Investigations on Single and Multipolarization SAR Image Compression

by
Jing Wang

B.Sc. (Computer Science) HeFei University of Technology, 1991
M.Sc. (Computer Science) HeFei University of Technology, 1994

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
December 2002
© Jing Wang, 2002
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
Vancouver, Canada

Date Dec 19th, 2002
Abstract

Synthetic Aperture Radar (SAR) images provide important information about our living earth. However, there are problems associated with the storage and transmission of that data that are critical to extending their potential applications. To solve this problem we must find a suitable compression approach that can significantly decrease the volume of the data without losing any useful information that SAR images may provide.

There are two main parts to this thesis. The first part is concerned with single channel SAR image compression. Here, single channel represents single polarization, which is used to obtain images, since for single channel SAR image compression, only one image is used, and this image is the amplitude image. We focus our investigation on transform coding, which is a very popular data compression approach; the particular transform we are interested in is the Discrete Wavelet Transform (DWT). We want to find a way to improve the DWT based compression method to make it more suitable for SAR image compression. Based on experimental results, we find out that our goal can be achieved by adaptively adopting techniques, such as wavelet packet and block coding.

The second part of the thesis involves the investigation of the compression of multipolarization SAR images, which include three intensity images and two phase-difference images as a whole data set. The compression method we are concerned with is called the Principal Component Analysis (PCA), a standard compression technique for hyperspectral image compression. Our experiment results show that PCA is less efficient for multipolarization image compression than for multiple hyperspectral images. This is because PCA is more efficient at compressing multiple images with higher correlation, which is not true for multipolarization SAR images. In this thesis, we suggest multipolarization SAR images should be compressed separately to achieve the best compression performance, instead of grouping them together.
## Table of Contents

Abstract ii
Table of Contents iii
List of Figures vii
List of Tables ix
List of Abbreviations x
Acknowledgements xii

### Chapter 1 Introduction 1

1.1 Introduction to Synthetic Aperture Radar Images 1
1.2 Introduction to SAR Image Compression 3
1.3 Overview of the Thesis 5

### Chapter 2 Polarimetric SAR Data 7

2.1 Introduction 7
2.2 Maxwell's Equations for Electromagnetic Waves 8
2.3 Wave Polarization 9
2.4 Scattering Matrix 11
2.5 Stokes Vector and Stokes Matrix 13
2.6 Covariance Matrix 14
2.7 Coherency Matrix 15
2.8 Testing Data 16
   2.8.1 Single Channel SAR Data Format 16
4.7.1 Case One
   4.7.1.1 Variable Ratio 54
   4.7.1.2 1-D Spectrum 57
   4.7.1.3 MSE and Bit Allocation Results 59
4.7.2 Case Two
   4.7.2.1 Variable Ratio 61
   4.7.2.2 1-D Spectrum 63
   4.7.2.3 MSE and Bit Allocation Results 63
4.8 Summary 66

Chapter 5 Polarimetric SAR Image Compression 70
5.1 Introduction 70
5.2 Questions Associated with Polarimetric Data Compression 71
   5.2.1 The First Question 71
   5.2.2 The Second Question 72
5.3 Mathematical Background for PCA 73
5.4 Experiment Method 74
5.5 Experimental Results for Three Intensity Images 76
   5.5.1 Covariance Matrix 76
   5.5.2 Correlation Matrix 76
   5.5.3 Quantizer 78
      5.5.3.1 Optimal Quantizer for Gaussian Distribution 78
      5.5.3.2 Optimal Quantizer for Laplacian Distribution 79
   5.5.4 PCA Property: Orthogonal and Unitary Transform 83
   5.5.5 Wavelet Transform 84
   5.5.6 Comparison: PCA vs Wavelet Transform 85
   5.5.7 Appropriate Solution 86
Chapter 6 Conclusions and Future Work

Bibliography

Appendix

Compression of RADARSAT Data with Block Adaptive Wavelets
List of Figures

Figure 2.1 Coordinate system for radar polarimetry 9
Figure 2.2 Polarization ellipse in the x-y plane 11
Figure 2.3 Single channel SAR image 19
Figure 2.4 Intensity images of multichannel SAR data 20
Figure 2.5 Phase difference images of multichannel SAR data 21
Figure 3.1 Two-channel perfect reconstruction filter bank scheme 27
Figure 3.2 The full binary tree for the two-scale wavelet packet transform 28
Figure 3.3 Frequency response for the two-band wavelet packet filter bank 29
Figure 3.4 Signal symmetries. (a) Whole-sample symmetry, (b) Half-sample symmetry 31
Figure 3.5 Parent-child dependencies of the subbands 33
Figure 3.6 Scanning order of the subbands 34
Figure 4.1 General block diagram of WPEB. (a) Encoding diagram, (b) Decoding diagram. 41
Figure 4.2 The difference between wavelet transform and wavelet packet. (a) Wavelet Transform, (b) Wavelet packet (to LH1 only) 44
Figure 4.3 Rate-distortion curves for three different cases. (a) The whole image is not divided, (b) The whole image is divided into two blocks, (c) The whole image is divided into three blocks 50
Figure 4.4 Denoising to wavelet coefficients. (a) Hard thresholding, (b) Soft thresholding. 53
Figure 4.5 Two scale wavelet decomposition 56
Figure 4.6 1-D spectrum of the first test image 58
Figure 4.7 MSE results of SPIHT and WPEB for the first test image 60
Figure 4.8 Rate-distortion curve for the first test image 61
Figure 4.9 1-D spectrum for the second test image  
Figure 4.10 Rate-distortion curve for the whole image and two blocks  
Figure 4.11 Optimal compression searching curve  
Figure 4.12 MSE results of SPIHT and WPEB for the second test image  
Figure 4.13 Test images. (a) Case one, (b) Case Two  
Figure 4.14 Compression results for SPIHT and WPEB at 1.0bpp for case one  
Figure 4.15 Compression results for SPIHT and WPEB at 1.0bpp for case two  
Figure 5.1 Three principal components with eigenvalues as 97, 16 and 597  
Figure 5.2 Three reconstructed image using an average of 1bpp (R=3)  
Figure 5.3 Three reconstructed image using an average of 2bpp (R=6)  
Figure 5.4 Compression results for three intensity images at 1bpp and 2bpp  
Figure 5.5 Frequency response of PCA eigenvectors  
Figure 5.6 PCA filters in multiple image direction  
Figure 5.7 Comparison of SNR for two components and original signal
List of Tables

Table 2.1  Dynamic data range for HH, HV, VV channel images 18
Table 4.1  Variance ratio for the first test image 55
Table 4.2  Variance ratio for the second test image 62
Table 4.3  Optimal compression ratio for two blocks 64
Table 5.1  SNR and MSE for three compressed intensity images 79
Table 5.2  RMS for three reconstructed intensity images (Gaussian distribution) 79
Table 5.3  SNR, MSE and RMS for three reconstructed intensity images (Laplacian distribution) 79
Table 5.4  Compression results for three intensity images using SPIHT at 1bpp 84
Table 5.5  Some related data for three principal components 85
Table 5.6  Compression results (without quantizer) 86
List of Abbreviations

ADC  Analog-to-digital Converter
BAQ  Block Adaptive Quantization
BPP  Bit Per Pixel
CCRS Canada Center for Remote Sensing
DCT  Discrete Cosine Transform
DFT  Discrete Fourier Transform
DWT  Discrete Wavelet Transform
EBCOT Embedded Block Coding with Optimal Truncation
EM   ElectroMagnetic
EZW  Embedded Zerotree Wavelet
GIS  Geographical Information System
HS   Half-sample Symmetry
IID  Independent Identical Distribution.
JPEG Joint Photographic Experts Group
JPL  Jet Propulsion Laboratory
KL   Karhuene - Loeve
LIS  List of Insignificant Sets
LIP  List of Insignificant Pixels
LSP  List of Significant Pixels
MSE  Mean Squared Error
PR   Perfect Reconstruction
QMF  Quadrature Mirror Filter
RADAR Radio Detection and Ranging
RMS  Root Mean Square
ROI  Region Of Interest
<table>
<thead>
<tr>
<th>Acronym</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>SAR</td>
<td>Synthetic Aperture Radar</td>
</tr>
<tr>
<td>ScanSAR</td>
<td>Scanning Synthetic Aperture Radar</td>
</tr>
<tr>
<td>SIR-C</td>
<td>Spaceborne Image Radar-C band</td>
</tr>
<tr>
<td>SNR</td>
<td>Signal Noise Ratio</td>
</tr>
<tr>
<td>SPIHT</td>
<td>Set Partitioning in Hierarchical Trees</td>
</tr>
<tr>
<td>WP</td>
<td>Wavelet Packet</td>
</tr>
<tr>
<td>WS</td>
<td>Whole-sample Symmetry</td>
</tr>
</tbody>
</table>
Acknowledgements

First, I would like to thank my supervisor Dr. Ian Cumming for his consistent and patient supervision. My thesis would not have been finished without him. He is also the first person in my life who makes me realize that clear writing is very important and helpful to any kind of thinking.

Second, my thanks is given to many of the friends I met during my studies at UBC. Especially, I would like to thank HuiQun Deng for her valuable discussions and suggestions, Bernd Scheuchl for providing me with polarimetric data for my experiments, and Mehran Azimi for his help when I encountered computer problems.

Finally, I wish to thank my family for their deep understanding and love. I could not have done anything without them. This is why I would like to devote my thesis to them.

Jing Wang

The University of British Columbia
September 2002
Chapter 1

Introduction

1.1 Introduction to Synthetic Aperture Radar Images

Synthetic Aperture Radar (SAR) images provide an efficient mechanism for earth observation. SAR is an active system that transmits a beam of electromagnetic (EM) wave and receives the backscattered electromagnetic wave as raw data for imaging the target. As an active system, SAR provides its own illumination and is not dependent on light from the sun, thus permitting continuous day and night operation. Furthermore, neither clouds nor precipitation have a significant effect on microwaves, thus permitting all-weather imaging. In many cases, radar is the only way scientists can explore inaccessible regions of the Earth's surface.

Conditions on the Earth's surface influence how much radar energy is reflected back to an antenna. An area with a variety of surface types, such as hills, trees, and large rocks, generally reflects more energy back to the radar than a less complex area, such as desert. The resulting radar image of the varied terrain is brighter overall, than the image of the simpler area.

SAR image data provide important information for many applications, such as the following:

*Agriculture: Agriculture plays a dominant role in the economy of every nation. It represents a substantial trading industry for an economically strong country and sustenance for a hungry, overpopulated one. SAR data are used as mapping tools to classify crops, examine their health and viability, and monitor farming practices. Agricultural applications of remote sensing include crop type classification, crop condition...
assessment, crop yield estimation, mapping of soil characteristics, mapping of soil management practices, and so forth.

*Forestry: Forests are a valuable resource providing food, shelter, wildlife habitat, fuel, and daily supplies, such as medicinal ingredients and paper. Forestry applications where remote sensing data can be utilized include sustainable development, biodiversity, deforestation, reforestation monitoring and managing, commercial logging operations, wildlife habitat assessment, and other environmental concerns.

*Geology: Geology involves the study of landforms, structures, and the subsurface, to understand the physical processes creating and modifying the earth's crust. Remote sensing is used as a tool to extract information about the land surface structure, composition or subsurface. SAR data provide an expression of surface topography and roughness, and thus is extremely valuable. Remote sensing is not limited to direct geology applications, it is also used to support logistics, such as route planning for access into a mining area, reclamation monitoring, and generating base maps upon which geological data can be referenced or superimposed.

*Hydrology: Hydrology is the study of water on the Earth's surface, whether flowing above ground, frozen in ice or snow, or retained by soil. Hydrology is inherently related to many other applications of remote sensing, particularly forestry, agriculture and land cover, since water is a vital component in each of these disciplines. Remote sensing data offers a synoptic view of the spatial distribution and dynamics of hydrological phenomena, often unattainable by traditional ground surveys. SAR radar has brought a new dimension to hydrological studies with its active sensing ability. Typical hydrological applications include soil moisture estimation, river and lake ice, flood mapping, irrigation scheduling, and so forth.

A general example of an application associated with utilizing SAR images lies in disaster monitoring, such as the flood disaster which happened in 1998 in China along the Yangtze River; RADARSAT ScanSAR acquired the data, and the Canada Centre for Remote Sensing (CCRS) geocoded, enhanced and classified the data with respect to the
GIS data were overlaid on the RADARSAT image to provide a map reference for the normal water level. The monitored information is very important for making decisions and designing a plan to prevent the spread of loss caused by the flood disaster.

Other applications can be found in sea ice information, land cover and land use, mapping, oceans, and coastal monitoring. Details of many application examples can be easily obtained at the CCRS website in the Technology R&D subdirectory [31].

1.2 Introduction to SAR Image Compression

Synthetic aperture radar is a very efficient instrument for obtaining a better understanding of the earth’s environment. SAR data represent an important source of information for a large variety of scientists around the world. Generally, there are two types of image acquiring systems. One is called the airborne system in which the earth is treated as flat; the other is the spaceborne radar system, in which earth curvature is important.

The spaceborne remote sensing data, complemented by aircraft data, provide detailed information of the Earth’s surface. Each system collects large amount of data every year, so the data volume is so enormous and data compression is inevitable. The great importance associated with SAR image compression lies in the following aspects.

- SAR Raw Data Downlink

Modern spaceborne SAR systems have on-board hardware consisting of a transmitting and receiving unit and analog-to-digital (A/D) conversion, followed by a real time downlink and/or storage facility. One of the major constraints in the design and operation of current spaceborne SAR systems is the non-availability of a downlink with a high data rate.

At the downlink, the data are raw or are signal data, only after ground processing can they be viewed as imagery. The process associated with imagery calibration is out of the
scope of this thesis. Details regarding raw data processing and image calibration can be found both in [1] and [2].

Until now, the block adaptive quantizer (BAQ) [34] [35] has been selected for on-board data compression due to its simplicity for coding and decoding. The real time implementation of on-board encoding has reduced the data rate for downlink. In this case, less power is required for transmission to the ground station, and more information channels (e.g., polarimetric and multifrequency applications) can be incorporated in the downlink. In addition, more raw data are stored on board before transmission to ground stations, which allows SAR imaging during longer orbit segments, where no downlink is available.

• Imagery Archive/Storage

A general idea about how much memory is needed to store data acquired by SAR at CCRS in one year can be found below. Each year, the Canada Centre for Remote Sensing collects and archives 30 terabytes for satellite imagery [32]. One terabyte is

* 1024 gigabytes, or
* 1,048,576 megabytes, or
* 1,073,741,824 kilobytes, or
* 1,099,511,627,776 bytes.

The data the CCRS now has is approximately 300 terabytes. If you put about 640 megabytes of the data onto one compact CD, it takes 430,000 CDs to store all the information the CCRS has collected so far. If the CDs were placed side by side, they would span 51.6 km. That is a lot of data. Since most of the stored data is calibrated as image data, an efficient algorithm for SAR image data compression has become an important tool in reducing the amount of stored data.

• SAR Image Transmitting
The ability to transmit SAR data to end-users is normally impaired by the limited channel bandwidth. Compression of SAR image data is necessary in shortening transmission time, especially for large volume data with an urgent delivering deadline.

1.3 Overview of the Thesis

For a multichannel and multipolarization case, SAR data volume associated with downlink, archive and transmission is even larger than in a single channel, single polarization case. The objective of this research is to investigate SAR polarimetric image data (not raw data) compression by using wavelet packets (WP). As a generalization of the wavelet basis, the WP, which is a rich family of orthonormal/biorthonormal bases, can be more suited to match the non-stationary statistics of the images and the significant medium and high frequency components.

There are two problems associated with practical polarimetric SAR image compression. One is the speckle phenomenon. Speckle results from the necessity of creating the image with coherent radiation. A fully developed speckle pattern appears chaotic and unordered. When detailed information of the image is important, speckles can be viewed as noise degradation of the image. Therefore, speckle reduction is an essential procedure before the application of automatic target detection and recognition. Another problem is the data representation associated with multichannel polarimetric SAR data. As we know, for single channel SAR data, only the intensity image needs to be compressed, and the phase information is so random that it is not necessary to be compressed at all. For polarimetric SAR data that are inherently complex, the intensity images as well as the phase images should be compressed to make the information complete.

Chapter 2 introduces the radar polarimetry mechanism. It answers such questions as how the polarimetric SAR data are formed, how to fully represent polarimetric data, and so forth. Although this chapter deals with the fundamentals of radar polarimetry, it is
enough for us to understand the polarimetric data stored in CEOS format. As a matter of fact, test image data are obtained by extracting useful information from this data format.

In Chapter 3, wavelet transforms in a continuous time format, as well as a discrete time format, are introduced, followed by filter bank and wavelet packets. Some standard image compression algorithms are introduced to demonstrate the great success of utilizing wavelet transforms for the purpose of compression.

Chapter 4 is the main menu of this thesis. The main features of single channel SAR image compression are presented in this chapter. The new compression algorithm based on wavelet packets and a conditional block coding scheme is emphasized. It is known that multichannel polarimetric SAR data compression can be treated as several single channel SAR data compressions separately, which is also true for the phase data for each polarization channel data.

In Chapter 5, polarimetric data representation is examined in order to find the best data representation for polarimetric data compression. Since polarimetric data are complex, they not only contain intensity images, but also phase images. The decorrelation method called principal component analysis is examined to find out whether it is suitable for multichannel polarimetric data decorrelation or not. Although the final answer is negative, it is still a very interesting result that has never been stressed before. In addition, the conditions to successfully utilize principle component analysis for denoising is proposed and proved mathematically.

Finally, the conclusions are given in the last chapter, as well as some future concerns.
Chapter 2

Polarimetric SAR Data

2.1 Introduction

Conventional SAR operates with a single fixed polarization antenna for both the transmission and reception of the signals. In this way, a single scattered coefficient is measured for a specific transmitting and receiving polarization combination, for many thousands of points in a scene. A result of this implementation is that only one component of the scattered wave is measured, and the additional information about the surface contained in the polarization properties of the reflected signal is lost.

The concept of utilizing imaging radar polarimetry involves measuring the polarization of the scattered wave for any or all transmitting polarizations. This technology enables the complete measurement of a target’s polarization properties, thus permitting a much more detailed understanding of the electromagnetic scattering process. Radar polarimetry, where the complete, complex scattering matrix for every point in a high-resolution radar image is measured, opens a realm of new applications for imaging radar systems.

Radarsat-2 will be the first satellite capable of providing fully polarimetric data on an operational basis and it is going to be launched in 2003. The advanced polarization capability of RADARSAT-2 can be expected to facilitate the discrimination of the observed features and considerably enhance the overall application potential. For fully polarized SAR, the backscattered data are acquired by transmitting vertical and horizontal polarized radar waves, and receiving the reflected wave in a vertical and horizontal direction as well. Thus, we have four channel data: the HH, VH, HV, and VV
multichannel data. The first capital letter denotes the transmitted wave channel; the second denotes the received wave channel.

This chapter introduces the following: (1) the fundamental theory of radar polarimetry; (2) polarimetric data representation, such as scattering matrix, covariance matrix, coherency matrix, and Stokes matrix (also known as the Muller matrix); and (3) the test data format.

2.2 Maxwell’s Equations for Electromagnetic Waves

Maxwell’s equations are the fundamentals for electromagnetic wave propagation. Maxwell’s equations are listed below:

\[ \nabla \cdot D = \rho \]
\[ \nabla \times E = -\frac{\partial B}{\partial t} \]
\[ \nabla \cdot B = 0 \]
\[ \nabla \times H = J + \frac{\partial D}{\partial t} \]  \hspace{1cm} (2.1)

Where

\( E(r, t) \): Electric field intensity.
\( H(r, t) \): Magnetic field intensity.
\( B(r, t) \): Magnetic flux density.
\( D(r, t) \): Electric displacement.
\( \rho(r, t) \): Charge density.
\( J(r, t) \): Current density.

The law of conservation of charge is as follows:

\[ \nabla \cdot J + \frac{\partial \rho}{\partial t} = 0 \]  \hspace{1cm} (2.2)

Fundamental relations associated with the Maxwell Equation are as follows:
\[ D = \varepsilon' E \]
\[ B = \mu H \]
\[ J = \sigma E \]

Where
\[ \varepsilon' : \text{Permittivity} \]
\[ \mu : \text{Permeability} \]
\[ \sigma : \text{Conductivity} \]

### 2.3 Wave Polarization

The polarization of a plane wave describes the shape and locus of the E-vector (in a plane orthogonal to the direction of propagation) as a function of time. In the general case, the locus of the E-vector in a plane orthogonal to the direction of propagation and the wave is called elliptically polarized [3]. Under certain conditions, the ellipse may degenerate into a segment of a straight line or a circle, and the polarization is then called linear or circular. While in many cases, the most frequently used linear polarization is referred to as horizontal or vertical polarization, as shown in Figure 2.1 [36].

![Coordinate system for radar polarimetry.](image)

Transmitted waves travel in the \( \hat{z} \) direction, and received waves in the \(-\hat{z}\) direction. Horizontally polarized waves have an electric field vector parallel to \( \hat{x} \), and
vertically polarized waves have an electric field vector perpendicular to \( \hat{x} \), which lies in the \( \hat{y} \) direction.

Since the electric field is transverse, we can represent any time harmonic plane wave solution to Maxwell's equations by the following vector:

\[
E = \text{Re} \left[ (E_x \hat{x} + E_y \hat{y}) e^{-j(\omega t - kz)} \right]
\]  
(2.4)

Where

\[
E_x = a_x e^{-j\delta_x},
\]  
(2.5)

\[
E_y = a_y e^{-j\delta_y},
\]  
(2.6)

The positive amplitudes in the \( x \) and \( y \) direction are \( a_x \) and \( a_y \), and the corresponding phase are \( \delta_x \) and \( \delta_y \) relative to the phase factor \( \omega t - kz \), where \( \omega \) is radian frequency, \( t \) is time, \( k \) is wave number, and \( z \) is the distance traveled in the \( \hat{z} \) direction. Any wave can be uniquely represented by the complex vector pair \( (E_x, E_y) \).

We refer to the relative amplitude and phase relationships of the components of a given wave at the polarization state. A geometric interpretation of polarization follows if we rewrite (2.4) as the following:

\[
E = a \cos \chi \sin(\omega t - kz + \alpha) \hat{\xi} + a \sin \chi \cos(\omega t - kz + \alpha) \hat{\eta}
\]  
(2.7)

Where \( a^2 = a_x^2 + a_y^2 \) is the intensity of the wave, \( \alpha \) is a phase angle, and \( \hat{\xi}, \hat{\eta} \) are the unit vectors in a coordinate system rotated by angle \( \psi \) with respect to \( \hat{x} \). This is a parametric form of the equation of an ellipse, hence the term elliptic polarization. Such an ellipse is shown in Figure 2.2.
Figure 2.2 Polarization ellipses in the x-y plane.

The wave is traveling in the z direction, which is off the page.

\( \alpha \) is an auxiliary angle defined as the following:

\[
\tan \alpha = \frac{a_x}{a_y}
\]  
(2.8)

For the ellipticity angle \( \chi \), which specify the shape of the ellipse, it has the following limits \([-45^\circ, 45^\circ]\), and is defined as the following:

\[
\tan \chi = \frac{a_y}{a_x}
\]  
(2.9)

For the rotation angle \( \psi \), which is the angle between the major axis of the ellipse and a reference direction, chosen here to be the x axis, it has the range of \([-90^\circ, 90^\circ]\). If \( \chi = 0^\circ \) (linear polarization), \( \psi = 0^\circ \) represents horizontal polarization, and \( \psi = 90^\circ \) represents vertical polarization.

### 2.4 Scattering Matrix

A polarimeter is a radar transmitting in two orthogonal polarizations, and receiving four available channels, two co-polarised and two cross-polarised. Since the data
measured by a polarimeter are complex, they not only contain the amplitude, but also the phase. A Scattering matrix is a general mathematical representation of SAR polarimetry measurement in the sense that each element of the scattering matrix is a complex value. It relates to the scattered electric field $E^s$ and the incident electric field $E^i$.

Considering a scatterer illuminated by an electromagnetic plane with the incident electric field $E^i$, we have the following:

$$E^i = E^i_v \hat{v}_i + E^i_h \hat{h}_i \quad (2.10)$$

Here the transmitted wave hits the point objects, and the scattered wave is radiated and received by the antenna. In the far zone of the scatterer, the scattered wave is an outgoing, spherical wave, and can be approximated by a plane wave over the relatively small area occupied by the receiving antenna. The electric field of the scattered wave $E^s$ can be written as the following:

$$E^s = E^s_v \hat{v}_s + E^s_h \hat{h}_s \quad (2.11)$$

The scattering matrix $S$ is defined as the following:

$$\begin{pmatrix} E^i_v \\ E^i_h \end{pmatrix} = \frac{\epsilon^{ikr}}{r} \begin{pmatrix} E_{vv} & E_{vh} \\ E_{hv} & E_{hh} \end{pmatrix} \begin{pmatrix} E^s_v \\ E^s_h \end{pmatrix} \quad (2.12)$$

Or

$$E^s = \frac{\epsilon^{ikr}}{r} SE^i \quad (2.13)$$

Here, $r$ is the distance between the scatter and the receiving antenna, and $k$ is the wave number of the transmitting pulse wave. From (2.12), (2.13), it is clear that the scattering matrix has the following properties:

1. It is a direct matrix relating the components of $E^s$ and $E^i$.

2. Each element of the scattering matrix may be a function of frequency, the scattering and illuminating angles, and the orientation of the scatterer relative to the coordinate system.

12
In many cases, the two cross-polarised components $S_{hv}$ and $S_{vh}$ are assumed to be identical as the reciprocity theorem holds. This assumption for most natural targets results in the reduction of data volume associated with the system. However this assumption also relies strongly on the thorough calibration of the data, as channel imbalance and crosstalk may severely affect measurement.

2.5 Stokes Vector and Stokes Matrix

The Stokes vector is another mathematical representation of the electric field of radar polarimetry.

$$F = \begin{bmatrix} |E_v|^2 + |E_h|^2 \\ |E_v|^2 - |E_h|^2 \\ 2 \text{Re}(E_v E_h^*) \\ 2 \text{Im}(E_v E_h^*) \end{bmatrix}$$

(2.14)

The Stokes matrix is a transform related to the incident wave Stokes vector $F^i$ and the scattered wave Stokes vector $F^s$, which can be written as follows:

$$F^s = MF^i$$

(2.15)

where the Stokes matrix $M$ is a 4X4 real valued matrix whose elements are a linear combination of the cross products of the four basic elements of the scattering matrix. The reciprocity indicates that $S_{hv}$ and $S_{vh}$ are equal, resulting in a symmetrical scattering matrix, also referred to as a monostatic mode. For a monostatic Stokes matrix, each element is defined as follows:

$$M_{11} = \frac{[S_{hh} \cdot S_{hh}^* + S_{vv} \cdot S_{vv}^* + 2 S_{hv} \cdot S_{hv}^*]}{4}$$

$$M_{12} = \frac{[S_{hh} \cdot S_{hh}^* - S_{vv} \cdot S_{vv}^*]}{4}$$

$$M_{13} = \Re [S_{hh} \cdot S_{hv}] + \Re [S_{hv} \cdot S_{vv}] / 2$$

$$M_{14} = -\Im [S_{hh} \cdot S_{hv}] / 2 - \Im [S_{hv} \cdot S_{vv}] / 2$$

13
\[ M_{22} = \frac{\text{Shh} \cdot \text{Shh}^* + \text{Svv} \cdot \text{Svv}^* - 2 \text{Shv} \cdot \text{Shv}^*}{4} \]
\[ M_{23} = \Re\left[ \text{Shh} \cdot \text{Shv} \right]/2 - \Im\left[ \text{Shv} \cdot \text{Svv} \right]/2 \]
\[ M_{24} = \Im\left[ \text{Shh} \cdot \text{Shv} \right]/2 - \Re\left[ \text{Shv} \cdot \text{Svv} \right]/2 \]
\[ M_{33} = \frac{\text{Shv} \cdot \text{Shv}^*}{2} + \frac{9\text{Shh} \cdot \text{Svv}^*}{2} \]
\[ M_{34} = \frac{3\text{Shh} \cdot \text{Svv}^*}{2} \]
\[ M_{44} = \frac{\text{Shv} \cdot \text{Shv}^*}{2} - \Re\left[ \text{Shh} \cdot \text{Svv} \right]/2 \]

2.6 Covariance Matrix

The covariance matrix \( C \) is a Hermitian matrix, which can be written as the following:

\[ C = K_c K_c^+ \quad (2.16) \]

where + denotes conjugate transpose, and \( K_c \) is a vector. Assuming reciprocity, this is a three elements vector, given as the following:

\[ K_c = \begin{bmatrix} \text{Shh} \\ \sqrt{2} \text{Shv} \\ \text{Svv} \end{bmatrix} \quad (2.17) \]

Substituting \( K_c \) in (2.17) into the representation of (2.16), we get the covariance matrix written as the following:

\[ C = \begin{pmatrix} \text{ShhShh}^* & \sqrt{2}\text{ShhShv}^* & \text{ShhSvv}^* \\ \sqrt{2}\text{ShvShh}^* & 2\text{ShvShv}^* & \sqrt{2}\text{ShvSvv}^* \\ \text{SvvShh}^* & \sqrt{2}\text{SvvShv}^* & \text{SvvSvv}^* \end{pmatrix} \quad (2.18) \]

The covariance matrix is a positive semidefinite Hermitian matrix, and consists of nine independent real parameters, with three real coefficients in the diagonal and three complex (6 real) coefficients in the non-diagonal position. It not only fully describes the scatterer, but also immediately shows all measurable properties of the target. In addition, the covariance matrices have the advantage that the collective properties of a group of
resolution elements can be expressed using a single matrix rather than requiring the
individual scattering matrices for each element [27].

It should be noted that there are certain applications for which the scattering
matrices are preferable for either the covariance or Stokes matrix representation. In
particular, when maximum resolution is desired, the scattering matrix is a straightforward
representation of the measured target. The major advantage of either the covariance or
Stoke matrices is that they may be averaged to reduce data volume. However, it is
impossible to uniquely invert from a single covariance or Stokes matrices for the original
scattering matrices that characterized the observation. Applications requiring the original
data are better served by more voluminous scattering matrix data sets.

2.7 Coherency Matrix

Another useful approach to represent polarimetric data is by adopting the coherency
vector $K_i$, which is defined below as the following:

$$
K_i = \frac{1}{\sqrt{2}} \begin{bmatrix}
S_{hh} + S_{vv} \\
S_{hh} - S_{vv} \\
2 S_{vv}
\end{bmatrix}
$$

and $T_i = K_i \cdot K_i^\dagger$ \hspace{1cm} (2.19)

where $T_i$ is the coherency matrix, $^*$ represents conjugate, and $^T$ the transpose. It is well
known that polarimetric data are frequently multilook processed for speckle reduction
and data volume compression by averaging $n$ neighboring pixels. The multilooked
coherency matrix becomes the following:

$$
\langle T \rangle = \frac{1}{n} \sum_{i=1}^{n} K_i \cdot K_i^\dagger
$$

(2.20)

The coherency matrix is frequently used for classification based on target
decomposition by eigenvalue analysis of the matrix [28] [29] [30]. Another interesting
link between the covariance matrix and the coherency matrix is that they are linearly related by a linear transform [29].

\[ \langle T \rangle = N \langle C \rangle N^T \text{ where } N = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 0 & 1 \\ 1 & 0 & -1 \\ 0 & \sqrt{2} & 0 \end{bmatrix} \] (2.21)

2.8 Testing Data

There are mainly two types of SAR data available in this thesis for verifying the algorithm experimentally, and evaluating the results for image compression. The first data set is an X-band high-resolution airborne SAR image (Figure 2.3). The second data set is a multichannel data set containing three intensity and two phase difference images under the reciprocity assumption (Figure 2.4 and Figure 2.5).

2.8.1 Single Channel SAR Data Format

The SAR test data of a single channel (single polarization HH) shown in Figure 2.3, contains a large rural area that is part of Bedfordshire in southeast England. The river runs across the top of the image with other water features (dark areas). The built-up area in the lower center is the village of Stanwick. The dark stripe crossing the image from left to right just above the image is a major road. The remainder of the image is largely composed of isolated trees and fields bordered by hedges.

The total image size is 768X538, and the left corner of the image size 512X512 is the real data used in this thesis for convenience. Each pixel is represented with an 8-bit integer, and the data range is [0, 255].
2.8.2 Multichannel SAR Data Format

The test data were acquired during a SIR-C mission that provides spaceborne multifrequency, multi-polarisation SAR data during two space shuttle flights on April 18, 1994. The SAR data shows the west coast of Newfoundland with the areas of sea ice, open water and land. The landmass in the scene is part of Gros Morne National Park in Newfoundland. The incident angle ranges from 25.4 to 30.0. Refer to Figure 2.4.

The polarimetric multichannel (HH, HV, VV polarization mode) data for this thesis come from JPL with CEOS format. This format intends to reduce the data disk storage requirements. Details regarding all available data types, such as singlelook, multilook, quad-pol, and dual-pol, can be easily obtained from the following website http://southport.jpl.nasa.gov/software/dcomp/dcomp.html#RTFToC10.

The test data type is multilooked complex data. Each pixel is represented using 10 bytes, with 8 bits/byte. For each pixel at HH, HV, VV polarization mode, the intensity and phase value can be recovered from those 10 bytes. The data obtained are 32-bit float point data with a total image size of 900X200. The definition for 10 bytes of each pixel is listed below.

Byte (1)=int{log2(ShhShh*+2ShvShv*+2SvvSvv*)}

Byte (2)=nint{254[Mantissa – 1.5]}  where nint means near integer

Mantissa=(ShhShh*+2ShvShv*+SvvSvv*) / (2^Byte(1))

qsca=[(Byte(2)/254) + 1.5] (2^Byte(1))

Byte (3)=nint{255 sqrt(ShvShv* / qsca)} – 127

Byte (4)=nint{255 (SvvSvv* / qsca)} – 127

Byte (5)=nint{sign[Re(ShhShv*)] 127sqrt(2 | Re(ShhShv*)| / qsca)}

Byte (6)=nint{sign[Im(ShhShv*)] 127sqrt(2 | Im(ShhShv*)| / qsca)}

Byte (7)=nint{127(2Re(ShhSvv*) / qsca)}

Byte (8)=nint{127(2Im(ShhSvv*) / qsca)}
Byte (9) = nint\{\text{sign}(\text{Re}(ShvSvv*)) \times 127 \sqrt{2 \mid \text{Re}(ShvSvv*) \mid / \text{qsc}}} \\
Byte (10) = nint\{\text{sign}(\text{Im}(ShvSvv*)) \times 127 \sqrt{2 \mid \text{Im}(ShvSvv*) \mid / \text{qsc}}\}\}

Normally we can recover all the intensity and phase information from these ten bytes for each pixel. This leads to 9 values of 32-bit float point data for each pixel.

The useful terms are extracted by the following

\[
\begin{align*}
\text{ShvShv*} &= \text{qsc} \left[ \frac{\text{Byte}(3) + 127}{255} \right]^2 \\
\text{SvvSvv*} &= \text{qsc} \left[ \frac{\text{Byte}(4) + 127}{255} \right] \\
\text{ShhShh*} &= \text{qsc} - \text{SvvSvv*} - 2\text{ShvShv*} \\
\text{Re} (\text{ShhShv*}) &= 0.5 \text{qsc} \left[ \text{sign} (\text{Byte} (5)) \times \frac{\text{Byte} (5) / 127}{2} \right] \\
\text{Im} (\text{ShhShv*}) &= 0.5 \text{qsc} \left[ \text{sign} (\text{Byte} (6)) \times \frac{\text{Byte} (6) / 127}{2} \right] \\
\text{Re} (\text{ShhSvv*}) &= \text{qsc} \left[ \frac{\text{Byte} (7) / 254}{2} \right] \\
\text{Im} (\text{ShhSvv*}) &= \text{qsc} \left[ \frac{\text{Byte} (8) / 254}{2} \right] \\
\text{Re} (\text{ShvSvv*}) &= 0.5 \text{qsc} \left[ \text{sign} (\text{Byte} (9)) \times \frac{\text{Byte} (9) / 127}{2} \right] \\
\text{Im} (\text{ShvSvv*}) &= 0.5 \text{qsc} \left[ \text{sign} (\text{Byte} (10)) \times \frac{\text{Byte} (10) / 127}{2} \right]
\end{align*}
\]

The dynamic range for the 32-bit float point data is listed in Table 2.1. In this table, it shows that the dynamic ranges of HH and VV co-polarization channels are larger than that of an HV cross-polarization channel. This is normally true for multichannel polarimetric data.

<table>
<thead>
<tr>
<th>Channel</th>
<th>HH</th>
<th>HV</th>
<th>VV</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-25.5(db)~1.29(db)</td>
<td>-27.4(db)~9.86(db)</td>
<td>-25.2(db)~0.2(db)</td>
</tr>
</tbody>
</table>

Before wavelet transform, we need a quantizer to map the 32 float point to 8 bit or 16 bit integer for compression purposes. The quantizer is not easy to choose because of the dynamic data range for each channel data, and the histograms of HH, HV, VV data show that all of three channel data have skewed distribution. One possible choice for the mapping is called a nonuniform quantizer, that is explained in Chapter 4 at the preprocessing step.
Figure 2.3  Single channel SAR image.
Figure 2.4 Intensity images of multichannel SAR data.
Figure 2.5 Phase difference images of mulitchannel SAR data.
Chapter 3

Discrete Wavelet Transform

3.1 Introduction

The wavelet transform is a powerful analytical tool for signal analysis, which provides a theoretical approach and a potentially practical method for solving real world problems. It is especially efficient for transient signal and non-stationary signal analysis. The significant properties of wavelet transform lie in two aspects. One is the property of time frequency representation providing localization both in the time and frequency domain. Another is that of multiresolution where the decomposition of a signal is in terms of the resolution of details.

Before we start, some notations which are used in this chapter are simply explained as follows:

\[ Z \]: set of all the integer numbers
\[ R \]: domain of all the real numbers
\[ L^2 \]: all the functions with finite integral of the square.

3.2 Wavelet Expansion

Before wavelet transform attracts public attention, Fourier transform has been fundamental for signal analysis for a very long time. The basic goal of Fourier series is to decompose the signal into its various frequency components. The basis of Fourier expansion is the sine and cosine function.
Suppose we have a signal \( f(t) \), its Fourier decomposition in terms of sine and cosine is presented as the following:

\[
f(t) = \sum_n \left( a_n \sin(nt) + b_n \cos(nt) \right)
\]  

(3.1)

For a wavelet transform, the basis is a wavelike function with finite support in time, which is called the mother wavelet \( \psi_{jk}(t) \). The signal \( f(t) \) can be expressed as follows:

\[
f(t) = \sum_j \sum_k a_{j,k} \psi_{j,k}(t)
\]

(3.2)

\[
\psi_{j,k}(t) = 2^{j/2} \psi(2^j t - k) \quad j, k \in \mathbb{Z}
\]

(3.3)

Formula (3.3) is a general expression of a wavelet expansion function or wavelet basis. Where \( \mathbb{Z} \) is the set of all integers, and the factor \( 2^{j/2} \) maintains a constant norm independent of scale \( j \).

In addition to the wavelet function \( \psi(t) \) for signal expansion, another function called the scaling function \( \varphi(t) \) is needed for the wavelet expansion as well. The reason for needing this function is related to the filter bank which is essentially the basic scheme for implementing perfect reconstruction and the fast discrete wavelet transform. On the other hand, the requirement of the scaling function makes \( j \) in equation (3.3) change from negative infinity to \( j \geq 0 \) in equation (3.4). The scaling function and wavelet function to wavelet expansion is like the sine and cosine function to Fourier expansion. The combination of these scaling and wavelet function allows a large class of signals to be represented by the following:

\[
f(t) = \sum_{k=-\infty}^{\infty} c_k \varphi(t - k) + \sum_{j=0}^{\infty} \sum_{k=-\infty}^{\infty} d_{j,k} \psi(2^j t - k)
\]

(3.4)

The properties that make the wavelet expansion and wavelet transform very efficient and effective include the following:
The wavelet expansion coefficients $a_{j,k}$ in (3.2), or $d_{j,k}$ in (3.4) drop off rapidly with $j$ and $k$ for a large class of signals. This property is called "being in an unconditional basis" and it is why wavelets are so effective in signal and image compression, denoising.

Wavelet expansion allows for a more accurate local description and separation of signal characteristics. Fourier expansion has an obvious limitation for decomposing only periodic signals that last for all time, and is poor for expressing temporary events and non-periodic signals. It is the localization property of wavelets that allows the wavelet expansion of a transient event to be modelled with a small number of coefficients. This turns out to be very useful in applications.

### 3.3 Continuous Wavelet Transform

If the signal is a function of a continuous variable, the continuous wavelet transform is defined by the following:

$$F(s,T) = s^{-1/2} \int f(t) w(t - \frac{T}{s}) dt$$  
(3.5)

with the inverse transform as

$$f(t) = K \int \frac{1}{s^2} F(s,T) w(t - \frac{T}{s}) dsdT$$

with the normalized constant given by

$$K = \int \frac{|W(\omega)|^2}{|\omega|} d\omega$$  
(3.7)

where $w(t)$ is the basic wavelet, and $s, \tau \in \mathbb{R}$ are real continuous variables. $W(\omega)$ is the Fourier transform of the wavelet $w(t)$. In order for the wavelet to be admissible, $K$ should be less than infinity. The admissibility conditions for the wavelet $w(t)$ for supporting this invertible transform is discussed by Daubechies [16].
For a wavelet with scale \( s \) and time shifting \( \tau \), a signal can be decomposed at different resolution levels and time-frequency properties can be observed. Any wavelet system shows three general characteristics relating to the scale \( s \) and the shifting \( \tau \).

* A high scale leads to a stretched wavelet, and the wavelet transform results correspond to low frequency components and low resolution.
* A low scale leads to a compressed wavelet, and the wavelet transform results correspond to high frequency components and high resolution.
* Shifting simply means delaying the wavelet function \( w(t) \) in time by \( \tau \), the outcome is \( w(t - \tau) \).

### 3.4 Discrete Wavelet Transform

The coefficients in the wavelet expansion (3.4) are called the discrete wavelet transform of the signal \( f(t) \). A more general statement of the expansion (3.4) is given by the following:

\[
f(t) = \sum_{j_0} c_{j_0}(k) \varphi_{j_0,k}(t) + \sum_{j=j_0}^\infty \sum_{k} d_{j}(k) \psi_{j,k}(t)
\]

where \( j_0 \) could be zero, a positive value, or it could be even negative infinity when no scaling functions are used. The choice of \( j_0 \) sets the coarsest scale whose space is spanned by \( \varphi_{j_0,k}(t) \). The rest of the \( L^2(R) \) is spanned by the wavelets which provide the highest resolution details of the signal.

In practice what is given is the samples of a signal, not the continuous signal itself, and the highest resolution is also given when the finest scale is at the sample level.

If the wavelet systems are orthogonal, the coefficients in (3.8) can be calculated by the following inner products:

\[
c_{j}(k) = \langle f(t), \varphi_{j,k}(t) \rangle = \int f(t) \varphi_{j,k}(t) dt
\]
\[ d_j(k) = \langle f(t), \psi_{j,k}(t) \rangle = \int f(t)\psi_{j,k}(t)dt \tag{3.10} \]

If the scaling function is well behaved, then at a high scale, the scaling is similar to a Dirac delta function, and the inner product simply samples the function. In other words, at a high enough resolution, samples of the signal are very close to the scaling coefficients. This is a very useful fact that makes a discrete wavelet transform of any sampled signal possible. While in practice, most engineers deal with a discrete signal instead of a continuous one.

### 3.5 Orthogonal/Biorthogonal Wavelet Transform.

In a wavelet transform domain, there exists two type of wavelets, one is called orthogonal, and the other is biorthogonal. The difference between them includes the following:

1. For an orthogonal wavelet, the analysis scaling function is orthogonal to the analysis wavelet function. Only one set of basis is needed for perfect reconstruction. For a biorthogonal wavelet, the analysis scaling function is not orthogonal to the analysis wavelet function, but to the synthesis wavelet scaling function. Thus two sets of basis are needed for perfect reconstruction.

2. For a biorthogonal wavelet, both the scaling filter and wavelet filter have different lengths, and those filters have the property of the linear phase.

3. For an orthogonal wavelet, the scaling and wavelet filters have the same length, and the linear phase is not achievable.

4. The normalized scaling filters generated from orthogonal wavelets have the unit norm. Therefore the energy is preserved after the transform.

5. The scaling filter generated from the biorthogonal wavelet can not achieve unit energy. The total energy of the transformed coefficients can not have the same amount as that of the input signal, though they can be very close.

6. Only Harr wavelet is an orthogonal wavelet, but can still achieve linear phase.
3.6 Filter Bank and Perfect Reconstruction

The technique of filter banks existed in digital signal processing, especially in filter design for perfect reconstruction, before wavelet theory was proposed. Perfect reconstruction is achieved when the output of the system is a delayed and scaled version of the input signal. Quadratic mirror filters (QMF) are used for this perfect reconstruction scheme, which allows a signal to be split into two downsampled subband signals, and then reconstructed without aliasing. The easiest way to understand what the filter banks mean and how perfect reconstruction is achieved is to study Figure 3.1, which shows a two-channel perfect reconstruction filter bank scheme.

![Diagram of a two-channel perfect reconstruction filter bank scheme.](Image)

Figure 3.1 Two-channel perfect reconstruction filter bank scheme.

In Figure 3.1, H1 and H0 are analysis filters with H1 representing a high pass filter and H0 a low pass one. Similarly, G1 and G2 represent synthesis filters, which correspond to H1 and H0, respectively. We would not like to define G1 and G2 as low pass or high pass filters, but use them only to achieve the output signal to be the same as the delayed input signal. $\downarrow 2$ and $\uparrow 2$ represent downsampling and upsampling by the factor of 2. After downsampling, the output signal $y_0$ and $y_1$ are only half the length of $x$. Similarly, after upsampling, the output data length becomes the same as that of $x$. Later in this chapter, we discuss about how to deal with image edges to make the transform samples exactly half the size of an image.
Before wavelet theory was constructed, filter banks were applied to many applications. The close relationship between wavelet transform and the filter banks lies in the fact that filtering by the low pass filter $H_0$ is the same method as obtaining the scaling coefficients $c_j(k)$, and filtering by a high pass filter $H_1$ is the same method as obtaining the wavelet coefficients $d_j(k)$. It is proven that the analysis filter bank scheme is an efficient way to implement discrete-time wavelet transform.

Although the two-channel filter bank has been studied intensively, the $M$ channel perfect reconstruction filter bank is also available, and its application can be found in the communication system.

### 3.7 Wavelet Packets

A wavelet packet differs from a wavelet transform in that it decomposes the signal not only to the low frequency bands, but also to the medium and high frequency bands. Through this, the frequency response of the signal can be better understood, and signals containing rich medium and high frequency behaviours can be completely investigated. Figure 3.2 below shows the full binary tree for two level wavelet packet decomposition.

![Figure 3.2](image-url)
When both the lowpass and highpass bands at all stages are split up, the resulting filter bank structure is similar to a full binary tree, as in Figure 3.2. Notice the meaning of the subscripts in the signal space. The first integer subscript is the scale of the space. Each following subscript is a zero or one. A “zero’ indicates going through a lowpass filter (scaling function decomposition), and a “one” indicates going through a lowpass filter (wavelet decomposition).

The frequency response for the two-band wavelet packet decomposition is shown in Figure 3.3.

![Figure 3.3 Frequency response for the two-band wavelet packet filter bank.](image)

Full packet decomposition generates a valid packet basis system and allows for a flexible tiling of the time scale plane. The pruning algorithm used for best basis selection can be done based on the entropy at each level. The least entropy tree branch is the best basis for compression purposes, based on the information theory.

### 3.8 Wavelet Image Compression

Data compression algorithms can usually be classified as lossless and lossy compression. Lossless compression such as Huffman coding, run-length coding and predictive coding are used when exact reconstruction of the original data is required. Lossy compression algorithms such as transform coding, and vector quantization are used for applications where some levels of the degradation of the data are tolerable. This
section deals with some successful compression algorithms, such as EZW, SPIHT and JPEG2000, all belonging to wavelet-based image coder.

### 3.8.1 Symmetric Extension

Here we discuss the image border extension issue existing in wavelet based image compression algorithms. As we know, the finite length of an input signal or image usually produce problems when being processed by a filter bank. The difficulty is how to handle filtering at the image borders and making the transform result non-expansive.

There are two general approaches that can be used to address this problem. The first is called periodic extension. The result of applying a linear translation-invariant (LTI) filter to the input signal produces the same period as the input signal. This approach is equivalent to circular convolution. Although a periodic extension is nonexpansive, it has two drawbacks. First, it introduces an artificial discontinuity into the input signal at the edges. This generally increases the variance of each subband and forces the coding scheme to waste bits coding preprocessed artifacts: the effects on transform coding gain are more pronounced if the decomposition involves several levels of a filter cascade. Second, the applicability of the periodic extension is limited to signals whose length is divided by the decimation rate, and if the filter bank decomposition involves L levels of cascade then the input length must be divisible by $M^L$. This causes problems in applications for which the coding algorithm designer is not allowed to change the size of the input signal.

The second approach is called symmetric extension. This approach is based on the use of linear phase filters, which impose a restriction on the allowable filter banks (biorthogonal) for subband coding applications. The improvement in rate-distortion performance achieved by going from periodic to symmetric extension was first demonstrated in [38] in the range of 0.2-0.8 dB PSNR for images coded around 1 bpp.
With the symmetric extension approach, we generate a linear phase input signal by forming a symmetric extension of the finite-length input vector. A signal can either be symmetric about one of its samples, or about a point midway between two samples. These two cases are referred to as whole-sample symmetry (WS) and half-sample symmetry (HS), respectively. Examples of both are given in Figure 3.4.

![Whole-sample symmetry](image1)

(a) Whole-sample symmetry

![Half-sample symmetry](image2)

(b) Half-sample symmetry

Figure 3.4. Signal symmetries.

Using the convolution theorem, it is clear that applying a linear phase filter to a symmetric signal results in a filtered signal that is also symmetric at the edges. The key to implementing the symmetric extension method in a linear phase sense without increasing the size of the signal is to ensure that the subbands remain symmetric after downsampling. When this condition is met, half of each symmetric subband will be redundant and can be discarded with no loss of information. If everything is set up correctly, the total number of transform domain samples that must be transmitted will be exactly equal to the length of the input signal, meaning that the transform is nonexpansive. For different types of signal symmetry and filter symmetry, detailed processing methods can be found in [37], [38]. We use whole-sample symmetry in our compression algorithm to handle the image edges, as shown in next chapter.
3.8.2 Embedded Zerotree Wavelet Algorithm (EZW)

The embedded zerotree wavelet algorithm [8] was proposed by J. M. Shapiro in 1993. This coding scheme is a simple and effective image compression algorithm. The distinguished property of EZW is that the bits in the bit stream are generated in order of importance, and it yields a fully embedded code. It does not require any kind of training like vector quantization does, and it is a universal coding scheme which does not require the prior knowledge of the image source. The EZW algorithm is based on the following features.

- A discrete wavelet transform is utilized to provide a compact multiresolution representation of the image.
- Zerotree coding provides a compact multiresolution representation of significant maps, which are binary maps indicating the positions of significant coefficients. Zerotrees allow the successful prediction of insignificant coefficients across scales to be efficiently represented as part of exponentially growing trees.
- Successive approximation provides a compact multiprecision representation of significant coefficients and facilitates the embedding algorithm.
- Large wavelet coefficients are deemed more important than smaller coefficients regardless of their scale, as well as spatial locations.
- Adaptive multilevel arithmetic coding [39] provides a fast and efficient method for the entropy coding of symbols, and requires no training or prestored tables.
- The algorithm runs sequentially and stops whenever a target bit rate or a target distortion is met. A target bit rate can be met exactly, and an operational rate vs distortion function can be computed point by point.

A wavelet coefficient x is said to be insignificant with respect to a given threshold T if |x|<T. The zerotree is based on the hypothesis that if a wavelet coefficient at a coarse
scale is insignificant with respect to a given threshold, then all wavelet coefficients in the
same spatial location at finer scales are likely to be insignificant with respect to T. More
specifically, in a hierarchical subband system, except for the highest frequency subbands,
every wavelet coefficient at a given scale can be related to a set of coefficients at the next
finer scale of similar orientation. The coefficient at the coarser scale is called the parent,
and all coefficients corresponding to the same spatial location at the next finer scale of a
similar orientation are called children. For typical subband decomposition, the parent-
children dependencies are shown in Figure 3.5.

![Figure 3.5 Parent–child dependencies of the subbands.](image)

The arrows in Figure 3.5 points from the subband of the parents to the subband of
the children. The lowest frequency subband is at the top left, and the highest frequency
subband is at the bottom right. Also shown is a wavelet tree consisting of all the
descendants of a single coefficient in subband HH3.
A scanning of the coefficients is performed in such a way that no child node is scanned before its parent. The scanning pattern for three-scale pyramid can also be seen in Figure 3.6. Note that each coefficient within a given subband is scanned before any coefficient in the next subband. For wavelet packets, the scanning pattern is the same.

Figure 3.6 Scanning order of the subbands.

Given a threshold $T$, a coefficient $x$ is said to be an element of a zerotree for threshold $T$ if itself and all of its descendents are insignificant with respect to $T$. The significant map can then be efficiently represented by using four symbols: 1) zerotree root; 2) isolated zero, which means the coefficient is insignificant but has some significant descendent; 3) positive significant; and 4) negative significant.

EZW is designed to reduce the cost of encoding the significant map so that for a given bit budget, more bits are available to encode expensive significant coefficients. In
general, large fractions of the insignificant coefficients are efficiently encoded as part of a zerotree.

3.8.3 Set Partitioning in Hierarchical Trees (SPIHT)

Embedded zerotree wavelet (EZW) coding, introduced by J. M. Shapiro, is a very effective and computationally simple technique for image compression. Moreover, a different implementation, based on set partitioning in an hierarchical tree (SPIHT) [9], provides even better performance than EZW by partial ordering of the coefficients by magnitude with a set partitioning sort algorithm, ordered bit plane transmission, and exploitation of self-similarity across different scales.

After the wavelet transform, all the wavelet coefficients are said to be significant or insignificant according to their magnitude, if they are is larger or less than a threshold. All the significant information is stored in three ordered lists, called list of insignificant sets (LIS), list of insignificant pixels (LIP), and list of significant pixels (LSP).

During the sorting pass, the pixels in the LIP are tested, and those become significant are moved to LSP. Similarly, sets are sequentially evaluated following the LIS order, and when a set was found to be significant, it is removed from the list and partitioned. The new subsets with more than one element are added back to the LIS, while the single-coordinate sets are added to the end of the LIP or to the LSP, depending on whether they are significant or insignificant, respectively. The LSP contains the coordinates of the pixels that are visited in the refinement pass.

The logarithmic distributed quantization is used in the coding algorithm by setting a threshold at each sorting pass and half the threshold at the next pass. In order to obtain the initial threshold, we need to first find the maximum absolute magnitude wavelet coefficient $C_{ij}$. Second, we need to calculate the maximum total quantization level $N$ by the following:

$$N = \text{floor}(\log_2(C_{ij})) \quad \text{floor means lower integer.}$$

(3.11)
Thus, the initial threshold $T_0$ is given as follows:

$$T_0 = 2^N$$

Compared to the initial threshold used in EZW $T_0 = M2^E$, with $M$, a parameter that needs to be determined, and $E$, an integer, the threshold is easily obtained in SPIHT without any unknown parameters.

Since SPIHT is a very successful image compression algorithm using the wavelet transform, our coding scheme is compared to it for the compression performance evaluation.

### 3.8.4 JPEG2000

JPEG [40] is the acronym for Joint Photographic Experts Group. It is an international standard for the digital compression and coding of continuous-tone still images. Based on the Discrete Cosine Transform (DCT), JPEG can produce a sharp degradation in the quality of the reconstructed image at low bit rate. To correct this and other shortcomings, JPEG2000 is a new standard for still image compression. Based on the Discrete Wavelet Transform (DWT), JPEG2000 represents advances in image compression technology where the image coding system is optimised not only for efficiency, but also for scalability and interoperability in network and mobile environments.

The JPEG2000 [41], [42], [43] standard provides a set of features that are important for many image applications, such as internet, colour facsimile, printing, scanning, digital photography, remote sensing, mobile, medical imagery, digital libraries and archiving, and E-commerce. Each application area imposes some requirements that the standard, up to a certain degree, should fulfil. Some of the most important features are the following:

* The source image is decomposed into components. Then, the image components are (optionally) decomposed into rectangular tiles. The tile-component is the basic unit of the original or reconstructed image.
The term tiling refers to the partition of the original (source) image into rectangular non-overlapping blocks (tiles), which are compressed independently. All operations, including wavelet transform, quantization, and entropy coding are performed independently on the image tiles. Tiling reduces memory requirements, and since they are reconstructed independently, they can be used for decoding specific parts of the image instead of the whole image.

As expected, tiling affects the image quality both subjectively and objectively. Smaller tiles create more tiling artifacts compared to large tiles. In other words, large tiles perform visually better than smaller tiles.

* A wavelet transform is applied on each tile. The tile is decomposed into different resolution levels. There are two types of wavelet transforms that are supported by JPEG2000. One is convolution based, and the other is lifting based. Convolution based filtering consists of performing a series of dot products between filter masks and the boundary extended 1-D signal. Lifting based filtering is also called the integer wavelet transform, and it is implemented by very simple operations for which alternately odd sample values of the signal are updated with a weighted sum of even sample values, and even sample values are updated with a weighted sum of odd sample values. The integer wavelet transform is reversible, and this property makes it a good candidate for lossless compression.

The decomposition levels are made up of subbands of coefficients that describe the frequency characteristics of local areas of the tile components, rather than across the entire image component.

* The subbands of coefficients are quantized and collected into rectangular arrays of code blocks. Uniform scalar quantization with a dead-zone around the origin and trellis coded quantization are two candidates in the JPEG2000 standard. For scalar quantization, the quantization step size is represented relative to the dynamic range of each subband. In other words, the JPEG 2000 standard supports a separate quantization step-size for each subband. For reversible compression, the quantization step-size is required to be one.
*The bit planes of the coefficients in a code block (i.e., the bits of equal significant across the coefficients in a code block) are entropy coded. Entropy coding is achieved by means of an arithmetic coding system that compresses binary symbols.

* The encoding can be done in such a way that certain regions of interest (ROI) can be coded at a higher quality than the background. The functionality of ROI is important in applications where certain parts of the image are of higher importance than others. In such a case, these regions need to be encoded at a high quality than the background. During the transmission of the image, these regions need to be transmitted first, or at a higher priority.

* Overall, the JPEG 2000 standard offers the richest set of features in a very efficient way, and within a unified algorithm. However, this comes at the price of additional complexity as compared to JPEG. This might be perceived as a disadvantage for some applications.
Chapter 4

Single Channel SAR Image Compression

4.1 Introduction

SAR image data can provide information about the surface of the earth. The data volume associated with the earth information is so huge that SAR data compression becomes important in transmitting and archiving.

At present, most image compression algorithms are designed for standard test images or optical images. The most popular compression standard for a still image is the JPEG/JPEG2000 standard that aims to compress traditional images. Applying those compression schemes to SAR data generally does not lead to ideal compression results in terms of the signal to noise ratio. What hinders SAR data compression is that those compression schemes do not take into consideration SAR data characteristics, such as the speckle phenomena, and significant high frequency components that are highly related to terrain boundaries and terrain textures.

In this paper, a new algorithm called the Wavelet Packet-based Embedded Block coding (WPEB) scheme is proposed for SAR data compression. This algorithm combines the properties listed below.

(1) Wavelet packet decomposition is adopted to exploit middle and high frequency components associated with SAR data.
(2) The block coding scheme is adaptively utilized to improve the DWT coefficient coding efficiency by allocating more bits to regions of importance with higher information content. (e.g. more contrast, edges).

(3) Speckle reduction is built into the bit allocation scheme by using a wavelet transform denoising method.

This chapter is organized in a sequence where the overview of the algorithm is introduced first, followed by the experimental results. Finally, the comments are given based on the experimental results, as well as some concerns.

4.2 Compression Algorithm Overview

WPEB belongs to the class of lossy transform compression algorithm. The main steps of the coding scheme include: wavelet transform, scalar quantization, and entropy coding. Some ideas are borrowed from SPIHT [9] and EBCOT [10]. Since SPIHT performs very well for many image compression applications, and recently EBCOT (Embedded Block Coding with Optimal Truncation) shows many interesting properties, such as it can produce the bit-streams which are SNR and resolution scalable, and which also have a random access attribute whereby independent portions of the bit-stream correspond to different spatial regions of the image. Both algorithms are combined into our coding scheme for SAR data compression. It turns out that SAR data has different signal-noise ratio for different image blocks, which makes WPEB more appropriate for this type of data compression.

Our coding scheme includes the wavelet packet transform, embedded block coding, speckle reduction, and optimal bit allocation. We consider the SAR data spectrum property in bit allocation in order to achieve better compression performance and make this algorithm more flexible when compressing different types of SAR data (sea ice, city, mountain, etc), and meet different compression evaluation criteria. The encoding and
decoding diagram for WPEB is shown in Figure 4.1. At the encoder, the discrete wavelet packet is first applied on the source image data. The transform coefficients are then quantized and entropy coded before forming the output code sequence. The decoder is the reverse of the encoder. The decoding stream consists of entropy decoding, inverse quantization, and an inverse wavelet packet.

Figure 4.1 General block diagram of WPEB. (a) Encoding diagram, (b) Decoding diagram.
4.3 Data Preprocessing

Data preprocessing aims at preparing the source data for the convenience of core processing (wavelet/wavelet packet transform). Some operations which are required at this stage may include the following:

(1) Logarithmic transform.

For SAR images, speckle noise typically can be modeled as a multiplicative identical independent distribution (iid) Gaussian noise [26]. The logarithmic transform of a SAR image converts the multiplicative noise into additive noise.

For a digitized image, $z(i,j)$ is defined as the $(i,j)$th pixel of the image. The pixel level of a SAR image can be written as follows:

$$z = x e + \lambda$$

where $x$ is the desired information, $e$ is the multiplication noise, and $\lambda$ is the receiving additive noise, which is usually smaller than the first term $xe$ and ignored. For a logarithmically transformed SAR image, the speckle is approximately Gaussian additive noise.

$$\bar{z} = \bar{x} + \bar{e}$$

Based on the fact that the Gaussian additive noise model is also assumed for the wavelet denoising method, logarithmic transform is necessary to prepare the SAR data for the wavelet transform.

(2) Quantization

As shown in Chapter 2, some SAR image data may be in the format of 32 float-point data. For compression algorithms, such as EZW, SPIHT, integer is the only data format for input. Quantization may be required in this case to transform the 32 float-point data to 8-bit data.
Our method of transforming the CEOS format 32 float-point data into the 8 bit integer is implemented by applying a log transform to the original 32 float-point data. Since the log transform is a type of companded quantization [12], it makes the interval of the decision boundary small in which the input lies with high probability. It also tends to equalize the probability of the signal being in any given quantization level.

Coincidently, at our case, the logarithmic is not only utilized to change the multiplicative speckle noise to additive noise, but is also used as a quantizer.

(3) Normalization

Image data normalization is applied to make the input image mean approximately zero, so that at each wavelet decomposition level, each subband has approximately zero mean. In this case, the statistical distribution of wavelet coefficients is symmetric around zero. The symmetric wavelet coefficients benefit encoding in a way that the probability of positive and negative sign is equal with the maximum entropy for the sign symbol.

4.4 Wavelet Packet Decomposition

Since SAR data usually have significant middle and high frequency components, as in texture regions, this makes the wavelet packet suitable for SAR data compression. Wavelet packet decomposition differs from standard wavelet transform by allowing the decomposition of the upper frequency bands as well as the lower ones. Through this method, wavelet packets can be used to achieve a more accurate representation of medium and high frequency information in SAR images.

The difference between wavelet transform and wavelet packet is shown in Figure 4.2. Here, we consider an image that is the 2-d signal with two levels of decomposition. In Figure 4.2(b), only LH1 is decomposed further. As a matter of fact, every subband, such as HL1, HH1, could be decomposed further if it is required to do so.
In previous work of SAR image compression [11], [44], [45], texture analysis for the compression is based on subband energy at different decomposition levels. A constant $C$ is needed as the criterion for applying wavelet packet at each level. Energy $E$ is defined as the following:

$$E = \frac{1}{MN} \sum_{m=1}^{M} \sum_{n=1}^{N} |x(m,n)|^2$$

(4.3)

where $x(m,n)$ is the wavelet coefficient at any subband at any decomposition level. $E_{\text{max}}$ is the maximum wavelet coefficient energy at a given decomposition level. It is known that at any decomposition level, the most energy resides in the low-low band, so $E_{\text{max}}$ represents the energy of the low-low band at each level. The general idea of applying a wavelet packet is that when $E$ is less than $C E_{\text{max}}$, wavelet packet is applied.
In [11], the orthonormal wavelet DB4 is used in image decomposition. For the orthogonal wavelet, Parseval's theorem holds so that energy can be a criterion for wavelet decomposition. The disadvantage of using an orthogonal wavelet for image compression is that this type of wavelet cannot achieve a linear phase so that symmetric extension is not available. Perfect reconstruction is not fully achievable because of the distortion at the border of each decomposition level.

Here, a biorthogonal wavelet is selected to achieve the linear phase and symmetrical extension. However, Parseval's theory does not hold in this case, so that the criteria for a wavelet packet should be changed to the variance for each subband. Since all the test images are normalized to a zero mean before the discrete wavelet transform is applied, variance becomes a valid criterion.

Normally threshold C should be a value below 1. This is because the variance of the low-low subband is normally larger than those of LH, HL, and HH subbands. As for how to choose C, we propose that, for different types of testing data, C should be chosen so that significant middle or high frequency components can be preserved at reconstruction. The larger this constant is, the lower the compression ratio that can be achieved.

For HL, LH, and HH, we set different $C_{HL}$, $C_{LH}$, and $C_{HH}$. Normally, we choose the following:

$$C_{HL} = C_{LH}, \quad C_{HL} > C_{HH}.$$  

LH and HL usually have the same variance, and this is why $C_{HL}$ is supposed to be the same as $C_{LH}$. Since for normal images, the variance of HH subband is much lower than LL, and lower than LH and HL as well, then the threshold for HH is supposed to be less than $C_{HL}$ and $C_{LH}$.

Therefore, at each decomposition stage, the wavelet packet subroutine proceeds as follows:

(1) Initialize $C_{HL}$, $C_{LH}$, $C_{HH}$, set maximum decomposition level.
(2) Apply wavelet transform to the data, and get four subband coefficients at LL, LH, HL, and HH.

(3) Calculate the variance for each subband;

(4) If variance (LH) / variance (LL) > $C_{LH}$, then apply the wavelet packet to LH.

(5) If variance (HL) / variance (LL) > $C_{HL}$, then apply the wavelet packet to HL.

(6) If variance (HH) / variance (LL) > $C_{HH}$, then apply the wavelet packet to HH.

(7) Apply wavelet transform to LL.

(8) If decomposition level is reached, stop; otherwise go to (3).

4.5 Embedded Block Coding

It is known that block transform coders enjoy success because of their low complexity in implementation and their reasonable performance. The most popular block transform coder leads to the image compression standard JPEG which utilizes the 8x8 Discrete Cosine Transform (DCT). While for the upcoming JPEG2000 [40], 64x64 and 32x32 code-block size are recommend for the Discrete Wavelet Transform (DWT).

The general idea for embedded block coding is that we can divide the image into blocks, so that wavelet decomposition, quantization and coding can be applied to smaller sections of the image, rather than to the whole image at once. In this way, the mean squared error can be optimized within each block, which in turn brings the capability of flexible signal noise ratio and resolution.

Generally, large image blocks lead to smaller mean squared error in the sense that the correlation among larger block samples could be exploited better in the compression. This is true for optical images whose correlation coefficient is as high as 0.95 [24]; however, for SAR images the correlation is not as high as that of the optical image, and the correlation coefficient is probably about 0.7.

By using block coding, the advantages include the following:
(1) Different statistics of each block can be recognized. The statistics, such as the dynamic data range and entropy, affect the coding performance in the way that the data range is associated with the quantization step and entropy is associated with the maximum compression ratio for the given image block.

(2) By adopting the idea of block coding, it is possible to achieve the properties of flexible resolution for different blocks. This is done by assigning different compression ratio to different scene content under the constraint that the overall compression ratio should be the given.

On the other hand, the disadvantage of blocking is the discontinuities or artifacts across the block boundaries, especially at very low bit rates. Although some techniques, such as boundary overlapping and boundary extension, can overcome this problem to a certain extent, this increases the computational complexity greatly.

Despite the disadvantage of its block artifacts, block coding might still work for SAR image compression because a frequently used compression ratio for SAR image compression is around 10, and under this compression ratio, block artifacts are not very significant. Higher compression ratio is not acceptable for SAR image compression because the applications of those images, such as object detection, scientific study is impossible when information is mostly lost for highly compressed SAR images. Thus the main problem associated with SAR compression becomes how to preserve as much information at a given compression ratio as possible instead of determining what the best compression ratio level is, as the latter question is already known. From the applications point of view, this problem is related to how to preserve as much application-based information as possible at a given compression ratio, or how to make the compressed image contain the maximum specific information for the specific application.

This is a real problem for real world. For example, if the SAR image is used for sea-ice classification, the best compression scheme should preserve texture regions that represent different types of ice (first year, multi year, etc). Some regions, where the
image consists of land or ground, are out of our main concern, and thus we can assign less bits to those regions.

The basic idea associated with applying an embedded blocking algorithm is that we can design a bit allocation scheme to adaptively assign different numbers of bits to image regions based on the regions of importance.

There are two ways to apply blocking. One is to use some existing algorithm to automatically divide the images into many blocks. The other is that users can define blocks of interest as the prior knowledge for compression. Here in our experiments, the image blocks are defined by hand to demonstrate the blocking effect.

The most frequently used blocking algorithm is the quadtree decomposition that involves subdividing an image into blocks that are more homogeneous than the image itself. It is very useful as the first step in adaptive compression algorithms. In addition, image blocking could be done manually in order to choose regions of interest easily.

The most important disadvantage of block coding is the block artifacts at the block edges. It needs a serious concern on how to define the block size. The advantages of using smaller blocks include preserving the block variability better and the efficiency of computation, but the artifacts may be more apparent.

4.6 Bit Allocation Scheme With Speckle Reduction.

This scheme is designed to assign different numbers of bits to blocks, which are the result of embedded block coding introduced in Section 4.5. Since each encoding design is done at a bit plane, we focus on finding out how many bits should be assigned to each block.

Now, let us explain the bit allocation problem mathematically.

Since the image is composed of a collection of code blocks, \( B_i \), whose embedded bit stream may be truncated at rate \( R_i^n \), the corresponding contribution to the distortion in
the reconstructed image is denoted as $D_i^n$ for each truncation point $n_i$. The relevant distortion matrix is additive as follows:

$$D = \sum_i D_i^n \quad \text{with constraint } R$$

(4.4)

where $D$ is the distortion for the reconstructed image, and $n_i$ is the truncation point for $B_i$. From the rate distortion optimization point of view, the problem is to find the optimal selection of $n_i$ so as to minimize the overall distortion $D$, subject to a constraint, $R$, where $R$ is the given bit rate for the whole image.

If the image is compressed as a whole, the number of blocks is one, and the overall distortion is equal to the block distortion. For a specific encoding and decoding algorithm, the rate-distortion is a curve with only one parameter, being the compression ratio (bpp) or the truncation point $n_i$, as shown in Figure 4.3 (a). Given a specific compression ratio (1bpp) for the overall image, only one point or one distortion value can be obtained.

If the image is divided into two blocks with same size, then overall optimization can be obtained by searching the minimum MSE in two dimensions defined as the compression ratio of each image block, bpp1 and bpp2, as shown in Figure 4.3(b). If bpp1 and bpp2 are the total allowed bpp, then we have $\text{bpp} = (\text{bpp1} + \text{bpp2}) / 2$.

It is found out that the distortion curve for a specific compression ratio for the whole image is in the shape of “U”, with one minimum at the bottom of “U”. The reason for the U shape is that for each block, the rate-distortion curve is convex, which means the slope of the curve decreases as the bpp value increases, so the slope decreases at the left side of the U shape, and increases at the right side of the U shape. The other property is that there is no guarantee the U shape is symmetric. Symmetry happens when two blocks have the same rate-distortion curve.
If the image is divided into 3 blocks, then the overall optimization can be obtained by searching the minimum MSE on a surface, as shown in Figure 3.4 (c). Also we have \( b_{ppt} = \frac{b_{ppl} + b_{pp2} + b_{pp3}}{3} \).

From mathematical point of view, the distortion function can be a continuous function. However, in applications when the numerical method is applied to find the minimum value, a specific step (in bpp) should be given at the initial stage of the search.

Figure 4.3. Rate-distortion curves for three different cases. (a) The whole image is not divided. (b) The whole image is divided into two blocks. (c) The whole image is divided into three blocks.
There are some properties that require consideration when assigning bits to different blocks. The bit allocation scheme can count on some features presented in the upcoming sections.

4.6.1 Blocks of Importance

For the block, which contains important features for classification or detection, more bits should be assigned to it.

\{B_1, B_2, \ldots, B_n\} is the bit set to all image blocks, B_n is the number of bits for block n. Given overall bits R, we then have the following:

\[ R = B_1 + B_2 + \ldots + B_n \]  \hspace{1cm} (4.5)

Suppose that each block has the same image size, and each block is compressed at the same compression ratio, then we could expect \( B_1 = B_2 = \ldots = B_n \). However, it is possible in some applications (such as the region of interest feature in JPEG2000) that the number of bits to every block is different. Especially when a different image block has totally different local statistics. In this case, we could define the weight \( S_n \) for block n to represent its importance as follows:

\[ S_n = \frac{R}{(N*B_n)} \]  \hspace{1cm} (4.6)

When the parameter \( S_n \) is greater than 1, it tells that this block is assigned more bits than the average, its activity level (using variance as the criterion) is probably higher than the average, and this image block is more difficult to compress. As a matter of fact, the latter two are pairs that would normally occur simultaneously.

4.6.2 Speckle Reduction

For some blocks contaminated heavily by speckle noise, we allocate less bits to it so that we can save precious bits for blocks with high information content.
In wavelet transform domain, speckle reduction is equal to denoising. The basic idea of thresholding the wavelet transform coefficients are proposed by Donoho [13]. Assume a finite length of the observed signal has the following form:

\[ y_i = x_i + \varepsilon n_i \quad i = 1, \ldots, N \]  \hspace{1cm} (4.7)

Where \( x_i \) is a finite length signal, \( n_i \) is an i.i.d. zero mean, white Gaussian noise with standard deviation \( \varepsilon \). The goal is to recover the signal \( x \) from the noisy observation \( y \). Let \( W \) be left invertible wavelet transformation matrix of the discrete wavelet transform, then (4.7) can be written in the transformation domain as follows:

\[ Y = X + N. \]  \hspace{1cm} (4.8)

where \( Y \) is the wavelet transform of \( y \), and \( Y = Wy \).

Let \( \hat{X} \) denote an estimate of \( X \), based on the observation of \( Y \). We consider the following diagonal linear projections:

\[ \Delta = \text{diag}(\delta_1, \ldots, \delta_N), \quad \delta_i \in \{0, 1\}, \quad i = 1, \ldots, N \]  \hspace{1cm} (4.9)

Which means that the estimate \( \hat{X} \) is obtained by simply keeping or zeroing the individual wavelet coefficients. Therefore, we have the estimate:

\[ \hat{x} = W^{-1} \hat{X} = W^{-1} \Delta Y = W^{-1} \Delta Wy \]  \hspace{1cm} (4.10)

The frequently used denoising methods are called hard-thresholding and soft thresholding, which are shown in Figure 4.4. Both thresholding methods are applied to the wavelet coefficients of all subbands.

The hard thresholding is applied as follows:

1. Compute DWT \( Y = Wy \) \hspace{1cm} (4.11)

2. Perform thresholding in wavelet domain, where \( t \) is the threshold with a non-negative value.

\[ \hat{X} = \begin{cases} Y, & |Y| \geq t \\ 0, & |Y| < t \end{cases} \]  \hspace{1cm} (4.12)
(3) Compute the inverse DWT. \( \hat{x} = W^{-1} \hat{X} \).

In fact, the hard thresholding is obtained at the encoding process depending on the given compression ratio.

The soft thresholding is implemented by changing the second step, as follows:

\[
\hat{X} = \begin{cases} 
\text{sgn}(Y)(|Y| - t), & |Y| \geq t \\
0, & |Y| < t 
\end{cases}
\]  

(a) Hard Thresholding \hspace{1cm} \text{(b) Soft Thresholding}

Figure 4.4 Denoising to wavelet coefficients.

In summary, hard thresholding yields better results than soft thresholding in terms of rms error. It is clear that the observation values \( y_i \) is a better estimate for the real value \( x_i \) than a shrunk value in a zero mean noise scenario.

4.6.3 Dynamic Data Range

It is intuitively correct that each block has a different dynamic data range, we associate another Dynamic Data Range (DDR) set \{d1, d2, ...dn\} to represent this
information. Since wavelet transform is a linear transform, the image block, which has a large data range, also contains a large range of wavelet coefficients. In order to code those coefficients with as little distortion as possible, we need to first increase the quantization level by decreasing the decision step, and second allocate more bits to this block.

4.7 Experiment Results

Experiments are mainly done to two test images, referred to Figure 4.13. The comparison between WPEB and SPIHT algorithms is used, and the rate distortion curve is drawn to fully demonstrate the merits of our compression algorithm. The first test image has two image blocks with the same contrast, while the second test image has two image blocks with different contrast. These two test images are chosen to investigate how our method can improve compression performance by using blockings.

4.7.1 Case One

In the first experiment case, the test image with the image size of 512X256 is divided into two blocks of the same size (256X256). These two blocks have very similar variance, with the first block 2800, and the second 2840. In addition, the whole test image variance is 2850.

4.7.1.1 Variance Ratio

The variance ratios for the test image using a three level biorthogonal wavelet transform are listed in Table 4.1.
Table 4.1 Variance ratio for the first test image.

<table>
<thead>
<tr>
<th>LH1/LL1</th>
<th>0.114</th>
</tr>
</thead>
<tbody>
<tr>
<td>HL1/LL1</td>
<td>0.056</td>
</tr>
<tr>
<td>HH1/LL1</td>
<td>0.015</td>
</tr>
<tr>
<td>LH2/LL2</td>
<td>0.179</td>
</tr>
<tr>
<td>HL2/LL2</td>
<td>0.138</td>
</tr>
<tr>
<td>HH2/LL2</td>
<td>0.059</td>
</tr>
<tr>
<td>LH3/LL3</td>
<td>0.246</td>
</tr>
<tr>
<td>HL3/LL3</td>
<td>0.265</td>
</tr>
<tr>
<td>HH3/LL3</td>
<td>0.133</td>
</tr>
</tbody>
</table>

In Table 4.1, the ratio for the associated LL subband is calculated at each level. For example, the ratio for HL1 is defined as the variance of HL1 over the variance of LL1, and the ratio for HL2 is the variance of HL2 over the variance of LL2, and so forth.

From Table 4.1, some properties can be observed:

(1) It is clear that the ratio increases as the decomposition level increases. This suggests that $C_{HL}$, $C_{LH}$ and $C_{HH}$ are level dependent; for example, the threshold variables may be multiplied by a factor of two as the level is increased by one.

(2) The ratio for HL is almost equal to that of LH at the third level.

(3) The LL band still contains the most significant information at each level as compared to three other bands at the same level.

The best way to explain why those properties exist is to check the spectrum of the test images. As we know the wavelet transform is identical to the hierarchical subband system, where the subbands are logarithmically (base 2) spaced in frequency and represent an octave-band decomposition (refer to Figure 4.5).
To begin the wavelet decomposition, the image is divided into four subbands which are denoted as LL1, LH1, HL1 and HH1. Each coefficient represents a spatial area, all of which cover the size of the original image. The LL1 represents the low frequency bandwidth corresponding to $0 < |\omega| < \pi$, where the high frequencies represent the band from $\pi/2 < |\omega| < \pi$. The four subbands arise from the separable application of vertical and horizontal filters. The subband LL1 is further decomposed and we obtain four subbands labeled as LL2, LH2, HL2 and HH2. The LL2 represents the low frequencies approximately corresponding to $0 < |\omega| < \pi/4$, both in the vertical and horizontal direction. HL2 represents $\pi/4 < |\omega| < \pi/2$ in a vertical direction corresponding to the first letter H, and $0 < |\omega| < \pi/4$ in a horizontal direction corresponding to the second letter L. The frequency bandwidth for other subbands can be discerned in the same way.
4.7.1.2 1-D Spectrum

Since the variance is the same amount as the energy, and Fourier transform also preserves the energy, therefore, the spectrum of the test image can reflect the variance. To check the data in Table 4.1, using matlab fft function and averaging along the horizontal and vertical direction, we can get the spectrum of test data in Figure 4.6.

The variance ratio of LH1 over LL1 is identical to the ratio of the area from cell 33 to 65, over the area from cell 1 to 32. The variance ratio of LH2 over LL2 is identical to the ratio of the area from cell 17 to 31, over the area from cell 1 to 16. This explains why the variance ratio increases as the decomposition level increases.

\[
\frac{\text{Variance (LH1)}}{\text{Variance (LL1)}} = \frac{\text{area} (\text{cell33} \sim \text{cell65})}{\text{area} (\text{cell1} \sim \text{cell32})}
\]  

(4.14)

Generally, for images with a lot of high frequency components, there is more benefit to applying further decomposition to the higher frequencies, and the wavelet packet can accomplish this function as opposed to the normal wavelet decomposition. In fact, wavelet packets make the wavelet coefficients more compact at those decomposition levels, so that when using magnitude ordering coding algorithm like EZW or SPIHT, the larger coefficients are coded first and the information at those subbands are preserved well at the reconstruction stage.

The disadvantage of using wavelet packet decomposition for data compression lies in computational complexity. Therefore there exists the trade-off between computational complexity and compression efficiency. In spite of the disadvantages of wavelet packet to the overall algorithm, it definitely provides an efficient tool for the analysis of signals with significant high frequency components, such as textures.
In Figure 4.6, the horizontal spectrum represents the SAR data in the range direction, and the vertical one represents the data in the azimuth direction. Since at the radar signal processing stage, the Kaiser window is used to taper the spectrum of the image, therefore the data spectrum drop off rapidly in order to depress the noise level associated with the original raw data. From Figure 4.6, the following properties can be easily discerned.

1. The horizontal and vertical spectrum are almost the same.
2. Low horizontal frequency energy is larger than that of high horizontal frequency energy, 79.75 vs 14.26. This is also true for the vertical spectrum, 40.80 vs 3.34

By applying 2-D Fourier transform to different types of images, such as optical images and SAR images, we found that for the SAR image spectrum, the ratio of the
power of high frequency bandwidth over low frequency bandwidth is much higher than that for non-SAR images. It is around 10 times higher for some cases.

### 4.7.1.3 MSE and Bit Allocation Results.

Compression distortion is measured using the Mean Squared Error (MSE). In general, it is difficult to examine the difference on a term-by-term basis. Therefore, the average squared error is used to summarize the information in the difference of pixel value. The MSE is often represented by the following:

\[
MSE = \frac{1}{MN \sum_{i=1}^{M} \sum_{j=1}^{N}} (x(i, j) - \hat{x}(i, j))^2
\]

where \(x(i, j)\) is the original pixel at the \((i, j)\)th position, and \(\hat{x}(i, j)\) is the reconstructed pixel value at \((i, j)\)th position. In the case that \(M=2N\) with \(M\) and \(N\) representing the exact number of image rows or columns, the overall mean squared error is the average of the mean squared error of each same-sized block (NXN). In Equation (4.16), \(M\) represents the number of rows, and \(N\) represents the number of columns. Other situations, such as \(M\) representing the number of columns and \(N\) the number of rows can be obtained similarly. \(MSE_1, MSE_2\) in Equation (4.16) represent the mean squared error for block1 and block2, respectively.

\[
MSE = \frac{1}{MN \sum_{i=1}^{M} \sum_{j=1}^{N}} (x(i, j) - \hat{x}(i, j))^2
\]

\[
= \frac{1}{2NN} \left( \sum_{i=1}^{N} \sum_{j=1}^{N} (x(i, j) - \hat{x}(i, j))^2 + \sum_{i=1}^{M} \sum_{j=1}^{N} (x(i, j) - \hat{x}(i, j))^2 \right)
\]

\[
= \frac{1}{2} (MSE_1 + MSE_2)
\]

Figure 4.7 shows the MSE experimental results for the SPIHT and WPEB at four compression ratios. As the bpp increases, the compression ratio decreases. Assuming the
original image is of 8 bpp, the relationship between bpp to compression ratio (cr) is as follows.

\[
    cr = \frac{\text{number of bits for the original image}}{\text{number of bits for the reconstructed image}} = \frac{8 \text{ bpp}}{\text{number of bpp}} \quad (4.17)
\]

It is clear that WPEB achieves less mean squared error (36%), as compared to the SPIHT algorithm. The distortion curve for WPEB is always lying below the curve for SPIHT, which obviously demonstrates the merit of WPEB versus SPIHT at every compression ratio.

Figure 4.7 MSE results of SPIHT and WPEB for the first test image.

The bit allocation is simple in this case as each block has the same compression ratio as the whole image. This is because the variances for the two blocks are almost the same. A detailed rate-distortion curve for each block is shown in Figure 4.8(a). For 1.0bpp, the minimum distortion is obtained at 1.0bpp for each block, as referred to in Figure 4.8 (b).
4.7.2 Case Two

In the second experimental case, the test image with the image size of 512X256 is divided into two blocks of the same size (256X256). These two blocks have a completely different variance, with the first block 2.80e3, and the second 5.84e3. In addition, the whole test image variance is 4.6e3. Increasing the contrast of the second block causes its large variance.

4.7.2.1 Variance Ratio

The second test image is transformed using biorthogonal wavelet with three level decomposition. The variance ratio is shown in Table 4.2. It shows the same properties as Table 4.1. As we set the threshold for the wavelet packet as 0.2, so LH3 and HL3 will be decomposed further.
Table 4.2 Variance Ratio for the second test image

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>LH1/LL1</td>
<td>0.128</td>
</tr>
<tr>
<td>HL1/LL1</td>
<td>0.052</td>
</tr>
<tr>
<td>HH1/LL1</td>
<td>0.018</td>
</tr>
<tr>
<td>LH2/LL2</td>
<td>0.187</td>
</tr>
<tr>
<td>HL2/LL2</td>
<td>0.141</td>
</tr>
<tr>
<td>HH2/LL2</td>
<td>0.064</td>
</tr>
<tr>
<td>LH3/LL3</td>
<td>0.242</td>
</tr>
<tr>
<td>HL3/LL3</td>
<td>0.257</td>
</tr>
<tr>
<td>HH3/LL3</td>
<td>0.133</td>
</tr>
</tbody>
</table>

Figure 4.9 1-D Spectrum for the second test image.
4.7.2.2 1-D Spectrum.

The 1-D spectrum is also shown in Figure 4.9. Compared with Figure 4.6, the second image spectrum has the same shape as the first test image, and Table 4.2 also shows the same frequency property as Table 4.1.

4.7.2.3 MSE and Bit Allocation Results.

The Rate-Distortion curve for the whole image and the two blocks are shown in Figure 4.10. It shows that the R-D curve of block2 (the lower image block) is close to the curve of the whole image, and at the same time, the variance of block2 is close to the whole image.

By using optimal searching, the compression ratios for the two blocks are not the same as the compression ratio for the whole image. The details are shown in Table 4.3.
This can be obtained by searching the minima of the rate-distortion curve at a specific compression ratio, as shown in Figure 4.11.

Table 4.3 Optimal compression ratio for two blocks

<table>
<thead>
<tr>
<th></th>
<th>overall</th>
<th>0.5 bpp</th>
<th>1.0 bpp</th>
<th>1.5 bpp</th>
<th>2.0 bpp</th>
</tr>
</thead>
<tbody>
<tr>
<td>block1</td>
<td>0.3 bpp</td>
<td>0.6 bpp</td>
<td>0.9 bpp</td>
<td>1.5 bpp</td>
<td></td>
</tr>
<tr>
<td>block2</td>
<td>0.7 bpp</td>
<td>1.4 bpp</td>
<td>2.1 bpp</td>
<td>2.5 bpp</td>
<td></td>
</tr>
</tbody>
</table>

Figure 4.11 Optimal compression searching curve.

Four curves in Figure 4.11 show the Rate-Distortion at 0.5 bpp, 1.0 bpp, 1.5 bpp, and 2.0 bpp, respectively. The horizontal axis shows the compression ratio for block1,
and the vertical axis shows the MSE for the whole image. The compression ratio (bpp2) for block2 is as follows:

For 0.5 bpp, \( \text{bpp2} = 1.0 - \text{bpp1} \), \( \text{bpp1} \) is in the range of \([0, 1.0]\).

For 1.0 bpp, \( \text{bpp2} = 2.0 - \text{bpp1} \), \( \text{bpp1} \) is in the range of \([0, 2.0]\).

For 1.5 bpp, \( \text{bpp2} = 3.0 - \text{bpp1} \), \( \text{bpp1} \) is in the range of \([0, 3.0]\).

For 2.0 bpp, \( \text{bpp2} = 4.0 - \text{bpp1} \), \( \text{bpp1} \) is in the range of \([0, 4.0]\).

The MSE for SPIHT and WPEB is shown in Figure 4.12. The MSE for WPEB is 38% lower than the SPIHT. Without minima searching (each block is compressed at the same bit rate), the MSE is only 32% lower than SPIHT. This great achievement is produced by the blocking effect and the optimal searching based on MSE.

From Table 4.3, at each compression ratio, block1 is compressed at a higher ratio, while block2 at a lower ratio. This is because the R-D curve for block1 is lower than that of block2, which means that at the same distortion, block1 is at a higher ratio than block2. By finding the minima R-D curve for the two blocks, the optimal compression performance can be achieved.

![Figure 4.12  MSE results of SPIHT and WPEB for the second test image.](image-url)
4.8 Summary

In this chapter, two main properties have been investigated. First, by checking the spectrum of the SAR image, we found that this type of image is different from an optical image in that it has much more energy at the middle and high frequencies. Based on this, it is beneficial to apply wavelet packet decomposition instead of dyadic wavelet decomposition to better utilize the characteristics of SAR images in terms of its spectrum.

Secondly, the blocking effect on the compression of SAR images is investigated. It is well known that the blocked image can not achieve better compression performance in terms of mean squared error as the entropy of any data sequence tends to decrease as the length of sequence increases only if the image statistics are uniform. In the study of any random process, it is commonly assumed that the random process is wide-sense stationary and ergodic, thus any given sequence will represent the characteristics of the random process thoroughly. Whereas in real applications, the image data statistical model sometimes may deviate from what is assumed; thus applying image compression to the blocks may be better than when it is applied to the whole. However, a problem arises in this issue is that how to divide the image into blocks efficiently so that it can benefit the compression algorithm to the greatest extent. Although this problem has not been solved thoroughly, we agree that there is a trade off between block size and image variability. For example, if the block size is close to the distance over which the image statistics vary, then blocking will have the best performance in terms of being able to allocate bits in the most efficient fashion.

Our experimental results show that a 32% decrease in MSE is obtained by dividing the second test image into two blocks when statistical property of each block varies significantly. On the contrary, it does not do any good for the first test image to decrease the MSE when the statistical property of each block does not vary so much.
Figure 4.13 Test images. (a) Case One, (b) Case Two.
Figure 4.14 Compression results for SPIHT and WPEB at 1.0 bpp for Case One.
Figure 4.15 Compression results for SPIHT and WPEB at 1.0 bpp for case two.
Chapter 5

Polarimetric SAR Data Compression

5.1 Introduction

Polarimetric data have the potential to many of the remote sensing applications. With the launch of RADARSAT-II, more polarimetric data will be available for a large variety of application areas. The data volume associated with polarimetry is so significant that it is necessary to compress these data for storage, as well as transmission. The representation of the polarimetric data includes many choices. Each of them represents a complex vector in the co-pol and cross-pol channel.

Principal Component Analysis (PCA) is widely used as a standard tool for the compression and enhancement of remotely sensed multispectral data, such as Landsat, SPOT, and aircraft multispectral scanners. The existence of correlations among spectral bands permits PCA to condense the image information from a large number of bands into a small number of components with the additional advantage of noise reduction.

In this chapter, we analyse the possibility of utilizing principle component analysis to decorrelate multichannel polarimetry data, determine the eigenvalues of the covariance matrix of the multichannel data as fundamentals for an adaptive bit allocation scheme, encode/decode each transformed component separately, and compare the MSE by separately compressing each channel image.

Although the experimental results show that PCA is not as a good candidate for the compression of polarimetric data as that of multispectral data, this conclusion has never been addressed before. Based on our experiments, further conditions of using PCA for multichannel image decorrelation are studied; even mathematical analysis can explain the
merit of using a coherency matrix in the application of target decomposition in that the features of SAR polarimetric data are more apparent in this representation than others.

5.2 Questions Associated with Polarimetric SAR Data Compression

Possible representations of the polarimetric data include the scattering matrix, Stokes matrix, covariance matrix and coherency matrix. In order to compress the polarimetric data in an efficient way, the question associated with it may have the following concerns:

* What is the appropriate data representation for compression? Is it a covariance matrix, or a scattering matrix, or even a coherency matrix?
* Given a polarimetric data set, what is the best lossy compression scheme?

5.2.1 The First Question

One commonly used data volume reduction method for polarimetric SAR data is described in [18] by Pascale Dubois and Lynne Norikane of Jet Propulsion Laboratory in the California Institute of Technology. The general idea of their method is that 4-look averaging of the Stokes matrix instead of the scattering matrix introduces less information error. Lots of polarimetric data provided by JPL is compressed using this method, and the data format available is based on the Stokes matrix.

It is easy to obtain each element of the Stokes matrix from 9 elements of SaaSbb*, aa and bb represent hh, vv, and hv. As the covariance matrix also consists of these 9 elements, it is considered as an efficient representation for radar polarimetry. In addition, because of the availability of most polarimetry data formats (such as the CEOS format), the covariance matrix is probably the best candidate for polarimetric data compression.
5.2.2 The Second Question

For any polarimetric data, a general investigation based on Principle Component Analysis (PCA) [19] to a feature set, such as \{HHHH^*, HVHV^*, VVVV^*, \text{Re}(HHVV^*), \text{Im}(HHVV^*), \text{Re}(HHHV^*), \text{Im}(HHHV^*), \text{Re}(HVVV^*), \text{Im}(HVVV^*)\} showed the contribution difference of each feature to the overall scene variance. Larger contribution means the correspondent component contains more important information in this feature set, and vice versa.

In this chapter, we choose the feature scene that contains two data sets, the first data set includes three intensity scenes of HH, VV, and HV, and the second data set consists of three intensity images and two phase-difference images of HH-VV and HV-VV.

Phase HV-VV can be easily discerned from the differences of phases HH-VV and HH-HV, which are available from the covariance matrix. Actually, the absolute phase information is lost in the covariance matrix, and only the phase difference is preserved.

Since the second data set contains all the polarimetry information of the target, and it becomes a feature set we are interested in for compression.

The reason to set up two data sets lies in the fact that we would like to find out how to apply PCA to polarimetric data and what data set is the best for PCA.

(1) Is it appropriate to apply PCA only to three intensity images, or to five images together?

(2) Is there any difference between the effect of PCA to three intensity images than to that of five images?

Since the intensity image and phase difference image have completely different statistical properties and scene content, the total amount of information is scattered in those images, and redundancies exists in different forms inside both types of images. The
following sections investigate the decorrelation ability of PCA to both intensity and phase difference images.

5.3 **Mathematic Background for PCA**

The principle component analysis technique is related to the Karhunen-Loeve (KL) transform. It is known that KL transform is the most optimal transform for decorrelating data in image compression based on the eigenvalue and eigenvector of the covariance matrix of the image. Images after the KL transform will have the most energy compaction.

Suppose we have a sample space $S$, which includes all of the 2-d images. The $i$th experimental outcome corresponds to a sample in $S$. If we define the experiment as a random process, then the $i$th experiment is a random variable $X_i$.

Assume we have $N$ random variables, $\lambda_i$ is the $i$th eigenvalue of the covariance matrix (N by N) of the random variables $X=[X_1, X_2, \ldots X_N]$ listed in columns, and $e_i$ is the $i$th eigenvector corresponding to $\lambda_i$; then we get the transform matrix (N by N) consisting of all the eigenvectors. $e_{ij}$ is the $j$th component of the $i$th eigenvector.

$$A = \begin{bmatrix}
e_{11} & e_{12} & \ldots & e_{1N} \\
e_{21} & e_{22} & \ldots & e_{2N} \\
\vdots & \vdots & \ddots & \vdots \\
e_{N1} & e_{N2} & \ldots & e_{NN}
\end{bmatrix}$$

(5.1)

PCA transform is done by the following:

$$Y = AX$$

(5.2)

The inverse transform is as follows

$$X = A^{-1}Y$$

(5.3)
The important property associated with this type of transform is that the covariance of Y is diagonal:

$$C_Y = \begin{bmatrix}
\lambda_1 \\
\lambda_2 \\
\vdots \\
\lambda_N
\end{bmatrix}$$  \hspace{1cm} (5.4)

Other properties include the following [46]:

1. It completely decorrelates the signal in the transform domain.
2. It minimizes the mean-squared-error or maximizes the signal-noise-ratio in data compression.
3. It contains the most variance (energy) in the fewest numbers of transform coefficients.

### 5.4 Experiment Method

For the first data set, we are going to compress three intensity images. The first step in compression is to apply PCA to the three intensity images.

The decorrelation method associated with PCA is closely related to the KL transform in the sense that it is applied to multiple input images. After PCA, the transformed outcomes will be a totally uncorrelated data set without any information redundancy among them, and thus can be compressed efficiently with maximum transform coding gain. The only disadvantage of using the KL transform lies in its heavy computational requirement without any fast algorithm for the implementation.

For each group, the PCA can be done in a sequence as follows:
* Measuring the covariance matrix of the input of the three intensity images
* Calculating the eigenvalue and eigenvector of the covariance matrix
* Transforming the three input images to uncorrelated components.

The covariance of the transformed results have only non-zero values at the diagonal positions that correspond to the eigenvalues of the covariance matrix of the input images, as referred to in the equations in Section 5.3.

After PCA, we obtained the transformed images and their variances. If the variances are different, it makes sense to assign a different number of bits to different transformed images. We assume the variance of the principle components corresponds to the amount of information contained in each component. Thus, the principle component with a higher variance is assigned more bits than that with a smaller variance.

The algorithm [12] for bit allocation works as follows:
(1) Set $R_n=0$ for $n=1,2,3$. $R$ is the total number of bits available for distribution.
(2) Compute variance $\sigma_n$ for each principle component.
(3) Sort the variance,
(4) If $\sigma_k$ is the maximum, increase $R_k$ by 1, divide $\sigma_k$ by 2.
(5) Decrement $R$ by 1.
(6) If $R=0$, stop; otherwise go to step 3.

Since the bit allocation for each PCA transformed component is complete, then we can use this information for further coding. There are many candidates for the coding scheme. We would like to compare two coding methods as shown below.
* Uniform quantizer.
* Transform coding using discrete wavelet transform and zero tree coding.

Image reconstruction can be easily achieved by obtaining the inverse of the eigenvector of the covariance matrix, and applying the inverse transform to the principal
components. The performance of the compression is evaluated by Signal-to-Noise Ratio (SNR) and Mean Squared Error (MSE).

5.5 Experimental Result for Three Intensity Images.

5.5.1 Covariance Matrix

For three intensity images shown in Figure 2.4, the symmetric covariance matrix $C$ is listed below.

\[
\begin{array}{ccc}
     & HH & HV & VV \\
HH & 300.5 & 173.9 & 229.6 \\
HV & 200.0 & 114.5 & \\
VV & & 210.1 \\
\end{array}
\]

For two random variables $X_1, X_2$, the covariance is defined as follows:

\[
\text{Cov}(X_1, X_2) = E \left[ (X_1 - u_1) * (X_2 - u_2) \right] \tag{5.5}
\]

where $u_1$ is the mean or expected value of random variable $X_1$, and $u_2$ is the mean value of random variable $X_2$.

In the covariance matrix, $C(1,1)$ is the variance of $X_1$ itself, and $C(1,2)$ is the covariance of $X_1$ and $X_2$, and so forth. It is easy to discern that $C(1,1)$ is just the variance of the HH image, $C(2,2)$ the variance of the second HV image, and $C(3,3)$ the variance of the third VV image. HH image has the largest variance or energy, VV has the second, and HV has the least variance and energy. Since all the input images are subtracted by their means, the variance has the same amount as the energy.

5.5.2 Correlation Matrix

In order to check the correlation among those three images, we can use the correlation coefficient matrix for clear observation. The correlation coefficient of two random variables is defined as the following:
Cor(X₁,X₂) = \frac{\text{Cov}(X₁,X₂)}{\sqrt{\text{D}(X₁) \cdot \text{D}(X₂)}} \tag{5.6}

\text{D}(X₁) \text{ is the variance of random variable } X₁. \text{ D}(X₂) \text{ is the variance of random variable } X₂.

The symmetric correlation coefficient matrix for three original images is as follows:

\[
\begin{array}{ccc}
\text{HH} & \text{HV} & \text{VV} \\
\hline
\text{HH} & 1.00 & 0.713 & 0.91 \\
\text{HV} & 1.00 & 0.55 & \\
\text{VV} & 1.00 & &
\end{array}
\]

The correlation coefficient between HH and VV (0.91) are higher than any other two pairs. This means that the redundancy between HH and VV are the highest among that of any two pairs.

The eigenvectors of the covariance matrix are listed below:

\[
\begin{array}{ccc}
0.19 & -0.69 & 0.70 \\
-0.86 & 0.23 & 0.46 \\
0.48 & 0.68 & 0.55 \\
\end{array}
\]

The inverse of the eigenvector matrix is the transpose of the eigenvector matrix, as follows.

\[
\begin{array}{ccc}
0.19 & -0.86 & 0.48 \\
-0.69 & 0.23 & 0.68 \\
0.70 & 0.46 & 0.55 \\
\end{array}
\]

After PCA, the covariance matrix for three principle components is as follows:

\[
\begin{array}{ccc}
97.1 & 0.0 & 0.0 \\
0.0 & 16.1 & 0.0 \\
0.0 & 0.0 & 597.4
\end{array}
\]
The three transformed components are shown in Figure 5.1. The correlation matrix for three components is as follows:

\[
\begin{pmatrix}
1.0 & 0.0 & 0.0 \\
0.0 & 1.0 & 0.0 \\
0.0 & 0.0 & 1.0 \\
\end{pmatrix}
\]

This demonstrates that there is no correlation between each component. It shows that the principle component transform can generate uncorrelated coefficients. In a geometrical sense, it rotates the highly correlated features in N-dimensions to a more favourable orientation in the feature space, orthogonal to each other. This process is viewed as information compression into a small number of components from a large number of features by discarding redundant information of the components with small variance.

5.5.3 Quantizer

The Quantizer is the only step where loss is caused by the compression process. As we know, the transformed results are real valued coefficients; after quantization the three transformed components are all integer coefficients. The optimal quantizer for a given distribution has decision levels such that the probability of the occurrence is equal.

5.5.3.1 Optimal Quantizer for Gaussian Distribution

Given the compression ratio of 1 bpp, the bit allocation algorithm assigns 3 bpp to the third principle component that is of the largest variance (597.4), and zero bpp to both first and second components.

Suppose that by applying optimal quantizer to Gaussian distributed random variables, we get 3 reconstructed images with SNR and MSE, listed below in Table 5.1.
Table 5.1. SNR and MSE for three compressed intensity images.

<table>
<thead>
<tr>
<th></th>
<th>HH</th>
<th>HV</th>
<th>VV</th>
</tr>
</thead>
<tbody>
<tr>
<td>SNR</td>
<td>22.4</td>
<td>10.3</td>
<td>22.4</td>
</tr>
<tr>
<td>MSE</td>
<td>41.0</td>
<td>100.2</td>
<td>45.0</td>
</tr>
</tbody>
</table>

Root Mean Square (RMS) is another criteria for performance measurement when the original image and its reconstructed results are compared. It is defined as follows:

\[
\text{RMS} = \sqrt{\frac{\sum_{j=1}^{N} X_j^2}{N}}
\]  

(5.7)

Table 5.2. RMS for reconstructed three intensity images (Gaussian distribution).

<table>
<thead>
<tr>
<th>RMS</th>
<th>HH</th>
<th>HV</th>
<th>VV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Original</td>
<td>84.8</td>
<td>32.9</td>
<td>88.1</td>
</tr>
<tr>
<td>Reconstructed</td>
<td>83.0(97.9%)</td>
<td>30.5(92.8%)</td>
<td>86.8(98.5%)</td>
</tr>
</tbody>
</table>

5.5.3.2 Optimal Quantizer for Laplacian Distribution

Using optimal quantizer for Laplacian distributed random variables [22] [25], at 1bpp compression ratio, we get three reconstructed images (Figure 5.2) with SNR and RMS listed in Table 5.3.

Table 5.3. SNR, MSE and RMS for three intensity images (Laplacian distribution).

<table>
<thead>
<tr>
<th></th>
<th>HH</th>
<th>HV</th>
<th>VV</th>
</tr>
</thead>
<tbody>
<tr>
<td>SNR</td>
<td>22.7</td>
<td>10.4</td>
<td>22.5</td>
</tr>
<tr>
<td>MSE</td>
<td>38.3</td>
<td>99.8</td>
<td>43.3</td>
</tr>
<tr>
<td>RMS</td>
<td>83.0(97.9%)</td>
<td>30.7(93.2%)</td>
<td>86.7(98.4%)</td>
</tr>
</tbody>
</table>
Figure 5.1. Three principal components with eigenvalues 97, 16 and 597.
Figure 5.2. Three reconstructed images using an average of 1bpp (R=3).
Figure 5.3. Three reconstructed images using an average of 2bpp (R=6).
Laplacian uniform quantizer is better than Gaussian uniform quantizer in terms of compression performance (SNR). This means the each principal component is more likely to have Laplacian distribution than Gaussian distribution.

Another example of average 2bpp is studied as well. The bit allocation scheme results in assigning 2bpp to the first principal component, 1bpp to the second principal component, and 3bpp to the third principal component. The reconstructed images are shown in Figure 5.3. The overall compression results are summarized in Figure 5.4.

![SNR for three reconstructed images](image)

Figure 5.4. compression results for three intensity images at 1bpp and 2bpp.

At 1 bpp, the bits assigned to 3 principal components are: 0, 0, 3.
At 2 bpp, the bits assigned to 3 principal components are: 2, 1, 3.

### 5.5.4 PCA Property: Orthogonal and Unitary Transform

The inverse of the eigenvector matrix is the transpose of the eigenvector matrix. That is the reason why the transform matrix is orthogonal and unitary. In addition, the
energy of the images and their transformed results are always the same. However, the
energy distribution for three images and their transformed parts are completely different
because of the energy compaction capability of the transform. The efficiency of a
transform depends on how much energy compaction is provided by the transform. One
way of measuring the amount of energy compaction by a particular uniform transform is
to take the ratio of the arithmetic mean of the variances of the transform coefficients to
their geometric means. This ratio is also referred to as the transform coding gain $G$:

$$G = \frac{1}{N} \sum_{i=1}^{N} \sigma_i^2 \left( \prod_{i=1}^{N} \sigma_i^2 \right)^{1/N}$$ (5.8)

5.5.5 Wavelet Transform

Since wavelet transform coding is an efficient coding scheme for most image
compression, experiments based on DWT coding for three intensity polarimetric images
are conducted here for further comparison in the following section. In this case, each
image is compressed with 1bpp, respectively. The compression results using the SPIHT
compression algorithm is listed in Table 5.4.

<table>
<thead>
<tr>
<th></th>
<th>HH</th>
<th>HV</th>
<th>VV</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>SNR</strong></td>
<td>23.1</td>
<td>20.7</td>
<td>24.2</td>
</tr>
<tr>
<td><strong>MSE</strong></td>
<td>34.8</td>
<td>9.9</td>
<td>29.9</td>
</tr>
</tbody>
</table>
5.5.6 Comparison: PCA vs. Wavelet Transform.

In our experiment, wavelet transform coding for individual intensity images is better than the PCA method for three intensity images in terms of SNR.

For compression of the second image (HV), wavelet transform coding is much better than the PCA method. This is because at 1bpp, the PCA method allocates first and second components with zero bits, and HV image information is mostly transformed into these two components; therefore, most of the HV image information is lost and cannot be reconstructed very well.

A more detailed investigation is conducted by checking the transform gain of the wavelet transform to three principal components, and the possibility of improving the previous PCA method.

The transform gain for three principal components and some related data is listed in Table 5.5.

<table>
<thead>
<tr>
<th></th>
<th>#1 pc</th>
<th>#2pc</th>
<th>#3pc</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variance</td>
<td>97.1</td>
<td>16.1</td>
<td>597.4</td>
</tr>
<tr>
<td>Mean</td>
<td>32.1</td>
<td>8.8</td>
<td>119.3</td>
</tr>
<tr>
<td>Wavelet transform gain</td>
<td>3.67</td>
<td>1.37</td>
<td>5.15</td>
</tr>
</tbody>
</table>

The table above shows the benefit of applying the wavelet transform to three components according to the transform gain.

Using the principle component, which corresponds to the maximum eigenvalue, the best achievable compression results without applying a quantizer to principal components, are listed in Table 5.6.
Table 5.6. Compression results (without quantizer)

<table>
<thead>
<tr>
<th></th>
<th>HH</th>
<th>HV</th>
<th>VV</th>
</tr>
</thead>
<tbody>
<tr>
<td>SNR</td>
<td>28.0</td>
<td>11.8</td>
<td>24.1</td>
</tr>
<tr>
<td>MSE</td>
<td>11.4</td>
<td>72.2</td>
<td>29.9</td>
</tr>
</tbody>
</table>

From the table above, SNR for HH and VV intensity images are almost twice the value of HV images. This proves the fact that the third component, which corresponds to the maximum eigenvalue, contains most of the information included in HH and VV intensity images.

This is the maximum SNR the principal component algorithm can achieve at 1bpp; because the quantizer is not used, the quantization error is minimized to zero, and therefore its contribution to overall distortion is also zeroed. Even at this extreme condition, the algorithm using PCA is not the best for compressing the HV intensity image. No matter how hard you tried, for the HV intensity image under the compression rate of 1bpp, it is impossible to get a better SNR result using PCA than using wavelet transform directly.

If it is necessary to increase the SNR for the HV intensity image, we should at least include other principle components. The most straightforward way includes the following:

* Modify the bit allocation scheme to assign bits to at least two components.
* Modify the uniform quantizer to a nonuniform one to achieve minimum distortion.

These two methods are not very practical because they are closely related to rate distortion theory, which needs a numerical method to obtain a better result. In this way, the computational complexity is usually high and not for real time applications.

5.5.7 One Possible Solution

The easiest ways to improve compression performance for three intensity images is to group only HH and VV intensity images together, and make HV a separate image for
compression. In this way, the SNR for HH is 29.2, for the VV image it is 27.9 using only one component, which is much higher than the situations when HV is included in the group. The transform gain for a two-image datum set case is 2.5, a little bit higher than that of a three-intensity image case of 2.4.

5.5.8 Phase-Difference Image

Adding two phase-difference images to three intensity images and applying PCA to these five images, we get the following correlation coefficient matrix:

<table>
<thead>
<tr>
<th></th>
<th>HH</th>
<th>HV</th>
<th>VV</th>
<th>phase (HH-VV)</th>
<th>phase (HV-VV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HH</td>
<td>1</td>
<td>0.71</td>
<td>0.91</td>
<td>0.18</td>
<td>0.26</td>
</tr>
<tr>
<td>HV</td>
<td></td>
<td>1</td>
<td>0.56</td>
<td>-0.18</td>
<td>0.29</td>
</tr>
<tr>
<td>VV</td>
<td></td>
<td></td>
<td>1</td>
<td>0.28</td>
<td>0.25</td>
</tr>
<tr>
<td>Phase(HH-VV)</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>-0.16</td>
</tr>
<tr>
<td>Phase(HV-VV)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
</tbody>
</table>

Properties associated with the upper correlation matrix include the following:

(1) Two phase-difference images are not as highly correlated with a coefficient of -0.16, as compared to the correlation among three intensity images. This may be caused by the speckle-like noise of those two phase-difference images referred to Figure 2.5. Since those two phase-difference images are not correlated very well, there is little benefit in applying PCA to two phase-difference images.

(2) The compression of these two phase-difference images requires more bits to preserve the information contained in them. PCA is not an appropriate method because of
the high degradation it causes at the image reconstruction. Each phase difference image should be compressed separately at an appropriate ratio to maintain the data quality for further analysis and processing.

(3) For three intensity images, it is generally acknowledged that the correlation between HH and VV are much higher than that between HV and VV, or between HV and HH. In addition, it is generally assumed that the speckles in HH and VV images are correlated, while cross-polarized images HV and VH are uncorrelated with either HH or VV images [20]. Based on this knowledge, we propose the compression algorithm to combine the HH and VV image together, apply PCA only to these two images, and compress the HV image separately.

In summary, for polarimetric data compression, a very high compression ratio is not achievable because of the large amount of information and speckle phenomenon associated with them. In addition, since two phase-difference images are not highly correlated with three intensity images, it is better not to combine the intensity and phase-difference images together in the same data set for compression. In many cases, the phase-difference scene shows the property of almost pure noise; therefore, its information content is completely different from that of intensity images, and the compression algorithm should be different as well, for each type of image.

5.6 Frequency Response for Three Eigenvectors

It is interesting that three eigenvetors in our experiment show the properties of lowpass, bandpass and high pass effect, referred to in Figure 5.5.
From Figure 5.5, it is discerned that the third eigenvector has the property of a low pass filter, so that the corresponding component has the largest energy and variance in our experiment. This can also be discerned from the fact that the normal signal spectrum shows that most of the energy of the signal resides in low frequency part of the spectrum. The second eigenvector corresponds to the second component, and it shows the filtering property of a band pass. The first eigenvector has the property of a high pass filter.

This figure tells us that applying the principle component transform to multiple images is equal to applying filters in the multiple image direction (Figure 5.6).
5.7 Where Do Speckles Go After PCA?

It is generally acknowledged that co-polarized (HH, VV) terms represent backscattered signals mainly contributed by single bounce scatters, while the cross-polarized terms (HV, VH) are mainly due to multiple bounce behavior. The similarity in scattering mechanisms makes the speckles in HH and VV highly correlated, and the co-polarized term and cross-polarized term statistically uncorrelated.

Speckles in the intensity SAR images have the characteristic of multiplicative noise, in the sense that speckle noise is signal dependent. After logarithmic transformation, the speckle noise becomes additive and signal independent, and its probability density distribution is approximately Gaussian [23] [26].

Let us look at the following examples with two random variables $z_1$ and $z_2$, which are corrupted by speckles $e_1$ and $e_2$.

$$z_1 = x_1 + e_1$$  \hspace{1cm} (5.9)

$$z_2 = x_2 + e_2$$  \hspace{1cm} (5.10)

Because $e_1$ and $x_1$ are statistically independent, so we have the following:

$$D(z_1) = D(x_1) + D(e_1)$$  \hspace{1cm} (5.11)

For $e_2$ and $x_2$, we also have the following:
\[ D(z2) = D(x2) + D(e2) \]  
(5.12)

where \(D\) represents variance.

We assume \(x1, x2, e1, \) and \(e2\) are all zero mean variables, with standard deviation \(\sigma_{x1}, \sigma_{x2}, \sigma_{e1}, \) and \(\sigma_{e2}.\) From the matrix theory [24], we know that for any Hermitian matrix \(R,\) there exists a unitary matrix \(\Phi,\) such that the following holds:

\[ \Phi^* R \Phi = \Lambda \]  
(5.13)

where \(\Lambda\) is a diagonal matrix containing the eigenvalues of \(R.\) An alternate form of the above equation is as follows:

\[ R \Phi = \Phi \Lambda \]  
(5.14)

which is the following set of eigenvalue equations:

\[ R \phi_k = \lambda_k \phi_k \quad k = 1, \ldots, N \]  
(5.15)

where \(\{ \lambda_k \}\) and \(\{ \phi_k \}\) are the eigenvalues and eigenvectors of \(R,\) respectively.

For Hermitian matrices, the eigenvectors corresponding to distinct eigenvalues are orthogonal. Since the covariance matrix in our case is also a Hermitian matrix, there exists a unitary and orthogonal matrix, which consists of the eigenvalue of this covariance matrix.

In real applications, the transform matrix is highly dependent on the input data, and cannot be obtained before the input data are known. Since the transform itself aims to decorrelate a set of correlated data to its uncorrelated components (such as the principal component), we can use the unitary and orthogonal transform matrix \(C,\) as shown below for investigation.

\[ C = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \]  
(5.16)
Let $Z=[z_1, z_2]'$, $Y=[y_1, y_2]'$.

Transform $Z$ to $Y$: $Y=CZ$

Let $kk=1/\sqrt{2}$, then $kk^2=0.5$.

$y_1=kk(z_1+z_2)$; $y_2=kk(z_1-z_2)$; \hspace{1cm} (5.17)

Then we can obtain the variance for $y_1$ as follows:

$D(y_1)=D(kk(z_1+z_2))$

$=0.5*D(x_1+e_1+x_2+e_2)$

Because $x_1$ and $e_1$ are independent, and $x_2$ and $e_2$ are independent, we have the following:

$D(y_1)=0.5*(D(x_1+x_2)+D(e_1+e_2))$

$=0.5*(\sigma_{x_1}^2+\sigma_{x_2}^2+2\rho_s\sigma_{x_1}\sigma_{x_2})+0.5*(\sigma_{e_1}^2+\sigma_{e_2}^2+2\rho_e\sigma_{e_1}\sigma_{e_2})$ \hspace{1cm} (5.18)

$=\text{power of signal + power of speckle noise}$

$\rho_s$ is the correlation coefficients of $x_1$ and $x_2$.

$\rho_e$ is the correlation coefficients of $e_1$ and $e_2$.

$\sigma_{x_1}$ is the variance of signal $x_1$.

$\sigma_{x_2}$ is the variance of signal $x_2$.

$\sigma_{e_1}$ is the variance of noise $e_1$.

$\sigma_{e_2}$ is the variance of noise $e_2$.

The signal-noise-ratio for $x_1$ and $x_2$ are defined as the power of the signal to the power of speckle noise, giving the following:

$sx_1=\sigma_{x_1}^2/\sigma_{e_1}^2$, \hspace{1cm} sx_2=\sigma_{x_2}^2/\sigma_{e_2}^2$ \hspace{1cm} (5.19)
The signal-noise-ratio for $y_1$ and $y_2$ are obtained as follows:

$$s_{y1} = \frac{\left(\sigma_{x_1^2} + \sigma_{x_2^2} + 2\rho_{x_1} \sigma_{x_1} \sigma_{x_2}\right)}{\left(\sigma_{e_1^2} + \sigma_{e_2^2} + 2\rho_{e_1} \sigma_{e_1} \sigma_{e_2}\right)}$$  \hspace{1cm} (5.20)

$$s_{y2} = \frac{\left(\sigma_{x_1^2} + \sigma_{x_2^2} - 2\rho_{x_1} \sigma_{x_1} \sigma_{x_2}\right)}{\left(\sigma_{e_1^2} + \sigma_{e_2^2} - 2\rho_{e_1} \sigma_{e_1} \sigma_{e_2}\right)}$$  \hspace{1cm} (5.21)

Assuming $s_{x1} = s_{x2} = k$, and letting $k$ change from 1 to 100 with step 1, we get the SNR results for the two components and original signal, as shown in Figure 5.7.

Figure 5.7 Comparison of SNR for two components and original signal.

(a) correlation for signal is 0.15, for noise is 0.9;
(b) correlation for signal is 0.3, for noise is 0.15;
(c) correlation for signal is 0.95, for noise is 0.8;
(d) correlation for signal is 0.95, for noise is 0.15.
Since component 1 (y1) is obtained by adding two input signals, and thus making the feature more apparent, it contains more important information than component 2. Therefore, using SNR1 to represent the signal-noise-ratio of the first component y1 and SNR2 to represent the signal-noise-ratio of the second component y2, what we expect PCA to do is to make SNR1 much greater than SNR2. However, from Figure 5.7, we discover that it is not always true.

(1) In case (a), SNR2 is greater than SNR1, which means when noise is highly correlated, it becomes the dominant information, and noise reduction is absolutely impossible.

(2) In case (b), SNR1 is close to SNR2. Since both the signal and the noise are not highly correlated, after PCA, both noise and signal will resides in two principal components and noise reduction is impossible.

(3) In case (c), (d), SNR1 is much greater than SNR2. At this case, input signals are highly correlated (\(\rho_x = 0.95\)), and PCA achieves better results for the purpose of data compression and denoising, than in cases (a) and (b).

(4) In case (d), since the noise correlation coefficient is small, the PCA denoising effect is more apparent than in case (c).

In summary, the experimental results illustrate the conditions for applying PCA for noise reduction and data compression. The condition is that input signals should be highly correlated and noises in each signal should not. This condition also explains that PCA to HH and VV performs better than PCA to HH, HV and VV.

It is acknowledged that for polarimetric intensity and phase-difference images, PCA makes the speckle noise reside in every principal component; therefore, it is not an efficient method for data volume reduction, as well as speckle reduction.
Chapter 6

Conclusions and Future Work

The objective of this thesis is to investigate the application of Wavelet Transform and Principal Component Analysis to SAR polarimetric data compression. SAR image compression is a research topic that has been studied for many years from DCT based to wavelet based compression schemes. Recent studies show that wavelet based compression algorithm works better than others.

In this paper, we investigated SAR image properties, such as speckles and rich middle and high frequency spectrums. We proposed the compression algorithm based utilizing wavelet packets and subdividing the image into blocks.

The wavelet packets take advantage of the extensive medium and high frequency information in SAR images. The block aspect involves dividing the image into blocks so that wavelet decomposition, quantization, and coding can be applied to a smaller section of the image rather than to the whole image at once. In this way, the different statistics of each block can be recognized (especially the activity level), and an optimal bit allocation scheme can be applied to improve the coding efficiency of the whole image. Speckle reduction can also be applied simultaneously during the encoding/decoding process.

Experiments on single channel images with two blocks demonstrate that the ratio of the optimal number of bits to each image block is nearly proportional to the variance ratio of each image block.

Further study of single channel SAR image compression may include about the following:

(1) What is the best block size to be used? This is a critical issue in block coding since small block size leads to serious block artifacts.
(2) What is the range of compression ratio for which our compression method works best? At present, the SAR image has been compressed at the ratio of 10, which is at the medium level. High compression ratio causes large information loss and is not applicable for SAR image applications. Our aim is to ensure our compression method works best at the middle compression range such as from 0.8 bpp to 1.2 bpp.

Compared to single-channel SAR data, polarimetric data have HH, VV and HV channels of radar intensity images, plus the phase-difference among them. We examined the information redundancy of multiple data using the PCA method. The experimental results show that PCA is not appropriate for polarimetric data denoising as well as compression; although, it has been successfully used for multispectral data. However, this result was never stressed before. At present, it seems that polarimetric images have to be compressed separately.

Further study for polarimetric data compression can be focused on the following:

(1) Design a statistical measure that will give a near optimal bit allocation scheme when there are a large number of blocks available.

(2) Find out the best block size given a type of image with certain typical scene contents, such as forests, agricultures, urban areas, etc.

(3) Find the compression results by applying PCA to blocks of polarimetric images.

(4) Find out the compression results by applying wavelet compression to PCA coefficients.

(5) Using different compression schemes to intensity images and phase-difference images, as those two types of images have different statistical properties.

Finally, new compression performance evaluation criteria, such as phase accuracy, or terrain classification accuracy, will be more appropriate than SNR or MSE measurement.
Bibliography


Appendix

Compression of RADARSAT Data with Block Adaptive Wavelets
Ian Cumming and Jing Wang
Department of Electrical and Computer Engineering
The University of British Columbia
2356 Main Mall, Vancouver, BC, Canada V6T1Z4
ianc@ece.ubc.ca jingwa@ece.ubc.ca
Tel: (604) 822-4623 Fax: (604) 822-5949

Abstract: This paper proposes a new algorithm referred to as the Wavelet Packet-based
Embedded Block coding (WPEB) scheme for SAR data compression. This algorithm
combines the following properties: (1) wavelet packet decomposition is adopted to
exploit middle and high frequency components associated with SAR data; (2) block
coding is utilized to improve DWT coefficient coding efficiency by adaptively allocating
more bits to regions of importance with higher information content (e.g. more contrast,
edges); (3) speckle reduction is built into the bit allocation scheme by using wavelet
transform denoising. Examples are given using RADARSAT data, which show that the
compression performance is better than conventional wavelet methods and visual image
interpretation is acceptable at 1 bpp.

1. Introduction

SAR image data can provide unique information about the surface of the Earth [1].
The volume of data associated with Earth information is so huge that the compression of
SAR data becomes crucial to transmission and archiving.

At present, most image compression algorithms are designed for standard test
images, usually optical images [2, 4, 5]. The most popular algorithms for still image
compression are those included in the JPEG/JPEG2000 standard. These algorithms are
designed to compress traditional images, and do not lead to ideal compression results for
SAR images. This is because these compression schemes are not designed to account for
SAR data characteristics, such as high dynamic range, the speckle phenomena, and the
presence of significant high frequency components arising from terrain texture and edges.

WPEB belongs to the class of lossy transform compression algorithms. Our coding
scheme includes the wavelet packet transform [3], embedded block coding, speckle
reduction, and optimal bit allocation. We consider the SAR data spectrum in wavelet
analysis and bit allocation in order to achieve better compression performance and make
this algorithm more flexible when compressing different types of SAR data (ocean, snow,
ice, city, forest, agriculture, mountains, etc), and satisfying different compression
evaluation criteria. A block diagram for WPEB encoding and decoding is shown in
Figure 1. At the encoder, the discrete wavelet packet transform is applied to the source
image after it is divided into blocks. The transform coefficients are then quantized and
entropy coded to form the output sequence. The decoder is the reverse of the encoder,
consisting of entropy decoding, inverse quantization, and an inverse wavelet packet
transform.

1This paper was submitted to Data Compression Conference on Nov 15th, 2002.
2. Wavelet Packet Decomposition

SAR data usually have significant middle and high frequency components, as in regions with textures and edges; this makes the wavelet packet transform a better choice than the standard discrete wavelet transform (DWT) for SAR data compression. The typical spectrum for optical and SAR images is shown in Figure 2, in which we can find out that the spectrum of SAR image is above the spectrum of the optical image.

Wavelet packet decomposition differs from the standard wavelet transform by allowing the decomposition of the upper frequency bands as well as the lower ones. By this method, wavelet packets can be used to achieve a more accurate representation and compression of the medium and high frequency information in SAR images.

In previous work of SAR image compression [6], texture analysis is based on sub-band energy at different decomposition levels. The total energy $E_{SB}$ of each sub-band is defined as:

$$E_{sb} = \frac{1}{MN} \sum_{m=1}^{M} \sum_{n=1}^{N} |x(m,n)|^2$$

(1)

where $x(m,n)$ is the wavelet coefficient set at a given decomposition level. Because the low-low sub-band normally has the maximum energy, we call its energy $E_{ref}$ and compare the energy of other sub-bands with it. A constant $C_{sb} < 1$ is used as the criterion for
applying another level of wavelet packet decomposition. When the energy \( E_{SB} \) of a given sub-band is greater than \( C_{SB} E_{ref} \), wavelet packets are applied to further decompose the higher frequencies.

![1-D spectrum](image)

Figure 2. Typical spectrums for SAR images and optical images

For our algorithm, a biorthogonal wavelet [3] is selected because of its linear phase property, and because edge symmetry can be obtained using symmetrical extension. Since all test images are normalized to a zero mean before the DWT is applied, the energy measure \( E_{SB} \) is simply the sub-band variance. At each decomposition stage, the wavelet packet analysis proceeds as follows:

1. Initialize \( C_{HL}, C_{LH}, C_{HH} \), and set the maximum decomposition level.
2. Apply a DWT, to obtain the four sub-band coefficients, LL, LH, HL and HH.
3. Calculate the variance for each sub-band, \( E_{SB} \).
4. If \( E_{LH} > C_{LH} E_{ref} \), then apply a DWT to LH.
5. If \( E_{HL} > C_{HL} E_{ref} \), then apply a DWT to HL.
6. If \( E_{HH} > C_{HH} E_{ref} \), then apply a DWT to HH.
7. Apply a DWT to LL.
8. If the decomposition level is reached, stop; otherwise go to step (3).

3 Embedded Block Coding

It is known that block transform coders enjoy success because of their low complexity and their effective performance. The most popular block transform coder is JPEG, which utilizes the 8x8 Discrete Cosine Transform (DCT).
The general idea behind block coding is that the image can be divided into blocks, so that wavelet decomposition, quantization and coding can be applied to smaller sections of the image, rather than to the whole image at once. In this way, different number of bits can be allocated to each block so that the overall image SNR is maximized.

Generally, large image blocks lead to smaller mean squared error in the sense that the correlation within larger block samples can be exploited more fully in compression. This is true for optical images with correlation coefficients as high as 0.95, but not so true for SAR images where the correlation coefficients are usually below 0.7.

The advantages of using block coding include the following:

1. Different statistics of each block can be recognized. The statistics, such as dynamic data range and entropy, affect coding performance in the sense that the data range is associated with quantization steps, and entropy is associated with the maximum compression ratio for a given image block.

2. The flexibility of assigning different number of bits to each block is possible, within the constraint of a specified total number of bits per image.

On the other hand, the disadvantages of blocking include discontinuities and artifacts across the block boundaries, especially at very low bit rates. The choice of block size is a compromise between (1) obtaining good compression within a block and avoiding block artifacts (favors larger block sizes), and (2) adaptation to block statistics and computing efficiency (favors smaller block sizes). In practice, we found that a block size between 128 and 256 to be appropriate for SAR images.

Generally, we apply an embedded blocking algorithm by designing a bit allocation scheme to adaptively assign different numbers of bits to image regions based on importance. This is a practical problem when dealing with images containing many types of scene features. For example, if the SAR image is used for sea-ice classification, an ideal compression scheme preserves texture regions that represent different types of ice (first year ice, multi year ice, etc.). Some regions of an image contain features not of interest, such as land areas in sea ice images, and thus we can assign fewer bits to these regions.

4 Bit Allocation Scheme With Speckle Reduction

In this subsection, we explain the bit allocation problem mathematically. Since the image is composed of a collection of coded blocks $B_i$, with embedded bit streams that may be truncated at a rate $R_i$, the corresponding contribution to the distortion in the reconstructed image is denoted as $D_i^{n_i}$ for each truncation point $n_i$. The relevant distortion matrix is additive:

$$D = \sum_i D_i^{n_i}$$ (2)

where $D$ is the total distortion for the reconstructed image, and $n_i$ is the truncation level giving $B_i$ bits for block $i$. From the rate distortion optimization point of view, the problem is to find the optimal selection of $n_i$, so as to minimize the overall distortion $D$, subject to a constraint, $R$, where $R$ is the allowed bit rate for the whole image.
If the image is compressed as a single block, the overall distortion is equal to the block distortion. For a specific encoding algorithm, the rate-distortion curve is monotonic with only one parameter: the compression ratio or the truncation point \( n_t \), as shown in Figure 3(a).

If the image is divided into two blocks, then overall optimization can be obtained by searching the minimum MSE as shown in Figure 3(b). If \( \text{bpp}_T \) is the total allowed bpp, and \( \text{bpp}_i \) is the bpp of block \( i \), then we search for the optimal \( \text{bpp}_i \) given the constraint that

\[
\text{bpp}_T = \left( \text{bpp}_1 + \text{bpp}_2 \right) / 2
\]

The vertical axis in Figure 3(b) is the distortion for the whole image at a given \( \text{bpp}_T \). It is interesting to note that the distortion curve for a specific compression ratio for the whole image is a “U” shape, with a unique minimum. The U shape arises because for each block, the rate-distortion curve is convex, which means the slope of the curve decreases as the bpp value increases. Therefore for the overall distortion curve, the slope decreases on the left side of the U shape, and increases on the right side of the U shape. In other words, if we assign too many bits (or too few bits) to block 1, the overall compression performance will drop. Also there is no guarantee the U shape is symmetric. Symmetry occurs when two blocks have the same rate-distortion curve.

Another parameter affecting bit allocation is speckle reduction. We use hard-thresholding to reduce speckle, depending upon the block dynamic range [7]. For blocks with high speckle noise, more aggressive hard thresholding reduces the number of bits used.

### 4.1 Blocks of Importance

Blocks that contain important features for classification or detection should be assigned more bits. If \( \{ B_1, B_2, ..., B_n \} \) is the number of bits for each image block, and \( B_T \) is the total number of bits specified by \( R \), then:

\[
B_T = B_1 + B_2 + ... + B_n
\]

Suppose that each block is the same size, and that each is compressed at the same ratio; then we could expect \( B_1 = B_2 = ... = B_n \). However, it is possible in some applications (such as the “region of interest” feature in JPEG2000) that the number of bits
assigned to each block differs, especially when different image blocks have completely different local statistics.

4.2 Dynamic Data Range

As each block can have a different dynamic range, we define a Data Dynamic Range (DDR) set \{d_1, d_2, ... d_n\} to represent this information. Since the wavelet transform is a linear, the image block that has a large data range also contains a large range of wavelet coefficients. In order to encode these coefficients with as little distortion as possible, we need to increase the quantization level by decreasing the decision step, and allocate more bits to this block.

5 Experimental Results

Experiments are done on a test image shown in Figure 6(a). The raw data were acquired in February 1998 by RADARSAT-1, and processed in July 1999 by Radarsat International. The scene includes Vancouver airport and surrounding urban areas. The comparison between the WPEB and SPIHT algorithms [4] is used to demonstrate the merits of our compression algorithm.

In the experiment, the test image of 512x1024 pixels is divided into two blocks of the same size (512x512). The variance of the left block is 1330, while the variance of the right block is 5340, and the variance for the whole image is 4350.

5.1. Bit Allocation Results

The optimal bit allocation can be obtained by searching the minima of the rate-distortion curve at specific overall compression ratios, as shown in Figure 4. The search results are shown in Table 1. It was found that the optimal compression ratio for each block is not the same as for the whole image for the 2-block scene.

In practice, the search is inefficient when the image is divided into many blocks. However, it was found that a simple statistical measure such as block standard deviation was able to give a near-optimal bit allocation.

| Table 1 Optimal compression ratio for two blocks |
|-----------------|-----------------|-----------------|-----------------|-----------------|
| Overall         | 0.5 bpp         | 1.0 bpp         | 1.5 bpp         | 2.0 bpp         |
| Left block      | 0.3 bpp         | 0.6 bpp         | 1.1 bpp         | 1.6 bpp         |
| Right block     | 0.7 bpp         | 1.4 bpp         | 1.9 bpp         | 2.4 bpp         |
The four curves in Figure 4 show the rate-distortion curves for average bit rates of 0.5 bpp, 1.0 bpp, 1.5 bpp, 2.0 bpp, respectively. The horizontal axis shows the compression ratio for Block 1, and the vertical axis is the MSE for the whole image. The bpp value for Block 2 is given by equation (3).

5.2 MSE Results

In order to compare the performance of the WPEB algorithm fairly, it is useful to compare it against the SPIHT algorithm applied over the whole image, and the SPIHT applied using two blocks (but with the same number of bits in each block). The MSE results for the SPIHT, averaged 2-block SPIHT (ASPIHT) and WPEB algorithms are shown in Figure 5. The MSE for WPEB is 25% lower than that for 1-block SPIHT, and 12% lower than the 2-block SPIHT. The improvement of the 2-block SPIHT over the 1-block SPIHT is further evidence of the advantages of blocking.

The reconstructed images at 1.0 bpp for the SPIHT and WPEB algorithms are shown in Figure 6. The WPEB image is visually better than the SPIHT image. In the left block, more details are observed in the water in the WPEB case, even though fewer bits are used. Also block artifacts can be seen in the water in the SPIHT case. In the right block, the city and airport details of the WPEB case are closer to those in the original image than the SPIHT case.
6. Conclusions

Our compression algorithm has two main strengths. First, we use wavelet packet decomposition to better represent SAR imagery's significant middle and high frequency components. Second, we utilize a block coding scheme to exploit statistical properties, such as the activity level or energy compaction, of each block. Our experimental results show that this coding scheme gives a lower MSE than traditional wavelet methods and is promising for SAR image compression.

![Figure 5. MSE results of SPIHT and WPEB encoding of the test image.](image)

7. References


Figure 6. Compression results for SPIHT and WPEB at 1.0 bpp.