

Resolution Scalable Image Coding using Rational Wavelet Transforms

by

Felix Petngang, B.Eng.

A thesis submitted to the
Faculty of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of

Master of Applied Science in Electrical Engineering

Ottawa-Carleton Institute for Electrical and Computer Engineering

Department of Systems and Computer Engineering

Carleton University

Ottawa, Ontario

August, 2009

©Copyright

Felix Petngang, 2009



Library and Archives
Canada

Bibliothèque et
Archives Canada

Published Heritage
Branch

Direction du
Patrimoine de l'édition

395 Wellington Street
Ottawa ON K1A 0N4
Canada

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file *Votre référence*
ISBN: 978-0-494-60227-0
Our file *Notre référence*
ISBN: 978-0-494-60227-0

NOTICE:

The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

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

AVIS:

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

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

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.


Canada

Abstract

Wavelet-based image and video compression methods have received an increased interest since the advent of the JPEG2000 image coding standard. Among other advantages, wavelet decompositions offer the inherent capability of multi-resolution decoding, a feature that allows the development of image and video coders with spatial resolution scalability. However, the traditionally used two-dimensional wavelet decompositions have a dilation factor of two and are only capable of producing lower spatial resolutions obtained from the original resolution by a dyadic ratio, i.e. a power of $1/2$. In this thesis, we propose to increase the set of lower spatial resolutions available from a wavelet-based image compression system, by introducing a two-dimensional wavelet decomposition made of combining wavelet transforms of non-integer dilation factors. We describe how the proposed decomposition can produce spatial resolutions related to the original resolution by ratios of dyadic and non-dyadic nature, and evaluate its performance in a wavelet-based image coding framework.

Acknowledgments

I would like to thank my two co-supervisors, professor Richard Dansereau and professor Chris Joslin, for their assistance during the course of this research. Their advice and suggestions have consistently helped to steer this work in a productive direction, and have greatly contributed to the quality of this thesis.

Table of Contents

Abstract	iii
Acknowledgments	iv
Table of Contents	v
List of Tables	ix
List of Figures	xi
1 Introduction	1
1.1 Wavelet-based image compression and the notion of spatial resolution scalability	2
1.2 Rational wavelet transforms and their potential for finer spatial resolution scalability	8
1.3 Motivation and problem description	10
1.4 Contributions of the research	13
1.5 Organization of the thesis	14
2 Background literature review on non-dyadic spatial resolution scalability and rational wavelet transforms	15

2.1	A review of recent work on non-dyadic spatial resolution scalability in wavelet-based image coding	15
2.1.1	Increased spatial resolution scalability with full-tree dyadic wavelet decomposition	16
2.1.2	M-Band filter banks for spatial resolution scalability by non-dyadic factors	18
2.1.3	Non-dyadic spatial resolution scalability with a wavelet transform of rational dilation factor	20
2.2	Overview of the current literature on the construction of wavelet filters for rational wavelet transforms	21
3	A new wavelet decomposition structure with combined rational wavelet transforms	23
3.1	Combined rational wavelet decompositions for increased spatial resolution scalability with dyadic and non-dyadic factors	25
3.2	A parent-child tree construction scheme for compatibility with zerotree-based subband coders	30
3.3	Combining dyadic wavelet decomposition and combined rational wavelet decomposition	36
3.3.1	An extension of the combined rational wavelet decomposition to a wavelet packet decomposition	37
3.3.2	Combining dyadic wavelet decomposition and combined rational wavelet decomposition with the defined wavelet packet extension	40
3.4	Compatibility of the proposed decompositions with other existing subband coders	43

4	An algorithm structure for spatial resolution scalability with SPIHT and SPECK	44
4.1	Converting SPIHT to support spatial resolution scalability	46
4.1.1	An overview of the coding algorithm	46
4.1.2	Conversion to a subband coder supporting spatial resolution scalability	47
4.2	Converting SPECK to support spatial resolution scalability	50
4.2.1	An overview of the coding algorithm	52
4.2.2	Conversion to a subband coder supporting spatial resolution scalability	53
4.3	Limitations of the proposed conversions	54
5	A wavelet-based image coder using combined rational wavelet decomposition and its performance evaluation at different spatial resolutions	56
5.1	General structure of the wavelet-based image coding system	59
5.2	Presentation of the wavelet filters for the rational wavelet transforms	61
5.3	Experimental setup for the performance evaluation	62
5.4	Evaluation of compression performance at full size resolution	66
5.4.1	Experimental results	69
5.4.2	Analysis	75
5.5	Evaluation of compression performance at intermediate spatial resolutions	76
5.5.1	A description of the experimental procedure to generate images at non-dyadic spatial resolutions	76
5.5.2	Experimental results	82

5.6	Discussion	102
6	An investigation of image coding using an overcomplete combined rational wavelet decomposition	104
6.1	Overcomplete filter banks for rational wavelet transform	105
6.2	Overcomplete combined rational wavelet decomposition	107
6.2.1	Extending the parent-child tree construction scheme to overcomplete combined wavelet decompositions	110
6.3	Handling the redundancy in the overcomplete representation	113
6.4	Evaluation of performance at full size resolution of a wavelet-based image coder using overcomplete combined wavelet decomposition . . .	116
6.5	Analysis and discussion	121
7	Conclusion and future directions	123
7.1	Concluding remarks	123
7.2	Discussion on potential future research directions	124
	List of References	126

List of Tables

3.1	Native size ratios from multiple levels of wavelet decomposition with dilation factors of 2, 3/2 and 4/3.	24
3.2	The set of native spatial resolutions from an original image of size 704×576 , with wavelet transforms of dilation factors of 2, 3/2 and 4/3. Decimal sizes are rounded to the highest integer.	24
3.3	Native size ratios from a two-level combined rational wavelet decomposition, with different sets of dilation factors.	29
3.4	Native spatial resolutions from an original image of size 704×576 when using a two-level combined rational decomposition against a regular dyadic decomposition. Decimal sizes are rounded to integer values.	29
3.5	Sizes of the subbands produced at the m -th level of a combined decomposition with dilation factors $\{(q_i/p_i)\} = \{4/3, 3/2\}$. The original image size is $(w, h) = (w_0, h_0)$	33
5.1	Native spatial resolutions produced by each level of decomposition of an original image of size 704×576 , when using a combined rational wavelet decomposition with dilation factors $\{(q_i/p_i)\} = \{4/3, 3/2\}$ against a regular dyadic wavelet decomposition.	57
5.2	SSIM scores of the images decoded at full size resolution. The original image serves as reference image.	69

5.3	Description of the spatial resolutions of interest r_1 and r_2 for the three test images, along with the corresponding lower and higher dyadic spatial resolutions.	82
5.4	SSIM scores of the images from the rational wavelet-based image coder decoded at 3/4 the original size. The reference image is an uncompressed approximation subband from the combined wavelet decomposition.	86
5.5	SSIM scores of the images from the rational wavelet-based image coder decoded at 3/8 the original size. The reference image is an uncompressed approximation subband from the combined wavelet decomposition.	86
5.6	SSIM scores of the output images obtained at 3/4 the original size. The reference image is a decimated version of the original image. . . .	91
5.7	SSIM scores of the output images obtained at 3/8 the original size. The reference image is a decimated version of the original image. . . .	92
5.8	SSIM scores of the output images obtained at full size resolution. The original image serves as reference image.	92
6.1	Objective quality scores at full resolution for the images <i>Boat</i> and <i>Girl</i> .	118

List of Figures

1.1	A typical image compression system.	2
1.2	Traditional wavelet transform filter bank structures for one-dimensional and two-dimensional signals.	3
1.3	Illustration of two levels of two-dimensional wavelet decomposition. The details subbands LH, HL and HH produced at each level correspond respectively to the horizontal, vertical and diagonal highpass subbands.	4
1.4	An illustration of multi-resolution decoding with a wavelet-based image compression system.	5
1.5	Structure of a resolution-scalable bitstream from an encoding stage with two levels of wavelet decomposition.	5
1.6	Spatial resolution scalability with a resolution-scalable image coder.	7
1.7	Spatial resolution scalability using spatial resolution transcoding with a non-scalable image coder.	7
1.8	A one-dimensional (p, q) wavelet transform filter bank.	8
1.9	Illustration of the subbands from one level of two-dimensional (p, q) wavelet decomposition.	9
1.10	Illustration of the native resolutions produced by different (p, q) wavelet decompositions.	10

2.1	Illustration of a full-tree two-dimensional dyadic wavelet decomposition.	16
2.2	Reconstruction procedure from the full-tree decomposition proposed by Nakachi <i>et al.</i> , to obtain an image of non-dyadic spatial resolution.	17
2.3	An illustration of a one-dimensional M -band wavelet filter bank. . . .	18
2.4	Subband decomposition obtained from one level of two-dimensional wavelet decomposition with a 3-band or 4-band wavelet transform. . .	19
2.5	Illustration of a one-dimensional M -band wavelet transform with a modified synthesis filter bank to achieve a ratio of P/M , as proposed by Pau <i>et al.</i>	19
3.1	A combined rational wavelet decomposition with rational wavelet transforms of dilation factors $4/3$ and $3/2$	26
3.2	Illustration of the second level of combined decomposition with rational wavelet transforms of dilation factors $4/3$ and $3/2$	27
3.3	Construction of the parent-child tree structure for a dyadic wavelet decomposition.	32
3.4	An illustration of the difficulty of completing the parent-child tree structure for a $(2, 3)$ rational wavelet decomposition.	32
3.5	Subband association in a regular parent-child tree structure.	34
3.6	Subband association in the proposed parent-child tree structure. . . .	34
3.7	Comparison between a regular parent-child tree construction scheme in a dyadic wavelet decomposition and the proposed parent-child tree construction scheme for a combined rational wavelet decomposition. .	35
3.8	Illustration of the use of dyadic decomposition in conjunction with combined decomposition: $n_1 = 2$ levels of combined decomposition are followed by $n_2 = 2$ levels of dyadic decomposition.	37

3.9	Illustration of the wavelet packet extension of the combined rational wavelet decomposition with dilation factors $\{(q_i/p_i)\} = \{4/3, 3/2\}$. . .	38
3.10	Illustration of the use of dyadic decomposition in conjunction with combined decomposition in wavelet packet extension mode: $n_1 = 2$ levels of combined decomposition are followed by $n_2 = 2$ levels of dyadic decomposition.	40
3.11	An illustration of the parent-child tree construction scheme in a dyadic wavelet packet decomposition structure, as suggested by Kim <i>et al.</i> . .	41
3.12	A slight variant of the parent-child tree construction scheme for a dyadic wavelet packet decomposition structure.	41
3.13	An illustration of the proposed parent-child tree construction scheme when a dyadic decomposition is used in conjunction with a combined decomposition.	42
4.1	An example of bitstream with SNR scalability.	45
4.2	An illustration of bitstream with spatial resolution scalability and SNR scalability within each resolution.	45
4.3	A simplistic overview of the SPIHT coding algorithm.	47
4.4	SPIHT coding algorithm with resolution-dependent lists, producing a bitstream with SNR scalability.	48
4.5	SPIHT coding algorithm with resolution-dependent lists, producing a bitstream with resolution scalability.	49
4.6	A resolution-scalable bitstream with all possible resolutions and one bit budget specified for the full bitstream.	49
4.7	A resolution-scalable bitstream with two resolutions of interest and a bit budget specified for each portion of the bitstream.	49

4.8	Proposed SPIHT coding algorithm structure with support for spatial resolution scalability on two resolutions of interest.	51
4.9	A simplistic overview of the SPECK coding algorithm.	52
4.10	SPECK coding algorithm with resolution-dependent lists, producing a bitstream with SNR scalability.	53
4.11	Proposed SPECK coding algorithm structure with support for spatial resolution scalability on two resolutions of interest.	55
5.1	General structure of the wavelet-based image compression system. . .	59
5.2	An expanded view of the subband coding phase with RS-SPIHT or RS-SPECK.	60
5.3	Analysis wavelet filters for the (2, 3) rational wavelet transform. . . .	63
5.4	Analysis wavelet filters for the (3, 4) rational wavelet transform. . . .	64
5.5	Test image set	67
5.6	Comparison of objective quality scores between the images decoded at full size resolution.	70
5.7	A part of the decoded image from <i>Cafe</i> , compressed at 0.5 bpp with SPIHT.	72
5.8	A part of the decoded image from <i>Cafe</i> , compressed at 1.0 bpp with SPIHT.	73
5.9	A part of the decoded image from <i>Boat</i> , compressed at 0.3 bpp with SPIHT.	74
5.10	Proposed experimental procedure to generate the output images at a non-dyadic spatial resolution r and at full size resolution, from the image coders \mathcal{A} and \mathcal{B}	79
5.11	An illustration of the structure of the bitstreams from the three separate encoding stages of the experimental procedure.	80

5.12	An illustration of the structure of the bitstreams from the encoding stages of the experimental procedure, when two intermediate resolutions of interest are specified.	83
5.13	Comparison of objective quality scores between images from the rational wavelet-based image coder decoded at 3/4 the original size.	87
5.14	Comparison of objective quality scores between images from the rational wavelet-based image coder decoded at 3/8 the original size.	88
5.15	Comparison of objective quality scores between the output images obtained at 3/4 the original size.	93
5.16	Comparison of objective quality scores between the output images obtained at 3/8 the original size.	94
5.17	Comparison of objective quality scores between the output images obtained at full size resolution.	95
5.18	A part of the output images from <i>Cafe</i> , encoded with RS-SPIHT at a bitrate of 0.5 bpp and decoded at 3/4 the original size.	97
5.19	A part of the output images from <i>Cafe</i> , encoded with RS-SPIHT at a bitrate of 1.0 bpp and decoded at 3/4 the original size.	98
5.20	A part of the output images from <i>Boat</i> , encoded with RS-SPIHT at bitrates of 0.3 bpp (left) and 1.0 bpp (right), and decoded at 3/4 the original size.	99
5.21	Output images from <i>Girl</i> , encoded with RS-SPIHT at a bitrate of 1.0 bpp and decoded at 3/8 the original size.	101
6.1	Critically-sampled filter bank structure for a one-dimensional (p, q) rational wavelet transform.	106
6.2	Overcomplete filter bank structure proposed by Bayram and Selesnick for a one-dimensional (p, q) rational wavelet transform.	106

6.3	Analysis filter coefficients and frequency responses of the wavelet filters for the overcomplete (3, 4) and (2, 3) rational wavelet transforms. . .	108
6.4	Subband decompositions obtained from a critically-sampled (p, q) rational wavelet transform and an overcomplete (p, q) rational wavelet transform. The illustrated structures correspond to one level of two-dimensional rational wavelet decomposition.	109
6.5	Subband decomposition from one level of overcomplete combined rational wavelet decomposition with dilation factors $\{(q_i/p_i)\} = \{4/3, 3/2\}$.	110
6.6	An illustration of the parent-child tree construction rules.	112
6.7	Description of the iterative noise shaping method by Reeves and Kingsbury.	113
6.8	Objective quality scores of the reconstructed image after each iteration, for a target threshold of 8 and a starting threshold of 64.	115
6.9	Percentage of non zeros coefficients in the highpass subbands after each iteration, for a target threshold of 8 and a starting threshold of 64. .	115
6.10	A description of the two possible encoding stages for the wavelet-based image coder with overcomplete combined rational wavelet decomposition.	117
6.11	Compression performance at full resolution for the images <i>Boat</i> and <i>Girl</i>	119
6.12	A part of the decoded image from <i>Boat</i> , compressed at 0.5 bpp with SPIHT.	120

Chapter 1

Introduction

In our daily applications, images and videos are usually carried in compressed forms. The area of data compression defines the field of methodologies and techniques that are used to represent data in a compact fashion. Throughout the years, the knowledge and methods for image and video compression have gradually improved, as confirmed by the constant evolution of image and video coding standards: from JPEG [1] to JPEG2000 [2, 3] in image compression and from MPEG [4] to H.264/AVC [5, 6] in video coding. In addition to better compression efficiency, the recent compression standards have brought new properties ranging from better robustness to transmission errors to enhanced security features and scalability on different levels. The notion of *scalability*, in particular, designates the idea that an image (or video) can be compressed once and decompressed in various ways to produce images of different visual quality and different spatial resolutions, or videos of different temporal resolutions. The inclusion of scalability in an image or video compression scheme provides an efficient way to accommodate various users with different bandwidth requirements, storage capabilities, or display capabilities, all from one single compressed file. This particular property was at the heart of the development of the recent JPEG2000 image coding standard [2].

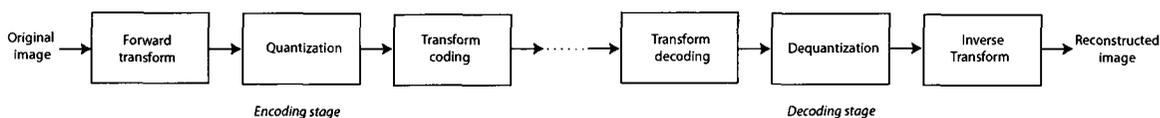
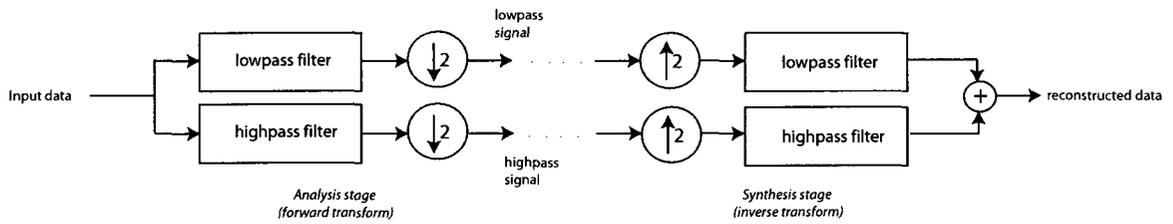


Figure 1.1: A typical image compression system.

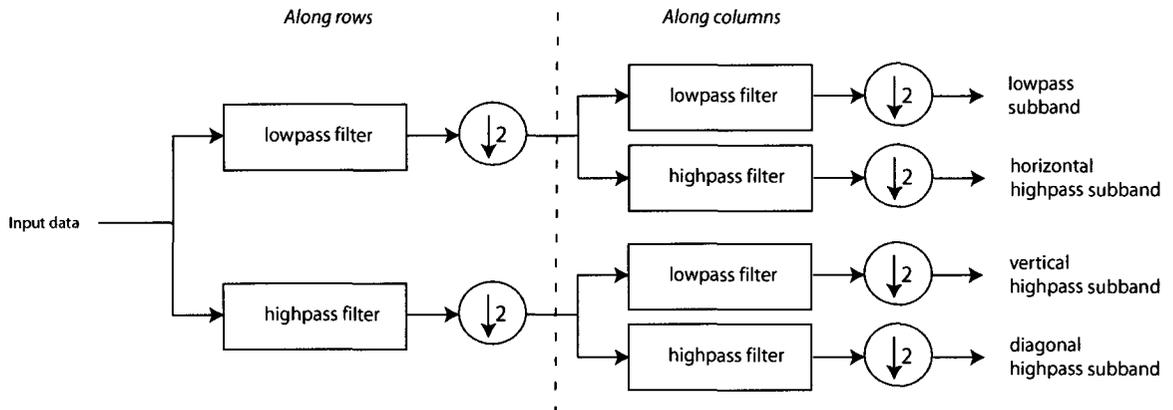
1.1 Wavelet-based image compression and the notion of spatial resolution scalability

A typical image compression system (or image coding system, or image coder)¹ consists at the encoding stage of an image transform operation followed by quantization and coding steps, as illustrated in Figure 1.1. The image transform operation produces coefficients that are amenable to compression, by separating low frequency components that represent perceptually significant data of the image from high frequency components which are perceived as details. The Discrete Cosine Transform (DCT) and the Discrete Wavelet Transform (DWT) are the most popular image transform operations used in image compression. DCT is used in the widely known JPEG image compression standard, while the discrete wavelet transform is at the core of the state-of-the-art JPEG2000 image coding standard. As traditionally defined, a wavelet transform convolves an input signal with a lowpass filter and a highpass filter to produce respectively a lowpass signal and a highpass signal, after a subsequent downsampling operation by a factor of two. Figure 1.2 illustrates the filter bank structures of a one-dimensional and a two-dimensional wavelet transform. To apply the transform on a two-dimensional data like an image, the operation is performed along the rows and the columns, and produces one *lowpass subband* (the *approximation subband*) and three *highpass subbands* (the *details subbands*). Multiple levels of decomposition are constructed by successively performing a two-dimensional wavelet

¹The three expressions will be used interchangeably in the current document.



(a) One-dimensional analysis and synthesis filter banks



(b) Two-dimensional analysis filter bank

Figure 1.2: Traditional wavelet transform filter bank structures for one-dimensional and two-dimensional signals.

transform on the approximation subband. Figure 1.3 illustrates the subbands from a two-level wavelet decomposition² on an input image. The reconstruction operation is the inverse process of the forward decomposition; the original image is reconstructed from the approximation and details subbands.

The approximation subband produced by a two-dimensional wavelet decomposition contains low frequency components of the image and is a good quality version of the original image at a smaller *spatial resolution*. Throughout this document, the term *spatial resolution* is used to denote the size of an image, expressed in width and

²The terms transform and decomposition are used interchangeably herein.

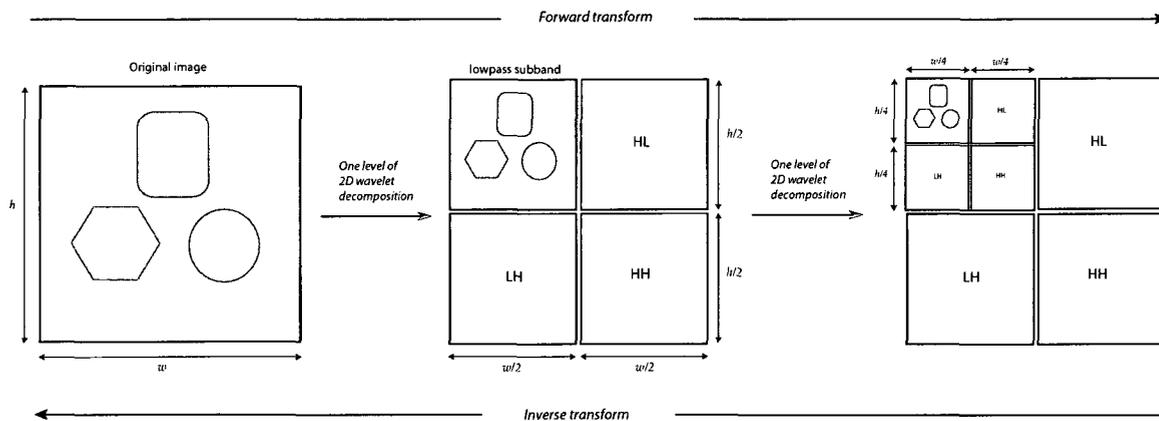


Figure 1.3: Illustration of two levels of two-dimensional wavelet decomposition. The details subbands LH, HL and HH produced at each level correspond respectively to the horizontal, vertical and diagonal highpass subbands.

height. In a multi-level wavelet decomposition, the spatial resolution of the approximation subband is divided by a factor of two at each level, producing increasingly smaller versions of the original image. We define by *native* spatial resolutions the successive spatial resolutions of the approximation subband during the multi-level wavelet decomposition process (this includes the original resolution of the image). For example, the two-level wavelet decomposition of Figure 1.3 produces two native spatial resolutions $(w/2, h/2)$ and $(w/4, h/4)$ in addition to the original spatial resolution (w, h) of the image. During the reconstruction process, the approximation subband starts with a resolution $(w/4, h/4)$ and is successively reconstructed to an image of resolution $(w/2, h/2)$ and then to the original image. This property of the wavelet transform to produce smaller versions of the original image during the reconstruction process is used in wavelet-based image compression systems to decode the original image at multiple spatial resolutions, from the same encoded bitstream.

Figure 1.4 illustrates an example of multi-resolution decoding with a wavelet-based image compression system where the original image is transformed using two

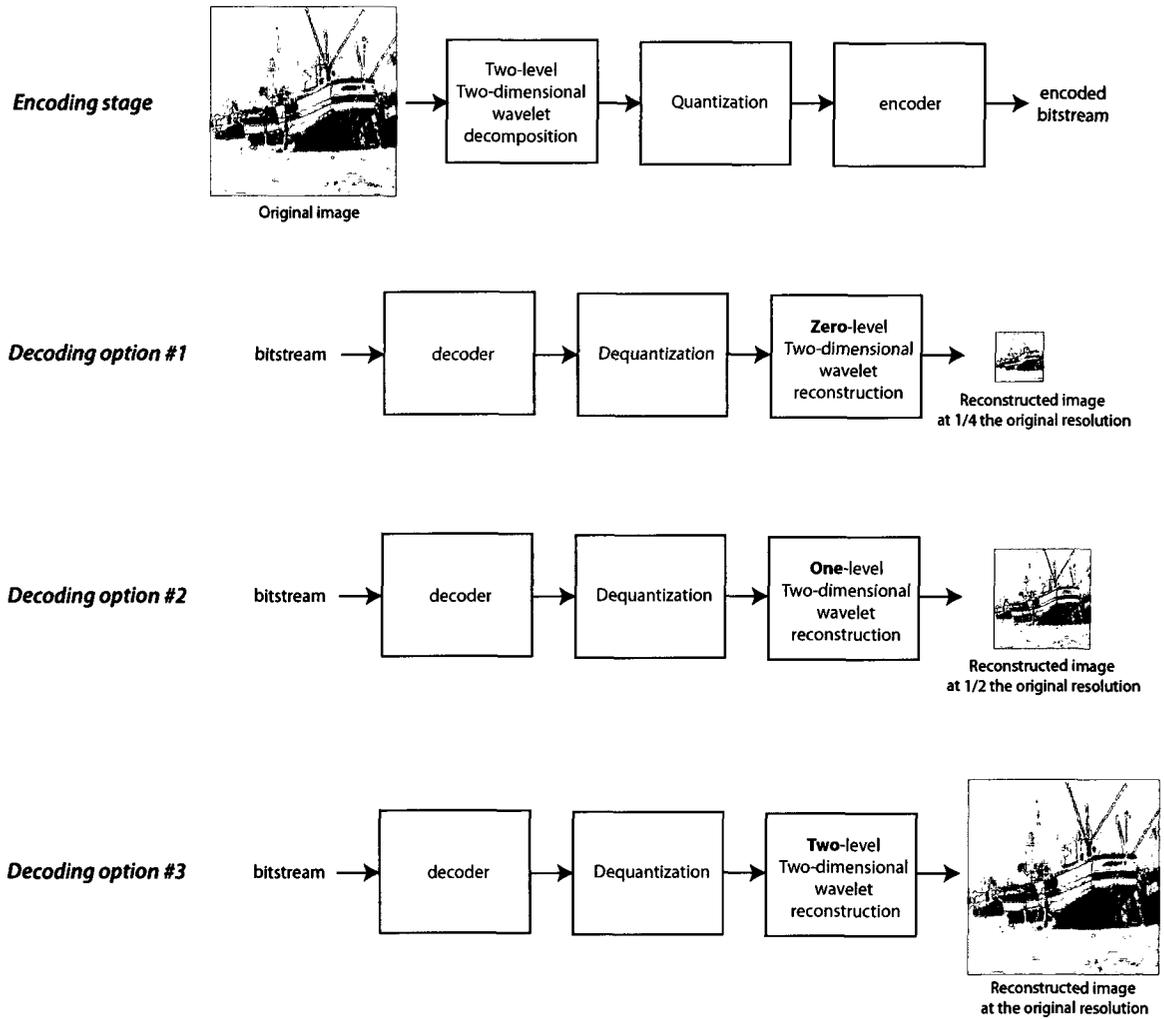


Figure 1.4: An illustration of multi-resolution decoding with a wavelet-based image compression system.

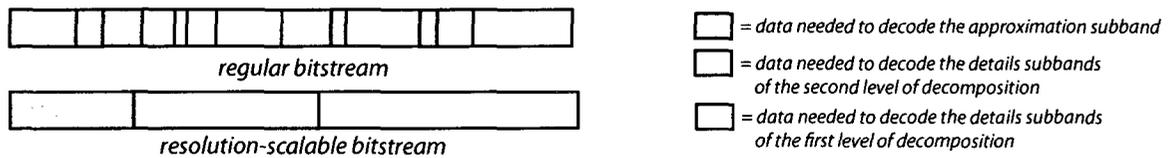


Figure 1.5: Structure of a resolution-scalable bitstream from an encoding stage with two levels of wavelet decomposition.

levels of two-dimensional wavelet decomposition at the encoding stage. In addition to the original image resolution, the image can be decoded at the native spatial resolutions produced during the forward wavelet decomposition. This capability of decoding the image at multiple spatial resolutions from the same encoded bitstream is defined as *spatial resolution scalability*. Spatial resolution scalability is a very efficient way to accommodate various devices or end-users with different display capabilities. The original image is encoded only once, but can be decoded at different spatial resolutions according to the user's needs. Furthermore, in an image coding system with support for spatial resolution scalability, the bitstream produced by the encoding stage is made *resolution-scalable*. This means the bitstream data is arranged so that the bitstream can be truncated to only comprises the necessary information needed to decode the image at a given spatial resolution. This is illustrated in Figure 1.5 for a bitstream produced by an encoding stage with two levels of two-dimensional wavelet decomposition. A *resolution-scalable* image coder is defined as one capable of producing a resolution-scalable bitstream at the encoding stage.

With a resolution-scalable bitstream, a bitstream parser can truncate the original bitstream and send different data to users with various resolution needs. This is illustrated in Figure 1.6. The motivation behind the bitstream parsing operation is simple: only send a user the data that it actually needs to decode the image up to its preferred spatial resolution. This is quite natural when you consider that in practice, a device with smaller display capabilities usually comes with smaller storage capabilities as well. To achieve the same goals, *non-scalable*³ image compression systems often make use of *spatial resolution transcoding*: in order to produce a bitstream for a given spatial resolution, a transcoder first decodes the original bitstream, then performs an image resizing operation in the spatial or transform domain, and finally encodes the

³By *non-scalable*, we implicitly refer to scalability on spatial resolution.

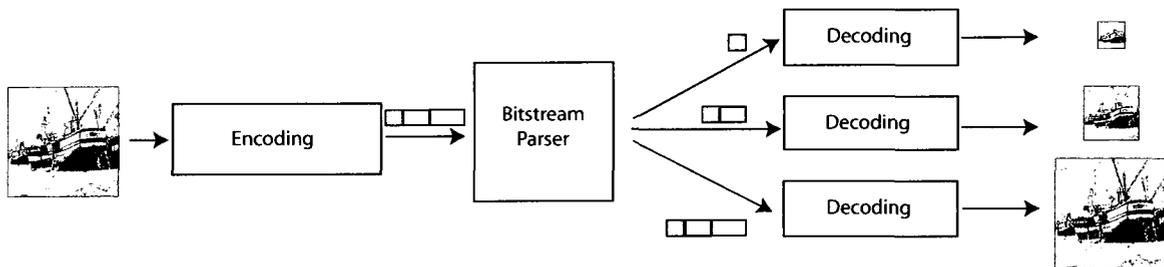


Figure 1.6: Spatial resolution scalability with a resolution-scalable image coder.

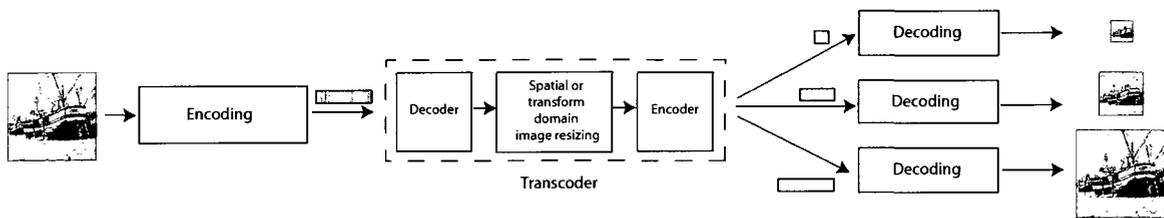


Figure 1.7: Spatial resolution scalability using spatial resolution transcoding with a non-scalable image coder.

resulting data. An illustration of the use of spatial resolution transcoding to achieve spatial resolution scalability is displayed in Figure 1.7. Spatial resolution transcoding has been often used with non-scalable DCT-based image coders, and image resizing techniques of various computational complexity have been developed [7,8]. However, it bears the inconveniences of higher computing load and potential delay in real-time applications. The success of wavelet transforms in image compression is in part due to the inherent capability of producing resolution-scalable image coding systems. The additional spatial resolutions are produced during the forward wavelet decomposition process, and because the inverse operation naturally reconstructs the spatial resolutions in progressive fashion, it eases the formation of resolution-scalable bitstreams and the subsequent parsing operation.

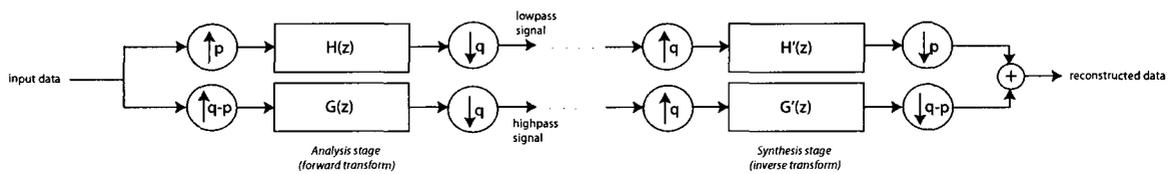


Figure 1.8: A one-dimensional (p, q) wavelet transform filter bank.

1.2 Rational wavelet transforms and their potential for finer spatial resolution scalability

The wavelet transform filter bank described in Section 1.1 contained upsampling and downsampling factors of two. This is representative of the widely used definition of a wavelet transform. It can however be generalized to a (p, q) wavelet transform, with the filter bank structure displayed in Figure 1.8. In a (p, q) wavelet transform, the *sampling coefficients* p and q are integer values such that $p < q$. The ratio (q/p) is referred to as the *dilation* (or *scaling*) *factor*. The traditional definition of a wavelet transform, as introduced and discussed in the previous section, corresponds to a (p, q) wavelet transform with sampling coefficients $p = 1$ and $q = 2$, and a dilation factor of 2. In the rest of the document, the expression *wavelet transform* will denote a (p, q) wavelet transform. A wavelet transform with a dilation factor of 2 will be qualified a *dyadic* wavelet transform. In the same idea, we denote as *rational* wavelet transform, a wavelet transform with a non-integer (rational) dilation factor q/p . Rational wavelet transforms have been often used in areas like audio coding and communications [9,10], because they produce a *finer* frequency selectivity. In this section, we look at them from the perspective of *finer* spatial resolution scalability.

A two-dimensional (p, q) wavelet decomposition produces an approximation sub-band of size $(\frac{p}{q} \times w, \frac{p}{q} \times h)$, where (w, h) represent the width and height of the original

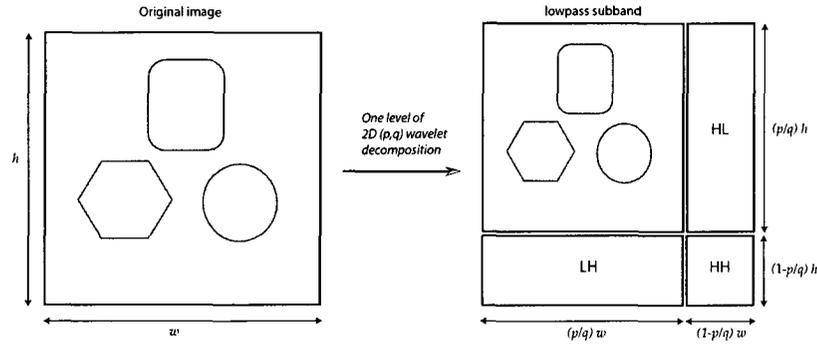


Figure 1.9: Illustration of the subbands from one level of two-dimensional (p, q) wavelet decomposition.

image. This is illustrated in Figure 1.9 for one level of decomposition. Multiple levels of decomposition produce native spatial resolutions that are related to the original resolution by a power of (p/q) . For example, a dyadic wavelet decomposition produces native spatial resolutions of $1/2$ the original size, $1/4$ the original size, $1/8$ the original size, etc..., in addition to the original size of the image. In general, for n levels of two-dimensional decomposition, the set of native spatial resolutions that are produced by a (p, q) wavelet transform is determined by the set of ratios $\{(\frac{p}{q})^0, (\frac{p}{q})^1, (\frac{p}{q})^2, \dots, (\frac{p}{q})^n\}$ (also denoted as $\{(\frac{p}{q})^i, i = 0, \dots, n\}$ ⁴). The specific ratios of $(\frac{1}{2})^i$ relating the native resolutions produced by a dyadic wavelet decomposition to the original size of the image are referred to as *dyadic ratios* or *dyadic factors*. Correspondingly, the native resolutions obtained by dyadic factors are referred to as the *dyadic* native resolutions.

In a two-dimensional decomposition, the partitioning of the two-dimensional space formed by the original image is determined by the dilation factor q/p : a smaller dilation factor induces a finer partitioning and results in a higher number of possible native resolutions on a given space. Figure 1.10 illustrates the successive spatial

⁴The notation $i = a, \dots, b$ means that i can take any integer value between a and b , inclusively.

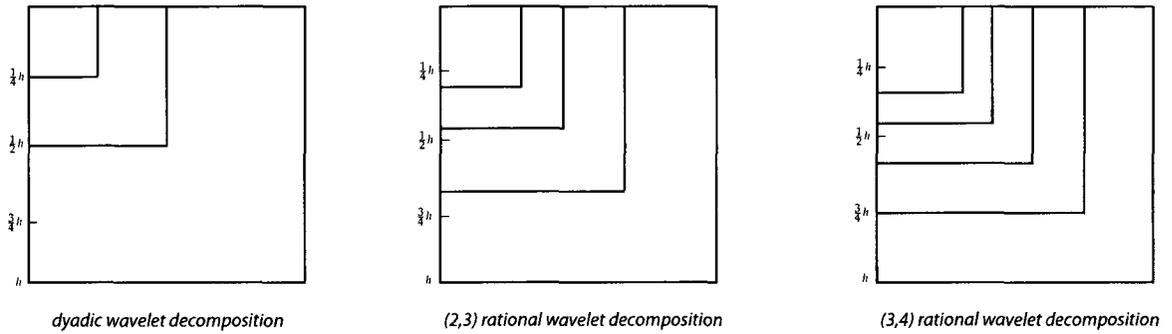


Figure 1.10: Illustration of the native resolutions produced by different (p, q) wavelet decompositions.

resolutions produced by wavelet transforms of different dilation factors. In the illustration, the wavelet transform with the smallest dilation factor $q/p = 4/3$ produces five native spatial resolutions that are greater or equal to one quarter of the original image size. The transforms with higher dilation factors $q/p = 3/2$ and $q/p = 2$ produce respectively four and three native spatial resolutions greater or equal to one quarter of the original image size.

1.3 Motivation and problem description

In this thesis, we are interested in the notion of spatial resolution scalability and the ability to achieve it with a resolution-scalable image coding system. The traditionally used dyadic wavelet transform provides spatial resolution scalability with *dyadic* ratios, and the possible resolutions are related to the original resolution by a power of $1/2$. As described in Section 1.2, a finer partitioning of the two-dimensional space can be produced with a wavelet transform of smaller dilation factor than the dyadic wavelet decomposition. In this research, we are interested in increasing the spatial resolution scalability potential of wavelet-based image compression systems, by providing a richer set of possible decoding resolutions at the decoding stage. With spatial resolution scalability on a larger set of resolutions, the image compression

system would have the ability to efficiently accommodate a wider range of end-users with various display/memory capabilities. This is naturally motivated by the increasing development of devices with display terminals of various sizes, equipped with the capability of running image and video applications.

We illustrated in the previous section that a rational wavelet decomposition with a dilation factor $q/p < 2$ can produce more native resolutions than a traditional dyadic wavelet decomposition. The use of rational wavelet transforms is therefore, in theory, a better option than the dyadic wavelet decomposition from the perspective of *fine* spatial resolution scalability in the image compression system. However, rational wavelet decompositions have seldom been used in image compression [11, 12]. This can be attributed to many reasons. On one hand, until recently, earlier efforts to define spatially scalable extensions for existing image and video coding standards had received very little success, and spatial resolution transcoding was still a preferred option. The recent advent of the spatially scalable extension of the H.264/AVC video coding standard [13], including additional support for *non-dyadic* configurations [14], has contributed to the increased interest for image and video coders with support for spatial resolution scalability by dyadic and non-dyadic factors. On the other hand, very few propositions of wavelet filters for rational wavelet transforms are available in the current literature. The field of filter design for wavelet transforms of rational dilation factors does not exhibit the level of maturity that can be observed in the design of wavelet filters for dyadic wavelet decompositions. Research work on the construction of wavelet filters for rational wavelet transforms is very scarce, and has not produced the kind of diversity in the choice of filters that is offered for dyadic wavelet transforms. However, significant advances have been made recently on the design of wavelet filters for rational wavelet transforms [15–17]. Encouraged by these recent developments, the research presented in the current document proposes to use

rational wavelet decompositions with dilation factors $q/p < 2$ in order to increase the set of native spatial resolutions available from wavelet-based image compression systems. The objective of the research work is to build a wavelet-based image coder capable of producing a richer set of decoding resolutions than a traditional image coder with dyadic wavelet decomposition. Furthermore, we are interested in building a wavelet-based image coder whose set of decoding resolutions can include spatial resolutions obtained from dyadic and non-dyadic ratios. This last requirement is defined in order to make the resulting image coder an improvement over a traditional wavelet-based image coder only capable of dyadic spatial resolution scalability.

An obvious challenge that arises from the stated objectives is how to derive a wavelet decomposition structure capable of producing native spatial resolutions of dyadic and non-dyadic nature. Additionally, image coding with rational wavelet decompositions presents the challenge of compatibility with some of the existing wavelet subband coders, because of the different partitioning of the two-dimensional space. In the current research, the problem of compatibility with existing wavelet subband coders is specifically targeted at *zerotree-based* subband coders, a family of transform coders that are widely used in wavelet-based image compression. Because these subband coders take advantage of the specific partitioning from a dyadic wavelet decomposition, the use of alternative decomposition structures requires that a proper adaptation scheme should be defined. The work presented in the current document not only introduces a wavelet decomposition structure capable of producing spatial resolutions obtained from dyadic and non-dyadic factors, it also describes an adaptation scheme for compatibility with subband coders of the zerotree coding family.

Finally, the scope of the current research has been limited to still image coding. However, the work described can potentially be extended to wavelet-based video compression, where the inclusion of spatial resolution scalability presents additional

challenges [13, 18, 19].

1.4 Contributions of the research

In order to achieve the objectives described, different contributions are proposed in this research:

- a) *A two-dimensional wavelet decomposition structure made of combining multiple rational wavelet decompositions of different dilation factors (q_i/p_i) , selected such that: $\prod_i(q_i/p_i) = 2$.*

The proposed *combined rational wavelet decomposition* has the advantage of producing native spatial resolutions obtained from dyadic *and* non-dyadic ratios.

- b) *A parent-child tree construction scheme for the resulting subband decomposition.*

Subband coders of the zerotree coding family use a parent-child tree structure in their coding algorithm. Along with the proposed wavelet decomposition, we define an adaptation scheme for the construction of the parent-child tree structure, in order to ensure compatibility with zerotree-based subband coders.

- c) *An adaptation of two wavelet subband coders, namely SPIHT and SPECK, into subband coders capable of producing resolution-scalable bitstreams.*

SPIHT(Set Partitioning In Hierarchical Trees) [20] and SPECK (Set Partitioning Embedded bloCK coder) [21] are two popular subband coders used in wavelet-based image compression. The current research presents a simple change to the structure of their respective coding algorithm in order to produce a bitstream amenable to spatial resolution scalability.

- d) *An experimental procedure for the evaluation of performance at non-dyadic spatial resolutions against a wavelet-based image coder only capable of producing dyadic native spatial resolutions.*

In the current literature on non-dyadic spatial resolution scalability in wavelet-based image compression, the authors have used various ways of comparing their proposed image coder to an image coder only capable of producing dyadic native spatial resolutions. In the course of the current research, a different experimental procedure has been defined for the evaluation of compression performance at non-dyadic spatial resolutions between a wavelet-based image coder capable of non-dyadic spatial resolution scalability and a traditional image coder based on dyadic wavelet decomposition.

1.5 Organization of the thesis

The present document is organized as follows: Chapter 2 provides a review of existing work on spatial resolution scalability by non-dyadic factors in wavelet-based image coding, along with an overview of various contributions on the construction of wavelet filters for rational wavelet transforms. Chapter 3 introduces the new two-dimensional wavelet decomposition structure obtained by combining multiple rational wavelet decompositions, and describes the adaptations proposed for zerotree-based subband coders. Chapter 4 describes how SPIHT and SPECK can be converted to produce bitstreams with spatial resolution scalability. In Chapter 5, the compression performance of an image coder built with the proposed wavelet decomposition is evaluated at different spatial resolutions. Chapter 6 presents an extension of our work to *over-complete* filter banks for rational wavelet transforms. Finally, Chapter 7 concludes the document and discusses potential future research directions.

Chapter 2

Background literature review on non-dyadic spatial resolution scalability and rational wavelet transforms

2.1 A review of recent work on non-dyadic spatial resolution scalability in wavelet-based image coding

Research work on non-dyadic spatial resolution scalability with wavelet-based image coders is fairly limited. The current literature on the area is very scarce, and only few contributions are available. In this section, we present the work performed by different authors to realize spatial resolution scalability by non-dyadic factors in wavelet-based image compression systems. The techniques described vary from *full-tree* dyadic decompositions [22], *M-band* filter banks [23] and the use of a wavelet transform of rational dilation factor [11].

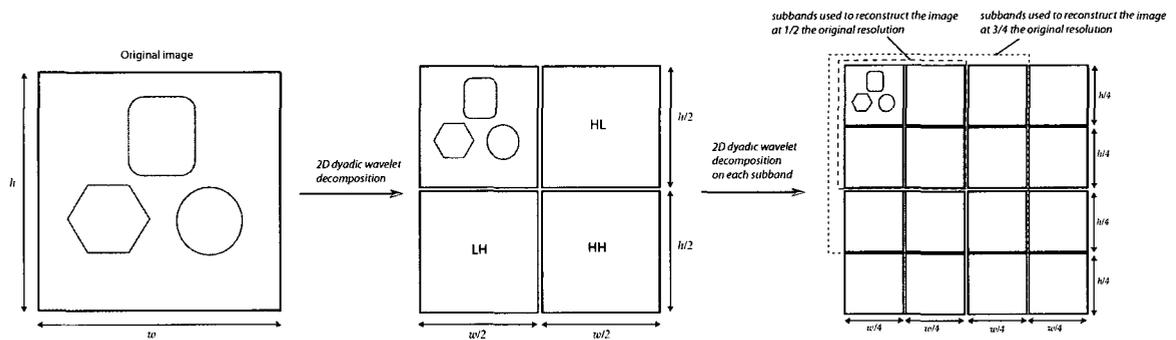


Figure 2.1: Illustration of a full-tree two-dimensional dyadic wavelet decomposition.

2.1.1 Increased spatial resolution scalability with full-tree dyadic wavelet decomposition

To achieve spatial resolution scalability with resolutions obtained from dyadic and non-dyadic ratios, Nakachi *et al.* [22,24] propose the use of a *full-tree* two-dimensional dyadic wavelet decomposition. At each level of dyadic wavelet decomposition, with the exception of the first, the two-dimensional wavelet transform is applied on the approximation subband *and* on each of the three details subbands from the previous level of decomposition. The resulting subband decomposition forms a two-dimensional grid, as illustrated in Figure 2.1 for two levels of decomposition. The authors determine from this grid structure what subbands to select in order to reconstruct the image at a given spatial resolution. The subbands selected for a reconstruction at $1/2$ the original size of the image and at $3/4$ the original size of the image are indicated on Figure 2.1. While the subbands selected to reconstruct the half-size resolution are the subbands obtained from a regular dyadic wavelet decomposition, the subbands selected for a reconstruction at $3/4$ the original size include additional subbands produced by extending the decomposition to a full-tree decomposition. The reconstruction process to obtain an image at a non-dyadic spatial resolution is illustrated

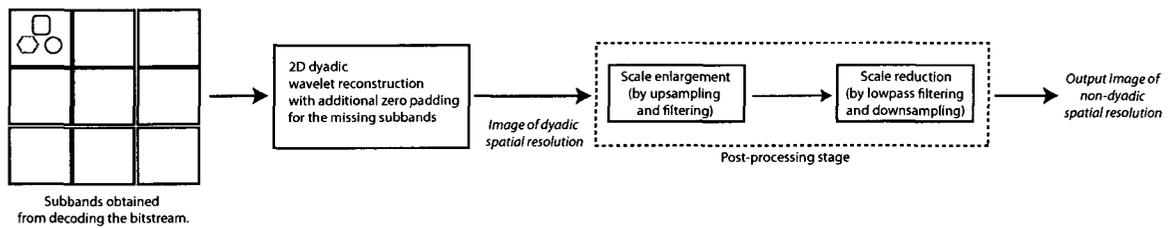


Figure 2.2: Reconstruction procedure from the full-tree decomposition proposed by Nakachi *et al.*, to obtain an image of non-dyadic spatial resolution.

in Figure 2.2. In the wavelet reconstruction operation, missing subbands from the full-tree decomposition are replaced by zeros. The output of the wavelet reconstruction is an image of dyadic spatial resolution, precisely the dyadic resolution that is immediately higher than the requested spatial resolution. After the wavelet reconstruction, a post-processing stage is performed in order to resize the output image from the dyadic scale to the target non-dyadic resolution. The post-processing stage consists of a scale enlargement followed by a scale reduction, both operations performed through the use of filtering and sampling. The output of the post-processing stage is the output image of non-dyadic spatial resolution.

An advantage of the method by Nakachi *et al.* is the ability to produce additional spatial resolutions obtained from non-dyadic factors, while keeping the traditional native spatial resolutions from a dyadic wavelet decomposition. However, the post-processing stage to obtain non-dyadic spatial resolutions is an additional complexity at the decoding side and a computational burden for any image compression system. Another concern is the fact that the subband coefficients of discarded subbands from the full-tree decomposition are replaced by zeros before the wavelet reconstruction is performed: this means that the output of the wavelet reconstruction is an image with limited amount of high frequency details and consequently limited visual quality.

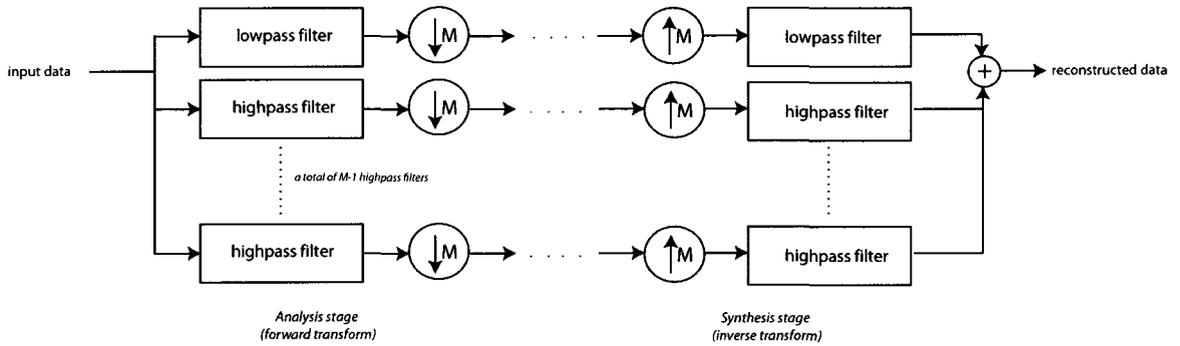


Figure 2.3: An illustration of a one-dimensional M -band wavelet filter bank.

2.1.2 M -Band filter banks for spatial resolution scalability by non-dyadic factors

Pau *et al.* [23, 25] have proposed the idea of using M -band wavelet transforms for increased spatial resolution scalability in wavelet-based image coding. An M -band wavelet filter bank is illustrated in Figure 2.3. It is a generalization of the 2-band filter bank used in the traditional dyadic wavelet decomposition. A two-dimensional M -band wavelet decomposition is formed by applying the one-dimensional filter bank along the rows and the columns of the image. It produces an approximation subband and $((M - 1)^2 - 1)$ details subbands. This is illustrated in Figure 2.4 for one level of two-dimensional decomposition with a 3-band wavelet transform and a 4-band wavelet transform. The method proposed by Pau *et al.* attempts to exploit the two-dimensional partitioning obtained from M -band decompositions with $M > 2$, in order to produce spatial resolutions obtained from different ratios. To that end, the authors propose to use a modified synthesis filter bank in the manner illustrated in Figure 2.5. By changing the upsampling factors from M to P in the synthesis stage and by modifying the synthesis filters, an output signal of $(\frac{P}{M} \times N)$ samples is obtained from an input signal of N samples. A resampling filter is used to modify

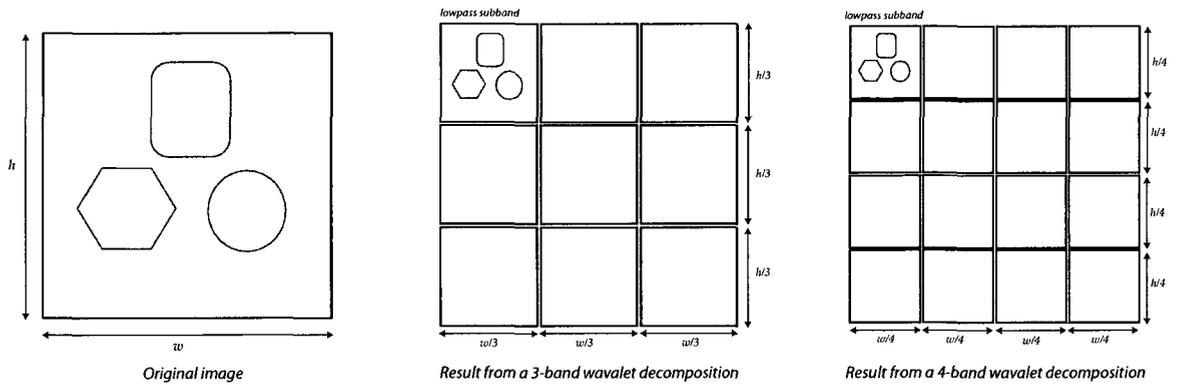


Figure 2.4: Subband decomposition obtained from one level of two-dimensional wavelet decomposition with a 3-band or 4-band wavelet transform.

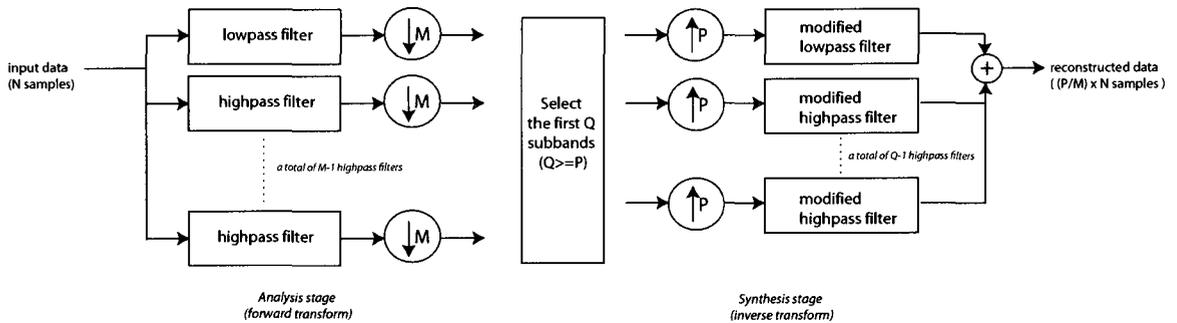


Figure 2.5: Illustration of a one-dimensional M -band wavelet transform with a modified synthesis filter bank to achieve a ratio of P/M , as proposed by Pau *et al.*

the filters of the synthesis stage.

For a target size ratio of (P/M) , the method specifies to select a number Q of subbands ($P \leq Q \leq M$) out the M subbands produced by the analysis stage. For a two-dimensional data, this translates into selecting a $Q \times Q$ group of subbands out of the $M \times M$ group of subbands produced by the forward wavelet decomposition. Ideally, to obtain an image of spatial resolution equal to P/M the original size of the image, we would like the value Q to be equal to P , since a $P \times P$ group of subbands would have the same number of samples as a decoded image at the requested resolution. The results reported by the authors suggest, however, that a higher value of Q is preferable, as it produces better results in terms of visual quality. This

is certainly a limitation of the described method. An additional concern with the approach presented is the need to build a different set of synthesis filters for each possible resizing ratio.

2.1.3 Non-dyadic spatial resolution scalability with a wavelet transform of rational dilation factor

In their contribution, Xiong *et al.* [11] propose the use of a wavelet transform of rational dilation factor to obtain non-dyadic spatial resolution scalability. The work described by the authors mainly focuses on implementing a rational wavelet transform under a *lifting-based* transform structure. The use of a *lifting* implementation [26,27] for the wavelet transform is very popular in wavelet-based image compression because it results in significantly reduced computational complexity, in comparison to the traditional convolution-based filter bank implementation. Xiong *et al.* propose a lifting framework for a wavelet transform of dilation factor $3/2$ and present various prediction models to generate the lifting coefficients. However, the authors do not attempt to combine spatial resolution scalability by both dyadic and non-dyadic ratios. Furthermore, their results are constrained to wavelet filters of one *vanishing moment*, a property of the wavelet filters which is reported to have an impact on the compression performance [28, 29].

2.2 Overview of the current literature on the construction of wavelet filters for rational wavelet transforms

The notion of rational wavelet transform was introduced and briefly discussed in Chapter 1. A main reason for the limited use of rational wavelet transforms is the limited results from the research effort on the construction of wavelet filters for rational wavelet transforms. The difficulty of designing perfect-reconstruction filters for rational wavelet transforms is well described by Blu [30] and by Bayram and Selnick [17]. The problems posed by the filter design process in the case of rational wavelet transforms cannot be solved by the same approaches that are used to design wavelet filters for dyadic wavelet decompositions. As a consequence, very little flexibility is available in the design, and only few constraints can be imposed on the filters. Although the issue of filter design for rational wavelet transforms is not specifically addressed in the current research, we present in this section an overview of different contributions and recent developments made in this area. Examples of wavelet filters for rational wavelet transforms are presented in Chapter 5 and Chapter 6.

Earlier contribution from Kovacevic and Vetterli [31] and a more recent work by Blu [30] are often cited references in the construction of wavelet filters for rational wavelet transforms. Kovacevic and Vetterli [31] were the first to propose an extension of the dyadic filter bank to sampling coefficients forming a rational dilation factor. Blu [30] has proposed an algorithm to design filter banks with integer or rational dilation factors satisfying the constraints of *orthonormality* and *regularity*. Orthonormal filter banks are characterized by synthesis filters being time-reversed versions of the analysis filters. The regularity property rejoins the concept of *vanishing moment* mentioned

at the end of Section 2.1. From an image compression perspective a higher regularity order is preferable. The algorithm proposed by Blu is an iterative algorithm which reaches convergence for arbitrary regularity orders only when the filter bank has an integer dilation factor. As reported by the author, convergence cannot be achieved for more than one regularity order when a rational dilation factor is requested.

To our knowledge, Bayram and Selesnick [15] were the first to propose wavelet filters of higher regularity order for orthonormal filter banks of rational dilation factors. However, in their design, the authors used a powerful but computationally expensive method to generate the filter coefficients. In a very recent work, the same authors have proposed a design algorithm with an iterative approach of smaller complexity to compute the filter coefficients [16]. The wavelet filters used in the current research, which are described in Chapter 5, were obtained from this contribution.

Another interesting contribution on rational wavelet transforms is proposed by Baussard *et al.* [12]. The authors describe a construction of the wavelet filters in the Fourier domain and propose a fast algorithm to compute the rational wavelet transform. Compared to other work, their design does not involve an iterative algorithm but the proposed filters are infinitely-based, a property often judged impractical in image compression systems.

Finally, Bayram and Selesnick [17, 32] have recently proposed the use of an *overcomplete* filter bank of rational dilation factor. The filter bank is said to be overcomplete because the total number of samples in the output subbands of the analysis stage is higher than the number of samples in the input signal. The authors propose the use of an overcomplete filter bank to introduce more flexibility in the design process, and reduce the overall complexity of the design algorithm. In the current research, we have investigated the use in image compression of an overcomplete filter bank of rational dilation factor. This is described in Chapter 6.

Chapter 3

A new wavelet decomposition structure with combined rational wavelet transforms

In Chapter 1, we demonstrated how the use of wavelet transforms of rational dilation factors $q/p < 2$ can produce a finer partitioning of the two-dimensional space than a dyadic wavelet decomposition. We established that a two-dimensional wavelet transform with a rational dilation factor of q/p (which will also be referred to as a (p, q) *rational wavelet transform*) produces a set of native resolutions determined by the set of ratios $\{(\frac{p}{q})^m, m = 0, \dots, n\}$, where n represents the total number of levels of decomposition. We denote this set of ratios as the *native size ratios* produced by the wavelet decomposition. Given an original image of size (w, h) , the native resolution produced after m levels of decomposition corresponds to a native size ratio of $(\frac{p}{q})^m$ and an image of size $((\frac{p}{q})^m \times w, (\frac{p}{q})^m \times h)$.

Table 3.1 lists the native size ratios produced by multiple levels of dyadic wavelet decomposition, against the use of rational wavelet decompositions of dilation factors of $3/2$ or $4/3$. For the sake of simplicity in the example, we have assumed that the resolution corresponding to a quarter of the original image size is the lowest desired spatial resolution, and consequently restricted ourselves to native size ratios that are

$(p, q)=(1, 2)$ $n = 2$	$(p, q)=(2, 3)$ $n = 3$	$(p, q)=(3, 4)$ $n = 4$
1	1	1
1/2	2/3	3/4
1/4	4/9	9/16
	8/27	27/64
		81/256

Table 3.1: Native size ratios from multiple levels of wavelet decomposition with dilation factors of 2, 3/2 and 4/3.

$(p, q)=(1, 2)$ $n = 2$	$(p, q)=(2, 3)$ $n = 3$	$(p, q)=(3, 4)$ $n = 4$
704×576	704×576	704×576
352×288	470×384	528×432
176×144	313×256	396×324
	209×171	297×243
		223×183

Table 3.2: The set of native spatial resolutions from an original image of size 704×576 , with wavelet transforms of dilation factors of 2, 3/2 and 4/3. Decimal sizes are rounded to the highest integer.

greater or equal to 1/4. With this constraint, the total number of decomposition levels n is dependent on the sampling coefficients (p, q) and calculated as the highest integer such that $(\frac{p}{q})^n \geq \frac{1}{4}$. Given the native size ratios of Table 3.1, Table 3.2 lists the corresponding native spatial resolutions for an original image at 4CIF resolution¹. It can be observed that the number of possible native spatial resolutions increases with the use of rational wavelet decompositions of dilation factors $q/p < 2$. However, apart from the original resolution, there is no other common resolution between the sets proposed by each (p, q) rational wavelet decomposition and the set of possible resolutions obtained from a dyadic wavelet transform. What is illustrated is that although a rational wavelet decomposition produces a richer set of possible native resolutions, this gain comes with a loss of the dyadic native resolutions that are traditionally expected from a wavelet-based image compression system.

In the current chapter, we present a two-dimensional wavelet decomposition structure made of rational wavelet transforms, and able to produce native resolutions obtained by dyadic factors. The proposed decomposition is a combination of rational

¹CIF=Common Intermediate Format. CIF resolution corresponds to a size 352×288 . 4CIF resolution corresponds to a size 704×576 . QCIF resolution corresponds to a size 176×144 .

wavelet transforms with different dilation factors, which results in a set of native size ratios that include dyadic and non-dyadic ratios. We address the issue of compatibility with subband coders of the zerotree coding family, a family of transform coders that are widely used in wavelet-based image compression. Our work introduces a simple compatibility scheme for the proposed wavelet decomposition, by defining a different way to build the *parent-child tree structure* used in zerotree-based subband coders. Finally, a proposition on how to combine dyadic wavelet decomposition and rational wavelet decomposition is formulated, and a compatibility scheme for zerotree-based subband coding is defined for the resulting subband decomposition.

3.1 Combined rational wavelet decompositions for increased spatial resolution scalability with dyadic and non-dyadic factors

In this section, we look at the problem of generating native size ratios of dyadic nature with wavelet transforms of rational dilation factors. Doing this with a single rational wavelet transform means finding the right sampling coefficients (p, q) such that the set of native size ratios $\{(\frac{p}{q})^m, m = 0, \dots, n\}$ include dyadic factors of the type $(\frac{1}{2})^j$. In other terms, it means finding the right integers p, q such that there exist two integers m_1 and m_2 for which the relationship $(\frac{p}{q})^{m_1} = (\frac{1}{2})^{m_2}$ is verified. With the constraint that q/p forms a rational value, no solution exists to that relationship. However, the problem can be stated differently if we combine multiple rational wavelet transforms of different dilation factors in the multi-level two-dimensional wavelet decomposition.

To illustrate the idea of combining rational wavelet transforms of different dilation factors, consider decomposing an original image using successively a two-dimensional

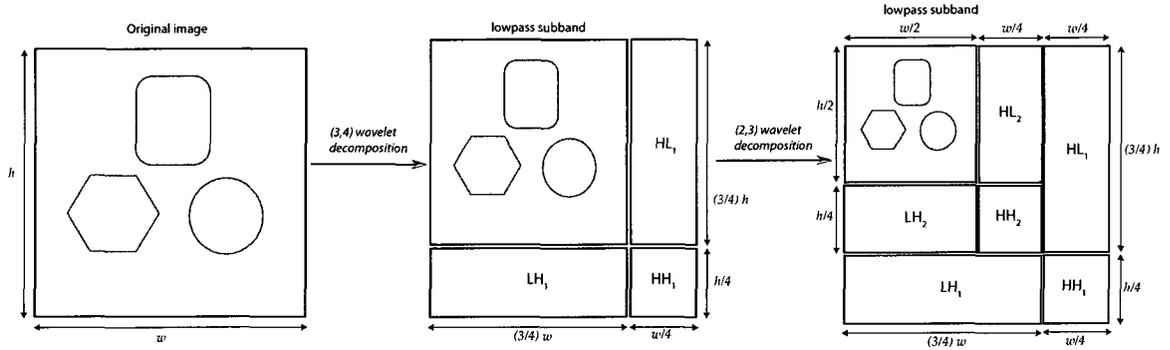


Figure 3.1: A combined rational wavelet decomposition with rational wavelet transforms of dilation factors $4/3$ and $3/2$.

$(3, 4)$ rational wavelet decomposition and a two-dimensional $(2, 3)$ rational wavelet decomposition, as follows:

1. The original image of size (w, h) is transformed using a two-dimensional $(3, 4)$ rational wavelet decomposition. The wavelet decomposition produces an approximation subband of size $(w_1, h_1) = (\frac{3}{4}w, \frac{3}{4}h)$ and three details subbands.
2. The approximation subband of size (w_1, h_1) is transformed using a two-dimensional $(2, 3)$ rational wavelet decomposition. The wavelet decomposition results in a new approximation subband of size $(w_2, h_2) = (\frac{2}{3}w_1, \frac{2}{3}h_1) = (\frac{1}{2}w, \frac{1}{2}h)$ and three additional details subbands.

This alternation of two-dimensional rational wavelet transforms of different dilation factors ($4/3$ and $3/2$ in the example) is what we denote a *combined* rational wavelet decomposition. Figure 3.1 illustrates the described combined rational wavelet decomposition, with rational wavelet transforms of sampling coefficients $(3, 4)$ and $(2, 3)$. Although the illustrated decomposition is made of two levels of wavelet decomposition, each with a different wavelet transform, we will often refer to it as *one* level of *combined* rational wavelet decomposition. To apply the combined decomposition²

²In the text, the expression *combined rational wavelet decomposition* will often be simplified to combined decomposition or combined wavelet decomposition

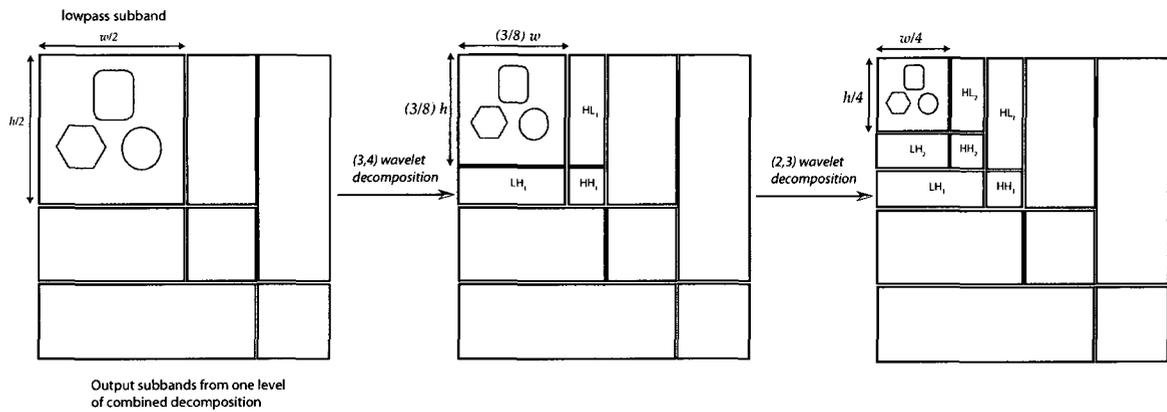


Figure 3.2: Illustration of the second level of combined decomposition with rational wavelet transforms of dilation factors $4/3$ and $3/2$.

on multiple levels, the described scheme is iterated on the approximate subband. Figure 3.2 illustrates a second level of the combined rational wavelet decomposition described in Figure 3.1. The second level of decomposition produces successively an approximation subband of size $(w_3, h_3) = (\frac{3}{8}w, \frac{3}{8}h)$, then $(w_4, h_4) = (\frac{1}{4}w, \frac{1}{4}h)$. The specific dilation factors of $4/3$ and $3/2$ are only used as example, and the notion of combined rational wavelet decomposition can be generalized to any set of rational wavelet transforms with different dilation factors. As a matter of fact, we characterize a combined rational decomposition by the set of dilation factors $\{(q_i/p_i), i = 1, \dots, j\}^3$ of the j different rational wavelet transforms used in the decomposition. *One* level of *combined* decomposition represents a cascade of two-dimensional wavelet decompositions with each of the different rational wavelet transforms involved in the combined decomposition. With this characterization, the examples illustrated in Figure 3.1 and Figure 3.2 represent respectively one level and two levels of combined wavelet decomposition with dilation factors $\{(q_i/p_i)\} = \{4/3, 3/2\}$, in that exact order.

The interesting aspect of the combined rational wavelet decomposition is the native spatial resolutions that are produced during the decomposition process. We can

³This will often be denoted simply as $\{(q_i/p_i)\}$

determine from Figure 3.1 and Figure 3.2 that a two-level combined decomposition with dilation factors $\{(q_i/p_i)\} = \{4/3, 3/2\}$ produces a set of native size ratios equal to $\{1, \frac{3}{4}, \frac{1}{2}, \frac{3}{8}, \frac{1}{4}\}$. As a generality, the set of native size ratios produced by n levels of combined rational wavelet decomposition with two rational wavelet transforms of dilation factors q_1/p_1 and q_2/p_2 is given by the union of the two partial sets:

- *set 1* = $\{(\frac{p_1}{q_1})^m (\frac{p_2}{q_2})^{m-1}, m = 1, \dots, n\}$, and
- *set 2* = $\{(\frac{p_1}{q_1})^m (\frac{p_2}{q_2})^m, m = 0, \dots, n\}$

For example, with $q_1/p_1 = 4/3$, $q_2/p_2 = 3/2$, and $n = 2$:

- *set 1* = $\{(\frac{3}{4})(\frac{2}{3})^0, (\frac{3}{4})^2(\frac{2}{3})^1\} = \{\frac{3}{4}, \frac{3}{8}\}$
- *set 2* = $\{(\frac{3}{4})^0(\frac{2}{3})^0, (\frac{3}{4})^1(\frac{2}{3})^1, (\frac{3}{4})^2(\frac{2}{3})^2\} = \{1, \frac{1}{2}, \frac{1}{4}\}$

The union of the two partial sets gives the complete set of native size ratios derived earlier. Similarly, if three different dilation factors are used in the combined decomposition, the possible sizes ratios are given by the union of the three partial sets:

- *set 1* = $\{(\frac{p_1}{q_1})^m (\frac{p_2}{q_2})^{m-1} (\frac{p_3}{q_3})^{m-1}, m = 1, \dots, n\}$,
- *set 2* = $\{(\frac{p_1}{q_1})^m (\frac{p_2}{q_2})^m (\frac{p_3}{q_3})^{m-1}, m = 1, \dots, n\}$, and
- *set 3* = $\{(\frac{p_1}{q_1})^m (\frac{p_2}{q_2})^m (\frac{p_3}{q_3})^m, m = 0, \dots, n\}$,

It can be observed that the set of native size ratios produced by two levels of combined decomposition with dilation factors $\{(q_i/p_i)\} = \{4/3, 3/2\}$ include dyadic and non-dyadic ratios. The inclusion of dyadic ratios was made possible by appropriately selecting the dilation factors $\{(q_i/p_i)\}$ of the wavelet transforms involved in the combined rational wavelet decomposition. Indeed, from the formulation above, the partial set of native size ratios defined by $\{\prod_i (p_i/q_i)^m, m = 0, \dots, n\}$ is always

$\{(q_i/p_i)\} = \{(4/3), (3/2)\}$	$\{(q_i/p_i)\} = \{6/5, 5/4, 4/3\}$
1	1
3/4	5/6
1/2	2/3
3/8	1/2
1/4	5/12
	1/3
	1/4

Table 3.3: Native size ratios from a two-level combined rational wavelet decomposition, with different sets of dilation factors.

Dyadic decomposition	$\{(q_i/p_i)\} = \{(4/3), (3/2)\}$	$\{(q_i/p_i)\} = \{6/5, 5/4, 4/3\}$
704×576	704×576	704×576
352×288	528×432	587×480
176×144	352×288	470×384
	264×216	352×288
	176×144	294×240
		235×192
		176×144

Table 3.4: Native spatial resolutions from an original image of size 704×576 when using a two-level combined rational decomposition against a regular dyadic decomposition. Decimal sizes are rounded to integer values.

included among the possible native size ratios from an n -level combined decomposition. Specifically, it corresponds to *set 2* in the case of two dilation factors and *set 3* when three different dilation factors are involved in the combined decomposition. Considering that the set of possible native size ratios achievable with an n -level *dyadic* decomposition is given by $\{(1/2)^m, m = 0, \dots, n\}$, the combined decomposition will be able to produce native spatial resolutions obtained by dyadic factors if:

$$\prod_i (q_i/p_i) = 2 \quad (3.1)$$

This relationship is verified for a combined rational wavelet decomposition with dilation factors $\{(q_i/p_i)\} = \{4/3, 3/2\}$, which it's why we are able to obtain native resolutions of dyadic factors from this specific combined decomposition.

Table 3.3 lists the possible size ratios from a two-level combined rational wavelet decomposition with the combination of dilation factors $4/3$ and $3/2$ on one hand, and

dilation factors $6/5$, $5/4$ and $4/3$ on the other hand. Note that in both cases the dilation factors satisfy the condition cited in (3.1). Table 3.4 indicates the native spatial resolutions produced by each combined decomposition for an original image at 4CIF resolution. The native resolutions produced by a dyadic wavelet decomposition are also indicated for comparison. As demonstrated, the two combined rational wavelet decompositions produce a richer set of possible resolutions, *and* include native resolutions obtained by dyadic factors.

3.2 A parent-child tree construction scheme for compatibility with zerotree-based subband coders

Subband coders of the zerotree coding family [20, 33] have been widely used in the image coding literature. Said and Pearlman's SPIHT (Set Partitioning In Hierarchical Trees) [20], an example of zerotree-based subband coder, is considered one of the state-of-the-art coders in wavelet-based image compression thanks to the relative simplicity of the coding algorithm along with excellent reported performance. Likewise, we have a keen interest in investigating the potential use of zerotree-based subband coders to encode the approximation and details subbands from the proposed combined rational wavelet decomposition. As with many other subband coding techniques targeted for image applications, zerotree-based subband coders were initially developed for a dyadic wavelet decomposition and take advantage of the dyadic splitting of the approximation subband across consecutive levels of decomposition. A difficulty naturally arises in trying to adapt the subband coding scheme to support a (p, q) rational wavelet transform that performs a different partitioning of the two-dimensional space.

Zerotree-based subband coders exploit the correlation between subband coefficients at the same spatial location across different levels of wavelet decomposition. At the core of the technique employed is the construction of a parent-child tree structure that associates every pixel at a higher level of decomposition (the parent node) to a group of pixels at the next finer (or smaller) level (the children nodes). In the case of a two-dimensional dyadic decomposition, every parent node has exactly four children nodes grouped as a 2×2 block of pixels. To build the complete tree structure, every node is considered a potential parent node until the remaining nodes all belong to the finest level of decomposition. Figure 3.3 illustrates the construction of the parent-child tree structure for a three-level dyadic wavelet decomposition. A pixel at the highest level of decomposition is a parent node of a 2×2 block of pixels in the next finer level, which in turn are the parent nodes of a 4×4 block of pixels in the finest level of decomposition. If every pixel in the details subbands at the highest level is considered a root node, then the complete tree structure can be built. To extend that scheme to any (p, q) rational wavelet decomposition, one can generalize the rule by considering an $m \times m$ block of pixels as the parent nodes of a $(\frac{q}{p} \times m) \times (\frac{q}{p} \times m)$ block of pixels in the next finer level of decomposition. For the scheme to work, the value $(\frac{q}{p} \times m)$ must be kept as an integer value, something difficult to achieve with a rational dilation factor q/p when many levels of decomposition are performed. For example, Figure 3.4 displays an attempt to use that method in a case of a three-level rational wavelet decomposition with a wavelet transform of dilation factor $3/2$. Starting with a block of 2×2 pixels at the highest level ($m = 2$), the ratio $(\frac{q}{p} \times m)$ gradually increases to $\frac{3}{2} \times 2 = 3$ for the next finer level, then $\frac{3}{2} \times 3 = \frac{9}{2}$ for the finest level. With a non-integer value, the size of the child nodes block cannot be determined for the finest level of decomposition. In order to complete the tree structure, a different method would have to be defined.

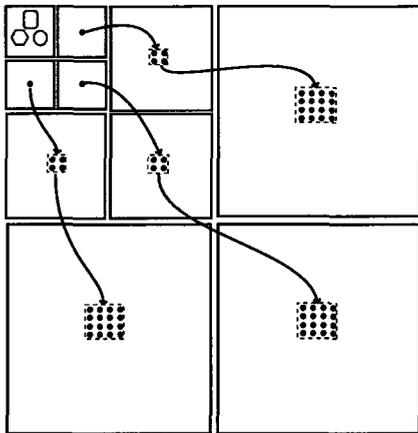


Figure 3.3: Construction of the parent-child tree structure for a dyadic wavelet decomposition.

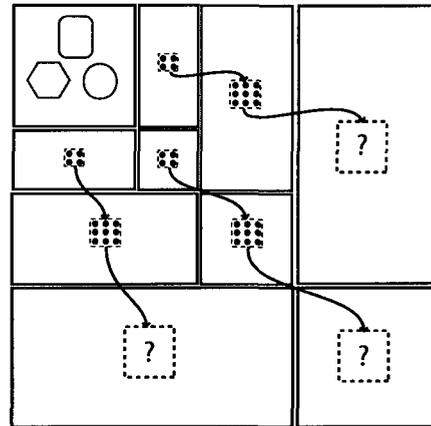


Figure 3.4: An illustration of the difficulty of completing the parent-child tree structure for a $(2,3)$ rational wavelet decomposition.

The difficulty in building a complete parent-child tree structure for subband decompositions obtained from rational wavelet transforms originates from the non-integer nature of the dilation factor q/p . The dilation factor dictates the *size increase ratio* between the size of a subband at a higher level of decomposition and the size of the corresponding subband at the next finer level. A non-integer size increase ratio inevitably presents a challenge for the construction of the parent-child tree. The combined rational wavelet decomposition introduced in Section 3.1 does not produce a constant size increase ratio across successive levels of wavelet decomposition, because of the different dilation factors of the wavelet transforms. While this seems to be an additional complexity, we demonstrate in the current section how this property can be taken advantage of.

Consider for example the combined wavelet decomposition with dilation factors $\{(q_i/p_i)\} = \{4/3, 3/2\}$ used as example in Section 3.1. Let's denote by LH_k , HL_k and HH_k the respective horizontal, vertical and diagonal highpass subbands obtained from a two-dimensional (p_k, q_k) rational wavelet transform performed within one level

Subband	LL	LH_2^m	HL_2^m	HH_2^m	LH_1^m	HL_1^m	HH_1^m
width	$w_m = (\frac{1}{2})w_{m-1}$	$(\frac{1}{2})w_{m-1}$	$(\frac{1}{4})w_{m-1}$	$(\frac{1}{4})w_{m-1}$	$(\frac{3}{4})w_{m-1}$	$(\frac{1}{4})w_{m-1}$	$(\frac{1}{4})w_{m-1}$
height	$h_m = (\frac{1}{2})h_{m-1}$	$(\frac{1}{4})h_{m-1}$	$(\frac{1}{2})h_{m-1}$	$(\frac{1}{4})h_{m-1}$	$(\frac{1}{4})h_{m-1}$	$(\frac{3}{4})h_{m-1}$	$(\frac{1}{4})h_{m-1}$

Table 3.5: Sizes of the subbands produced at the m -th level of a combined decomposition with dilation factors $\{(q_i/p_i)\} = \{4/3, 3/2\}$. The original image size is $(w, h) = (w_0, h_0)$.

of combined decomposition, and by LL the approximation subband obtained from the complete combined decomposition. For details subbands produced when multiple levels of combined decomposition are performed, we use the terminology LH_k^m , HL_k^m and HH_k^m , to indicate the subbands produced after m levels of combined decomposition. The sizes of each subband produced at the m -th level of combined decomposition are indicated in Table 3.5.

In a multi-level combined wavelet decomposition, a natural way to build the parent-child tree would be to associate parent nodes in a given detail subband to a set of children nodes in the corresponding detail subband from the previous rational wavelet decomposition. This type of subband association is illustrated in Figure 3.5 for a two-level combined decomposition. In the illustrated scheme, a node in subband LH_2^2 would have children nodes in subband LH_1^2 , which themselves would be parents to a set of nodes in subband LH_2^1 . Looking at the sizes of the subbands as indicated by Table 3.5, the size increase ratio between subbands LH_2^2 and LH_1^2 is $3/2$ along the width and 1 along the height; the same ratio between subbands LH_1^2 and LH_2^1 is $4/3$ along the width and 2 along the height. The described subband association produces non-integer size increase ratios along the width, which results in the same challenge introduced earlier with the use of a single rational wavelet decomposition.

Instead, we propose to build the parent-child tree structure by associating any subband \mathcal{S}_k^m at the m -th level of decomposition to the corresponding subband \mathcal{S}_k^{m-1} at the next finer level of combined decomposition $m - 1$. The resulting subband

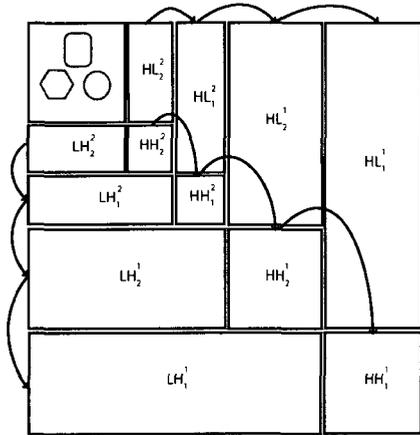


Figure 3.5: Subband association in a regular parent-child tree structure.

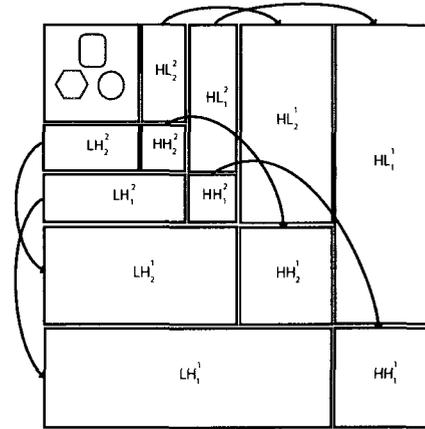


Figure 3.6: Subband association in the proposed parent-child tree structure.

association is illustrated in Figure 3.6. A node in subband LH_2^2 is parent to a set of children nodes in subband LH_2^1 . Similarly, a node in subband LH_1^2 is parent to children nodes in subband LH_1^1 . The size increase ratio for both pairs of subbands becomes 2 along the width and 2 along the height, all integer values.

Figure 3.7 compares the construction of the parent-child tree structure in a dyadic case against the proposed construction for a combined rational wavelet decomposition. In both cases, one pixel node is the parent of four nodes at a finer level of decomposition. The main difference is that for a dyadic decomposition the trees are rooted in the *three* details subbands of the highest level of wavelet decomposition, whereas in the case of a combined decomposition with two different wavelet transforms, the trees are rooted in the *six* details subbands of the highest level of *combined* wavelet decomposition.

The integer size increase ratio obtained by using the proposed parent-child tree structure is a consequence of the dilation factors of the rational wavelet transforms involved in the combined decomposition. Indeed, it can be determined that the size increase ratio between any subband \mathcal{S}_k^{m+1} and the corresponding subband \mathcal{S}_k^m at

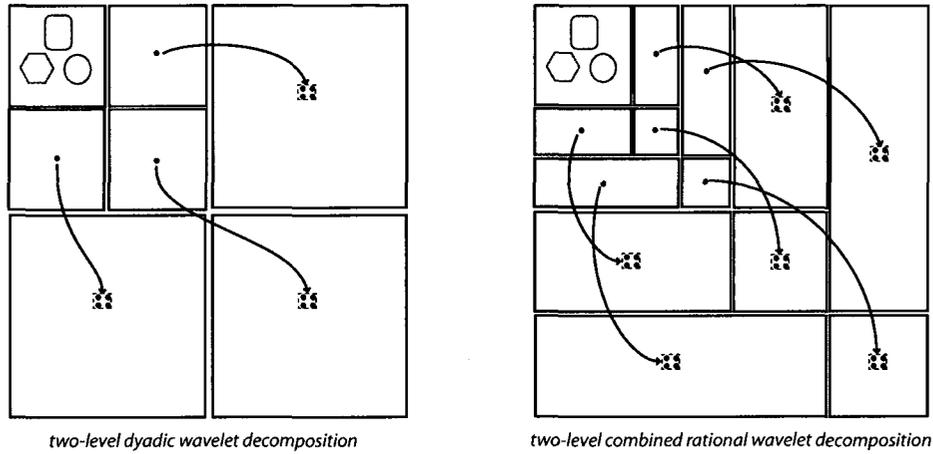


Figure 3.7: Comparison between a regular parent-child tree construction scheme in a dyadic wavelet decomposition and the proposed parent-child tree construction scheme for a combined rational wavelet decomposition.

the next finer level is equal to the ratio w_m/w_{m+1} (respectively h_m/h_{m+1}) between the width (resp. height) of the approximation subband after m levels of combined decomposition and the width (resp. height) of the approximation subband after $m + 1$ levels. For a combined rational wavelet decomposition with dilation factors coefficients $\{(q_i/p_i)\}$, the sizes of the approximation subband across consecutive levels of combined decomposition are related by :

- $w_{m+1} = \prod_i \left(\frac{p_i}{q_i}\right) w_m$, and
- $h_{m+1} = \prod_i \left(\frac{p_i}{q_i}\right) h_m$

An integer size increase ratio is kept along the width and height if the dilation factors are selected such that $\prod_i (q_i/p_i)$ is an integer value. Furthermore , if $\prod_i (q_i/p_i) = 2$ as suggested in Section 3.1, then a size increase ratio of 2 is achieved, just like in a regular dyadic wavelet decomposition.

In this section, a combined wavelet decomposition with two dilation factors has been used only as an example. The proposed method to build the parent-child tree structure can be extended to any number of dilation factors in the combined wavelet decomposition. If a total number of j different rational wavelet transforms are used,

the parent-child tree construction scheme can be implemented with trees rooted in the $3 \times j$ details subbands of the highest level of combined decomposition. The only constraint on the dilation factors is that they should be chosen such that their cumulative product $\prod_i (q_i/p_i)$ is an integer value.

3.3 Combining dyadic wavelet decomposition and combined rational wavelet decomposition

In this section, we are interested in using a dyadic wavelet decomposition in conjunction with a combined rational wavelet decomposition. The idea is to follow n_1 levels of combined wavelet decomposition with n_2 levels of two-dimensional dyadic wavelet transform. This is illustrated in Figure 3.8 for $n_1 = 2$, $n_2 = 2$, and a combined wavelet decomposition with dilation factors $\{(q_i/p_i)\} = \{4/3, 3/2\}$. The motive behind this decomposition is to use the dyadic wavelet transform to increase the overall compression performance by taking advantage of the rich base of wavelet filters used in image coding with dyadic wavelet decompositions. The performance results of this structure will be analyzed and discussed in the chapters ahead; in the present chapter we only focus on defining a way to adapt the parent-child tree construction scheme proposed in Section 3.2 to a two-dimensional decomposition made of dyadic and combined rational wavelet transforms.

In Section 3.2, the parent-child tree structure for the combined wavelet decomposition is built by rooting the parent-child trees in the *six* highpass subbands at the highest level of combined decomposition. In the subband decomposition of Figure 3.8, however, the highest level of decomposition is obtained from a two-dimensional dyadic wavelet decomposition that produces *three* highpass subbands. Obviously, the method

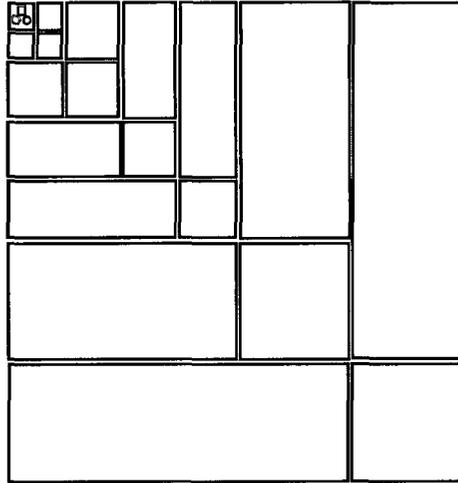


Figure 3.8: Illustration of the use of dyadic decomposition in conjunction with combined decomposition: $n_1 = 2$ levels of combined decomposition are followed by $n_2 = 2$ levels of dyadic decomposition.

described in Section 3.2 cannot be applied in that context. In the rest of this section, we define a different technique to build the complete parent-child tree structure, inspired from wavelet packet decompositions. In contrast to the method presented in Section 3.2 which can be extended to different combination of rational wavelet transforms, the proposed construction scheme is very specific to a combined decomposition with dilation factors $\{(q_i/p_i)\} = \{4/3, 3/2\}$. We will justify that restriction later in the section.

3.3.1 An extension of the combined rational wavelet decomposition to a wavelet packet decomposition

For the rest of the section, we'll consider specifically the combined rational wavelet decomposition with dilation factors $\{(q_i/p_i)\} = \{4/3, 3/2\}$. The terminology introduced in Section 3.2 will be used to denote the different subbands produced by the wavelet decomposition. We propose to extend the two-dimensional structure obtained

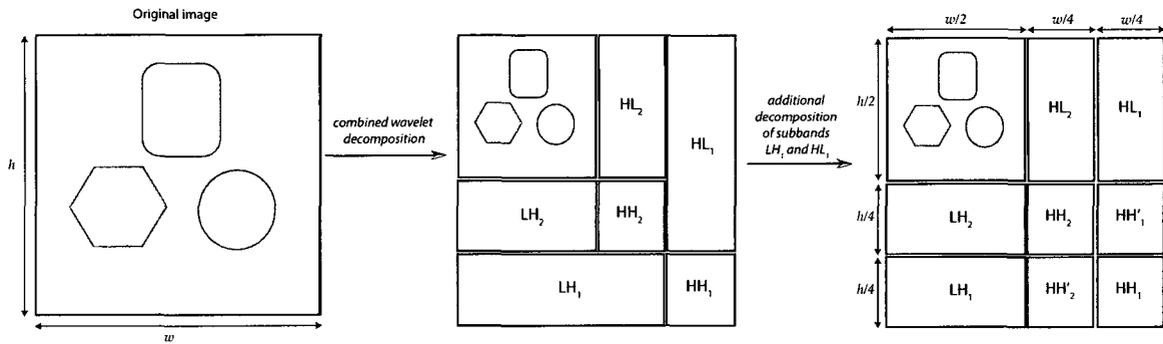


Figure 3.9: Illustration of the wavelet packet extension of the combined rational wavelet decomposition with dilation factors $\{(q_i/p_i)\} = \{4/3, 3/2\}$.

after one level of combined decomposition in the following manner:

1. Perform a one-dimensional $(2, 3)$ rational wavelet transform along the columns of the vertical subband HL_1 .
2. Perform a one-dimensional $(2, 3)$ rational wavelet transform along the rows of the horizontal subband LH_1 .

This produces two additional diagonal subbands HH'_1 and HH'_2 as illustrated in Figure 3.9. In the specific case of a combined decomposition with dilation factors $\{(q_i/p_i)\} = \{4/3, 3/2\}$, the two additional subbands HH'_1 and HH'_2 have the same sizes as the diagonal subbands HH_1 and HH_2 . The resulting structure is very similar to a type of decomposition traditionally referred to as a *wavelet packet decomposition*. A wavelet packet decomposition designates an extension of the two-dimensional dyadic wavelet decomposition where each highpass subband can be further split by the means of one or two-dimensional dyadic wavelet transforms to create alternative two-dimensional decomposition structures. For example, a decomposition structure similar to the one obtained in Figure 3.9 can be generated by the following wavelet packet decomposition:

1. Perform a two-dimensional dyadic wavelet decomposition on the original image

2. Perform a one-dimensional dyadic wavelet decomposition along the columns of the horizontal subband LH .
3. Perform a one-dimensional dyadic wavelet decomposition along the rows of the vertical subband HL .
4. Perform a two-dimensional dyadic wavelet decomposition on the diagonal subband HH .

For clarity, wavelet packet decompositions generated from dyadic wavelet transforms will be appended the mention *dyadic*. The decomposition illustrated in Figure 3.9 will be qualified as the *wavelet packet extension* of the combined rational wavelet decomposition. Dyadic wavelet packet decompositions have been often used in image coding for their efficiency with highly textured images [34, 35]. They have received a good interest in the image coding literature since the support for alternative decomposition structures was included in Part 2 of the JPEG2000 image coding standard. Several existing frameworks and subband coders developed for wavelet-based image compression naturally offer support for dyadic wavelet packet decompositions and the use of zerotree-based coders have been investigated in that context as well [36, 37].

With its particular structure, the decomposition illustrated in Figure 3.9 offers more flexibility than the regular combined decomposition, with respect to compatibility with existing wavelet-based image coding frameworks or subband coders. However, as defined, the wavelet packet extension of the combined decomposition is only similar to a dyadic wavelet packet decomposition when the dilation factors are specifically set to $4/3$ and $3/2$. Indeed, two other dilation factors, or even the same dilation factors used in reverse order, would produce details subbands of different sizes than what a dyadic wavelet packet decomposition would produce. Finally, it can be verified that a combined decomposition in wavelet packet extension mode is fully compatible with the parent-child tree construction scheme described in Section 3.2. In the case

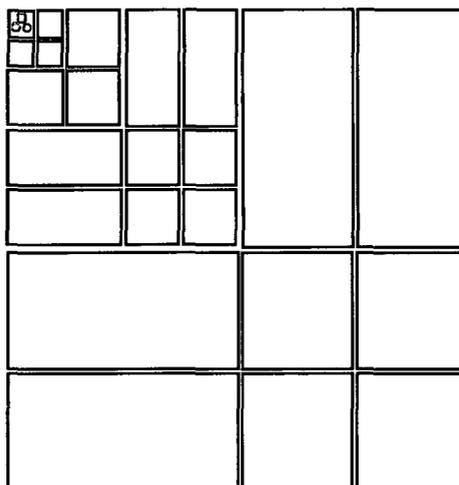


Figure 3.10: Illustration of the use of dyadic decomposition in conjunction with combined decomposition in wavelet packet extension mode: $n_1 = 2$ levels of combined decomposition are followed by $n_2 = 2$ levels of dyadic decomposition.

of two dilation factors, the parent-child trees are rooted in the *eight* details subbands at the highest level of combined decomposition (all levels of decomposition should be performed using the wavelet packet extension).

3.3.2 Combining dyadic wavelet decomposition and combined rational wavelet decomposition with the defined wavelet packet extension

Figure 3.10 illustrates the subbands obtained when $n_1 = 2$ levels of combined decomposition in wavelet packet extension mode are followed by $n_2 = 2$ levels of two-dimensional dyadic wavelet transform. We are interested in building a parent-child tree structure across the different subbands by taking advantage of the similarity with two-dimensional dyadic wavelet packet decompositions. As we pointed out earlier, the use of zerotree-based subband coders has been considered for dyadic wavelet

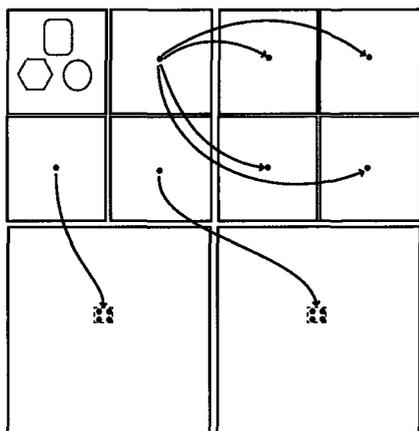


Figure 3.11: An illustration of the parent-child tree construction scheme in a dyadic wavelet packet decomposition structure, as suggested by Kim *et al.*

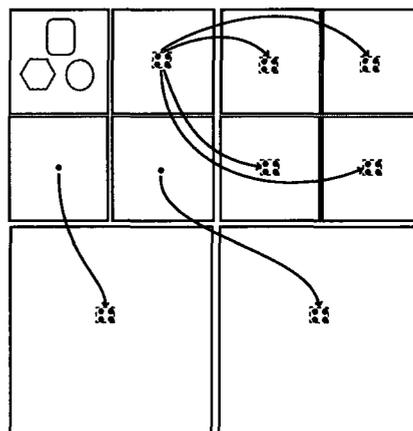


Figure 3.12: A slight variant of the parent-child tree construction scheme for a dyadic wavelet packet decomposition structure.

packet decompositions, by redefining the way the parent-child tree structure is constructed [36,37]. Kim *et al.* [36] propose to build the trees in the manner described in Figure 3.11. In the illustrated example, the dyadic wavelet packet decomposition consists of two levels of two-dimensional dyadic decomposition, plus a two-dimensional dyadic wavelet transform to the vertical highpass subband from the first level of decomposition. To build the parent-child tree structure on the vertical highpass subbands, the children nodes from one parent at the second level of decomposition are split in each of the quadrants obtained from decomposing the first-level subband. We elect to add a slight change to this approach, similar to a strategy used by Sprljan *et al.* [37]: instead of splitting the children nodes from one parent in each of the four quadrants, parent nodes are grouped in 2×2 blocks and each parent node is defined to have all four children in one quadrant, as illustrated in Figure 3.12.

The same idea can be transposed to the subband decomposition of Figure 3.10 and used to build the parent-child tree at the transition between a subband from the dyadic decomposition and subbands from the combined wavelet decomposition.

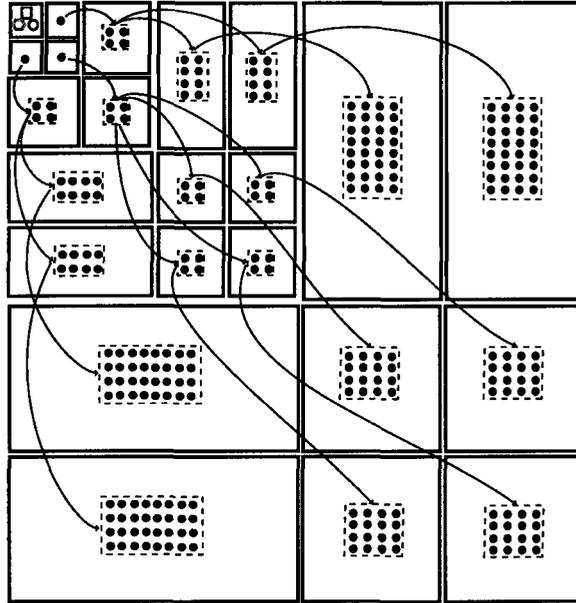


Figure 3.13: An illustration of the proposed parent-child tree construction scheme when a dyadic decomposition is used in conjunction with a combined decomposition.

Before that transitional step, the parent-child tree structure is built as with any two-dimensional dyadic wavelet decomposition. Past that transitional step, the rest of the parent-child tree structure is completed using the technique presented in Section 3.2 for combined rational wavelet decompositions. Figure 3.13 is an illustration of the proposed parent-child tree construction scheme, illustrated for a subband decomposition consisting of two levels of combined wavelet decomposition and two levels of dyadic wavelet decomposition.

The proposed technique is designed specifically for a combined wavelet decomposition with dilation factors $\{(q_i/p_i)\} = \{4/3, 3/2\}$. Although the definition of the wavelet packet extension mode of the combined decomposition can be generalized to any set of dilation factors, only this specific selection of dilation factors can create a structure similar to a dyadic wavelet packet decomposition. This similarity is exploited in the parent-child tree construction scheme.

3.4 Compatibility of the proposed decompositions with other existing subband coders

By offering a different partitioning of the two-dimensional space than the dyadic wavelet decomposition, the combined rational wavelet decomposition introduced in this chapter naturally presents a challenge for subband coders traditionally used in wavelet-based image coding. In Section 3.2 and Section 3.3, we have addressed this challenge for the particular family of zerotree-based subband coders, by redefining the parent-child tree structure to adapt to the sizes of the subbands produced. Intra-subband coders like SPECK (Set Partitioning Embedded bloCK coder) [21], EZBC (Embedded coding using ZeroBlocks of wavelet coefficients and Context modeling) [38] and EBCOT (Embedded Block Coding with Optimized Truncation) [39] offer an alternative to zerotree-based coding. By working only within a given subband and treating each subband independently, these coders allow for more flexibility in the sizes of the subbands produced by the wavelet decomposition, since there's no need for a cross-level subband relationship. SPECK, for example, uses a subband split partitioning technique in order to exploit the local correlation of coefficients within the subband, and does not require any cross-subband processing.

The wavelet packet extension described in Section 3.3 for a combined rational wavelet decomposition with dilation factors of $4/3$ and $3/2$ offers an even greater compatibility with existing wavelet-based image coding frameworks. Several subband coders in the image coding literature offer a natural support for dyadic wavelet packet decomposition, including the ones listed above. Very often, efficient implementations that are optimized to take advantage of the half-splitting inherent to a dyadic decomposition show support for dyadic wavelet packet decompositions as well.⁴

⁴The popular *kakadu* software for JPEG2000's EBCOT is a good example.

Chapter 4

An algorithm structure for spatial resolution scalability with SPIHT and SPECK

In this chapter, we introduce two popular subband coders used in wavelet-based image coding : SPIHT [20], a zerotree-based subband coder widely-used in the image coding literature, and SPECK [21], a more recent coder that exploits intra-subband correlation. We are interested in both wavelet subband coders for potential use in the construction of a wavelet-based image compression system with spatial resolution scalability. The two subband coders use a *bitplane coding* technique, where the subband coefficients are arranged into bitplanes and the coding algorithm processes the data one bitplane after another. As originally defined, SPIHT and SPECK produce an encoded bitstream with fine *Signal-to-Noise-Ratio (SNR) scalability*. That is, a bitstream in which the data is arranged by increasing layers of quality, as illustrated in Figure 4.1.

In their respective original forms, neither SPIHT or SPECK naturally offers the

<i>layer 1</i>	<i>layer 2</i>	<i>layer 3</i>	...
----------------	----------------	----------------	-----

Figure 4.1: An example of bitstream with SNR scalability.

<i>resolution 1</i>	<i>resolution 2</i>			<i>resolution 3</i>
	<i>layer 1</i>	<i>layer 2</i>	<i>layer 3</i>	...

Figure 4.2: An illustration of bitstream with spatial resolution scalability and SNR scalability within each resolution.

capability to produce a *resolution-scalable* bitstream for the support of spatial resolution scalability. Various works in the literature have proposed algorithmic modifications to both SPIHT [40–42] and SPECK [43, 44] in order to produce bitstreams with scalability on spatial resolution. The typical structure of the resulting bitstreams is shown in Figure 4.2: the bitstream is scalable in resolution, and also scalable in quality within each resolution. However, the different contributions in the current literature do not attempt to specify a different *bitrate* for every resolution included in the bitstream during the encoding process. In other words, they do not attempt to control the proportion of the bitstream allocated by the subband coder to each spatial resolution. Instead, the use of a bitstream parser equipped with a rate allocation model is typically suggested to extract a partial bitstream from the original encoded bitstream, if a bitstream of specific length is required for a specific resolution.

A reason for that is the fact that the previously cited works target the inclusion of fine SNR scalability and full spatial resolution scalability in the bitstream. The use of a bitstream parser with a rate allocation model is indeed a very efficient way to support both types of scalability, if the complexity of the rate allocation model can be afforded [40, 42]. In the current chapter, we present a simple way of specifying a bitrate for each resolution directly to the subband coder, by appropriately modifying

the structure of the coding algorithm. The next sections describe the algorithmic manipulations performed to achieve this goal. The resulting algorithm structures for SPIHT and SPECK allow for each resolution to be specified a target bitrate directly in the encoding process, so that the lengths of the different portions of the resolution-scalable bitstream produced by the subband coder can be controlled.

4.1 Converting SPIHT to support spatial resolution scalability

Said and Pearlman's SPIHT (Set Partitioning In Hierarchical Trees) [20] is one the most widely-used wavelet subband coder in the image compression literature. Based on the zerotree subband coding paradigm introduced by Shapiro [33], SPIHT is an inter-scale coder which exploits the relationship across wavelet subband coefficients at consecutive decomposition levels. The algorithm does it by constructing a parent-child tree structure rooted at the highest wavelet decomposition level, and by pruning the tree's branches, bitplane after bitplane, in search of *significant* coefficients. The relative simplicity of the algorithm along with its very high efficiency have made SPIHT a benchmark in wavelet subband coding.

4.1.1 An overview of the coding algorithm

Figure 4.3 gives a very simplistic general overview of the SPIHT coding algorithm. The algorithm is only described from a top-level perspective for the sake of simplicity, and the interested reader is referred to the work by Said and Pearlman [20] for additional details. The algorithm uses three lists: the List of Insignificant Pixels (LIP), the List of Insignificant Sets (LIS) and the List of Significant Pixels (LSP).

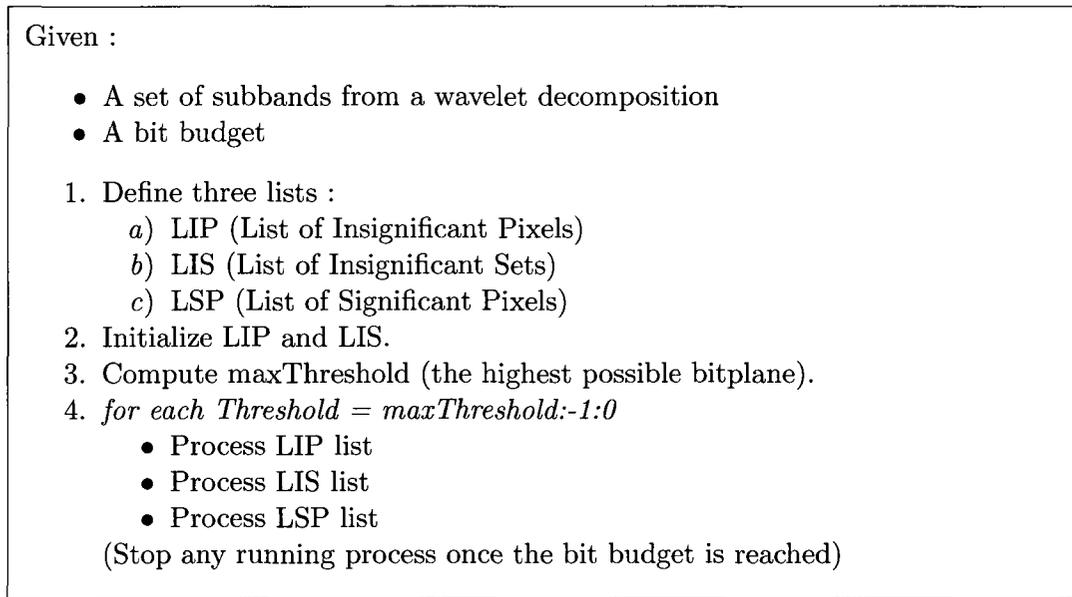


Figure 4.3: A simplistic overview of the SPIHT coding algorithm.

The LIS list is used to build and prune the complete parent-child tree structure. A set in the LIS represents a subtree of subband coefficients rooted at one higher level of decomposition down to the finest level. A pixel is either part of a set in the LIS or member of the LIP or LSP list. The algorithm processes every list, bitplane after bitplane, eventually moving the pixels from the sets in the LIS to the LIP and LSP lists, where their actual values are encoded.

4.1.2 Conversion to a subband coder supporting spatial resolution scalability

Previous works by Danyali and Mertins [40] and by Christophe and Pearlman [42] have proposed the use of *resolution-dependent* lists, in order to modify the coding algorithm of SPIHT into one that produces a resolution-scalable bitstream. Recall that a n -level dyadic wavelet decomposition produces $n + 1$ native spatial resolutions (including the full resolution). The idea of resolution-dependent lists is to define $n + 1$

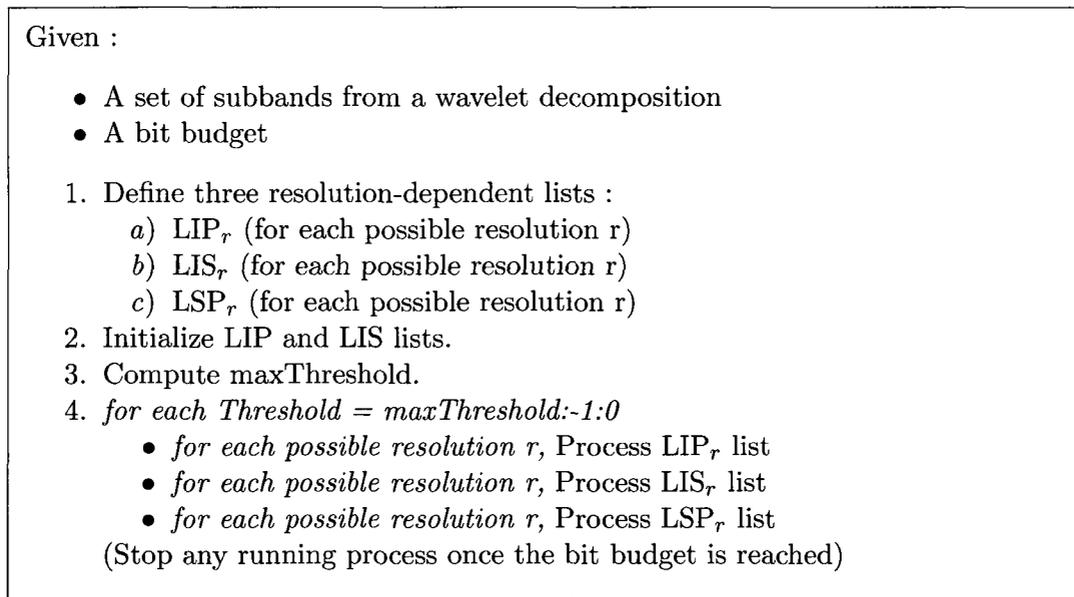


Figure 4.4: SPIHT coding algorithm with resolution-dependent lists, producing a bitstream with SNR scalability.

lists LIP_k , LIS_k and LSP_k , with $k = 0, \dots, n$. With the use of resolution-dependent lists, the coding algorithm can be re-expressed in the form described in Figure 4.4 to produce the same SNR-scalable bitstream. A simple change to the order of the loops at step 4 gives the algorithm shown in Figure 4.5, which produces a resolution-scalable bitstream. The interested reader is referred to Christophe and Pearlman's work [42] for details on the coding algorithms using resolution-dependent lists.

The algorithm described in Figure 4.5 produces a resolution-scalable bitstream with all possible resolutions included in the bitstream and one single bit budget specified: the bit budget for the entire bitstream. This results in the type of bitstream structure illustrated in Figure 4.6. We propose to change the structure of the algorithm in order to add the two following capabilities:

- a) the capability to decide which resolutions should be included in the bitstream.
- b) the capability to specify a bit budget for each resolution.

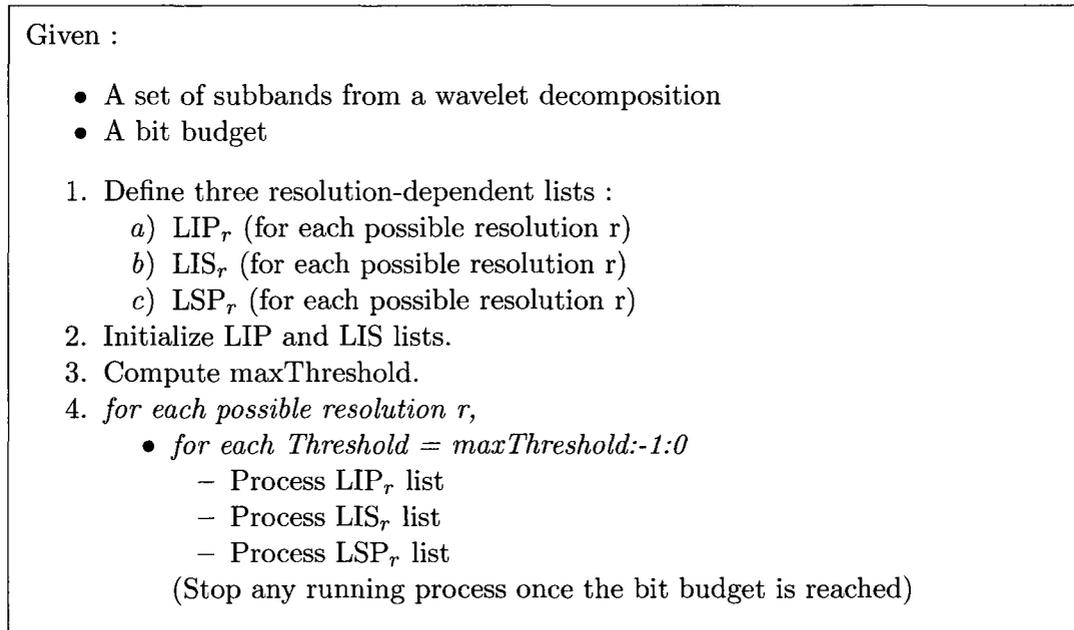


Figure 4.5: SPIHT coding algorithm with resolution-dependent lists, producing a bit-stream with resolution scalability.

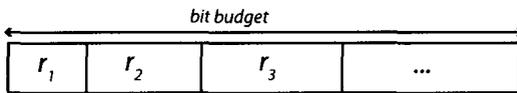


Figure 4.6: A resolution-scalable bit-stream with all possible resolutions and one bit budget specified for the full bit-stream.

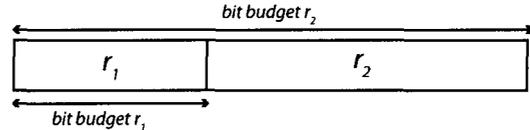


Figure 4.7: A resolution-scalable bit-stream with two resolutions of interest and a bit budget specified for each portion of the bitstream.

In other terms, we want to be able to specify which spatial resolutions are of interest to us and the bitrate at which we want these resolutions to be encoded. Figure 4.7 illustrates the type of flexibility targeted on the bitstream structure, represented for two spatial resolutions of interest.

In order to generate such a bitstream, we propose an algorithm structure derived from the one presented in Figure 4.4 to produce a SNR-scalable bitstream using resolution-dependent lists. At step 4 of the algorithm of Figure 4.4, the lists corresponding to all possible resolutions are processed under one threshold loop. If instead,

this loop is separated in two threshold loops such that the first loop only processes the spatial resolutions smaller or equal to a given resolution r_1 , and the second loop processes the resolutions between r_1 and a higher resolution r_2 , this would result in a bitstream with spatial resolution scalability on the two resolutions r_1 and r_2 . Additionally, if a mechanism is put in place to control the bit budget allocated to each spatial resolution, then the target bitstream illustrated in Figure 4.7 can be formed. Figure 4.8 describes the proposed structure of the coding algorithm, presented for two spatial resolutions of interest. The resulting bitstream also has the property of SNR scalability within each spatial resolution r_1 and r_2 . The described algorithm structure can be straightforwardly extended to a higher number of spatial resolutions of interest.

4.2 Converting SPECK to support spatial resolution scalability

SPECK (Set Partitioning Embedded bloCK coder) [21] belongs to a more recent family of subband coders which exploits the local clustering of coefficients within a wavelet subband by using a split partitioning technique. Other coders of the same family include BISK (BInary Set splitting with K-d trees) [45] and most notably EZBC [38], a SPECK variant, on which the successful MC-EZBC video coding framework [46, 47] is based. The coding algorithms of SPECK and SPIHT share notable similarities, though SPECK does not use any cross-subband processing and consequently does not use any cross-subband tree structure. Instead, SPECK uses a quad-tree partitioning technique, where each subband is split into four subsets and every subset can be further split into four subsets again, and so on and so forth. The algorithm uses

Given :

- A set of subbands from a wavelet decomposition
 - A bit budget for the resolution r_1 ($bitBudget_{r_1}$)
 - A bit budget for the resolution r_2 ($bitBudget_{r_2}$)
1. Define three resolution-dependent lists :
 - a) LIP_r (for each possible resolution r)
 - b) LIS_r (for each possible resolution r)
 - c) LSP_r (for each possible resolution r)
 2. Initialize LIP and LIS lists.
 3. Compute $maxThreshold$. Set $maxBudget = bitBudget_{r_2}$.
 4. for each $Threshold = maxThreshold:-1:0$
 - for each resolution r , $1 \leq r \leq r_1$, Process LIP_r list
 - for each resolution r , $1 \leq r \leq r_1$, Process LIS_r list
 - for each resolution r , $1 \leq r \leq r_1$, Process LSP_r list
 - if $bitBudget_{r_1}$ is reached or even exceeded,
 - set $minThreshold_{r_1} = Threshold$ and exit the loop
 (Stop any running process if $maxBudget$ is reached)
 5. for each $Threshold = maxThreshold : -1 : minThreshold_{r_1}$
 - for each resolution r , $r_1 < r \leq r_2$, Process LIP_r list
 - for each resolution r , $r_1 < r \leq r_2$, Process LIS_r list
 - for each resolution r , $r_1 < r \leq r_2$, Process LSP_r list
 (Stop any running process if $maxBudget$ is reached)
 6. if $bitBudget_{r_2}$ is not yet reached,
 - for each $Threshold = minThreshold_{r_1} - 1 : -1 : 0$
 - for each resolution r , $1 \leq r \leq r_2$, Process LIP_r list
 - for each resolution r , $1 \leq r \leq r_2$, Process LIS_r list
 - for each resolution r , $1 \leq r \leq r_2$, Process LSP_r list
 (Stop any running process once $bitBudget_{r_2}$ is reached)

Figure 4.8: Proposed SPIHT coding algorithm structure with support for spatial resolution scalability on two resolutions of interest.

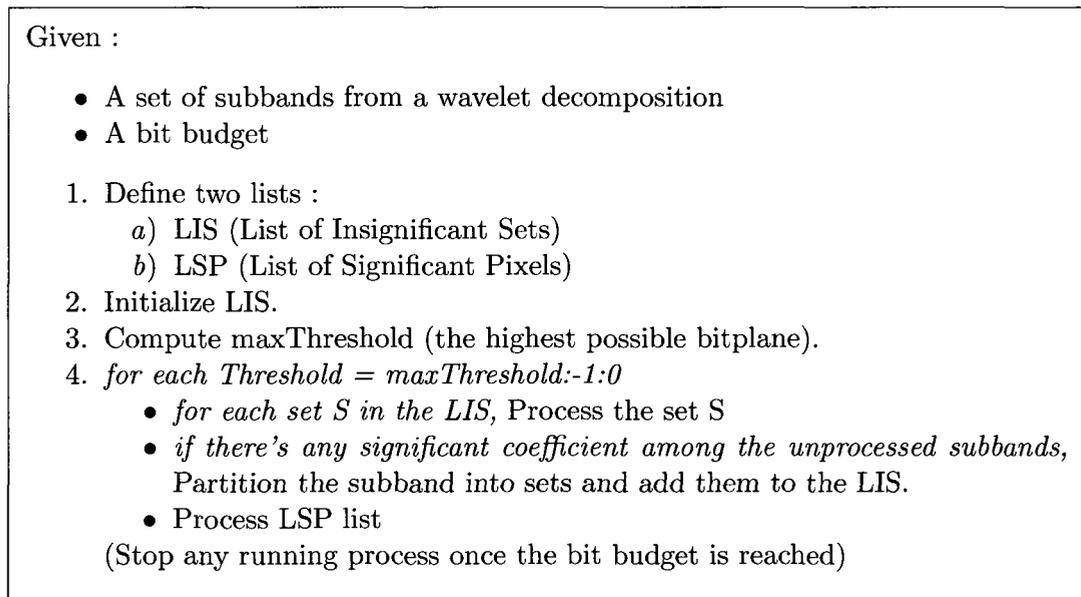


Figure 4.9: A simplistic overview of the SPECK coding algorithm.

that set partitioning technique to locate significant coefficients within each subband, bitplane after bitplane.

4.2.1 An overview of the coding algorithm

Figure 4.9 gives a simplistic general overview of the SPECK coding algorithm, from a top-level perspective. The interested reader is referred to the work by Pearlman *et al.* [21] for additional algorithm details. The algorithm uses two lists: the List of Insignificant Sets (LIS) and the List of Significant Pixels (LSP). A set in the LIS represents a group of coefficients obtained after partitioning of a bigger set (which could be a subband). Sets are gradually added to the LIS in decreasing order of decomposition level (from the highest to the finest level) and significant coefficients found while processing the sets in the LIS are moved to the LSP list.

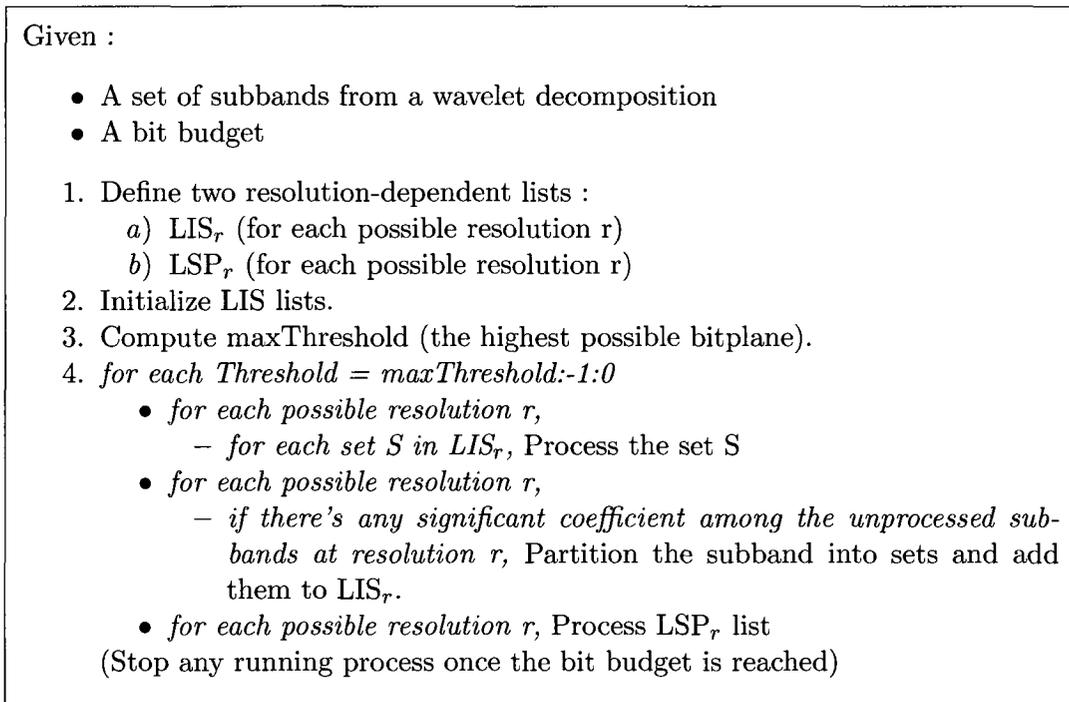


Figure 4.10: SPECK coding algorithm with resolution-dependent lists, producing a bit-stream with SNR scalability.

4.2.2 Conversion to a subband coder supporting spatial resolution scalability

Converting the original coding algorithm of SPECK to support spatial resolution scalability can be done following the same methodology described in Section 4.1 for SPIHT. The first step is to convert the regular algorithm to an equivalent algorithm using resolution-dependent lists LIS_k and LSP_k . The resulting algorithm structure is described in Figure 4.10. From that result, an algorithm structure with the capability of producing a resolution-scalable bitstream and specifying a bit budget for each spatial resolution can be derived. The idea is the same used in Section 4.1 in the case of SPIHT: the threshold loop at step 4 of the algorithm in Figure 4.10 can be separated in two loops to introduce spatial resolution scalability on two resolutions of interest r_1 and r_2 . The final algorithm structure including mechanisms to control

the bit budget allocated to each spatial resolution is described in Figure 4.11. It can be extended to a higher number of spatial resolutions in a straightforward manner.

4.3 Limitations of the proposed conversions

The different algorithm structures proposed for SPIHT and SPECK both require to be specified a set of spatial resolutions of interest and the corresponding bit budgets. In practice, however, the actual bitstream could be allocated a bit budget larger than the requested length for each resolution r_i smaller than the highest possible resolution. This is one tradeoff for simplicity in the algorithm. Because SPIHT and SPECK are bitplane coders, every threshold (or bitplane) that is being processed for a given resolution has to be fully completed before the next higher resolutions can be processed at that same threshold. So, in the proposed algorithm structures, the processing of a bitplane is not halted for any resolution other than the highest possible resolution, even if the requested bit budget is reached. The immediate consequence is that for a set of two resolutions of interest r_1 and r_2 , the portion of the bitstream allocated to r_1 will be a bit larger than requested. In theory, this could have an impact on the quality of the next resolution r_2 since part of the additional bit budget allocated to r_2 is consumed by the extra data. However, it should be observed that in a typical wavelet decomposition pyramid, subband coefficients at finer levels of decomposition are usually smaller in values than those belonging to the higher levels. Hence, more often than not, the details subbands needed for the bigger resolution r_2 will have to be processed at all the bitplanes at which the details subbands at higher levels had been processed, and the extra data produced from processing these higher level subbands will actually be used during the processing of the subbands at finer levels of decomposition.

Given :

- A set of subbands from a wavelet decomposition
- A bit budget for the resolution r_1
- A bit budget for the resolution r_2

1. Define two resolution-dependent lists :
 - a) LIS_r (for each possible resolution r)
 - b) LSP_r (for each possible resolution r)
2. Initialize LIS lists.
3. Compute $maxThreshold$. Set $maxBudget = bitBudget_{r_2}$
4. *for each Threshold = maxThreshold:-1:0*
 - *for each resolution r , $1 \leq r \leq r_1$,*
 - *for each set S in LIS_r , Process the set S*
 - *for each resolution r , $1 \leq r \leq r_1$,*
 - *if there's any significant coefficient among the unprocessed subbands at resolution r , Partition the subband into sets and add them to LIS_r .*
 - *for each resolution r , $1 \leq r \leq r_1$, Process LSP_r list*
 - *if bitBudget $_{r_1}$ is reached or even exceeded,*
 - *set $minThreshold_{r_1} = Threshold$ and exit the loop.*

(Stop any running process if $maxBudget$ is reached)
5. *for each Threshold = maxThreshold : -1 : minThreshold $_{r_1}$*
 - *for each resolution r , $r_1 < r \leq r_2$,*
 - *for each set S in LIS_r , Process the set S*
 - *for each resolution r , $r_1 < r \leq r_2$,*
 - *if there's any significant coefficient among the unprocessed subbands at resolution r , Partition the subband into sets and add them to LIS_r .*
 - *for each resolution r , $r_1 < r \leq r_2$, Process LSP_r list*

(Stop any running process if $maxBudget$ is reached)
6. *if bitBudget $_{r_2}$ is not yet reached,*
 - *for each Threshold = minThreshold $_{r_1} - 1 : -1 : 0$*
 - *for each resolution r , $1 \leq r \leq r_2$,*
 - * *for each set S in LIS_r , Process the set S*
 - *for each resolution r , $1 \leq r \leq r_2$,*
 - * *if there's any significant coefficient among the unprocessed subbands at resolution r , Partition the subband into sets and add them to LIS_r .*
 - *for each resolution r , $1 \leq r \leq r_2$, Process LSP_r list*

(Stop any running process once $bitBudget_{r_2}$ is reached)

Figure 4.11: Proposed SPECK coding algorithm structure with support for spatial resolution scalability on two resolutions of interest.

Chapter 5

A wavelet-based image coder using combined rational wavelet decomposition and its performance evaluation at different spatial resolutions

In Chapter 3, we have introduced the notion of combined rational wavelet decomposition: a two-dimensional wavelet decomposition made from combining rational wavelet transforms of different dilation factors. We have demonstrated that the set of native spatial resolutions produced by the combined wavelet decomposition can include spatial resolutions obtained from dyadic and non-dyadic ratios, if the dilation factors $\{(q_i/p_i)\}$ of the rational wavelet transforms involved in the decomposition are chosen such that $\prod_i(q_i/p_i) = 2$. In Chapter 4, we have described how SPIHT and SPECK, two popular subband coders used in wavelet-based image compression, can be converted into subband coders producing bitstreams with spatial resolution scalability. The present chapter introduces a wavelet-based image compression system using the combined rational wavelet decomposition introduced in Chapter 3 and the resolution-scalable subband coders derived in Chapter 4. We are interested in

Number of (combined) decomposition levels	Dyadic wavelet decomposition	Combined wavelet decomposition
0	704×576	704×576
1	352×288	528×432 352×288
2	176×144	264×216 176×244

Table 5.1: Native spatial resolutions produced by each level of decomposition of an original image of size 704×576 , when using a combined rational wavelet decomposition with dilation factors $\{(q_i/p_i)\} = \{4/3, 3/2\}$ against a regular dyadic wavelet decomposition.

evaluating the visual quality of the images at different spatial resolutions obtained from the image coder against the output images from a similar wavelet-based image compression system built with a dyadic wavelet decomposition. A prime interest is the evaluation of the compression performance (from a visual quality perspective) at *intermediate* spatial resolutions obtained from non-dyadic factors. We qualify as *intermediate* resolutions the native spatial resolutions produced by successive levels of two-dimensional wavelet decomposition, not including the resolution of the original image. For the construction of the wavelet-based image coder, we consider the use of a combined rational wavelet decomposition with dilation factors $\{(q_i/p_i)\} = \{4/3, 3/2\}$. As shown in Table 5.1 for an original image at 4CIF resolution, this produces one additional native spatial resolution than a dyadic wavelet decomposition for each level of (combined) decomposition. In the example of Table 5.1, the combined rational wavelet decomposition produces intermediate spatial resolutions of 528×432 and 264×216 that are not available with a dyadic wavelet decomposition. To produce an image of size 528×432 from a typical wavelet-based image coding system that uses a dyadic wavelet decomposition, one could either choose to interpolate the lower dyadic

spatial resolution of 352×288 , or to decimate the higher dyadic spatial resolution of 704×576 . Similarly, a spatial resolution of 264×216 would be obtained by either interpolating the lower dyadic resolution of 176×144 or decimating the next higher dyadic resolution of 352×288 .

The main objective of the present chapter is to evaluate the compression performance of a wavelet-based image coder in which a combined rational wavelet decomposition is used in place of a dyadic wavelet decomposition. Because we are mainly interested in the capacity of the image coder to produce additional native spatial resolutions, the performance evaluation is primarily targeted at images produced at intermediate spatial resolutions of non-dyadic nature. The visual quality of these images of non-dyadic resolution is evaluated against images of the same size obtained from a dyadic wavelet-based image coding system using one of the two following scenarios:

- a) Interpolate the lower dyadic spatial resolution
- b) Decimate the higher dyadic spatial resolution

To encode the subbands from the wavelet decomposition, we use the resolution-scalable versions of SPIHT and SPECK described in Chapter 4, with an algorithm structure that allows the assignment of a specific bitrate to each resolution. For SPIHT in particular, the parent-child tree construction scheme defined in Chapter 3 is implemented for a combined rational wavelet decomposition of two dilation factors. The use of SPECK does not require any particular adaptation. In order to differentiate the resolution-scalable versions of the subband coders from their respective original versions only capable of SNR scalability, we'll denote them as RS-SPIHT and RS-SPECK.

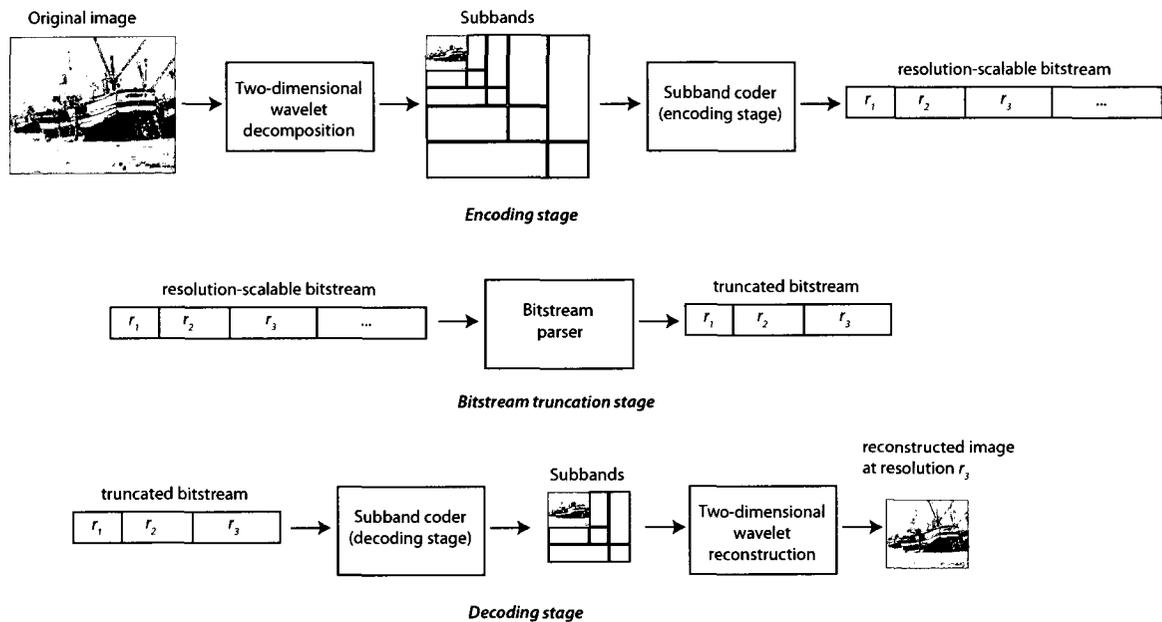


Figure 5.1: General structure of the wavelet-based image compression system.

5.1 General structure of the wavelet-based image coding system

Figure 5.1 illustrates the general structure of the typical wavelet-based image coding system that will be built in the current chapter. At the encoding stage, a two-dimensional wavelet decomposition is performed on the original image. The resulting subbands are inputs to a wavelet subband coder which produces a bitstream with spatial resolution scalability. In order to decode the image up to a given spatial resolution, a bitstream truncation stage is first used to truncate the encoded bitstream so that the resulting bitstream only comprises the data relevant to the spatial resolution of interest. The decoding stage consists of an inverse subband coding operation to produce a set of wavelet subbands from the input bitstream and a two-dimensional wavelet reconstruction to build an output image at the selected spatial resolution.

The framework illustrated in Figure 5.1 does not specify any particular kind of

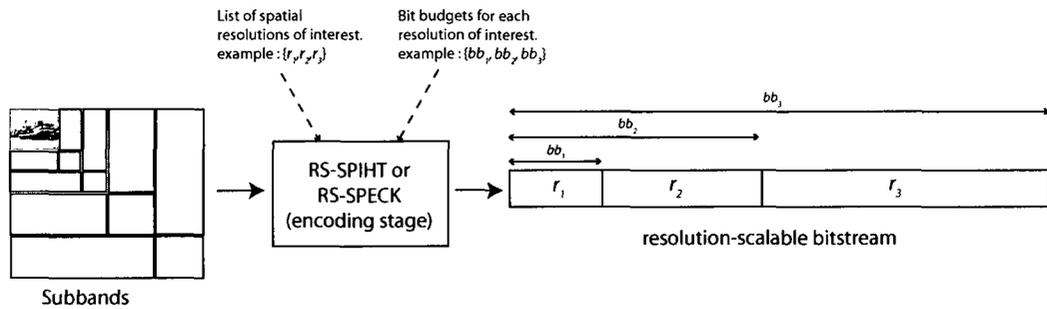


Figure 5.2: An expanded view of the subband coding phase with RS-SPIHT or RS-SPECK.

wavelet decomposition. In fact, the wavelet decomposition and reconstruction phases can be performed by using either dyadic wavelet decomposition, combined rational wavelet decomposition or a combination of both. Similarly, there's no restriction on the wavelet subband coders, with the only requirement being the capability to produce a bitstream with spatial resolution scalability. As such, this structure can be used to build a wavelet-based image compression system with dyadic or combined wavelet decomposition, and using RS-SPIHT or RS-SPECK to encode the subbands from the decomposition. In addition to the wavelet subbands, both subband coders require two complementary inputs at the encoding stage:

- The list of spatial resolutions of interest: this determines what spatial resolutions are available in the output bitstream.
- The bit budget for each spatial resolution of interest: this determines the length of each portion of the bitstream.

Figure 5.2 illustrates the output bitstream produced by the subband coders when three resolutions of interest $\{r_1, r_2, r_3\}$ are specified along with their respective bit budgets $\{bb_1, bb_2, bb_3\}$ ¹. At the bitstream truncation stage, the point of truncation

¹As discussed in Chapter 4, the actual bit lengths allocated to resolutions r_1 and r_2 in the bitstream could be larger than bb_1 and bb_2 .

p is selected according to the target spatial resolution of the output image at the decoding stage:

- if the target resolution of the output image is r_1 , select p such that $0 < p \leq bb_1$
- if the target resolution of the output image is r_2 , select p such that $bb_1 < p \leq bb_2$
- if the target resolution of the output image is r_3 , select p such that $bb_2 < p \leq bb_3$

Since both RS-SPECK and RS-SPIHT support SNR scalability within each spatial resolution, the quality of the output image at any given resolution increases as the truncation point approaches the higher end of the interval.

5.2 Presentation of the wavelet filters for the rational wavelet transforms

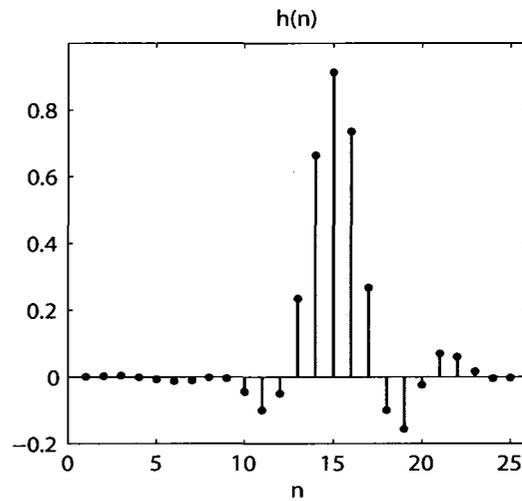
At the center of the wavelet-based image coding system is the two-dimensional wavelet decomposition/reconstruction. For a combined rational wavelet decomposition this translates into a succession of two-dimensional wavelet transforms with one set of wavelet filters designed to perform a $(3, 4)$ rational wavelet transform and a second set of wavelet filters designed to realize a $(2, 3)$ rational wavelet transform. Contrary to their dyadic counterparts, wavelet filters for rational wavelet transforms (which will be qualified as *rational* wavelet filters, in opposition to *dyadic* wavelet filters) do not enjoy the same type of popularity and only a few contributions are available in the current literature on the construction of such filters, as discussed in Chapter 2. In their most recent filter design algorithms, Bayram and Selesnick [15, 16] have been able to construct near-perfect reconstruction rational wavelet filters with a high number of vanishing moments, a feature that Blu [30] was not able to achieve in his often cited design algorithm. The impact of the number of vanishing moments on

the compression performance in wavelet-based image coding has been investigated by Rioul [28] and by Villasenor *et al.* [29]. For the $(3, 4)$ and $(2, 3)$ rational wavelet transforms, we select to use two sets of orthogonal rational wavelet filters of four vanishing moments designed by Bayram and Selesnick [16]. For comparison, the popular CDF9/7 dyadic wavelet filter [48] used in many dyadic wavelet-based image coding systems and included in the JPEG2000 still image coding standard also has four vanishing moments. Figure 5.3 and Figure 5.4 describe the sets of analysis filters used respectively for the $(2, 3)$ and $(3, 4)$ rational wavelet transform. The wavelet filters for the synthesis stage can be obtained by time reversing the analysis filters.

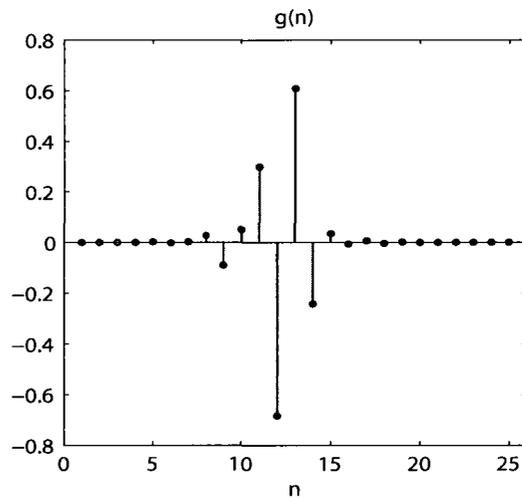
5.3 Experimental setup for the performance evaluation

To analyze the potential gain of using a combined rational wavelet decomposition, the compression performance of a wavelet-based image coder built with a combined rational wavelet decomposition (a *rational wavelet*-based image coder) will be evaluated against a similarly built wavelet-based image coder which only uses a dyadic wavelet decomposition (a *dyadic wavelet*-based image coder). We use the wavelet filters introduced in Section 5.2 for the rational wavelet transforms involved in the combined wavelet decomposition. The dyadic wavelet transforms in the dyadic wavelet-based image coder are performed with the popular CDF9/7 wavelet filters. The CDF9/7 wavelet filters are largely considered to be the state-of-the-art wavelet filters for still image compression and are commonly used as a benchmark for performance evaluation in wavelet-based image coding. A characterization of these filters has been performed by Unser and Blu [49]. Following the framework described in Section 5.1,

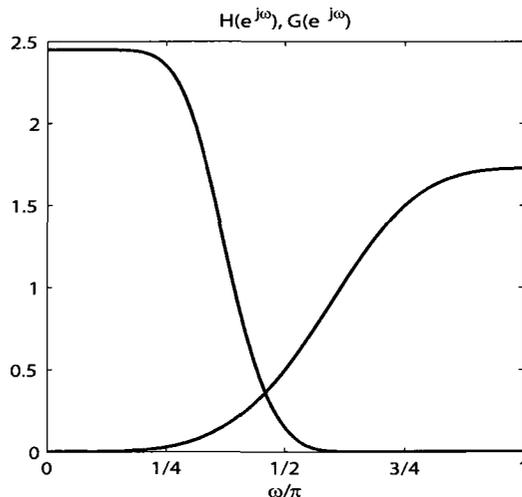
Lowpass filter $h(n)$ (25 taps)	Highpass filter $g(n)$ (24 taps)
0.000931996509026	0.000009153089499
0.002448598465988	0.000058618393965
0.002911535015330	0.000920490361886
-0.000599426506197	-0.003385498658335
-0.007829859416929	0.000920461218136
-0.012749076549185	0.026039038926867
-0.011155843247176	-0.091349120893691
-0.001820902202488	0.052010079644205
-0.004139850211368	0.298974832496901
-0.044664711110611	-0.684042986950223
-0.099670585730747	0.609060549740868
-0.049159632000169	-0.242171171905598
0.234998445166731	0.035240913946682
0.663483664409675	-0.006581467948113
0.913100742673968	0.006764158002450
0.735646034493276	-0.004319789768347
0.267354907116959	0.002519156934273
-0.100148260876629	-0.000804524917989
-0.155868638202700	0.000223941791444
-0.023896300373026	-0.000115246676411
0.070249999867410	0.000033702220187
0.059773760632585	-0.000010135502215
0.016427058658293	0.000007128236802
-0.003568876991631	-0.000002281783245
-0.002565036807206	



(a) Analysis lowpass filter



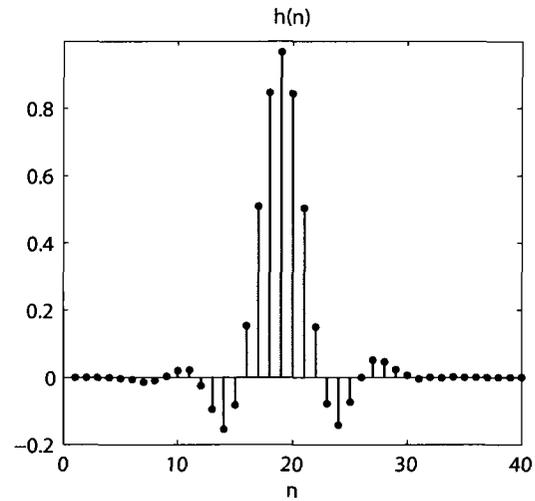
(b) Analysis highpass filter



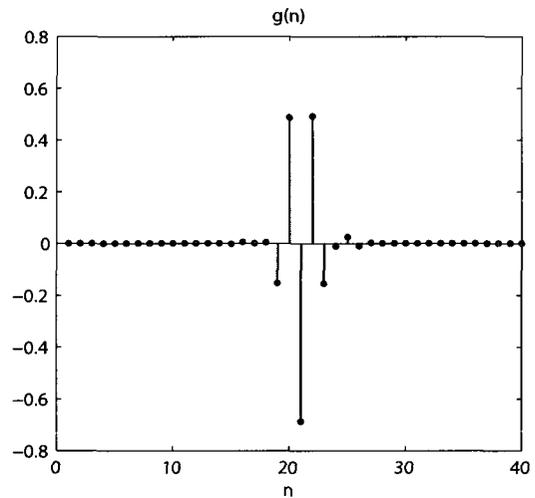
(c) Frequency responses (magnitude)

Figure 5.3: Analysis wavelet filters for the (2, 3) rational wavelet transform.

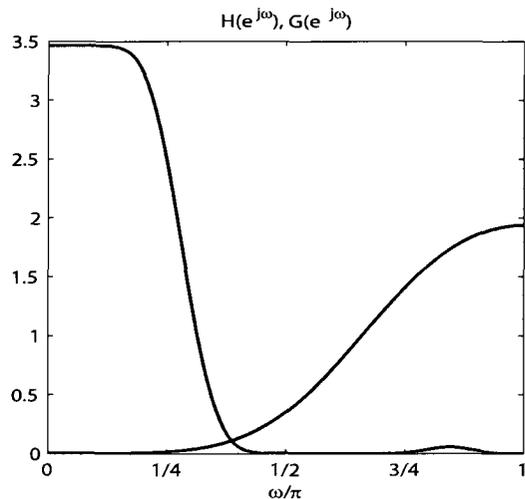
Lowpass filter $h(n)$ (40 taps)	Highpass filter $g(n)$ (38 taps)
0.000295756059765	-0.000000000003380
0.000439272336941	0.000000000023621
0.000641612390600	-0.000000000014316
0.000244652220555	0.0000000000873515
-0.002958989136838	-0.000000002206159
-0.006508681977425	-0.000000011867319
-0.013044748324851	-0.000000000191607
-0.009063057605487	0.000000206565007
0.002859814951498	0.000000438342251
0.020444422306546	-0.000002818036552
0.022422193339646	-0.000006263268591
-0.025372847093214	0.000040826774992
-0.095539338931745	0.000112367440133
-0.153927280052234	-0.000182691367135
-0.082332449886683	-0.001946845978853
0.153021053257279	0.005418750268284
0.508257705929164	0.000968841845441
0.846319246227002	0.004927883570663
0.967821497487002	-0.153133241221851
0.843084898440553	0.486527243828209
0.503346007531451	-0.687498642472118
0.150871298948758	0.492498688817879
-0.077938405585723	-0.153934447413044
-0.142206706929821	-0.011006152934278
-0.073103827000819	0.025373108674778
0.000988538240335	-0.010358996525128
0.052194350630454	0.002358472425750
0.046677473308111	0.000029082311815
0.022965209944714	-0.000252591384277
0.006082053288775	0.000077248053877
-0.004909195281031	-0.000011049124307
-0.000539146762223	0.000000046162456
-0.000728876194226	0.000000750101275
0.001199116649938	-0.000000219677184
0.000852181474525	0.000000016336507
0.000088647900358	0.000000001255208
0.000631940631443	-0.000000000068946
0.000117417815844	0.000000000083348
0.000318367540505	
0.000090437048313	



(a) Analysis lowpass filter



(b) Analysis highpass filter



(c) Frequency responses (magnitude)

Figure 5.4: Analysis wavelet filters for the (3, 4) rational wavelet transform.

we build two wavelet-based image compression systems:

- image coder \mathcal{A} : a wavelet-based image coder that uses a combined rational wavelet decomposition with $\{(q_i/p_i)\} = \{4/3, 3/2\}$ at the encoding stage, potentially with additional levels of dyadic wavelet decomposition performed after the combined decomposition, as introduced in Chapter 3.
- image coder \mathcal{B} : a wavelet-based image coder that only uses a typical dyadic wavelet decomposition at the encoding stage.

To handle the image boundaries during the wavelet transform computation, *periodic* extension is used for the rational wavelet transforms performed with the filters introduced in Section 5.2. For the dyadic wavelet transforms computed with the CDF9/7 wavelet filters, *symmetric* extension is performed at the image boundaries by using a lifting implementation [27]. The use of symmetric extension as the boundary handling method is a common practice with the CDF9/7 filters and is typically preferred to periodic extension because it results in less distortion around the borders of the image [50]. The CDF9/7 filters (and symmetric extension) are also used for any dyadic wavelet decomposition performed in conjunction to the combined rational wavelet decomposition of the image coder \mathcal{A} . Finally, to encode the wavelet subbands, we alternatively use RS-SPIHT or RS-SPECK in our experiments. The alternate use of the two subband coders is guided by an interest in analyzing how well different types of correlations between subband coefficients are exploited. RS-SPIHT exploits a cross-scale subband correlation among the coefficients, while RS-SPECK exploits the frequency clustering of coefficients within each separate subband. An arithmetic coding stage is added to each subband coder according to the respective original specification of SPIHT [20] and SPECK [21] and using the context-adaptive arithmetic model by Witten *et al.* [51]. Implementation of the wavelet transform operations and the subband coding algorithms was realized with MATLAB (from MathWorks, Inc.).

The performance of the two types of wavelet-based image coders is evaluated at various native spatial resolutions, and especially at those additional spatial resolutions obtained by non-dyadic factors (i.e., part of the set of native resolutions produced by the image coder \mathcal{A} but not included in the set of native resolutions produced by the image coder \mathcal{B}). We use three test images illustrated in Figure 5.5 : the popular *Boat* image of size 512×512 from the USC-SIPI image database², an image of size 768×512 from the Kodak image set³ (denoted *Girl*), and a half-size version of the *Cafe* image from the ISO400 JPEG test images set, of size 1024×1280 ⁴.

5.4 Evaluation of compression performance at full size resolution

First, we want to evaluate the visual quality of the images reconstructed at the original image resolution. The evaluation at full size resolution serves two objectives:

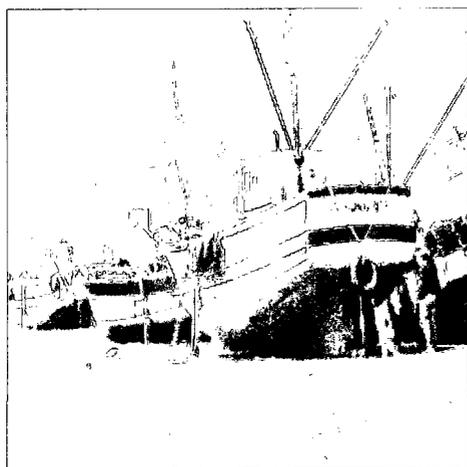
- a) validate the combined rational wavelet decomposition introduced in Chapter 3 along with the proposed parent-child tree construction schemes,
- b) assess the quality of the rational wavelet filters introduced in Section 5.2 for image compression.

We run the experimentation as follows: given a number of decomposition levels n and a specific bitrate, three decoded images are generated from the following three scenarios (we use the wavelet-based image coders \mathcal{A} and \mathcal{B} described in Section 5.3):

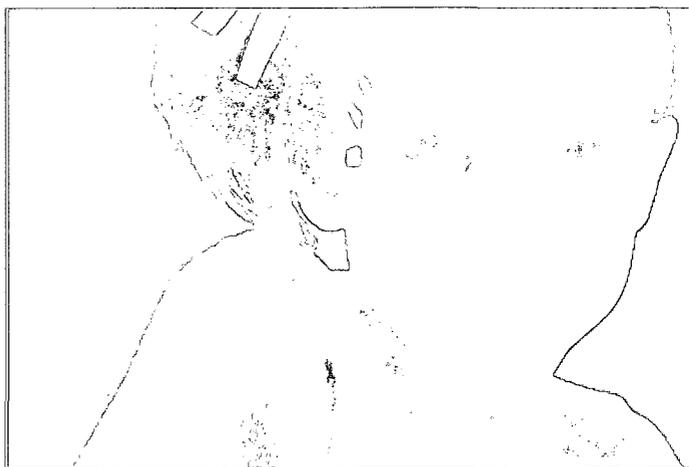
²The USC SIPI image database is available at <http://sipi.usc.edu/database/>

³The Kodak image set is available from the CIPR image database at <http://www.cipr.rpi.edu/resource/stills/kodak.html>

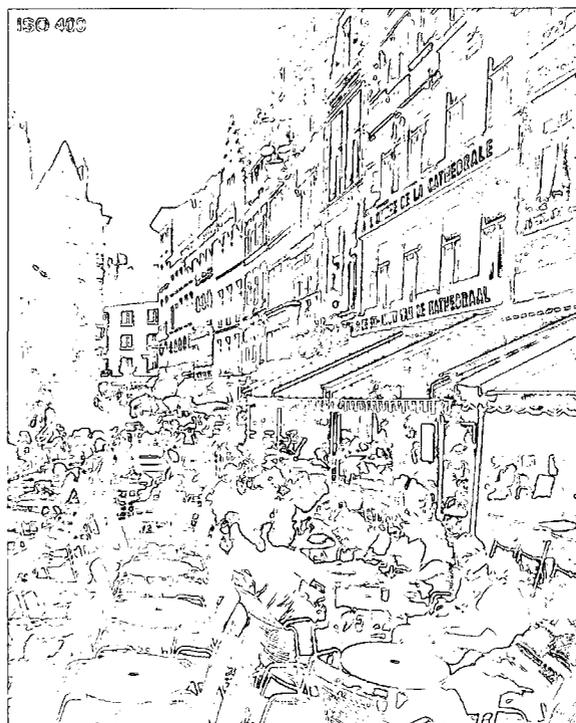
⁴The original *Cafe* image has a size of 2048×2560 . The test image is obtained by resizing the original using a *lanczos3* kernel.



(a) Boat (512 × 512)



(b) Girl (768 × 512)



(c) Cafe (1024 × 1280)

Figure 5.5: Test image set

- A_1 : The image coder \mathcal{A} is used with $n_1 = n$ levels of combined rational wavelet decomposition and no additional dyadic wavelet decomposition specified at the encoding stage.
- A_2 : The image coder \mathcal{A} is used with $n_1 = 2$ levels of combined rational wavelet decomposition and an additional $n_2 = n - 2$ levels of dyadic wavelet decomposition specified at the encoding stage.
- B : The image coder \mathcal{B} is used with n levels of dyadic wavelet decomposition specified at the encoding stage.

For this experiment, the original SPIHT and SPECK coding algorithms (including the arithmetic coding stage) are used to encode the subbands from the wavelet decomposition (this corresponds to using RS-SPIHT or RS-SPECK and specify the original image resolution as the single spatial resolution of interest). The quality of the images obtained after the decoding stages is expressed by using the *Structural Similarity Index Metric* (SSIM) by Wang *et al.* [52], an image quality assessment technique proposed as an alternative to the traditional *Peak Signal-to-Noise Ratio* (PSNR), and which has been very well received by the research community. The SSIM measure is one of several new quality metrics that have been reported to provide more accurate objective measures than mean squared error calculations. The reader interested in more discussion on alternative image quality metrics can find a potential interest in the article by Wang and Bovik [53]. Throughout our experiments, the SSIM measure has consistently given us objective quality scores closer to our subjective visual evaluation than the results from the traditionally-used PSNR. The Structural Similarity Index Metric gives objective quality scores on a scale of 0 to 1. In the experiments described in this section, the original image serves as reference image for the objective quality measurements.

bitrate (in bpp)		0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
<i>Boat</i> , SPIHT	A_1	0.788	0.82	0.849	0.865	0.879	0.891	0.9	0.908
	A_2	0.795	0.827	0.855	0.87	0.883	0.894	0.903	0.911
	B	0.803	0.836	0.862	0.878	0.89	0.899	0.905	0.912
<i>Boat</i> , SPECK	A_1	0.792	0.828	0.856	0.87	0.882	0.895	0.905	0.914
	A_2	0.798	0.835	0.861	0.874	0.886	0.898	0.907	0.916
	B	0.804	0.844	0.867	0.878	0.891	0.901	0.91	0.921

bitrate (in bpp)		0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
<i>Girl</i> , SPIHT	A_1	0.86	0.892	0.908	0.922	0.934	0.945	0.951	0.956
	A_2	0.871	0.897	0.911	0.923	0.935	0.943	0.948	0.953
	B	0.872	0.898	0.914	0.926	0.936	0.947	0.953	0.958
<i>Girl</i> , SPECK	A_1	0.869	0.896	0.909	0.926	0.937	0.946	0.952	0.957
	A_2	0.878	0.898	0.911	0.925	0.937	0.944	0.948	0.953
	B	0.879	0.899	0.913	0.928	0.937	0.948	0.953	0.958

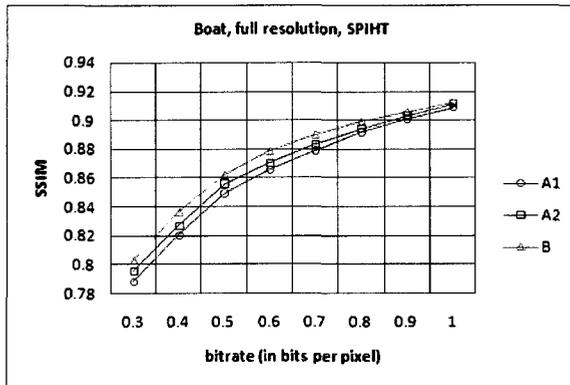
bitrate (in bpp)		0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
<i>Cafe</i> , SPIHT	A_1	0.689	0.736	0.777	0.806	0.831	0.854	0.871	0.887
	A_2	0.698	0.75	0.786	0.814	0.836	0.858	0.874	0.89
	B	0.709	0.763	0.797	0.829	0.85	0.874	0.891	0.906
<i>Cafe</i> , SPECK	A_1	0.698	0.741	0.781	0.814	0.837	0.857	0.874	0.889
	A_2	0.708	0.752	0.79	0.818	0.842	0.861	0.875	0.892
	B	0.716	0.763	0.8	0.83	0.853	0.876	0.891	0.905

Table 5.2: SSIM scores of the images decoded at full size resolution. The original image serves as reference image.

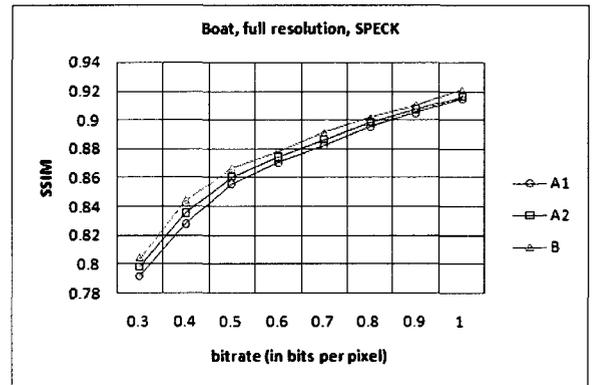
5.4.1 Experimental results

Table 5.2 and Figure 5.6 show the results from the objective quality evaluation of the three test images encoded at bitrates varying from 0.3 to 1 bit per pixel (bpp). The equivalent bit budget is equal to the product of the bitrate by the number of pixels in the original image. The total number of decomposition levels is specified to $n = 4$ for the images *Boat* and *Girl* and to $n = 5$ for the image *Cafe*.

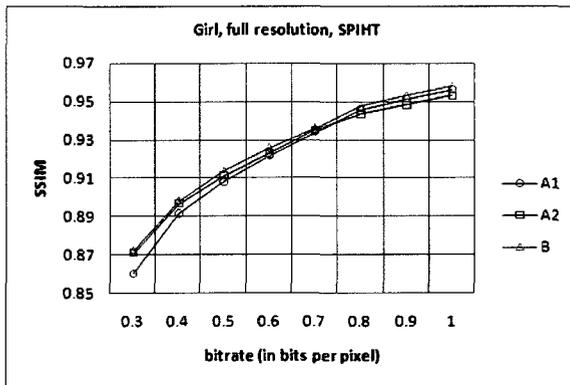
The output image from the dyadic wavelet-based image coder consistently obtains the highest SSIM score from the objective measurement, with both SPIHT and SPECK. This behavior is confirmed by visual inspection. At low bitrates, the decoded images from the rational wavelet-based image coder exhibit strong *ringing* artifacts, visible at bitrates ranging from 0.3 to 0.5 bpp for the image *Boat* and across the full range of bitrates for the image *Cafe* which has more details. For all three test images,



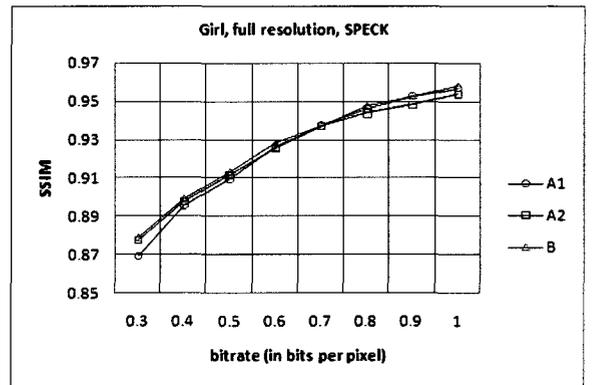
(a) *Boat*, SPIHT



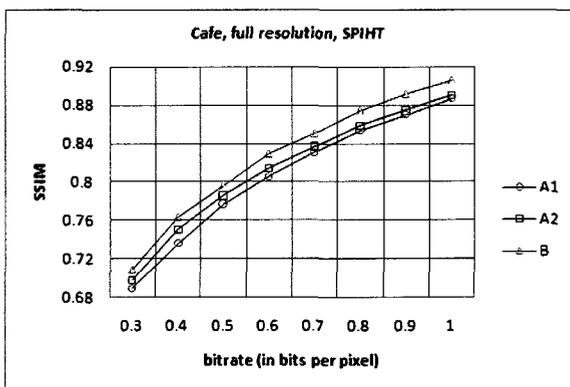
(b) *Boat*, SPECK



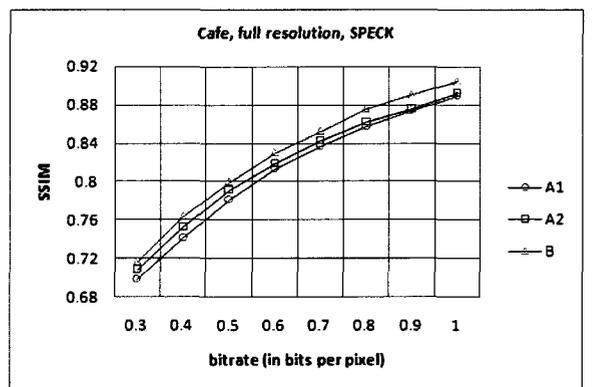
(c) *Girl*, SPIHT



(d) *Girl*, SPECK



(e) *Cafe*, SPIHT



(f) *Cafe*, SPECK

Figure 5.6: Comparison of objective quality scores between the images decoded at full size resolution.

the ringing distortion at low bitrates is less significant in the output images from the dyadic wavelet-based image coder. The two configurations A_1 and A_2 used with the rational wavelet-based image coder \mathcal{A} produce images of different visual quality. In fact, a notable characteristic observed when a dyadic wavelet transform is used for the highest levels of decomposition is the attenuation of the amount of ringing artifacts in the decoded image. Also, the decoded images obtained when the combined wavelet decomposition is used for *all* levels of decomposition show strong artifacts around the borders of the images. This distortion is also attenuated when dyadic wavelet decomposition and combined decomposition are used together. The difference of visual quality between the output images from both configurations becomes less significant as the bitrate increases, and eventually, the two output images become very similar. Finally, it can be observed from the SSIM scores that the use of SPIHT or SPECK to encode the subbands does not produce any significant difference in the results between the images from all three scenarios.

Figure 5.7 and Figure 5.8 illustrate a portion of the output images from *Cafe*, encoded at bitrates of 0.5 bpp and 1 bpp, respectively. Significant ringing can be observed in the two images obtained using the rational wavelet-based image coder. Figure 5.9 shows a portion of the output images from *Boat*, encoded at the bitrate of 0.3 bpp. The attenuating effect of using dyadic wavelet decomposition in conjunction with combined rational wavelet decomposition is noticeable, as far as the amount of ringing is concerned.



Figure 5.7: A part of the decoded image from *Cafe*, compressed at 0.5 bpp with SPIHT. *From top to bottom:* output image from the rational wavelet-based image coder with only combined decomposition, output image from the rational wavelet-based image coder with dyadic and combined decompositions, output image from the dyadic wavelet-based image coder.



Figure 5.8: A part of the decoded image from *Cafe*, compressed at 1.0 bpp with SPIHT. *From top to bottom:* output image from the rational wavelet-based image coder with only combined decomposition, output image from the rational wavelet-based image coder with dyadic and combined decompositions, output image from the dyadic wavelet-based image coder.

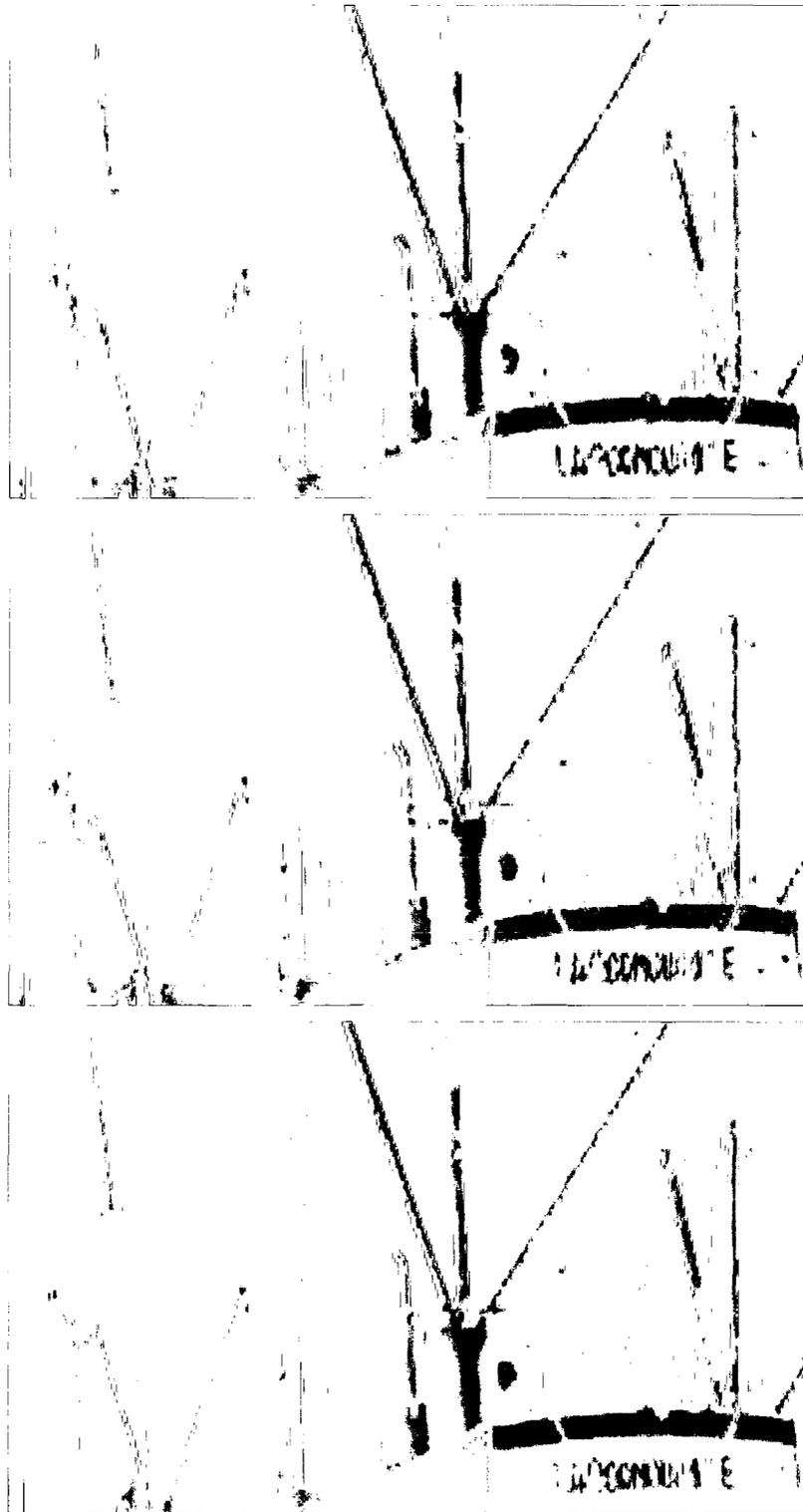


Figure 5.9: A part of the decoded image from *Boat*, compressed at 0.3 bpp with SPIHT. *From top to bottom*: output image from the rational wavelet-based image coder with only combined decomposition, output image from the rational wavelet-based image coder with dyadic and combined decompositions, output image from the dyadic wavelet-based image coder.

5.4.2 Analysis

The appearance of ringing artifacts at low bitrates is characteristic of wavelet-based image compression systems. However, it is believed to be more significant when wavelet filters with a relatively long highpass synthesis filter are used [50]. The length of the rational wavelet filters is, in theory, a likely cause of the noticeable ringing artifacts observed when a combined rational wavelet decomposition is performed at the encoding stage of the compression. The use of two-dimensional dyadic wavelet transform with the CDF9/7 filters instead of the combined rational wavelet transforms for the highest levels of decomposition contributes to attenuate the distortion at low bitrates. This can be explained by the fact that at low bitrates, the wavelet subbands that are mostly processed by the subband coder are those at the highest levels of decomposition. As the bitrate increases and the subbands at finer decomposition levels become more involved, the effect of using the dyadic wavelet decomposition in place of combined wavelet decomposition for the highest levels becomes less significant. Finally, the use of periodic extension at the image boundaries during the rational wavelet transform computation explains the perceptible distortion around the image borders when a combined rational wavelet decomposition is used. This distortion is attenuated when a dyadic wavelet decomposition is performed in addition to the combined decomposition, because a better boundary handling method with symmetric extension is implemented for the CDF9/7 dyadic wavelet filters.

One of the objectives of the quality evaluation at full size resolution was to assess the quality of the rational wavelet filters used in the combined rational wavelet decomposition and identify some of their limitations in the context of wavelet-based image compression. From that perspective, the interesting result from the experimentations at full size resolution is the gain that can be achieved by using a dyadic wavelet

decomposition with the CDF9/7 filters in conjunction with the combined wavelet decomposition. In fact, the results suggest that to mitigate some of the shortcomings of the rational wavelet filters and attenuate the amount of artifacts visible at low compression bitrates, the combined rational decomposition should only be used where there's a need for the additional intermediate spatial resolutions. For example, if the additional resolutions needed are the native resolutions corresponding to $3/4$ the original size and $3/8$ the original size, then it is preferable to use two levels of combined rational decomposition and an additional three levels of dyadic decomposition, than to use five levels of combined rational wavelet decomposition in the encoding process. How well this translates on the quality of the images at intermediate spatial resolutions is investigated in the next section.

Finally, within the context of the current research, the decrease in compression performance at full size resolution between the rational wavelet-based image coder and a similarly built dyadic wavelet-based image coder was deemed tolerable. It is judged an acceptable tradeoff for the inclusion of additional native spatial resolutions.

5.5 Evaluation of compression performance at intermediate spatial resolutions

5.5.1 A description of the experimental procedure to generate images at non-dyadic spatial resolutions

In this section, we want to evaluate the quality of a decoded image at an intermediate spatial resolution of non-dyadic nature produced by the rational wavelet-based image coder against an image of the same size obtained from the dyadic wavelet-based image coder by resizing an image decoded at the lower or higher dyadic spatial resolution.

Both image coders require three input arguments: the original image, the list of spatial resolutions of interest for the subband coder and the list of corresponding bit budgets for each spatial resolution. In the following paragraphs, we present an experimental procedure for the evaluation. The procedure describes how to generate the output images at the spatial resolutions of interest from each wavelet-based image coder, by specifying how the input arguments should be assigned and how the outputs of the encoding and decoding stages should be generated.

We denote by $R_{\mathcal{A}}$ the set of native spatial resolutions produced by the rational wavelet-based image coder \mathcal{A} and by $R_{\mathcal{B}}$ the set of native spatial resolutions produced by the dyadic wavelet-based image coder \mathcal{B} . We denote by r_{full} the spatial resolution corresponding to the size of the original image and by r , any spatial resolution such that $r \in R_{\mathcal{A}}$ and $r \notin R_{\mathcal{B}}$. Finally, for a given spatial resolution r , we denote by r^- and r^+ the two members of $R_{\mathcal{B}}$ corresponding respectively to the first dyadic resolution lower than r and the first dyadic resolution higher than r . To produce an output image of spatial resolution r using the dyadic wavelet-based image coder, two alternatives can be considered:

1. Interpolate a decoded image of spatial resolution r^- to resolution r .
2. Decimate a decoded image of spatial resolution r^+ to resolution r .

Consider an example where we use the image coder \mathcal{A} to encode an original image X , specifying two resolutions of interest $\{r, r_{full}\}$ ($r \in R_{\mathcal{A}}$ and $r \notin R_{\mathcal{B}}$) and the attached bit budgets $\{bitBudget_r, bitBudget_{full}\}$. For the evaluation of visual quality between the decoded image from the rational wavelet-based image coder and an image of the same size obtained from the dyadic wavelet-based image coder \mathcal{B} , the experimental procedure should clearly define how the output image from image coder \mathcal{B} should be generated. We propose to use two configurations:

B_1 : At the encoding stage of image coder \mathcal{B} , two resolutions of interest $\{r^-, r_{full}\}$ are specified with the attached bit budgets $\{bitBudget_r, bitBudget_{full}\}$. The decoded image is an image at resolution r^- , which is then interpolated to the non-dyadic resolution r .

B_2 : At the encoding stage of image coder \mathcal{B} , two resolutions of interest $\{r^+, r_{full}\}$ are specified with the attached bit budgets $\{bitBudget_r, bitBudget_{full}\}$. The decoded image is an image at resolution r^+ , which is then decimated to the non-dyadic resolution r .

Figure 5.10 describes the complete experimental procedure. The described procedure generates three set of output images $\{A^r, A^{full}\}$, $\{B_1^r, B_1^{full}\}$ and $\{B_2^r, B_2^{full}\}$ corresponding respectively to:

- the decoded image at spatial resolution r from the image coder \mathcal{A} , plus the decoded image at full size resolution r_{full} .
- the output image at spatial resolution r obtained from the image coder \mathcal{B} by interpolating an image decoded at the lower dyadic resolution r^- , plus the decoded image at full size resolution r_{full} .
- the output image at spatial resolution r obtained from the image coder \mathcal{B} by decimating an image decoded at the higher dyadic resolution r^+ , plus the decoded image at full size resolution r_{full} .

The procedure presented in Figure 5.10 comprises three pairs of encoding and decoding stages. First, the image coder \mathcal{A} is used to produce the output images $\{A^r, A^{full}\}$ which are obtained directly from the native set of resolutions proposed by the rational wavelet-based image coder. Then, two pairs of encoding/decoding stages are performed with the dyadic wavelet-based image coder \mathcal{B} , under the two configurations B_1 and B_2 described earlier, to produce the output images $\{B_1^r, B_1^{full}\}$ and

Given

- X , the original image,
 - r_{full} , the spatial resolution of the original image,
 - r , the intermediate spatial resolution of interest,
 - $bitBudget_{full}$, the required length of the full bitstream produced by the encoder,
 - $bitBudget_r$ the bit budget to assign to the spatial resolution r ,
1. Build the output images from image coder \mathcal{A} by performing the following steps:
 - a) Encode the original image with the image coder \mathcal{A} , specifying $\{r, r_{full}\}$ for spatial resolutions of interest with attached bit budgets $\{bitBudget_r, bitBudget_{full}\}$.
 - b) Truncate the original bitstream at length $bitBudget_r$ and decode the truncated bitstream to produce the output image A^r at spatial resolution r .
 - c) Decode the original bitstream to produce the output image A^{full} at the full size resolution r_{full} .
 2. Build the first output images from image coder \mathcal{B} by performing the following steps:
 - a) Encode the original image with the image coder \mathcal{B} , specifying $\{r^-, r_{full}\}$ for spatial resolutions of interest with attached bit budgets $\{bitBudget_r, bitBudget_{full}\}$.
 - b)
 - i. Truncate the original bitstream at length $bitBudget_r$ and decode the truncated bitstream to produce an image of spatial resolution r^- .
 - ii. Interpolate the result to spatial resolution r to form the output image B_1^r .
 - c) Decode the original bitstream to produce the output image B_1^{full} at the full size resolution r_{full} .
 3. Build the second output images from image coder \mathcal{B} by performing the following steps:
 - a) Encode the original image with the image coder \mathcal{B} , specifying $\{r^+, r_{full}\}$ for spatial resolutions of interest with attached bit budgets $\{bitBudget_r, bitBudget_{full}\}$.
 - b)
 - i. Truncate the original bitstream at length $bitBudget_r$ and decode the truncated bitstream to produce an image of spatial resolution r^+ .
 - ii. Decimate the result to spatial resolution r to form the output image B_2^r .
 - c) Decode the original bitstream to produce the output image B_2^{full} at the full size resolution r_{full} .

Figure 5.10: Proposed experimental procedure to generate the output images at a non-dyadic spatial resolution r and at full size resolution, from the image coders \mathcal{A} and \mathcal{B} .

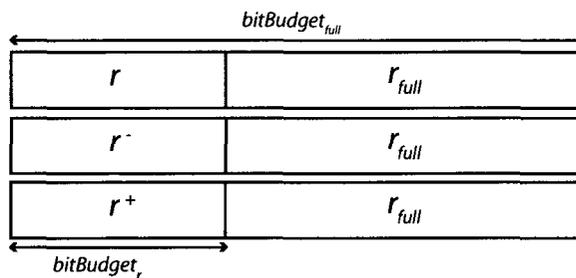


Figure 5.11: An illustration of the structure of the bitstreams from the three separate encoding stages of the experimental procedure.

$\{B_2^r, B_2^{full}\}$. The three encoding stages produce three output bitstreams with the same bit budget distribution, but a different set of spatial resolutions is available from each bitstream. Figure 5.11 illustrates the structure of the encoded bitstreams produced by each separate encoding stage. Before decoding the data up to a target spatial resolution r (respectively r^- or r^+), the original bitstream is truncated after $bitBudget_r$ bits and the resulting bitstream is decoded to build an image of spatial resolution r (resp. r^- or r^+). The decoded images from the dyadic wavelet-based image coder \mathcal{B} are resized from their initial resolution r^- or r^+ to match the non-dyadic spatial resolution r .

It can be observed that the described procedure takes advantage of the capability to control the structure of the bitstream and how many bits are assigned to each spatial resolution. Indeed, for a fair comparison between the output images from both wavelet-based image coders, it is important that the same bit budget is allocated to encode the spatial resolution r with image coder \mathcal{A} and the spatial resolution r^- or r^+ with image coder \mathcal{B} . This bitstream distribution can hardly be obtained without a subband coder capable of spatial resolution scalability. For example, if the original versions of SPIHT and SPECK are used in the wavelet-based image coding structure instead of RS-SPIHT and RS-SPECK, then the output image generation for each image coder would consist of :

1. Specify a bit budget for the full size resolution and encode the subbands from the wavelet decomposition.
2. Input the resulting bitstream to the decoding stage and generate a set of subbands for wavelet reconstruction.
3. Decide how many wavelet reconstruction levels to perform according to the spatial resolution of interest (r for the image coder \mathcal{A} , r^- or r^+ the image coder \mathcal{B}) and reconstruct the image up to that resolution.
4. For the image coder \mathcal{B} , resize the output images to the spatial resolution r .

The major disadvantage of this approach is that the subband coder controls how many bits are assigned to encode the data relevant to the spatial resolution r (resp. r^- or r^+). As a consequence, given a fixed bit budget for the full size spatial resolution, the number of bits used to encode the data up to spatial resolution r in the image coder \mathcal{A} might be much bigger than the number of bits used to encode the data up to spatial resolution r^- in the image coder \mathcal{B} .

In various contributions on non-dyadic spatial resolution scalability available in the literature [11, 25], the experimental procedure uses a subband coder with spatial resolution scalability but does not specify a bit budget for each spatial resolution. Instead, only the length of the full bit budget is specified to the subband coder, which leads to the same imbalance in bit budget allocation between output bitstreams from different image coders. Furthermore, the results reported by Pau *et al.* [25] and by Xiong *et al.* [11] only consider the option of interpolating the lower dyadic resolution to generate a resolution of non-dyadic nature from a dyadic wavelet-based image coder. In the experimental procedure described in this section, we have emphasized on having the same bit budget distribution for the image coders \mathcal{A} and \mathcal{B} , and considered two possible alternatives with the dyadic wavelet-based image coder. Although the procedure described is presented for a bitstream with only one non-dyadic spatial

	r_{full}	r_1	r_1^-	r_1^+	r_2	r_2^-	r_2^+
<i>Boat</i>	512×512	192×192	128×128	256×256	384×384	256×256	512×512
<i>Girl</i>	768×512	288×192	192×128	384×256	576×384	384×256	768×512
<i>Cafe</i>	1024×1280	384×480	256×320	512×640	768×960	512×640	1024×1280

Table 5.3: Description of the spatial resolutions of interest r_1 and r_2 for the three test images, along with the corresponding lower and higher dyadic spatial resolutions.

resolution, it can easily be generalized to any number of spatial resolutions obtained from non-dyadic factors.

5.5.2 Experimental results

For the experimentation at intermediate spatial resolutions, we are interested in two native spatial resolutions produced by the combined rational wavelet decomposition with dilation factors $\{(q_i/p_i)\} = \{4/3, 3/2\}$:

- a) the spatial resolution corresponding to $3/4$ the original size of the image
- b) the spatial resolution corresponding to $3/8$ the original size of the image

To use the terminology introduced in the description of the experimental procedure, we denote as r_1 and r_2 the spatial resolutions corresponding respectively to $3/8$ the original size of the image and $3/4$ the original size of the image. The lower dyadic spatial resolutions r_1^- and r_2^- are set to $1/4$ the original size of the image and $1/2$ the original size of the image, respectively. The higher dyadic spatial resolutions r_1^+ and r_2^+ are set respectively to $1/2$ the original size of the image and to the full size of the image. Table 5.3 lists the sizes corresponding to the spatial resolutions of interest for the three test images *Boat*, *Girl* and *Cafe*. The experimental procedure described earlier is extended to a set of resolutions of interest $\{r_1, r_2, r_{full}\}$ and produces encoded bitstreams of the structure illustrated in Figure 5.12. Furthermore, the image coder

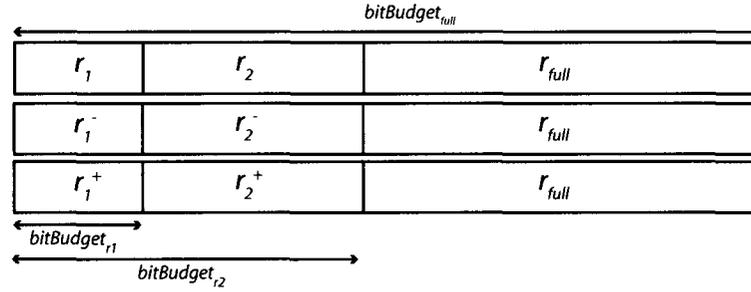


Figure 5.12: An illustration of the structure of the bitstreams from the encoding stages of the experimental procedure, when two intermediate resolutions of interest are specified.

\mathcal{A} is used under the two configurations presented in Section 5.4 for the evaluation of the compression performance at full size resolution :

A_1 : The image coder \mathcal{A} is used with $n_1 = n$ levels of combined rational decomposition and no additional dyadic wavelet decomposition specified at the encoding stage.

A_2 : The image coder \mathcal{A} is used with $n_1 = 2$ levels of combined rational decomposition and an additional $n_2 = n - 2$ levels of dyadic wavelet decomposition specified at the encoding stage.

In total, the extended procedure generates four sets of output images: $\{A_1^{3/8}, A_1^{3/4}, A_1^{full}\}$, $\{A_2^{3/4}, A_2^{3/8}, A_2^{full}\}$, $\{B_1^{3/8}, B_1^{3/4}, B_1^{full}\}$ and $\{B_2^{3/8}, B_2^{3/4}, B_2^{full}\}$. The same total number of decomposition levels n is used with both image coders \mathcal{A} and \mathcal{B} . The bit budget assigned to each intermediate spatial resolution is set to be proportional to the length of the full bitstream. Precisely for this experimentation, we set the intermediate bit budgets as :

$$bitBudget_{3/8} = (3/8) \times bitBudget_{full},$$

$$bitBudget_{3/4} = (3/4) \times bitBudget_{full},$$

where $bitBudget_{full}$ is the length requested for the full bitstream. The ratios of 3/4 and 3/8 used for the intermediate bit budgets are arbitrary numbers and could be set differently according to the image. During the course of the research, different set of

ratios have been experimented for the intermediate bit budgets; the ratios presented were selected because they have consistently given us satisfying results in terms of the balance of visual quality between images decoded at the three different spatial resolutions. The experimentation is run with the three test images at compression bitrates for the full size resolution ranging from 0.3 to 1.0 bits per pixel. For consistency, the results at all spatial resolutions are presented against that range of bitrates, with the understanding that

$$\text{bitrate} = \text{bitBudget}_{full} / (\text{number of pixels in the original image}),$$

and the corresponding intermediate bit budgets can be derived using the defined ratios.

Evaluating the effects of using dyadic and combined wavelet decompositions in the rational wavelet-based image coder

In the performance evaluation at full size resolution presented in Section 5.4, the use of dyadic wavelet transform with the CDF9/7 filters in addition to the combined rational wavelet decomposition resulted in an improvement of image quality at low bitrates. We are interested to see if the same positive effect is observed at intermediate spatial resolutions. To that end, we compare the visual quality of the output images produced by the rational wavelet-based image coder under the two configurations A_1 and A_2 . Recall that for the configuration A_1 , all levels of subband decomposition are obtained from the combined rational wavelet decomposition, whereas the configuration A_2 uses only two levels of combined decomposition and performs additional levels of decomposition with a dyadic wavelet transform. The Structural Similarity Index Metric is used to express the objective quality scores of the output images. The reference image is set as the uncompressed approximation subband of the same

size obtained while performing two levels of combined wavelet decomposition on the original image. Indeed, during the combined wavelet decomposition process, an approximation subband of spatial resolution corresponding to $3/4$ the original size and $3/8$ the original size is successively produced. The quality of that approximation subband represents, in theory, the highest image quality achievable at the given spatial resolution by any decoded image from the rational wavelet-based image coder.

Table 5.4 lists the objective quality scores of the output images decoded at $3/4$ the original size under both configurations. A graphical representation of the SSIM scores is displayed in Figure 5.13. As observed at full size resolution, the output images from the rational-wavelet based image coder exhibit ringing artifacts at low bitrates. These artifacts are visible for the image *Boat* between bitrates of 0.3 and 0.5 bpp. The decoded images from *Cafe* exhibit ringing artifacts across the full range of bitrates; in the case of the *Girl* image, the distortion is visible around the contour of the face, between bitrates of 0.3 and 0.6 bpp. The visual effects of using a dyadic wavelet transform for the highest levels of decomposition are similar to those observed at the full size resolution : attenuated ringing artifacts at low bitrates and less significant distortion around the borders of the image. The same positive effect is observed on the images decoded at $3/8$ the original size. The objective quality scores obtained are indicated in Table 5.5 and illustrated in Figure 5.14. It confirms the benefit of using dyadic wavelet decomposition in conjunction with combined rational wavelet decomposition for the rational wavelet-based image coder. Finally, the use of RS-SPECK in place of RS-SPIHT to encode the subbands does not have any visible impact on the difference in quality between the output images from both configurations.

bitrate (in bpp)		0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
<i>Boat</i> , RS-SPIHT	A_1	0.818	0.852	0.881	0.904	0.917	0.928	0.938	0.945
	A_2	0.831	0.86	0.886	0.91	0.921	0.93	0.94	0.947
<i>Boat</i> , RS-SPECK	A_1	0.829	0.857	0.889	0.911	0.922	0.931	0.941	0.949
	A_2	0.838	0.864	0.896	0.914	0.924	0.934	0.944	0.95

bitrate (in bpp)		0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
<i>Girl</i> , RS-SPIHT	A_1	0.864	0.893	0.915	0.93	0.94	0.951	0.957	0.963
	A_2	0.875	0.902	0.921	0.933	0.942	0.951	0.958	0.963
<i>Girl</i> , RS-SPECK	A_1	0.869	0.902	0.918	0.932	0.944	0.953	0.96	0.964
	A_2	0.876	0.908	0.922	0.933	0.945	0.952	0.955	0.963

bitrate (in bpp)		0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
<i>Cafe</i> , RS-SPIHT	A_1	0.727	0.774	0.814	0.844	0.869	0.887	0.903	0.917
	A_2	0.738	0.782	0.82	0.849	0.873	0.889	0.905	0.919
<i>Cafe</i> , RS-SPECK	A_1	0.737	0.784	0.819	0.854	0.876	0.893	0.909	0.923
	A_2	0.749	0.793	0.826	0.857	0.878	0.895	0.91	0.924

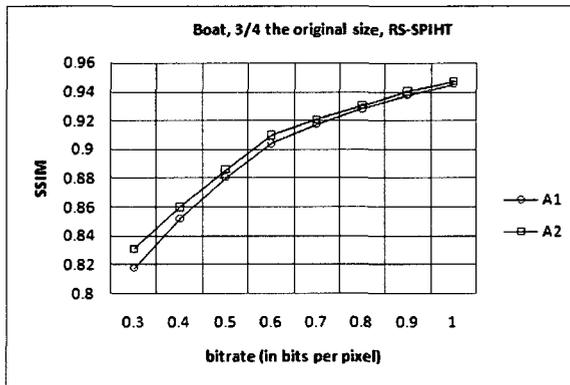
Table 5.4: SSIM scores of the images from the rational wavelet-based image coder decoded at 3/4 the original size. The reference image is an uncompressed approximation subband from the combined wavelet decomposition.

bitrate (in bpp)		0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
<i>Boat</i> , RS-SPIHT	A_1	0.851	0.89	0.921	0.941	0.956	0.967	0.976	0.982
	A_2	0.864	0.9	0.929	0.947	0.958	0.97	0.978	0.982
<i>Boat</i> , RS-SPECK	A_1	0.863	0.903	0.929	0.947	0.961	0.972	0.979	0.984
	A_2	0.867	0.911	0.937	0.95	0.964	0.974	0.979	0.985

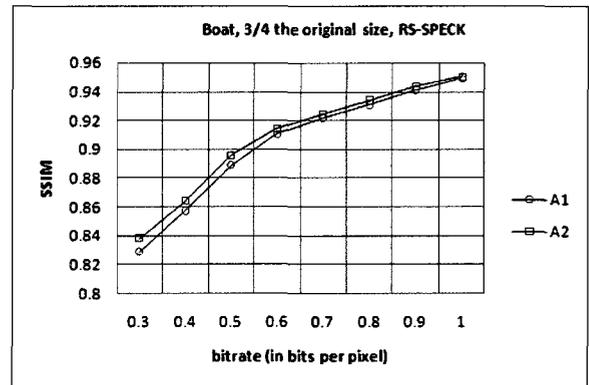
bitrate (in bpp)		0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
<i>Girl</i> , RS-SPIHT	A_1	0.899	0.935	0.952	0.967	0.976	0.982	0.987	0.99
	A_2	0.914	0.942	0.96	0.973	0.979	0.985	0.989	0.991
<i>Girl</i> , RS-SPECK	A_1	0.912	0.941	0.959	0.972	0.979	0.985	0.989	0.991
	A_2	0.92	0.946	0.964	0.974	0.981	0.987	0.99	0.992

bitrate (in bpp)		0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
<i>Cafe</i> , RS-SPIHT	A_1	0.823	0.869	0.904	0.928	0.946	0.96	0.972	0.978
	A_2	0.832	0.879	0.912	0.933	0.949	0.963	0.973	0.978
<i>Cafe</i> , RS-SPECK	A_1	0.835	0.883	0.913	0.936	0.952	0.964	0.974	0.981
	A_2	0.844	0.892	0.922	0.94	0.955	0.968	0.975	0.982

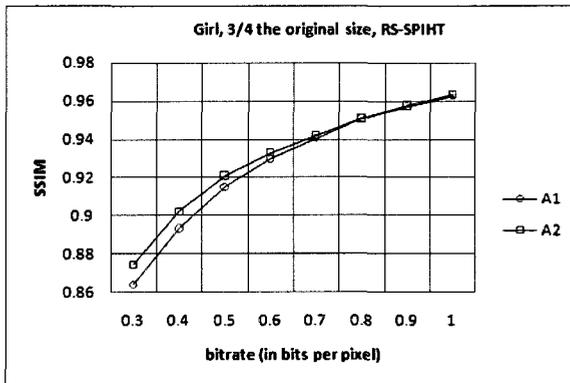
Table 5.5: SSIM scores of the images from the rational wavelet-based image coder decoded at 3/8 the original size. The reference image is an uncompressed approximation subband from the combined wavelet decomposition.



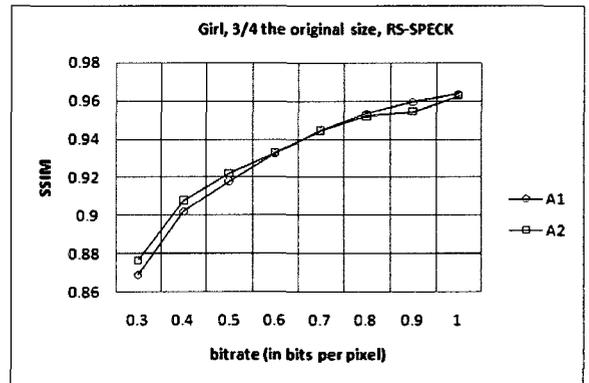
(a) *Boat*, RS-SPIHT



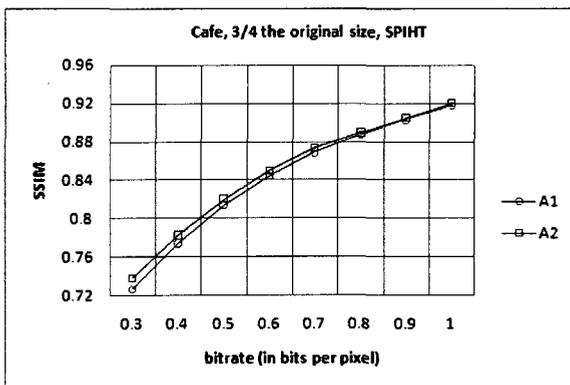
(b) *Boat*, RS-SPECK



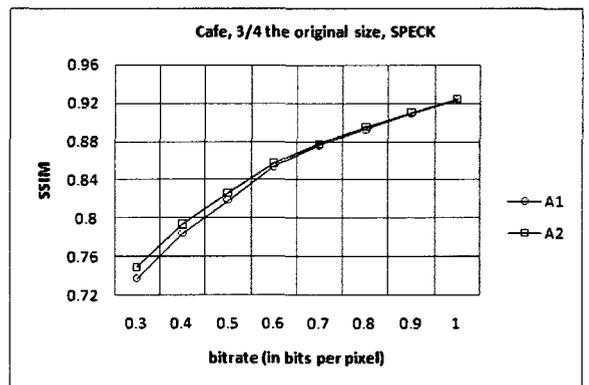
(c) *Girl*, RS-SPIHT



(d) *Girl*, RS-SPECK

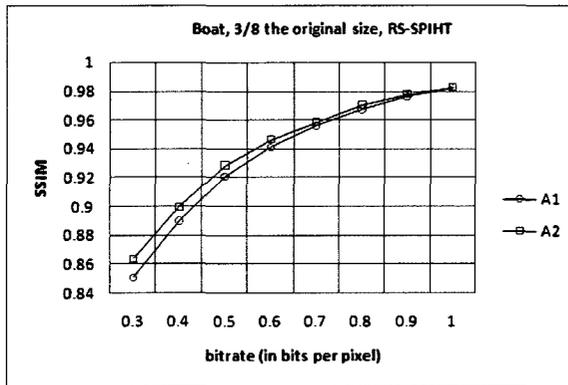


(e) *Cafe*, RS-SPIHT

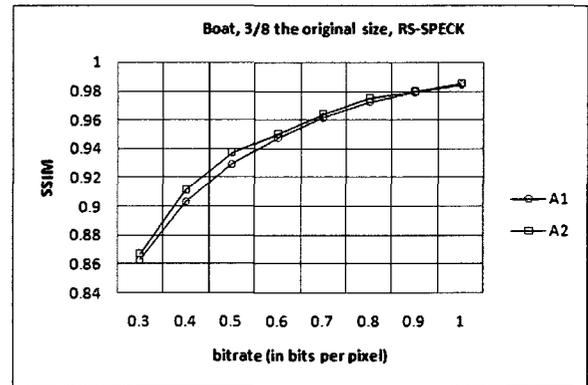


(f) *Cafe*, RS-SPECK

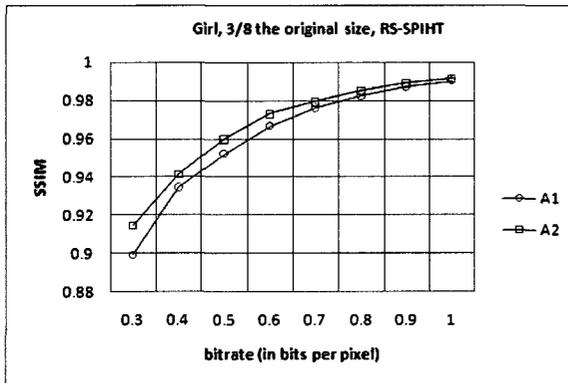
Figure 5.13: Comparison of objective quality scores between images from the rational wavelet-based image coder decoded at 3/4 the original size.



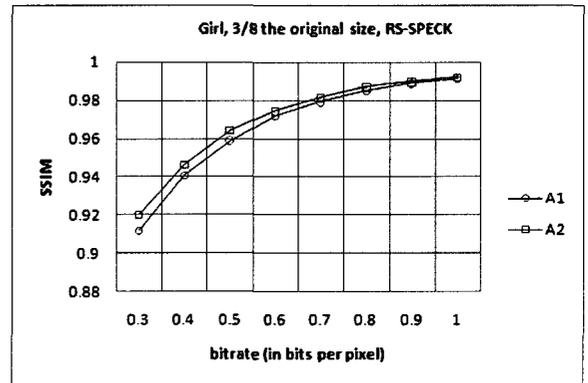
(a) *Boat*, RS-SPIHT



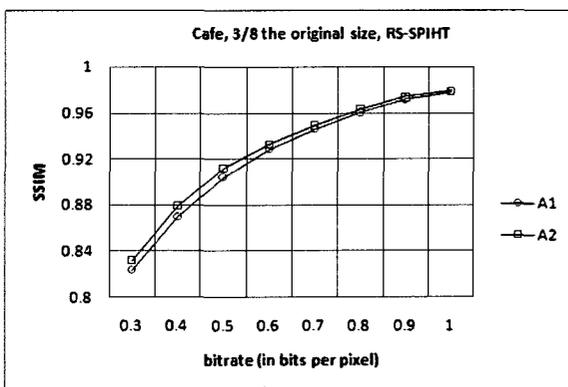
(b) *Boat*, RS-SPECK



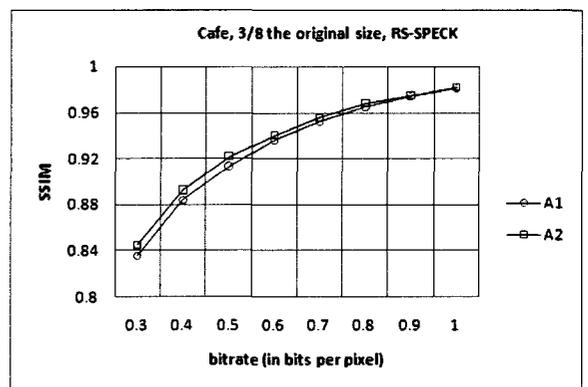
(c) *Girl*, RS-SPIHT



(d) *Girl*, RS-SPECK



(e) *Cafe*, RS-SPIHT



(f) *Cafe*, RS-SPECK

Figure 5.14: Comparison of objective quality scores between images from the rational wavelet-based image coder decoded at 3/8 the original size.

Evaluating the quality of the output images from both image coders

We now want to evaluate the visual quality of the images at intermediate spatial resolutions obtained from the rational wavelet-based image coder \mathcal{A} against the quality of the images obtained from the dyadic wavelet-based image coder \mathcal{B} by resizing images decoded at the lower or higher dyadic resolution. To that end, we compare the sets of images $\{A_2^{3/4}, A_2^{3/8}, A_2^{full}\}$, $\{B_1^{3/8}, B_1^{3/4}, B_1^{full}\}$ and $\{B_2^{3/8}, B_2^{3/4}, B_2^{full}\}$ corresponding respectively to :

- the output images obtained from the rational wavelet-based coder when a dyadic wavelet transform is used at the highest levels of decomposition,
- the output images obtained from the dyadic wavelet-based image coder by interpolating images decoded at the lower dyadic spatial resolution,
- the output images obtained from the dyadic wavelet-based image coder by decimating images decoded at the higher dyadic spatial resolution,.

The quality of the output images at full size resolution is evaluated in order to determine how the introduction of spatial resolution scalability affects the quality of the images decoded at the original size.

As with the previous experiments, we use the Structural Similarity Index Metric to express the objective quality measures of all images. A brief discussion has to be made about the reference image used for the computation. For images at full size resolution, the original image is used as the reference image, since it is in theory the best quality image that any decoding stage (from image coder \mathcal{A} or \mathcal{B}) can produce. At intermediate spatial resolutions, however, it is not always possible to determine such an absolute best quality image. When we compared the output images at intermediate resolutions produced by different configurations of the rational wavelet-based image coder, we selected to use as the reference images the uncompressed approximation

subbands from the combined rational wavelet decomposition. At the given intermediate resolutions, the uncompressed approximation subbands represent, in theory, the best quality achievable by any image obtained from the rational wavelet-based image coder. The same idea cannot be transposed to output images from two different image coders which use different wavelet decompositions and different techniques to generate the intermediate spatial resolutions. To compromise, we have decided to use a decimated version of the original image as reference for the objective quality measurements. Practically, the reference images are obtained by resizing the original image using a *lanczos* kernel with MATLAB. Another option was to use the uncompressed approximation subband from the combined rational wavelet decomposition. However, with this choice of reference image, the objective quality measure seemed to always give higher scores to the decoded images from the rational wavelet-based image coder, and lower scores to the images obtained from the dyadic wavelet-based image coder. This was not reflective of the subjective results from visual inspection. The probable reason is because the objective quality measurement finds more correlation between the output images from the rational wavelet-based image coder and a reference image selected as the approximation subband from multiple levels of combined rational wavelet decomposition. The compromising solution of using a decimated version of the original image for reference in the SSIM scores computation has resulted more consistently in objective quality scores that were reflective enough of the visual observations, with few exceptions experienced for some images used during the course of this research. It was also noticed that at intermediate spatial resolutions, the Peak Signal-to-Noise-Ratio measure was less consistent in its accuracy to reflect the visual observations; this can potentially be attributed to the distorted nature of the reference image, to which the PSNR seems to be more sensitive than the SSIM.

bitrate (in bpp)		0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
<i>Boat</i> , RS-SPIHT	A_2	0.841	0.868	0.894	0.919	0.928	0.937	0.947	0.953
	B_1	0.851	0.875	0.897	0.913	0.917	0.925	0.93	0.932
	B_2	0.847	0.87	0.896	0.922	0.93	0.937	0.944	0.949
<i>Boat</i> , RS-SPECK	A_2	0.849	0.872	0.904	0.923	0.931	0.941	0.951	0.956
	B_1	0.853	0.878	0.902	0.913	0.92	0.923	0.93	0.934
	B_2	0.857	0.873	0.904	0.925	0.932	0.936	0.946	0.955

bitrate (in bpp)		0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
<i>Girl</i> , RS-SPIHT	A_2	0.881	0.909	0.924	0.937	0.945	0.953	0.959	0.964
	B_1	0.886	0.911	0.919	0.93	0.937	0.937	0.94	0.941
	B_2	0.878	0.905	0.922	0.93	0.941	0.949	0.956	0.961
<i>Girl</i> , RS-SPECK	A_2	0.883	0.915	0.926	0.938	0.949	0.954	0.958	0.964
	B_1	0.89	0.911	0.921	0.93	0.934	0.939	0.939	0.941
	B_2	0.877	0.913	0.923	0.929	0.943	0.953	0.956	0.961

bitrate (in bpp)		0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
<i>Cafe</i> , RS-SPIHT	A_2	0.752	0.795	0.825	0.856	0.878	0.891	0.904	0.918
	B_1	0.757	0.797	0.822	0.846	0.86	0.865	0.872	0.877
	B_2	0.753	0.799	0.82	0.849	0.866	0.892	0.902	0.91
<i>Cafe</i> , RS-SPECK	A_2	0.764	0.802	0.832	0.864	0.883	0.896	0.908	0.922
	B_1	0.767	0.803	0.826	0.842	0.864	0.864	0.867	0.879
	B_2	0.764	0.802	0.821	0.851	0.884	0.895	0.903	0.911

Table 5.6: SSIM scores of the output images obtained at 3/4 the original size. The reference image is a decimated version of the original image.

For the three test images used throughout the experiments, the objective quality scores have been reflective enough of the subjective visual evaluation, at least within an acceptable margin. The resizing operation for the output images of image coder \mathcal{B} is performed with MATLAB using a *bicubic* kernel⁵ for both the interpolation and decimation operations. The objective quality scores obtained with the described evaluation model are listed in Table 5.6 and Table 5.7 for the spatial resolutions corresponding respectively to 3/4 the original size and 3/8 the original size. Table 5.8 lists the objective quality scores obtained from the images decoded at full size resolution. The scores are illustrated graphically in Figure 5.15, Figure 5.16 and Figure 5.17.

⁵Note that this is different from the *lanczos* kernel used for the reference image. In our experiments, the use of two different kernels resulted in better accuracy of the objective quality scores.

bitrate (in bpp)		0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
<i>Boat</i> , RS-SPIHT	A_2	0.863	0.899	0.926	0.943	0.954	0.964	0.971	0.976
	B_1	0.858	0.877	0.893	0.899	0.905	0.907	0.91	0.91
	B_2	0.858	0.889	0.918	0.933	0.945	0.953	0.961	0.966
<i>Boat</i> , RS-SPECK	A_2	0.867	0.912	0.932	0.947	0.959	0.967	0.973	0.978
	B_1	0.862	0.88	0.896	0.9	0.906	0.907	0.91	0.911
	B_2	0.862	0.904	0.921	0.936	0.948	0.958	0.963	0.968

bitrate (in bpp)		0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
<i>Girl</i> , RS-SPIHT	A_2	0.919	0.942	0.958	0.968	0.975	0.98	0.983	0.986
	B_1	0.918	0.933	0.94	0.942	0.945	0.946	0.947	0.947
	B_2	0.909	0.928	0.947	0.959	0.966	0.97	0.975	0.979
<i>Girl</i> , RS-SPECK	A_2	0.923	0.946	0.962	0.971	0.977	0.981	0.984	0.986
	B_1	0.921	0.934	0.941	0.943	0.945	0.946	0.947	0.947
	B_2	0.916	0.93	0.952	0.96	0.967	0.971	0.977	0.979

bitrate (in bpp)		0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
<i>Cafe</i> , RS-SPIHT	A_2	0.833	0.879	0.909	0.927	0.943	0.954	0.962	0.969
	B_1	0.828	0.85	0.87	0.876	0.882	0.886	0.888	0.891
	B_2	0.821	0.863	0.897	0.91	0.925	0.937	0.942	0.951
<i>Cafe</i> , RS-SPECK	A_2	0.846	0.894	0.915	0.934	0.949	0.957	0.965	0.972
	B_1	0.836	0.854	0.872	0.877	0.883	0.887	0.889	0.891
	B_2	0.827	0.88	0.9	0.913	0.922	0.943	0.948	0.952

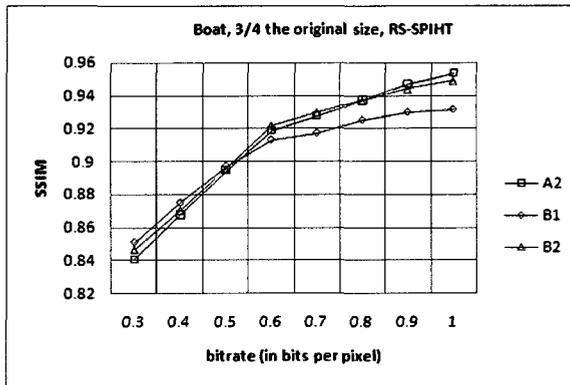
Table 5.7: SSIM scores of the output images obtained at 3/8 the original size. The reference image is a decimated version of the original image.

bitrate (in bpp)		0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
<i>Boat</i> , RS-SPIHT	A_2	0.795	0.827	0.855	0.87	0.883	0.894	0.903	0.911
	B_1	0.803	0.838	0.861	0.874	0.885	0.891	0.902	0.905
	B_2	0.803	0.836	0.862	0.878	0.89	0.899	0.905	0.912
<i>Boat</i> , RS-SPECK	A_2	0.798	0.837	0.858	0.874	0.886	0.891	0.902	0.916
	B_1	0.804	0.843	0.864	0.88	0.885	0.893	0.904	0.908
	B_2	0.804	0.844	0.867	0.878	0.891	0.901	0.91	0.921

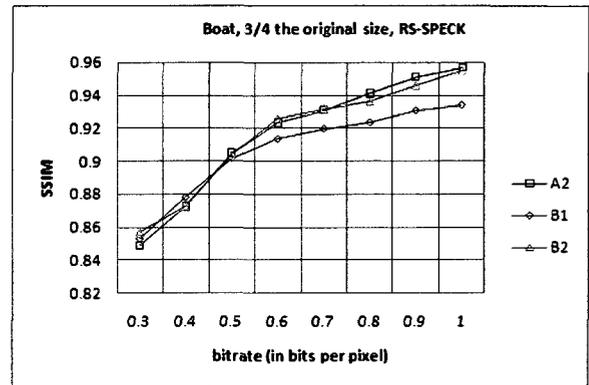
bitrate (in bpp)		0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
<i>Girl</i> , RS-SPIHT	A_2	0.871	0.897	0.911	0.923	0.936	0.943	0.948	0.953
	B_1	0.879	0.896	0.91	0.924	0.931	0.936	0.941	0.943
	B_2	0.872	0.898	0.914	0.926	0.936	0.947	0.953	0.958
<i>Girl</i> , RS-SPECK	A_2	0.878	0.898	0.911	0.925	0.937	0.944	0.949	0.953
	B_1	0.88	0.897	0.911	0.924	0.931	0.935	0.938	0.942
	B_2	0.879	0.899	0.913	0.928	0.937	0.948	0.953	0.958

bitrate (in bpp)		0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
<i>Cafe</i> , RS-SPIHT	A_2	0.696	0.75	0.779	0.814	0.835	0.858	0.874	0.89
	B_1	0.709	0.756	0.794	0.824	0.844	0.853	0.867	0.876
	B_2	0.709	0.763	0.797	0.829	0.85	0.874	0.891	0.906
<i>Cafe</i> , RS-SPECK	A_2	0.707	0.752	0.784	0.809	0.841	0.861	0.875	0.892
	B_1	0.716	0.762	0.789	0.817	0.844	0.847	0.855	0.878
	B_2	0.716	0.763	0.8	0.83	0.853	0.876	0.891	0.905

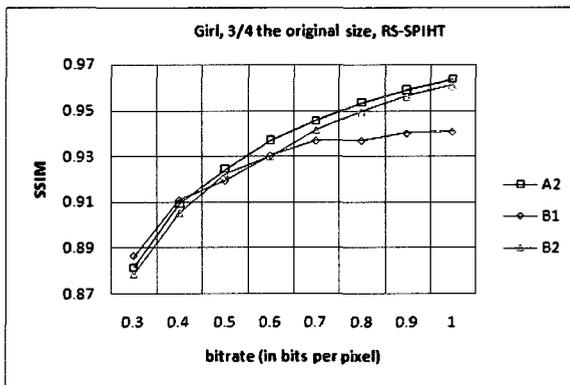
Table 5.8: SSIM scores of the output images obtained at full size resolution. The original image serves as reference image.



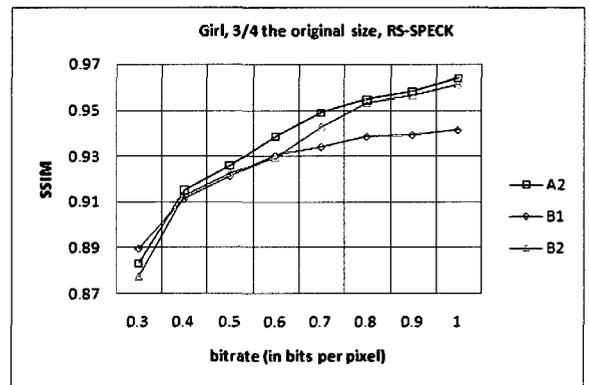
(a) *Boat*, RS-SPIHT



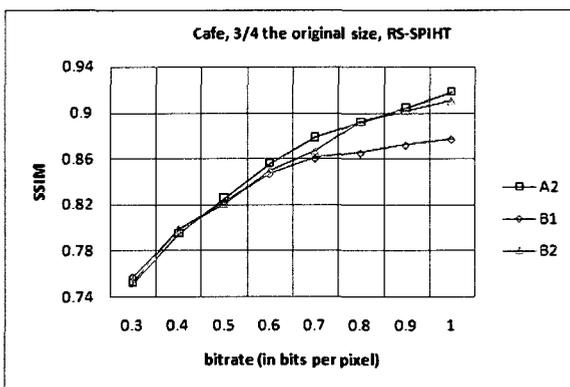
(b) *Boat*, RS-SPECK



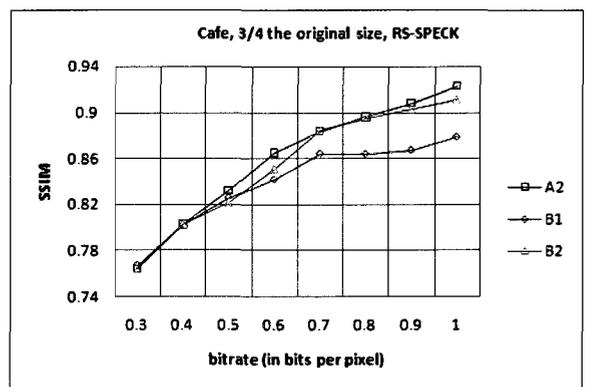
(c) *Girl*, RS-SPIHT



(d) *Girl*, RS-SPECK

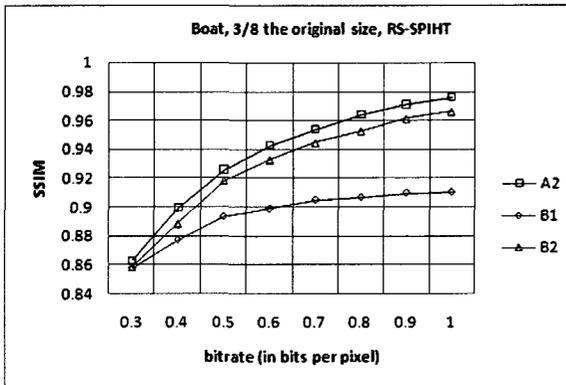


(e) *Cafe*, RS-SPIHT

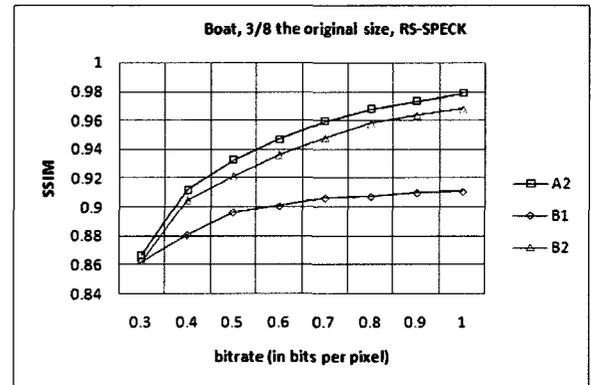


(f) *Cafe*, RS-SPECK

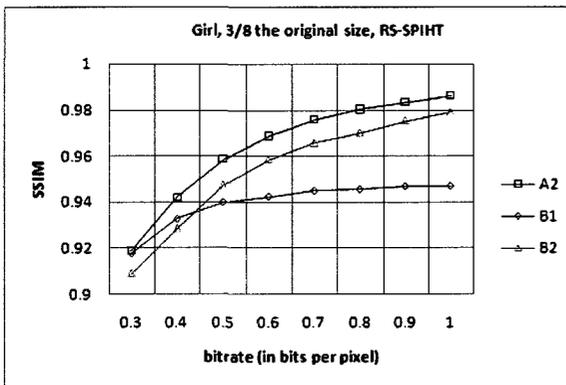
Figure 5.15: Comparison of objective quality scores between the output images obtained at 3/4 the original size.



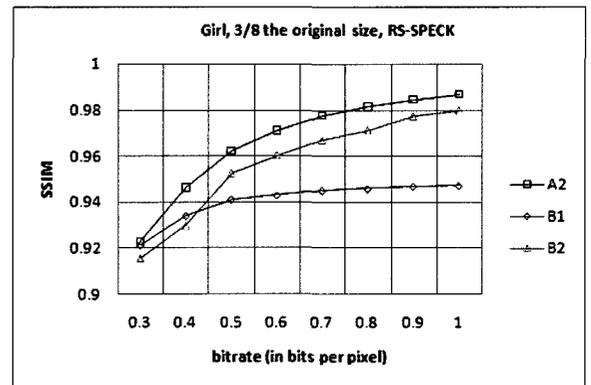
(a) *Boat*, RS-SPIHT



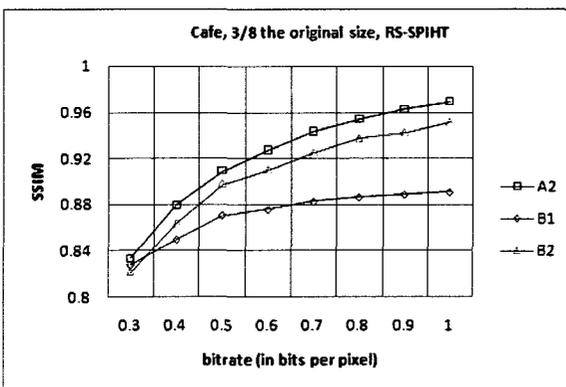
(b) *Boat*, RS-SPECK



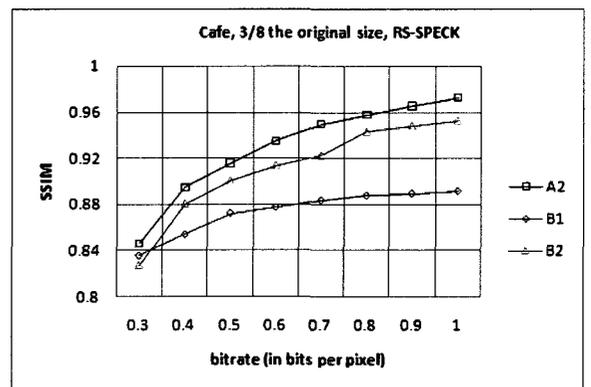
(c) *Girl*, RS-SPIHT



(d) *Girl*, RS-SPECK

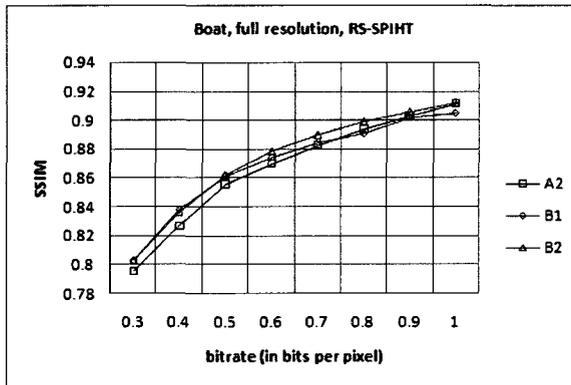


(e) *Cafe*, RS-SPIHT

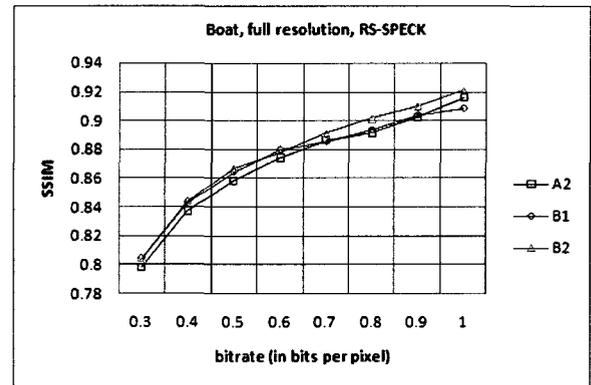


(f) *Cafe*, RS-SPECK

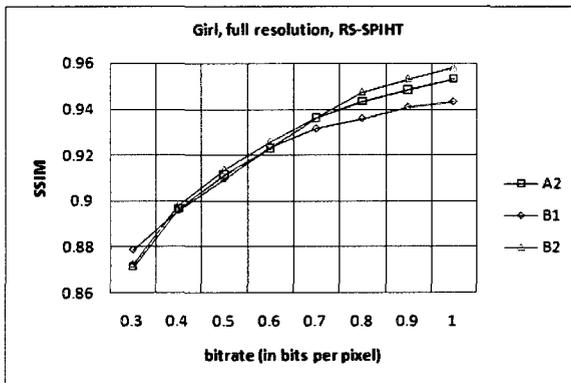
Figure 5.16: Comparison of objective quality scores between the output images obtained at 3/8 the original size.



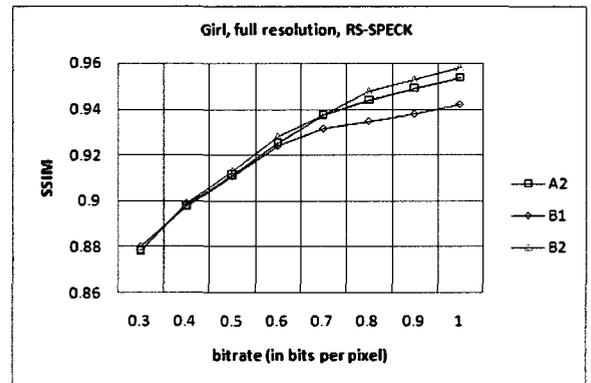
(a) *Boat*, RS-SPIHT



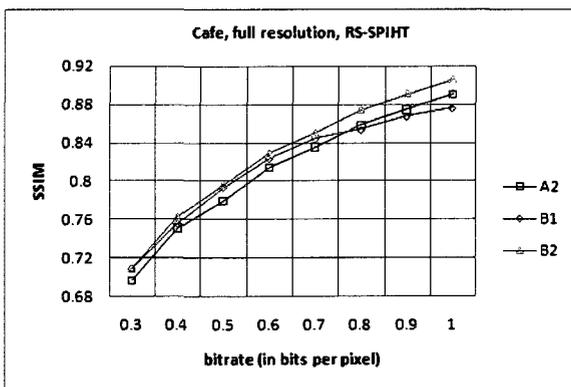
(b) *Boat*, RS-SPECK



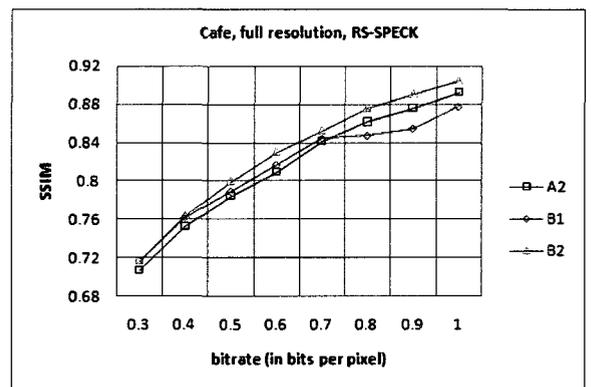
(c) *Girl*, RS-SPIHT



(d) *Girl*, RS-SPECK



(e) *Cafe*, RS-SPIHT



(f) *Cafe*, RS-SPECK

Figure 5.17: Comparison of objective quality scores between the output images obtained at full size resolution.

At $3/4$ the original size, the output image obtained from the dyadic wavelet-based image coder by interpolating the lower dyadic spatial resolution has the better visual quality of all three output images at low bitrates. At these bitrates, the image obtained from interpolating the lower dyadic resolution looks smoother and is almost free of some of the ringing artifacts that are displayed in both the decoded image obtained from the rational wavelet-based image coder and the image obtained from decimating the higher dyadic resolution. However, as the bitrate increases, the latter two images have more sharpness than the image obtained from the lower dyadic resolution and it visually compensates for some of the artifacts. Eventually, as the bitrate gets closer to 1.0 bpp, the output image obtained by interpolating the lower dyadic resolution looks blurry while the other two images exhibit more details. For all three test images, and across all the bitrates with the exception of the lower ones, the visual difference between the image obtained by decimating the higher dyadic resolution and the image decoded from the native resolutions of the rational wavelet-based image coder remains relatively small. At lower bitrates, both images exhibit some ringing artifacts, which are more noticeable in the output from the rational wavelet-based image coder. As the bitrate increases, the most consistent visual difference between the two output images is the sharpness, which always seem to be better for the image obtained from the rational wavelet-based image coder. Figure 5.18 and Figure 5.19 illustrate examples of output images from *Cafe* obtained at $3/4$ the original resolution, for encoding bitrates of 0.5 bpp and 1.0 bpp, respectively. Images from *Boat*, also obtained at $3/4$ the original resolution, are illustrated in Figure 5.20 for encoding bitrates of 0.3 bpp and 1.0 bpp. At the lower bitrate, significant ringing artifacts can be observed in both the output image from the rational-wavelet based image coder



Figure 5.18: A part of the output images from *Cafe*, encoded with RS-SPIHT at a bitrate of 0.5 bpp and decoded at 3/4 the original size.

From top to bottom: output image from the rational wavelet-based image coder, output image from interpolating the lower dyadic resolution, output image from decimating the higher dyadic resolution.



Figure 5.19: A part of the output images from *Cafe*, encoded with RS-SPIHT at a bitrate of 1.0 bpp and decoded at $3/4$ the original size.

From top to bottom: output image from the rational wavelet-based image coder, output image from interpolating the lower dyadic resolution, output image from decimating the higher dyadic resolution.

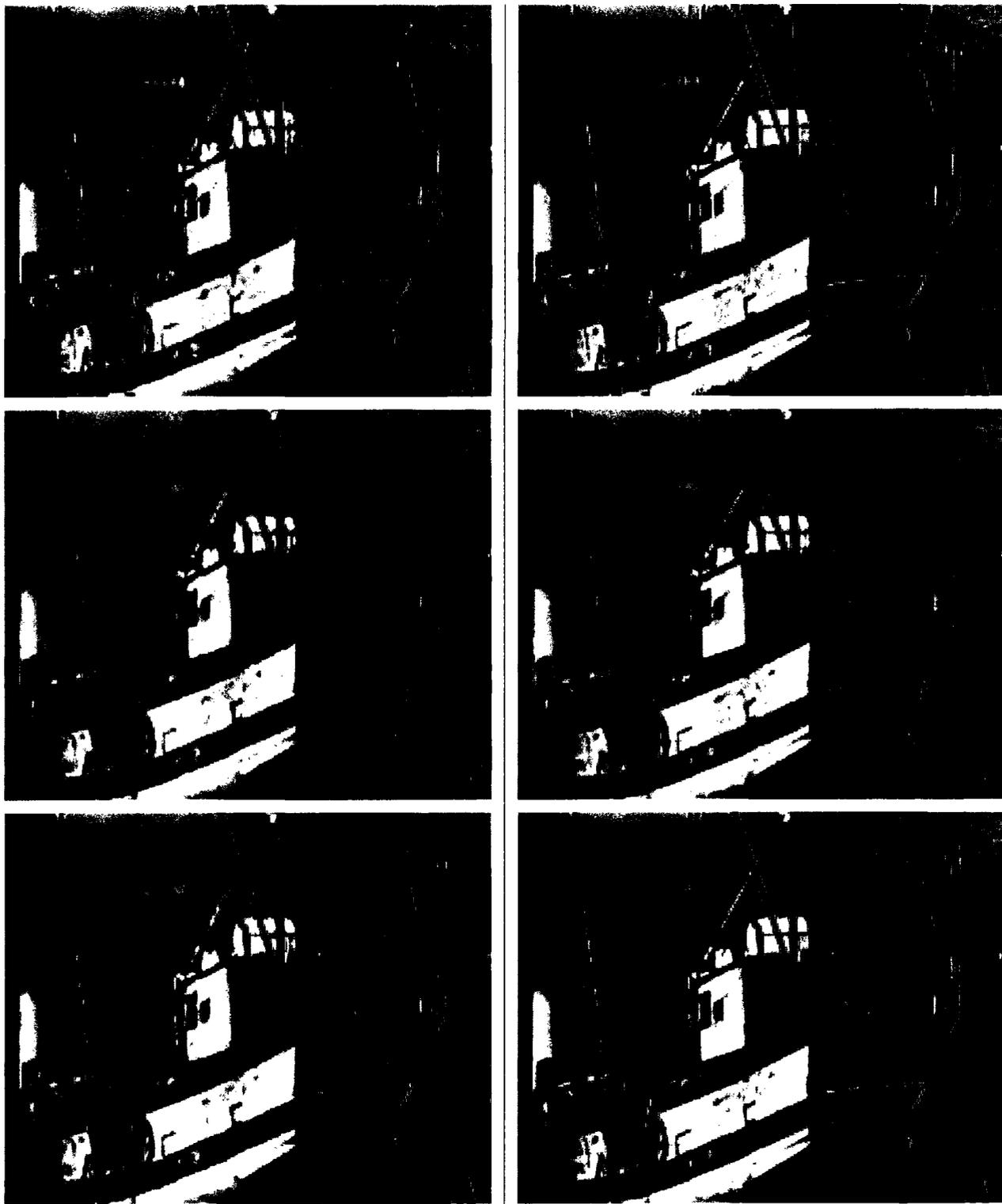


Figure 5.20: A part of the output images from *Boat*, encoded with RS-SPIHT at bitrates of 0.3 bpp (left) and 1.0 bpp (right), and decoded at $3/4$ the original size.

From top to bottom: output image from the rational wavelet-based image coder, output image from interpolating the lower dyadic resolution, output image from decimating the higher dyadic resolution.

and the output image obtained from decimating the higher dyadic resolution, while very little ringing is visible in the output image obtained by interpolating the lower dyadic resolution. However, at the higher bitrate, the image obtained by interpolating the lower dyadic resolution looks blurry in comparison to the other two.

For output images of $3/8$ the original size, the image obtained from interpolating the lower dyadic resolution looks blurry, exhibits fewer details and is consistently of lesser visual quality across the full range of bitrates. The image obtained from the native resolutions of the rational wavelet-based image coder is the sharpest of all three output images and consistently shows more details. Ringing artifacts are less significant for the images obtained at $3/8$ the original size, even at low bitrates. Figure 5.21 illustrates the output images from *Girl*, encoded at a bitrate of 1.0 bpp. A difference in the amount of details can be observed on the texture of the girl's clothing.

The objective quality results and the visual evaluations performed on the output images decoded at full size resolution are in line with the results presented in Section 5.4, in the cases of the output image from the rational wavelet-based image coder and the image obtained from the dyadic wavelet-based image coder when the intermediate resolutions bit budgets are distributed among the higher dyadic resolutions. On the other hand, a noticeable decrease in quality at higher bitrates is visible on the image at full size resolution obtained when the intermediate resolution bit budgets are distributed among the lower dyadic resolutions. This is a good illustration of the impact of spatial resolution scalability on the full resolution image. Because the bit budgets are not well balanced between the intermediate resolutions and the full size resolution (in this configuration of the dyadic wavelet-based image coder, $3/4$ of the full bit budget is allocated to the half-size resolution), this impacts the quality of the

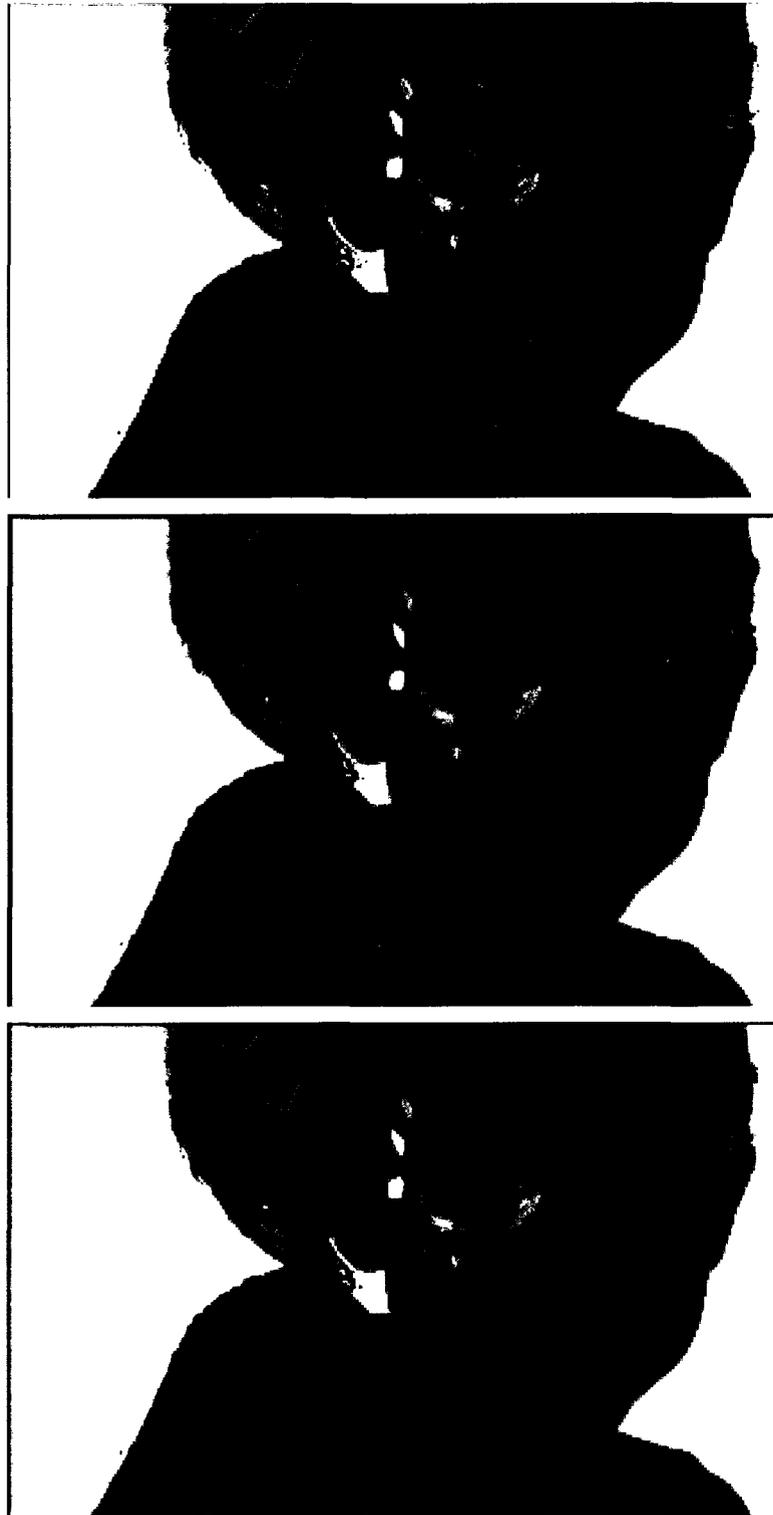


Figure 5.21: Output images from *Girl*, encoded with RS-SPIHT at a bitrate of 1.0 bpp and decoded at $3/8$ the original size.

From top to bottom: output image from the rational wavelet-based image coder, output image from interpolating the lower dyadic resolution, output image from decimating the higher dyadic resolution.

image obtained at full size resolution. The impact is less significant for the other two output images because of a better balance in the bit budget distribution.

Finally, in our experiments, the results at all spatial resolutions have looked very similar with both subband coders : the difference in visual quality between the output images encoded with RS-SPECK has consistently looked similar to the observations made when RS-SPIHT was used instead. At no point did it look like one type of correlation was better exploited in the case of the rational wavelet-based image coder over the dyadic wavelet-based image coder, or vice-versa.

5.6 Discussion

The appearance of strong ringing artifacts at low bitrates is the most notable characteristic observed in the images obtained from the rational wavelet-based image coder. These artifacts are an illustration of the limitations of the rational wavelet filters used in the combined wavelet decomposition. Although they are results from recent advances in rational wavelet filter design, the filters described in Section 5.2 and used throughout the experiments were not specifically designed for image coding applications and had not been tested in that context before. Various researchers have characterized how different properties of the wavelet filter can impact the overall compression performance [28, 29]. The current research did not include designing a set of rational wavelet filters with specific properties targeted for image coding applications. Instead, an effort was made on finding corrective measures that can be used to overcome the limitations of the set of rational wavelet filters. The conjunction of dyadic wavelet transform and combined rational wavelet decomposition was developed as a corrective measure, with the idea of taking advantage of the much richer family of wavelet filters for dyadic decomposition to improve the compression

performance. The results observed from the experimentation were quite satisfactory from that perspective, as the use of dyadic wavelet transform for the higher levels of decomposition consistently resulted in better visual quality and attenuated distortion at all spatial resolutions.

Another limitation of the rational wavelet-based image coder presented in this chapter is the method used to handle the image boundaries during the wavelet transform computation. The use of periodic extension at image boundaries, although being relatively simple, is known to produce visible artifacts around the borders of the image. To correct that, symmetric extension at image boundaries is typically performed. However, it is a more complex option for which techniques have been defined for different family of filters. For odd-length wavelet filters, which are very popular in dyadic wavelet-based image coding, the use of symmetric extension has been made very simple with the advent of the lifting scheme [26,27] and other elegant techniques proposed in the literature [54]. The CDF9/7 filters used for the dyadic wavelet transforms in our experiments belong to this family of filters, and we have used a symmetric extension for the dyadic wavelet transforms performed with these filters. On the other hand, the rational wavelet filters presented in Section 5.2 are even-length non-symmetric filters and in the current work, we have settled for the use of periodic extension for the rational wavelet transforms performed with these filters. A slight improvement on compression performance can be expected if an alternate boundary handling method is applied during the rational wavelet transform computation.

Chapter 6

An investigation of image coding using an overcomplete combined rational wavelet decomposition

In their most recent work on rational wavelet filter design, Bayram and Selesnick [17,32] have contemplated the idea of an *overcomplete* filter bank for rational wavelet transform. The filter bank is said to be *overcomplete*, in opposition to *critically-sampled*, to indicate that the total number of samples in the output signal is larger than the number of samples in the input signal. An interesting advantage of overcomplete filter banks is reduced complexity in the design of the rational wavelet filters, which ultimately results in less constraints and more flexibility in the design process. In the present chapter, we consider the potential use of an overcomplete filter bank for the rational wavelet transforms performed under the combined rational wavelet decomposition. Overcomplete wavelet transforms have been used in applications such as image and video denoising [55,56], but the idea of using an overcomplete transform for an image compression scheme certainly sounds paradoxical, since the purpose of image compression is to reduce the volume of data. However, recent work on image coding with the dual-tree wavelet transform [57,58], an overcomplete transform in

nature [59], has resulted in compression performances competitive with the state-of-the-art wavelet-based image coders. Positive results from the use of the dual-tree wavelet transform in video coding applications have also been reported [60,61]. Considering the potential avenue of overcomplete filter banks for rational wavelet filter design, the present chapter describes an attempt to extend our current work with rational wavelet decompositions to overcomplete representations. We present a set of highly redundant overcomplete filter banks proposed by Bayram and Selesnick [17] for rational wavelet transforms, and illustrate how the notion of combined rational wavelet decomposition can be extended to *overcomplete* combined rational wavelet decomposition. Accordingly, we describe how the parent-child tree construction scheme introduced in Chapter 3 can be extended to an overcomplete decomposition. Finally, we test the resulting framework in a wavelet-based image compression scheme that includes an overcomplete combined rational wavelet decomposition and an iterative thresholding method to reduce the total volume of significant data. Although the compression performance falls short of matching that of a wavelet-based image coder using a critically-sampled combined rational wavelet decomposition, the work described in the next sections is presented as a base for potential future research direction.

6.1 Overcomplete filter banks for rational wavelet transform

Figure 6.1 illustrates a typical filter bank structure for a one dimensional (p, q) rational wavelet transform, with the sampling coefficients p and q chosen such that $q = (p+1)$. Given an input signal of length l samples, the transform outputs a lowpass signal of

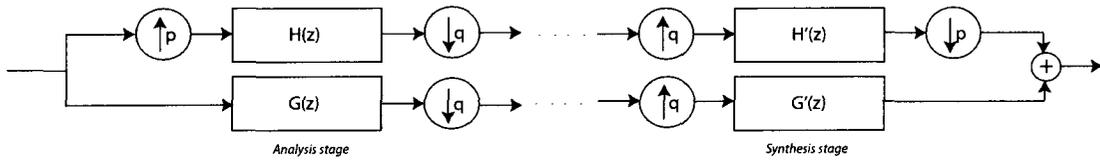


Figure 6.1: Critically-sampled filter bank structure for a one-dimensional (p, q) rational wavelet transform.

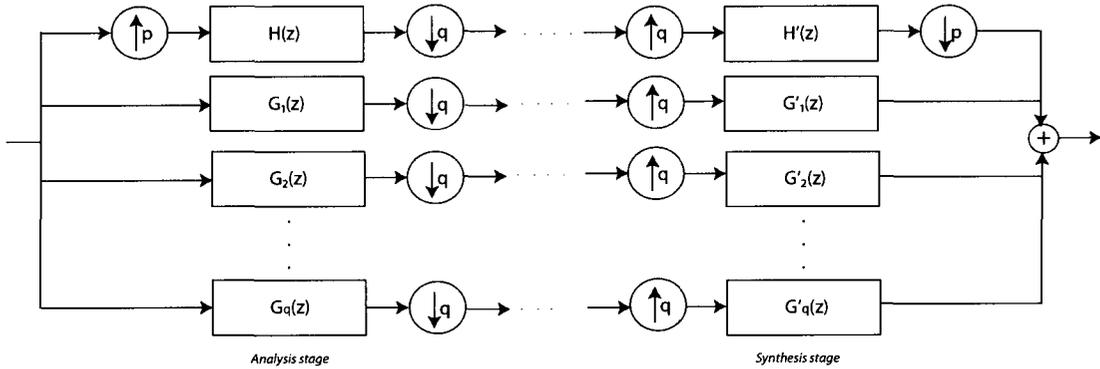


Figure 6.2: Overcomplete filter bank structure proposed by Bayram and Selesnick for a one-dimensional (p, q) rational wavelet transform.

length $(p/q) \times l$ samples and a highpass signal of length $(1/q) \times l$ samples. Because the total number of samples in the lowpass and highpass signals equals the number of samples in the input signal, the transform is said to be *critically-sampled*. On the other hand, an *overcomplete* (or *oversampled*) transform produces a higher number of samples in the output signals than the original number of samples of the input signal. Figure 6.2 illustrates an overcomplete filter bank designed by Bayram and Selesnick [17] for a wavelet transform of rational dilation factors and sampling coefficients p and q related by $q = (p + 1)^1$. Given an input signal of l samples, the illustrated filter bank produces a lowpass signal of $(p/q) \times l$ samples and multiple highpass signals of $(1/q) \times l$ samples. The total number of highpass signals from the analysis stage of the overcomplete filter bank is equal to the sampling coefficient q . Note that all the highpass signals have the same number of samples, which also corresponds to the

¹This constraint is specified by the authors

number of samples of the highpass signal from a critically-sampled filter bank. The *redundancy factor* is defined as the ratio between the total size (in number of samples) of the output signals against the size of the input signal. A critically-sampled filter bank has a redundancy factor of 1. The overcomplete filter bank of Figure 6.2 has a redundancy factor of $(p + q)/q$.

In the rest of the chapter, we use the overcomplete filter bank structure of Figure 6.2 to implement an overcomplete (3, 4) rational wavelet transform and an overcomplete (2, 3) rational wavelet transform. The wavelet filters of four vanishing moments used for each overcomplete rational wavelet transform were designed by Bayram and Selesnick [17]. The filters used in the analysis stage are presented in Figure 6.3. The synthesis filters can be obtained by time reversing the analysis filters.

6.2 Overcomplete combined rational wavelet decomposition

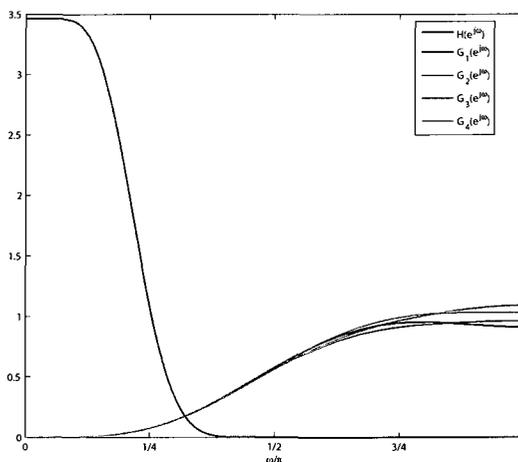
As with a critically-sampled transform, a two-dimensional overcomplete wavelet decomposition is performed by applying the one-dimensional filter bank along each of the two dimensions of the input data. Figure 6.4 illustrates the two-dimensional subband decompositions resulting from a regular (p, q) rational wavelet decomposition and from an overcomplete (p, q) rational wavelet decomposition. The overcomplete rational wavelet transform uses the filter bank described in Section 6.1, which produces a total of q highpass subbands along one dimension. The two-dimensional overcomplete wavelet decomposition produces three *sets* of *redundant* highpass subbands LH , HL and HH . The *sets* LH and HL corresponding respectively to the horizontal and vertical subbands are made of q *redundant subbands*. The set HH

Lowpass filter $h(n)$		Highpass filter $g_1(n)$	Highpass filter $g_2(n)$	Highpass filter $g_3(n)$	Highpass filter $g_4(n)$
0.000313630547715	-0.214691331005942	-0.134298355333610	0	0	0
0.002501601867151	-0.162882089732557	0.436980987431394	0.132367085534232	0	0
0.011340252235284	-0.071305028293876	-0.430855170548277	-0.432001608232518	0.136016795265619	0
0.036795337617036	0.007404262358666	0.036214978582619	0.432209700026642	-0.451246020553820	0.137889702104353
0.093577199932053	0.046231688411269	0.124418524354172	-0.053294952352157	0.472254300176830	-0.455885899775789
0.195496057705621	0.047108818244635	-0.016348508101007	-0.101564883653243	-0.094034244165871	0.470172202343323
0.343976713179003	0.028483694131348	-0.013033261735138	0.002514555822835	-0.081914839874902	-0.076257956154662
0.516737260320247	0.008881282567191	-0.002987363024772	0.017405035497270	-0.000414017869247	-0.105031274243844
0.665431205921812	-0.002473353703495	-0.000090457218301	0.002196931104627	0.016955196056301	0.013445647587266
0.730427507780420	-0.005411670507021	-0.000001374407078	0.000168136252311	0.002224222514909	0.012537603844824
0.669334638121312	-0.004026899706661			0.000158608450180	0.003055056455895
0.484014330946288	-0.001902375066090				0.000073056215738
0.227990923078991	-0.000588528800963				0.000001861622896
-0.015551529699564	-0.000102613701190				
-0.173009369610927					

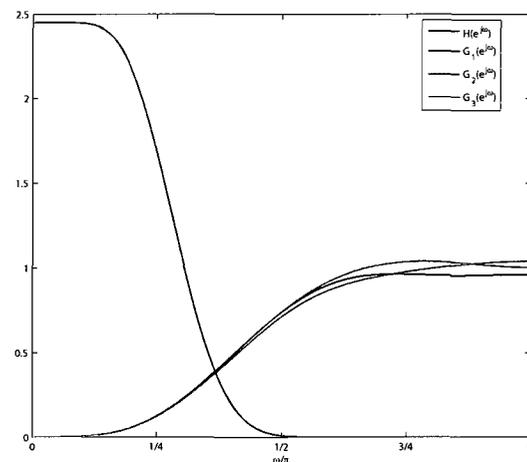
(a) Analysis filter coefficients for the (3, 4) rational wavelet filters

Lowpass filter $h(n)$		Highpass filter $g_1(n)$	Highpass filter $g_2(n)$	Highpass filter $g_3(n)$
0.000622483137494	-0.198115036854266	0.173458043509241	0	0
0.006363703439349	-0.148267271330018	-0.515396437209461	0.173273952965473	0
0.032226149358361	-0.038155951442624	0.408644819750306	-0.521464763976249	0.177228861658550
0.106412940144046	0.034556718341859	0.069719244041806	0.434348291278732	-0.534166110591272
0.253805876085218	0.043995155328702	-0.125561687346898	0.031985321833195	0.442183902226036
0.457744071728248	0.021145062251923	-0.042479819757916	-0.104237406402781	0.048533723851258
0.633664536932725	0.001171543346060	0.020311184951384	-0.041737143476859	-0.133066256724004
0.663647079800624	-0.004948288516924	0.008839261717061	0.014440740085950	-0.025612128702896
0.489965223179834	-0.003335659330251	0.002339316661416	0.011531844857770	0.012083025183805
0.185161407078212	-0.001052088172371	0.000122298089073	0.001652261118396	0.010625351962653
-0.086973529876513	-0.000144381846511	0.000003775593989	0.000206901716374	0.002064674343853
				0.000119772516473
				0.000005184275544

(b) Analysis filter coefficients for the (2, 3) rational wavelet filters



(c) Frequency responses, $p = 3, q = 4$



(d) Frequency responses, $p = 2, q = 3$

Figure 6.3: Analysis filter coefficients and frequency responses of the wavelet filters for the overcomplete (3, 4) and (2, 3) rational wavelet transforms.

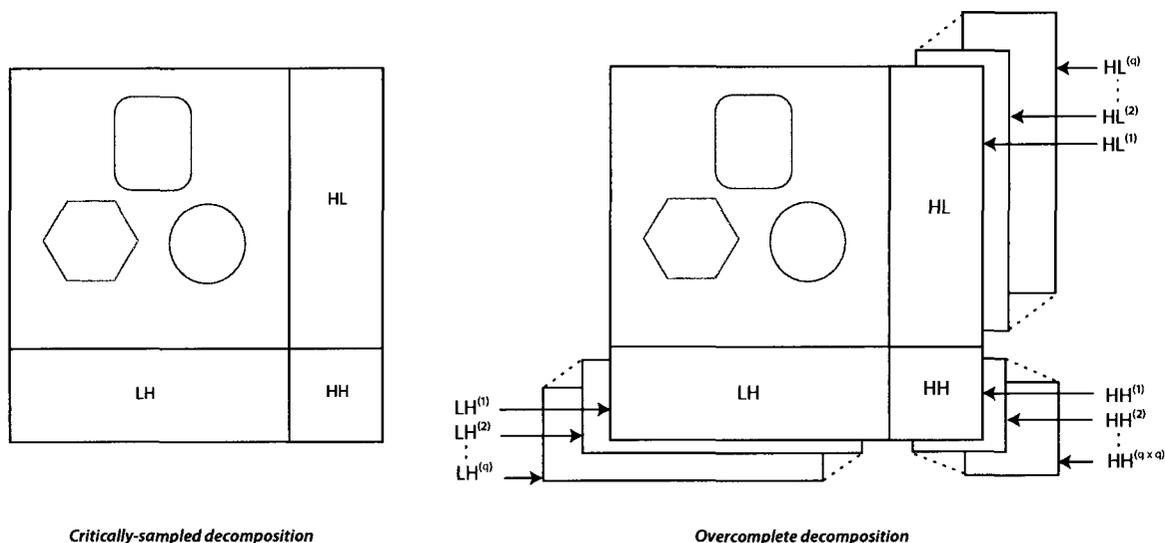


Figure 6.4: Subband decompositions obtained from a critically-sampled (p, q) rational wavelet transform and an overcomplete (p, q) rational wavelet transform. The illustrated structures correspond to one level of two-dimensional rational wavelet decomposition.

corresponding to the diagonal subbands contains $q \times q$ redundant subbands (the first decomposition along the rows produce q redundant subbands; then for each of the q subbands, another q redundant subbands are formed when the transform is applied along the columns, giving a total of $q \times q$ diagonal subbands).

We extend the notion of two-dimensional combined rational wavelet decomposition to *overcomplete* two-dimensional combined rational wavelet decomposition, where the critically sampled (p_i, q_i) rational wavelet transforms are replaced by the corresponding (p_i, q_i) overcomplete rational wavelet transforms. Figure 6.5 illustrates the subband decomposition obtained when the overcomplete combined rational wavelet decomposition is performed in *regular* mode and in the *wavelet packet extension* mode introduced in Chapter 3. The illustrated example is for an overcomplete combined decomposition with dilation factors $\{(q_i/p_i)\} = \{4/3, 3/2\}$. In the subband decomposition obtained with the wavelet packet extension mode, the extra sets of diagonal

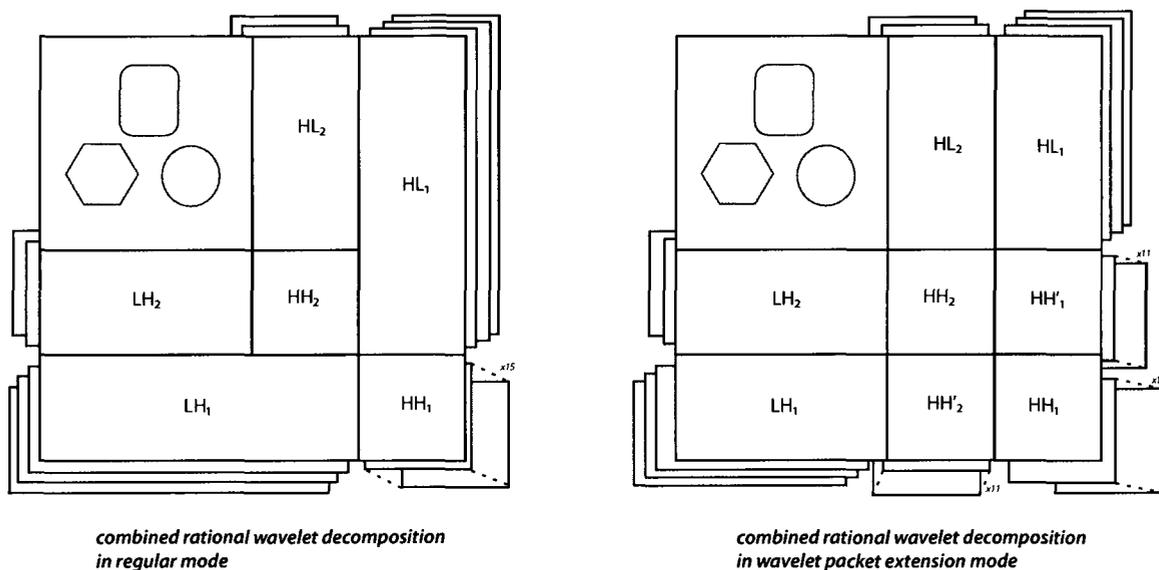


Figure 6.5: Subband decomposition from one level of overcomplete combined rational wavelet decomposition with dilation factors $\{(q_i/p_i)\} = \{4/3, 3/2\}$.

subbands HH'_1 and HH'_2 are each made of $q_1 \times q_2$ redundant subbands.

6.2.1 Extending the parent-child tree construction scheme to overcomplete combined wavelet decompositions

In Chapter 3, we described a different construction scheme of the parent-child tree structure (used by zerotree-based subband coders like SPIHT) for the subband decomposition obtained from a combined rational wavelet decomposition. We propose to extend that scheme to an overcomplete combined rational wavelet decomposition.

Consider a subband \mathcal{S}^m obtained after m levels of combined rational decomposition. With a critically sampled decomposition, the subband \mathcal{S}^m is a two-dimensional block of subband coefficients. With an overcomplete decomposition, \mathcal{S}^m is instead a set of redundant subbands $\mathcal{S}^m = \{\mathcal{S}^{m(1)}, \mathcal{S}^{m(2)}, \dots, \mathcal{S}^{m(t)}\}$ where t is the total number of redundant subbands. Each redundant subband $\mathcal{S}^{m(k)}$ is a two-dimensional block of data of the same size as a subband from a critically-sampled decomposition. In

the description below, a set of redundant subbands will often be viewed as a three-dimensional block of data. The method used to build the parent-child tree structure in the case of a critically-sampled wavelet decomposition can be expressed in the following rule:

Critically-Sampled rule :

A node p in subband \mathcal{S}^m is the parent of a 2×2 block of children nodes in subband \mathcal{S}^{m-1} . Each node of the 2×2 block is in turn a parent node of a 2×2 block of children nodes in subband \mathcal{S}^{m-2} . The scheme is applied to all subbands until all the nodes have been covered.

We propose two possible adaptations to extend that rule to an overcomplete wavelet decomposition. In one instance, we decide that for any set \mathcal{S}^m made of multiple redundant subbands, only the nodes located in the redundant subband $\mathcal{S}^{m(1)}$ can be parent nodes. The nodes in the other redundant subbands $\mathcal{S}^{m(k)}$ ($1 < k \leq t$) cannot be parent nodes. The tree construction rule can be expressed as follows :

Overcomplete Adaptation 1 :

A node p in the redundant subband $\mathcal{S}^{m(1)}$ is the parent of a $2 \times 2 \times t$ block of children nodes in the set \mathcal{S}^{m-1} . Each node of the $2 \times 2 \times t$ block that belongs to the redundant subband $\mathcal{S}^{m-1(1)}$ is in turn a parent node of a $2 \times 2 \times t$ block of children nodes in the set \mathcal{S}^{m-2} . The scheme is applied to all sets of redundant subbands until all the nodes have been covered.

The second possible adaptation considers any node from any redundant subband of the set \mathcal{S}^m as a potential parent node. The tree construction rule is a direct extension of the critically-sampled rule, and is expressed as follows:

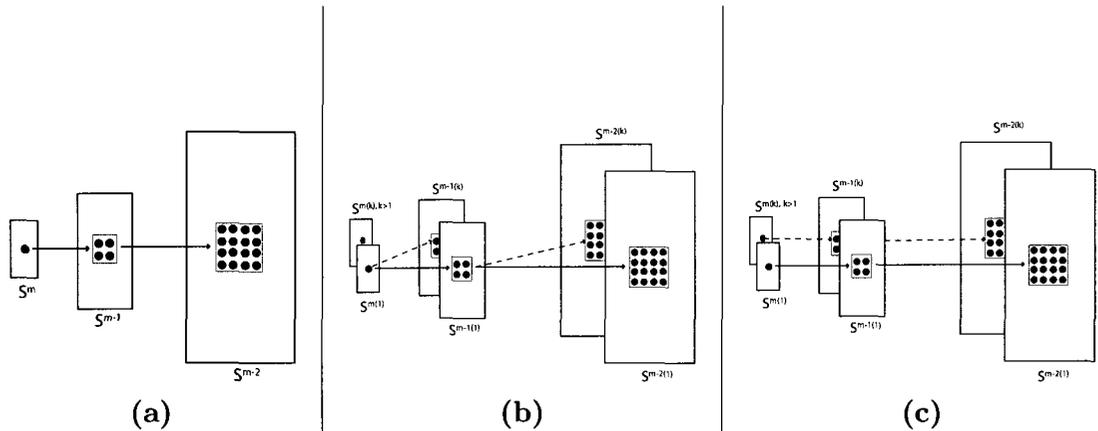


Figure 6.6: An illustration of the parent-child tree construction rules.
 (a) Critically-Sampled (b) Overcomplete Adaptation 1 (c) Overcomplete Adaptation 2

Overcomplete Adaptation 2 :

A node p in the redundant subband $\mathcal{S}^{m(k)}$ ($1 \leq k \leq t$) is the parent of a 2×2 block of children nodes in the redundant subband $\mathcal{S}^{m-1(k)}$ of the set \mathcal{S}^{m-1} . Each node of the 2×2 block is in turn a parent node of a 2×2 block of children nodes in the redundant subband $\mathcal{S}^{m-2(k)}$ of the set \mathcal{S}^{m-2} . The scheme is applied to all sets of redundant subbands until all the nodes have been covered.

Figure 6.6 illustrates the three rules described for the construction of the parent-child tree structure. The first adaptation proposed for overcomplete decompositions keeps the same total number of parent nodes as a critically-sampled decomposition and uses three-dimensional blocks of children nodes to handle the additional redundant subbands. This type of structure can be taken advantage of if the transform results in a high correlation between subband coefficients at the same spatial location on all the redundant subbands of a given set. The second adaptation creates additional parent nodes by having a separate parent-child tree for each redundant subband, and does not require the use of three-dimensional blocks of children nodes. It is a straightforward extension of the critically-sampled rule.

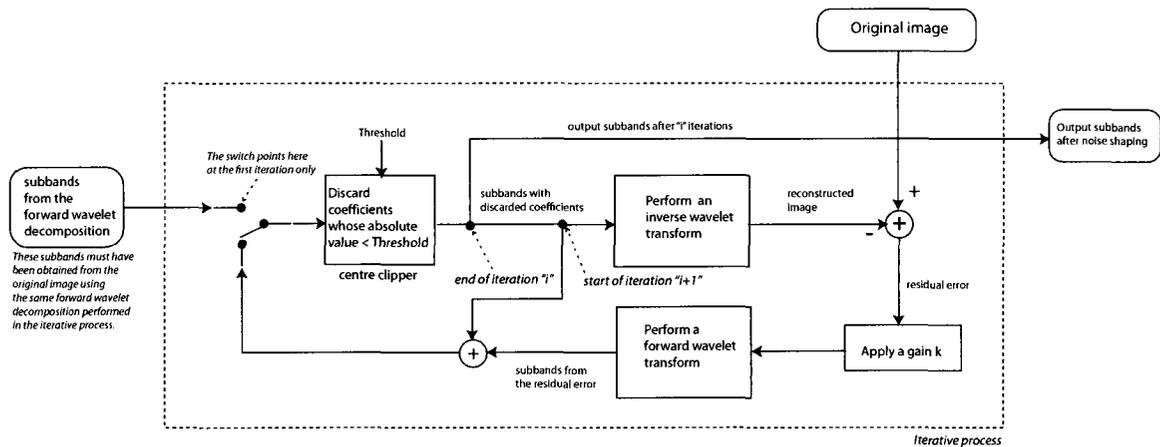


Figure 6.7: Description of the iterative noise shaping method by Reeves and Kingsbury.

6.3 Handling the redundancy in the overcomplete representation

An obvious challenge that arises from the use of an overcomplete transform in image compression is how to deal with the bigger volume of subband coefficients. In recent works on image and video coding using overcomplete complex wavelet transforms, an iterative method proposed by Reeves and Kingsbury, called *iterative noise shaping* [62], is used to reduce the number of *significant* coefficients by discarding low magnitude coefficients. The basic idea of iterative noise shaping is to discard coefficients (by setting their value to zero) that are below a given threshold and compensate for their loss by iteratively increasing the value of the remaining coefficients of larger magnitude, until a target reconstruction quality is achieved. Figure 6.7 describes the iterative process of noise shaping. At the first iteration, the wavelet subbands obtained from decomposing the original image are used as inputs to a centre clipper that discards any subband coefficient whose absolute value is less than the given threshold. From that point, every iteration consists of the following steps:

1. A reconstructed image is built from the output of the centre clipper; the residual

error between the reconstructed image and the original image is computed.

2. A gain k ($1 \leq k < 2$) is applied to the residual error. The effect of the gain k is to calibrate the amount of residual error that is reinserted for the next iteration. A higher gain results in more error reinserted, which ultimately leads to better reconstruction quality after the next iteration, but at the expense of more non-zeros coefficients.
3. A forward wavelet transform is performed on the residual error; the obtained subband coefficients are added to the initial output of the centre clipper.
4. The resulting subbands are given as input to the centre clipper. The output of the centre clipper is the output produced by the iteration, and the input to the next iteration.

The operation is repeated iteratively until a target quality is reached for the reconstructed image obtained after the inverse wavelet transform following the centre clipper. The target quality could be unreachable if the threshold used by the centre clipper is not selected appropriately. In practice, identification of an appropriate threshold for a given target image quality is done on an *ad hoc* basis. Once a target threshold has been identified, three thresholding strategies can be used [63] :

- a) Keep a constant threshold at each iteration: the threshold is set to the target threshold at the first iteration and kept constant for all iterations.
- b) Decrease the threshold at each iteration: at the first iteration, the threshold is set to a greater value than the target threshold, and is decreased at every iteration until its value reaches the target threshold.
- c) Increase the threshold at each iteration: at the first iteration, the threshold is set to a smaller value than the target threshold, and is increased at every iteration until its value reaches the target threshold.

We test the iterative noise shaping method using two levels of overcomplete combined rational wavelet decomposition at the wavelet transform stages. The experiment is run with the image *Boat* presented in Chapter 5, and using a decreasing threshold strategy

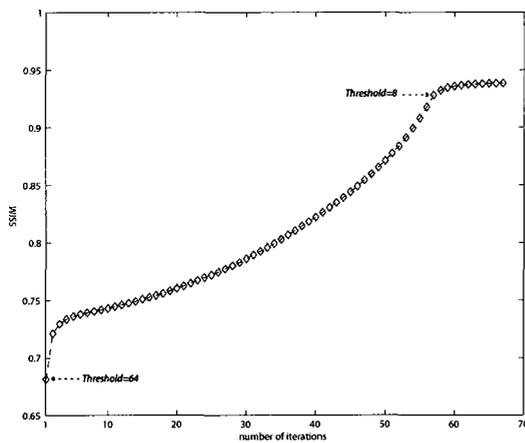


Figure 6.8: Objective quality scores of the reconstructed image after each iteration, for a target threshold of 8 and a starting threshold of 64.

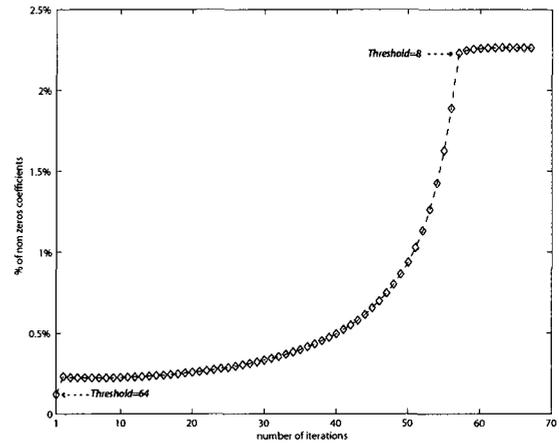


Figure 6.9: Percentage of non zeros coefficients in the highpass subbands after each iteration, for a target threshold of 8 and a starting threshold of 64.

with a target threshold of 8, a starting threshold of 64 and a decreasing step of 1. Once the threshold reaches the target number, an extra 10 iterations are performed for the results to converge. The starting and target thresholds selected are experimental numbers: the target threshold has been selected so that the reconstructed image quality is of reasonably good visual quality. The choice of a decreasing threshold strategy is also based on experimental results. Finally, the gain applied to the residual error in the iterative process is set to an *ad hoc* value of 1.6.

Figure 6.8 and Figure 6.9 illustrate how the quality of the reconstructed image improves after each iteration, at the expense of more significant data in the highpass subbands. The objective quality scores are expressed by using the Structural Similarity Index Metric (SSIM). We consider the percentage of non-zeros coefficients in the highpass subbands as a measure of the volume of *significant* data in the subbands. The different parameters selected for the iterative process, that is the starting and target thresholds and the gain, were experimentally selected on the criteria of balance between the quality of the reconstructed image and the volume of significant data in the highpass subbands. The effects of these parameters have been investigated by Reeves and Kingsbury [63]. With a starting threshold

of 64 and a target threshold of 8, the reconstructed image obtains an objective quality score of 0.938 and the highpass subbands contain about $1/40$ of the initial number of significant (non-zeros) coefficients. As a comparison, if a simple thresholding at a threshold of 8 is applied to the subbands coefficients from the overcomplete combined wavelet decomposition, it produces a reconstructed image with an objective quality score of 0.899, and the resulting highpass subbands would contain about $1/15$ of the initial number of significant coefficients.

6.4 Evaluation of performance at full size resolution of a wavelet-based image coder using overcomplete combined wavelet decomposition

We extend the wavelet-based image coder with combined rational wavelet decomposition introduced in Chapter 5 to an image coder that uses an overcomplete combined rational wavelet decomposition at the encoding and decoding stages. We built two variants of the new image coder: one that does not include any redundancy handling method, and another one that includes the described iterative noise shaping method to reduce the volume of significant data in the highpass subbands from the overcomplete decomposition. Figure 6.10 illustrates the two possible encoding stages. The structure of the decoding stage is the same for both variants. We alternatively use the original SPIHT and SPECK subband coders to encode the wavelet subbands. For SPIHT, we define two configurations according to the two parent-child tree construction rules presented in Section 6.2:

- *SPIHT-v1*, which uses the overcomplete adaptation 1 to build the parent-child tree structure.
- *SPIHT-v2*, which uses the overcomplete adaptation 2 to build the parent-child tree structure.

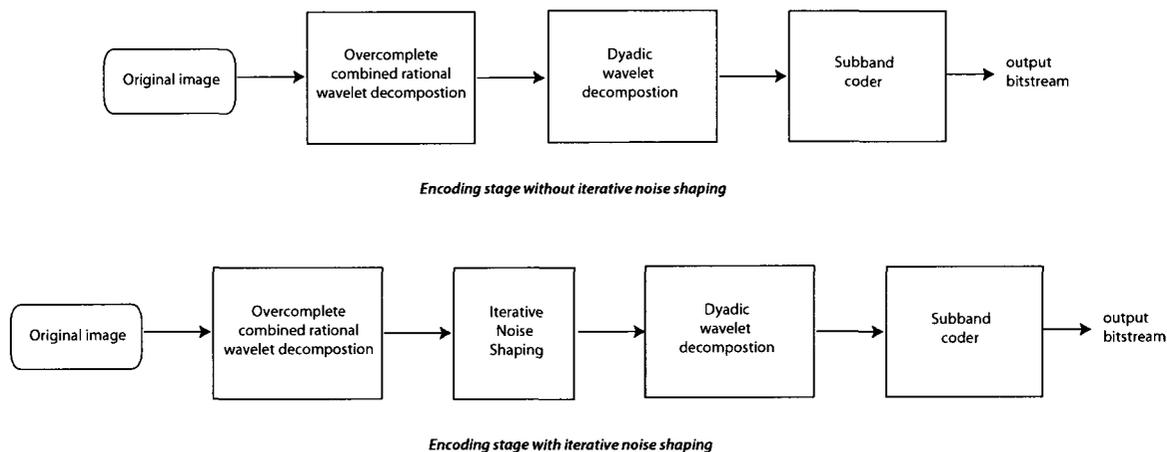


Figure 6.10: A description of the two possible encoding stages for the wavelet-based image coder with overcomplete combined rational wavelet decomposition.

With SPECK, since each subband is treated independently, support for the overcomplete combined decomposition is introduced easily by considering each redundant subband as an independent subband. The overcomplete rational wavelet transforms are performed with the rational wavelet filters described in Section 6.1, and periodic extension is used to handle the image boundaries. We use for the experiment the two test images *Boat* and *Girl* presented in Chapter 5, and a range of compression bitrates from 0.3 to 1.0 bits per pixel. A total of four levels of decomposition is specified, distributed in two levels of overcomplete combined wavelet decomposition and two levels of dyadic wavelet decomposition with the CDF9/7 filters. The results obtained at full size resolution are compared to those obtained in Chapter 5 with an image coder that uses a critically-sampled combined rational wavelet decomposition at the encoding stage.

Table 6.1 lists the objective quality scores obtained from decoding the images at full size resolution. The results are represented graphically in Figure 6.11. The output images *OC*, *OC + NS* and *CS* designate respectively:

- the decoded image obtained from the image coder using overcomplete combined decomposition and no iterative noise shaping,

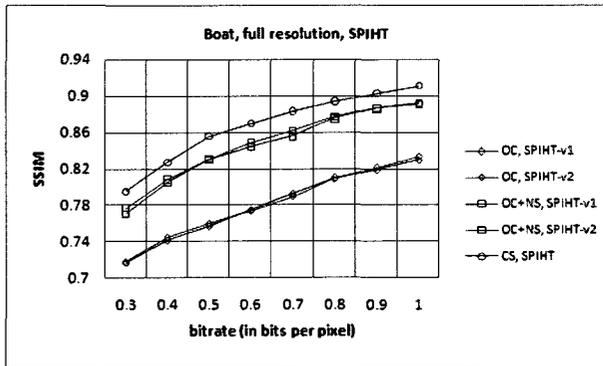
bitrate (in bpp)		0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
<i>Boat</i> , SPIHT	<i>OC, SPIHT-v1</i>	0.718	0.744	0.759	0.773	0.789	0.81	0.821	0.833
	<i>OC, SPIHT-v2</i>	0.716	0.741	0.757	0.774	0.792	0.808	0.818	0.829
	<i>OC + NS, SPIHT-v1</i>	0.776	0.807	0.831	0.848	0.862	0.877	0.886	0.892
	<i>OC + NS, SPIHT-v2</i>	0.771	0.805	0.830	0.843	0.856	0.875	0.886	0.891
	<i>CS, SPIHT</i>	0.795	0.827	0.855	0.870	0.883	0.894	0.903	0.911
<i>Boat</i> , SPECK	<i>OC</i>	0.720	0.745	0.762	0.783	0.801	0.813	0.824	0.834
	<i>OC + NS</i>	0.784	0.809	0.838	0.853	0.869	0.879	0.889	0.896
	<i>CS</i>	0.798	0.835	0.861	0.874	0.886	0.898	0.907	0.916

bitrate (in bpp)		0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
<i>Girl</i> , SPIHT	<i>OC, SPIHT-v1</i>	0.804	0.821	0.838	0.852	0.857	0.866	0.872	0.877
	<i>OC, SPIHT-v2</i>	0.802	0.823	0.835	0.848	0.856	0.862	0.869	0.877
	<i>OC + NS, SPIHT-v1</i>	0.853	0.873	0.895	0.911	0.918	0.927	0.935	0.938
	<i>OC + NS, SPIHT-v2</i>	0.851	0.868	0.892	0.908	0.915	0.924	0.935	0.937
	<i>CS, SPIHT</i>	0.871	0.897	0.911	0.923	0.935	0.943	0.948	0.953
<i>Girl</i> , SPECK	<i>OC</i>	0.807	0.827	0.838	0.848	0.854	0.864	0.872	0.881
	<i>OC + NS</i>	0.856	0.879	0.897	0.912	0.921	0.929	0.935	0.939
	<i>CS</i>	0.878	0.898	0.911	0.925	0.937	0.944	0.948	0.953

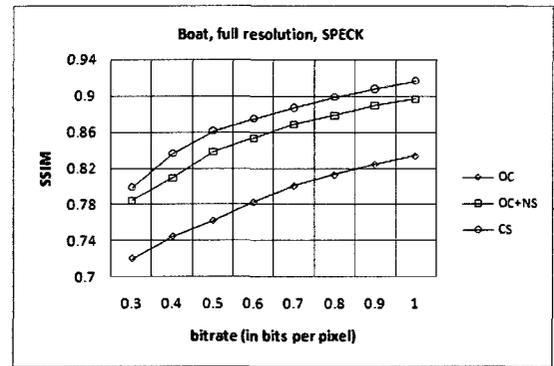
Table 6.1: Objective quality scores at full resolution for the images *Boat* and *Girl*.

- the decoded image obtained from the image coder using overcomplete combined decomposition with iterative noise shaping,
- the decoded image obtained from the image coder using critically-sampled combined decomposition.

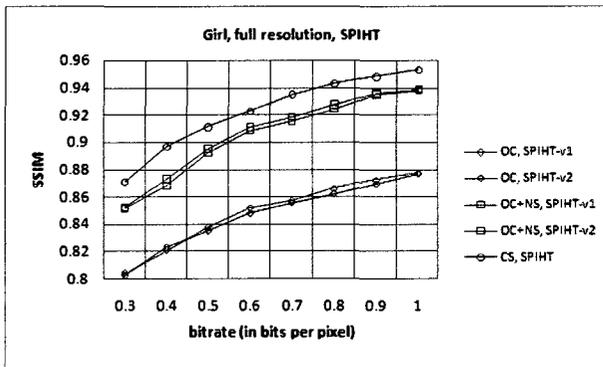
From visual inspection, a significant increase in image quality is observed at all bitrates when the iterative noise shaping is performed after the overcomplete decomposition. It is clearly illustrated in Figure 6.12 on an output image from *Boat*, encoded at a bitrate of 0.5 bits per pixel. In the illustration, the decoded image obtained from the image coder without iterative noise shaping lacks high frequency details, a consequence of the high volume of significant data in the highpass subbands. However, the decoded image obtained from the image coder using an overcomplete combined decomposition and iterative noise shaping does not match the quality of the image obtained from the image coder with critically-sampled decomposition. This can be observed across the whole range of bitrates from the objective quality measures.



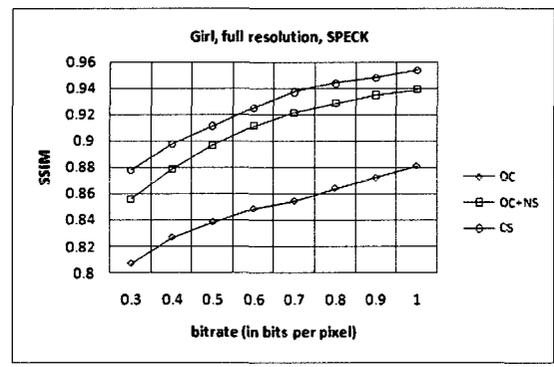
(a) *Boat*, SPIHT



(b) *Boat*, SPECK



(c) *Girl*, SPIHT



(d) *Girl*, SPECK

Figure 6.11: Compression performance at full resolution for the images *Boat* and *Girl*.



Figure 6.12: A part of the decoded image from *Boat*, compressed at 0.5 bpp with SPIHT. *From top to bottom:* output image from the image coder with overcomplete decomposition and no iterative noise shaping, output image from the image coder with overcomplete decomposition and iterative noise shaping, output image from the image coder with critically-sampled decomposition.

6.5 Analysis and discussion

The experimental results from Section 6.4 show that, although the inclusion of iterative noise shaping improves the results at full size resolution, the quality of the images obtained from the wavelet-based image coder with overcomplete combined decomposition is still far from matching what can be obtained if a critically-sampled decomposition is used instead. A subsequent analysis of the output bitstream from the encoding stage of the image coder shows that despite a reduction in the volume of significant data, the overcomplete decomposition still poses the problem of creating a significant coding overhead for the *position bits*. A subband coder encodes the subband coefficients by allocating bits to code the value of the coefficients (the *value bits*) as well as the location of the coefficients in the subband decomposition structure (the *position bits*). A subband decomposition with a higher number of samples ultimately results in more bits allocated to encode the position information for each significant coefficient. An analysis of the data produced by SPIHT in the case of the image coder with overcomplete decomposition and iterative noise shaping indeed reveals that a significant number of bits are produced during the *sorting pass* of the coding algorithm, which represents the stage in the algorithm where position bits are mostly produced. The overcomplete filter bank for rational wavelet transform introduced in Section 6.1 has a redundancy factor of 1.75 and 1.67 for the (3, 4) and (2, 3) rational wavelet transform, respectively. For overcomplete rational wavelet decompositions to be practical in image compression applications, a smaller redundancy factor is needed. Very recent work by Bayram and Selesnick [32] has produced an overcomplete filter bank for rational wavelet transform with a smaller redundancy factor than the filter bank used in the present chapter. However, the proposed filters have infinite impulse response, and have given us very poor compression performances.

In additional experiments conducted in the current research, we have also analyzed the visual quality of the images obtained at intermediate spatial resolutions with an image coder using overcomplete combined wavelet decomposition. However, the results were generally

poor, even with the use of iterative noise shaping. The inclusion of these results in the current chapter was not found to be of any benefit. As presented, the iterative process to reduce the volume of significant data is calibrated according to the quality of the reconstructed image at full size resolution *only*. To obtain more competitive results at intermediate spatial resolutions, the structure of iterative scheme should be adapted such that the quality of the reconstructed images at intermediate spatial resolutions can be controlled.

The work presented in this chapter is proposed to be an investigation of the potential use of an overcomplete filter bank for the combined rational wavelet decomposition discussed throughout this document. From the nature of the results, this has a potential for future research direction, if an overcomplete filter bank with smaller redundancy can be built, and if the overcomplete nature of the filter bank can be taken advantage of in the construction of rational wavelet filters of good compression properties.

Chapter 7

Conclusion and future directions

7.1 Concluding remarks

The main objective of the research presented in this thesis was to increase the set of possible decoding resolutions available from a wavelet-based image compression system built under a framework allowing the support of spatial resolution scalability. Indeed, a current limitation of the typical wavelet-based image coders discussed in the literature is the fact that the different spatial resolutions available to the decoder are related to the original resolution of the image by a dyadic ratio. In this research, we have proposed a new wavelet decomposition structure made of a combination of wavelet transforms of rational dilation factors. We showed that an adequate selection of the dilation factors of the different rational wavelet transforms can result in a wavelet decomposition that produces native spatial resolutions related to the original resolution by dyadic and non-dyadic factors. The subsequent implementation of a resolution-scalable image coder based on the proposed wavelet decomposition confirmed the added capability of scalability on a richer set of spatial resolutions. An evaluation of compression performance against a similarly built image coder based on a typical dyadic wavelet decomposition showed the potential of our proposed image coder, and also highlighted its current limitations.

Experimental results at different spatial resolutions illustrated some of the shortcomings

of the wavelet filters used for the rational wavelet transforms. In early chapters, we have briefly described how the field of wavelet filters for rational wavelet transforms has not yet matured to the level that can be observed in the current literature on wavelet filters for dyadic wavelet transforms. Our research did not bring any particular contribution to this field. Instead, we have presented how the rich base of dyadic wavelet filters can be taken advantage of by using a decomposition structure made of rational wavelet decompositions and dyadic wavelet decompositions. Subsequent performance evaluations have shown the positive effects of this combination. However, at the current stage of the research, an improvement in the quality of the rational wavelet filters will be a welcome addition to the work presented.

7.2 Discussion on potential future research directions

As mentioned, one of the challenges faced in this research was the quality of the rational wavelet filters for image coding applications. Significant developments have been made recently in the design of wavelet filters for rational wavelet transforms. A future work on building a set of filters with specific constraints defined to improve image coding performance is obviously a strong potential future direction for the current research. For instance, biorthogonal and symmetric filters have recently been the preferred choice in wavelet-based image or video compression. Inclusion of these two constraints in the construction of rational wavelet filters could be a first step in this research direction.

In addition, further analysis on the impact of combining different wavelet transforms in one subband decomposition structure could certainly be beneficial. A typical subband decomposition only involves one set of analysis filters. Analyzing the effect of alternating different analysis filters in the subband decomposition can lead to valuable information that can be exploited in the design of the wavelet filters or even in the subband coding phase.

Indeed, if the combination of different analysis filters results in a different distribution of magnitude between subband coefficients from many levels of decomposition, then an improvement can be obtained from an adequate adaptation of the subband coding method. As such, an in-depth study on the effects of combining wavelet transforms of different dilation factors would also be a future direction of interest for the research.

Finally, in Chapter 6, we have considered the use of an overcomplete filter bank for the rational wavelet transforms. Among the recent advances in the construction of rational wavelet filters, the use of an overcomplete filter bank is believed to introduce more flexibility in the design process, and allow more constraints to be imposed on the filters. However, from an image compression perspective, it presents the challenge of additional redundancy and induces the need for a method designed to reduce the volume of significant data. Our current work used an iterative thresholding scheme to limit the amount of significant coefficients in the subband. However, different alternatives could be potentially explored. For instance, the overcomplete filter bank described in Chapter 6 produces sets of redundant details subbands; an additional transform operation on these redundant subbands can potentially concentrate the significant data on fewer subbands, without the need for a lossy thresholding operation. The construction of such a transform, which would exploit the correlation between the redundant subbands, is a worthy area for future research targeted at handling the additional redundancy.

List of References

- [1] W. B. Pennebaker and J. L. Mitchell, *JPEG Still Image Data Compression Standard*. New York: Van Nostrand Reinhold, 1993.
- [2] A. Skodras, C. Christopoulos, and T. Ebrahimi, "The JPEG 2000 still image compression standard," *IEEE Signal Processing Magazine*, vol. 18, pp. 36–58, Sep 2001.
- [3] M. Rabbani and R. Joshi, "An overview of the JPEG 2000 still image compression standard," *Signal Processing: Image Communication*, vol. 17, no. 1, pp. 3 – 48, 2002.
- [4] C. Fogg, D. LeGall, J. Mitchell, and W. Pennebaker, *MPEG Video Compression Standard*. International Thomson Publishing, 1997.
- [5] T. Wiegand, G. Sullivan, G. Bjontegaard, and A. Luthra, "Overview of the H.264/AVC video coding standard," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 13, pp. 560–576, July 2003.
- [6] D. Marpe, T. Wiegand, and G. Sullivan, "The H.264/MPEG4 advanced video coding standard and its applications," *IEEE Communications Magazine*, vol. 44, pp. 134–143, Aug. 2006.
- [7] R. Dugad and N. Ahuja, "A fast scheme for image size change in the compressed domain," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 11, pp. 461–474, Apr 2001.
- [8] Y. S. Park and H. W. Park, "Arbitrary-ratio image resizing using fast DCT of composite length for DCT-based transcoder," *IEEE Transactions on Image Processing*, vol. 15, pp. 494–500, Feb. 2006.
- [9] L. Yu and L. White, "Complex rational orthogonal wavelet and its application in communications," *IEEE Signal Processing Letters*, vol. 13, pp. 477–480, Aug. 2006.
- [10] G. F. Choueiter and J. R. Glass, "An implementation of rational wavelets and filter design for phonetic classification," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 15, pp. 939–948, March 2007.

- [11] R. Xiong, J. Xu, and F. Wu, "A lifting-based wavelet transform supporting non-dyadic spatial scalability," in *IEEE International Conference on Image Processing*, pp. 1861–1864, Oct. 2006.
- [12] A. Baussard, F. Nicolier, and F. Truchetet, "Rational multiresolution analysis and fast wavelet transform: application to wavelet shrinkage denoising," *Signal Processing*, vol. 84, no. 10, pp. 1735 – 1747, 2004.
- [13] C. Segall and G. Sullivan, "Spatial scalability within the H.264/AVC scalable video coding extension," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 17, pp. 1121–1135, Sept. 2007.
- [14] E. Francois and J. Vieron, "Extended spatial scalability : A generalization of spatial scalability for non dyadic configurations," in *IEEE International Conference on Image Processing*, pp. 169–172, Oct. 2006.
- [15] I. Bayram and I. W. Selesnick, "Design of orthonormal and overcomplete wavelet transforms based on rational sampling factors," in *Wavelet Applications in Industrial Processing V* (F. Truchetet and O. Lalignant, eds.), vol. 6763, p. 67630H, SPIE, 2007.
- [16] I. Bayram and I. W. Selesnick, "Orthonormal FBs with rational sampling factors and oversampled DFT-modulated FBs: A connection and filter design," *IEEE Transactions on Signal Processing*, vol. 57, pp. 2515–2526, July 2009.
- [17] I. Bayram and I. Selesnick, "Overcomplete discrete wavelet transforms with rational dilation factors," *IEEE Transactions on Signal Processing*, vol. 57, pp. 131–145, Jan. 2009.
- [18] N. Adami, A. Signoroni, and R. Leonardi, "State-of-the-art and trends in scalable video compression with wavelet-based approaches," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 17, pp. 1238–1255, Sept. 2007.
- [19] N. Mehrseresht and D. Taubman, "A flexible structure for fully scalable motion-compensated 3-D DWT with emphasis on the impact of spatial scalability," *IEEE Transactions on Image Processing*, vol. 15, pp. 740–753, March 2006.
- [20] A. Said and W. Pearlman, "A new, fast, and efficient image codec based on set partitioning in hierarchical trees," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 6, pp. 243–250, Jun 1996.
- [21] W. Pearlman, A. Islam, N. Nagaraj, and A. Said, "Efficient, low-complexity image coding with a set-partitioning embedded block coder," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 14, pp. 1219–1235, Nov. 2004.

- [22] T. Nakachi, T. Sawabe, J. Suzuki, and T. Fujii, "A Study on Non-octave Scalable Image Coding and Its Performance Evaluation Using Digital Cinema Test Material," *IEICE Transactions on fundamentals of electronics, communications and computer science*, vol. E89-A, no. 9, pp. 2405–2414, 2006.
- [23] G. Pau, A. Pesquet-Popescu, and G. Piella, "Modified M-band synthesis filter bank for fractional scalability of images," *IEEE Signal Processing Letters*, vol. 13, pp. 345–348, June 2006.
- [24] T. Nakachi, T. Sawabe, J. Suzuki, and T. Fujii, "A study on non-octave scalable coding with filter bank and its performance evaluation using EBCOT," in *IEEE International Symposium on Communications and Information Technology*, vol. 2, pp. 841–844 vol.2, Oct. 2004.
- [25] G. Pau and B. Pesquet-Popescu, "Image coding with rational spatial scalability," in *Proceedings of EUSIPCO, Florence, Italy*, sept. 2006.
- [26] W. Sweldens, "The lifting scheme: a construction of second generation wavelets," *SIAM Journal on Mathematical Analysis*, vol. 29, no. 2, pp. 511–546, 1998.
- [27] I. Daubechies and W. Sweldens, "Factoring wavelet transforms into lifting steps," *Journal of Fourier Analysis and Applications*, vol. 4, pp. 247–269, May 1998.
- [28] O. Rioul, "On the choice of 'wavelet' filter for still image compression," in *IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 5, pp. 550–553 vol.5, Apr 1993.
- [29] J. Villasenor, B. Belzer, and J. Liao, "Wavelet filter evaluation for image compression," *IEEE Transactions on Image Processing*, vol. 4, pp. 1053–1060, Aug 1995.
- [30] T. Blu, "A new design algorithm for two-band orthonormal rational filter banks and orthonormal rational wavelets," *IEEE Transactions on Signal Processing*, vol. 46, pp. 1494–1504, Jun 1998.
- [31] J. Kovacevic and M. Vetterli, "Perfect reconstruction filter banks with rational sampling factors," *IEEE Transactions on Signal Processing*, vol. 41, pp. 2047–2066, June 1993.
- [32] A. Bayram and I. W. Selesnick, "Frequency-domain design of overcomplete rational-dilation wavelet transforms," *IEEE Transactions on Signal Processing*, vol. 57, pp. 2957–2972, Aug. 2009.
- [33] J. Shapiro, "Embedded image coding using zerotrees of wavelet coefficients," *IEEE Transactions on Signal Processing*, vol. 41, pp. 3445–3462, Dec 1993.

- [34] Z. Xiong, K. Ramchandran, and M. Orchard, "Wavelet packet image coding using space-frequency quantization," *IEEE Transactions on Image Processing*, vol. 7, pp. 892–898, Jun 1998.
- [35] F. Meyer, A. Averbuch, and J.-O. Stromberg, "Fast adaptive wavelet packet image compression," *IEEE Transactions on Image Processing*, vol. 9, pp. 792–800, May 2000.
- [36] B.-J. Kim, Z. Xiong, and W. Pearlman, "Low bit-rate scalable video coding with 3-D set partitioning in hierarchical trees (3-D SPIHT)," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 10, pp. 1374–1387, Dec 2000.
- [37] N. Sprljan, S. Grgic, and M. Grgic, "Modified SPIHT algorithm for wavelet packet image coding," *Real-Time Imaging*, vol. 11, no. 5-6, pp. 378 – 388, 2005. Special Issue on Multi-Dimensional Image Processing.
- [38] S.-T. Hsiang and J. Woods, "Embedded image coding using zeroblocks of sub-band/wavelet coefficients and context modeling," in *IEEE International Symposium on Circuits and Systems*, vol. 3, pp. 662–665, May 2000.
- [39] D. Taubman, "High performance scalable image compression with EBCOT," *IEEE Transactions on Image Processing*, vol. 9, pp. 1158–1170, Jul 2000.
- [40] H. Danyali and A. Mertins, "Fully spatial and SNR scalable, SPIHT-based image coding for transmission over heterogenous networkss," *Journal of Telecommunications and Information Technology*, vol. 2, pp. 92–98, 2003.
- [41] H. Danyali and A. Mertins, "Flexible, highly scalable, object-based wavelet image compression algorithm for network applications," *IEE Proceedings on Vision, Image and Signal Processing*, vol. 151, pp. 498–510, Dec. 2004.
- [42] E. Christophe and W. A. Pearlman, "Three-dimensional SPIHT coding of volume images with random access and resolution scalability," *EURASIP Journal on Image and Video Processing*, vol. 2008, 2008.
- [43] G. Xie and H. Shen, "Highly scalable, low-complexity image coding using zeroblocks of wavelet coefficients," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 15, pp. 762–770, June 2005.
- [44] X. Tang and W. A. Pearlman, "Progressive resolution coding of hyperspectral imagery featuring region of interest access," in *Visual Information Processing XIV* (Z. ur Rahman, R. A. Schowengerdt, and S. E. Reichenbach, eds.), vol. 5817, pp. 270–280, SPIE, 2005.
- [45] J. Fowler, "Shape-adaptive coding using binary set splitting with k-d trees," in *International Conference on Image Processing*, vol. 2, pp. 1301–1304 Vol.2, Oct. 2004.

- [46] S.-T. Hsiang and J. W. Woods, “Embedded video coding using invertible motion compensated 3-D subband/wavelet filter bank,” *Signal Processing: Image Communication*, vol. 16, no. 8, pp. 705 – 724, 2001.
- [47] P. Chen and J. Woods, “Bidirectional MC-EZBC with lifting implementation,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 14, pp. 1183–1194, Oct. 2004.
- [48] M. Antonini, M. Barlaud, P. Mathieu, and I. Daubechies, “Image coding using wavelet transform,” *IEEE Transactions on Image Processing*, vol. 1, pp. 205–220, Apr 1992.
- [49] M. Unser and T. Blu, “Mathematical properties of the JPEG2000 wavelet filters,” *IEEE Transactions on Image Processing*, vol. 12, pp. 1080–1090, Sept. 2003.
- [50] G. Strang and T. Nguyen, *Wavelets and Filter Banks*. Wellesley-Cambridge Press, 1995.
- [51] I. H. Witten, R. M. Neal, and J. G. Cleary, “Arithmetic coding for data compression,” *Communications of the ACM*, vol. 30, no. 6, pp. 520–540, 1987.
- [52] Z. Wang, A. Bovik, H. Sheikh, and E. Simoncelli, “Image quality assessment: from error visibility to structural similarity,” *IEEE Transactions on Image Processing*, vol. 13, pp. 600–612, April 2004.
- [53] Z. Wang and A. Bovik, “Mean squared error: Love it or leave it? a new look at signal fidelity measures,” *IEEE Signal Processing Magazine*, vol. 26, pp. 98–117, Jan. 2009.
- [54] D. Taubman and M. W. Marcellin, *JPEG2000 Image compression Fundamentals, Standards and Practice*. Kluwer Academic Publishers, 2002.
- [55] J.-L. Starck, E. Candes, and D. Donoho, “The curvelet transform for image denoising,” *IEEE Transactions on Image Processing*, vol. 11, pp. 670–684, Jun 2002.
- [56] I. W. Selesnick and K. Y. Li, “Video denoising using 2D and 3D dual-tree complex wavelet transforms,” in *Wavelets: Applications in Signal and Image Processing X* (M. A. Unser, A. Aldroubi, and A. F. Laine, eds.), vol. 5207, pp. 607–618, SPIE, 2003.
- [57] J. Yang, Y. Wang, W. Xu, and Q. Dai, “Image coding using dual-tree discrete wavelet transform,” *IEEE Transactions on Image Processing*, vol. 17, pp. 1555–1569, Sept. 2008.
- [58] J. E. Fowler, J. B. Boettcher, and B. Pesquet-Popescu, “Image coding using a complex dual-tree wavelet transform,” in *Proceedings of the European Signal Processing Conference (EUSIPCO), Poznan, Poland, sept 2007*.

- [59] I. Selesnick, R. Baraniuk, and N. Kingsbury, “The dual-tree complex wavelet transform,” *IEEE Signal Processing Magazine*, vol. 22, pp. 123–151, Nov. 2005.
- [60] J. Boettcher and J. Fowler, “Video coding using a complex wavelet transform and set partitioning,” *IEEE Signal Processing Letters*, vol. 14, pp. 633–636, Sept. 2007.
- [61] B. Wang, Y. Wang, I. Selesnick, and A. Vetro, “Video coding using 3D dual-tree wavelet transform,” *EURASIP Journal on Image and Video Processing*, vol. 2007, no. Article ID 42761, 2007.
- [62] T. H. Reeves and N. G. Kingsbury, “Overcomplete image coding using iterative projection-based noise shaping,” in *Proc. of IEEE Intern. Conf. on Image Processing*, vol. 3, pp. 597–600, June 2002.
- [63] N. G. Kingsbury and T. Reeves, “Iterative image coding with overcomplete complex wavelet transforms,” in *Visual Communications and Image Processing* (T. Ebrahimi and T. Sikora, eds.), vol. 5150, pp. 1253–1264, SPIE, 2003.