Multi-Region Multi-Quality Coding Schemes for Medical Image Compression

by

Sinan Tasdoken, B.S.

A thesis submitted to the
Faculty of Graduate Studies and Research
in partial fulfilment of the requirements
for the degree of Master of Applied Science
in Electrical Engineering

Ottawa-Carleton Institute of Electrical and Computer Engineering
Department of Systems and Computer Engineering
Carleton University,
Ottawa, ON, Canada

December 16, 2002

©Sinan Tasdoken, 2002

National Library
of Canada

Acquisitions and
Bibliographic Services

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque nationale
du Canada

Acquisitions et
services bibliographiques

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

Our file Notre référence

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-79738-4

Canada

The undersigned recommend to the Faculty of Graduate
Studies and Research acceptance of the thesis entitled

Multi-Region Multi-Quality Coding Schemes for Medical Image Compression

submitted by Sinan Tasdoken,
in partial fulfilment of the requirements for
the degree of Master of Applied Science in Electrical Engineering.

Thesis Supervisor
Chairman,
Department of Systems and
Computer Engineering

Ottawa-Carleton Institute of Electrical and Computer Engineering,
Department of Systems and Computer Engineering
Faculty of Engineering
Carleton University
December 16, 2002

Abstract

Region-based coding schemes, employed in JPEG-2000 and MPEG-4 standards, are promising for medical image compression. In most medical imaging applications, the diagnostically significant information is localized over multiple, arbitrary shape regions which require high quality coding and the rest of the image might be encoded with lower fidelity. Motivated by this fact, in this thesis, we investigate multi-region multi-quality (MRMQ) coding of medical images. Proposed methods differ from existing region-of-interest (ROI) coding schemes in such a way that maximally decimated wavelet transforms are used for representation of regions rather than the identification of the region information in the transform domain. Two new MRMQ coders based on SPIHT, JPEG-2000 and EZBC algorithms are presented. These coders offer superior rate-distortion performance with additional features such as distortion scalability, spatial scalability and flexibility in precise rate control. Furthermore, a novel scheme which addresses the region size sensitivity problem of tree-based wavelet coders is proposed. Experimental results verify the strength of the proposed MRMQ coders in the utilization of the available bit rate.

Acknowledgements

I would like to thank my thesis supervisor Dr. Aysegul Cuhadar, whose critical thinking and experience has been a great source of inspiration for this study. Beyond her supervision, her kindness and support throughout my days in Canada has been invaluable.

I also would like to thank the faculty members in Ottawa-Carleton Institute of Electrical and Computer Engineering and in Bilkent University who enriched my experience in this field, especially to Prof. Enis Cetin of Bilkent University for introducing me the interesting topics in signal processing.

This short acknowledgement is very little to thank Cihan and my family, to whom I dedicate this thesis, for their constant love and support all along.

To Cihan and my family

Contents

Abstract	iii
Acknowledgements	iv
Contents	vi
List of Tables	viii
List of Figures	ix
List of Acronyms	xii
1 Introduction	1
1.1 Motivation	1
1.2 Thesis Objectives	4
1.3 Thesis Contributions	5
1.4 Thesis Outline	7
2 Wavelet Representations	8
2.1 Wavelet Representations over Countable Domains	9
2.2 Wavelet Representations over Finite Support	13
2.3 Wavelet Representations over AS-ROS	16

CONTENTS	vii
2.3.1 Multiresolution Property	18
2.3.2 Vanishing Moments	20
2.3.3 White Noise Sensitivity	20
2.4 Concluding Remarks	21
3 Wavelet Coefficient Plane Coding	23
3.1 Image Compression Goals	23
3.1.1 Bit Stream Scalability	24
3.1.2 Wavelet Coefficient Plane Coders	28
3.2 Efficient Wavelet Coefficient Plane Coders	29
3.2.1 SPIHT	29
3.2.2 JPEG-2000	33
3.2.3 EZBC	37
3.3 Concluding Remarks	39
4 Multi-region Multi-quality Image Coding	41
4.1 Overview	42
4.1.1 Previous Approaches	42
4.1.2 Our Approach	44
4.2 Wavelet Representation for Multiple AS-ROS	46
4.2.1 Constraints on Wavelet Basis Selection	46
4.2.2 Representation with Odd Length Biorthogonal Wavelet Basis	48
4.3 MRMQ Coding based on SPIHT Algorithm	50
4.3.1 MRMQ-SPIHT with Sets of Descendent Labels	50
4.3.2 MRMQ-SPIHT with Unbalanced Spatial Orientation Trees	53
4.4 MRMQ Coding based on JPEG-2000 Algorithm	55
4.4.1 Modified Tier 1	55
4.4.2 Modified Tier 2	58

CONTENTS	viii
4.5 MRMQ Coding based on EZBC Algorithm	59
4.5.1 MRMQ-EZBC with Sets of Descendent Labels	60
4.5.2 MRMQ-EZBC with Unbalanced Quadtrees	63
4.6 Performance Evaluation of MRMQ Coders	64
4.6.1 Prototype Basis Vector Selection	64
4.6.2 Comparison of Region Size Sensitivity	67
4.6.3 Evaluation of R-D Performance	69
4.6.4 MRMQ vs. ROI Coding	76
4.6.5 Summary	78
5 Conclusion	81
5.1 Summary	81
5.2 Future Work	83
Bibliography	90

List of Tables

4.1	Sorted table which contains the contribution order of code block bit streams.	58
4.2	Comparison of basis performances for encoding left lung region of chest radiograph shown in Figure 1.1 with MRMQ-SPIHTa.	65
4.3	Comparison of basis performances for encoding microcalcification cluster region of breast mammogram shown in Figure 1.2 with MRMQ-J2000.	66

List of Figures

1.1	Anatomic region partitioning of chest radiograph	2
1.2	Region partitioning of breast mammogram based on diagnostic significance	4
2.1	Conventional two channel analysis filter bank.	10
2.2	Conventional wavelet tree decomposition	12
2.3	Computation of finite length wavelet transform by symmetric extension and odd-length symmetric biorthogonal wavelet basis	17
2.4	Multiresolution reconstruction of an arbitrary shape region within <i>Lena</i> image.	19
2.5	White noise sensitivity characteristics of 9/7 filter collected for various N, L and H values values. The solid line shows the ideal condition imposed by equation (2.10).	21
3.1	Comparison of tree structures used in EZW and SPIHT algorithms .	30
3.2	Two tiered coding mechanism employed in JPEG-2000	36
4.1	Proposed MRMQ Coding Mechanism	45
4.2	An example wavelet decomposition which satisfies equation 4.1. . . .	49
4.3	Flowchart for redefinition of parent-child relationships in spatial orientation trees.	54
4.4	Comparison of labels of code block samples in conventional JPEG-2000 coding and MRMQ coding.	56

4.5	The <i>Angio</i> image used in region size sensitivity comparison and the corresponding circular region selection with radius R	67
4.6	Region size sensitivity of proposed MRMQ-SPIHTa and MRMQ-SPIHTb: Comparison of bit rate required to attain 40 dB. ROI-PSNR for circular regions of radius R	68
4.7	Region size sensitivity of proposed MRMQ-EZBCa and MRMQ-EZBCb: Comparison of bit rate required to attain 40 dB. ROI-PSNR for circular regions of radius R	69
4.8	R-D performance comparison of conventional wavelet coefficient plane coders SPIHT, JPEG-2000, and EZBC for the chest radiograph. . . .	70
4.9	R-D curves for regions Ω_i , $i \in \mathcal{Z}_4$ shown in Figure 1.1.	71
4.10	Illustration of transparency of MRMQ-J2000 to conventional JPEG-2000 coder	73
4.11	Original breast mammogram image	74
4.12	Conventional (JPEG-2000) coding of mammogram image at 0.6 bpp. In this case 16% of the bit rate is allocated to encode the background and 84% to encode the breast region.	74
4.13	MRMQ (MRMQ-J2000) coding of mammogram image at 0.6 bpp. In this case 100% of the bit rate is used to encode breast region. Diagnostically significant breast region encoded at 4 dB. better PSNR value than that of conventional coding shown in Figure 4.12.	75
4.14	R-D performance comparison of ROI-SPIHT [21] and MRMQ-SPIHTa. (a) ROI-PSNR vs. bit rate performance comparison for Lena and Baboon images. (b) Overall PSNR vs. bit rate performance comparison for Lena and Baboon images.	77
4.15	R-D performance comparison of ROI coding unit of JPEG-2000 and MRMQ-J2000. (a) ROI-PSNR vs. bit rate performance comparison for Goldhill image. (b) Background-PSNR vs. bit rate performance comparison for Goldhill image.	78

List of Acronyms

AS-ROS	-	Arbitrary Shape Region of Support
CB-AC	-	Context-Based Arithmetic Coding
DCT	-	Discrete Cosine Transform
EZBC	-	Embedded Zero-Block Coding
JPEG	-	Joint Photographic Experts Group
MPEG	-	Moving Picture Experts Group
MRMQ	-	Multi-Region Multi-Quality
MSE	-	Mean Square Error
OL-BO	-	Odd-Length Biorthogonal
PCRD	-	Post-Compression Rate-Distortion
PSNR	-	Peak Signal to Noise Ratio
R-D	-	Rate-Distortion
ROI	-	Region of Interest
SPIHT	-	Set Partitioning in Hierarchical Trees

Chapter 1

Introduction

1.1 Motivation

Medical image compression applications are *quality-driven* applications, that is, high quality should be maintained at least for certain regions in the image, since, even subtle artifacts introduced by lossy coding might alter subsequent diagnosis, which might cause severe legal consequences. Research in medical image compression is mostly concentrated on coding of the entire image plane as a single entity either losslessly or in a lossy fashion, depending on the imposed fidelity constraint. However this homogeneous quality constraint for the entire image may prevent high-performance compression. We propose that, for medical image archiving and transmission, use of heterogeneous (or region-dependent) fidelity constraints yields better utilization of available bit rate, since, for medical images, the diagnostically significant information

Figure 1.1: Anatomic region partitioning of chest radiograph

is typically localized over relatively small regions, where other regions can be encoded at a lower fidelity.

Region-based coding, which was originally investigated to incorporate object-based coding functionality to video coding standards, is found to be promising for compression of medical images. Some possible applications are:

- **Compression of ultrasound images:** A typical ultrasound image is composed of several distinct parts, involving, a nonrectangular ultrasound-scanned region, a region surrounding the Doppler plot, regions consisting of text and the background. These regions may be independently encoded at different fidelities,

depending on the the tolerable distortion constraints.

- **Compression of images with diagnostically significant information localized over arbitrary shaped regions:** Medical images, such as the chest radiograph shown in Figure 1.1, typically consist of several anatomic regions. As illustrated in Figure 1.1, the image plane for the chest radiograph can be divided into four (arbitrary shaped) anatomic regions, denoted by $\Omega_k, k = 0, \dots, 3$, namely, the background, left lung, right lung and mediastinum, respectively. Each of these regions may be independently encoded with differing fidelity constraints which can be determined by the significance of the diagnostic information contained within each region. Furthermore, diagnostically significant information may be located in a relatively small sub-region of an anatomic region, as shown for a mammogram in Figure 1.2. In this case, microcalcification clusters are localized within region Ω_1 , which should be coded with the best fidelity. Other regions, specifically, remaining part of breast region, Ω_2 , pectoral muscle, Ω_3 , and background, Ω_0 , may be encoded at a lower fidelity to preserve the overall texture of the image.

In addition to the aforementioned applications, region-based coding also enables region based querying of medical image databases and provides better visual quality than other image coding schemes if region boundaries coincide with the boundaries of anatomic objects. Moreover, nonlocal memory access problem in non-block based image codecs can be overcome by using the proposed coding schemes rather than encoding the entire image at once.

Figure 1.2: Region partitioning of breast mammogram based on diagnostic significance

1.2 Thesis Objectives

The goal of this thesis is to develop image coding algorithms which allow heterogeneous fidelity encoding of different image regions. Heterogeneous fidelity image coding typically consists of two parts. First, the image is partitioned into regions, as shown in Figure 1.1, either automatically or in a manual fashion. This is followed by encoding of extracted regions with differing fidelities. Without loss of generality, we extract the image regions manually without employing an automatic segmentation algorithm. Our primary objective is to investigate the second part, namely, Multi-region Multi-

quality (MRMQ) coding of these partitioned images. There exist numerous ways to accomplish this, one is based on object coding functionality of MPEG-4 and the other is found in the region-of-interest (ROI) coding unit of JPEG-2000. The problem with the first approach is that, due to its computational efficiency, block-based discrete cosine transform (DCT) is used for arbitrary shape object representation. However, it is well known that block DCT yields blocking artifacts for low bit rate image coding. The problem with the latter approach is that *overcomplete* representation of regions results in substantial degradation in rate-distortion (R-D) performance and impedes effective region quality control. To overcome the aforementioned problems, we use maximally decimated arbitrary shape wavelet transforms to represent multiple image regions independently. The region-based wavelet transform retains most of the features of the conventional wavelet transform, including multiresolution property and locality. Motivated by this fact, we propose MRMQ coding schemes based on the state of the art wavelet coders SPIHT, JPEG-2000 and EZBC. The proposed coders provide distortion scalability, spatial scalability and rate control flexibility for efficient communication over networks with various bandwidth conditions and computational capabilities, namely, the heterogeneous networks. In addition to this, we address the region size sensitivity problem typically encountered in tree based wavelet coders, by introducing the concept of *unbalanced trees*.

1.3 Thesis Contributions

The contributions of the thesis can be summarized as follows:

- We evaluate the extent to which wavelet representations over arbitrary shape region of support (AS-ROS) with symmetric odd-length biorthogonal (OL-BO) basis retains the properties of conventional wavelet decomposition experimentally. Our experimental results verify that the aforementioned representations over AS-ROS preserve the multiresolution, locality and noise sensitivity properties of the conventional wavelet transforms, and thereby allows use of conventional coders for efficient region-based image coding.
- We propose novel schemes to overcome the limitations of the state of the art image coding schemes for medical image compression applications. First, we develop an MRMQ coding algorithm based on the SPIHT encoder. Unlike the previously reported ROI-based SPIHT coders, the proposed scheme do not result in loss of detail information when region-based wavelet representations are encoded [29, 31]. Next, we propose a novel MRMQ coder based on JPEG-2000 coding tools. The PCRD optimization tier of JPEG-2000 is modified in such a way that the proposed algorithm enables distortion-scalable interleaving of region bit streams. Finally, we describe another MRMQ coding algorithm based on EZBC algorithm [30]. The advantage of this coder over MRMQ-J2000 is its flexibility in precise rate control without comprising the R-D performance.
- The concept of *unbalanced* spatial orientation trees and quadrees is introduced to overcome the region size sensitivity problem observed in MRMQ coding with SPIHT and EZBC encoders.

1.4 Thesis Outline

In Chapter 2, we review the conventional wavelet transform and its extension to arbitrary shape region representation based on symmetric extension and linear phase wavelet filters. We examine their similarities and differences which will be exploited in the development of efficient MRMQ coding schemes.

In Chapter 3, we define the important image compression requirements for efficient communication over heterogeneous networks. This is followed by the description of the state of the art, highly scalable, wavelet coefficient plane coders SPIHT, JPEG-2000 and EZBC, which are found to be promising for MRMQ coding of images.

In Chapter 4, we propose various schemes for MRMQ coding based on the aforementioned conventional wavelet coders. This is followed by the performance evaluation of MRMQ coders, in terms of R-D performance and region size sensitivity.

Finally, Chapter 5 summarizes the results of this thesis, highlights its contributions and future applications.

Chapter 2

Wavelet Representations

This chapter provides an overview of wavelet/subband representation of signals. The wavelet transform is described in vector space framework, which furnishes insights for representation of signals defined over arbitrary shape domains. Although arbitrary shape wavelet representations exhibit similarities to conventional wavelet transform, there are important differences which should be examined to improve the performance of the subsequent region-based encoding unit.

The chapter is organized as follows: In section 2.1, wavelet representations of infinite extent signals is described. In section 2.2, extension of infinite extent wavelet transform to finite length signal representation is discussed, which is followed by representation of signals defined over arbitrary shape region of support (AS-ROS). The chapter concludes with Section 2.4.

2.1 Wavelet Representations of Infinite Extent Signals

The basic unit used for obtaining the wavelet decomposition of a signal is the two channel filter bank structure shown in Figure 2.1. The input signal $x[n]$, defined over a set of lattice points Ψ is fed into two channels, each of which performs a filtering and subsequent downsampling operation. The downsampling operation retains only the samples positioned at lattice points Γ_i , $i = 0, 1$ and discards the remaining, where Γ_i is either a sublattice of Ψ or a coset of this sublattice. For one-dimensional case, the lattice Ψ is generally the set of integers, $\mathcal{Z}$, and Γ_i is either the set of even integers or the set of odd integers where the selection depends on the type of the decomposition filter. Thus, the signal $x[n]$ is represented by two half-rate subband signals $y_0[n]$ and $y_1[n]$, which are said to be the L-band and H-band signals respectively since the upper channel performs a lowpass filtering operation ($a_0[n]$) and the lower channel performs a highpass filtering operation ($a_1[n]$).

The input signal and the subband signals are related by

$$y_i[k] = \langle x[n] \mid a_i[2k - n] \rangle, \quad i = 0, 1 \quad (2.1)$$

where $\langle a[n] \mid b[n] \rangle$ denotes the inner product of two sequences (or vectors) $a[n]$ and $b[n]$. Henceforth, we will use sequence notation $a[n]$ and vector notation $\mathbf{a}$ interchangeably where $\mathbf{a} = [\cdots a[-1] \ a[0] \ a[1] \ \cdots]^T$ and T denotes the transpose of the corresponding vector (or matrix). For compact representation, let's define the interleaved subband signal $y[n]$ such that $y[2n + i] = y_i[n]$, $i = 0, 1$. Since filtering and

Figure 2.1: Conventional two channel analysis filter bank.

downsampling operations are linear operations, the mapping from $x[n]$ to $y[n]$ can be described by a linear operator $\mathbf{A}$:

$$\mathbf{y} = \underbrace{\begin{bmatrix} \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \\ \dots & a_0[N] & a_0[N-1] & \dots & \dots & a_0[0] & \dots & \dots & \dots \\ \dots & a_1[N] & a_1[N-1] & \dots & \dots & a_1[0] & \dots & \dots & \dots \\ \dots & a_0[N+2] & a_0[N+1] & a_0[N] & \dots & a_0[2] & a_0[1] & a_0[0] & \dots \\ \dots & a_1[N+2] & a_1[N+1] & a_1[N] & \dots & a_1[2] & a_1[1] & a_1[0] & \dots \\ \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}}_{\mathbf{A}} \mathbf{x} \quad (2.2)$$

Now consider the synthesis of the decomposed signal from the corresponding subband signals. In this phase, the subband signals are upsampled from Γ_i to Ψ , filtered by synthesis filters $s_i[n]$, $i = 0, 1$ and superposed to obtain the reconstructed signal, $\hat{x}[n]$. Similar to decomposition case, the operation can be described by the

linear operator $\mathbf{S}$:

$$\hat{\mathbf{x}} = \underbrace{\begin{bmatrix} \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \\ \dots & s_0[N] & s_0[N-1] & \dots & \dots & s_0[0] & \dots & \dots & \dots \\ \dots & s_1[N] & s_1[N-1] & \dots & \dots & s_1[0] & \dots & \dots & \dots \\ \dots & s_0[N+2] & s_0[N+1] & s_0[N] & \dots & s_0[2] & s_0[1] & s_0[0] & \dots \\ \dots & s_1[N+2] & s_1[N+1] & s_1[N] & \dots & s_1[2] & s_1[1] & s_1[0] & \dots \\ \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}}_{\mathbf{S}}^T \mathbf{y} \quad (2.3)$$

An important requirement for wavelet representation is the *perfect reconstruction* constraint that is the reconstructed signal $\hat{x}[n]$ is the same as the decomposed signal $x[n]$. This can be accomplished if column vectors of $\mathbf{S}$ span the space of signals defined over Ψ . If the representation is *maximally decimated*, then the column vectors provide a *basis* for that space.

In summary, the wavelet basis construction process consists of finding *prototype* lowpass and highpass vectors $s_0[n]$ and $s_1[n]$, shifted versions of which provide a basis for the vector space of signals defined over the countable domain Ψ . For this basis to yield an efficient representation, the energy of the input signal should be compactly represented by a few samples, i.e., the resulting representation has as low *entropy* as possible (typically, the representation is as *sparse* as possible). The conventional wavelet decomposition of a signal is obtained by recursively applying two-channel analysis scheme to the produced L-band signals and in case of wavelet packet representation, the two-channel analysis scheme is recursively applied to either the L-band or H-band signals. The conventional wavelet decomposition of one dimensional signal

Figure 2.2: Conventional wavelet tree decomposition

for three levels is shown in Figure 2.2. Representation for multidimensional signals can be obtained by using tensor products of one dimensional. For two dimensional case, the wavelet decomposition is computed by applying the two channel filter bank structure along vertical and horizontal dimensions, resulting in four subband signals, namely LL, HL, LH, and HH-band signals. Nonseparable methods are also proposed for the representation of multidimensional signals, but they have high computational complexity and they do not offer significant improvement in compression efficiency [12].

2.2 Wavelet Representations over Finite Support

In this section, we examine the representation of finite length signals with the two channel filter bank structure introduced in the previous section. This can be accomplished either by appropriate extrapolation [4, 15, 13, 10] or boundary filter construction [7, 18, 1, 28]. Our formulation is based on the extrapolation methods, due to their lower computational complexity than boundary filter construction.

Let the signal $x[n]$ be limited to a finite support of length L . The representation of this signal requires L basis vectors, thus the analysis operator used for decomposition of this signal will be of size L -by- L . Therefore, we need to truncate doubly infinite matrix $\mathbf{A}$, given in equation (2.2), to an L -by- L matrix $\mathbf{A}_{\mathbf{q}}$ in such a way that the $\mathbf{A}_{\mathbf{q}}$ is invertible[†]. Direct truncation of $\mathbf{A}$ does not necessarily produce an invertible matrix since even though a matrix is nonsingular, this does not imply all L -by- L submatrices within this matrix are nonsingular.

Without loss of generality, assume that the analysis and synthesis filters are of finite impulse response (FIR), restricted to domain $\mathcal{Z}_N = \{0, 1, \dots, N-1\}$. To construct the L -by- L matrix, $\mathbf{A}_{\mathbf{q}}$, we first choose L rows of $\mathbf{A}$ since rows of $\mathbf{A}$ are

[†]This is equivalent to saying that the columns of the inverse of $\mathbf{A}_{\mathbf{q}}$ provide a basis for space of finite support signals of length L .

linearly independent. The resulting matrix has the form

$$\tilde{\mathbf{A}} = \begin{bmatrix} \dots & a_0[N-1] & a_0[N-2] & \dots & \dots & a_0[0] & 0 & 0 & 0 & \dots & \dots \\ \dots & a_1[N-1] & a_1[N-2] & \dots & \dots & a_1[0] & 0 & 0 & 0 & \dots & \dots \\ \dots & 0 & 0 & a_0[N-1] & \dots & a_0[2] & a_0[1] & a_0[0] & 0 & \dots & \dots \\ \dots & 0 & 0 & a_1[N-1] & \dots & a_1[2] & a_1[1] & a_1[0] & 0 & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & 0 & 0 & 0 & \dots & \dots & \dots & \dots & \dots & a_1[0] & \dots \end{bmatrix}$$

Since rows of $\tilde{\mathbf{A}}$ are linearly independent, the matrix formed by discarding the all-zero columns of $\tilde{\mathbf{A}}$,

$$\tilde{\tilde{\mathbf{A}}} = \begin{bmatrix} a_0[N-1] & a_0[N-2] & \dots & \dots & a_0[0] & 0 & 0 & 0 & \dots & 0 \\ a_1[N-1] & a_1[N-2] & \dots & \dots & a_1[0] & 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & a_0[N-1] & \dots & a_0[2] & a_0[1] & a_0[0] & 0 & \dots & 0 \\ 0 & 0 & a_1[N-1] & \dots & a_1[2] & a_1[1] & a_1[0] & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & 0 \\ 0 & 0 & 0 & 0 & \dots & \dots & \dots & \dots & \dots & a_1[0] \end{bmatrix}$$

will also have rank L. Consider representation of signal $x[n]$ with the vectors given as rows of $\tilde{\tilde{\mathbf{A}}}$. Since the vectors have L+Q elements (where Q is an integer depending on the support of analysis filters), we need to extend signal $x[n]$ by Q samples to compute its inner products with these vectors. The extension signal $x_{ex}[n]$ is defined as a function of $x[n]$, i.e., $\mathbf{x}_{ex} = f(\mathbf{x})$. In case of linear extensions:

$$\mathbf{x}_{ex} = \mathbf{E}_a \mathbf{x} \quad (2.4)$$

where $\mathbf{E}_a$ is Q-by-L matrix. Thus

$$\mathbf{y} = \underbrace{\begin{bmatrix} \mathbf{A}_0 & \mathbf{A}_1 \end{bmatrix}}_{\tilde{\tilde{\mathbf{A}}}} \begin{bmatrix} \mathbf{x} \\ \mathbf{x}_{\text{ex}} \end{bmatrix} \quad (2.5)$$

where $\tilde{\tilde{\mathbf{A}}}$ is partitioned into submatrices $\mathbf{A}_0$ and $\mathbf{A}_1$ of size L-by-L and L-by-Q, respectively. This leads to

$$\mathbf{y} = (\mathbf{A}_0 + \mathbf{A}_1 \mathbf{E}_a) \mathbf{x} = \mathbf{A}_q \mathbf{x} \quad (2.6)$$

which completes the derivation of truncated matrix $\mathbf{A}_q$. The representation associated with extension matrix $\mathbf{E}_a$ is invertible if and only if $(\mathbf{A}_0 + \mathbf{A}_1 \mathbf{E}_a)$ is nonsingular.

To construct the truncated synthesis operator, $\mathbf{S}_q$, we follow the steps given above to obtain $\tilde{\tilde{\mathbf{S}}}$ from $\mathbf{S}$ and compute the inverse transform by appending an extension to the subband signal $y[n]$, i.e., $\mathbf{y}_{\text{ex}} = \mathbf{E}_s \mathbf{y}$. Similar to partitioning of $\tilde{\tilde{\mathbf{A}}}$ into $\mathbf{A}_0$ and $\mathbf{A}_1$, we partition $\tilde{\tilde{\mathbf{S}}}$ into two submatrices $\mathbf{S}_0$ and $\mathbf{S}_1$ and this gives

$$\mathbf{x} = (\mathbf{S}_0 + \mathbf{S}_1 \mathbf{E}_s) \mathbf{y} \quad (2.7)$$

Therefore, $\mathbf{E}_a$ and $\mathbf{E}_s$ are related by

$$(\mathbf{S}_0 + \mathbf{S}_1 \mathbf{E}_s) = (\mathbf{A}_0 + \mathbf{A}_1 \mathbf{E}_a)^{-1} \quad (2.8)$$

In this formulation, we did not mention any constraints on extension type, thus it provides a general framework for obtaining a representation of finite support signals with any type of prototype filter selection, such as orthogonal, linear phase, nonlinear

phase. Additionally, if the extension operation uses only boundary samples, the solution to the problem of determining optimal[†] matrix $\mathbf{E}_a$ is also a solution for optimal boundary basis vector construction problem. In the following section, an efficient method based on *symmetric extension* and odd length biorthogonal filters will be described for the representation of signals defined over arbitrary shape region of support (AS-ROS).

2.3 Wavelet Representations over Arbitrary Shape Region of Support

A representation scheme that trades off computational complexity and representation sparsity on equal footing is based on symmetric extension and linear phase decomposition filters [15]. The following fact provides the underlying mechanism of the proposed method:

Fact 1 *Consider extension of a finite-length signal $x[n]$ to an infinite length signal $\tilde{x}[n]$ by symmetric extension. The extended signal will have the form*

$$\tilde{x}[n + l(2N - 2)] = \begin{cases} x[n], & \text{if } n \in \mathcal{Z}_N; \\ x[2N - 2 - n], & \text{if } n \in \mathcal{Z}_{2N-2} \setminus \mathcal{Z}_N. \end{cases} \quad l \in \mathcal{Z} \quad (2.9)$$

where support of $x[n]$ is $\mathcal{Z}_N$. The inner products of $\tilde{x}[n]$ with odd length symmetric

[†]Optimal in the sense that the extension method preserves the vanishing moment and frequency selectivity properties of interior basis vectors.

wavelet basis vectors results in a subband signal $\tilde{y}[n]$ which has the same form as $\tilde{x}[n]$, i.e., it can be identified by a length N signal, $y[n]$, as given in equation (2.9).

Figure 2.3: Computation of finite length wavelet transform by symmetric extension and odd-length symmetric biorthogonal wavelet basis

The fact is depicted in Figure 2.3 to find the wavelet representation of signal $x[n] = \{a, b, c, d, e\}$. In the figure, the prototype lowpass basis vector is given as $[\cdots 0 \chi \beta \alpha \beta \chi 0 \cdots]^T$ and prototype highpass basis vector is given as $[\cdots 0 \varepsilon \delta \varepsilon 0 \cdots]^T$. Recall that all basis vectors are even shifted versions of these prototype vectors. As can be seen, due to symmetry in $\tilde{x}[n]$ and prototype basis vectors, the produced subband signal $\tilde{y}[n]$ exhibits the same type of symmetry seen in $\tilde{x}[n]$ and thus can be expressed in terms of the signal $y[n] = \{k, l, m, n, o\}$. The signal $y[n]$ is said to be the *subband decomposition* of $x[n]$. Notice that in case of symmetric extension, there is no need to exploit equation (2.8) to compute synthesis extension matrix $\mathbf{E}_s$ from $\mathbf{E}_a$, since the extension matrices $\mathbf{E}_a$ and $\mathbf{E}_s$ are equal.

An important question is how efficient the above described method is for representation of signals defined over AS-ROS. To address this this question, we need to consider the important properties provided by conventional wavelet transforms, i.e., wavelet representations over countable domains such as 2-D plane and whether the representations over AS-ROS retain these properties or not.

2.3.1 Multiresolution Property

The proposed representation over AS-ROS maintains the multiresolution property since the proposed transform maps the input signal into a set of subband signals, subsets of which can be used for lower resolution reconstruction of the signal defined over AS-ROS. This is illustrated in Figure 2.4. In the figure, progressive higher resolution reconstruction of an arbitrary shape region within *Lena* image is obtained

by using more and more subbands.

Figure 2.4: Multiresolution reconstruction of an arbitrary shape region within *Lena* image.

2.3.2 Vanishing Moments

Another property of the conventional wavelet transform is that the low frequency signals are completely represented by the L-band signal. A wavelet basis is said to provide K vanishing moments if all polynomial signals $x[n] = n^k$ can be perfectly reconstructed by L-band samples only, $\forall k \in \mathcal{Z}_K$. For $k = 0$, this means there is no DC leakage in the representation. The symmetric extension prevents DC leakage but does not necessarily provide higher order vanishing moments. However, in [18], it is observed that, the extensions which preserve all vanishing moment conditions of prototype basis vectors do not yield superior R-D performance compared to the symmetric extension.

2.3.3 White Noise Sensitivity

To prevent adaptive bit allocation, it is desirable that the samples within different subbands have equal sensitivity to quantization noise, i.e., the white noise introduced during quantization should not be observed as colored noise during reconstruction. A condition to obtain this property is given by

$$\sum_{i \in \mathcal{Z}_L} \|\mathbf{s}_{l,i}\|^2 \approx \sum_{i \in \mathcal{Z}_H} \|\mathbf{s}_{h,i}\|^2 \quad (2.10)$$

where $\mathbf{s}_{l,i}, i \in \mathcal{Z}_L$ denotes the lowpass synthesis vectors used for reconstruction and $\mathbf{s}_{h,i}, i \in \mathcal{Z}_H$ are the corresponding highpass vectors. Notice that $L + H = N$ where N is the length of the decomposed signal. The comparison of right hand side and left hand side of equation (2.10) for biorthogonal linear phase 9/7 filter [36] is provided

Figure 2.5: White noise sensitivity characteristics of 9/7 filter collected for various N , L and H values. The solid line shows the ideal condition imposed by equation (2.10).

in Figure 2.5. In the figure, the deviation from the ideal occurs for small region sizes (or N) of odd length. In these situations, the signal energy may be represented in the H-band rather than the L-band, which violates the decaying spectrum assumption of conventional wavelet coders.

2.4 Concluding Remarks

In this section, we reviewed the conventional wavelet decomposition and its extension to arbitrary shape region representation by linear phase filters and symmetric extension. Wavelet representations over AS-ROS retains most of the properties of

conventional wavelet transform but exhibits different behavior when regions of relatively small size are represented, i.e., the energy of the signal might be represented in the H-band, unlike the conventional case. Therefore, an efficient region based coder should address this region size sensitivity problem.

Chapter 3

Wavelet Coefficient Plane Coding

In this chapter, we introduce the requirements for compression of images which allows efficient communication in environments with varying capabilities, and describe highly scalable, state of the art wavelet-based coders.

The chapter is organized as follows: In Section 3.1, important image coding goals are defined. In Section 3.2, the implementation details of wavelet coders, that are promising for MRMQ coding, are provided. The chapter concludes with section 3.3.

3.1 Image Compression Goals

An image compression scheme consists of an encoder which produces a bit stream providing a low bit rate description of the image and a decoder which recovers the

original image (in case of lossless compression) or the relevant information in the original image (in case of lossy compression) from the produced bit stream. The lossless compression schemes do not offer as high compression ratios as lossy compression methods [37]. Therefore, we will be mainly interested in lossy coding schemes in this thesis.

The primary goal in lossy image compression is to minimize the distortion observed in the reconstructed image due to a given (fixed) bit rate constraint for encoding the image. Since the human visual system (HVS) evaluates the quality of the reconstructed image, the distortion metric used for minimization should be defined as a measure mapping the subjective difference between the visual qualities of the reconstructed image and the original image to set of real numbers, $\mathcal{R}$. However, HVS is not a simple system to model and definition of such an objective visual metric is not a trivial problem. Instead, Mean Square Error (MSE) or Peak Signal-to-Noise Ratio (PSNR) are commonly used for measuring distortion. The resulting rate-distortion (R-D) curve is an important entity for evaluating the performance of a compression scheme. In the following section, we define important scalability requirements for efficient image communication. It is important to point out that the wavelet representation provides the necessary framework for achieving the defined goals.

3.1.1 Bit Stream Scalability

A picture archiving and communication system (PACS) establishes communication of medical data between medical experts with different processing capabilities and

network access bandwidths. There is an increasing demand to service each user optimally according to her/his access bandwidth and computational resources. Therefore, *scalability* of the bit stream is essential for communication over such an environment. Scalability means that the produced bit stream can be manipulated in a simple manner to obtain another bit stream suitable for decoding at a different bit rate, display resolution and/or computational complexity [34]. These important scalability requirements are defined below which is followed by a discussion of their importance for communication over heterogeneous networks.

Definition 1 (Distortion Scalability) A bit stream is said to have *distortion scalability* if the prefixes of the bit stream yields lower rate descriptions of the image at the highest possible levels of quality.

The definition is given by Ordentlich et al. [20]. In this case, a lower quality description of the image can be obtained simply by truncating the higher rate bit stream without re-executing the compression algorithm to obtain the lower quality description of the image, as in JPEG. A natural application of distortion scalable (or embedded) bit stream is progressive image transmission (PIT). In PIT, the image is transmitted over a low bandwidth channel, and the receiver updates the quality of the reconstructed image gradually as it receives more bits from the bit stream. Another important use of embedded encoding is prioritized data protection. In this framework, the idea is to vary the amount of error correction used for each part of the data, i.e., the important data is protected more strongly than the remaining part. This strategy is beneficial for communication over highly error prone channels, such

as wireless channels, since a reasonable description of the image can be recovered at the decoder side even in the presence of unanticipated noise in the communication link.

Definition 2 (Spatial Scalability) A bit stream is said to provide *spatial scalability* if transcoding of the bit stream into arbitrary spatial resolutions can be accomplished by simple parsing, i.e., each successive part of the bit stream contains the data required for reconstructing a higher resolution version of the image.

This feature is useful in any system that needs to transmit varying resolution versions of an image to output devices with different display resolution capabilities (such as monitors with different display resolutions or printers with varying resolutions). Unfortunately, distortion scalability and spatial scalability are conflicting requirements, i.e., if the bit stream is ordered according to spatial scalability, it will not necessarily be distortion scalable or vice versa. A practical solution to this problem is proposed by Taubman and Zakhor [33], namely the multilayer coding approach. In multilayer coding, the bit stream consists of layers (or chunks) where each successive layer can be used to reconstruct the image with higher fidelity. Additionally, each layer is ordered in a spatial scalable way, that is, each layer can be parsed efficiently to extract the subsets required for reconstructing the image at a different resolution.

Another important requirement for a compression scheme is *complexity scalability*. In applications where encoding has to be performed under varying complexity conditions (such as varying CPU power or varying memory resources), it is beneficial to design the algorithm to allow complexity scalability to overcome nonlocal memory

access and memory overflow problems or immoderate execution speed. It is important to mention that distortion and spatial scalability naturally lead to scalability in complexity, i.e., one can reduce the complexity of the algorithm by trading either reconstruction quality or resolution.

In addition to the above requirements, to enable *interactive encoding*, the bit stream should support region accessibility, that is, it should allow decoding an arbitrary cropped portion of the image without decoding the entire bit stream. In this framework, the image can be encoded by imposing heterogeneous (or region dependent) fidelities, e.g., in case of medical imaging, fidelity of each region may be determined considering diagnostic significance of each region.

In summary, the aforementioned goals enhance processing of the bit stream in an efficient way to obtain different rate/resolution descriptions of the image or to reconstruct some arbitrary shape sub-region of the image without decoding the entire image. Since these operations are performed on the compressed bit stream rather than the original image, they imply computation at a lower rate than processing the original image samples. Wavelet representations provide the necessary framework to achieve all of these goals efficiently, since:

1. Wavelet transform is a multiresolution transform, that is, it decomposes the image into subbands which can be encoded successively to achieve spatial scalability,
2. A multilayer bit stream can be obtained by encoding each subband in an embedded way followed by interleaving of these subband bit streams into multilayers

to achieve distortion scalability,

3. Wavelet basis is a local basis, so the wavelet coefficients corresponding to different regions can be differentiated efficiently, and can be encoded with differing fidelity constraints to achieve heterogeneous fidelity coding.

Additionally, wavelet-based coding techniques are promising for low bit rate coding, since they offer better visual quality than the block-based coding techniques, such as DCT [15].

3.1.2 Wavelet Coefficient Plane Coders

The wavelet coefficient coders can be coarsely grouped under two themes. In the first group, geometrically constrained data structures (such as blocks or trees) are used to efficiently describe low activity areas, i.e., areas consisting of small magnitude coefficients [25, 23, 39, 14]. The SPIHT algorithm of Said and Pearlman is one of the most efficient coders proposed within this framework [23]. In the second group, the wavelet coefficients are encoded sequentially (one after another) by using context based modeling and subsequent entropy coding [33, 38, 17, 32, 8]. The problem with these coders is the high computational cost due to model based arithmetic coding algorithm. Within this group, JPEG-2000 [35] and EZBC [8] are the coders with relatively reduced computational complexity and superior R-D performance. Therefore, we will investigate adaptation of highly scalable SPIHT, JPEG-2000 and EZBC coders for MRMQ coding due to their superior R-D performance and low

computational complexity.

3.2 Efficient Wavelet Coefficient Plane Coders

3.2.1 SPIHT

Said and Pearlman's SPIHT algorithm is a bit plane coding algorithm based on the EZW coder of Shapiro [25]. One of the main differences between EZW and SPIHT is the way the zerotrees are built. This is depicted in Figure 3.1. In EZW, *all* coefficients in the lowest frequency subband have three offsprings. Distinctively, the tree structures in SPIHT, namely the spatial orientation trees, have balanced structure, i.e., all coefficients either have four offsprings or do not have any offsprings. Combined with efficient management of significance information, this spatial orientation tree structure substantially improves EZW algorithm in both speed and R-D performance. The details of the algorithm are given below.

Definitions. Let Ω be the 2-D plane over which the image function is defined. Wavelet transform maps this data to a higher dimensional space, $\Lambda^\dagger$. Let $w[\mathbf{n}]$ denote the samples of the wavelet image, $\mathbf{n} \in \Lambda$. The wavelet image can be delineated as a union of K sub-images, Λ_k 's, $k \in \mathcal{Z}_K$.

[†]To simplify visual interpretation, wavelet representation is usually depicted on a 2-D plane, which we refer as wavelet coefficient plane and the function defined over this plane is said to be the wavelet image. Therefore, Λ will be defined as the set of indices $\{(k, l) | k \in \mathcal{Z}_N, l \in \mathcal{Z}_M\}$ for an N -by- M image.

Figure 3.1: Comparison of tree structures used in EZW and SPIHT algorithms

Due to availability of fixed-precision representation, we first quantize the coefficients to signed 16-bit (or more generally to signed $\tilde{m}$ -bit format), $\hat{w}[\mathbf{n}]$, by

$$\hat{w}[\mathbf{n}] = \text{sgn}(\hat{w}[\mathbf{n}]) \left\lfloor \frac{|w[\mathbf{n}]|}{\Delta} \right\rfloor \quad (3.1)$$

where $\text{sgn}(\cdot)$ function returns the sign of the coefficient and Δ is the quantization step size. In practice, Δ is chosen in such a way that the coefficient with maximum magnitude is confined to $\tilde{m}$ bit registers. The binary representations of the magnitudes of these quantized coefficients will be given by:

$$|\hat{w}[\mathbf{n}]| = \sum_{m=0}^{\tilde{m}-1} b_m[\mathbf{n}] 2^m \quad (3.2)$$

To achieve fine embedding, $b_{m+1}[\mathbf{n}]$ is described $\forall \mathbf{n} \in \Lambda$ before all $b_m[\mathbf{n}]$'s, and so on. To efficiently encode these bit planes spatial orientation trees are used. A spatial

orientation tree is built in a recursive way, i.e., if coefficient $w[\mathbf{n}]$ is in a spatial orientation tree, so is its offsprings, $\mathcal{O}[\mathbf{n}]$, defined as

$$\mathcal{O}[\mathbf{n}] = \{w[2\mathbf{n} + (k, l)] | k \in \mathbb{Z}_2, l \in \mathbb{Z}_2\} \quad (3.3)$$

Therefore, the index of the root node is a sufficient parameter to refer to entire set of coefficients within the spatial orientation tree.

To describe the SPIHT algorithm, it is instructive to define four types of coefficient sets: $\mathcal{T}[\mathbf{n}]$, the set of wavelet coefficients within a spatial orientation tree with root index $\mathbf{n}$; $\mathcal{D}[\mathbf{n}]$, the set of descendent coefficients defined as $\mathcal{T}[\mathbf{n}] \setminus \mathbf{n}$; $\mathcal{L}[\mathbf{n}]$, the set of descendents of offsprings defined as $\mathcal{D}[\mathbf{n}] \setminus \mathcal{O}[\mathbf{n}]$, and $\mathcal{H}$, the set of wavelet coefficients within lowest frequency subband (which is equivalent to Λ_0). Additionally, the significance indicator function $\sigma^{(\tilde{m})}$ is defined as

$$\sigma^{(\tilde{m})}(\mathcal{F}) = \bigvee_{\mathbf{n} \in \mathcal{F}} \bigvee_{m=\tilde{m}}^{\tilde{m}} b_m[\mathbf{n}] \quad (3.4)$$

where $\bigvee$ denotes bitwise or operation and $\mathcal{F}$ is a set of wavelet coefficients, $\mathcal{F} \subseteq \Lambda$. The encoding proceeds by computing the significance of sets of wavelet coefficients (within spatial orientation trees) for each bit plane level and partitioning of these sets based on the computed significance information. Three lists are maintained to record the partitioning information, namely, the list of insignificant pixels (LIP), list of insignificant sets (LIS) and the list of significant pixels (LSP). The encoding algorithm is summarized below.

Algorithm.

1. Initialization: The encoder first transmits the value of Δ to acknowledge the

real dynamic range of wavelet coefficients. LIP is initialized by set of indices $\mathbf{n} \in \mathcal{H}$. If $\mathbf{n} \in \mathcal{H}$ is a spatial orientation tree root, that is, if $\mathcal{D}[\mathbf{n}] \neq \emptyset$, then $\mathbf{n}$ is used also used in LIS initialization as type A entry[†]. LSP is initialized to empty list.

2. Main Part: The encoding proceeds from bit plane level $m = \ddot{m}$ to $m = 0$. For each level, initially, LIP is parsed to determine whether its entries become significant or not (one bit per entry is consumed to describe this information). List entries that are found to be significant are moved to LSP. Next, entries of LIS are parsed. When LIS entry of type A is being processed, $\sigma^{(m)}(\mathcal{D}[\mathbf{n}])$ is transmitted. If this value is one, then significance values of each $|w[\mathbf{n}]|$, $\mathbf{n} \in \mathcal{O}(\mathbf{n})$ is encoded and these offspring coefficients are moved either to LIP or LSP depending on their significance. Additionally, this type A LIS entry is moved to the end of LIS as type B entry. Alternatively, when an entry of type B is being processed, the encoder transmits $\sigma^{(m)}(\mathcal{L}[\mathbf{n}])$ and if this value is equal to one, spatial orientation tree with root index $\mathbf{n}$ is partitioned into four sub-trees each with root indices $\mathbf{n} \in \mathcal{O}(\mathbf{n})$ and each sub-tree is appended to the end of LIS as type A entry. Finally, LSP is parsed to send the refinement bits (i.e. n^{th} bit of the entry if bit plane n is being encoded) for coefficients found to be significant at some higher bit plane level.

[†]If an LIS entry, $\mathbf{n}$, is of type A, then subsequent significance test is performed on $\mathcal{D}[\mathbf{n}]$. Otherwise, significance test is performed on $\mathcal{L}[\mathbf{n}]$.

3.2.2 JPEG-2000

JPEG-2000 is the recent still image compression standard reported by Joint Photographic Experts Group (JPEG) [35]. JPEG-2000 adopts model based method for encoding of wavelet coefficient bit planes, namely the EBCOT algorithm of Taubman, rather than tree-based methods since it yields excellent R-D performance while allowing spatial scalability with lower complexity.

In EBCOT, each subband Λ_k , $k \in \mathcal{Z}_K$, is partitioned into (fixed size) code blocks, $\mathcal{B}_k[\mathbf{i}]$, where $\mathbf{i}$ is the index of the code block within subband k . A separate bit stream is obtained for each code block by employing Ordentlich's fractional bit plane coding method and context based arithmetic coding. R-D optimal interleaving (or bit stream layer formation) is achieved by PCRD scheme, which trades rate-control flexibility slightly to achieve optimal embedding. In this section, we will describe the bit plane coding modules of JPEG-2000 which is followed by a brief overview of PCRD optimization method.

Bit Coding Modules. To efficiently encode bits coming from coefficients within $\mathcal{B}_k[\mathbf{i}]$, EBCOT uses context-based arithmetic coding (CB-AC) machinery. Three classes of binary information to be encoded by CB-AC engine are:

- *Significance information* coming from each coefficient,
- Once a coefficient becomes significant, its *sign information*,
- If a coefficient is found to be significant at some bit plane level, its *refinement information* at subsequent bit plane levels.

To reduce the complexity due to arithmetic coding, EBCOT first partitions a code block, $\mathcal{B}_k[\mathbf{i}]$, into sub-blocks, $\mathcal{S}_k[\mathbf{i}](\mathbf{j})$ and employs quadtree coding to locate the sub-blocks which become significant at some bit plane level. In doing so, the number of bits to be processed by CB-AC engine is significantly reduced. Once a sub-block is found to be significant, the coefficients within that sub-block are scanned according to fractional bit plane coding order, and each information bit is encoded by one of following bit coding modules according to which aforementioned classes it belongs:

1. Zero Coding Module (ZCM). The significance bits are encoded with this module. The coding context is determined by the significance of neighboring coefficients and the orientation of subband in which the code block resides. The lookup table for identifying the coding contexts is provided in [32].
2. Run-length Coding Module (RLM). If a run of four insignificant coefficients with insignificant neighbors is observed, the significance of this low activity region is represented by a single bit. A separate coding context is used to describe this bit. This module further reduces the amount of information processed by the CB-AC engine.
3. Sign Coding Module (SCM). The sign information is encoded by this module, by exploiting the signs of neighboring coefficients.
4. Magnitude Refinement Module (MRM). The refinement bits are processed by this module by using three coding contexts.

To ensure embedding within a bit plane level, the bits coming from coefficients are

ordered according to idea of fractional bit planes [20]. In other words, the coefficients that are expected to become significant are encoded prior to other significance and refinement bits. To accomplish this, for each insignificant coefficient, the significance of its neighbors is examined, and insignificant coefficients with significant neighbors are described before magnitude refinement bits. Refinement bit encoding is followed by description of remaining (insignificant) bits that has not been processed during first two passes. In summary, each bit plane is processed in multiple passes, namely, *Forward Significance Propagation Pass*, *Magnitude Refinement Pass* and *Normalization Pass*.

Post Compression R-D (PCRD) Optimization. Assume that the bit streams of each code block are obtained by using above bit coding modules. To achieve a specified fidelity with minimum bit rate consumption, we need to truncate these bit streams in a certain way and efficiently interleave the resulting truncated bit streams. EBCOT adopts the convention of using ending points of each successive fractional bit plane coding phases as possible truncation points. Therefore, for each code block, if $\tilde{m}$ bit planes are encoded, then there are $3\tilde{m} - 2$ possible truncation points.

During the bit plane coding, the slope of R-D curve is recorded for each fractional bit plane coding phase and this information is used to determine the *feasible* truncation points. The set of *feasible* truncation points is simply the subset of *possible* truncation points lying on the *convex hull* of the R-D curve. These points can be identified by employing convex hull analysis [26]. Finally, the bit streams coming from each code block are interleaved according to the R-D slopes for each feasible

truncation point, ordered from high to low R-D slope.

Figure 3.2: Two tiered coding mechanism employed in JPEG-2000

Algorithm. In summary, the coding engine of JPEG-2000 consists of two tiers as shown in Figure 3.2. In the first tier, a separate bit stream is obtained for each code block by bit plane coding modules and meanwhile the (operational) R-D slopes for each fractional bit plane coding phase are collected. The second tier analyzes the collected R-D curves of compressed bit streams and determines the contribution of each code block to the overall bit stream in R-D optimal sense. This two-tiered scheme results in excellent R-D performance with low computational cost. However, since the optimization is performed after collecting the compressed descriptions of *all*

code blocks, the precise rate control efficiency is traded to PCRD optimization.

3.2.3 EZBC

Similar to JPEG-2000, EZBC is a model based coder with relatively low computational cost compared to that of other model-based sequential wavelet image coders without comprising R-D performance. This is accomplished by employing quadtree coding to efficiently zoom the clusters of high energy coefficients (at the coefficient level) and thereby reducing the size of information to be processed by the CB-AC unit. Unlike spatial orientation trees, quadtrees do not have nodes fed by coefficients from different spatial resolutions. Consequently, their use for significance coding naturally leads to spatial scalability. The advantage of EZBC over JPEG-2000 is the precise rate control flexibility, that is, produced bit stream is already interleaved in order of importance and there is no need for a second tier to organize the bits coming from different information sources after the termination of bit stream formation phase. The implementation details of the EZBC algorithm is summarized below.

Definitions. Recall that, wavelet transform maps the image into K subbands. EZBC makes use of quadtree coding to describe the significance partitioning for each subband separately. Let QT_k be the quadtree of k^{th} subband, with depth D_k . A node in this tree is denoted by $QT_k[l](\mathbf{n})$ where l is the tree depth (level) in which the node resides and $\mathbf{n}$ is the index of the coefficient within that level. The leaf-nodes, $QT_k[0](\mathbf{n})$ store the absolute values of *quantized* wavelet coefficients within subband k and the nodes at the higher levels of the quadtree, $QT_k[l](\mathbf{n})$, $l > 0$, contain the

maximum of the values stored in its four offsprings, $QT_k[l+1](\mathbf{n} + (p, q))$, $p, q \in \mathcal{Z}_2$.

Similar to the SPIHT algorithm, the tree partitioning information is maintained in lists of nodes, namely, the lists of insignificant nodes (LIN's) and the lists of significant pixel (LSP's). To obtain better embedding, EZBC uses different LIN's for each subband and quadtree level, denoted by $LIN_k[l]$, and different LSP's for each subband, denoted by LSP_k . The outline of the algorithm is given below. **Algorithm.**

1. Initialization: Similar to SPIHT, the encoder first transmits the quantization step size Δ . Specifically, $LIN_k[D_k]$'s are initialized to $QT_k[D_k](\mathbf{0})$ (or the root node of the quadtree QT_k) and all other LIN's and LSP's are initialized to empty list.
2. Main Part: The encoding consists of two phases. In the first phase, LIN's are parsed. The processing order of $LIN_k[l]$'s is l ranging from 0 to $\max_k D_k$ and for fixed l , k from 0^{th} subband to $(K-1)^{th}$ subband. This processing order choice is due to distortion scalability requirement. For the m^{th} bit plane, each list $LIN_k[l]$ is processed as follows: For each entry $\mathbf{n}$ in the list, we first transmit $\sigma^{(m)}(\mathbf{n})$. If the node is significant, and moreover if the node is a leaf node, the information about corresponding subband coefficient is encoded by *CB-AC engine* where context models are based on the significance and signs of neighboring subband samples as in JPEG-2000. Otherwise, the node is removed from the list and information about its offsprings are encoded. In the second phase LSP's are processed and refinement bits for entries of LSP_k is encoded for $k = 0, \dots, K-1$. The coding contexts for encoding bits produced during LIN parsing and LSP

parsing can be found in [9].

3.3 Concluding Remarks

In this chapter, we reviewed highly scalable wavelet coefficient coders which will be utilized in MRMQ coding. We are mainly focused on three popular wavelet coefficient plane coding algorithms, namely, SPIHT, JPEG-2000 and EZBC since they offer superior R-D performance without sacrificing distortion and resolution scalability. Moreover, these important features are satisfied in a low complexity framework. Each of these coders have different advantages and disadvantages, which can be summarized as follows:

- The SPIHT algorithm yields superior R-D performance to that of EZW and R-D/embedding optimization techniques does not result in significant improvement of this performance. The only problem with SPIHT coder is that it needs a buffer size proportional to size of input image which might lead to non-local memory access problems. This problem is recently addressed in [27] and an efficient real time low memory version of SPIHT is proposed. Hence, its excellent R-D performance and availability of efficient hardware platform makes SPIHT a promising choice for MRMQ image coding.
- EZBC and JPEG-2000 encoders yields comparable R-D performance to other high complexity model based wavelet coders while preserving embeddedness and distortion scalability. EZBC offers precise rate control flexibility but its

bit stream is not necessarily R-D optimal and JPEG-2000 trades off rate control efficiency to R-D optimality. Real-time hardware implementations with low memory is possible and together with aforementioned properties, they are promising candidates for MRMQ image coding.

Chapter 4

Multi-region Multi-quality Image Coding

In this chapter, we present MRMQ coding with highly scalable SPIHT, JPEG-2000 and EZBC algorithms for medical image compression applications. The chapter is organized as follows: section 4.1 provides an overview of existing heterogeneous fidelity image coding techniques and briefly describes our approach. In section 4.2, we describe how region-based wavelet transform can be employed to efficiently represent the partitioned image plane. This is followed by MRMQ coding algorithms based on SPIHT, JPEG-2000 and EZBC encoders respectively. Finally, in section 4.6, we present the experimental results which verify the power of proposed MRMQ coding techniques.

4.1 Overview

In MRMQ coding, our objective is to encode the image by imposing heterogeneous (or region-dependent) fidelity constraints. Two ways to accomplish this are: 1) Object/region-based coding schemes studied under MPEG-4 standard, 2) ROI-based coding methods investigated under JPEG-2000 standard. In this section we briefly review these techniques and describe our approach.

4.1.1 Previous Approaches

Object/Region-based Image Coding

In this case, a region-based transform is typically used to obtain *independent* representations of each object/region, that is, each object/region is encoded *sequentially* as a single entity without using any information from other regions.

In MPEG-4, a padding method is used to extend an arbitrarily shaped region to a rectangular grid that is suitable for applying hybrid block-based DCT to obtain a region-based representation [11]. MPEG-4 scheme is computationally efficient for object based coding of video but the application of this method to region-based still image compression results in blocking artifacts which impedes high quality compression at low bit rates. Maximally decimated, region-based wavelet representations, described in Chapter 2, can be employed to overcome this problem and recently, they attracted much attention in object based coding. In [15], the EZW coder is modified to encode the region-based wavelet representations, in which the zerotree based coding

algorithm is adapted to encode trees with *don't care* nodes. Similar modifications are performed in [19] to incorporate region-based coding functionality to SPIHT coder. Recently, in [16], this modified SPIHT scheme is used for region-based chromosome image compression. The problem with these approaches stems from the fact that the decaying spectrum assumption of zerotree coders may contradict with the properties of the region-based wavelet transform. Specifically, a wavelet coefficient which is low-pass in nature might be moved to highpass band to preserve the phase alignment [13]. Consequently this results in deficiency in the coding performance, especially when regions of small size are encoded. This is also pointed out in [24], however the proposed solution requires bookkeeping of the decomposition strategy performed at the *decoder side*, which increases the complexity of the decoder substantially. Therefore, efficient encoding of region-based wavelet representations still deserves further investigation.

ROI-based Coding

In this case, the conventional transform is applied to the entire image plane without considering the region partitioning information. Typically, a local basis is used for image representation. Thus, the transform coefficients required for the reconstruction of a region can be differentiated in the transform domain and heterogeneous reconstruction fidelity can be achieved by appropriately partitioning the available bit rate among transform coefficients, i.e., allocating more bits to transform coefficients that are required for reconstruction of a high fidelity region (or ROI) and less bits for remaining coefficients (or background).

In [2], a simple scaling method is employed to describe an ROI at high fidelity using the SPIHT coder. The coefficients belonging to an ROI are scaled up to encode these coefficients prior to the ones belonging to background. Similar scaling method is used in JPEG-2000 [5]. In [21], a parent ROI mask is introduced in addition to an ROI mask to improve the compression performance of the coder described in [2]. The advantage of these schemes is that they allow *interactive* encoding possibility, that is, after sending some low resolution version of the entire image, the encoder and decoder may interact to modify the priorities of coefficients. On the other hand, the wavelet coders proposed for ROI coding are highly inferior to region-based coders in terms of R-D performance, due to the use of overcomplete representations in description of regions. Furthermore, precise measurement of region fidelities in transform domain is not possible with aforementioned ROI coders.

Thus, in this thesis, we will use region-based wavelet representations to accomplish heterogeneous fidelity encoding. In the following section, we present the proposed encoding scheme.

4.1.2 Our Approach

The proposed encoding mechanism is shown in Figure 4.1. The encoder is a hybrid version of region-based MPEG-4 and JPEG-2000 encoders. That is, initially, the image plane is demultiplexed into multiple regions, Ω_i 's, and separate representation, Λ_i of each region is obtained by using region-based wavelet transform, similar to the mechanism used in MPEG-4. However, rather than encoding the region representa-

tions sequentially, our encoder allows concurrent encoding of multiple regions as in JPEG-2000, with different distortion constraints, D_i 's. An important advantage of concurrent encoding is, if complete independence of region bit streams is not required, the R-D performance can be further improved by exploiting inter-region spatial dependencies. This feature imposes a moderate constraint on wavelet basis construction as will be described in the following section.

Figure 4.1: Proposed MRMQ Coding Mechanism

4.2 Wavelet Representation for Multiple AS-ROS

4.2.1 Constraints on Wavelet Basis Selection

In MRMQ coding setting, the image plane Ω is partitioned into N disjoint subsets, Ω_i $i = 0, \dots, N - 1$, as shown in Figure 1.1 or in Figure 1.2. The union of these subsets gives the entire image plane, Ω . During the encoding process, for each region, a separate distortion constraint is employed, i.e.,

$$\mathcal{D}^{(\Omega_i)} \leq \mathcal{D}_i, \quad 0 \leq i < N$$

where $\mathcal{D}^{(\Omega_i)}$ is the distortion due to lossy compression and $\mathcal{D}_i$ is the maximum allowable distortion. As mentioned in Chapter 3, we use MSE (or some related measure) to measure distortion. Each region is represented separately by region-based wavelet transforms. In other words, the region-based transform maps the image samples of each region Ω_i into a set of transform coefficients defined over Λ_i with $\Lambda_i \subseteq \Lambda$ where Λ is the wavelet coefficient plane corresponding to conventional wavelet representation of image plane Ω . In practice, it is desirable to measure distortion in the transform domain rather than the image domain. To accomplish this we should use unitary (or approximately unitary) wavelet basis which yields:

$$\mathcal{D}^{(\Omega_i)} \approx \mathcal{D}^{(\Lambda_i)}$$

where $\mathcal{D}^{(\Lambda_i)}$ is the distortion of region i measured in the transform domain. Therefore, to measure the region distortions precisely without additional computational cost, we use unitary (or approximately unitary) wavelet basis for representation.

Note that, although region-based transform provides a maximally-decimated representation of the region samples, it is possible that

$$\Lambda_i \cap \Lambda_j \neq \emptyset, \text{ for } i \neq j$$

In other words, critically sampled representations of completely separated regions does not guarantee complete separation in the transform domain. However, by choosing appropriate transforms, it is possible to make subset Λ_i 's mutually disjoint with the property

$$\bigcup_{i=0}^{N-1} \Lambda_i = \Lambda \quad (4.1)$$

The region-based transforms that satisfy this property have two important features. First, conventional transform coefficient coders are usually defined to encode a rectangular grid of coefficients and their modification to region-based coding usually starts with encapsulating the arbitrarily shaped subset Ω_i to a bounding rectangular frame which is followed by encoding of corresponding region representation, one region at a time. The restriction imposed by equation (4.1) precludes adaptation of coding mechanism to varying bounding frame sizes and thereby facilitates region-based coding via conventional transform coefficient plane coding mechanisms with small modifications only. Also, it provides flexibility in efficient interleaving of bit streams produced for different regions, without the need of a post-processing unit. Approximately unitary, symmetric, odd-length biorthogonal (OL-BO) wavelet basis selection provides the necessary framework to satisfy both of aforementioned constraints with comparable R-D performance to that of orthogonal or even length biorthogonal wavelet basis. Representation of multiple regions with OL-BO wavelet

basis is described in the following section.

4.2.2 Representation with Odd Length Biorthogonal Wavelet Basis

Consider the wavelet basis construction for representation of arbitrary shape region of support (AS-ROS) with symmetric extension. For OL-BO case, each basis vector is centered at different sample locations. In other words, each signal sample is associated with a basis vector. For constructing the basis, the convention is to use lowpass basis vectors centered at even-indexed sample locations and highpass basis vectors centered at odd indexed sample locations where the basis vectors associated with boundary samples can be found by symmetric extension. An alternative basis can be built by choosing highpass basis vectors centralized at even-indexed sample locations and lowpass basis vectors centered at odd-indexed sample locations. The alternative basis helps to obtain *region-based wavelet transforms* that satisfy the condition given in equation (4.1) simply by using 'conventional basis' for regions starting at even-indexed samples locations and 'alternative basis' for regions starting at odd-indexed samples.

This basis construction process is illustrated in Figure 4.2. First, the one-dimensional signal consisting of 16 samples is demultiplexed into 4 regions, denoted by Ω_k , $k = 0, \dots, 3$. A separate basis is used for representation of each region. The basis vectors (and the corresponding wavelet coefficients) are denoted by $LP_i^{(l)}$ and $HP_i^{(l)}$, where l is the label of the region for which the basis is constructed and

Figure 4.2: An example wavelet decomposition which satisfies equation 4.1.

i is the index of the sample upon which the basis vector is centered. For example, Ω_0 starts with even-index, so conventional basis is used for its representation, which is formed by choosing low-pass basis vectors centered at even-indexed samples, $\{LP_0^{(0)}, LP_2^{(0)}, LP_4^{(0)}\}$, and high-pass basis vectors at odd-indexed samples, $\{HP_1^{(0)}, HP_3^{(0)}\}$. However, Ω_1 starts with odd-indexed sample and alternative basis is used for its representation, as shown in Figure 4.2. Note that the condition imposed by equation (4.1) is satisfied by the proposed decomposition strategy.

We can represent multiple regions, which constitute the entire image plane, by the wavelet basis constructed by the aforementioned method. The resulting wavelet coefficients can be stored in a two dimensional array of size same as the conven-

tional wavelet representation. This property facilitates use of conventional wavelet coefficient plane coders for multiple, arbitrary-shape region coding.

4.3 MRMQ Coding based on SPIHT Algorithm

We first investigate MRMQ coding within the framework of SPIHT algorithm. Application of proposed methods to other zerotree based wavelet coders is straightforward. Hereafter, the definitions and notations developed in Section 3.2.1 will be used.

4.3.1 MRMQ-SPIHT with Sets of Descendent Labels

Since representations of each region constitute the conventional wavelet coefficient plane, we can construct the spatial orientation trees as in the original algorithm. However, in MRMQ setting, we may encounter coefficients from different regions within the same spatial orientation tree. Consider encoding of the transform coefficients pertaining to region l in this case. A straightforward implementation (such as one given in [2]) can be obtained by partitioning all spatial orientation trees as in the conventional case and during LIP and LSP parsing phases, transmitting the information about the wavelet coefficients of region l only. However we can save more bits, since there is no need to transmit the partitioning information for (sub-)trees which do not have any descendent nodes pointing to wavelet coefficients of region l . This is also pointed out in [21] and the authors defined a parent ROI mask to overcome the problem of increased redundancy. However, the proposed scheme may lead to loss of

details when critically sampled wavelet representations of arbitrary shape regions are being encoded.

To describe our approach, it is instructive to introduce the notion of *sets of descendent labels*, which is defined as:

$$\Upsilon(\mathbf{i}) = \bigcup_{\mathbf{k} \in \mathcal{O}(\mathbf{i})} \{v(\mathbf{k}) \cup \Upsilon(\mathbf{k})\} \quad (4.2)$$

where $v(\mathbf{i})$ denotes the label of region to which coefficient with index $\mathbf{i}$ belongs, i.e., if $\mathbf{i} \in \Lambda_l$, then $v(\mathbf{i}) = l$. Recall from SPIHT algorithm that, list of insignificant pixels (LIP) and list of insignificant sets (LIS) are initialized by using the set of wavelet coefficients within the lowest resolution subband, $\mathcal{H}$. However, when region l is to be encoded, these lists should be initialized by using

$$\mathcal{H}_l^{LIP} = \{\mathbf{i} \mid \mathbf{i} \in (\mathcal{H} \cap \Lambda_l)\} \quad (4.3)$$

$$\mathcal{H}_l^{LIS} = \{\mathbf{i} \mid l \in \Upsilon(\mathbf{i}), \mathbf{i} \in \mathcal{H}\} \quad (4.4)$$

where LIP is initialized by using the coordinates, $\mathbf{i} \in \mathcal{H}_l^{LIP}$ and LIS is initialized by the zerotree roots with indices, $\mathbf{i} \in \mathcal{H}_l^{LIS}$. Another modification to original SPIHT algorithm occurs at the set partitioning stage, i.e., only the zerotrees (represented by root nodes $\mathbf{i}$) for which $l \in \Upsilon(\mathbf{i})$ are partitioned and no bits are transmitted regarding the partitioning information of remaining trees. Details of the proposed approach are provided in Algorithm 4.1. An important feature of this scheme is that, it is relatively straightforward to extend the algorithm to encode multiple regions at a desired quality altogether, which might be beneficial if complete independence of region bitstreams is

not required and a context based arithmetic coder is used to exploit the dependencies among neighboring wavelet coefficients.

In the following section, we describe the region size sensitivity of the proposed MRMQ-SPIHT algorithm and a low-complexity solution is presented based on the concept of *unbalanced spatial orientation trees*.

1) **Modified Initialization:**
 Output $\tilde{m}^{(l)} = \lfloor \log_2 \left(\max_{i \in \Lambda_l} |w[i]| \right) \rfloor$ and initialize current bit plane level to be processed, namely, $m^{(l)}$ to $\tilde{m}^{(l)}$;
 Add the indices $i \in \mathcal{H}_l^{LIP}$ to LIP;
 Add the tree roots with indices $i \in \mathcal{H}_l^{LIS}$ to LIS;
 Initialize LIS to empty list.

2) **Modified Sorting Pass:**

2.1) for each entry i in the LIP do:
 2.1.1) output $\sigma^{(m^{(l)})}(i)$;
 2.1.2) if $\sigma^{(m^{(l)})}(i) = 1$, then move i to the LSP and output the sign of $w[i]$;

2.2) for each i in the LIS do:
 2.2.1) if entry is of type A, then
 \$ output $\sigma^{(m^{(l)})}(\mathcal{D}(i))$;
 \$ if $\sigma^{(m^{(l)})}(\mathcal{D}(i)) = 1$, then
 * for each $k \in \mathcal{O}(i)$ do:
 # if $v(k) = l$, then
 - output $\sigma^{(m^{(l)})}(k)$;
 - if $\sigma^{(m^{(l)})}(k) = 1$, then append k to the LSP and output the sign of $w[k]$;
 - if $\sigma^{(m^{(l)})}(k) = 0$, then append k to the LIP;
 * if $\mathcal{L}(i) \neq \emptyset$, then
 # if $\exists k \in \mathcal{O}(i)$ such that $l \in \Upsilon(k)$, then
 - move i to the end of the LIS, as an entry of type B, and go to step 2.2.2);
 - otherwise, remove entry i from the LIS;

2.2.2) if the entry is of type B, then
 \$ output $\sigma^{(m^{(l)})}(\mathcal{L}(i))$;
 \$ if $\sigma^{(m^{(l)})}(\mathcal{L}(i)) = 1$, then
 * for each $k \in \mathcal{O}(i)$ do:
 # if $l \in \Upsilon(k)$, then
 - add k to the end of the LIS as an entry of type A;
 * remove i from the LIS;

3) **Refinement Pass:** for each entry i in the LSP, except those included in the last sorting pass, output the $m^{(l)}$ th significant bit of $|w[i]|$;

4) **Quantization-Step Update:** decrement $m^{(l)}$ by 1 and go to step 2.

Algorithm 4.1: Encoding of the wavelet representation of region l by MRMQ-SPIHT with Set of Descendent Labels

4.3.2 MRMQ-SPIHT with Unbalanced Spatial Orientation Trees

To further improve the performance of MRMQ-SPIHT algorithm described above, we provide a solution based on the following facts:

1. When a region of relatively small size is encoded, the region's wavelet coefficients are localized over a very small portion of the exponentially growing spatial orientation quadtree. In this case, if we enable the use of nodes from lower levels of zerotrees, i.e., nodes other than roots of zerotrees, upon LIS initialization, we can save the redundant bits transmitted until the sub-tree comprising the wavelet coefficients of the region is reached from the root of the zerotree.
2. If we can establish homogeneity in tree structures, i.e., if all nodes of the tree have a common label, then we do not need bookkeeping of sets descendent labels during encoding. This reduces the computational cost of the MRMQ coding scheme.

To benefit from these facts, we introduce the concept of *unbalanced spatial orientation trees* as an alternative to conventional quadtree spatial orientation trees.

Assume that the spatial orientation trees are built as quadtrees as described in the original SPIHT algorithm. This construction results in mismatch between labels of some nodes and the labels of their parent nodes. To obtain homogeneity in the tree structure, the following tree reconstruction scheme can be used: If the label of a node in a tree does not agree with the label of its parent, the parent node is said to

Figure 4.3: Flowchart for redefinition of parent-child relationships in spatial orientation trees.

be pseudo-parent, and in this case, the node is disconnected from its pseudo-parent and a new (real) parent is sought for this node. The real parent is defined as one of the eight-connected neighbors of the pseudo-parent and has the same label as the corresponding node. If such a parent node is found, the node is connected to its real parent, otherwise the node itself is used as a zerotree root upon LIS initialization. The flowchart of the algorithm is given in Figure 4.3. The proposed tree structure moves the introduced complexity to the initialization stage, no longer requires bookkeeping of the sets of descendent labels and the comparisons regarding the sets of descendent labels. Furthermore, it performs well when regions of relatively small sizes are to be

encoded.

4.4 MRMQ Coding based on JPEG-2000 Algorithm

As described in Section 3.2.2, the encoding mechanism of JPEG-2000 consists of two tiers. In the first tier, the bit streams of each code block $\mathcal{B}_k[\mathbf{i}]$ is obtained by fractional bit plane coding and context-based arithmetic coding (CB-AC). The second tier employs PCRD optimization to interleave bit streams of each code block to achieve distortion and spatial scalability. MRMQ coding based on JPEG-2000 suggests modifications in these two tiers such that the algorithm should allow transmitting different amount of information for different regions. These modifications are discussed below.

4.4.1 Modified Tier 1

In tier 1, the subbands are initially partitioned into code blocks. Unlike the conventional case, in MRMQ setting, some of the code blocks encompass coefficients from different regions as shown in Figure 4.4, where in conventional case each coefficient has label p and in MRMQ coding the coefficients within the code block has either label p , q or l . Consider encoding of region l in this case. If independence of region bit streams is desired, we cannot exploit the information provided by neighboring regions to describe region l . Therefore, during fractional bit plane passes, the encoder should

only process the coefficients of region l whereby skipping the remaining coefficients within the code block.

p	p	p	p
p	p	p	p
p	p	p	p
p	p	p	p

(a) Conventional

q	q	p	p
q	p	p	l
l	p	l	l
l	l	l	l

(b) MRMQ

Figure 4.4: Comparison of labels of code block samples in conventional JPEG-2000 coding and MRMQ coding.

The Modified Forward Significance Propagation Pass is described in Algorithm 4.2. In the algorithm, $\mathcal{N}^{(l)}[\mathbf{n}]$ denotes the set of eight-connected neighbors of coefficient $\mathbf{n}$ which lie within Λ_l . Similarly in the Magnitude Refinement Pass, only the refinements bits of $\mathbf{n} \in \Lambda_l$ are processed by the CB-AC engine.

```

 $\forall \mathbf{n} \in \mathcal{B}_k[i]$ 
  if  $\mathbf{n} \in \Lambda_l$ 
    if  $\sigma^{(m+1)}(\mathbf{n}) = 0$  and  $\sigma^{(m)}(\mathcal{N}^{(l)}\{\mathbf{n}\}) = 1$  then
      Code  $\sigma^{(m)}(\mathbf{n})$  with CB-AC engine;
    else
      Skip coefficient;
    endif;
  endif.

```

Algorithm 4.2: Modified Forward Significance Propagation Pass.

The main distinction between Normalization Pass and Forward Significance Pass is that in Normalization pass, the RLM can be invoked instead of ZCM when a run of zeros is encountered. However this suggests processing of coefficients in groups of four rather than individually. In MRMQ framework, this increases implementation complexity due to multi-region accumulation within the code block. Therefore, we do not use RLM when heterogeneous code blocks are encoded. In other words, if a code block consists of coefficients from different regions, during Modified Normalization Pass, all bits are encoded by ZCM even if a run of insignificant coefficients are encountered. For other code blocks, the conventional bit stream formation algorithm is employed.

In addition to bit stream formation, in Tier 1, we collect R-D summary information which will be used by PCRD optimization module. In MRMQ, we use approximately unitary wavelet basis which allows measurement of region distortions in transform domain. Unlike conventional case, for heterogeneous code blocks, we record the R-D slopes of each region separately. In the next section, the interleaving of bit streams based on this summary information will be discussed.

4.4.2 Modified Tier 2

For multi-region blocks, tier 1 produces a separate bit stream for each region within the code block. Let $\{\eta[\mathbf{i}](t)\}_{0 \leq t < T}$ be the set of *feasible* truncation points for region l and block $\mathbf{i}$, obtained by convex hull analysis. The R-D slope between two adjacent feasible truncation points is denoted by

$$\left(\frac{\delta D}{\delta R}\right)(\eta[\mathbf{i}](t), \eta[\mathbf{i}](t+1)) \quad (4.5)$$

For optimal embedding, we need to order these R-D slopes in decreasing order. This can be visualized as a sorted table, in which each successive row provides information about the region and block index, and corresponding truncation data which exhibits maximum R-D slope. An example is given in Table 4.4.2. In the table, the information in rows is ordered according to R-D slopes.

Processing Order	Region Label	Code Block Index	Number of Additional Passes	Number of Bytes Appended	$\left(\frac{\delta D}{\delta R}\right)$
1	l	(0,0)	11	417	0.796
2	l	(0,1)	3	119	0.701
3	k	(0,0)	7	341	0.667
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$

Table 4.1: Sorted table which contains the contribution order of code block bit streams.

Once this table is obtained by simple sorting, the distortion scalable interleaving of code block bit streams can be achieved by employing Algorithm 4.3. In the

algorithm, the set comparison $(a, b, \dots) > (m, n, \dots)$ returns 1 if $(a > m)$ and $(b > n)$, and so on. Otherwise, it returns zero. Upon final bit stream formation, we use small packet headers to uniquely identify the amount of contribution of each code block as in original JPEG-2000 algorithm.

```

r = 0;
while ( $\mathcal{D}^{(\Lambda_0)}, \mathcal{D}^{(\Lambda_1)}, \dots, \mathcal{D}^{(\Lambda_{N-1})} > (\mathcal{D}_0, \mathcal{D}_1, \dots, \mathcal{D}_{N-1})$ )
    Input  $r^{th}$  row of sorted table. Let  $i_r$  be the retrieved block index, and  $l_r$  be the retrieved region label and  $\delta b$  be the number of
    bytes to be appended;
    if  $\mathcal{D}^{(\Lambda_{l_r})} > \mathcal{D}_{l_r}$ 
        • Append header which identifies the retrieved information at the decoder;

        • Append  $\delta b$  bytes to the final bit stream from the bit stream of region  $l_r$  coming from code block  $i_r$ ;

    endif
    r ← r + 1
endwhile

```

Algorithm 4.3: Algorithm for R-D optimal interleaving of region bit streams.

Accordinging the above formulation, bit streams of different regions are interleaved according to embedding and resolution. The image samples corresponding to different regions, resolutions and rates can be extracted from the produced bit stream simply by skipping the unrelated information.

4.5 MRMQ Coding based on EZBC Algorithm

Similar to JPEG-2000, EZBC uses CB-AC to encode the wavelet coefficient plane. However, to reduce the complexity due to adaptive arithmetic coding, it employs quadtree coding to locate the clusters of significant coefficients and uses CB-AC to encode information about these coefficients. Its advantage over JPEG-2000 is the

precise rate control flexibility, i.e., there is no need for a post-compression processing unit to organize the bit stream into a distortion scalable one. Similar to SPIHT algorithm, MRMQ coding with EZBC can be accomplished either by using sets of descendent labels or unbalanced quadtrees. In the following sections, we will describe MRMQ coding within EZBC framework.

4.5.1 MRMQ-EZBC with Sets of Descendent Labels

The main distinction between the spatial orientation trees and the quadtree structure is that the spatial orientation trees are *fed from all nodes*, i.e., all nodes in the tree has one to one correspondence with a wavelet coefficient, but the quadtree structure is solely *fed from the leaf nodes*. Therefore, only the leaf nodes of the quadtree have corresponding labels. Let $v_k(\mathbf{i})$ be the label of the wavelet coefficients associated with $QT_k[0](\mathbf{i})$. The set of descendent labels associated with node $QT_k[l](\mathbf{i})$ is defined as:

$$\Upsilon_k[l](\mathbf{i}) = \begin{cases} \emptyset & \text{if } l = 0; \\ \bigcup_{\mathbf{n} \in \mathcal{O}(\mathbf{i})} v_k(\mathbf{n}) & \text{if } l = 1; \\ \bigcup_{\mathbf{n} \in \mathcal{O}(\mathbf{i})} \Upsilon_k[l-1](\mathbf{n}) & \text{otherwise.} \end{cases} \quad (4.6)$$

where $\mathcal{O}(\mathbf{i})$ denotes the direct descendent(offspring or children) nodes spreading from $QT_k[l](\mathbf{i})$. When the wavelet representation of region with label p is being encoded, only the information of descendents of nodes, $QT_k[l](\mathbf{i})$, for which $p \in \Upsilon_k[l](\mathbf{i})$ are encoded. Encoding of region p with modified EZBC algorithm is presented in Algorithm 4.4. The details of the modified LIN parsing procedure is given in Algorithm 4.5.

1. **Modified Initialization:**

- Output $\tilde{m}^{(p)} = \lfloor \log_2 \left(\max_{\mathbf{n} \in \Lambda_p} |w[\mathbf{n}]| \right) \rfloor$ and initialize current bit plane level to be processed, namely, $m^{(p)}$ to $\tilde{m}^{(p)}$;
- Initialize LIN's according to

$$LIN_k[l] = \begin{cases} \{(0, 0)\} & \text{if } l = D_k; \\ \emptyset & \text{otherwise.} \end{cases} \quad (4.7)$$
- Initialize LSP's to empty set.

2. **Modified LIN Parsing:**

$\forall l \in \mathcal{Z}_{D_{max}}$

$\forall k \in \mathcal{Z}_K$

MOD_ParseLIN(k,l);

3. **LSP Parsing**

$\forall k \in \mathcal{Z}_K$

For each entry $\mathbf{n}$ in the LSP_k , except those included in the last sorting pass, output the $m^{(p)}$ th significant bit of $|w[\mathbf{n}]|$;

4. **Quantization Step Update:** Decrement $m^{(p)}$ by 1 and go to step 2.

Algorithm 4.4: Main algorithm for MRMQ coding with EZBC encoder.

```

MOD_ParseLIN(k,l){
    •  $\forall \mathbf{n} \in LIN_k[l]$ 
        - Code  $\sigma^{(m^{(p)})}(\mathbf{n})$ ;
        - if  $\sigma^{(m^{(p)})}(\mathbf{n}) = 0$ , then node  $\mathbf{n}$  remains in  $LIN_k[l]$ ;
        - else
            * if  $(l = 0)$ ,
                · if  $\mathbf{n} \in \Lambda_p$ , then code sign bit of  $w[\mathbf{n}]$  and add  $\mathbf{n}$  to  $LSP_k$ ;
                · else skip this node;
            * else  $CodeDescendents(k, l - 1, \mathbf{n})$ 
    }

CodeDescendents(k, l,  $\mathbf{n}$ ){
    •  $\forall \mathbf{k} \in \mathcal{O}[\mathbf{n}]$ 
        - if  $p \in \Upsilon(\mathbf{k})$  or if  $(l = 1)$  and  $v_k(\mathbf{k}) = p$ 
            * Code  $\sigma^{(m^{(p)})}(\mathbf{k})$ ;
            * If  $\sigma^{(m^{(p)})}(\mathbf{k}) = 0$ , add  $\mathbf{k}$  to  $LIN_l[l - 1]$ ;
            * else
                · if  $l = 1$ 
                    1. if  $v_k(\mathbf{k}) = p$ , code sign bit of  $w[\mathbf{n}]$  and add  $\mathbf{k}$  to  $LSP_k$ ;
                · else  $CodeDescendents(k, l - 1, \mathbf{k})$ 
    }

```

Algorithm 4.5: Pseudocode for *MOD_ParseLIN* procedure introduced in Algorithm 4.4.

Similar to MRMQ-SPIHT with Sets of Descendent Label and JPEG-2000, this algorithm is not very efficient when regions of small size are being encoded. A computationally efficient solution to this problem is proposed by using *unbalanced quadtrees*.

4.5.2 MRMQ-EZBC with Unbalanced Quadrees

The region size sensitivity problem with EZBC algorithm is more severe than that of SPIHT, since the quadrees comprise the entire set of subband coefficients where the exponential growth of spatial orientation trees is limited to the number of levels used in wavelet decomposition. The unbalanced tree structure introduced in Section 4.3.1 can be used to overcome this problem but this definition requires modification since for the quadtree structure, a node at the higher levels of the tree does not have one-to-one correspondence to a wavelet coefficient so it does not have a *unique* label. The construction of proposed unbalanced quadtree structure is as follows:

We start with assigning labels to the nodes at the higher levels of the tree, which is defined as the most common label encountered in the offsprings of the node, i.e., if most of the offspring nodes have label m , then the label of the node is m . If the label of the node does not agree with its parent then this node is disconnected from its (pseudo-)parent and a new (real-)parent is sought for this node. The real parent is defined as one of the neighboring (eight) nodes of the pseudo-parent and has the same label as the corresponding node. If such a parent is found, the node is connected to its real parent, otherwise the node itself is used in LIN initialization. This tree definition guarantees homogeneity of labels within the tree structure, but the trees are no longer balanced. The proposed scheme moves the introduced complexity to initialization stage, no longer requires bookkeeping of sets of descendent labels and the compare operations about the sets of descendent labels, and is very efficient when regions of relatively small size are to be encoded.

4.6 Performance Evaluation of MRMQ Coders

In this section, the performance of proposed MRMQ coders is discussed. For brevity, MRMQ-SPIHT with Sets of Descendent Labels will be referred to as MRMQ-SPIHTa, MRMQ-SPIHT with Unbalanced Spatial Orientation Trees as MRMQ-SPIHTb, MRMQ coding based on JPEG-2000 as MRMQ-J2000, MRMQ-EZBC with Sets of Descendent Labels as MRMQ-EZBCa, and MRMQ-EZBC with Unbalanced Quadrees as MRMQ-EZBCb.

4.6.1 Prototype Basis Vector Selection

Recall that, the wavelet basis vectors are even-shifted versions of two prototype basis vectors, namely, the prototype *lowpass* basis vector ($\mathbf{a}_0$) and prototype *highpass* basis vector ($\mathbf{a}_1$). In MRMQ coding, to construct the wavelet basis for the partitioned image, we use symmetric, odd-length prototype basis vectors. We further constrain the basis to be (approximately) unitary to enable precise distortion measurement in the transform domain. The 9/7, 13/11, and 17/11 basis[†], which offer superior R-D performance in conventional image plane coding [3] and satisfy the aforementioned partitioned image representation constraints, are promising for MRMQ coding applications.

In Table 4.2 and Table 4.3, the R-D performance of 9/7, 13/11 and 17/11 wavelet

[†]The lowpass decomposition filter due to N/M wavelet basis has N nonzero samples, and correspondingly the highpass filter has M nonzero samples.

Region PSNR(dB.)	Bit rate(bpp.)		
	17/11	13/11	9/7
20	0.0017	0.0037	0.0017
25	0.0050	0.0111	0.0056
27	0.0087	0.0160	0.0093
29	0.0151	0.0263	0.0152
31	0.0260	0.0414	0.0264
33	0.0462	0.0640	0.0460
35	0.0815	0.1047	0.0806
40	0.2633	0.2956	0.2604
45	0.7102	0.7668	0.7261
50	1.5211	1.6000	1.5499
55	3.4218	3.4644	3.4368

Table 4.2: Comparison of basis performances for encoding left lung region of chest radiograph shown in Figure 1.1 with MRMQ-SPIHTa.

basis is compared. The basis selection which achieves the required region fidelity at the lowest bit rate is printed in boldface. The results can be summarized as follows:

- Similar to the conventional case, the 9/7 and 17/11 bases perform consistently better than 13/11 basis, but the difference is at most 0.05 bits per pixel (bpp.).
- The 9/7 basis and 17/11 basis perform almost equally well. This means, the 9/7 basis achieves the decorrelation obtained by 17/11 basis with significantly

Region PSNR(dB.)	Bit rate(bpp.)		
	17/11	13/11	9/7
20	0.0015	0.0037	0.0021
25	0.0040	0.0087	0.0051
27	0.0076	0.0181	0.0084
29	0.0276	0.0446	0.0268
31	0.0892	0.1144	0.0866
33	0.2039	0.2347	0.2054
35	0.3874	0.4283	0.3933
40	1.1526	1.2083	1.1717
45	2.0986	2.1506	2.1183
50	3.0852	3.1211	3.0894
55	4.4979	4.5381	4.5144

Table 4.3: Comparison of basis performances for encoding microcalcification cluster region of breast mammogram shown in Figure 1.2 with MRMQ-J2000.

reduced computational complexity compared to that of 17/11.

Consequently, the popular 9/7 basis [36] is also promising for MRMQ coding applications. Thus, in the comparisons presented in following sections, we use 9/7 basis for multi-region representation.

4.6.2 Comparison of Region Size Sensitivity

In this section, we compare R-D performance of MRMQ-SPIHTa with MRMQ-SPIHTb and MRMQ-EZBCa with MRMQ-EZBCb with respect to varying region sizes. The 512-by-512 *Angio* image, shown in Figure 4.5 is used for comparison. The

Figure 4.5: The *Angio* image used in region size sensitivity comparison and the corresponding circular region selection with radius R .

arbitrary shape region is selected as a circular region centered at center pixel of the *Angio* image and the objective PSNR for this region is fixed as 40 dB. Then, corresponding bit rates achieving this region PSNR are measured. The comparison of MRMQ-SPIHTa and MRMQ-SPIHTb is shown in Figure 4.6. As can be seen from the figure, MRMQ-SPIHTb outperforms MRMQ-SPIHTa for circular region sizes up

Figure 4.6: Region size sensitivity of proposed MRMQ-SPIHTa and MRMQ-SPIHTb: Comparison of bit rate required to attain 40 dB. ROI-PSNR for circular regions of radius R .

to 10 pixel radius, and after this point, the two methods perform equally well.

Similarly, the comparison of MRMQ-EZBCa with MRMQ-EZBCb is shown in Figure 4.7. As can be seen from the figure, MRMQ-EZBCb outperforms MRMQ-EZBCa for circular region sizes up to 20 pixel radius, and similar to MRMQ-SPIHT case, after this point, the methods perform equally well.

The results show that the unbalanced tree structures outperform the balanced ones when images partitioned into regions of relatively small size are to be encoded.

Figure 4.7: Region size sensitivity of proposed MRMQ-EZBCa and MRMQ-EZBCb: Comparison of bit rate required to attain 40 dB. ROI-PSNR for circular regions of radius R .

4.6.3 Evaluation of R-D Performance

In this section, we compare the R-D performance of the proposed MRMQ coders. We use the chest radiograph shown in Figure 1.1 for comparison. The radiograph is partitioned into 4 regions, namely, the left lung, right lung, mediastinum and background as depicted in the figure.

First consider encoding of the image as a single entity, i.e., without considering the region partitioning. The R-D performance comparison in this case, for SPIHT,

Figure 4.8: R-D performance comparison of conventional wavelet coefficient plane coders SPIHT, JPEG-2000, and EZBC for the chest radiograph.

JPEG-2000 and EZBC coders is shown in Figure 4.8. As can be seen from the figure, all encoders perform almost the same where JPEG-2000 has a very slight advantage over EZBC and SPIHT encoders, both of which perform equally well. Consider encoding the partitioned image plane with MRMQ versions of aforementioned conventional coders. In this case, the R-D curves for each region, Ω_i , $i \in \mathcal{Z}_4$, is shown in Figure 4.9. Similar to the conventional case, MRMQ-J2000 offers slightly better R-D performance compared to that of MRMQ-SPIHT and MRMQ-EZBC coders.

Figure 4.9: R-D curves for regions Ω_i , $i \in \mathbb{Z}_4$ shown in Figure 1.1.

An important property in evaluating the efficiency of a region-based coder is its *transparency* to the corresponding conventional coder. In other words, if an image is coded at a given fidelity by encoding each region independently at that fidelity, the visual quality of the resulting description should be comparable to encoding of the image at that quality without considering the region partitioning. The property is illustrated in Figure 4.10. In the figure, the chest radiograph is encoded at 0.5 bpp. first by the conventional JPEG-2000 algorithm and then by the MRMQ-J2000 coder where each of four regions, Ω_i , $i \in \mathcal{Z}_4$ is encoded independently at the same quality with overall bit rate 0.5 bpp. As can be seen from the figure, the MRMQ coder does not introduce any annoying artifacts due to use of independent region-based representations and is transparent to JPEG-2000.

It is important to mention that, for medical image coding applications, the proposed MRMQ coding schemes utilize the available bit rate better than the conventional coders. In other words, unlike conventional coders, the MRMQ coders allow encoding of diagnostically significant regions with best possible fidelity while allocating less bits to remaining regions. This strategy will be referred as *pseudo-lossless coding*. The advantage of pseudo-lossless coding is illustrated for the breast mammogram shown in Figure 4.11. In Figure 4.12, the wavelet coefficient plane of breast mammogram is encoded with conventional JPEG-2000 algorithm at 0.6 bpp. This corresponds to a fidelity of 45.7 dB. PSNR for the breast region and 46.1 dB. PSNR

JPEG-2000 at 0.5 bpp.

MRMQ-J2000 at 0.5 bpp.

Figure 4.10: Illustration of transparency of MRMQ-J2000 to conventional JPEG-2000 coder

Figure 4.11: Original breast mammogram image

Figure 4.12: Conventional (JPEG-2000) coding of mammogram image at 0.6 bpp. In this case 16% of the bit rate is allocated to encode the background and 84% to encode the breast region.

Figure 4.13: MRMQ (MRMQ-J2000) coding of mammogram image at 0.6 bpp. In this case 100% of the bit rate is used to encode breast region. Diagnostically significant breast region encoded at 4 dB. better PSNR value than that of conventional coding shown in Figure 4.12.

for the background region. In Figure 4.13, the image is encoded by assigning all available bit rate to representation of breast region only. In this case, the breast region is encoded at 50 dB. PSNR and no bits are consumed due to encoding of the irrelevant information contained in the background region. This example illustrates the importance of MRMQ coding for medical image compression applications, in which the diagnostically significant region is typically localized over relatively small regions.

4.6.4 MRMQ vs. ROI Coding

In this section, the performance of MRMQ-SPIHT coder is compared with the ROI-SPIHT in [21] and MRMQ-J2000 coder with the ROI coding unit of JPEG-2000 [6]. Recall that, aforementioned ROI coders do not use completely independent representation for each region and the discrimination of ROI/background qualities is achieved by allocating more bits to wavelet coefficients used for the reconstruction of ROI, and less bits for the remaining coefficients.

First consider ROI-SPIHT of Park et. al. [21]. This coder is taken as reference since it outperforms all previously proposed SPIHT ROI-coding schemes in terms of compression efficiency. The 256-by-256 *Lena* image with 51-by-51 ROI and 512-by-512 *Baboon* image with 101-by-151 ROI are used for comparison. The PSNR vs. bit rate values of ROI-SPIHT are taken from [21] and the bit rates that achieve the same ROI/background qualities is measured for MRMQ-SPIHTa. The ROI-PSNR vs. bit rate curves is provided in Figure 4.14.(a). In terms of ROI-PSNR, MRMQ-SPIHTa performs 10 dB.(/7 dB.) better than ROI-SPIHT when ROI of *Lena*(/*Baboon*) image is encoded. This is due to use of maximally decimated wavelet representation for the ROI rather than use of wavelet coefficients required for reconstruction of ROI, since the number of wavelet coefficients required for reconstruction of the ROI is usually far greater than number of samples within the ROI. Hence, according to Figure 4.14.(b), it is possible to achieve 0.2(/0.15) bpp. gain in the overall bit rate to attain the same ROI/background quality even though the area of ROI is only 4%(/5.8%) of the image area.

Figure 4.14: R-D performance comparison of ROI-SPIHT [21] and MRMQ-SPIHTa.

(a) ROI-PSNR vs. bit rate performance comparison for Lena and Baboon images.

(b) Overall PSNR vs. bit rate performance comparison for Lena and Baboon images.

The R-D performance of MRMQ-J2000 is compared with the ROI encoder in JPEG-2000. The 720-by-576 *Goldhill* image with 320-by-200 ROI is used for comparison. JPEG-2000 does not provide a precise control over the ROI/background distortions and the discrimination of ROI/background qualities is achieved by simply down-scaling wavelet coefficients pertaining to background. The ROI coding performance of JPEG-2000, given in Figure 4.15, is taken from [6] for scale = 11, and the performance of MRMQ-J2000 is evaluated by measuring the bit rate that achieves same ROI/background qualities observed in JPEG-2000. In terms of ROI-PSNR, MRMQ-J2000 performs 4 dB better than JPEG-2000, on average, when ROI

of Goldhill image is encoded. Thus, from Figure 4.15, on average, MRMQ-J2000 provides a 0.3 bpp. gain in the overall bit rate to attain the same ROI/background quality when the ROI area is 15% of the overall image area.

Figure 4.15: R-D performance comparison of ROI coding unit of JPEG-2000 and MRMQ-J2000. (a) ROI-PSNR vs. bit rate performance comparison for Goldhill image. (b) Background-PSNR vs. bit rate performance comparison for Goldhill image.

4.6.5 Summary

The experimental results can be summarized as follows:

- When regions of relatively small size are being encoded, the MRMQ coders

based on unbalanced tree structures, namely MRMQ-SPIHTb and MRMQ-EZBCb significantly outperform the MRMQ coders based on sets of descendent labels, namely MRMQ-SPIHTa and MRMQ-EZBCa. This verifies the power of unbalanced tree structures when regions of relatively small size are to be encoded.

- The R-D performance of MRMQ coders exhibit similar performance to corresponding conventional coefficient plane coders. That is, similar to conventional case, MRMQ-coder based on JPEG-2000 algorithm slightly outperforms MRMQ-coders based on EZBC algorithm, which offers slightly better R-D performance than MRMQ-coders based on SPIHT. This means, the choice of appropriate MRMQ coder for a given application should be based on requirements other than R-D performance, such as spatial scalability, complexity, precise rate control flexibility, and/or region size sensitivity.
- The proposed MRMQ-coders are transparent to their conventional versions. In other words, encoding each region independently with MRMQ coders does not introduce annoying artifacts at the boundaries and if the regions in the image are encoded at a given fixed fidelity, the resulting compressed image exhibits similar visual (subjective) quality to one obtained by conventional coefficient plane coding.
- The MRMQ coders offer significant R-D advantage to ROI coding schemes. This is due to use of maximally decimated region-based wavelet transforms for region representation. For example, it is possible to save 0.3 bpp. with the

MRMQ-J2000 coder, while encoding a specified ROI/background at the same fidelity as the ROI coding unit of JPEG-2000.

Chapter 5

Conclusion

5.1 Summary

In most medical imaging applications, an image can be either partitioned into several anatomic regions that vary in terms of diagnostic significance or the diagnostically significant information is localized over a relatively small size regions of interest. Conventional image plane coders cannot exploit this property due to use of a single fidelity constraint for describing the entire image. Therefore, we focused on heterogeneous fidelity coding of medical images, similar to emerging region-based or ROI-based coding approaches.

The region-based wavelet transform with symmetric OL-BO wavelet basis is used to represent each region independently. Other approaches for wavelet representation of arbitrary shape regions is based on optimal boundary basis construction.

However, the experiments show that the optimal boundary filters do not offer better R-D performance than that of OL-BO wavelet basis, but increase the system complexity significantly.

The proposed region-based decomposition strategy produces a partitioned wavelet coefficient plane where coefficients corresponding to different regions are completely separated. The partitioned coefficient plane is encoded with proposed MRMQ coders, i.e., by assigning various fidelity constraints for each region. Our MRMQ coders are based on the state of the art conventional wavelet coefficient plane coders SPIHT, JPEG-2000 and EZBC. The advantage of these coders is that their real-time implementation with low complexity is available, and they offer comparable R-D performance to that of highly complex wavelet coders reported in the literature while preserving distortion and spatial scalability.

The most important advantage of MRMQ coding over ROI-based coding is the significant improvement in R-D performance. Moreover, the proposed coders have the flexibility of encoding coefficients of different regions at a given fidelity altogether, which enables producing a distortion scalable bit stream with efficient interleaving of contributions of different regions.

Finally, we developed a novel scheme to address the problem of region size sensitivity observed in tree-based coding of wavelet representations of arbitrary shape regions. Experimental results verify the power of the proposed tree structure, when the image is partitioned into regions of relatively small size.

5.2 Future Work

- Considering the scalability properties offered by proposed MRMQ coders, a potential research topic can be the evaluation of their efficiency for communication environments with various bandwidth requirements and different computational capabilities, i.e., for communication over heterogeneous networks including wireless networks.
- In this thesis, we were mainly focused on source coding of medical images without considering whether the communication channel is error prone or not. An emerging joint source-channel coding method is prioritized bit stream protection, i.e., more important bits in the bit stream are protected more than the remaining information. Proposed MRMQ coders order the bit stream according to importance of bits, and therefore are suitable for prioritized bit stream protection. A potential research topic is the incorporation of prioritized bit stream protection techniques to MRMQ bit stream formation module and evaluation of the performance of resulting coder for communications over highly error prone channels.
- Our performance evaluation of MRMQ-coders is mainly based on the objective quality measure, PSNR. In this case, the bit streams are organized in such a way that the initial bits offer greater reduction in MSE than the remaining. However, it is widely recognized that a MSE in smooth areas of the image can result in much lower *visual* quality than same MSE introduced in textured

image areas [22]. In other words, organization of bit stream according to MSE does not necessarily yield ordering of information from higher visual quality to lower visual quality. Therefore, a higher-performance (in terms of subjective visual quality) coder can be obtained by using a subjective quality measure for better bit allocation. Incorporation of subjective quality measures to proposed MRMQ coders remains as a future research topic.

- We developed encoding schemes for coding the partitioned image plane efficiently. According to our implementation, a medical expert can manually partition the medical image into regions with different diagnostic properties and then assign various fidelity constraints to these regions to compress the image. The incorporation of automatic region extraction algorithms to the compression system and evaluation of the resulting picture archiving and communication system for telemedicine applications is another potential research topic.

Bibliography

- [1] J. G. Apostolopoulos and J. S. Lim. Critically sampled wavelet representations for multidimensional signals with arbitrary regions of support. In *IEEE International Conference on Acoustics, Speech and Signal Processing*, volume 3, pages 1509–1512, 1998.
- [2] E. Atsumi and N. Farvardin. Lossy/lossless region-of-interest image coding based on set partitioning in hierarchical trees. In *IEEE International Conference on Image Processing*, volume 1, pages 87–91, 1998.
- [3] S. Basu and B. Levy, editors. *Multidimensional Filter banks and Wavelets Research Developments and Applications*, chapter Low Bit-Rate Design Considerations for Wavelet-Based Image Coding. Kluwer Academic Publishers, 1997.
- [4] C. M. Brislawn. Preservation of subband symmetry in multirate signal coding. *IEEE Transactions on Signal Processing*, 43(12):3046–3050, December 1995.
- [5] C. Christopoulos, J. Askelf, and M. Larsson. Efficient methods for encoding regions of interest in the upcoming JPEG2000 standard. *IEEE Signal Processing*

Letters, 78:247–249, September 2000.

- [6] R. Grosbois, D. Santa-Cruz, and T. Ebrahimi. New approach to JPEG 2000 compliant region of interest coding. In *SPIE's 45th annual meeting, Applications of Digital Image Processing XXIV*, volume 4472 of *Proc. of SPIE*, pages 267–275, San Diego, California, August 2001.
- [7] C. Herley. Boundary filters for finite-length signals and time-varying filter banks. *IEEE Transactions on Circuits and Systems–II: Analog and Digital Signal Processing*, 42(2):102–114, February 1995.
- [8] S. Hsiang. Embedded image coding using zeroblocks of subband/wavelet coefficients and context modeling. In *Data Compression Conference*, pages 83–92, 2001.
- [9] S. Hsiang. *Highly Scalable Subband/Wavelet Image and Video Coding*. PhD thesis, Rensselaer Polytechnic Institute, May 2002.
- [10] H. Katata, N. Ito, T. Aono, and H. Kusao. Object wavelet transform for coding of arbitrary shaped image segments. *IEEE Transactions on Circuits and Systems for Video Technology*, 7(1):234–237, February 1997.
- [11] A. Kaup. Object-based texture coding for moving video in MPEG-4. *IEEE Transactions on Circuits and Systems for Video Technology*, 9(1):5–15, February 1999.

- [12] J. Kovacevic and M. Vetterli. Nonseparable two- and three-dimensional wavelets. *IEEE Transactions on Signal Processing*, 43(5):1269–1273, May 1995.
- [13] J. Li and S. Lei. Arbitrary shape wavelet transform with phase alignment. In *IEEE International Conference on Image Processing*, volume 3, pages 683–687, 1998.
- [14] J. Li and S. Lei. An embedded still image coder with rate-distortion optimization. *IEEE Transactions on Image Processing*, 8(7):913–924, July 1999.
- [15] S. Li and W. Lei. Shape-adaptive discrete wavelet transforms for arbitrarily shaped visual object coding. *IEEE Transactions on Circuits and Systems for Video Technology*, 10(5):725–743, August 2000.
- [16] Z. Liu, Z. Xiong, Q. Wu, Y. Wang, and K. Castleman. Cascaded differential and wavelet compression of chromosome images. *IEEE Transactions on Biomedical Engineering*, 49(4):372–382, April 2002.
- [17] S. M. LoPresto, K. Ramchandran, and M. T. Orchard. Image coding based on mixture modeling of wavelet coefficients and a fast estimation-quantization framework. In *Data Compression Conference*, pages 241–250, 1997.
- [18] A. Mertins. Boundary filter optimization for segmentation-based subband coding. *IEEE Transactions on Signal Processing*, 49(8):1718–1727, August 2001.

- [19] G. Minami, Z. Xiong, A. Wang, and S. Mehtora. 3-d wavelet coding of video with arbitrary regions of support. *IEEE Transactions on Circuits and Systems for Video Technology*, 11(9):1063–1068, September 2001.
- [20] E. Ordentlich, M. Weinberger, and Gadiel Seroussi. A low-complexity modeling approach for embedded coding of wavelet coefficients. In *Data Compression Conference*, pages 408–417, 1997.
- [21] K. Park and H. W. Park. Region-of-interest coding based on set partitioning in hierarchical trees. *IEEE Transactions on Circuits and Systems for Video Technology*, 12(2):106–113, February 2002.
- [22] M. Ramos and S. Hemami. Activity selective SPIHT coding. In *SPIE Conf. on Visual Communications and Image Processing, San Jose, CA*, volume 3653, pages 315–326, January 1999.
- [23] A. Said and W. A. Pearlman. A new, fast, and efficient image codec based on set partitioning in hierarchical trees. *IEEE Transactions on Circuits and Systems for Video Technology*, 6(3):243–249, June 1996.
- [24] H. Schwarz and E. Muller. Object-based 3-D wavelet coding using layered object representation. In *IEEE International Conference on Image Processing*, volume 1, pages 1000–1003, 2000.
- [25] J. M. Shapiro. Embedded image coding using zerotrees of wavelet coefficients. *IEEE Transactions on Signal Processing*, 41(12):3445–3462, December 1993.

- [26] Y. Shoham and A. Gersho. Efficient bit allocation for an arbitrary set of quantizers. *IEEE Transactions on Acoustics, Speech and Signal Processing*, pages 1445–1453, September 1988.
- [27] Y. Sun, H. Zhang, and G. Hu. Real-time implementation of a new low-memory SPIHT image coding algorithm using DSP chip. *IEEE Transactions on Image Processing*, 11(9):1112–1116, September 2002.
- [28] W. Sweldens. A construction of second generation wavelets. *SIAM J. Math. Anal.*, 29(2):511–546, 1997.
- [29] S. Tasdoken and A. Cuhadar. ROI coding with region-based integer wavelet transforms and unbalanced spatial orientation trees. In *SPIE Conference on Image and Video Communications and Processing*, Proc. of SPIE, Santa Clara, California, January 2003.
- [30] S. Tasdoken and A. Cuhadar (submitted). Quadtree-based multi-region multi-quality image coding. *IEEE Signal Processing Letters*, December 2002.
- [31] S. Tasdoken and A. Cuhadar (submitted). Multiple, arbitrary shape ROI coding with zerotree based wavelet coders. In *International Conference on Acoustics, Speech, and Signal Processing*, ICASSP, Hong Kong, April 2003.
- [32] D. Taubman. High performance scalable image compression with EBCOT. *IEEE Transactions on Image Processing*, 9(7):1058–1070, July 2000.

- [33] D. Taubman and A. Zakhor. Multirate 3-D subband coding of video. *IEEE Transactions on Image Processing*, 3(5):572–588, September 1994.
- [34] D. Taubman and A. Zakhor. A common framework for rate and distortion based scaling of highly scalable compressed video. *IEEE Transactions on Circuits and Systems for Video Technology*, 6(4):329–354, August 1996.
- [35] D. S. Taubman and M.W. Marcellin. JPEG2000: Standard for interactive imaging. *Proceedings of the IEEE*, 90(8):1336–1357, August 2002.
- [36] J. D. Villasenor, B. Belzer, and J. Liao. Wavelet filter evaluation for image compression. *IEEE Transactions on Image Processing*, 4(8):1053–1060, August 1995.
- [37] S. Wong, L. Zaremba, D. Gooden, and K. Huang. Radiologic image compression—a review. *Proceedings of the IEEE*, 83(2):194–219, February 1995.
- [38] X. Wu. High-order context modeling and embedded conditional entropy coding of wavelet coefficients for image compression. In *Proceedings of 31st Asilomar Conference on Signals, Systems and Computers*, pages 1378–1382, November 1997.
- [39] Z. Xiong, K. Ramchandran, and M. Orchard. Space-frequency quantization for wavelet image coding. *IEEE Transactions on Image Processing*, 6:677–693, May 1997.