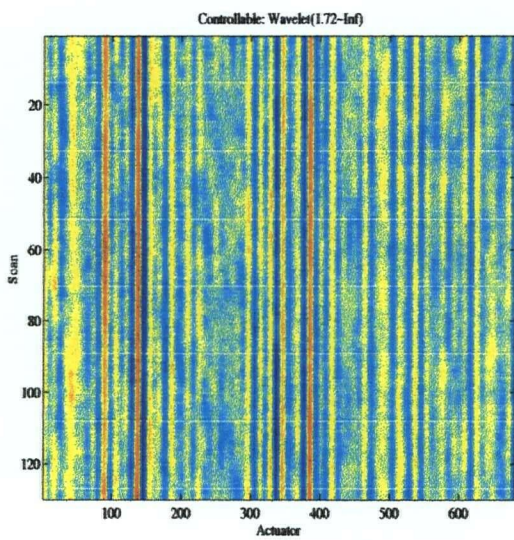


Figure 4.41: Controllable: Wavelet

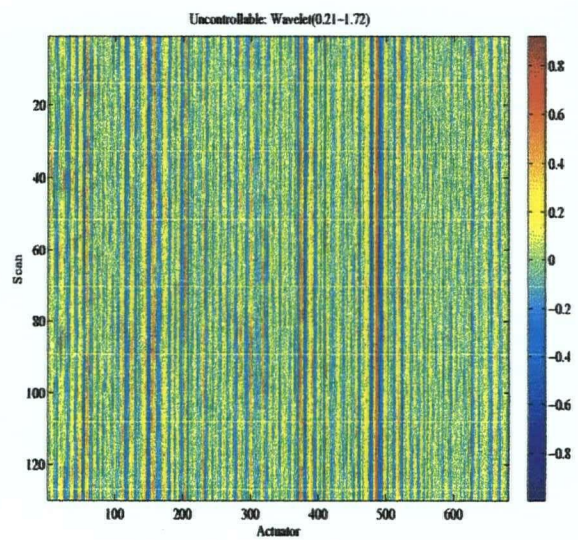


Figure 4.42: Uncontrollable: Wavelet

### 4.3 Performance Assessment Using Wavelet and Wavelet Packet Analysis

After the controllable and uncontrollable separation, the potential improvement in control that can be expected for the given actuator spacing can be evaluated in terms of performance index defined as follows.

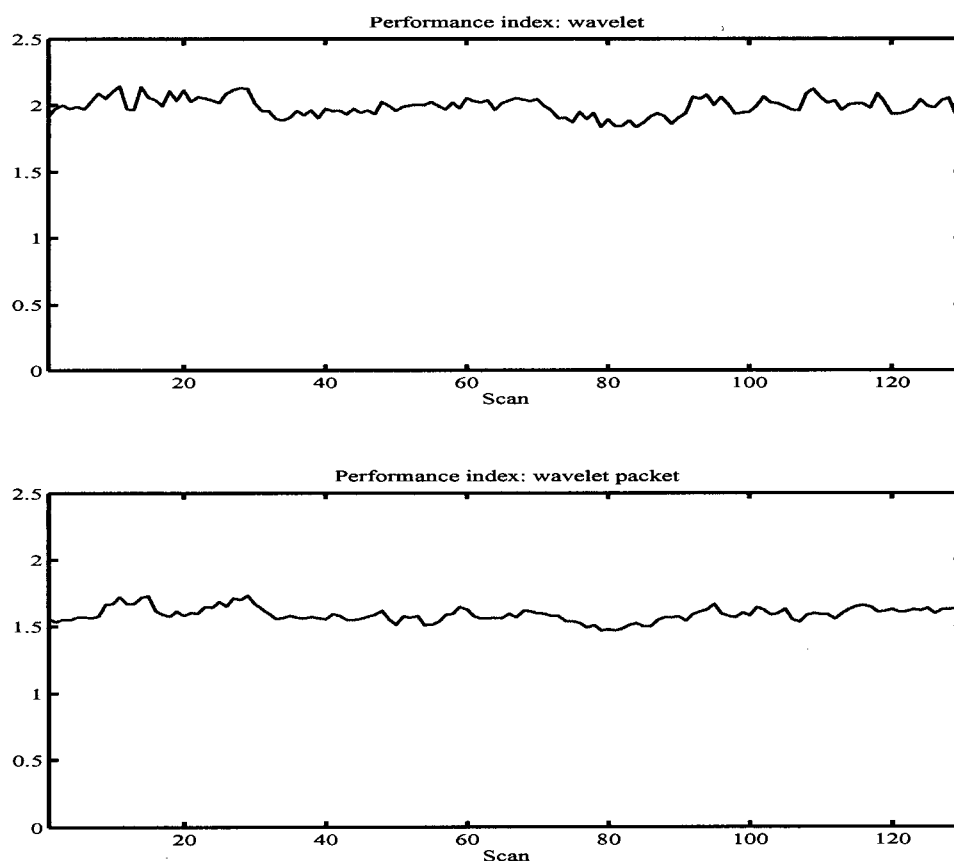


Figure 4.43: Performance index using wavelet and wavelet packets

Let  $\sigma_{pr}^2$  be the noise-free overall profile variance and  $\sigma_{contr}^2$  be the variance of the controllable component. The performance index can be calculated as [29]:

$$C = \frac{\sigma_{pr}^2}{\sigma_{pr}^2 - \sigma_{contr}^2} \quad (4.24)$$

If  $C = 1$ , perfect control is achieved, meaning that all controllable variations have been removed.  $C < 1$  is not possible since  $\sigma_{contr}^2 < \sigma_{pr}^2$ .  $C > 1$  means the controllable variations are still present in the profile. The following cases of combination of  $\sigma_{pr}^2$  and  $C$  are considered:

- If both  $\sigma_{pr}^2$  and  $C$  are low, the control performance is good.
- If  $\sigma_{pr}^2$  is low but  $C$  is high, the potential of improving control is marginal.
- If  $\sigma_{pr}^2$  is high but  $C$  is low, profile variability can only be reduced through process correction.
- If both  $\sigma_{pr}^2$  and  $C$  are high, there is a large potential to improve control.

The performance index provides a straightforward quantitative assessment of the process operating condition. It can be calculated on-line once the scan data is available. Together with the profile decomposition plots, the operator can be provided with a quick and effective assessment of manufactured reel and the control performance. Before calculating the performance index, first the very high frequency component due to formation noise is removed.

Figure 4.43 shows the performance index plot for basis weight using wavelet (a) and wavelet packets (b). It can be seen that the performance index value obtained using wavelet packet analysis is smaller (around 1.6) than the performance index value obtained using wavelet (around 2). This is because most of the uncontrollable variations (wavelength  $1.72 \sim 1.93$ ) improperly included in the controllable variations when using

wavelet decomposition are properly considered to be uncontrollable variations when using wavelet packet analysis. That is, the denominator in (4.24) increases, so the value of  $C$  is smaller in this case.

To summarize, wavelet packet analysis yields an accurate assessment of process performance. The value of  $C$  could be smaller or larger than when using wavelet analysis, depending on the characteristics of the profile property and the actuator spacing. For example, after carrying out similar analysis for the caliper data, wavelet packet analysis gives a larger value for the performance index  $C$  than wavelet analysis (see Figure 4.44). Figure 4.45 and Figure 4.46 also give the decomposition tree for caliper data using wavelet and wavelet packet analysis respectively.

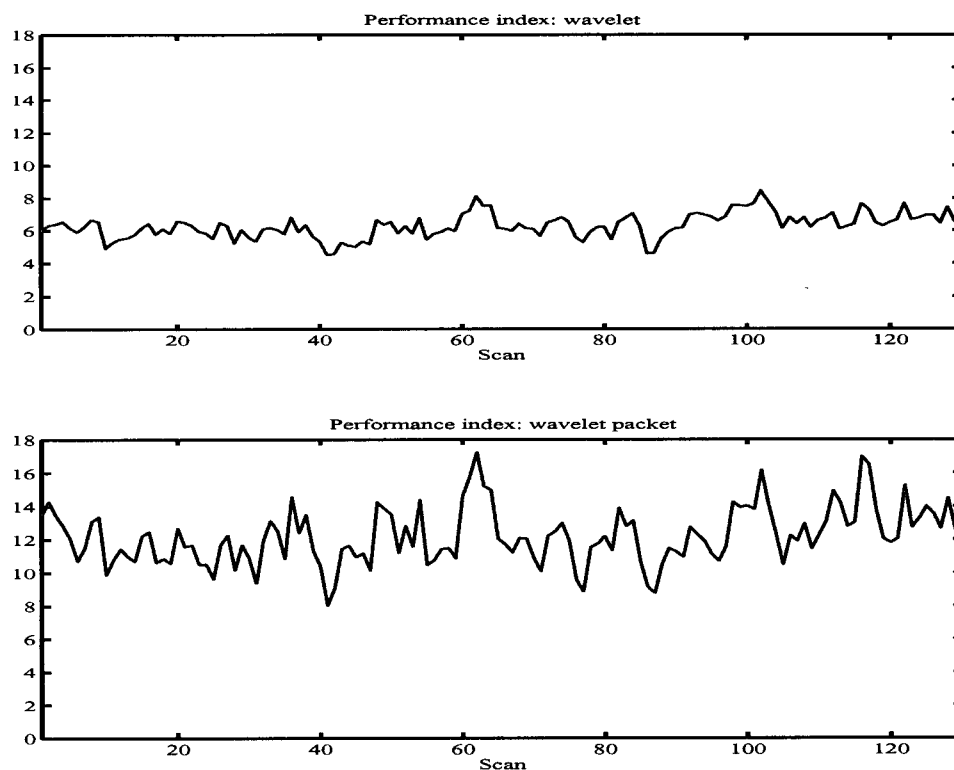


Figure 4.44: Performance index using wavelet and wavelet packets(Caliper)



Figure 4.45: Wavelet decomposition tree and wavelength at each node(Caliper)

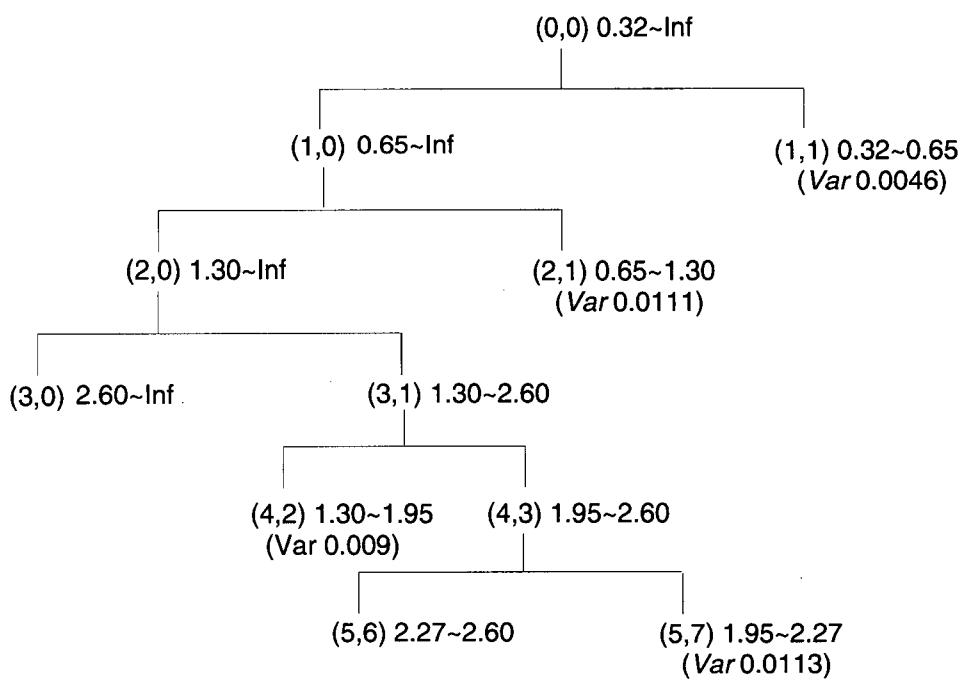


Figure 4.46: Wavelet packet decomposition tree(Caliper)

#### 4.4 Denoising

After wavelet and wavelet packet decomposition, the coefficients are used to estimate the noise at each level. MultiMAD level dependent thresholding method is used to select the threshold at each resolution level based on the noise strength estimation because the noise is often correlated in practice. Only the detail coefficients (wavelets) and uncontrollable nodes (wavelet packet) are thresholded. The threshold for wavelet packet is  $\sqrt{2 \log_e(n \log_2(n))}$ . After thresholding, the denoised wavelet and wavelet packet coefficients are reconstructed to give the denoised profile.

No distinct difference is found between the denoised profile using wavelets and wavelet packets.

#### 4.5 Compression of Industrial Data

In investigating the wavelet and wavelet packet compression, the basis weight data is tested. The theoretical background for wavelet compression and wavelet packet compression has been discussed in Chapter 2. The compression can be summarized into three steps: decomposition, thresholding and reconstruction.

Figure 4.47 plots the wavelet coefficients before and after thresholding for scan 128. The decomposition level is 3, and the wavelet coefficients are stored in a vector of length 1024, according to the order of *app3*, *det3*, *det2*, *det1*, where *app3* represents the wavelet approximation coefficients at level 3, *det3*, *det2*, *det1* represent the detail wavelet coefficients at level 3, level 2, level 1 respectively. MultiMAD thresholding method is used here because it can adjust the threshold value at different resolution levels.

An 8-bit quantizer is used. Let the variable *BITS* represent the number of bits for the quantizer. For the data set X, the quantization can be achieved according to the following procedure:

1. find  $X_{min} = \min(X)$ ,  $X_{max} = \max(X)$
2.  $\Delta X = (X_{max} - X_{min}) / (2^{BITS} - 1)$
3.  $X1 = X - X_{min}$
4.  $X_{quan} = \text{floor}(X1 / \Delta X)$

*floor* is the operation which rounds the given number to the nearest integer smaller or equal to the number.

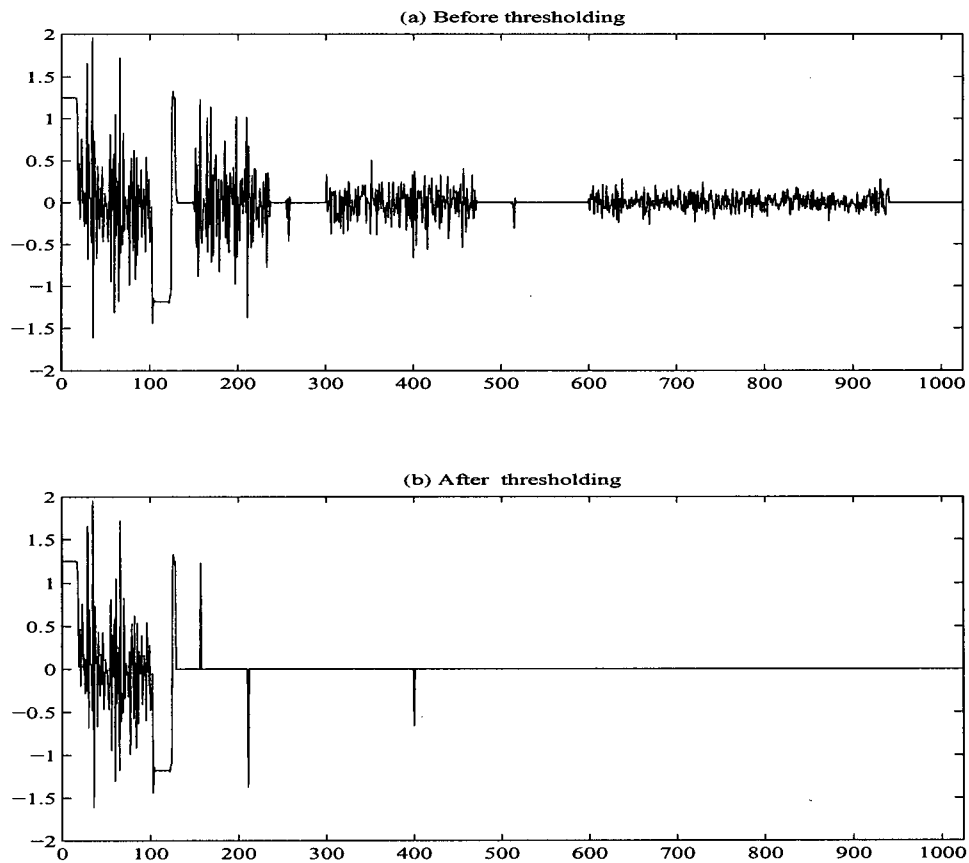


Figure 4.47: Wavelet coefficients before and after thresholding

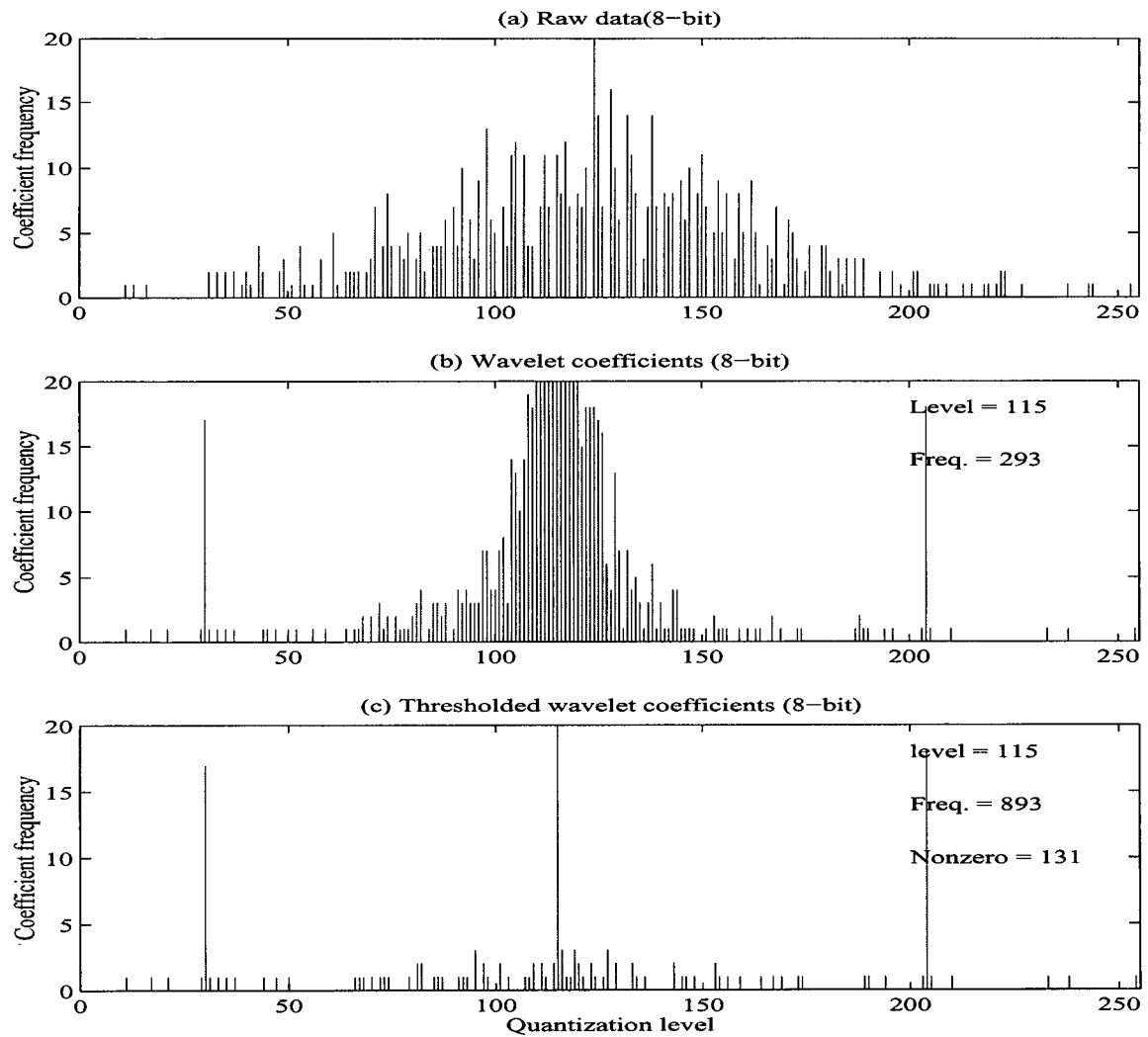


Figure 4.48: Scan 128: Energy distribution



In Figure 4.48, (a) is the energy distribution of scan 128 of basis weight data. (b) and (c) show the quantized wavelet coefficients before and after thresholding respectively. The wavelet coefficients are much concentrated at quantization level 105 to 125. There are 293 coefficients at the quantization level 115, which correspond to the zero values in the sparse matrix for the wavelet coefficients. After thresholding the occurring frequency at level 115 is 893 and there are only 131 nonzero coefficients. So a high compression is achieved.

In the case of wavelet packet compression, the entropy-based best basis algorithm is used here. Results on a paper machine simulation data done in [4] show that wavelet packets compressor can achieve a more efficient compression than wavelets in terms of compression ratio and mean square error.

Because wavelet packets can represent the signal at least as efficiently as wavelets, they normally gives a better compression than wavelets.

Table 4.4: Compression of industrial data

Properties	Wavelet	Wavelet packet
Compression ratio	17.6	23.86
Energy retained	82.63	87.02

Table 4.4 lists the compression result of the two-dimensional basis weight CD profile using wavelets and wavelet packets in terms of compression ratio and energy retained. Wavelet packets analysis results in a better compression performance.

## **Chapter 5**

### **Trim-Loss Optimization**

#### **5.1 Introduction**

The trim-loss problem arises when a large stock is cut into small pieces of order sizes required by customers. This problem may arise in many manufacturing industries and is particularly common in the paper, glass and steel industries. The cutting process can have severe impact on the company's profit, especially when material of high value is involved. Because poor cutting may result in a large amount of trim-loss, which means that the material and production resources must be recycled.

This chapter does not emphasize the optimization solving methods, however it uses a simple optimization model and addresses the paper quality issue during the optimization by dividing the reel into usable and un-usable sections using the measured sheet quality. The approach builds a connection between the sheet profile data analysis and production optimization in the paper manufacturer. First, based on the data analysis result from Chapter 3 and Chapter 4, some criteria are used to evaluate the paper quality. Then the usable sections in the paper reel are identified and the trim-loss problem is solved while avoiding the un-usable sections of the jumbo reel.

The outline of this chapter is: Section 2 gives an overview of the trim-loss problem followed by the description of the optimization model in Section 3 . Section 4 describes the Lingo Modeling language and solver which we are going to use. Section 5 and 6 proceed to solve the trim-loss optimization problem with quality consideration based on

the filtered sheet property. Three optimization schemes are studied and compared.

## 5.2 Overview of Trim-Loss Optimization

The trim-loss problem is among the earliest problems in the literature of operational research. The problem is: given the stock size, how should the stock be cut into pieces of dimensions required by the customers, the order sizes?

According to the dimension of the problem, the trim-loss problem can be categorized into 1-dimensional problem,  $1\frac{1}{2}$ -dimensional problem and 2-dimensional problem [20]. 1-dimensional problem is one in which one dimension of the order pieces is significant, the solution is not relevant to the other dimension of the stock.  $1\frac{1}{2}$ -dimensional problem is referred to the problems in which the stock has one fixed, one varied dimension, both of which are relevant to the solution. While a 2-dimensional trim-loss problem is one in which the order pieces are rectangular and dimensions in the two orthogonal directions are significant in the determinations.

The methods for solving trim-loss problems fall into two groups: algorithmic methods and heuristic methods. The algorithmic method for trim-loss problems are linear programming, branch-and-bound and dynamic programming [20]. The best algorithmic methods are the linear programming procedures of Gilmore and Gomory [16, 15] which were adapted subsequently for developing better solutions. However these methods are computationally prohibitive and they only consider trim-loss as the cost function.

G. Wascher [35] extended the traditional linear programming formulation into a linear formulation with multiple objectives and solved it using Multiple Criteria Decision Making. This method used a total-cost minimization model which included the material costs, overrun costs, left-over costs and trim-loss costs.

If an algorithm is not available or the computational cost of the available algorithm

is large, heuristic methods are usually adopted. A heuristic method, on the other hand cannot be guaranteed to find the optimal solution and often will not. A heuristic is often regarded as acceptable once it is shown to be "good enough" [20].

The heuristic methods have the following features:

1. An aspiration level is used to determine whether a cutting pattern that was found should be used.
2. Once a cutting pattern has been used, it will be used as often as possible.
3. The use of heuristic values can take into account factors such as machine set-up.

However the solution is often problem-dependent, often two similar problems may require different techniques.

In recent years the trim-loss problem in the paper industry has been solved with the aid of linear programming combined with heuristic rules in order to handle the non-linearities and discrete decisions. Efforts to solve the non-linear problem with heuristic methods were made for example by Haessler(1971) and Johnston(1979). The complete survey can be found in Hinxman [20](1980) whose taxonomy and terminology is used widely here. Since then, there have been few subsequent articles which consider quality in the context of a cutting stock problem.

P.E. Sweeney and R. W. Haessler [33] proposed a procedure for solving one-dimensional cutting stock problem when both the master rolls and customer orders have multiple quality gradation using a two-stage sequential heuristic. The objectives are to minimize trim-loss, avoid production overrun and avoid unnecessary slitter setups. At the first stage, the slitting decisions were made for the non-perfect master rolls. At the second stage, an LP model is solved for the remaining demands using the available perfect master rolls.

In most cases, the practical problem formulation of the trim-loss problem are restricted by the fact that the solution methods should handle the entire problem. Thus

only a suboptimal solution has been obtained. Recently, I. Harjunkski and T. Westlund formulated the trim-loss problem in several formulas which minimize the raw material used, processing time and production waste. They compared different methods of describing the problem as an Integer Linear Programming/Multi-Integer Linear Programming problem [37, 36, 19].

### 5.3 Mathematical Model for Trim-Loss Problem

In the paper-converting industry, the objective is to produce a set of paper rolls from storage rolls such that a cost function is minimized. Different criteria can be used as cost function, including absolute trim-loss, cutter utilization and cutting time. The trim-loss problem is considered to be a  $1\frac{1}{2}$ -dimensional problem, since the length of order rolls is assumed to be fixed and the width is variable. For simplicity, here we only consider trim-loss as the cost function of the optimization.

Figure 5.49 illustrates a trim-loss example with three products and three cutting patterns. A cutting pattern  $j$  is defined as a set of integers  $n_{ij}$  that tells how many product rolls are to be cut from the raw paper reel using the pattern  $j$ .

Let  $J$  be the total number of cuts,  $n_{ij}$  be the number of rolls  $i$  in cut  $j$ . Each type of roll corresponds to a given width  $w_i$ . Let us assume the length of each roll is the same, the total width of each cut  $j$  is between  $W_{j,min}$  and  $W_{j,max}$  and the total number of rolls in each cut is at most  $N_{j,max}$ , which is a physical restriction to possibly separate the rolls from each other [19].

Based on the formulation used by [19], the trim-loss problem can be formulated as:

$$\text{optimize} \quad \sum_{j=1}^J (W_{j,max} - \sum_{i=1}^I n_{ij} W_i) m_j \quad (5.25)$$

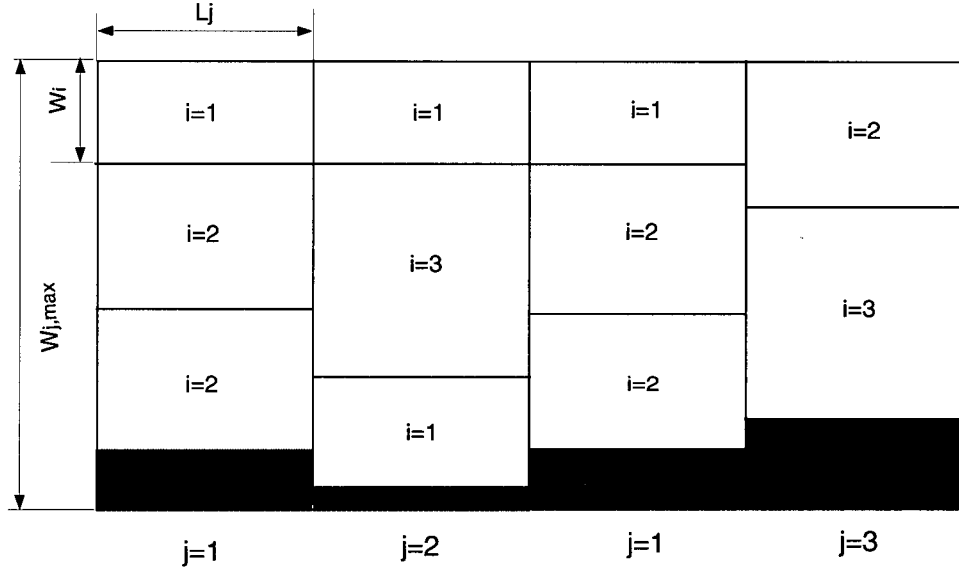


Figure 5.49: An example with 3 products ( $i=1,2,3$ ) and 3 cutting patterns ( $j=1,2,3$ )

$$\text{subject to } \sum_{i=1}^I W_i n_{ij} - W_{j,max} \leq 0 \quad \text{for all } j \quad (5.26)$$

$$- \sum_{i=1}^I W_i n_{ij} + W_{j,min} \leq 0 \quad \text{for all } j \quad (5.27)$$

$$\sum_{i=1}^I n_{ij} - N_{j,max} \leq 0 \quad \text{for all } j \quad (5.28)$$

$$- \sum_{j=1}^J m_j n_{ij} + N_{i,min} \leq 0 \quad \text{for all } i \quad (5.29)$$

$$\sum_{j=1}^J m_j n_{ij} - N_{i,max} \leq 0 \quad \text{for all } i \quad (5.30)$$

where

$W_i$ : width of roll  $i$ ;

$W_{j,max}$ : the maximal width that can be used in pattern  $j$ ;

$n_{ij}$ : the number of roll  $i$  in pattern  $j$ ;

$N_{j,max}$ : the maximal number of rolls per pattern;

$N_{j,min}$ : the minimal number of rolls per pattern;

$N_{i,max}$ : the maximal demand for roll  $i$ ;

$N_{i,min}$ : the minimal demand for roll  $i$ ;

$m_j$ : multiple for each cut type  $j$ .

This problem is non-convex mixed integer non-linear programming problem (MINLP) subjected to the bilinear constraints and bilinear cost function in the formulation above. The problem is often solved as a two-step optimization procedure in which the first step is to generate cutting patterns and the second step is a mixed integer linear programming [36]. Next we are going to use the Lingo optimization modeling environment to solve this model.

#### 5.4 Lingo Optimization Solver

In this chapter, the main concern is to avoid the unacceptable areas in the jumbo reel during the trim optimization process. The non-perfect areas in the reel are regarded as un-usable, and only those usable sections are available for optimization. The optimization is modeled and solved using the Lingo Optimization Modeling Language and Solver.

Lingo is an optimization modeling tool developed by Lindo System Inc., which allows users to utilize the power of linear and nonlinear optimization to formulate large problems concisely, solve them and analyze the solution. The model can be easily formulated, solved and modified. The correctness of the solution can also be assessed directly.

Windows version Lingo4 is used here because of its following features [1]:

##### *Powerful modeling environment*

Using Lingo, one can express a complex model in a simple and efficient manner.

The model is highly readable and easy to edit.

##### *Versatile lingo solver*

In order to solve different types of models, Lingo provides four solvers which are: a direct server, a linear server, a nonlinear server and a branch-and-bound server for integer restriction.

Lingo will call different solvers for different problems. It reads the model formulation and automatically selects the appropriate solver for the user. For example, if the model contains any integer restrictions, the branch-and-bound manager is invoked to enforce them. The branch-and-bound manager will, in turn, call either the linear or nonlinear solver depending upon the nature of the model.

#### *Easy data handling*

With Lingo, the data can be stored in a variety of convenient ways. The data can be embedded directly in the model or in a table or list from a separate file. So accessing the data is much easier. In the trim optimization of next section, Lingo reads the data evaluation result which is generated by Matlab application programs. It also saves the optimization results into a text file which is easy for Matlab to access.

#### *Links to other application*

Using Lingo, it is easy to interface with databases and spreadsheets. Together with its callable interfaces, it can be easily integrated with other Windows development tools. In this thesis, the Lingo Dynamic Link Library(Lingo.dll) is called by Visual Basic to perform the optimization solving. Also, by using real-time OLE(Object Linking and Embedding), Lingo imports the information from and exports the cutting result to Microsoft Excel spreadsheets.



### 5.5 Trim-Loss Optimization Using Visual Basic and Lingo

Given the entire jumbo reel, the task of the trim optimization here is to cut the reel into different rolls according to the order sizes based on the measured sheet properties. The paper quality is evaluated based on the two sigma criterion( 2 standard deviation) for the sheet property(basis weight).

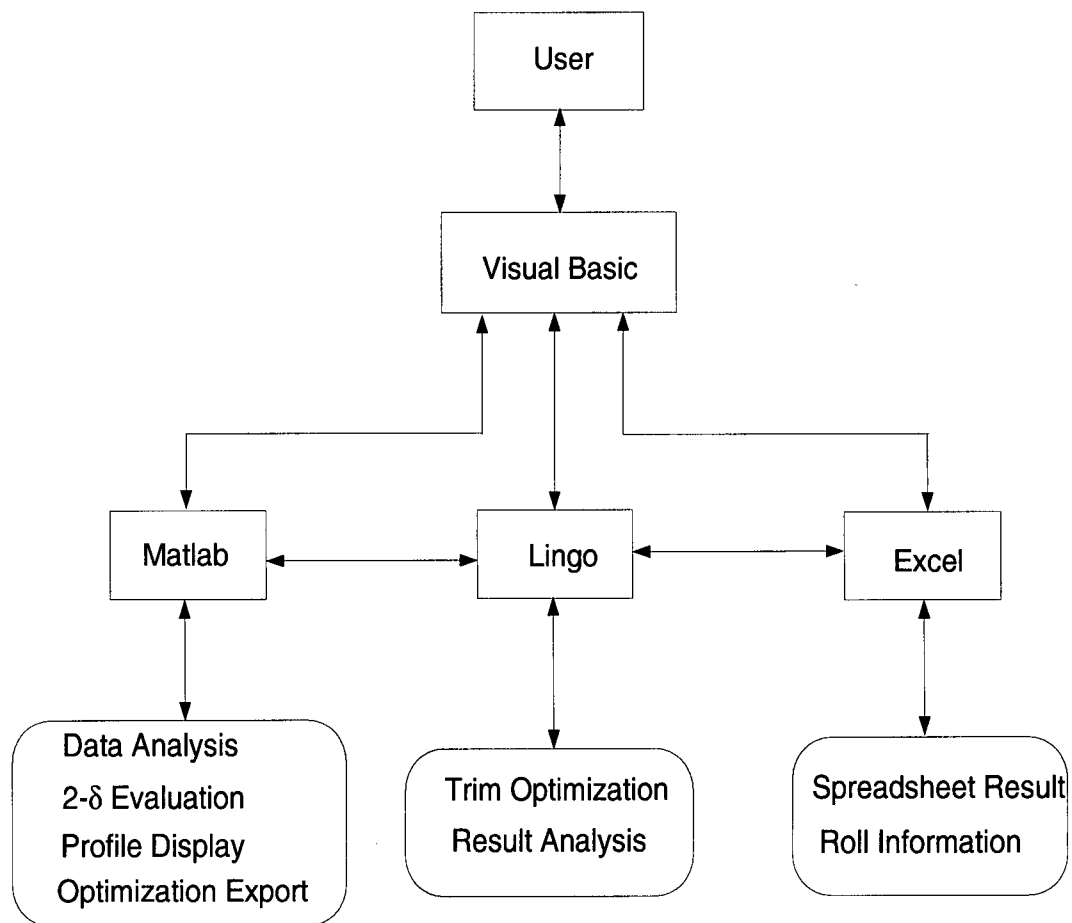


Figure 5.50: Trim optimization diagram

First, given the measured sheet property profile for the entire reel, one level wavelet decomposition is carried out and the very high frequency noise due to formation and

measurements is removed. Then the average property and the  $2\sigma$  value of the reel are calculated. Then the  $2\sigma$  criterion is used to evaluate the paper quality.

Three optimization schemes are studied. They are: trim-loss optimization without paper quality consideration, trim-loss optimization based on the  $2\sigma$  evaluation of sheet property before improving control and after improving control. The results from the three schemes are compared and the potential benefit of savings from better control is analyzed.

**Trim Optimization Example Using VB and the LINGO DLL**

File Quality Assess Solve Show Result

**Input Information**

	width	minDemand	maxDemand
Roll1	50	7	9
Roll2	30	5	7
Roll3	40	4	8

**Optimization Result**

Optimization Solution			
Section	Roll1	Roll 2	Roll 3
Section 1	0	1	0
Section 2	1	1	0
Section 3	0	1	0
Section 4	0	1	1
Section 5	1	0	0
Section 6	0	1	0
Section 7	0	1	1
Section 8	0	1	0

Load Calc Solve Result Export Quit

Figure 5.51: Trim optimization interface using Visual Basic

Several different application software tools are used together to accomplish the trim optimization. One is Lingo, which provides the optimization modeling environment and optimization solver. Another is Matlab, which implements the wavelet filtering of the sheet property profile to evaluate the paper quality. Matlab also performs some important functions including  $2\sigma$  evaluation, DWT, IDWT and graphical display. Furthermore, the roll order information needs to be handled by the user. The basic flow diagram is shown in Figure 5.50.

In order to integrate different application programs together and fully exploit the high computation performance and visualization tool of Matlab and the strong optimization facility in Lingo, an application interface using Visual Basic has been designed as shown in Figure 5.51.

The optimization model is written in Lingo optimization modeling language which specifies constraints and the cost function. The user can enter the roll order information such as roll width, minimal demand, maximal demand, from the input information frame of the interface. This information is passed to the Lingo solver through the use of Visual Basic script. The optimization is solved by Lingo solver and the results are saved in the data files and also into Microsoft Excel spreadsheets through Object Link Embedding.

Several command buttons are also available, the functions of which are introduced briefly according to the general operating procedure:

*Load* Load the optimization model, including:

1. Model1—Model for traditional trim-loss optimization, in which the sheet quality is not considered in the optimization.
2. Model2—Trim loss optimization model based on the sheet property evaluation before improving control.
3. Model3—Trim loss optimization model based on the sheet property evaluation

after improving control.

*Calc* Calculate the  $2\sigma$  and evaluate paper quality based on the  $2\sigma$  criterion.

*Solve* Solve trim-loss optimizations for the corresponding model using Lingo solver, each time the Lingo Dynamic Link Library(DLL) is called.

*Result* Show the optimization result in Excel spreadsheet.

*Export* Plot the trimming optimization result graphically from Matlab with the sheet property displayed.

*Quit* Exiting from the current application.

## 5.6 Trim Optimization Schemes

Three optimization schemes shown in Figure 5.52 are studied in this section. First the traditional trim-loss problem is studied without paper quality consideration. Second the trim-loss problem is studied based on the  $2\sigma$  evaluation of the basis weight. The third scheme is similar to the second one except the paper quality is assumed to have been improved through better control to illustrate the impact of improved basis weight control on the trim-loss optimization results. The optimization results from the three schemes are compared in terms of the total amount of waste paper, so the potential saving from improving control can be easily observed.

Three different roll order sizes are used and the roll information including the width, minimal and maximal demand for each roll are specified in Table 5.5. All rolls are of the same length.

Using the application programs and interface described earlier in the last section, the three trim-loss optimization schemes in Figure 5.52 are tested.

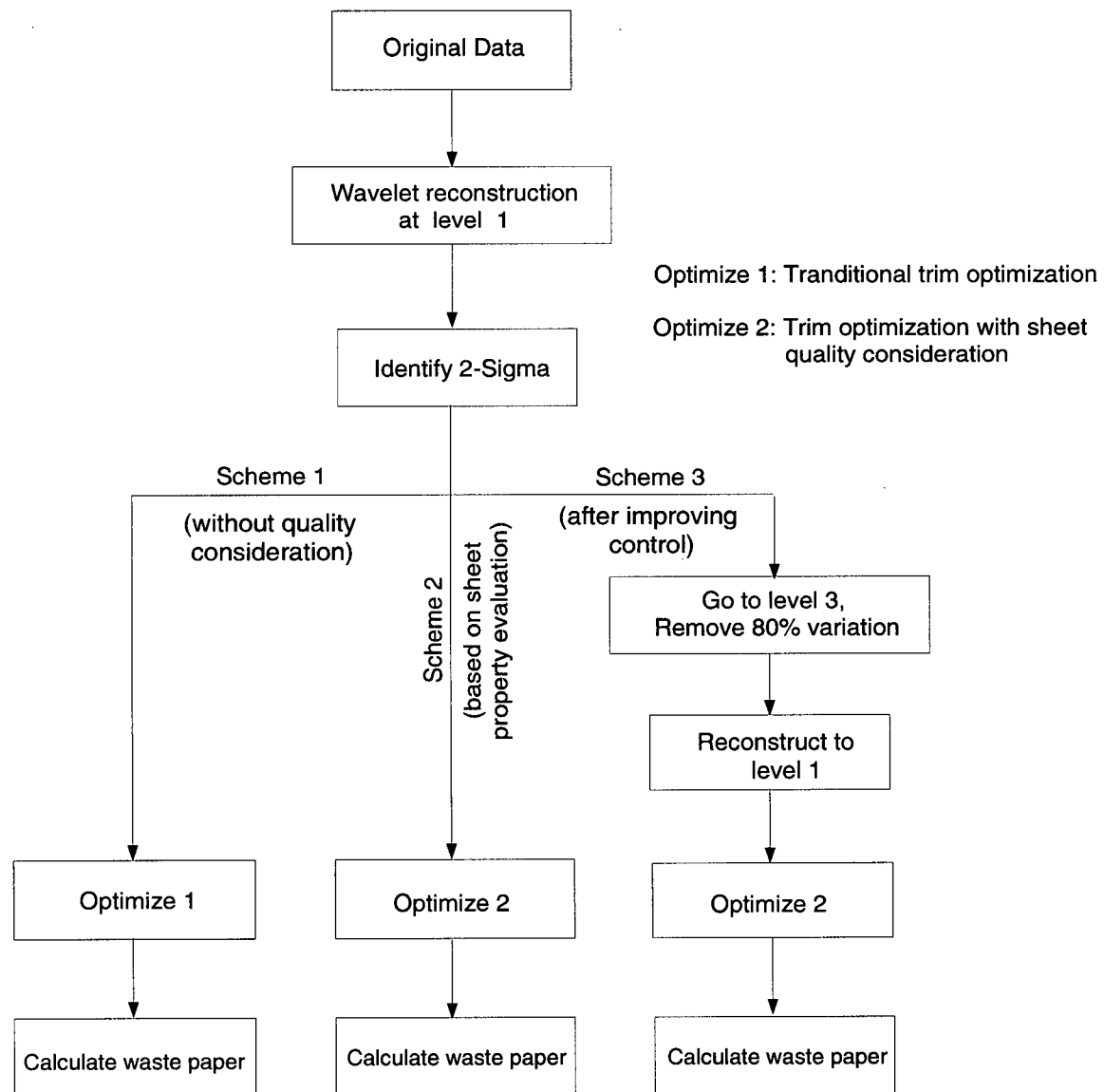


Figure 5.52: Three trim optimization schemes

Table 5.5: Roll order information

Roll	Width	Minimal Demand	Maximal Demand
Roll 1	50	7	9
Roll 2	30	5	7
Roll 3	40	4	8

### 5.6.1 Scheme 1: Trim Optimization without Paper Quality Consideration

In this scheme, the traditional trim-loss problem is solved without considering the paper quality.

Table 5.6: Trim optimization result: Scheme 1

Cut	Number of Roll 1	Number of Roll 2	Number of Roll 3
Cut 1	4	0	1
Cut 2	0	7	1
Cut 3	5	0	0
Cut 4	0	0	6
Total Number	9	7	8

This scheme seems to be the best if we ignore the paper sheet quality specification, because the number of produced rolls of each size is 9, 7, 8 respectively (see Table 5.6). However, many rolls produced using this scheme need to be discarded because the quality does not meet the specification ( $2\sigma$  here). It can be seen clearly from Figure 5.53: all the four *Roll 1* produced in the first cut and four *Roll 2* produced in the second cut will be discarded due to the quality problem. So the amount of waste paper is huge. Detailed information can be found in Table 5.9.

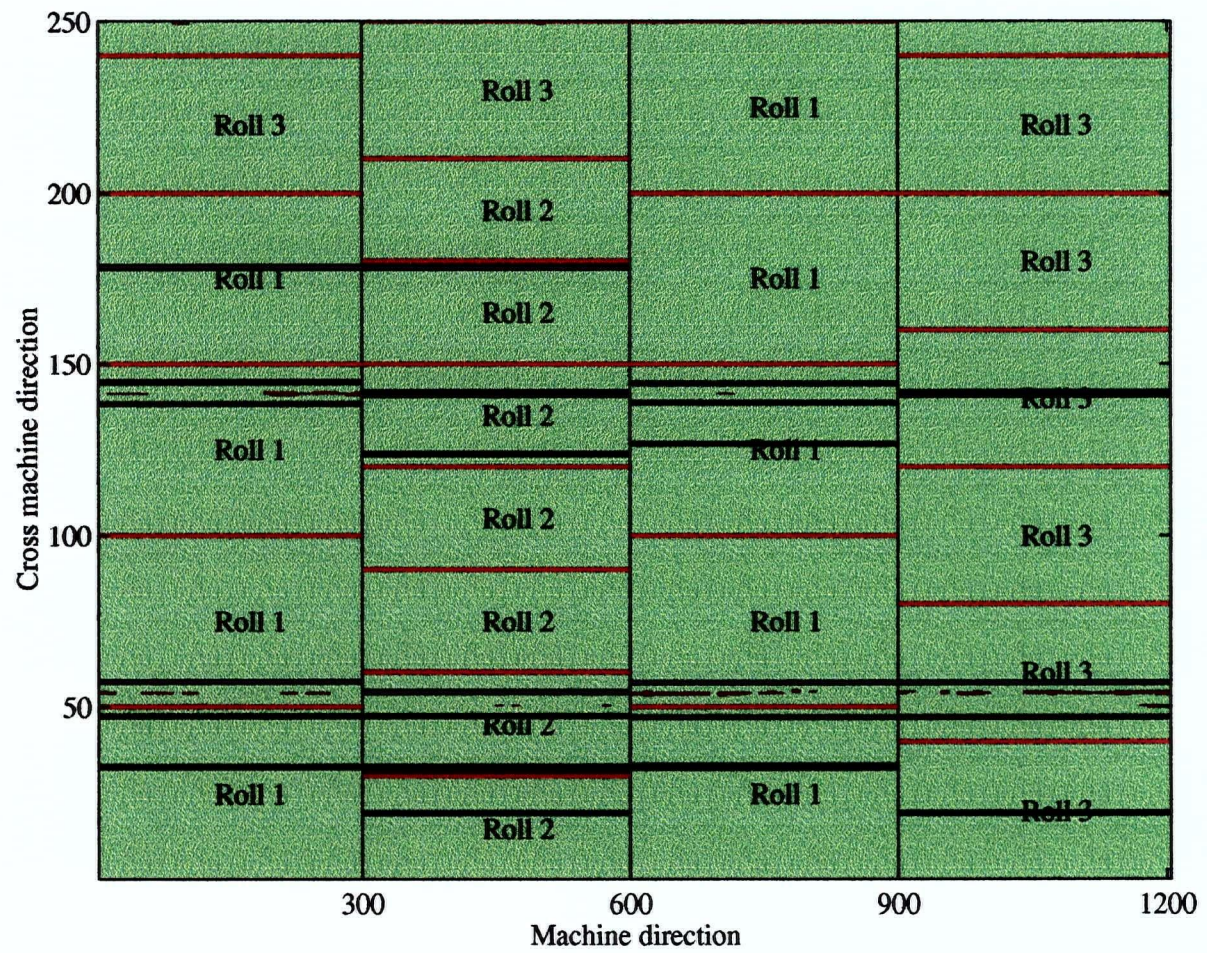


Figure 5.53: Trim optimization: Scheme 1



### 5.6.2 Scheme 2: Trim Optimization with Paper Quality Consideration

In this scheme, the paper quality is taken into consideration based on the  $2\sigma$  evaluation of basis weight property. First the measured raw profile is decomposed to level 1 using discrete wavelet transform and high frequency noise is removed. Then the reconstructed profile at level 1 is evaluated using  $2\sigma$  criteria. The areas with property variation within  $2\sigma$  are regarded as usable sections. The areas with property variation beyond  $2\sigma$  are regarded as unacceptable sections and thus will not be used during the cutting process. The optimization result is shown in Table 5.7 and Figure 5.54.

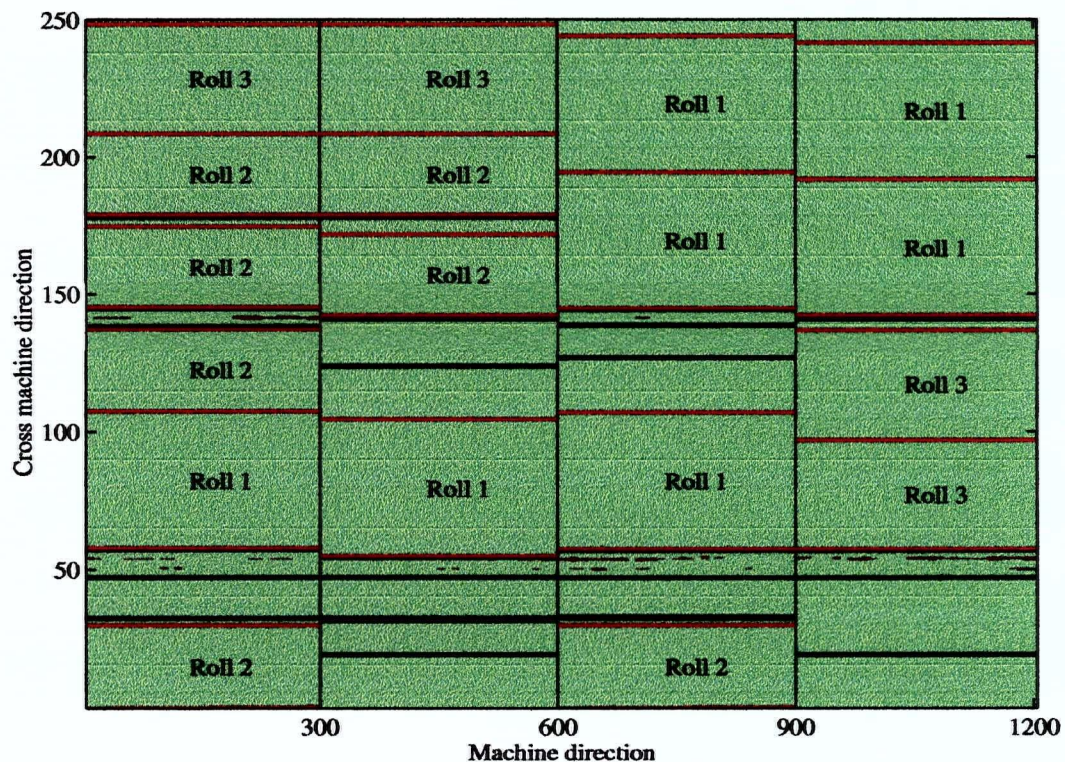


Figure 5.54: Trim optimization: Scheme 2



Table 5.7: Trim optimization result: Scheme 2

Usable section	Number of Roll 1	Number of Roll 2	Number of Roll 3
1	0	1	0
2	1	1	0
3	0	1	0
4	0	1	1
5	1	0	0
6	0	1	0
7	0	1	1
8	0	1	0
9	1	0	0
10	2	0	0
11	0	0	0
12	0	0	2
13	2	0	0
Total Number	7	7	4

### 5.6.3 Scheme 3: Trim Optimization after Improving Control

In this scheme, first the raw profile is decomposed into level 3 because level 3 is the controllable level(see chapter 3 and 4), and 80% of the controllable variations are removed by means of improving control. Then IDWT is performed to obtain the profile reconstruction at level 1. The reconstructed profile now is much smoother than the one in scheme 2, and thus there is high percentage of usable area in the jumbo reel after improving control. Considering that in practice, the specification is set according to customer's requirements, the same  $2\sigma$  value is used as in scheme 2. Then the optimization problem is solved. The optimization result is shown in Table 5.8 and Figure 5.55.

Table 5.8: Trim optimization result: Scheme 3

Usable section	Number of Roll 1	Number of Roll 2	Number of Roll 3
1	2	0	1
2	0	1	0
3	0	1	1
4	2	0	0
5	0	2	1
6	0	1	0
7	3	0	1
8	0	0	0
9	1	2	2
Total Number	8	7	6

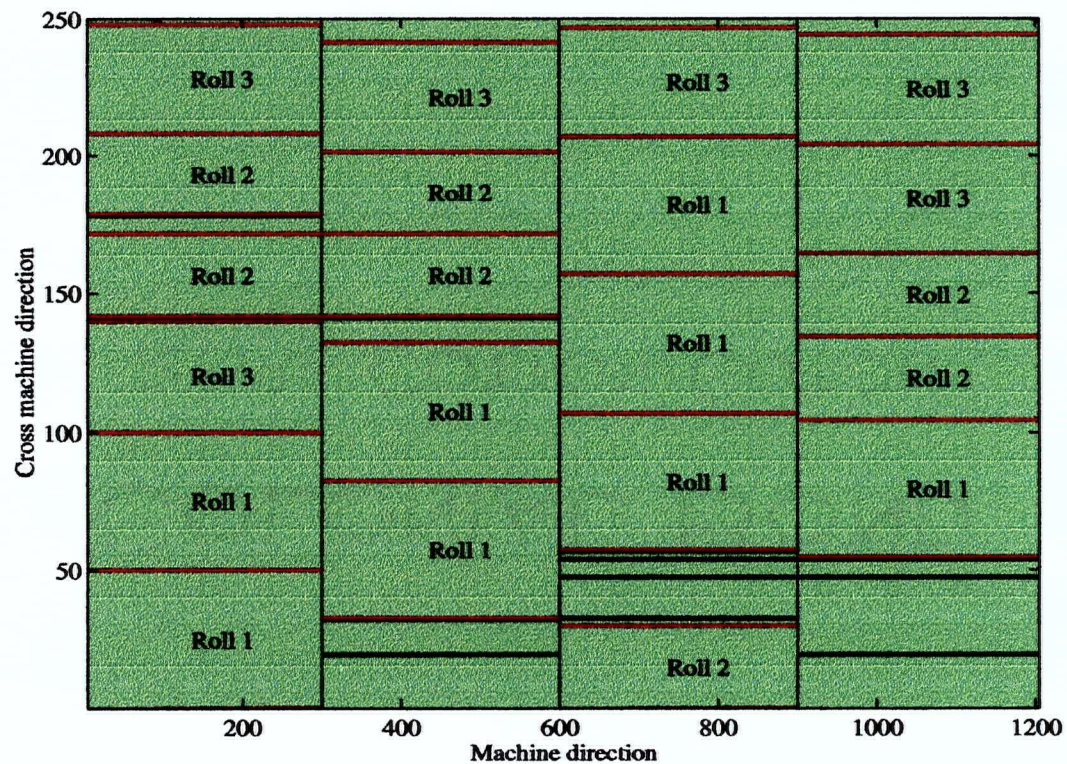


Figure 5.55: Trim optimization: Scheme 3

#### 5.6.4 Comparison and Conclusion

Based on the optimization results earlier, the amount of total waste paper is calculated and the three schemes are evaluated in terms of waste paper. The comparisons of these three optimization schemes are summarized in Table 5.9. In this table, the number of rolls produced for each roll order is listed, the total saleable paper is calculated.

Table 5.9: Result comparison for three schemes

No.	Total	Usable	Unusable	Roll1	Roll2	Roll3	Cut Rolls	Saleable	Saleable(%)
1	1000	823.8	176.2	9	7	8	980	390	39
2	1000	823.8	176.2	7	7	4	720	720	72
3	1000	917	83	8	7	6	850	850	85

It can be seen that if we just look at the numbers of produced rolls, Scheme 1 seems satisfactory. However, because paper sheet quality is ignored, as a result those entire rolls containing the un-usable area need to be discarded. The amount of cut rolls is 980, however the amount of saleable rolls is only 390. So the total waste (61%) will be the largest among the three schemes. For Scheme 2, the paper quality is taken into consideration. Instead of discarding the entire rolls later, the unusable areas are bypassed during the optimization. All the cut rolls are saleable and the saleable percentage is 72%, the total waste is 28%, much less than the 61% in Scheme 1. The total waste for Scheme 3 is the least (15%) among the three schemes. Since after improving control, the quality for the entire reel is improved, the amount of unusable area is reduced and thus trim-loss is minimized. This illustrates the potential benefit of improving control in terms of paper savings during trim-loss optimization.

## **Chapter 6**

### **Conclusions**

#### **6.1 Conclusions**

Aimed at improving paper machine efficiency, this thesis first described the use of wavelets and wavelet packets in paper machine data analysis. After wavelet decomposition, the data were analyzed to determine the controllable and uncontrollable components. Finally trim-loss optimization was studied as an attempt to maximize production quality while minimizing trim-loss.

First the validity and effectiveness of the wavelet transform analysis were confirmed by applying the discrete wavelet transform to both the paper machine on-line scanner data and off-line analyzer data.

Then, process monitoring and control performance assessment were achieved. By separating the controllable and uncontrollable variations in the cross machine direction profile using multi-resolution analysis, the performance and the control potential of the system were evaluated. Both wavelet and wavelet packets were used and the results were compared.

The wavelet and wavelet packets analyses have the following advantages:

- Wavelet filtering can separate the cross machine direction variation and machine direction variation accurately.
- Wavelet transform can represent the paper machine process data economically without loss of detail, it can also provide excellent visualization to the operator.

- Wavelet and wavelet packets provide high compression that allows the efficient storage of historical data and the transfer of data between different mills.
- Wavelet packet analysis can achieve an accurate separation of CD controllable and uncontrollable variation and control performance assessment.
- The performance index can be calculated on-line to provide the operator with a quick assessment of the process.

Finally, the processed paper machine profile obtained through wavelet or wavelet packet analysis was further used for trim-loss optimization in roll cutting. The method in this thesis, which took into account the quality of the paper sheet during optimization, presented a new concept in trim-loss optimization. Three different optimization schemes have been discussed and compared. The potential savings through optimization after improving control were also analyzed.

The results have shown that:

- By addressing the paper quality during trim optimization, the trim-loss is reduced.
- After improving control, improvements in the quality of the entire reel result in great savings through trim-loss optimization.
- The individual roll information such as  $2\sigma$  variation, mean value is displayed, which helps to separate different quality rolls according to different demands.

## 6.2 Further Work

There are some possible extensions to the research presented here, steps that might be taken to improve the methods in this thesis are described as following:

- For best denoising, Adaptive Waveform Analysis might need to be used to extract the coherent features of the signal adaptively using libraries of orthonormal waveforms.
- Before the separation of the CD controllable and uncontrollable profiles, the actuator response needs to be identified to exactly extract the controllable profile.
- Instead of minimizing the absolute trim-loss in the trim-loss optimization, a more comprehensive cost function including other important factors such as slit movement can be used.

## Bibliography

- [1] "LINGO: The Modeling Language and Optimizer". *Lindo System Inc.*, 1998.
- [2] A. Bruce and H.Y. Gao. "WaveShrink: Shrinkage Functions and Thresholds". *StatSci Division, MathSoft, Inc., Seattle, WA*, 41:909-996, 1986.
- [3] S. C. Chen. "Kalman Filtering Applied to Sheet Measurement". *7th American Control Conference, Atlanta, Georgia*, pages 643-647, 1988.
- [4] J. Chun. "Estimation and Control of Paper Machine Variables Using Wavelet Packets Analysis". *M.A.Sc. Thesis, The University of British Columbia*, 1997.
- [5] R. Coifman and Y. Meyer. "Signal Processing and Compression with Wavelet Packets". *Progress in Wavelet Analysis and Applications, Editions Frontiers, Toulouse, France*, pages 77-93, 1992.
- [6] K. A. Cutshall, G. E. Ilott, and J. H. Rogers. "Grammage Variation - Measurement and Analysis". *Grammage Variation Subcommittee, Process Control Committee, Technical Section, CPPA*, 1988.
- [7] E. Dahlin. "Computational Methods of a Dedicated Computer System for Measurement and Control on Paper Machines". *24th Engineering Conference, TAPPI, San Francisco, USA*, pages 62.1-62.42, Sept. 1969.
- [8] I. Daubechies. "Orthogonal Basis of Compactly Supported Wavelets". *Comm. Pure and Applied Math.*, 41:909-996, 1986.
- [9] I. Daubechies, A. Grossmann, and Y. Meyer. "Painless Non-Orthogonal Expansions". *J. Math. Phys.*, 27:293-309, 1986.
- [10] D. L. Donoho and I. M. Johnstone. "Minimax Estimation via Wavelet Shrinkage". *Technical Report 402, Department of Statistics, Stanford University*, July 1992.
- [11] D. L. Donoho and I. M. Johnstone. "Adapting to Unknown Smoothness via Wavelet Shrinkage". *J. Am. Stat. Ass.*, 1993.
- [12] D. L. Donoho and I. M. Johnstone. "Ideal Denoising in an Orthogonal Basis Chosen from a Library of Bases". *Technical Report 461, Department of Statistics, Stanford University*, Sept. 1994.

- [13] D. L. Donoho and I. M. Johnstone. "Ideal Spatial Adaptation via Wavelet Shrinkage". *Biometrika*, 81:425–455, July 1994.
- [14] G. A. Dumont, M. S. Davies, K. Natarajan, and C. Lindeborg. "An Improved Algorithm for Estimating Paper Machine Moisture Profiles Using Scanned Data". *30th IEEE Conference on Decision and Control, Brighton, England*, Dec. 1991.
- [15] H. Dyckhoff. "A New Linear Approach to the Cutting Stock Problem". *Operations Research*, 29:1092–1104, 1981.
- [16] P. G. Gilmore and R. E. Gomory. "A Linear Programming Approach to the Cutting-Stock Problem". *Operations Research*, pages 849–859, 1961.
- [17] A. Grossmann and J. Morlet. "Decomposition of Hardy Functions into Square Integrable Wavelets of Constant Shape". *SIAM J. Math. Anal.*, pages 724–736, 1983.
- [18] A. Haar. "Zur Theorie der Orthogonalen Funktionensysteme". *Mathematische Annalen*, 69:331–371, 1910.
- [19] I. Harjunoski and T. Westerlund. "Different Formulations for Solving Trim Loss Problems in a Paper-converting Mill with ILP". *Computers Chem. Engng*, 20, Suppl.:121–126, 1996.
- [20] A. I. HINXMAN. "The Trim-loss and Assortment Problems: A Survey". *European Journal of Operational Research*, 5:8–18, 1980.
- [21] B. Jawerth and W. Sweldens. "An Overview of Wavelet Based Multi-Resolution Analyses". *Technical Report 1993:1, Industrial Mathematics Initiative, Department of Mathematics, University of South Carolina*, 1993.
- [22] I. M. Johnstone and B. W. Silverman. "Wavelet Threshold Estimators for Data with Correlated Noise". *Technical report, Statistics Department, University of Bristol, UK*, Sept. 1994.
- [23] C. Lindeborg. "A Nonlinear Algorithm for the Estimation of Moisture Characteristics in the Paper Process". *Proc. of 17th International Conference BIAS-81, Milan, Italy*, 3:139–158, Oct. 1981.
- [24] M. Misiti and Y. Misiti. "Wavelet Toolbox User's Guide". *The Maths Works Inc.*, March 1996.
- [25] S. T. Morgan. "Estimation and Identification for Machine Direction Control of Basis Weight and Moisture". *Master's Thesis, The University of British Columbia*, June 1994.



- [26] K. Natarajan, G. A. Dumont, and M. S. Davies. "An Algorithm for Estimating Cross and Machine Direction Moisture Profiles for Paper Machines.". *IFFAC/FORS Symposium, Beijing, PRC*, pages 27–31, 1988.
- [27] Z. Nesic. "Paper Machine Data Analysis Using Wavelets". *M.A.Sc. Thesis, The University of British Columbia*, 1996.
- [28] Z. Nesic, M. S. Davies, and G. A. Dumont. "Paper Machine Data Analysis and Compression Using Wavelets". *Tappi Journal*, 80:191–204, 1997.
- [29] Z. Nesic, M. S. Davies, G. A. Dumont, and D. Brewster. "CD Control Diagnostics Using a Wavelet Toolbox". *International CD Symposium '97, Finland*, June 1997.
- [30] R. T. Ogden. "Essential Wavelets for Statistical Applications and Data Analysis". *Birkhauser, Boston*, 1997.
- [31] V. Solo. "Wavelet Signal Estimation in Coloured Noise with Extension to Transfer Function Estimation". *Proceedings of the 37th IEEE Conference on Decision and Control, Tampa, Florida USA*, December 1998.
- [32] G. Strang and T. Nguyen. "Wavelets and Filter Banks". *Wellesley-Cambridge Press*, 1996.
- [33] P. E. Sweeney and R. W. Haessler. "One-dimensional Cutting Stock Decisions for Rolls with Multiple Quality Grades". *European Journal of Operational Research*, 44:224–231, 1990.
- [34] X. G. Wang, G. A. Dumont, and M. S. Davies. "Modeling and Identification of Basis Weight Variations in Paper Machines". *IEEE Transactions on Control System Technology*, June 1993.
- [35] G. Wascher. "An LP-based Approach to Cutting Stock Problem with Multiple Objectives". *European Journal of Operational Research*, 1990:175–184, 1990.
- [36] T. Westerlund and I. Harjunkski. "Solving a Production Optimization Problem in a Paper-converting Mill with MILP". *Computers Chem. Engng*, 22:563–570, 1998.
- [37] T. Westerlund and F. Pettersson. "An Extended Cutting Plane Method for Solving Convex MINLP Problems". *Computers Chem. Engng*, 19, Suppl.:131–136, 1995.
- [38] M. V. Wickerhauser. "Lecture on Wavelet Packet Algorithm". *Department of Mathematics, Washington University*, Nov. 1991.