PROGRESSIVE IMAGE TRANSMISSION USING FRACTAL AND WAVELET TECHNIQUES WITH IMAGE COMPLEXITY MEASURES

by

RICHARD M. DANSEREAU

A Ph.D. Thesis

Submitted to the Faculty of Graduate Studies

in Partial Fulfillment of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

Department of Electrical and Computer Engineering

University of Manitoba

Winnipeg, Manitoba, Canada

Thesis Advisor: W. Kinsner, Ph.D., P. Eng.

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

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

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

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

0-612-57505-5

Canadä

THE UNIVERSITY OF MANITOBA

FACULTY OF GRADUATE STUDIES
*****
COPYRIGHT PERMISSION PAGE

Progressive Image Transmission Using Fractal and Wavelet

Techniques with Image Complexity Measures

BY

Richard M. Dansereau

.

A Thesis/Practicum submitted to the Faculty of Graduate Studies of The University

of Manitoba in partial fulfillment of the requirements of the degree

of

Doctor of Philosophy

RICHARD M. DANSEREAU © 2001

ABSTRACT

This thesis presents the theoretical and experimental development of progressive image transmission techniques involving fractals and wavelets, with emphasis on progressive image complexity measures to evaluate and guide the image decomposition. A new and novel progressive image transmission technique is presented where textures are synthesized to recreate an image. The textures are synthesized by generating fractal surfaces such that they interpolate control points, resulting in a higher level representation of an image. From this work, it was conjectured that fractal and multifractal complexity measures can serve as quantitative quality measures, since these dimensions characterize object complexity. The framework and experimentation for a complexity measure is developed based on the Rényi generalized entropy, the Rényi dimension spectrum, and the Mandelbrot spectrum. This framework is extended to the newly introduced relative Rényi dimension spectrum, which forms a new class of measures referred to as relative multifractal dimensions. Experimental results show that these multifractal dimensions, and in particular the relative Rényi dimension spectrum, has properties consistent with an image quality measure and correlate well with psychovisual characteristics. It is shown that the relative Rényi dimension spectrum is more resilient to calculation errors as compared to the other image quality measures. These image complexity measures are used to analyze and identify of regions of complexity disparity in an image for wavelet based progressive image transmission. Finally, the theoretical framework is developed to extend the idea of additive information cost functions in wavelet packet best basis searches such that the Rényi generalized entropy can serve as an entropy based information cost function.

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS AND ACRONYMS

1D	**One** dimensional
2D	**Two** dimensional
bpp	**Bits per pixel**
CAD	Computer aided design
CWT	Continuous wavelet transform
DAUB4	**Daub**echies 4-tap wavelet filter bank [Daub92]
dB	Decibel
DCT	Discrete cosine transform
DWPT	Discrete wavelet packet transform
DWT	Discrete wavelet transform
EZW	Embedded zerotree wavelet coding [Shap93]
fBm	Fractional Brownian motion [MaVa68]
FSSI	Fractal surface segmentation and interpolation
GIF	Graphics interchange format [GIF87a], [GIF90]
HVS	Human visual system
Hz	Hertz
ICWT	Inverse continuous wavelet transform
IFS	Iterated function systems [Barn88], [Jacq90], [WoJa99]
IQM	Image quality measure
IWT	Inverse wavelet transform
IDWT	Inverse discrete wavelet transform

IZT	Inverse zero-tree wavelet coefficient coding [Shap93]
JPEG	Joint photographic expert group
JND	Just noticeable difference
LSB	Least significant bit
LZ77	Lempel-Ziv 1977 data compression algorithm [ZiLe77]
LZ78	Lempel-Ziv 1978 data compression algorithm [ZiLe78]
LZW	Lempel-Ziv-Welch data compression algorithm [Welc84]
mfBm	Multifractional Brownian motion [PeVe95]
MOS	Mean opinion score
MSB	Most significant bit
MSE	Mean squared error
pdf	Probability distribution function
PIC	Perceptual subband image coder [SaJo89]
pixel	Picture element
PQS	Picture quality scale [LuAE95]
PSNR	Peak signal-to-noise ratio
RGB	Red, green, and blue
RMPD	Random midpoint displacement [Carp80], [FoFC82]
RMSE	Root mean squared error
RCD	Regions of complexity disparity
SNR	Signal-to-noise ratio
SPIHT	Set partitioning in hierarchical trees [SaPe96]
vel	Volume element

WPT	Wavelet packet transform
WT	Wavelet transform
ZT	Zero-tree wavelet coefficient coding [Shap93]

LIST OF SYMBOLS

General

E	Embedding dimension, the dimension of the support for an object (*i.e.* a surface would have $E = 2$)
ε	Expectation or the expected value
e_i	Vector with all zeros except the ith component with unit impulse (*i.e.* e_5 as all zeros except for the 5th component which is one).
I	Light intensity level
$\bar{f}(x)$	The bar denotes the complex conjugate of $f(x)$
$n(x;\mu, \sigma)$	A normal or Gaussian random variable with mean μ and standard deviation σ. Default if not specified is $\mu = 0$ and $\sigma = 1$
$u(x)$	Uniformly distributed random variable over $[0, 1)$
$W(s)$	Wiener process (stochastic process)
x and y	Position axes indicators
z	Complex plane

Sets

$\mathbf{N}$	Set of all positive integers including 0
$\mathbf{Z}$	Set of all integers
$\mathbf{R}$	Set of all real numbers
$\mathbf{R}^+$	Set of all positive real numbers
$\mathbf{C}$	Set of all complex numbers

Images

$f(x, y)$ Pixel (x, y) from an image (generic array-based image)

$f_{orig}(x, y)$ Pixel (x, y) from an original image, generally before lossy compression

$f_{peak}(x, y)$ Pixel (x, y) from an image where each pixel is the maximum pixel value

$f_r(x, y)$ Pixel (x, y) from a reconstructed image, generally after lossy compression

Fractals/Multifractals

α Hölder exponent, singularity point for probabilities in Mandelbrot spectrum

α_q Hölder exponent for a specific moment order q of the probabilities p_j in the Rényi generalized entropy $H_q(X)$

B Volume element or box used in measuring fractal and multifractal dimensions

β Spectral exponent, the exponent needed in the $\dfrac{1}{freq^\beta}$ to fit a single fractal object power spectrum curve

$B(t, w)$ Brownian motion

B_{H*} Fractional Brownian motion with Hurst exponent $H*$

δ Delta value (used in Brownian motion definition)

D Fractal dimension of an object (critical exponent of power-law relation)

D_q Rényi fractal dimension spectrum, multifractal dimension measure that gives the spread of single fractals in an object

D_{HB} Hausdorff-Besicovitch fractal dimension, morphological fractal dimension generally measured with a vel mesh covering

D_I Information fractal dimension, fractal dimension measure based on Shan-

non entropy measures

$f(\alpha) \equiv f_q$	Mandelbrot spectrum, multifractal dimension measure that gives the spread of single fractals in an object
H^*	Hurst exponent, relates to fractal dimension and used in fBm
$H(X)$	Shannon entropy for a discrete random variable X
$H_q(X)$	Generalized Rényi entropy of order q for a discrete random variable X, generalization of Shannon entropy $H(X)$
l	Displacement in Brownian and fractional Brownian motion
N_s	Number of yardsticks needed at scale s
p_j	Probability of a specific fractal object occurrence inside the vel j
q	Moment order for the probabilities p_j used in the Rényi generalized entropy $H_q(X)$ and in the Rényi dimension spectrum D_q
r	Base yardstick used to measure a fractal
s	Scale of yardstick or vel being used to measure a fractal

Wavelets/Fourier Transforms

a	Approximation signal after a filter bank pass in the wavelet transform
d	Detail signal after a filter bank pass in the wavelet transform
$F(\omega)$	Fourier transform
$F[\cdot]$	Operator denoting the Fourier transform
$F^{-1}[\cdot]$	Operator denoting the inverse Fourier transform
H and **G**	Quadrature mirror filter bank pair used in the wavelet transform
ϕ_k	kth element in an orthonormal basis

CHAPTER I

INTRODUCTION

1.1   Problem Definition

Transmission and manipulation of large images such as aerial or satellite images of city streets and natural resources proves to be a difficult and time consuming problem. These images are often used by city planners, telephone companies, hydro companies, and natural resource institutions, and therefore require high resolution for project planning. One application for high resolution images includes planning tasks such as deciding where to place new sewer pipelines or telecommunication lines. Another application is in analyzing satellite images of farm land to determine the best application pattern of pesticides, herbicides, and nutrient supplements. Other applications include astronomic image data warehousing, medical image database transmission, and photographic cartography.

Unfortunately, images of the required resolution are often large, such as the example aerial rural image in Fig. 1.1 [LGI95] which is 25 megabytes in size. The sheer size and high resolution required for images of this class makes storage, transmission, and manipulation of these images prohibitive. To compound these problems there may be hundreds or thousands of these images in a database. This size and quantity makes perusal through a large database of these images and focusing only on images of interest even more difficult.

Fig. 1.1. Example high resolution image of a rural area
(8-bit greyscale 5000x5000 pixels) [LGI95].

These problems are addressed through a number of approaches. The first approach is to have every user maintain their own copy of the image database. This way the user does not have to worry about transmission of the images during database perusal. The user can download the images at *off-peak hours*. This approach requires that the user has enough storage space for the database, which is often an unrealistic assumption, particularly for large sets of high resolution images. Also, there may not be sufficient off-peak hours for this download to occur. Additionally, the notion of having to synchronize and maintain multiple, duplicate image databases for each user is unappealing since up-to-date images may be desired at all times from a central source image database.

A second approach is to transmit the image at the request of the user. This is a more dynamic approach where little data is stored at the user's end. From a storage and usage viewpoint, this is the preferred method. Unfortunately, the large size of the images requires a lot of bandwidth and a lot of time to transmit the image as the user sits and waits. With the exception of image transmission time, this is the preferred method for most of the mentioned applications.

Proper ordering and control of the flow of the transmitted data allows for further improvements which make the second approach using on-demand image transmission more appealing. First of all, image compression techniques can help reduce the amount of actual data that must be transmitted from source to destination. In addition, if image compression is done in a manner where increasing levels of image detail are transmitted and then reconstructed, then important features within the reconstructed image can be recognized even before all of the fine image details have been transmitted. At the early stages of image transmission and reconstruction, if it is noticed that the image is not desired then the transmission of the image can be terminated early, and other images in the database can be then downloaded for inspection. This extra control over the image transmission is a useful feature to help reduce total transmission time and improve usability of any such large database of high resolution images.

Progressive image transmission is therefore an important aspect of the image transmission methodologies sought in this thesis. This form of image transmission allows for increasing levels of detail to be transmitted at the request of the user and can allow for psychovisually relevant image features to be transmitted first.

1.2   Objectives and Goals

The purpose of this thesis is to research and develop methodologies for the efficient transmission of digital images, with particular interest in the progressive transmission of the image information and the subsequent evaluation of the image reconstruction quality. The primary research focus and goal of this thesis is fourfold.

1.   How an image can be decomposed into psychovisually relevant features.

2.   How to select and order these image features for transmission over narrow-band channels.

3.   How to reconstruct the image at the receiving end as the data and image information is received.

4.   How to measure the quality of the resulting reconstructed image and image transmission process.

As will be evidenced throughout this thesis, these four points are not mutually exclusive, and should be considered as a whole, as well as intertwined and interdependent, when doing research and development for the efficient transmission of digital images.

To address these goals, this thesis focuses on wavelet, wavelet packet, and fractal techniques as the primary analysis and decomposition tools for the image compression and progressive image transmission. In addition to the actual image compression and progressive image transmission, these tools will be used in the development of new image quality measures to assist in the selection of the relevant image features. These measures will help guide the selection of parameters for the progressive image transmission. This is

done with the goal of improving the image reconstruction quality at the earliest stages of the transmission. Also, these measures will be used to evaluate the progressive image reconstructions, from the start of the reconstruction to the end, to determine the change in image quality throughout the progressive transmission.

1.3   Organization of this Thesis

This thesis consists of eight chapters. Chapter 1 states the purpose for the thesis, discusses the major problems to be addressed by the thesis, and provides some motivation for the thesis.

Chapter 2 provides background information on image compression and progressive image transmission. This background is from a general point of view without focus on specific coding techniques. A brief description of current progressive image transmission techniques is also given in this chapter.

In Chapter 3, a background on fractals, multifractals, and fractional Brownian motion is given. This background is later used to develop an image compression and progressive image transmission scheme using fractal surface segmentation and interpolation [DaKi96], [DaKi97], [DaKi98a]. In addition, this background serves as the basis for developing image complexity measures to help in guiding the selection of parameters for image compression and progressive image transmission techniques [DaKi98b], [DaKi99a], [DaKi99b], [DaKi00].

Chapter 4 provides background on wavelets and wavelet packets to be used in some of the progressive image transmission schemes developed. Experiments in later

chapters use wavelet based progressive image transmission schemes as the base scheme for the evaluation of the developed image complexity measures.

Chapter 5 introduces a new method for image compression and progressive image transmission using fractal surface segmentation and interpolation [DaKi96], [DaKi97], [DaKi98a]. This technique is developed using two dimensional fractional Brownian motion to synthesize textures that are used to recreate the image.

Chapter 6 introduces new methods for measuring the quality of image reconstructions based on the Rényi generalized entropy [DaKi99a] and the Rényi dimension spectrum measures [DaKi98b] as well as the Mandelbrot spectrum [DaKi98b], [DaKi99a], [DaKi99b]. This measuring of quality is done through the measurement of the complexity characteristics of the image and comparing to the original. A new class of multifractal dimensions is also presented which will be referred to as relative multifractal dimension measures. With this new class of relative multifractals, a new measure is presented which will be referred to as the relative Rényi dimension spectrum. These measures are used to determine the quality of progressive image transmissions and to judge the quality of the image reconstructions through signal complexity differences. The development of these measures results from observations made with the fractal surface segmentation and interpolation methods presented in Chapter 5. With these measures, better selection of parameters for the various progressive image transmission techniques can be done to improve the overall transmission and reconstruction of the images, especially at the earliest stages of the transmission. These results are compared with mean opinion scores to establish psychovisual correlations [DaKi00].

Chapter 7 develops progressive image transmission techniques using wavelets and wavelet packets. A pivotal portion of Chapter 7 is not only how wavelets and wavelet packets can be used for progressive image transmission, but, also in how the selection of wavelet coefficients and wavelet packet bases can be done using the developed multifractal measures from Chapter 6. This selection of wavelet coefficients and wavelet packet bases is performed to improve the psychovisual representation of the image at the earliest stages of the transmission.

Finally, Chapter 8 gives concluding remarks about the progressive image transmission and image complexity measure research, with some recommendations for future research.

Please note that due to the size of this dissertation, a separate technical report will be released giving the source code used in the research described in the rest of this dissertation [Dans01].

CHAPTER II

BACKGROUND ON IMAGE COMPRESSION AND PROGRESSIVE IMAGE TRANSMISSION

This chapter gives an overview of a number of different theoretical and practical concepts needed for the rest of this thesis. The first issue discussed is general image compression concepts as well as some of the elementary ideas behind progressive image transmission. This will give a basis for the goal of the thesis in general terms.

2.1   Image Compression

This section gives a brief description of some of the concepts behind image compression that are needed as a background for this thesis. This material includes a brief introduction to digital image representation, how a general image compression scheme works, and how limitations in the human psychovisual system allow for improved image compression schemes. Taking advantage of these limitations can reduce bit rates even further than standard data compression techniques would allow.

2.1.1   Digital Image Representation in Computers

To start this section on image compression, an outline of digital image representation is useful to get an idea of what class and scope of images will be dealt with in this thesis. There are two basic methods for representing a digital image. The first and by far the most common method is to represent the image as a rectangular array of small *picture elements* known as pixels. These pixels vary in light intensity and in colour depending on the

nature of the image being represented. For example, computer monitors and television sets display images in this fashion. The second method of representing an image is through vector graphics/object-based representations. With this representation an image is broken down into a set of objects such as, but not limited to, points, lines, planes, squares, circles, and spheres. The position and colour/texture of these objects are stored, and the proper placement of these objects within a scene recreates the image. This technique is used most often in *computer aided design* (CAD), virtual reality, or computer generated graphics. This thesis focuses on natural images, such as the aerial image from Fig. 1.1, so the first image representation method will be used, where the image is digitized into an array of pixels. This choice is made because of the current difficulty of performing accurate and meaningful object segmentation as is necessary to make the vector graphics/ object-based approach useful on arbitrary images.

When dealing with array-based image representations, some important questions behind the image representation are "What is the pixel depth/dynamic range of the image?" and "What type of greyscale or colour pixel representation does the image use?". In other words, these questions ask how a pixel is represented. The following discusses general pixel representation in an uncompressed format. Note that these *uncompressed representations* are important to understand since it is this representation that acts as the yardstick in this thesis when determining what the compression rate or compression ratio is of a compressed version of the image.

The first and simplest pixel representation comes from a bitmap representation used for monochrome images. Monochrome images, or bi-level images, contain two

shade possibilities for each pixel. This is usually the colours black and white but may also be any other two colours used to represent the foreground and background colour of an image. A monochrome image requires only 1 *bit per pixel* (bpp) to represent the black or white since the bit can be either 0 or 1. Monochrome images therefore have a pixel depth or dynamic range of 1 since each pixel requires only 1 bit. These 1 bit pixels are normally strung together in horizontal scan lines that span vertically across the image to form the bitmap representation of the image.

Greyscale images are an extension to monochrome images. A greyscale image can consist of pixels that are black, white, and a number of shades of grey in between. The number of bits used for each pixel determines the number of shades of grey possible in an image. A fairly common pixel depth is 8, which allows for a total of $2^8 = 256$ different shades of grey. This bit depth is often used for images to be viewed by a human observer since *just noticeable difference* (JND) experiments for light intensity differences, $\Delta I / I$, in human vision gives a *Weber's fraction* (Ernst Weber, *De tacu*: "Concerning Touch" in 1834) of roughly

$$\frac{\Delta I}{I} \approx \frac{1}{60} \tag{2.1}$$

across most light intensity ranges [BuGe90]. Given an intensity range of 100-160 millil-amberts for the average human eye at different brightness adaption levels [GoWo92], this gives a total of

$$\# \text{ graylevels} = \frac{\log(\text{intensity range})}{\log(1 + \text{Weber's fraction})} = \frac{\log([100, 160])}{\log(1 + 1/60)} = [278, 307] \tag{2.2}$$

discernible light intensities which is just over $2^8 = 256$ shades of intensity. Of course, this is an experimental maximum according to specific controlled experiments using few intensities at one time. In practice, the number of discernible shades of grey in a pixel image is far less. For most people, only 64–128 shades of grey are distinguishable with the limited light intensity range of the average computer monitor. This range is important in image compression research since this means that for images viewed by human observers, the primary target of images in this thesis, reductions are often possible in the encoding of 8-bit greyscale images without sacrificing image quality from a psychovisual perspective. Of course, if these images are not solely for viewing or perhaps will be used for scientific calculations, then more bits may be required per pixel. For example, many forms of medical images require 12 bits or more for acceptable pixel representation, and other forms of images, such as astronomical images, may require as great a pixel depth as possible with the capture device used.

Colour images fall outside of the current scope of this thesis so this is sufficient background on image representation. Even though colour image compression is not considered in this thesis, it should be understood that the majority of research done in greyscale image compression can be extended to colour images. This is typically done by applying the image compression algorithms to each component in the colour model. For instance, applying a greyscale image compression algorithm to each of the *Hue, Saturation, Intensity* (HSI) components in the HSI colour model since HSI forms a fairly good decomposition of colour light components from a psychovisual standpoint.

2.1.2   Image Compression/Reconstruction

Image compression refers to the reduction in the number of bits required to represent an image compared to the raw uncompressed representation as discussed in Sec. 2.1.1. In general, image compression can be thought of as the processes illustrated in Fig. 2.1.
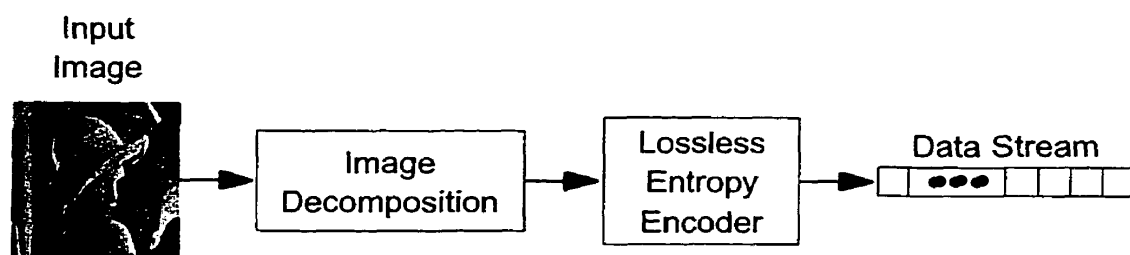
Fig. 2.1. General image compression data flow.

As Fig. 2.2 illustrates for image compression, an input image is taken and decomposed in some manner. This image decomposition is typically some reorder of the pixel information and/or some mathematical transformation of the pixel information into another domain that allows for a more efficient representation. This decomposition is generally performed with the goal of decorrelating local and global pixel information dependencies to improve the compression potential. In current image compression techniques, this decomposition can come in many forms, with the most basic form performing no decorrelation at all and other techniques doing difference or predictive coding. The decomposition can also take the form of transform based coding and subband coding using Fourier analysis, the discrete cosine transform [JPEG99], and the wavelet transform. Some decomposition techniques are based on neural network techniques such as vector

quantization [Gray84], [GeGr92], and progressive extensions [WaGo89], learned vector quantization [Koho90], self-organizing feature maps [Koho89], [Koho90], and some of our research group's implementations of these techniques [FeLK93], [DaKC95]. Other decomposition techniques use fractals, such as iterated function systems [Barn88], [Jacq90], [WoJa99]. Of course, there are many other methods and techniques that have been used with varying degrees of success for the image decomposition, but this presents some of the key techniques in use and being researched.

After the image decomposition step, the data representing this decomposition is then taken and lossless entropy encoding is done to remove redundancy in the symbol stream. If the decomposition of the image results in a completely decorrelated symbol set, then lossless 1st order entropy encoding can be performed with redundancy removed as dictated by the Shannon entropy [Shan48a], [Shan48b] minimum. If correlations within the symbol set still exist, then higher order statistics may be needed to remove high levels of redundancy in the data symbol stream. Lossless entropy encoding can include techniques such as Huffman coding [Huff52], dictionary methods such as LZ77 [ZiLe77], LZ78 [ZiLe78], and LZW [Welc84], and other techniques such as arithmetic coding [WiNC87]. The resulting data stream constitutes the final compressed version of the image.

Reconstructed
Image
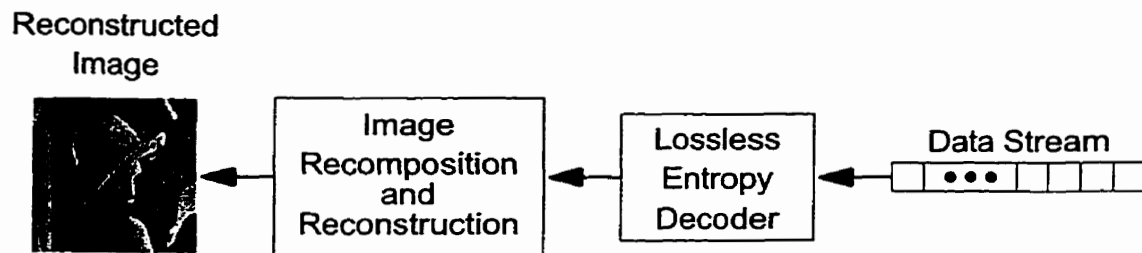
Fig. 2.2. General image reconstruction data flow.

The image reconstruction process is the reverse of the image compression process, as illustrated in Fig. 2.2. The compressed data stream is passed through the lossless entropy decoder which exactly rebuilds what the lossless entropy encoder removed in the compression phase. Reconstruction of the image through the reverse of the decomposition process then completes the image reconstruction. It is important to note that the inverse of the image decomposition process, depending on the method used, may not necessarily reconstruct an exact duplicate of the original image. This issue of lossy and lossless compression is discussed further in Sec. 2.1.3.

Looking at the entire image compression/reconstruction process, the primary question that must be answered is how to decompose the image to decorrelate the image information. This is the question that this thesis looks at with two additional objectives; how the most relevant psychovisual features can be extracted first from an image, and how these image features can be represented for the progressive transmissions and reconstruction of the image. The lossless entropy encoder is also an integral part of the image compression process. It completes the compression process by removing any additional data redundancy left by the image decomposition process. While work in this thesis could be

extended to investigate the effectiveness of different lossless entropy coders, only adaptive arithmetic coding [WiNC87] is used since it still remains one of the better lossless coding techniques in terms of compression rate.

2.1.3   Lossy Versus Lossless Compression

Data compression requires that the compressed representation of the data contains sufficient information to fully reconstruct the data. This compression is done by the removal of *data redundancy* from the data set with bounds given by information theory and Shannon entropy [Shan48a], [Shan48b]. This concept is known as *lossless* data compression.

While perfect data reconstruction is a requirement for pure data, source code, machine code, etc., often the final use of an image is for it to be viewed by a human observer. Keeping this use in mind, when image compression is performed, in general, it should be targeted for its intended final user. When the final user of an image is a human observer, an entirely new approach to image compression beyond data redundancy removal presents itself. This approach permits the removal of *psychovisual redundancy*. That is, it permits the removal of data from the image that does not add to the interpretation of the image by a human observer. In the case of images, the signal representation must be capable of reconstructing the psychovisual information contained in an image and not necessarily the exact bits from the original image.

This approach to image compression is known as *lossy* data compression since some of the original data may not be recoverable from the compressed representation.

Lossy data compression is acceptable as long as a human observer cannot distinguish between the original and reconstructed signals [JaJS93], in this case images. This removal of redundancy is possible for images since there are limits to a human's ability to distinguish between image differences (for instance, the limit in distinguishable greylevels as discussed in Sec. 2.1.1). Lossy data compression may also allow for a degradation in perceived reconstruction quality. This degradation is acceptable in image compression if the perceived difference is small and the image is still acceptable for its intended application. Therefore, the main idea behind a good lossy image compression scheme is the removal of *redundant information* from the perspective of an observer in addition to *redundant data* through information theoretic techniques.

2.1.4   Measuring Compression Rate

When looking at the performance of specific image compression techniques, the compression rate is generally used and is typically quoted in the number of bits per pixel required to represent the image. This measure can be expressed as follows

$$\text{bpp} = \frac{\text{\# of bits in compressed form}}{(\text{width}) \cdot (\text{height})} \qquad (2.3)$$

While bpp is a useful measure of compression performance, it is sometimes useful to factor out the pixel depth and quote the compression performance as a ratio of the original image size or as a percentage of the original size as follows

$$\text{compression \%} = \frac{\text{bpp}}{\text{bit depth of original image}} \times 100\% \qquad (2.4)$$

This then gives the compression ratio of the compressed image versus the raw uncompressed format of the original image as discussed in Sec. 2.1.1. It should be noted that

with Eq. 2.4 it is possible to have a compression scheme that for certain images results in the expansion of the bit stream as opposed to compression. This growth would result in a compression percentage greater than 100%.

2.1.5 Objective Fidelity Criteria of Images

In evaluating image reconstructions, perceptual comparisons are required, but a quantitative measure of the reconstruction is also useful. Quantitative measures can offer a method of evaluating images and their reconstructions. With these measures, quick comparisons can be made and quantitative limits established for the image reconstructions.

One means of evaluating the data lost resulting from a lossy compression is by determining the error between the original image and the reconstructed image. The error between the original image $f_{orig}(x, y)$ and the reconstructed image $f_r(x, y)$ can be expressed as follows

$$f_{error}(x, y) = f_{orig}(x, y) - f_r(x, y) \tag{2.5}$$

This gives the error within the reconstructed image for any specific value of x and y.

For a measure of the overall error between the images, the *mean squared error* (MSE) can be used. The MSE is expressed as follows

$$\text{MSE}(f_{orig}, f_r) = \frac{\displaystyle\sum_{\forall(x,y)} [f_r(x, y) - f_{orig}(x, y)]^2}{\displaystyle\sum_{\forall(x,y)} 1} \tag{2.6}$$

This metric is not used alone very often in image compression literature since researchers are generally more interested in what the distortion in the reconstructed image is compared to the original image. For this, the *signal-to-noise ratio* (SNR) can be used, which can be expressed as follows

$$\text{SNR} = \frac{\text{MSE}(0, f_r)}{\text{MSE}(f_{orig}, f_r)} = \frac{\sum_{\forall (x, y)} f_r(x, y)^2}{\sum_{\forall (x, y)} [f_{orig}(x, y) - f_r(x, y)]^2} \qquad (2.7)$$

or in *decibels* (dB)

$$\text{SNR (dB)} = 10\log_{10} \frac{\sum_{\forall (x, y)} f_r(x, y)^2}{\sum_{\forall (x, y)} [f_{orig}(x, y) - f_r(x, y)]^2} \qquad (2.8)$$

While the SNR is a somewhat more common measure, it is useful only for considering how the noise compares to the signal level of a specific signal, or for our purposes the level of distortion in a reconstructed image versus the average squared energy in the reconstructed image. Hence, comparison across different distorted or lossy compressed images is difficult since the respective original images likely have different average squared energies. To improve the comparison across different images, the *peak signal-to-noise ratio* (PSNR) is preferred and can be expressed as follows

$$\text{PSNR} = \frac{\text{MSE}(0, f_{peak})}{\text{MSE}(f_{orig}, f_r)} = \sum_{\forall (x, y)} \frac{f_{peak}(x, y)^2}{(f_r(x, y) - f_{orig}(x, y))^2} \qquad (2.9)$$

where $f_{peak}(x, y)$ is an image where each pixel has the maximum representation pixel value. The PSNR can similarly be expressed in decibels as

$$PSNR\ (dB)\ =\ 10\log_{10} \sum_{\forall(x, y)} \frac{f_{peak}(x, y)^2}{(f_r(x, y) - f_{orig}(x, y))^2} \qquad (2.10)$$

If the images being analyzed have a dynamic range of 8 bits, which is the case throughout this thesis, then Eq. 2.10 can be reduced to

$$PSNR_8\ (dB)\ =\ 10\log_{10} \sum_{\forall(x, y)} \frac{(2^8)^2}{(f_r(x, y) - f_{orig}(x, y))^2} \qquad (2.11)$$

While there is no concrete set of PSNR numbers that correlate to psychovisual quality, there are rough rules of thumb that researchers have noticed across a wide range of natural looking images compressed with a wide range of lossy image compression schemes. These rough rules of thumb can be roughly outlined as in Table 2.1.

Table 2.1 Rough rules of thumb of image quality for a corresponding PSNR.

PSNR value (dB)	Rough rule of thumb for image quality
< 25 dB	poor image quality, image likely unusable
25 dB	image is recognizable, but quality is perceptually inferior
28 dB	fair perceptual image quality, some noticeable artifacts but generally quite acceptable
30 dB	perceptually good image quality, few noticeable image artifacts
33 dB	broadcast quality images, very good with nearly no perceptual differences with the original image
> 35 dB	image virtually indistinguishable from the original

2.2   Progressive Image Transmission

This section gives an overview of progressive image transmission and outlines some of the current progressive image transmission technology and implementations available.

2.2.1   Image Transmission and Progressive Image Transmission Overview

This section describes some of the ideas behind image transmission, progressive image transmission, and progressive transmission as related to the goals of this thesis. Image transmission comes in many forms and flavours. The most basic form of image transmission is to transmit the entire image file from one machine to another and after all of the image file has been received, display the image. This form of image transmission is limiting since it is often beneficial to the user if they can get a sense of the image content while the image is being transferred instead of waiting until all of the data has been received. There are a number of reasons for having image features appear earlier:

- the user may wish to terminate transmission of the image if it is not of interest.

- image details may not be important to the user but basic image structure might be of interest.

- the user may wish to quickly browse through the images.

The last point can be solved through the use of thumbnails of the images (*i.e.* small versions of the image). Unfortunately, while thumbnails allow for quick flipping through

images they often do not contain sufficient detail for the user. A good progressive image

transmission scheme would allow the user to control that amount of detail they require.

| (a) 16,384 bytes | (b) 65,536 bytes | (c) 196,608 bytes |
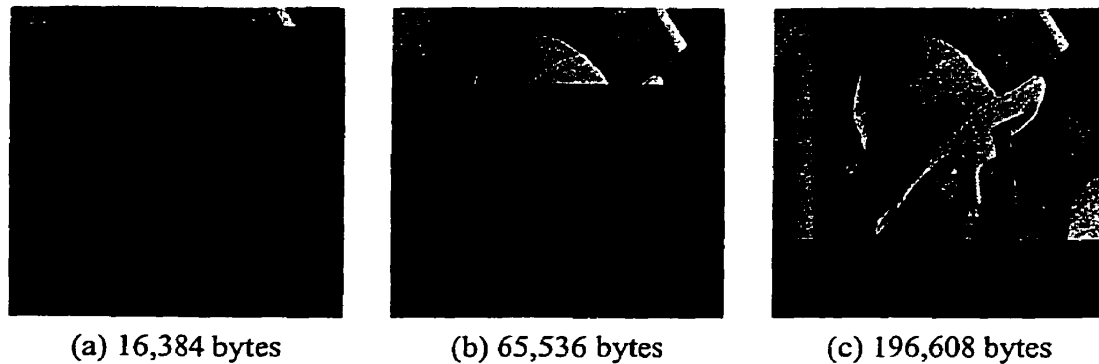
Fig. 2.3. Progressive transmission of horizontal scan lines of the image lena:
(a) after 16,384 bytes, (b) after 65,536 bytes, and (c) after 196,608 bytes.

One of the simplest forms of progressive image transmission, as illustrated in the

sample progressive image transmission in Fig. 2.3, works by sequentially transmitting

horizontal scan lines of the image starting from the top of the image. All details at each of

the horizontal scan lines are transmitted from left to right and the whole image is transmit-

ted in one pass. This is one of the simplest methods of transmitting (or storing) an image

but it has the disadvantage that the user must wait for the entire transfer to occur to get an

idea of what is at the bottom portion of the image. It should be noted that many people do

not consider this a progressive image transmission technique. In this thesis this will still

be classified as progressive image transmission since the user need not wait for the entire

image to be transmitted before they can view a portion of the image information. This

form of image transmission is typical of today's web browsers.

(a) 16,384 bytes　　　　　(b) 65,536 bytes　　　　　(c) 196,608 bytes
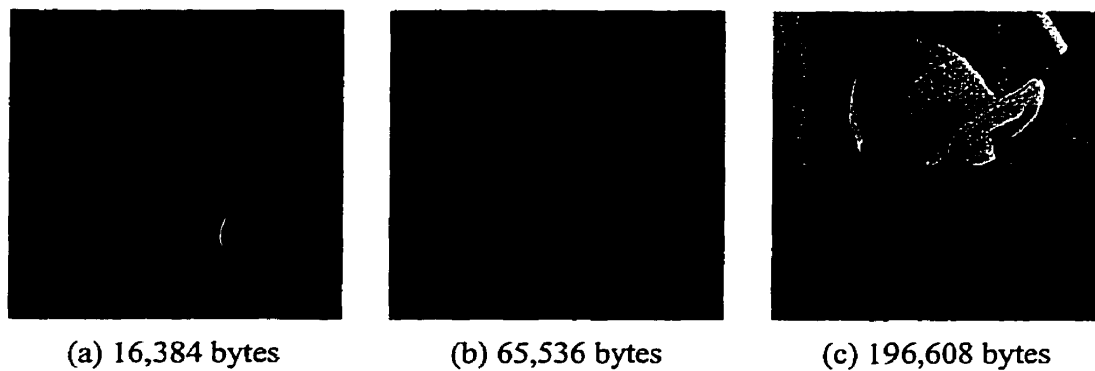
Fig. 2.4. Progressive transmission of every second scan line of the image lena:
(a) after 16,384 bytes, (b) after 65,536 bytes, and (c) after 196,608 bytes.

An extension to the above progressive image transmission technique is to transmit
every second horizontal scan line of the image in one pass and then to go back and trans-
mit the remaining scan lines. An example of this is shown in Fig. 2.4. This simple change
halves the time in which some of the details at the bottom of the image begin to appear.
Each horizontal scan line is still transmitted in full detail but a sense of the entire image is
obtained quicker. This is a better technique for progressive image transmission and begins
to show some of the many possibilities that can be used. This technique is still naive since
the structure of the image content itself is not considered in this type of approach. Better
techniques consider the content of the image and do not just blindly transmit the data as is
done with this approach.

(a) 16,384 bytes   (b) 65,536 bytes   (c) 196,608 bytes
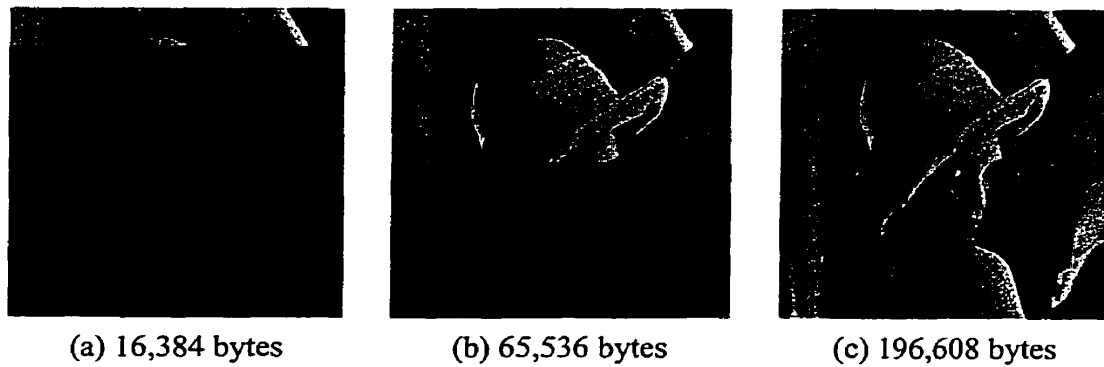
Fig. 2.5. Progressive transmission of every second scan line (filled) of the image lena: (a) after 16,384 bytes, (b) after 65,536 bytes, and (c) after 196,608 bytes.

It should be noticed that the missing lines in Fig. 2.4 can be annoying from a perceptual point of view. By using the same algorithm but duplicating the scan lines to the empty scan lines below, Fig. 2.5 is produced. While there is no change in the actual data transmitted, it is clear that this change corrects the perceptual problem of Fig. 2.4 having scan lines gaps every second scan line.

(a) 16,384 bytes   (b) 65,536 bytes   (c) 196,608 bytes
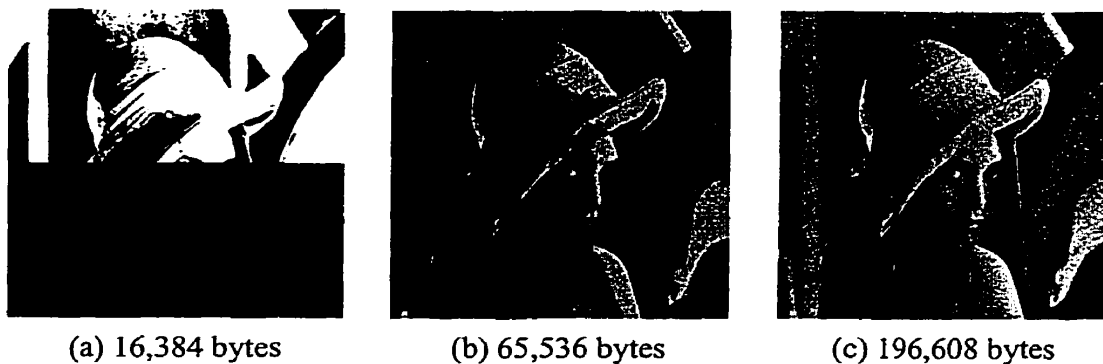
Fig. 2.6. Progressive transmission of bit planes starting from most significant for the image lena: (a) after 16,384 bytes, (b) after 65,536 bytes, and (c) after 196,608 bytes.

Another technique for progressive image transmission is to transmit the pixel bit layers one at a time, starting with the layer containing the *most significant bit* (MSB) of each pixel, and ending with the layer with the *least significant bit* (LSB). A sample progressive transmission using this technique is shown in Fig. 2.6. Conceptually this is an improvement over the previously discussed techniques since the image detail is ordered nicely according to the pixel intensity since much of the main energy components in an image can be extracted from the MSB first. Unfortunately, from a psychovisual viewpoint the results remain unfavourable at the early stages of the transmission since the portions of the image that are visible lack the details and greyscale intensity levels from the original image.

(a) 16,384 bytes   (b) 65,536 bytes   (c) 196,608 bytes
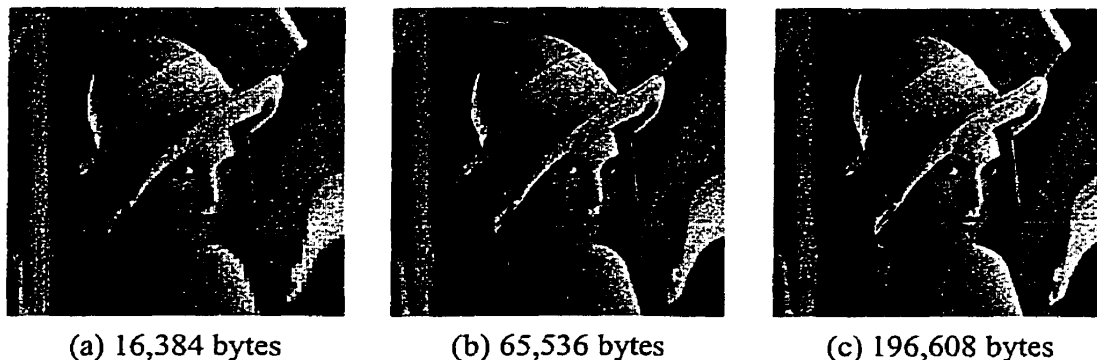
Fig. 2.7. Progressive transmission of quad-tree partitioned pixel averages for the image lena: (a) after 16,384 bytes, (b) after 65,536 bytes, and (c) after 196,608 bytes.

Another simple, yet less naive, technique is to do a quad-tree partitioning of the original image into four equal sized squares. The average pixel intensity of each of these squares is then calculated and these averages transmitted. The receiving machine then begins reconstruction of the image using the same square size and the transmitted pixel

intensity averages. Once this is done, each of the squares is further partitioned and the change in average pixel intensity from the previous average intensity for that square is transmitted. This process is repeated until the entire image has been transmitted. An example progressive image transmission using this technique is shown in Fig. 2.7. Conceptually, this technique is an improvement over the previous described techniques since some of the psychovisual content of the image (in terms of average greyscale intensity) is being considered. Unfortunately, this technique requires extra computations compared to the previous techniques in order to calculate the average greyscale values of groups of pixels. Techniques such as this would be what most people classify as a progressive image transmission technique.

Some other techniques require better knowledge or interpretation of the image content. For instance, one technique requires knowing what types of objects are in the image. This may include recognizing that there is a person in the image or perhaps a beachball. This information is transmitted along with their positions in the image. The next stage in the progressive image transmission is to find details about these objects and transmit these details for a further refined image reconstruction. For instance, the person might be wearing a green turtleneck and the beachball might have eight strips in alternating red and white colour. As can be seen from this technique this requires a very complete and possibly complicated representation of the image objects. Currently, this type of image representation for general images is unrealistic. This technique may become attractive one day when these types of computer models are more prevalent.

The general idea behind progressive image transmission, therefore, is the proper ordering of psychovisual image details so that important global details of interest are transmitted first, while small and minute details are transmitted last. This raises the question as to what is important in an image and how to decide what details are transmitted first. This thesis tries to answer some of these questions in later chapters during the investigation into wavelet and fractal techniques for progressive image transmission.

2.2.2   Current Progressive Image Transmission Implementations

A number of progressive image transmission schemes exist that were developed by other research groups and companies. The following is a list of the more common or unique progressive image transmission schemes available:

- interlaced GIF

- progressive JPEG

- progressive wavelets

- progressive fractal iterated function systems

Interlaced GIF (*graphics interchange format*), a format developed by CompuServe [GIF87a][GIF90], is currently the most well known and used progressive image transmission file format. This image compression technique is algorithmically simple, since it is based on linear predictive coding using a form of the *Lempel-Ziv-Welch* (LZW) [Welc84] algorithm. The GIF image format has the following features:

- lossless image compression only,

- works with bi-level images,

- works with 8-bit greyscale images,

- works with 8-bit colour images,

- does not work with 24-bit colour images (there is some push for
  GIF24 which also supports 24-bit colour images but little support
  for this currently exists).

Since GIF can do only lossless image compression, it does not take advantage of any psychovisual redundancy and therefore the compression rate is not as good as those methods that remove psychovisual redundancy. Also, the progressive nature of interlaced GIF does not use any of the image structure itself to form progressive steps. It only orders image pixels as a quadtree (similar to the partition into squares as discussed in the previous section) and then decodes the image from this quadtree representation.

Progressive JPEG is an extension of the JPEG still image compression standard [JPEG99], [PeMi92], [Wall91] which does lossy compression of bi-level, 8-bit greyscale and 24-bit colour images. The JPEG still image compression standard, described by the *Joint Photographics Expert Group*, is based on the *discrete cosine transform* (DCT) over 8×8 pixel blocks from the image. The progressive extension for JPEG reorders the DCT coefficients for transmission so that the most important coefficients from each 8×8 block in the image are transmitted first. This method attempts to remove psychovisual redundancy by giving preference to lower frequency image components, which contain rough details, over higher frequency components, which contain fine details. Also, the lossy nature of JPEG through the removal of smaller coefficients from the DCT means that higher compression ratios are achieved through small degradations of image quality. This method is not currently as widely used as interlaced GIF for progressive transmission but

is quickly gaining acceptance as a standard method for progressive image transmission. A newer version of JPEG known as JPEG2000 is currently under development that replaces the DCT with the wavelet transform of the image and does away with the idea of partitioning the image into 8×8 blocks of pixels.

Progressive wavelet image compression is the next logical step for image compression standards. Wavelet image compression retains many of the characteristics of DCT based image compression techniques (such as JPEG) but with improved compression in many situations. This improvement results from the spatial and frequency localization that wavelets enjoy whereas the DCT localizes only in the frequency domain, as will be expanded upon in Chapter 4. Currently, Infinop's Lightning Strike wavelet image compression [Infi99] which has been updated for progressive transmission is one of the better known progressive wavelet transmission implementations. Unfortunately, this is a commercial implementation with the standard not available publicly. A number of others have implemented progressive wavelet image transmission including Langi and Kinsner who have developed an experimental progressive wavelet image compressor that this thesis extends. Other implementations, such as DjVu from AT&T [DjVu99], use forms of a progressive wavelet transmission at their core, though DjVu is more a hybrid of a number of techniques. Other proposed standards include the upcoming JPEG2000 standard, ISO 15444 Part 1, which plans to incorporate progressive wavelet image transmission. Further discussion on wavelet and progressive wavelet techniques will be discussed in Chapter 4 and Chapter 7.

The fractal encoder and viewer from Iterated Systems, Inc. is another example of a progressive image transmission scheme [Iter99]. This method is a progressive image transmission scheme based on fractal *iterated functions systems* (IFS) [Barn88], [Jacq90]. This technique relies on the transformation and contractive mapping of sections of the image to other sections of the image. This transformation constitutes a self-similar mapping and consequently generates a fractal. The main problem with this technique is that no polynomial time algorithm is known that can find the contractive mapping for generating the image even though attempts at improving the complexity have been tried [Wall93]. While the decoding of an IFS compressed image can be done quickly, the actual encoding is an intractable problem, especially for large images. Some simple algorithms have been developed that can find contractive mappings of small images in a number of minutes, but these mappings are far from optimal and the compression rates are not as good as the IFS technique suggests is possible. This technique still requires more research in order to find the best contractive mappings for general images and therefore will not be a main focus for this thesis.

2.3   Summary

This chapter described some of the background image compression ideas needed for the rest of this thesis along with building a framework for progressive image transmission. General image representation was discussed as well as how monochrome and grey-scale images are represented. A high level image compression and reconstruction model was outlined with the absence of specific compression techniques, which will be discussed in the following chapters. The concept of progressive image transmission was touched

upon along with many of the reasons and concerns behind doing progressive image trans-

mission. The next chapters delve into specifics of the techniques developed and used for

this thesis, focusing on wavelet, wavelet packet, and fractal techniques.

CHAPTER III

BACKGROUND ON FRACTALS, MULTIFRACTALS, AND FRACTIONAL BROWNIAN MOTION

This chapter presents the necessary background behind fractals and multifractals for this thesis. Of primary interest is the fractal and multifractal measure of textures and image feature complexity. These measurements are important since they provide a means of looking at the complexity of features within an image over all resolutions and characterize these image features with fractal and multifractal dimensions. This background provides the setting for Chapter 5 where fractals are used to synthesize images using fractal surface segmentation and interpolation. The setting will also be set for Chapter 6 where multifractal image complexity measures are introduced based on some of the positive findings from the fractal surface segmentation and interpolation. Finally, Chapter 7 puts these multifractal image complexity measures to use in the selection of wavelet coefficients and wavelet packet bases to improve progressive image transmission.

3.1   Introduction to Fractals

An introduction to the concepts behind fractals is in order before presenting the fractal based progressive image transmission, as well as the multifractal based image complexity measures. In its simplest form, a fractal is a *self-similar* object, where its parts are similar to the whole and, correspondingly, the whole is similar to its parts. This characteristic means that a fractal has similar, if not identical, structure and complexity at all scales or magnifications [Mand82].

To illustrate the self-similar structure of fractals, two mathematically self-similar fractals are presented with the Koch curve [Koch04] in Fig. 3.1 and the Sierpinski gasket [Sier15] in Fig. 3.2. Each of these figures illustrate the development of their respective fractals beginning with an initiator and then repeatedly performing a copy/reduction process ad infinitum to produce the self-similar fractals.

Initiator $\qquad$ length $= 1 = \left(\dfrac{4}{3}\right)^0$

Generator
Step 1 $\qquad$ length $= 4 \cdot \dfrac{1}{3} = \left(\dfrac{4}{3}\right)^1$

Generator
Step 2 $\qquad$ length $= 16 \cdot \dfrac{1}{9} = \left(\dfrac{4}{3}\right)^2$

Generator
Step 3 $\qquad$ length $= 64 \cdot \dfrac{1}{27} = \left(\dfrac{4}{3}\right)^3$

Koch curve
Step ∞ $\qquad$ length $= \left(\dfrac{4}{3}\right)^\infty = \infty$
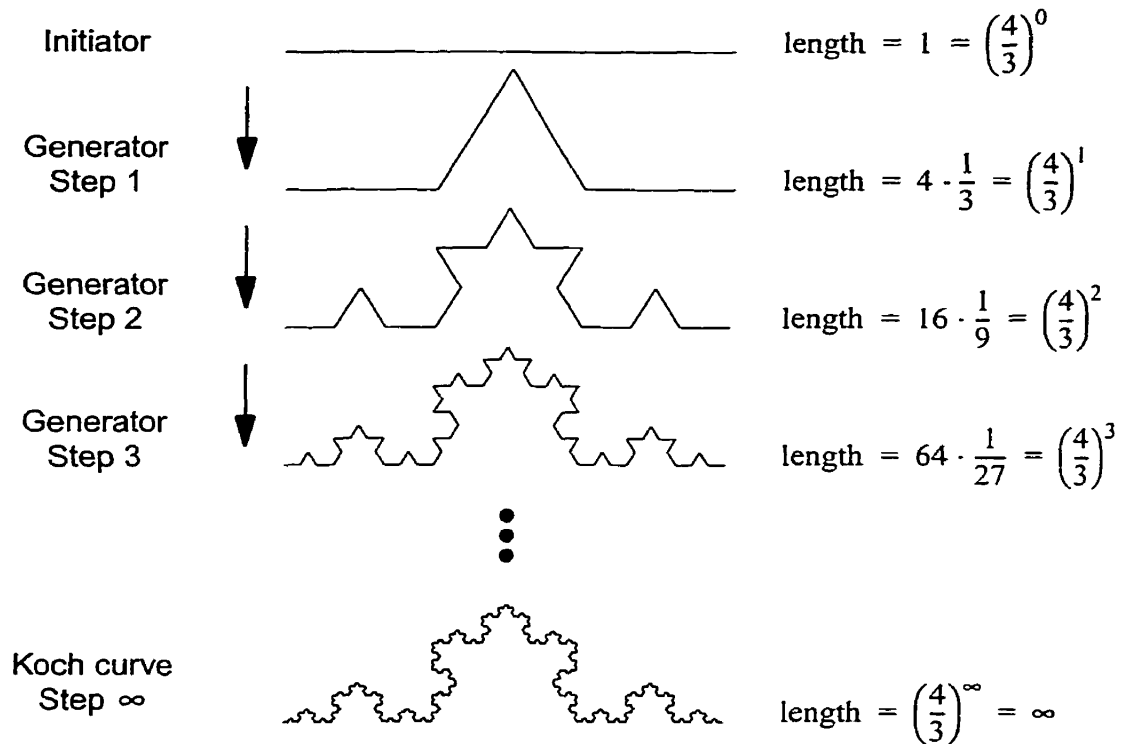
Fig. 3.1. Generation of the Koch curve fractal.

The generation of the Koch curve [Koch04] as illustrated in Fig. 3.1 starts with a line as an initiator which is then reduced to 1/3 of its size and copied 4 times in the pattern shown by the four lines in Step 1. This copy/reduction process is then repeated with the entire object in Step 1 where it is also reduced to 1/3 of its size and copied 4 times to pro-

duce the object in Step 2. This process is repeated an infinite number of times to produce

the final Koch curve fractal at the bottom of Fig. 3.1.
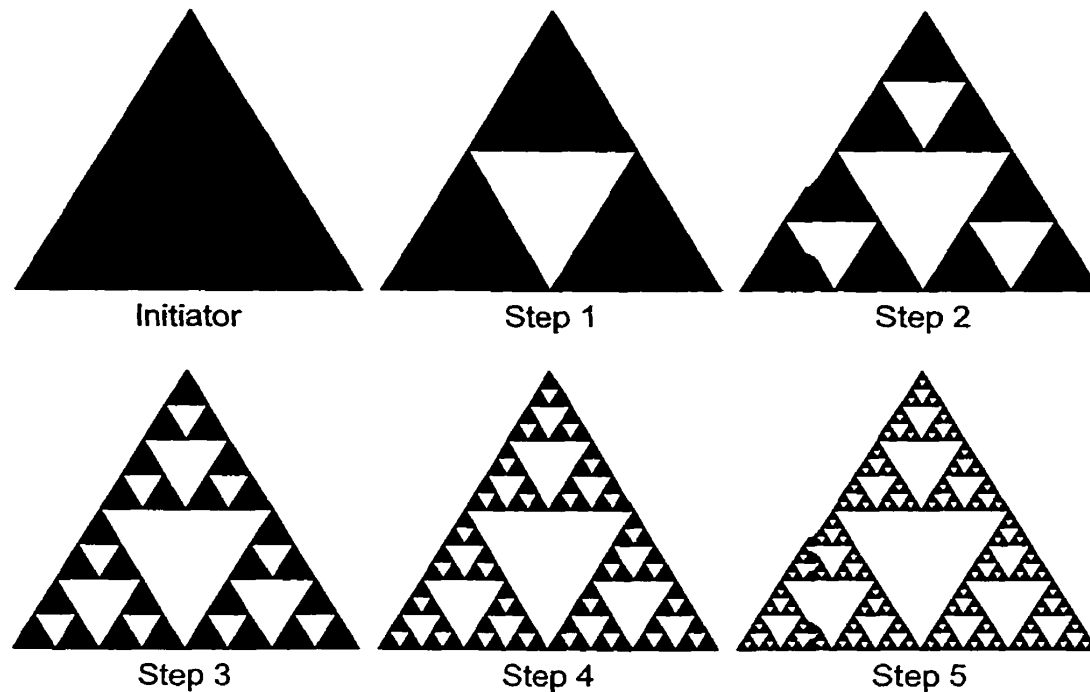
Fig. 3.2. Generation of the Sierpinski gasket fractal.

The generation of the Sierpinski gasket illustrated in Fig. 3.2 shows a similar type

of copy/reduction process, where a triangle as an initiator is reduced to $1/4$ of its size and

then copied 3 times in the pattern shown in Step 1. Step 2 of Fig. 3.2 is generated by tak-

ing the object in Step 1 (the three triangles), reducing its size by $1/4$, and copying the

resulting object 3 times to the positions as illustrated. This process is repeated an infinite

number of times, where the final Sierpinski gasket fractal is attained [Sier15]. Notice with

the Koch curve and the Sierpinski gasket, and any other fractal for that matter, that smaller

portions of the fractal object are replicas of the whole except for the scaling.

Initiator	length $= 1 = \left(\dfrac{4}{3}\right)^0$
Generator Step 1	length $= 4 \cdot \dfrac{1}{3} = \left(\dfrac{4}{3}\right)^1$
Generator Step 2	length $= 16 \cdot \dfrac{1}{9} = \left(\dfrac{4}{3}\right)^2$
Generator Step 3	length $= 64 \cdot \dfrac{1}{27} = \left(\dfrac{4}{3}\right)^3$
Koch curve Step ∞	length $= \left(\dfrac{4}{3}\right)^\infty = \infty$
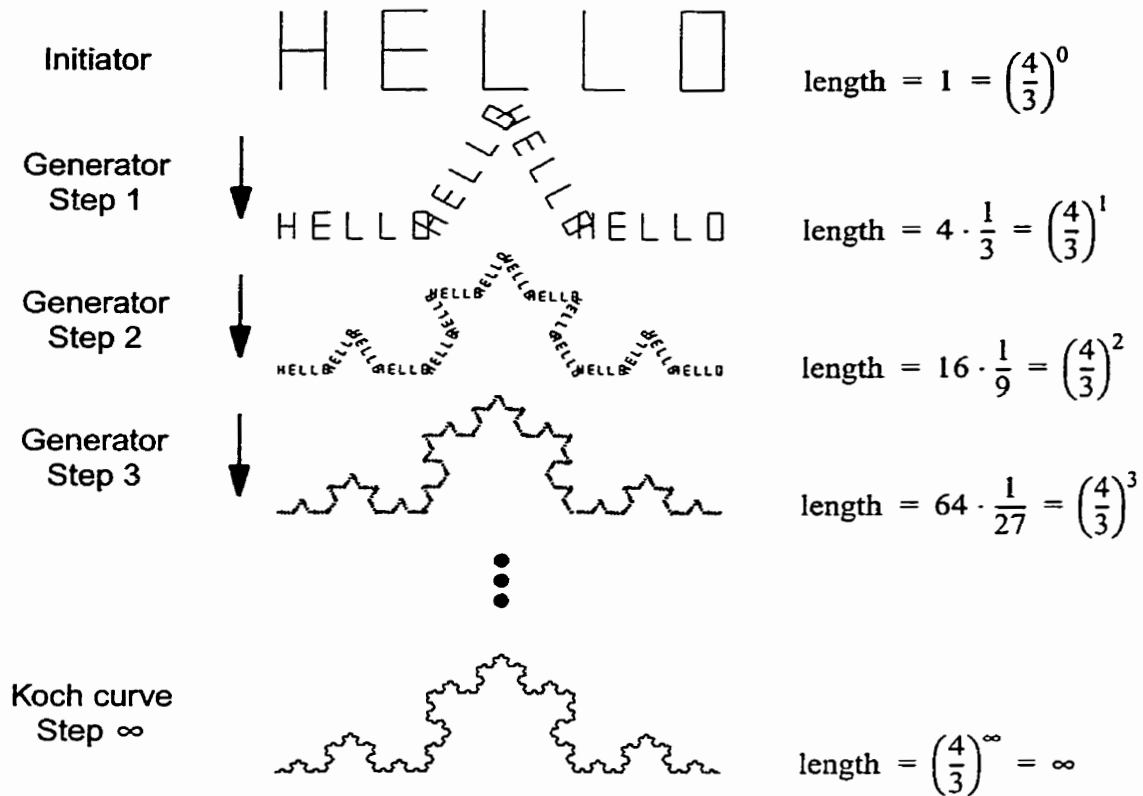
Fig. 3.3. Generation of the Koch curve fractal with "HELLO" initiator.

It should be noted that the generation of these two fractals is through the copy/ reduction process itself and does not rely on the initiator object at the beginning. This can be illustrated by generating the Koch curve fractal using the word "HELLO" as the initiator as shown in Fig. 3.3. With an infinite number of copy/reduction repetitions, the word "HELLO" disappears to a single point and only the Koch curve is left. This concept is important to realize since it shows that a fractal object is generated through a process, such as the copy/reduction process used for the Koch curve, as opposed to parameters of the initial object. This idea has ramifications that will be seen when performing fractal surface segmentation and interpolation later in this chapter.

3.2   Fractal Dimensions

This section presents background on fractal dimensions following the description given by Kinsner [Kins94a], [Kins95a] to unify and classify fractal dimensions. Fractals are named as such because they are classified as having a morphological dimension, or structural dimension, that is generally non-integral, that is, they have a *fractional dimension*. In simplest terms, this means that the complexity of fractal objects is such that the usual description of one, two or three dimensions is not accurate enough to describe these objects. The two objects presented in Sec. 3.1 are example fractals with fractal dimensions, as will be shown later in this section, of approximately 1.2619 and 1.5850 for the Koch curve and the Sierpinski gasket, respectively. Note that there are some special case objects where the dimension is actually integral, such as with space filling curves like the Hilbert curve [Hill91]. The Hilbert curve has a fractal dimension of 2 but it must be remembered that this curve is constructed using one dimensional line segments. The common feature behind all fractals therefore is that the complexity of the whole is greater than the initiating parts that form the object.

To see how this morphological complexity manifests itself, consider again the Koch curve in Fig. 3.1. The length of the generated curve at successive steps in units of the original initiator is $(4/3)^1$, $(4/3)^2$, $(4/3)^3$, and so on. In fact, the length of the curve can be generalized to $(4/3)^n$ at step n. Hence, if the generation of the curve continues with infinite steps producing the Koch curve, then the length of the Koch curve must be $(4/3)^\infty = \infty$. At first glance this result with the Koch curve having infinite length may be a bit disturbing since the Koch curve clearly has a starting point, a finishing

point, and all of the curve fits in a finite space. Another important characteristic of the

Koch curve is that each of the line segments within the curve becomes infinitely short, or

in other words each line segment shrinks to a point. With the way the Koch curve is con-

structed, it is therefore composed entirely of corners and contains no line segments. From

these realizations it is clear that the Koch curve is not of the same class of objects such as

lines, circles, parabolas, and other simple curves.

With some of the traditional rules of topology seemingly broken by the Koch

curve, as well as other fractal objects, some new way of atoning for these observations was

sought [Mand82]. One way is to re-evaluate how objects are measured.
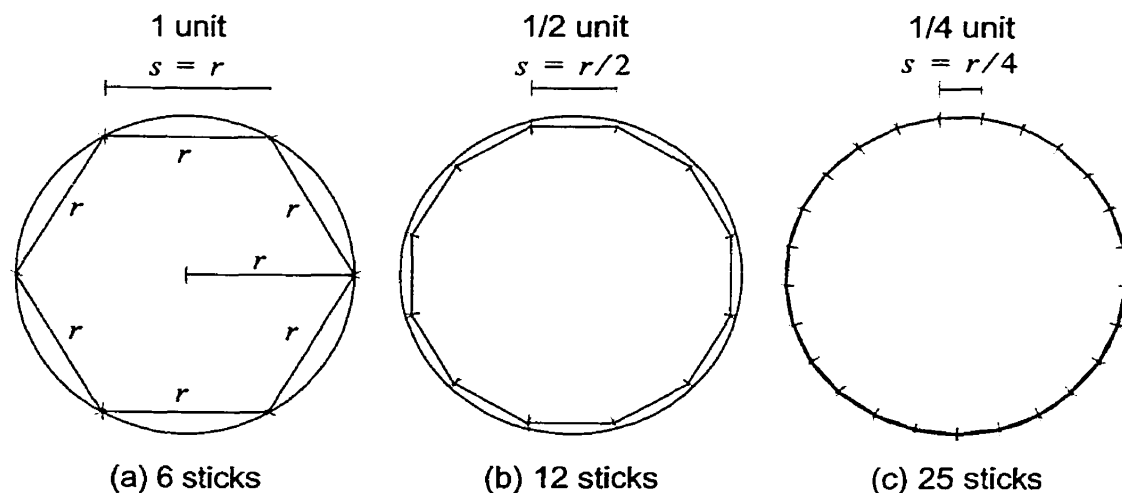
Fig. 3.4. Measuring the circumference of a circle with a measuring stick.
(a) $s = 1$ unit with 6 sticks,
(b) $s = 1/2$ unit with 12 sticks, and
(c) $s = 1/4$ unit with 25 sticks.

For instance, one method of measuring the circumference of a circle is to take a

measuring stick and go around the circle counting how many measuring sticks are

required, as illustrated in Fig. 3.4. As Fig. 3.4a illustrates, the radius, r, of the circle can be taken and then used to approximate the circumference of the circle as $6r$. It is known, of course, that the circumference is actually $2\pi r$, so $6r$, while accurate, is not overly precise. Better precision can be achieved by reducing the size of the measuring stick. Assuming the use of an *integral* number of measuring sticks for measuring the circumference of the circle, Table 3.1 can be produced.

**Table 3.1 Measuring sticks needed for measuring
the circumference of circle with radius r.**

Scale of measuring stick, s (in terms of radius r)	Integral number of measuring sticks, N_s, to approximate a circle's circumference	Total measured length
r	6	$6r$
$r/2$	12	$6r$
$r/4$	25	$6.25r$
$r/8$	50	$6.25r$
$r/16$	100	$6.25r$
$r/32$	201	$6.28125r$
$r/64$	402	$6.28125r$
$r/128$	804	$6.28125r$
$r/256$	1608	$6.28125r$
$r/512$	3216	$6.28125r$
$r/1024$	6433	$6.2822265625r$

Table 3.1 shows that as the size of the measuring stick is reduced, that the measured circumference of the circle approaches $2\pi \approx 6.2831853$. More importantly from a

morphological viewpoint, it should be noticed that the number of measuring sticks required versus the scale of the measuring sticks has the following relationship

$$N_s \propto \left(\frac{1}{s}\right)^1 \tag{3.1}$$

or in other terms

$$s^1 N_s \propto 1 \propto 2\pi \qquad \text{as } s \to 0 \tag{3.2}$$

This relation means that the measured length remains roughly constant as the size of the measuring stick is reduced since the number of measuring sticks required increases correspondingly.

1 unit   1/2 unit   1/4 unit
$s = r$ $s = r/2$ $s = r/4$

(a) 6 vels   (b) 12 vels   (c) 25 vels
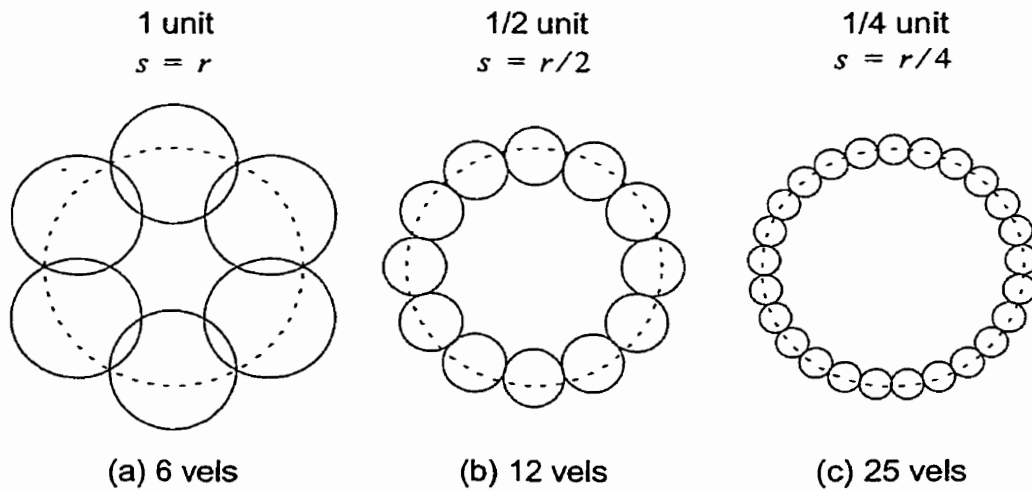
Fig. 3.5. Measuring the circumference of a circle with vels.
(a) $s = 1$ unit with 6 vels,
(b) $s = 1/2$ unit with 12 vels, and
(c) $s = 1/4$ unit with 25 vels.

To illustrate this measuring process further, consider extending the measuring stick to a measuring hypersphere or, more generally, a *volume element* (vel) [Kins94a] that

extends in all dimensional spaces. This same measuring process can be repeated for the circumference of the circle as illustrated in Fig. 3.5 by covering the circle with these vels over different measuring scales. The results obtained from these measurements are the same as those listed in Table 3.1. The advantage of using vels instead of a simple measuring stick is that now objects within any embedding dimension can be measured in this manner.

$$s = r \qquad\qquad s = r/2 \qquad\qquad s = r/4$$

(a) 9 vels   (b) 36 vels   (c) 144 vels
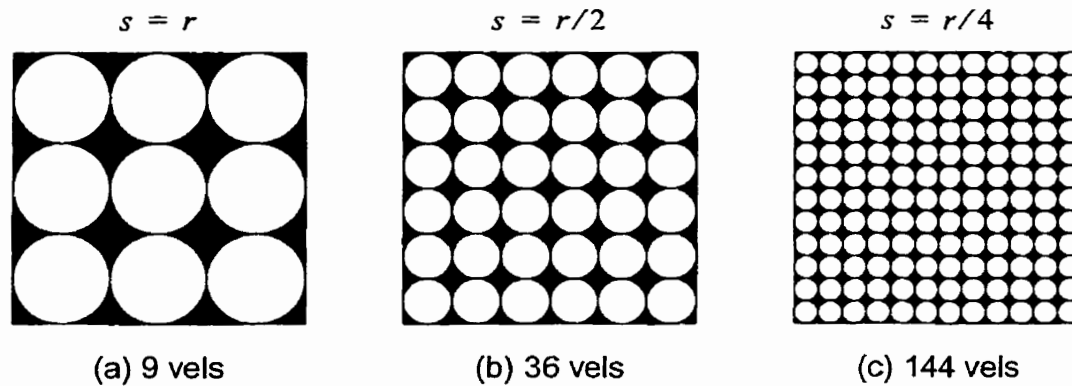
Fig. 3.6. Measuring the area of a square plane with non-overlapping vels.
(a) $s = 1$ unit with 9 vels,
(b) $s = 1/2$ unit with 36 vels, and
(c) $s = 1/4$ unit with 144 vels.

Take for example the square plane illustrated in Fig. 3.6. If vels are taken and used to cover and measure the area of a square plane as is illustrated in Fig. 3.6a, it is seen that roughly 9 non-overlapping vels are required. As the scale of the vels decreases, the number of vels required to cover the square plane increases as listed in Table 3.2.

Table 3.2 Vels needed to cover the square plane.

Scale of vel, s	Number of needed vels, N_s	Total measured area of the square plane
r	$3 \times 3 = 9$	$9r^2$
$r/2$	$6 \times 6 = 36$	$36(r/2)^2 = 9r^2$
$r/4$	$12 \times 12 = 144$	$144(r/4)^2 = 9r^2$
$r/8$	$24 \times 24 = 576$	$576(r/8)^2 = 9r^2$
$r/16$	$48 \times 48 = 2304$	$2304(r/16)^2 = 9r^2$
$r/32$	$96 \times 96 = 9216$	$9216(r/32)^2 = 9r^2$
$r/64$	$192 \times 192 = 36864$	$36864(r/64)^2 = 9r^2$
$r/128$	$384 \times 384 = 147456$	$147456(r/128)^2 = 9r^2$
$r/256$	$768 \times 768 = 589824$	$589824(r/256)^2 = 9r^2$
$r/512$	$1536 \times 1536 = 2359296$	$2359296(r/512)^2 = 9r^2$
$r/1024$	$3072 \times 3072 = 9437184$	$9437184(r/1024)^2 = 9r^2$

In analyzing the rate of growth of the number of needed vels versus the scale of the vels, it is noticed that the following relationship is followed

$$N_s \propto \left(\frac{1}{s}\right)^2 \tag{3.3}$$

or similarly

$$s^2 N_s \propto 1 \propto 9r^2 \qquad \text{as } s \to 0 \tag{3.4}$$

Therefore the rate of growth of the number of vels required to cover the square plane grows in correspondence to the square of how the vel size shrinks. To show that the measuring does not have to be done so carefully, consider Fig. 3.7 that uses overlapping vels to measure the planar square instead of non-overlapping vels as in Fig. 3.6. It can be seen

that the rate of growth of the number of vels N_s still follows the power-law, or exponential, relationship from Eq. 3.3.
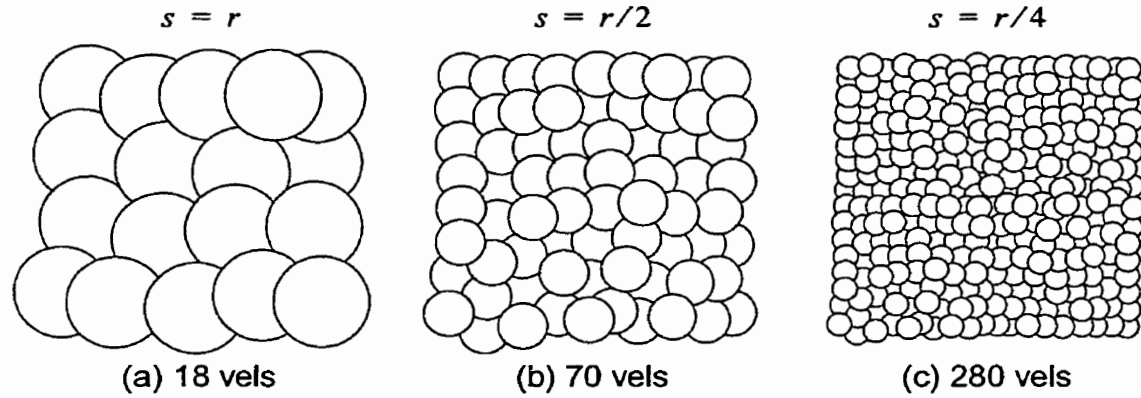
Fig. 3.7. Measuring the area of a square plane with overlapping vels.
(a) $s = 1$ unit with 18 vels,
(b) $s = 1/2$ unit with 70 vels, and
(c) $s = 1/4$ unit with 280 vels.

The important observation to make from Eq. 3.1 and Eq. 3.3 is that in both cases the exponent used in the power-law relationship between N_s and $1/s$ is precisely what would be referred to as the dimension of the object being measured, or the morphological dimension. As demonstrated, the boundary of the circle in Fig. 3.5 has a morphological dimension of 1 and the planar area of the square in Fig. 3.6 has a morphological dimension of 2.

Using this method of measuring objects through vel coverings, the morphological dimension of an object can be generalized through the following power-law relationship

$$N_{s_i} \propto \left(\frac{1}{s_i}\right)^D \qquad\qquad (3.5)$$

where N_{s_i} is the number of vels used at the scale s_i. The exponent D in Eq. 3.5 is known

as the critical exponent that stabilizes the change in N_{s_i} with the change of $(1/s_i)$.

Removing the proportionality by introducing the constant k, the following equation can

be expressed for Eq. 3.5.

$$N_{s_i} = k\left(\frac{1}{s_i}\right)^D \tag{3.6}$$

Taking Eq. 3.6 at two different scales s_i and s_j, and then dividing the two equations gives

the following equation

$$\frac{N_{s_i}}{N_{s_j}} = \frac{k\left(\frac{1}{s_i}\right)^D}{k\left(\frac{1}{s_j}\right)^D} \tag{3.7}$$

which can be simplified as follows

$$\frac{N_{s_i}}{N_{s_j}} = \left(\frac{s_j}{s_i}\right)^D . \tag{3.8}$$

Taking the log of both sides of Eq. 3.8 and solving for D gives

$$D = \frac{\log\dfrac{N_{s_i}}{N_{s_j}}}{\log\left(\dfrac{s_j}{s_i}\right)} . \tag{3.9}$$

From Eq. 3.9, as the scale becomes infinitely small, the ratios N_{s_i}/N_{s_j} and s_j/s_i will

reduce to a constant such that the following relation can be stated.

$$D = \lim_{s \to \infty} \frac{\log N_s}{\log\dfrac{1}{s}} \tag{3.10}$$

The exponent D is referred to as the fractal dimension associated with the object. This dimension is also commonly referred to as the *Hausdorff-Besicovitch dimension*, D_{HB}, [Mand85], [Kins94b] which is defined as follows.

Definition 3.1: Given N_s = inf(number of vels of scale s needed to cover object), the *Hausdorff-Besicovitch dimension*, D_{HB}, is defined as

$$D_{HB} = \lim_{s \to \infty} \frac{\log N_s}{\log \frac{1}{s}}. \tag{3.11}$$

❏

Using this concept of the Hausdorff-Besicovitch dimension, consider the same vel measurement scheme applied to the Koch curve from Fig. 3.1. Using the initiator in Fig. 3.1 as length r, Table 3.3 can be developed over a number of scales s.

Table 3.3 Vels needed to cover the Koch curve fractal.

Scale of vel, s	Number of needed vels, N_s	Total measured length of Koch curve
$r = r/3^0$	$1 = 4^0$	r
$r/3 = r/3^1$	$4 = 4^1$	$4r/3$
$r/9 = r/3^2$	$16 = 4^2$	$4^2 r/3^2$
$r/27 = r/3^3$	$64 = 4^3$	$4^3 r/3^3$
$r/81 = r/3^4$	$256 = 4^4$	$4^4 r/3^4$
$r/243 = r/3^5$	$1024 = 4^5$	$4^5 r/3^5$

In the case of the Koch curve, from Table 3.3, the following equation can be written for Eq. 3.9 when $s = r/3$ and $s = r/9$.

$$D = \frac{\log\left(\frac{16}{4}\right)}{\log\left(\frac{r/3}{r/9}\right)} = \frac{\log(4)}{\log(3)} \approx 1.2618595 \tag{3.12}$$

Notice that the r cancels out in Eq. 3.12 and that this is equivalent to using the Hausdorff-Besicovitch equation which can be solved as follows for $s = r/3$.

$$D_{HB} = \lim_{s \to \infty} \frac{\log N_s}{\log \frac{1}{s}} = \frac{\log(4)}{\log(3)} \approx 1.2618595 \tag{3.13}$$

It follows then that the morphological dimension of the Koch curve is approximately 1.26. This shows that the notion of integer dimensions needs to be expanded to fractional dimensions to characterize objects such as the Koch curve fractal.

To complete this section, the Sierpinski gasket illustrated in Fig. 3.2 can also be measured with the vel covering scheme with the results given in Table 3.4.

Table 3.4 Vels needed to cover the Sierpinski gasket fractal.

Scale of vel, s	Number of needed vels, N_s	Total measured length of Sierpinski curve
$r = r/2^0$	$1 = 3^0$	r
$r/2 = r/2^1$	$3 = 3^1$	$3r/2$
$r/4 = r/2^2$	$9 = 3^2$	$3^2 r/2^2$
$r/16 = r/2^3$	$27 = 3^3$	$3^3 r/2^3$
$r/32 = r/2^4$	$81 = 3^4$	$3^4 r/2^4$
$r/64 = r/2^5$	$243 = 3^5$	$3^5 r/2^5$

From Table 3.4, Eq. 3.9 can be written for $s = r/2$ and $s = r/4$ as follows.

$$D = \frac{\log\left(\frac{9}{3}\right)}{\log\left(\frac{r/2}{r/4}\right)} = \frac{\log(3)}{\log(2)} \approx 1.5849625 \tag{3.14}$$

Equivalently, the Hausdorff-Besicovitch dimension is expressed as follows for the Sierpinski gasket.

$$D_{HB} = \lim_{s \to 0} \frac{\log N_s}{\log\frac{1}{s}} = \frac{\log(3)}{\log\left(\frac{1}{1/2}\right)} \approx 1.5849625 \tag{3.15}$$

In summary, this section has shown a method for measuring objects, including fractals, using vels. This measurement allows for the Hausdorff-Besicovitch dimension of the object to be determined through a power-law relationship between the size of the vel and the number of vels needed to cover the object. Most importantly for the rest of this thesis is that the dimension measured is not necessarily integral and that this dimension value can be used to describe the level of complexity within an object. Simple curves result in a morphological dimension of 1, simple planes result in a morphological dimension of 2, and more complex objects may have morphological dimensions that are not integral. Note that besides the Hausdorff-Besicovitch dimension, there are other fractal dimensions and other similar approaches to measuring the fractal dimension of an object.

3.3   Multifractal Dimensions

This section gives some of the needed initial background on multifractals and multifractal dimension measures, following the descriptions of Kinsner [Kins94a], [Kins95a],

that are used later to develop image quality measures in the form of multifractal complexity measures. Multifractal (or inhomogeneous fractal) dimension measures are an extension of the fractal dimension measure described in Sec. 3.2 where these measures will be useful if there is more than one fractal dimension complexity within the measured object. As pointed out by Stanley and Meakin [StMe88], the majority of non-equilibrium, inhomogeneous phenomena in physics and chemistry exhibit complexities that single fractal dimension measures cannot characterize. Multifractal dimension measures are better suited for characterizing these complexities.

The following subsections lay down the preliminaries and groundwork for the Rényi generalized entropy, which is the basis for the Rényi dimension spectrum multifractal measure described in the following subsection. The Mandelbrot spectrum, which is an alternative representation to the Rényi dimension spectrum of the multifractal nature of an object, is then described.

3.3.1   Preliminaries

Some preliminary definitions and theorems are required for developing the multifractal dimensions for the discrete case in this section. Some introductory concepts in probability and information theory are presented next.

Let X be a discrete random variable with finite alphabet χ and probability mass function $p(x) \equiv \Pr\{X = x\}, \forall x \in \chi$

Definition 3.2: Let X be a discrete random variable with probability distribution $p(x)$. The *mean* or *expected value* of X is

$$\mu = \varepsilon(X) = \sum_{x \in \chi} xp(x) \qquad (3.16)$$

where ε denotes expectation.   ❑

Theorem 3.1: Let X be a discrete random variable with probability distribution $p(x)$. The mean or expected value of the random variable $g(X)$ is

$$\mu_{g(X)} = \varepsilon[g(X)] = \sum_{x \in \chi} g(x)p(x) \qquad (3.17)$$

❑

Theorem 3.2: If a function f has a second derivative which is non-negative (positive) everywhere, then the function is convex (strictly convex).   ❑

Theorem 3.3: *(Jensen's inequality)* If f is a convex function and X is a random variable, then

$$\varepsilon[f(X)] \geq f(\varepsilon[X]) \qquad (3.18)$$

where ε denotes expectation. Moreover, if f is strictly convex, then equality in Eq. 3.18 implies that $X = \varepsilon X$ with probability 1, i.e., X is a constant.   ❑

With these preliminary concepts in probability and information theory established, the next subsection introduces Shannon entropy, which is later generalized to the Rényi generalized entropy.

3.3.2 Shannon Entropy

The field of information theory owes most of its basis to Shannon's formulation of entropy [Shan48a], [Shan48b]. In its simplest terms, Shannon entropy is a measure of the uncertainty of a random variable X and is defined as follows.

Definition 3.3: The *entropy* or *Shannon entropy* $H(X)$ of a discrete random variable X is defined by

$$H(X) = -\sum_{x \in \chi} p(x)\log p(x) \tag{3.19}$$

where the convention of $0\log 0 = 0$ is used. ❑

In terms of expectation, the Shannon entropy can be rewritten as

$$H(X) = \varepsilon[-\log p(X)] = \varepsilon\left[\log\frac{1}{p(X)}\right]. \tag{3.20}$$

The general idea behind Shannon entropy is that the more regular or expected an event, the less information that can be derived from the event, and, conversely, the more irregular or random an event, the more information that can be derived from the event. This property is particularly useful in compression research since Shannon entropy provides a lower bound for first-order lossless compression of a data stream since the minimum number of events needed to represent a particular data stream probability distribution can be determined. Referring to Eq. 3.20, it is seen that Shannon entropy can also be referred to as the mean or average value of $\log\frac{1}{p(X)}$ [CoTh91]. Thus, Shannon entropy is effectively the average value of the information in the probability distribution

$p(x)$. This fact is useful when thinking about the Rényi generalized entropy which is described next. An important theorem to consider about Shannon entropy is as follows.

Theorem 3.4: $H(X) \leq \log|\chi|$ where $|\chi|$ is the cardinality of the set from which the random variable X is chosen. Equality is obtained if and only if X has a uniform distribution.☐

This theorem effectively states that $H(X)$ is concave with a maximum uncertainty only when all choices of the next symbol are uniformly probable.

3.3.3   Rényi Generalized Entropy

In 1955, Alfréd Rényi introduced a generalized form of the probability distributions where incomplete probability distributions are allowed [Rényi55]. This new formulation effectively removes the restriction that the summation $\sum\limits_{x \in \chi} p(x)$ must equal 1, though it still will for a normal probability distribution. Rényi's generalized probability distributions gave the basis for the following weight function

$$W(X) = \sum_{x \in \chi} p(x) \tag{3.21}$$

where $0 < W(X) \leq 1$. Clearly, when $W(X) = 1$ then the probability distribution of X is complete and is an ordinary probability distribution. When $0 < W(X) < 1$ then the probability distribution of X is incomplete.

In 1959, Rényi extended this generalized probability distribution to entropies and dimensions [Rény59]. With this extension and a paper in 1960, Rényi formalized a generalized form of Shannon entropy [Rény60] as follows

$$H(X) = \frac{-\sum_{x \in \chi} p(x) \log p(x)}{\sum_{x \in \chi} p(x)} = \frac{-\sum_{x \in \chi} p(x) \log p(x)}{W(X)} \tag{3.22}$$

which reduces to Eq. 3.19 when an ordinary distribution in $p(x)$ is used. Using this generalization it is apparent that Shannon entropy is a mean-value, where the mean-value is recognized only when the probability distribution is incomplete. Using this property of mean valued entropies, Rényi then searched for other mean valued quantities to generalize entropy instead of using the arithmetic mean using a linear function as in Eq. 3.22. Rényi showed that the only other admissible choice for mean value is with an exponential function. This form is defined as follows.

Definition 3.4: The *Rényi generalized entropy $H_q(X)$* (or Rényi entropy for short) of order q of a discrete random variable X is defined as

$$H_q(X) = \frac{1}{1-q} \log \frac{\sum_{x \in \chi} p^q(x)}{\sum_{x \in \chi} p(x)} \tag{3.23}$$

where $0 < q \le \infty$, $q \ne 1$, and $p(x)$ can be a complete or incomplete probability distribution. □

Certain liberties have been taken with the original definition of Rényi generalized entropy to extend the range of the order q to $-\infty \le q \le \infty$ and $q \ne 1$. For the purposes of

the multifractal measures developed, there is no loss of generality through the expansion

of this range.

To see how the Rényi generalized entropy is a generalization of Shannon entropy

with the inclusion of a moment of order q, consider when $q \rightarrow 1$.

Lemma 3.1: Let the moment order $q \rightarrow 1$. Then the Rényi generalized entropy reduces

to Shannon entropy in the limit as follows.

$$\lim_{q \rightarrow 1} H_q(X) = H(X) \tag{3.24}$$

Proof: Consider the Rényi generalized entropy as $q \rightarrow 1$.

$$\lim_{q \rightarrow 1} H_q(X) = \lim_{q \rightarrow 1} \frac{1}{1-q} \log \frac{\sum_{x \in \chi} p^q(x)}{\sum_{x \in \chi} p(x)} \tag{3.25}$$

Since

$$\lim_{q \rightarrow 1} 1 - q = 0$$

and

$$\lim_{q \rightarrow 1} \log \frac{\sum_{x \in \chi} p^q(x)}{\sum_{x \in \chi} p(x)} = \log \frac{\sum_{x \in \chi} p(x)}{\sum_{x \in \chi} p(x)} = \log 1 = 0$$

then l'Hôpital's rule can be used as follows

$$
\lim_{q \to 1} H_q(X) = \lim_{q \to 1} \frac{\dfrac{d}{dq}\left[\log \sum_{x \in \chi} p^q(x) - \log \sum_{x \in \chi} p(x)\right]}{\dfrac{d}{dq}(1-q)}
$$

$$
= -\lim_{q \to 1} \frac{\sum_{x \in \chi} \dfrac{d}{dq} p^q(x)}{\sum_{x \in \chi} p^q(x)}
$$

$$
= -\lim_{q \to 1} \left(\frac{\sum_{x \in \chi} p^q(x)\log(p(x))}{\sum_{x \in \chi} p^q(x)} \right)
$$

$$
= \frac{-\sum_{x \in \chi} p(x)\log p(x)}{\sum_{x \in \chi} p(x)}
$$

$$
= H(X) \tag{3.26}
$$

which is Shannon entropy.   ❏

This is an important result since this relates the Rényi generalized entropy back to the well studied Shannon entropy. Other interesting properties of the Rényi generalized entropy are at the order limits when $q \to \infty$ and $q \to -\infty$. The Rényi generalized entropy at these limits can be stated by the following two lemmata.

Lemma 3.2: Let the moment order $q \rightarrow \infty$. Then the Rényi generalized entropy $H_q(X)$

for an ordinary discrete random variable X becomes

$$H_\infty(X) \equiv \lim_{q \rightarrow \infty} H_q(X) = -\log\left\{ \frac{\sup(p(x))}{\sum\limits_{x \in \chi} p(x)} \right\}. \tag{3.27}$$

Proof: As $q \rightarrow \infty$, with $0 \leq p(x) < 1$ the sum $\sum\limits_{x \in \chi} p^q(x)$ is dominated by $\sup(p^q(x))$.

Therefore

$$H_\infty(X) = \lim_{q \rightarrow \infty} \frac{1}{1-q} \log\left\{ \frac{\sup(p^q(x))}{\sum\limits_{x \in \chi} p(x)} \right\}$$

Since $p(x)$ is positive, the power q can be moved as follows

$$H_\infty(X) = \lim_{q \rightarrow \infty} \frac{1}{1-q} \log\left\{ \frac{[\sup(p(x))]^q}{\sum\limits_{x \in \chi} p(x)} \right\}$$

$$= \lim_{q \rightarrow \infty} \frac{q}{1-q} \log\left\{ \frac{\sup(p(x))}{\sum\limits_{x \in \chi} p(x)} \right\} \tag{3.28}$$

And taking the limit gives

$$H_\infty(X) = -\log\left\{ \frac{\sup(p(x))}{\sum\limits_{x \in \chi} p(x)} \right\} \tag{3.29}$$

□

Lemma 3.3: Let the moment order $q \to -\infty$. Then the Rényi generalized entropy $H_q(X)$ for an ordinary discrete random variable X becomes

$$H_{-\infty}(X) \equiv \lim_{q \to -\infty} H_q(X) = -\log \left\{ \frac{\inf(p(x))}{\sum\limits_{x \in \chi} p(x)} \right\}. \tag{3.30}$$

Proof: As $q \to -\infty$, with $0 \le p(x) < 1$ the sum $\sum\limits_{x \in \chi} p^q(x)$ is dominated by $\inf(p^q(x))$. Therefore

$$H_{-\infty}(X) = \lim_{q \to -\infty} \frac{1}{1-q} \log \left\{ \frac{\inf(p^q(x))}{\sum\limits_{x \in \chi} p(x)} \right\} \tag{3.31}$$

Since $p(x)$ is positive, the power q can be moved as follows

$$H_{-\infty}(X) = \lim_{q \to -\infty} \frac{1}{1-q} \log \left\{ \frac{[\inf(p(x))]^q}{\sum\limits_{x \in \chi} p(x)} \right\}$$

$$= \lim_{q \to -\infty} \frac{q}{1-q} \log \left\{ \frac{\inf(p(x))}{\sum\limits_{x \in \chi} p(x)} \right\} \tag{3.32}$$

And taking the limit gives

$$H_{-\infty}(X) = -\log \left\{ \frac{\inf(p(x))}{\sum\limits_{x \in \chi} p(x)} \right\}. \tag{3.33}$$

$\square$

A final important property of the Rényi generalized entropy is that it is a monotonically non-increasing function. This characteristic is formulated and proven as follows.

Lemma 3.4: For a complete probability distribution $p(x)$, the Rényi generalized entropy $H_q(X)$ of order q is a monotonically non-increasing function in q.

Proof: Taking the first derivative of $H_q(X)$ gives

$$H'_q(X) = \frac{(1-q)\left(\dfrac{\sum\limits_{x \in \chi} p^q(x) \log p(x)}{\sum\limits_{x \in \chi} p^q(x)}\right) + \log \sum\limits_{x \in \chi} p^q(x)}{(1-q)^2}. \tag{3.34}$$

Eq. 3.34 can be rearranged to

$$H'_q(X) = \frac{(1-q)\sum\limits_{x \in \chi} p^q(x) \log p(x) + \sum\limits_{x \in \chi} p^q(x) \log \sum\limits_{x \in \chi} p^q(x)}{(1-q)^2 \sum\limits_{x \in \chi} p^q(x)}. \tag{3.35}$$

The denominator of Eq. 3.35 is clearly positive since $0 \le p(x) \le 1$.

Using Theorem 3.2 the function

$$f(x) = x\log x \tag{3.36}$$

which has first and second derivatives of

$$f'(x) = \log x + \frac{x}{x} = \log x + 1 \tag{3.37}$$

$$f''(x) = \frac{1}{x} \tag{3.38}$$

is convex since the second derivative is non-negative everywhere using the valid domain of $\log x$ as $x > 0$.

Using Jensen's inequality from Theorem 3.3 and with $x = p^{q-1}$ for the general convex function $f(x) = x\log x$, the following inequality can be stated

$$\varepsilon[p^{q-1}(x)\log(p^{q-1}(x))] \geq \varepsilon[p^{q-1}(x)]\log\{\varepsilon[p^{q-1}(x)]\} \tag{3.39}$$

This inequality can be expanded using Theorem 3.1 as follows.

$$\sum_{x \in \chi} p^q(x)\log(p^{q-1}(x)) \geq \sum_{x \in \chi} p^q(x)\log \sum_{x \in \chi} p^q(x) \tag{3.40}$$

Rearranging Eq. 3.40 gives

$$-\sum_{x \in \chi} p^q(x)\log(p^{q-1}(x)) + \sum_{x \in \chi} p^q(x)\log \sum_{x \in \chi} p^q(x) \leq 0 \tag{3.41}$$

and pulling down the exponent $q - 1$ gives

$$(1-q)\sum_{x \in \chi} p^q(x)\log p(x) + \sum_{x \in \chi} p^q(x)\log \sum_{x \in \chi} p^q(x) \leq 0. \tag{3.42}$$

The left hand side of Eq. 3.42 turns out to be the numerator of $H'_q(X)$ from Eq. 3.35. For $H'_q(X)$, with the denominator positive and the numerator zero or negative everywhere, the Rényi generalized entropy $H_q(X)$ is therefore a non-increasing function in q. $\qquad \square$

Lemmata 3.2, 3.3, and 3.4 define the range and general behaviour of the Rényi generalized entropy $H_q(X)$. From these lemmata it can be seen that $H_q(X)$ is bounded and is a monotonic non-increasing function. Plotted in Fig. 3.8 is a typical example of the Rényi generalized entropy $H_q(X)$ versus q for the probability distribution $p(x)$ listed in the right of the figure. This monotonic non-increasing S-curve is typical for $H_q(X)$. The Rényi generalized entropy curve will collapse to a horizontal line between the $\pm\infty$ limits only when the probability distribution $p(x)$ is uniform, as illustrated in the example in Fig. 3.9.
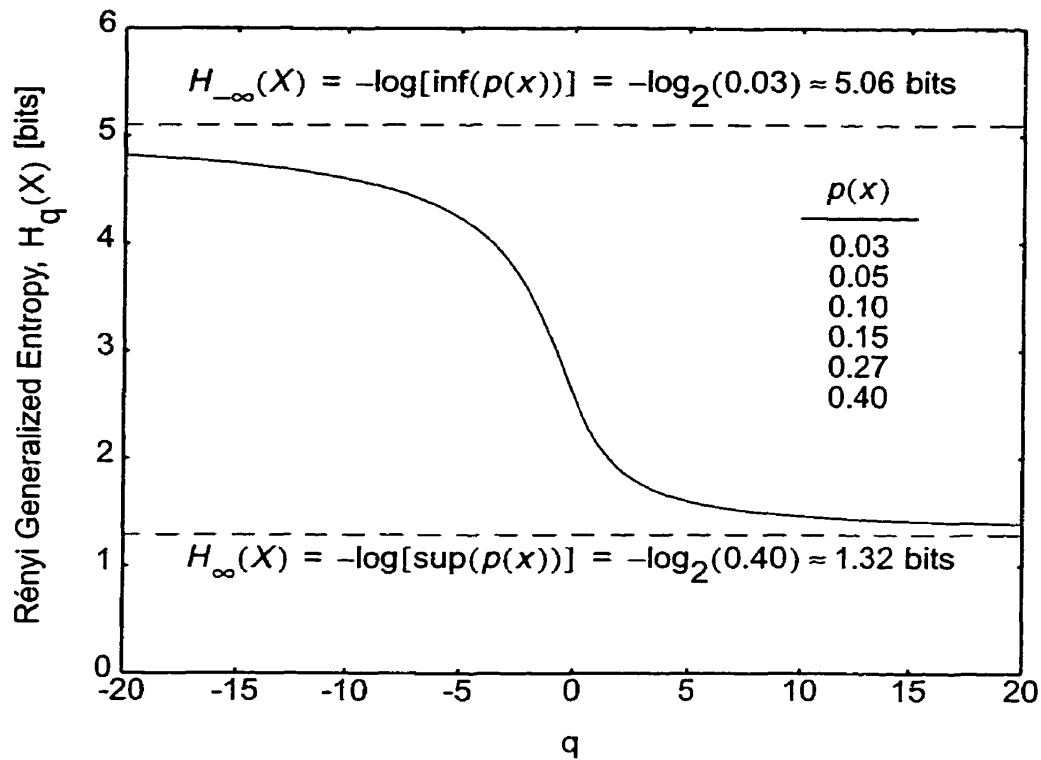
Fig. 3.8. Example 1 of Rényi generalized entropy $H_q(X)$ versus order q plot.

It should be noted that the Rényi generalized entropy is, in itself, not a dimension measure. The Rényi generalized entropy does not measure the probabilities to an infinite resolution of scale to find a critical exponent in a power-law relationship as demonstrated by the fractal dimension measures in Sec. 3.2. The Rényi generalized entropy does however serve as the basis for multifractal dimension measures. The next section describes the Rényi dimension spectrum, which is a multifractal measure based on the Rényi generalized entropy.
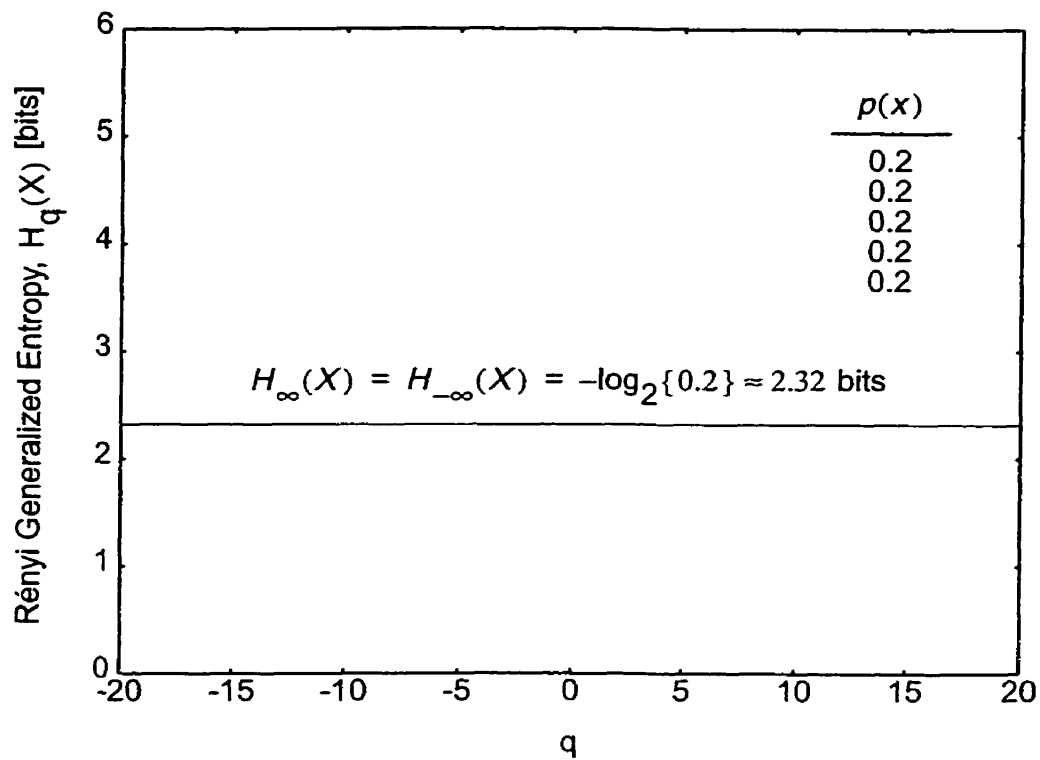
Fig. 3.9. Example 2 of Rényi generalized entropy $H_q(X)$ versus order q plot.

3.3.4   Rényi Dimension Spectrum

The Rényi dimension spectrum was first introduced by Hentschel and Procaccia [HePr83] but is named in honour of Rényi since the dimension is effectively an extension of his Rényi generalized entropy introduced in 1960 [Rény60]. The Rényi dimension spectrum is an infinite number of generalized fractal dimensions which covers the range of many previously known fractal dimension measures.

The Rényi dimension spectrum can be formulated as follows. If the Rényi generalized entropy from Eq. 3.23 is rewritten as

$$H_q(X) = \log\left(\frac{\sum\limits_{x \in \chi} p^q(x)}{\sum\limits_{x \in \chi} p(x)}\right)^{(1-q)^{-1}} \tag{3.43}$$

and using the portion inside the logarithm, then the following power-law relationship, in the spirit of Eq. 3.5, can be expressed

$$\left(\frac{\sum\limits_{x \in \chi_s} p_s^q(x)}{\sum\limits_{x \in \chi_s} p_s(x)}\right)^{(1-q)^{-1}} \propto \left(\frac{1}{s}\right)^{D_q(X)} \tag{3.44}$$

where $p_s(x)$ is the probability measurements using vels at scale s. Solving for the critical exponent $D_q(X)$, the following definition for the Rényi dimension spectrum is formed.

Definition 3.5: The *Rényi dimension spectrum* $D_q(X)$ (for brevity just D_q) of order q with measurements of probability $p_s(x)$ at scale s is defined as

$$D_q = D_q(X) \equiv \lim_{s \to \infty} \frac{1}{1-q} \frac{\log\dfrac{\sum\limits_{x \in \chi_s} p_s^q(x)}{\sum\limits_{x \in \chi_s} p_s(x)}}{\log\left(\dfrac{1}{s}\right)} = \lim_{s \to \infty} \frac{H_q(X)}{\log\left(\dfrac{1}{s}\right)} \tag{3.45}$$

where $H_q(X)$ is the Rényi generalized entropy for a discrete random variable X.   ❑

Assuming an ordinary discrete random variable X where $\sum\limits_{x \in \chi} p_s(x) = 1$, Eq. 3.45 reduces to the formulation of multifractal dimension spectrum as introduced by Hentschel and Procaccia [HePr83].

The Rényi dimension spectrum has many of the same characteristics as the Rényi generalized entropy as follows.

Lemma 3.5: Let the moment order $q \to \infty$. Then the Rényi dimension spectrum $D_q(X)$ for an ordinary discrete random variable X becomes

$$D_\infty(X) \equiv \lim_{q \to \infty} D_q(X) = \lim_{s \to \infty} \frac{H_\infty(X)}{\log\left(\frac{1}{s}\right)} = \lim_{s \to \infty} \frac{-\log\{\sup(p_s(x))\}}{\log\left(\frac{1}{s}\right)}. \qquad (3.46)$$

$\square$

Lemma 3.6: Let the moment order $q \to -\infty$. Then the Rényi dimension spectrum $D_q(X)$ for an ordinary discrete random variable X becomes

$$D_{-\infty}(X) \equiv \lim_{q \to -\infty} D_q(X) = \lim_{s \to \infty} \frac{H_{-\infty}(X)}{\log\left(\frac{1}{s}\right)} = \lim_{s \to \infty} \frac{-\log\{\inf(p_s(x))\}}{\log\left(\frac{1}{s}\right)}. \qquad (3.47)$$

$\square$

Lemma 3.7: The Rényi dimension spectrum $D_q(X)$ of order q is a monotonically non-increasing function of q.

Proof: Follows from $H_q(X)$ being monotonically non-increasing from Lemma 3.4. $\square$

Corollary 3.1: $D_q \geq D_{q'}$ for $q' > q$ with equality if and only if the object is homogeneous.

Proof: The inequality follows directly from Lemma 3.7. The equality follows from Lemma 3.5 and Lemma 3.6 when $D_{-\infty} = D_\infty$ with a more rigorous proof by Hentschel and Procaccia [HePr83]. $\square$
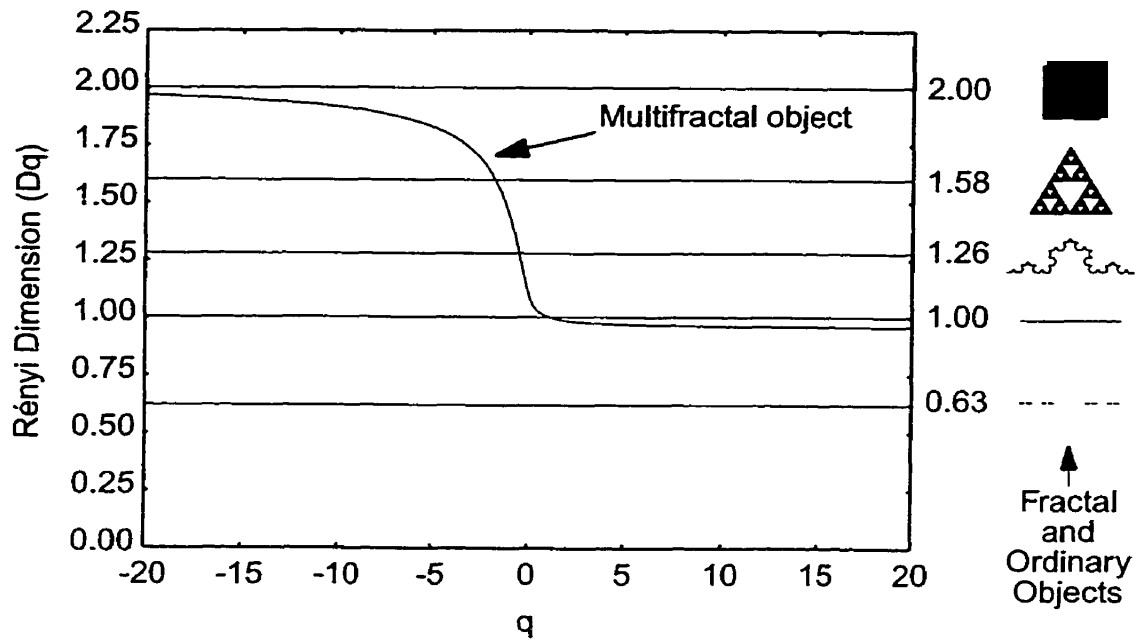
Fig. 3.10. Example Rényi dimension spectrum calculation for the Cantor set, a line, the Koch curve, the Sierpinski gasket, a box, and a multifractal object.

The important feature of $D_q(X)$ is that it does not give only one number for the measurement of the fractal dimension, but, rather an entire spectrum of dimensions. If Fig. 3.10 is considered, which has a plot of D_q versus q for a number of fractal objects, it is noticed that for simple objects such as a line and a box that the dimension D_q is 1 and 2, respectively, for all values of q. For self-similar single fractal objects such as the Koch curve, the Sierpinski gasket, and the Cantor set [Cant83] (not described in this text) the fractal dimensions are 1.26, 1.58, and 0.68, respectively, for all values of q. These objects contain only one level of complexity throughout the object. Therefore, regardless of the value of q, the fractal dimension remains constant. This is the same property as with Rényi generalized entropy illustrated in the example in Fig. 3.9 where all probabilities in $p_s(X)$ are equal. This equality follows from Corollary 3.1 because objects such as a line,

a box, the Koch curve, the Sierpinski gasket, and the Cantor set all have uniform probabilities within the set since the parts are always similar to the whole.

When more complex objects such as inhomogeneous fractals which contain more then one fractal complexity, otherwise known as multifractals, are measured, then S-curves similar to the one illustrated in Fig. 3.10 result. This curve follows partially from Lemma 3.7. In multifractals the probabilities $p_s(x)$ in X are not uniform between different regions of the object being measured. Therefore, the order q will emphasize the different probabilities $p_s(x)$ depending on the value of q. When $q = 1$ then the fractal measurement reduces to the morphological dimension as described in Sec. 3.2 and measures the dominant fractal feature in the multifractal. When $q \neq 1$, then the order q helps to suppress the dominant fractal and emphasize other inhomogeneous fractal features within the object. In actuality, the orders of q extract a spectrum of fractal dimensions from the object which can be used to characterize the object.

The exposition on the Rényi dimension spectrum is now complete. The following subsection describes an alternate form of multifractal measure, the Mandelbrot spectrum, which is a transformation of the Rényi dimension spectrum.

3.3.5   Mandelbrot Spectrum

The Mandelbrot spectrum (as referred to by Kinsner [Kins94a]) is another multifractal dimension measure that has some interesting analytical characteristics. One of the earlier formulations of what will be referred to as the Mandelbrot spectrum in this thesis was by Halsey *et al.* [HJKP86] and other useful descriptions have also been given

[AtSV88], [Kins94a], [Chen97]. The Mandelbrot spectrum can be developed by first considering local calculations in single vels of an object covering. In a completely homogeneous fractal object (*i.e.* $p_s(x)$ is constant at any scale) the local fractal dimension measures will be the same for all vels [HePr83], [HJKP86]. This is not the case in a multifractal object since it has a inhomogeneous probability distribution of X of fractal complexity throughout the object. Therefore, the local fractal dimension measures are likely different from vel to vel.

Since multifractals have inhomogeneous probabilities, it is useful to form a measure that can characterize the change in probabilities. With single fractals the probability $p_{s,\,i}(x)$ for the *i*th vel of size $r = 1/s$ follows the following power-law relationship

$$p_{s,\,i}(x) \propto r^D \tag{3.48}$$

Notice the correspondence with the power-law relationship of Eq. 3.5. With multifractals this probability will change depending on the scale as well as the vel, so the following power-law relationship can be made

$$p_{s,\,i}(x) \propto r^{\alpha_r(x)} \tag{3.49}$$

where $\alpha_r(x)$ is a scaling index that describes the variation of the probability versus the variation of r [AtSV88]. With Eq. 3.49, the moment order q can be taken for both sides of the proportionality to bring it more in the form of the Rényi generalized entropy as follows.

$$p_{s,\,i}^q(x) \propto r^{q\alpha_r(x)} \tag{3.50}$$

Now consider the number of times $N_r(\alpha)$ that α takes on a value in the range $\alpha \in [\alpha', \alpha' + d\alpha']$ having an object, or singularity, density $\rho(\alpha')$ in that range to be of the form

$$N_r(\alpha) = \rho(\alpha')r^{-f(\alpha')}d\alpha' \tag{3.51}$$

where $f(\alpha')$ is a continuous function [HJKP86]. Combining Eq. 3.51 with Eq. 3.50, gives the following probability moment summation similar to the $\sum\limits_{x \in \chi} p^q(x)$ term in Rényi dimension spectrum (Eq. 3.45). This combination can be written as follows

$$\int N_r(\alpha)p^q(x)d\alpha \propto \int \rho(\alpha')r^{-f(\alpha')}r^{q\alpha'(x)}d\alpha' \tag{3.52}$$

Seeing that $s \rightarrow \infty$ in Eq. 3.45 then $r \rightarrow 0$ (recall that $r = 1/s$), the right side of this proportionality is dominated by the value α' that makes $q\alpha' - f(\alpha')$ the smallest. Replacing α' by the value α_q that minimizes $q\alpha' - f(\alpha')$, a new formulation of D_q can be expressed as [HJKP86]

$$D_q = \frac{1}{q-1}[q\alpha_q - f(\alpha_q)] \tag{3.53}$$

Solving this equation for $f(\alpha_q)$ and α_q gives the following definition for the Mandelbrot spectrum.

Definition 3.6: The *Mandelbrot spectrum* $f_q(\alpha) \equiv D_{Man}$ of order q is defined as

$$f_q(\alpha) \equiv D_{Man} = q\alpha_q - (q-1)D_q.. \tag{3.54}$$

where α_q is the value that minimizes α' in $q\alpha' - f(\alpha')$ and is expressed as

$$\alpha_q = \frac{d}{dq}[(q-1)D_q] \tag{3.55}$$

$\square$

Taking the first and second derivatives of $q\alpha' - f(\alpha')$ and replacing α' by $\alpha(q)$ to find the extrema [HJKP86] gives

$$\frac{d}{d\alpha'}[q\alpha' - f(\alpha')]\bigg|_{\alpha' = \alpha(q)} = 0 \tag{3.56}$$

and

$$\frac{d^2}{d(\alpha')^2}[q\alpha' - f(\alpha')]\bigg|_{\alpha' = \alpha(q)} > 0 \tag{3.57}$$

From Eq. 3.56 and Eq. 3.57, solving for the first and second derivatives of $f(\alpha)$ with respect to α [HJKP86] respectively gives

$$\frac{df}{d\alpha} = q \tag{3.58}$$

and

$$\frac{d^2 f}{d\alpha^2} < 0 \tag{3.59}$$

This gives us some important properties for $f(\alpha)$ in that it is a concave function with an maximum extrema when $q = 0$. Some other important and useful properties should be noticed about the Mandelbrot as stated in the following lemmata.

Lemma 3.8: Let the moment order $q = 0$. Then the Rényi dimension spectrum D_q is related to $f(\alpha)$ in the Mandelbrot spectrum as follows

$$D_{q=0} = \sup f(\alpha). \tag{3.60}$$

Proof: With $q = 0$ then $D_{q=0}$ is

$$D_{q=0} = \frac{1}{q-1}[q\alpha_q - f(\alpha_q)]\bigg|_{q=0} = f(\alpha_0) \tag{3.61}$$

Using Eq. 3.58, $f(\alpha_0)$ is an extrema for $f(\alpha)$ since $q = 0$. Eq. 3.59 shows that $f(\alpha)$ is concave, therefore this extrema is a maximum. Therefore

$$D_{q=0} = f(\alpha_0) = \sup f(\alpha) \tag{3.62}$$

❑

Lemma 3.9: Let the moment order $q \rightarrow \infty$. Then the Rényi dimension spectrum D_q approaches

$$D_{q \rightarrow \infty} = \alpha_{min}. \tag{3.63}$$

Proof: Using the first and second derivatives of $f(\alpha)$ with respect to α in Eq. 3.58 and Eq. 3.59, respectively it is noticed that $f(\alpha)$ is a concave continuous function. As $q \rightarrow \infty$ the tangent to $f(\alpha)$ approaches infinity according to Eq. 3.58. This is only possible according to Eq. 3.59 if α is approaching its minimum.   ❑

Lemma 3.10: Let the moment order $q \rightarrow -\infty$. Then the Rényi dimension spectrum D_q approaches

$$D_{q \rightarrow -\infty} = \alpha_{max}. \tag{3.64}$$

Proof: Using the first and second derivatives of $f(\alpha)$ with respect to α in Eq. 3.58 and Eq. 3.59, respectively it is noticed that $f(\alpha)$ is a concave continuous function. As $q \rightarrow -\infty$ the tangent to $f(\alpha)$ approaches infinity according to Eq. 3.58. This is only possible according to Eq. 3.59 if α is approaching its maximum.   ❑

An example plot of the Mandelbrot spectrum $f(\alpha)$ versus α is shown in Fig. 3.11. The concavity $f(\alpha)$ is clearly seen in this plot along with the markings for $D_0 = f_{max}(\alpha)$, $D_\infty = \alpha_{min}$, and $D_{-\infty} = \alpha_{max}$.

$f(\alpha)$

$$D_0 = f_{max}(\alpha)$$

$$D_{-\infty} = \alpha_{max} \qquad D_{\infty} = \alpha_{min}$$

α
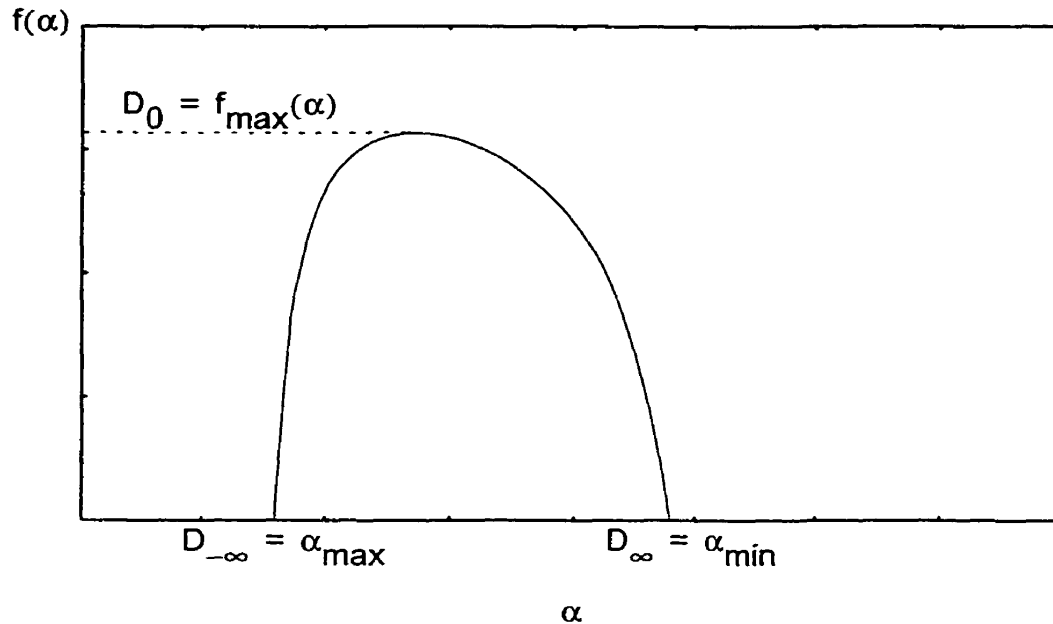
Fig. 3.11. Example plot of Mandelbrot spectrum $f(\alpha)$ versus α.

This completes the general background on multifractals. The next section moves into another area of fractals dealing with the generation of statistically self-similar or in other words self-affine fractals. This is all based on a concept known as fractional Brownian motion.

3.4   Brownian Motion and Fractional Brownian Motion

One question that may be asked is how to generate fractals, and more specifically how to generate random fractals that have a specified fractal dimension. The most prominent method is to generate *fractional Brownian motion* (fBm) [MaVa68], which is a conceptual extension of *Brownian motion* [Brow28]. The following subsections describe

Brownian motion and fractional Brownian motion, as well as an extension into multifractals with a brief discussion on multifractional Brownian motion [PeVe95].

3.4.1   Brownian Motion

In 1827 the botanist Robert Brown (1773-1858) observed irregular movements of pollen grains under a microscope, but, unfortunately did not have a good explanation for such behaviour [Brow28]. This observed phenomenon came to be known as *Brownian motion*. In the early 1900's, the Swedish chemist Theodor Svedberg suggested that this type of motion is due to the unequal bombardment of small particles by molecules. This hypothesis is now known to be the case and gave rise to some of the first measurements of atom size from a set of equations that Einstein developed for Brownian motion.

The motion of the grains of pollen that Robert Brown saw would have been very similar in nature to the path illustrated in Fig. 3.12a. A trail of a possible path of the pollen grain is shown in a two dimensional plane where the particle can collide with other particles that are uniformly randomly distributed throughout the plane. This motion can be described by making a record of the two dimensional random walk in the complex plane z as follows

$$z_j = z_{j-1} + n(x;0, 1)e^{i2\pi u(x)} \tag{3.65}$$

where $u(x)$ is a uniform random variable on $[0, 1)$ and $n(x;0, 1)$ is a zero mean Gaussian random variable with unity standard deviation. The Gaussian density function used with mean μ and standard deviation σ has the form

$$n(x;\mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \tag{3.66}$$

The initial condition z_0 is the position of the particle in the complex plane when observations start.

As Fig. 3.12a illustrates, the motion of the particle appears quite random as it travels, but some constraints and characteristics of its motion should be noted. The first constraint is that there is continuity in the motion of the particle. This property arises since the particle has finite velocity. With this constraint, the particle is still free to move in any direction, giving the particle two degrees of freedom in this illustration.

Another constraint on the movement of the particle is the density of the particles surrounding it. The particle will travel in a straight line at a constant speed until it collides with a neighbouring particle, causing it to change direction and speed. Therefore, the density of the particles affects the average distance that any particular particle can travel before colliding with another particle.
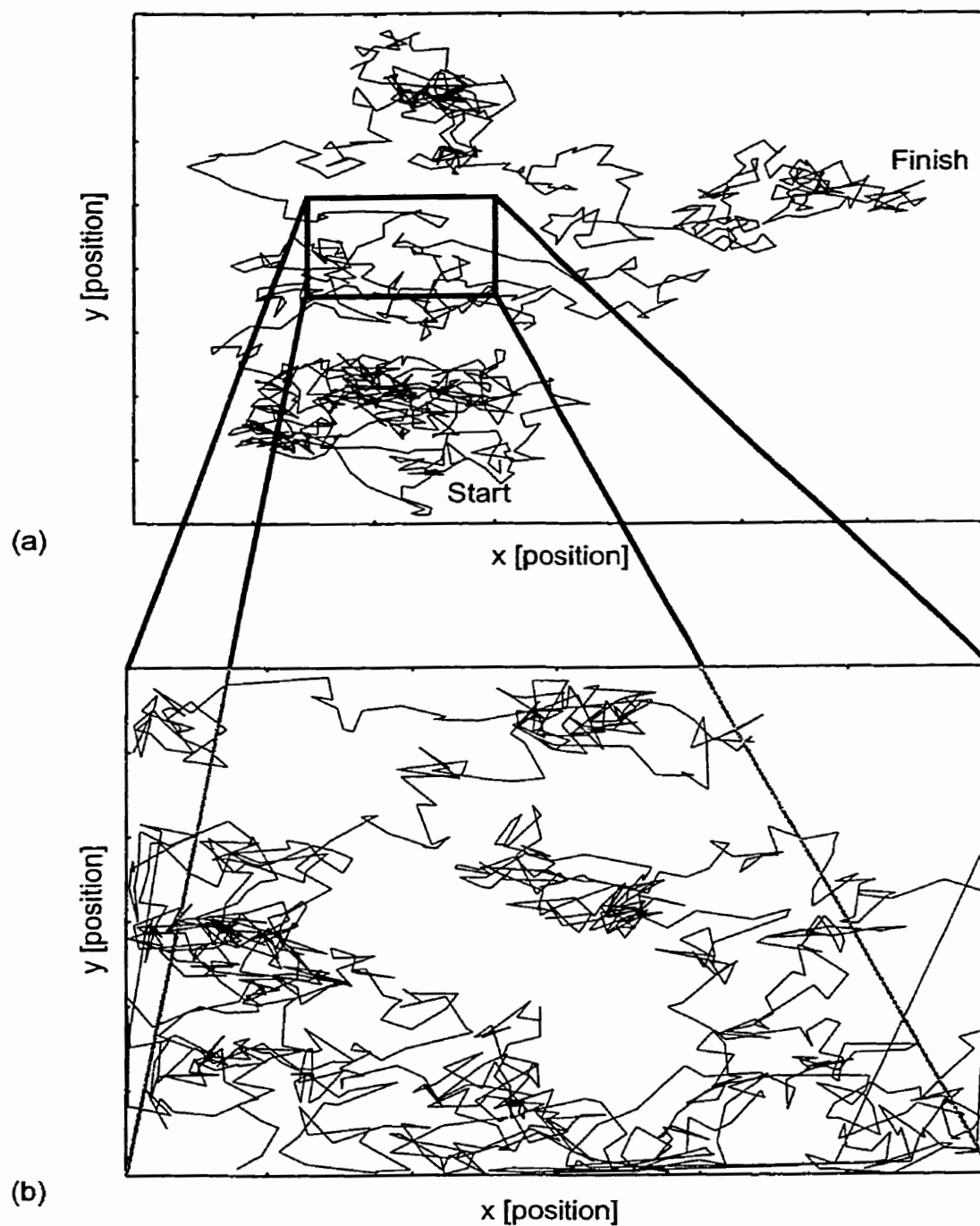
Fig. 3.12. Example 2D Brownian motion of a particle.
(a) particle's path and (b) magnified portion of the particle's path.

A few other characteristic attributes behind the particle motion illustrated in Fig. 3.12a should also be pointed out. First, consider the average distance that the particle travels in time t. If in time t the particle collides with n other particles, which are uniformly randomly distributed, and the displacement vector travelled between collisions is l_n, then the particle will have a displacement of

$$\text{displacement} = \sum_{\forall n} l_n \tag{3.67}$$

in time t. Using the Gaussian random variable from Eq. 3.65, Eq. 3.67 can be rewritten as

$$\text{displacement} = \sum_{\forall n} l_n = \sum_{\forall n} n(x;0,1)e^{i2\pi u(x)} \tag{3.68}$$

Another important note about the Brownian motion is that the direction travelled after each collision follows a uniformly distributed probability density function. This characteristic follows from the phasor notation of $e^{i2\pi u(x)}$ used in Eq. 3.65. The addition of these displacement vectors l_n in Eq. 3.67 results in an average displacement or mean of

$$\mu_{displacement} = \sum_{\forall n} l_n = \sum_{\forall n} n(x;0,1)e^{i2\pi u(x)} = 0 \tag{3.69}$$

Another characteristic of Brownian motion to consider is the mean square displacement of the particle in time t. Using the *central limit theorem*, the mean square displacement of the particle follows a Gaussian distribution. This is the case since the mean square displacement using l_n^2 is independent and identically distributed, and the variance is finite and non-zero because of the continuity in the particle's path as previously described. Therefore, at specific snapshots in time the particle will have moved by a zero mean Gaussian random variable from its current location even though its long-term mean displacement is still zero. Figure 3.12 was actually generated this way by displacing a

simulated particle by a Gaussian random variable along a directed vector in the current direction of motion. At each collision in the simulation, a new direction vector was determined from a uniformly distributed random variable.

Another point to be made is that the Brownian motion considered in this thesis will be from a mathematical standpoint and thus have infinite resolution. This condition is as opposed to a physical situation which is limited in resolution, as in the case of a pollen grain colliding with other particles at a finite resolution. In the mathematical case, any portion of the Brownian motion can be magnified to reveal still more Brownian motion as illustrated in Fig. 3.12b. This ability demonstrates the scale invariance that Brownian motion exhibits. Since the magnification scale is irrelevant, no numbers have been marked along the axes in Fig. 3.12a or Fig. 3.12b since the numbers have little meaning in this case. The emphasis is that the Brownian motion can be magnified to any level and exhibit the same Brownian motion characteristics.

Since the Brownian motion can be magnified to any level, Brownian motion also exhibits one of the important characteristics behind fractals. All fractals have the characteristic of self-similarity such as was discussed with the Koch curve. If the trace of Brownian motion can be magnified infinitely and always have the same Brownian motion characteristics at any scale, then Brownian motion follows the self-similarity trait exhibited by other fractals, where the parts are similar to the whole. Though, since the Brownian motion characteristics are statistically similar instead of perfectly similar, such as with the Koch curve, this property is generally referred to as *self-affinity* as opposed to self-similarity.

Another example of Brownian motion is to take a record over time of one ordinate of the two dimensional Brownian motion. An example of this record is plotted in Fig. 3.13. This plot shows Brownian motion at different samples through time. All of the previously discussed properties of Brownian motion still hold.
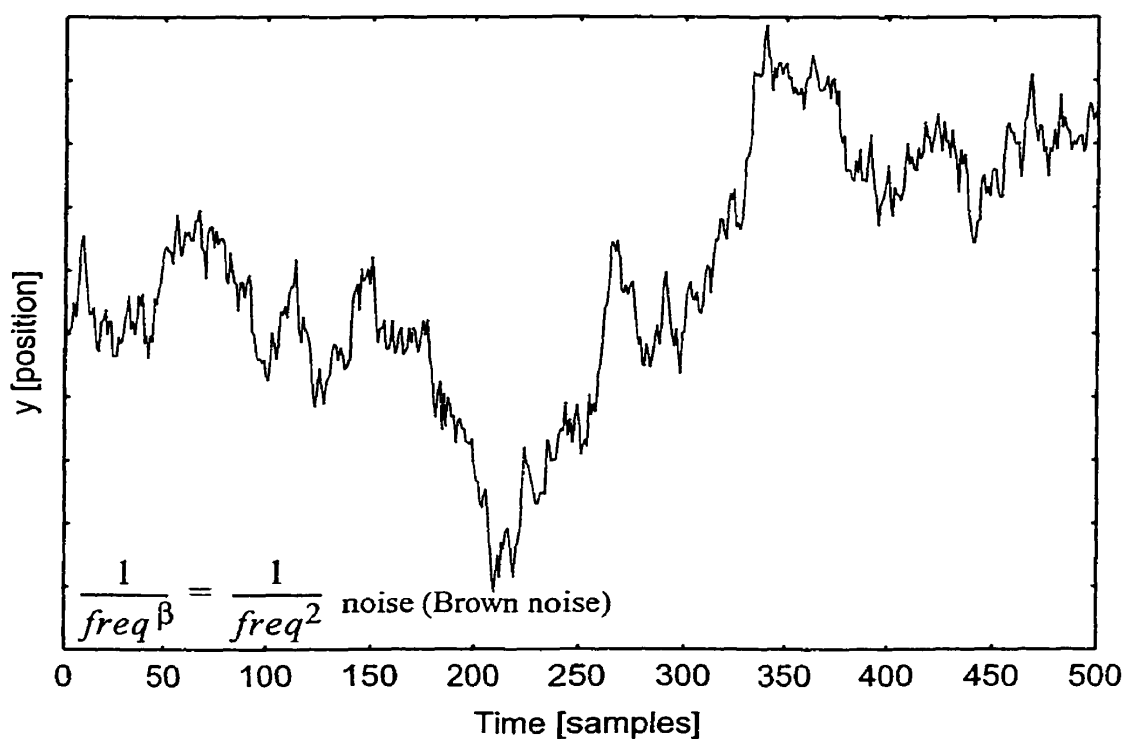
Fig. 3.13. Record of Brownian motion in one dimension.

To define the concept of Brownian motion a little more formally, consider the following definition for the one dimensional Brownian motion process.

Definition 3.7: Let $B(t, w)$ be a continuous-time stochastic process defined with time $-\infty < t < \infty$ and w the set of all values of a random function. Also, let the sequence of increments $B(t + \delta) - B(t)$ (defined for t a multiple of δ) be an independent Gaussian

random variable with zero mean and a variance of $|\delta|$. In addition, let any two non-inter-

secting intervals, $[t_1, t_1 + \delta_1] \cap [t_2, t_2 + \delta_2] = \varnothing$, of $B(t, w)$ be independent and have

the same joint distribution for all δ. If these conditions hold then $B(t, w)$ is known as a

Brownian motion process or also called a *Wiener process*.   ❏

This definition of Brownian motion brings together all of the previously discussed

points about Brownian motion. The following subsection expands on the idea of Brown-

ian motion to fractional Brownian motion. As will be seen, fractional Brownian motion is

a generalization of Brownian motion where the persistence in motion can be changed to

form objects with different statistical characteristics.

3.4.2   Fractional Brownian Motion

This section describes fractional Brownian motion and how it is a generalization of

Brownian motion. This discussion follows some of the computer experiments done by

Mandelbrot and Wallis [MaWa69a], [MaWa69b], [MaWa69c] to generate fBm and look at

some of its characteristics, as well as some of the compiled descriptions by Kinsner

[Kins94c], [Kins95a]. The next subsections give an overview of fBm including some of

fBm's characteristics and some examples of this motion. Techniques for generating fBm

curves and surfaces will be left to Chapter 5 where they are needed for one of the progres-

sive image transmission techniques in this thesis which uses fractal surface segmentation

and interpolation.

3.4.2.1   What is fractional Brownian motion?

Fractional Brownian motion, as introduced by Mandelbrot and van Ness [MaVa68], is a generalization of Brownian motion in terms of the level of persistence or anti-persistence that it possesses. Fractional Brownian motion can be defined as follows.

Definition 3.8: Let $B(t, w)$ be a Brownian motion or Wiener process with time $-\infty < t < \infty$, w the set of all values of a random function, and b_0 an arbitrary real number. For index H and $t > 0$, *fractional Brownian motion* $B_H(t, w)$ is defined by

$$B_H(0, w) = b_0$$

$$B_H(t, w) - B_H(0, w) =$$

$$\frac{1}{\Gamma(H + 1/2)} \left\{ \int_{-\infty}^{0} \left[(t - s)^{H - \frac{1}{2}} - (-s)^{H - \frac{1}{2}} \right] dB(s, w) \right.$$

$$\left. + \int_{0}^{t} (t - s)^{H - \frac{1}{2}} dB(s, w) \right\} \qquad (3.70)$$

where the gamma function is *Euler's second integral* which can be expressed as

$$\Gamma(x) = \int_{0}^{\infty} (t^{x - 1} e^{-t}) dt. \qquad (3.71)$$

□

As seen in Eq. 3.70, fractional Brownian motion is largely a summation over this Wiener process dependent on the index H. When the index $H = 1/2$ then Eq. 3.70 reduces to

$$B_{H=1/2}(t, w) - B_{H=1/2}(0, w) = \frac{\int_{-\infty}^{0} [1-1]dB(s, w)}{\Gamma(1)} + \frac{\int_{0}^{t} dB(s, w)}{\Gamma(1)}$$

$$= \frac{\int_{0}^{t} dB(s, w)}{\Gamma(1)} - 0 = \frac{\int_{0}^{t} dB(s, w)}{\int_{0}^{\infty} (t^{1-1}e^{-t})dt}$$

$$= \frac{\int_{0}^{t} dB(s, w)}{\int_{0}^{\infty} e^{-t}dt} = \int_{0}^{t} dB(s, w) = B(t, w) - B(0, w) \tag{3.72}$$

which is just ordinary Brownian motion $B(t, w)$ with an arbitrary starting point. At other values of the index H, $B_H(t, w)$ is a weighted moving average of $B(t, w)$ by the kernel $(t-s)^{H-1/2}$ [MaVa68].

The main property to notice with fBm is that the increments of $B_H(t, w)$ are stationary and self-similar. Formally, this means that the increments $B_H(t_0 + \delta, w) - B_H(t_0, w)$ have the same finite joint distribution as $h^{-H}[B_H(t_0 + h\delta, w) - B_H(t_0, w)]$, where δ is the increment size and h some multiple of the increment size. Therefore, magnifying any portion of fBm with a specific index H will result in fBm with the same characteristic index H. The following subsection shows some examples of fBm and attempts to give some further clarifications to the properties of fBm.

3.4.2.2   Characteristics of fractional Brownian motion

Before looking at examples of fractional Brownian motion, let us consider again the Brownian motion illustrated in Fig. 3.13. In examining Fig. 3.13, it is important to notice how the motion of the particle has persistence in time as it travels. This persistence can be thought of physically as the momentum the particle has between collisions with other particles. After each collision the momentum changes and a new state in its motion is achieved, so the persistence is limited. This persistence in time means that Brownian motion has a non-stationary mean as, in this case, the particle is moving slowly from one portion of the plot to another.

It should be noticed that Fig. 3.13 has no position values labeled on the ordinate. This is purposely done, as with Fig. 3.12, since the record of Brownian motion exhibits the same scaling characteristics found in the two dimensional Brownian motion illustrated in Fig. 3.12a and Fig. 3.12b. Therefore, Brownian motion has the same statistical form and characteristics at any scale of the motion.
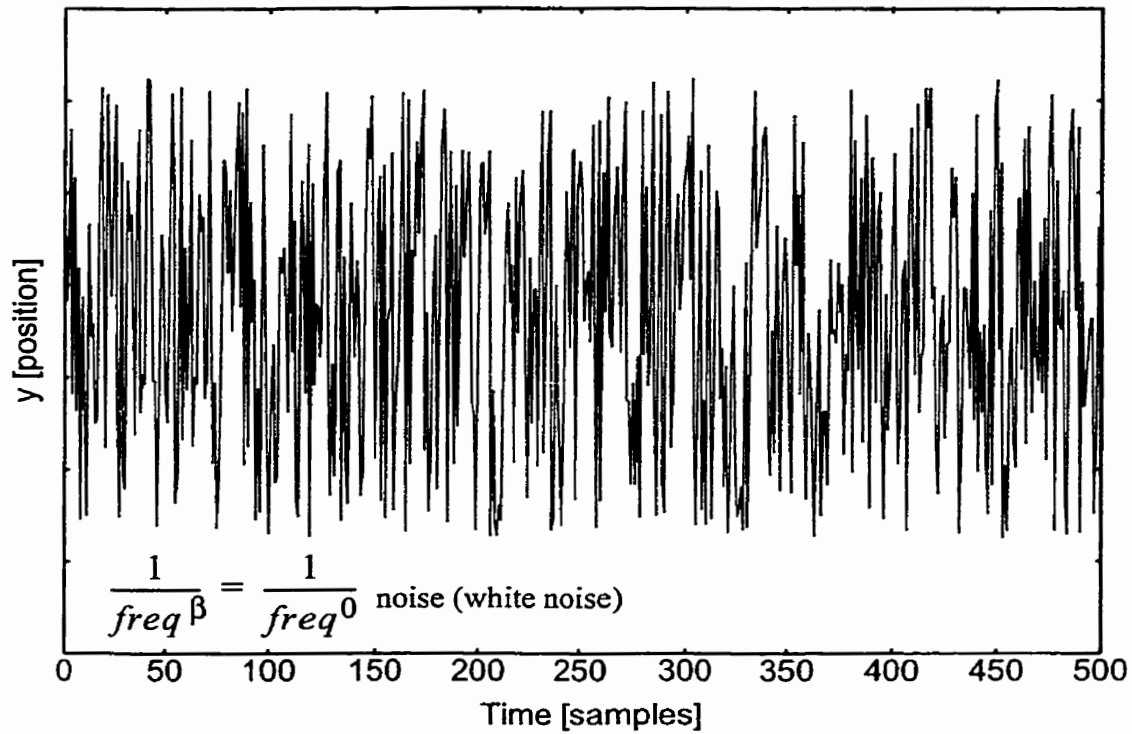
Fig. 3.14. Record of white noise motion in one dimension.

The Brownian motion's particle persistence from the last state is contrasted by the white noise illustrated in Fig. 3.14. This figure, generated by placing the particle in space at a location determined by a uniformly distributed random variable in each time step, illustrates uncorrelated and independent motion with a stationary mean. This type of uncorrelated and independent signal is generally referred to as white noise because its power spectrum, $1/freq^0$, is flat and contains all frequencies at an equal intensity. This signal is similar in concept to combining all the colours of the light spectrum to produce white light. The Brownian motion illustrated in Fig. 3.13, on the other hand, has a power spectrum density of $1/freq^2$ which means that the frequency components in the signal taper off with increases in the frequency.
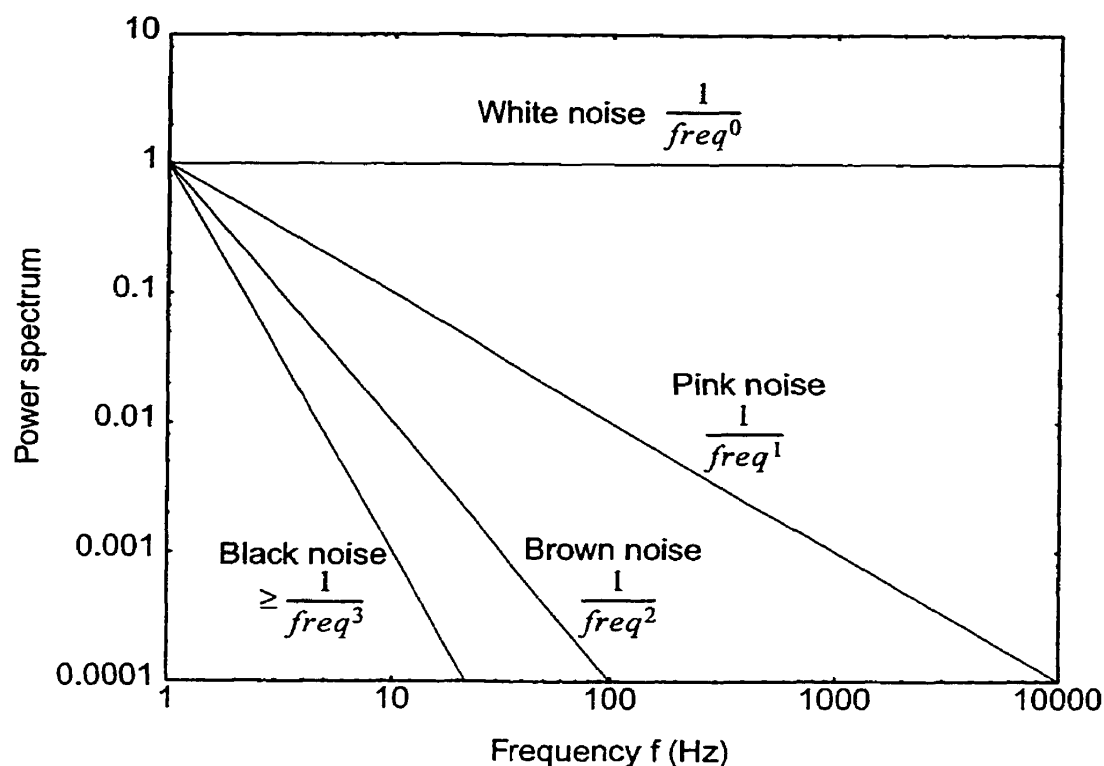
Fig. 3.15. White, pink, Brown, and black noise power spectrum plot.

Taking the colour idea of white noise a little further, a spectrum of noise or motion colours can be defined as shown in Fig. 3.15. This figure illustrates different power spectrum densities in the form $1/freq^\beta$ where β is the spectral exponent. Illustrated is white noise with $\beta = 0$, pink noise with $\beta = 1$, Brown noise with $\beta = 2$ (Brownian motion), and finally black noise with $\beta \geq 3$. These names and spectral exponents are taken as listed by Schroeder [Schr91], [Kins94c] though there is some debate over the names with many other additions such as orange, red, green, blue, purple, and grey noise [Wisn96].

$$\frac{1}{freq^{\beta}} = \frac{1}{freq^{1}} \text{ noise (pink noise)}$$
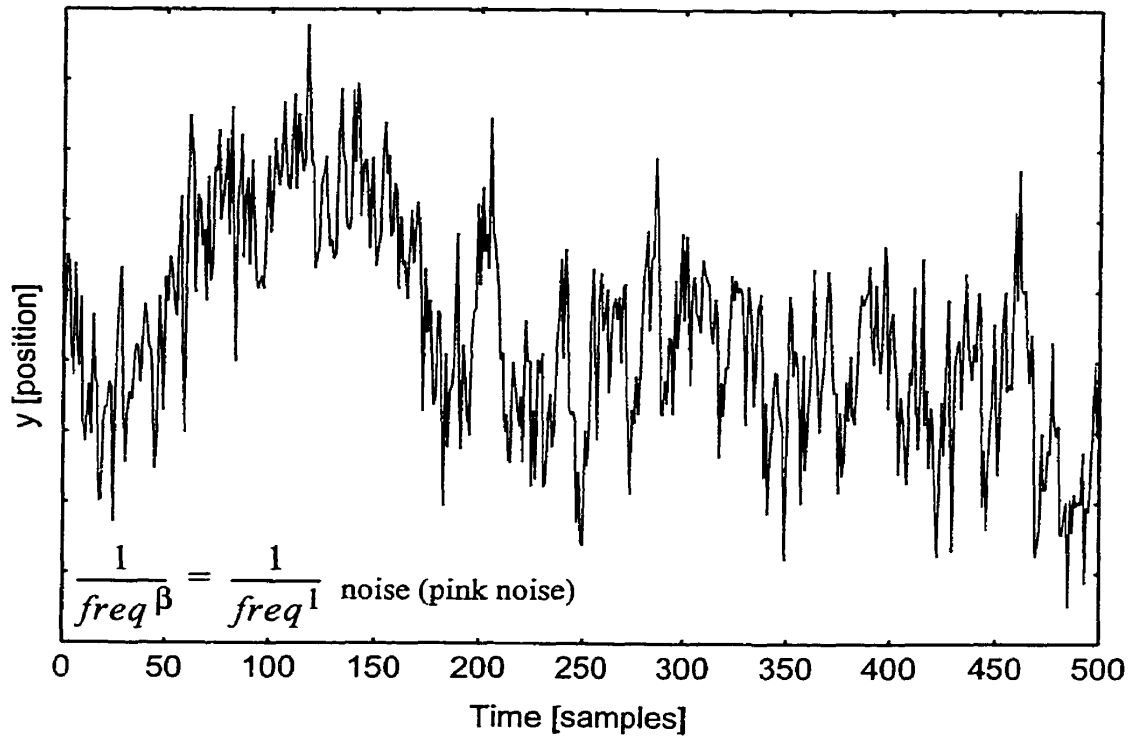
Fig. 3.16. Record of pink noise motion in one dimension.

Fractional Brownian motion extends the idea of pure Brownian motion to allow for different levels of persistence in the direction of the path over time. Changing the density of the power spectrum through changing the spectral exponent β for the signal creates different "fractional" noise values. For instance, pink noise, as illustrated in Fig. 3.16, is between white noise and Brown noise according to the power spectrum plot of Fig. 3.15. Following the idea of persistence, this says that pink noise has more persistence than white noise (which has no persistence), but, less persistence than Brown noise. With less persistence than Brown noise, pink noise is often referred to be anti-persistent in that, in the case of a time series, the system tends to revert the direction of its mean more often than not and in the limiting case of fBm will change direction of its mean at every step. Following

the idea of stationarity, pink noise is relatively more stationary than white noise but less stationary than Brown noise. Pink noise can also be referred to as quasi-stationary.

Another important issue to realize from the power spectrum plot in Fig. 3.15 is the distribution of power throughout the spectrum. For white noise the power for any fixed sized interval is equal. For example, the power in the range from 100 Hz to 200 Hz is the same as the power in the range from 10,000 Hz to 10,100 Hz. This is not the case for pink noise. For pink noise the power distribution is the same over logarithmic octaves. For example, the power in the range from 100 Hz to 200 Hz is the same as the power in the range from 10,000 Hz to 20,000 Hz.

Viewing some of these factors, it has been observed that pink noise models many natural phenomena better than white noise models, such as parallel relaxation processes [Schr91]. For example, many musical pieces have $1/f^1$ power spectra and pink noise can be used to generate interesting music, while white noise is too random and Brown noise is too correlated [VoCl78].

The figure shows a plot with y-axis labeled "y [position]" and x-axis labeled "Time [samples]" ranging from 0 to 500. Within the plot is the equation:

$$\frac{1}{freq^\beta} = \frac{1}{freq^{2.8}} \text{ noise}$$
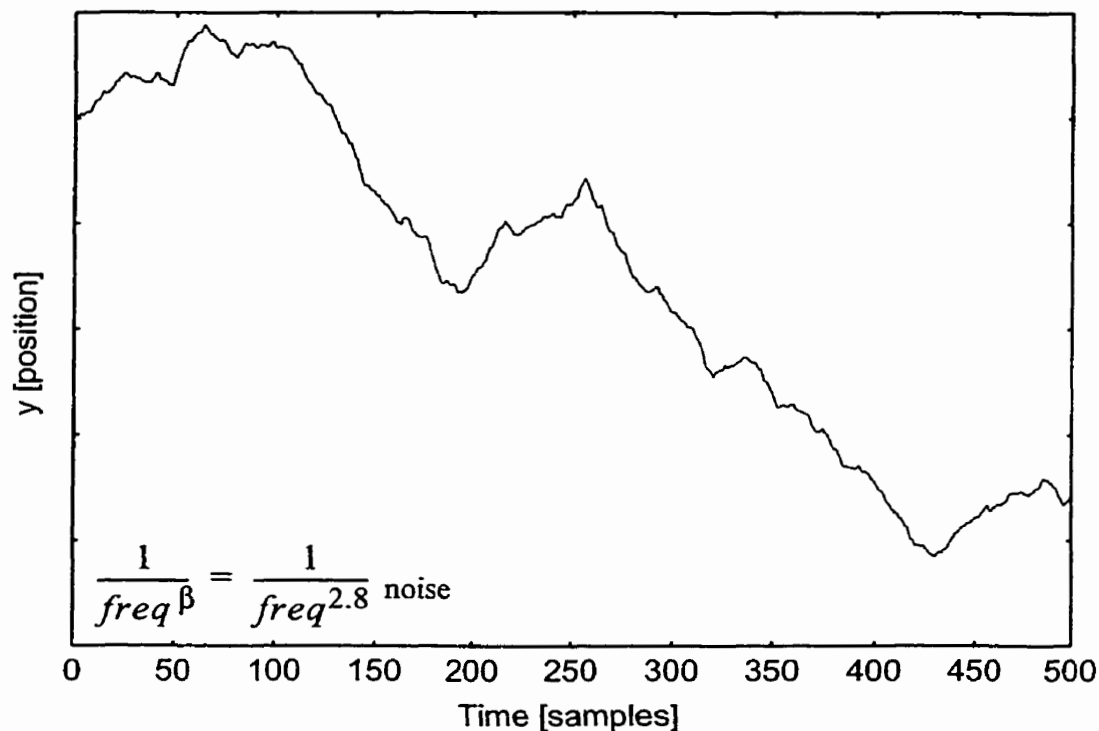
Fig. 3.17. Record of fractional Brownian motion in one dimension.

A final example to illustrate fBm is shown in Fig. 3.17. This example shows motion that is more persistent than Brownian motion since it has a power spectrum density of $1/freq^{2.8}$ with a spectral component of $\beta = 2.8$. This spectrum can be referred to as dynamically persistent in that the trend currently seen in the series is reinforced so that the trend in that direction of motion will likely continue. This form of behaviour exists in some natural phenomena such as natural disasters and flooding [Schr91].

To tie in the whole concept of persistence, Mandelbrot coined the term the *Joseph*

effect [Mand75, p. 105] after Joseph son of Jacob and son of Israel who said

> "Behold, there come seven years of great plenty throughout all the land of
> Egypt: And there shall arise after them seven years of famine; and all the
> plenty shall be forgotten in the land of Egypt; and the famine shall con-
> sume the land;"

> (Genesis 41:29-30, King James Version)

This idea of seven years of plenty and then seven years of famine in Egypt nicely illus-

trates how persistence in natural phenomena occurs.

Mandelbrot also coined the term the *Noah effect* [Mand75, p. 105] after the story

of Noah where discontinuity rules and things can change suddenly as when there was a

drought and then 40 days of rain came.

> "And it came to pass after seven days, that the waters of the flood were
> upon the earth. In the six hundredth year of Noah's life, in the second
> month, the seventeenth day of the month, the same day were all the foun-
> tains of the great deep broken up, and the windows of heaven were opened.
> And the rain was upon the earth forty days and forty nights."

> (Genesis 7:10-12, King James Version)

Accordingly, events do not necessarily gradually change but can change suddenly. These

happenings may be viewed more with the Brown to pink noise phenomena where wide

swings of change can occur and do occur suddenly.

This section outlining some of the additional properties and characteristics of fBm.

is now completed. The next subsection ties in the ideas of fBm with the work done by

Hurst, the Hurst exponent, and the spectral exponent for power spectra.

3.4.2.3   Fractional Brownian motion, the Hurst exponent, and the spectral exponent

The spectral exponent β used in the fBm power spectrum density function $1/freq^\beta$ can be related to what is known as the Hurst exponent H^* by the relationship [MaWa69b]

$$\beta = 2H^* + 1 \tag{3.73}$$

The Hurst exponent [Hurs51], [HuBS65] was developed by the British hydrologist Harold Edwin Hurst. Hurst was trying to determine if the historical yearly water level flows of the Nile showed any pattern, so that the long-term storage capacity of the Nile reservoirs could be appropriately built [Hurs51]. He found that instead of the flow levels being random each year, there was a clustering in that there would be runs of wet or dry years. Hurst's results showed that there is a non-random positive correlation in the water levels of the Nile over the 800 years of Nile flow records, as well as non-random positive correlations in some other natural phenomena such as rainfall, temperature and pressure, tree rings, varves, and sunspot activity [Hurs51], [HuBS65]. The Hurst exponent used to get these results is defined as follows

$$H^* \equiv \frac{\log\left(\frac{R(\tau)}{S(\tau)}\right)}{\log(\tau)} \tag{3.74}$$

where τ is the duration of the time series ξ_t and R/S is the rescaled range statistical analysis of the time series data. In R/S-statistics, R is the range of the time series data ξ_t and can be defined as follows

$$R(\tau) = \max_{1 \le t \le \tau} \sum_{u=1}^{t} (\xi_u - <\xi>_\tau) - \min_{1 \le t \le \tau} \sum_{u=1}^{t} (\xi_u - <\xi>_\tau) \tag{3.75}$$

while S in the R/S-statistics is the standard deviation in the time series data ξ_t and is defined as follows

$$S(\tau) = \sqrt{\left(\frac{1}{\tau}\right) \sum_{t=1}^{\tau} (\xi_t - <\xi>_\tau)^2} \tag{3.76}$$

with the operator $<\,>$ being the mean of the data as follows

$$<\xi>_\tau = \frac{1}{\tau} \sum_{t=1}^{\tau} \xi_t \tag{3.77}$$

In words, the Hurst exponent measures how the range of the time series data ξ_t changes in accordance to the standard deviation of the data over different slices of time. Therefore, the Hurst exponent gives a measure as to whether a trend will persist or if a mean will revert back to a historical average.

The Hurst exponent can be thought of as the H parameter in the fBm $B_H(t, w)$. So, we could effectively write $B_{H*}(t, w) = B_H(t, w)$. The Hurst exponent will be used in Chapter 5 for the generation of fractal surfaces using the midpoint displacement algorithm. Also, the Hurst exponent is related to the previously discussed Hausdorff-Besicovitch dimension, D_{HB}, [Mand85]. This relationship can be expressed as follows

$$D_{HB} = E + \frac{3 - \beta}{2} = E + 1 - H^* \tag{3.78}$$

where $E + 1$ is the Euclidean dimension or support dimension in which the object is embedded ($E = 0$ for a point, $E = 1$ for a curve, $E = 2$ for a surface, $E = 3$ for a volumetric object, etc.). Equation 3.78 shows that if fBm can be generated with a specific Hurst exponent, then this fBm will have a known Hausdorff-Besicovitch dimension.

3.4.2.4 *Multifractional Brownian motion*

While not a key focus of this thesis, mention should be made of an extension of

fBm to *multifractional Brownian motion* (mfBm) as introduced by Peltier and Lévy Véhel

[PeVe95]. The idea is similar to that of fBm except that instead of a single H value for

$B_H(t, w)$, a function H_t is used to represent the changes in dimension through time. This

can be seen by reformulating Eq. 3.70 to

$$mB_{H_t}(0, w) = b_0$$

$$mB_{H_t}(t, w) - mB_{H_t}(0, w) =$$

$$\frac{1}{\Gamma(H_t + 1/2)} \left\{ \int_{-\infty}^{0} \left[(t-s)^{H_t - \frac{1}{2}} - (-s)^{H_t - \frac{1}{2}} \right] dB(s, w) \right. \tag{3.79}$$

$$\left. + \int_{0}^{t} (t-s)^{H_t - \frac{1}{2}} dB(s, w) \right.$$

Using this formulation, curves can be created that are effectively concatenations of

fBm curves with different values of H.

3.5   Summary

This chapter covered the general background required for this thesis on fractals,

multifractals, and fractional Brownian motion. The chapter started by explaining an

approach to measuring the morphological topology of an object and then how this mea-

surement is used to formalize the idea of a fractional dimension. Extensions were then

made from homogeneous fractals to inhomogeneous fractals, or multifractals, where more

then one fractal complexity exists in an object. This extension is done with a generaliza-

tion of Shannon entropy using the Rényi dimension spectrum as well as the Mandelbrot

spectrum. Finally, Brownian motion and fractional Brownian motion were described mathematical for statistically self-similar, otherwise known as self-affine, fractal objects.

The next chapter provides background on wavelets and wavelet packets. This discussion will set the stage for the developed progressive image transmission in later chapters.

CHAPTER IV

BACKGROUND ON WAVELETS
AND
WAVELET PACKETS

This chapter provides the required theoretical background on wavelets and wavelet packets for this thesis. This theory will serve as the basis behind some of the progressive image transmission techniques presented in later chapters.

4.1   Preliminaries

This section gives some preliminary mathematics before the overview on wavelets and wavelet packets is provided.

4.1.1   Metric Spaces, the Hilbert Space, and Orthogonality

Wavelets and wavelet analysis techniques falls into a class of metric spaces referred to as the Hilbert space. While this section is not intended to be a complete reference on metric spaces, a brief introduction is given on preliminary concepts of metric spaces and the Hilbert space in context of wavelet analysis. Further information on metric spaces can be taken from various sources on functional analysis [Krey89], [Shil74] which served as references for this writing.

For this thesis, interest will be only on finite energy signals since anything recorded from a physical source will have finite energy. In general, a signal $f(x)$ having finite energy is square integrable as stated in the following definition and lemma.

Definition 4.1: A function $f:X \to \mathbf{R}$ is said to be *square integrable* if

$$\int_X |f|^2 < \infty. \tag{4.1}$$

❏

Lemma 4.1: A function $f:X \to \mathbf{R}$ has *finite energy* if it is square integrable.   ❏

Definition 4.2: Let $f:X \to \mathbf{R}$ be a square integrable function. Then f is referred to as an L^2-*function*. Also, the set of L^2-functions in a measurable space forms the L^2-*space*, often referred to as $L^2(\mathbf{R})$ for real valued vectors and functions.   ❏

Another important concept is that of a metric space. A metric space is a vector space (*i.e.* sets of vectors X that follow the commutative, associative, distributive, and identity properties) that also has a distance measure associated with the space. This distance measure gives a distance between two vectors in the space. More formally, a metric space is defined as follows.

Definition 4.3: A *metric space* is a pair (X, d), where X is a set and $d(x, y)$ is a metric on X, that is, the "distance from x to y" such that for all $x, y, z \in X$ the following axioms are satisfied:

1. d is real valued, finite, and nonnegative

2. $d(x, y) > 0$ if $x \neq y$, $d(x, x) = 0$ for every x

3. $d(x, y) = d(y, x)$ for every x and y

4. $d(x, z) \leq d(x, y) + d(y, z)$ for every x, y, and z

❏

Definition 4.4: A *normed space* is a vector space with a norm, denoted by the operator $\|\cdot\|$, defined on it. ❏

A norm is a measure that can be used to gauge the size of an element in a space. For the finite energy space $L^2(\mathbf{R})$ the following norm is often used

$$\|x\| = \left[\int_{-\infty}^{\infty} |f(x)|^2 \right]^{1/2} < \infty \tag{4.2}$$

and is referred to as the L^2-norm.

With this background, the Hilbert space can be defined as follows.

Definition 4.5: A *Hilbert space* is a normed vector space with a defined inner product operator $\langle \cdot, \cdot \rangle$ such that the norm

$$\|f\| = \sqrt{\langle f, f \rangle} \tag{4.3}$$

makes the vector space a metric space. ❏

The inner product in a Hilbert space can be defined in many different ways depending on the metric used for that space. For the finite energy space $L^2(\mathbf{C})$, the inner product of $f \in L^2(\mathbf{C})$ and $g \in L^2(\mathbf{C})$ is

$$\langle f, g \rangle = \int_{-\infty}^{\infty} f(x)\overline{g(x)}dx \tag{4.4}$$

where $\overline{g(x)}$ is the complex conjugate of $g(x)$. Of course, this equation reduces to

$$\langle f, g \rangle = \int_{-\infty}^{\infty} f(x)g(x)dx \tag{4.5}$$

when dealing only with the real valued functions $f \in L^2(\mathbf{R})$ and $g \in L^2(\mathbf{R})$. These inner products are valid in an infinite dimensional space so f and g can be of any dimensionality, such as one dimensional for time series and two dimensional for images.

One important characteristic of the Hilbert space is that the inner product can be used to see if any pair of vectors are orthogonal. Another important property which can be determined is whether a vector is orthonormal. These two characteristics of a Hilbert space are defined as follows.

Definition 4.6: A vector $f \in X$ is said to be *orthogonal* to a vector $g \in X$ if

$$\langle f, g \rangle = 0. \tag{4.6}$$

❑

Definition 4.7: A set of orthogonal vectors $\{\phi_k(x)\}_{k \in \mathbf{Z}}$ is said to be *orthonormal* if the norm of every vector is one, that is $\|\phi_k\| = 1$.   ❑

4.1.2   Signal Decomposition on an Orthogonal Basis

Using the idea of orthogonality, a vector can be decomposed into a set of orthogonal or orthonormal vectors. For instance given the orthonormal set $\{\phi_x, \phi_y\} \in L^2(\mathbf{R})$, the two dimensional vector v can be decomposed to $v = v_x + v_y = c_x \phi_x + c_y \phi_y$ where $\{c_x, c_y\}$ are the multiplicative coefficients in the decomposition. Graphically, this decomposition is illustrated in Fig. 4.1.
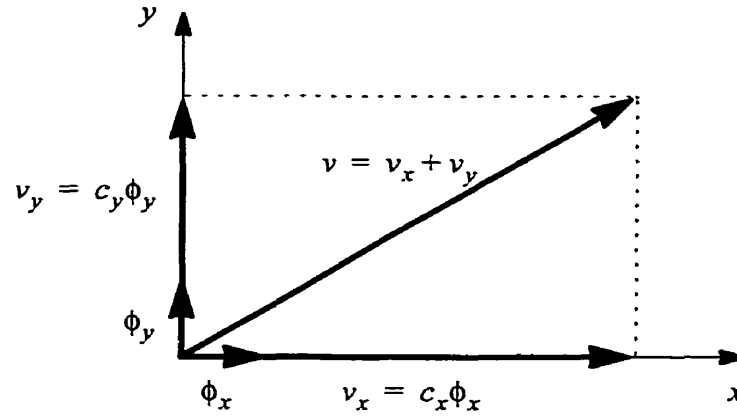
Fig. 4.1. Example orthogonal decomposition of the two dimensional vector v using the orthonormal basis $\{\phi_x, \phi_y\}$.

In the general case, the decomposition of a function $f(x)$ using an orthonormal basis $\{\phi_k(x)\}_{k \in \mathbb{Z}} \in L^2(\mathbf{C})$ is done as follows

$$f(x) = \sum_{k = -\infty}^{\infty} c_k \phi_k(x) \tag{4.7}$$

where c_k are the coefficients in the decomposition. In a Hilbert space the coefficients are found using the inner product in Eq. 4.4 which results in the coefficients c_k being

$$c_k = \langle f, \phi_k \rangle = \int_{-\infty}^{\infty} f(x)\overline{\phi_k(x)}dx \tag{4.8}$$

Therefore, given the orthonormal basis ϕ_k any function in a Hilbert space can be decomposed, or transformed, into the set of coefficients c_k. An important fact about the decomposition of a signal with an orthonormal basis is that it defines the signal *completely* and that the decomposition into the coefficients c_k is *unique*. The Fourier transform, wavelet transform, and wavelet packet transform are examples of signal decompositions

on different orthogonal bases. These decompositions will be discussed further later in this chapter.

4.1.3   Linear and Nonlinear Approximations

In some applications, such as signal compression, it is often desired to produce approximations of a signal as opposed to a full representation of the signal. This ability is important in areas such as image compression since it can allow for effective lossy image compression with high compression ratios. As seen in a Hilbert space, a signal f can be decomposed using an orthonormal basis $\{\phi_k\}_{k \in N} \in L^2(\mathbf{C})$ as follows

$$f = \sum_{k=0}^{\infty} \langle f, \phi_k \rangle \phi_k \tag{4.9}$$

Many forms of approximations exist for the signal f when inspecting Eq. 4.9. One such approximation would be to take only the first A components in the decomposition of f to form the approximation $f_A \approx f$ as follows

$$f \approx f_A = \sum_{k=0}^{A-1} \langle f, \phi_k \rangle \phi_k \tag{4.10}$$

This leaves a mean squared error in the representation of f by f_A of

$$\|f - f_A\|^2 = \sum_{k=A}^{\infty} |\langle f, \phi_k \rangle|^2 \tag{4.11}$$

With this *a priori* selection of coefficients to represent f_A, a linear approximation of the signal is obtained [Mall99].

While this approximation is often good, it can often be improved by selecting the set of orthonormal vectors to use from the basis depending on the signal being represented in an *a posteriori* manner. In general, the following approximation of f by f_A can be made

$$f_A = \sum_{k \in I_K} \langle f, \phi_k \rangle \phi_k \tag{4.12}$$

where I_K is the set of indices for the selected orthonormal basis vectors ϕ_k dependent on the characteristics of the signal f. The problem then reduces to how to select which of the orthonormal basis vectors ϕ_k to choose. This issue will be discussed further in the context of wavelets and wavelet packets later in this thesis.

4.2   Fourier Series and Fourier Transform

Before discussing wavelets and the wavelet transform, it is instructive to review the well known and well studied Fourier series and Fourier transform, which have many similarities to their wavelet counterparts. This foundation will give a basis for a discussion on wavelets and the wavelet transform for anyone familiar with the Fourier transform, and will also help in understanding the advantages of wavelets in image compression. This knowledge will be particularly important when considering the differences between Fourier transform based image compression, such as JPEG which is one of the better standards of the day, and wavelet transform based image compression.

The Fourier series is an expansion of a periodic function on a possibly infinite set of orthonormal sine and cosine functions (or the exponential function in the complex case). The Fourier transform takes the Fourier series expansion and develops forward and

inverse transforms for a signal from the time/space domain to the frequency domain. The Fourier series pair and the Fourier transform pair are defined as follows.

Definition 4.8: Let $\omega_0 = 2\pi/T_0$ be the fundamental frequency of a periodic finite energy signal $f(t) \in L^2(\mathbf{C})$ with period T_0. The *Fourier series* for the function $f(t)$ is defined as follows

$$f(t) = \sum_{k=-\infty}^{\infty} c_k e^{jk\omega_0 t} \tag{4.13}$$

where c_k are the *Fourier series coefficients* which are defined as

$$c_k = \langle f, e^{jk\omega_0 t} \rangle = \frac{1}{T_0} \int_{T_0} f(t) e^{-jk\omega_0 t} dt. \tag{4.14}$$

❑

Definition 4.9: Let $f(t) \in L^2(\mathbf{C})$ be a periodic finite energy signal, t be an independent variable in the time domain, and ω be an independent variable in the frequency domain. The *Fourier transform* $F(\omega)$ of the function $f(t)$ is then defined through to following transformation

$$F(\omega) = \int_{-\infty}^{\infty} f(t) e^{-j\omega t} dt \tag{4.15}$$

Also, let $F[\cdot]$ be an operator denoting the Fourier transform. ❑

Definition 4.10: Let $F(\omega) \in L^2(\mathbf{C})$ be a periodic finite energy signal, t be an independent variable in the time domain, and ω be an independent variable in the frequency

domain. The *inverse Fourier transform* $f(t)$ of the function $F(\omega)$ is then defined through

to following transformation

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\omega)e^{j\omega t}d\omega.$$

(4.16)

Also, let $F^{-1}[\cdot]$ be an operator denoting the inverse Fourier transform.   ❑
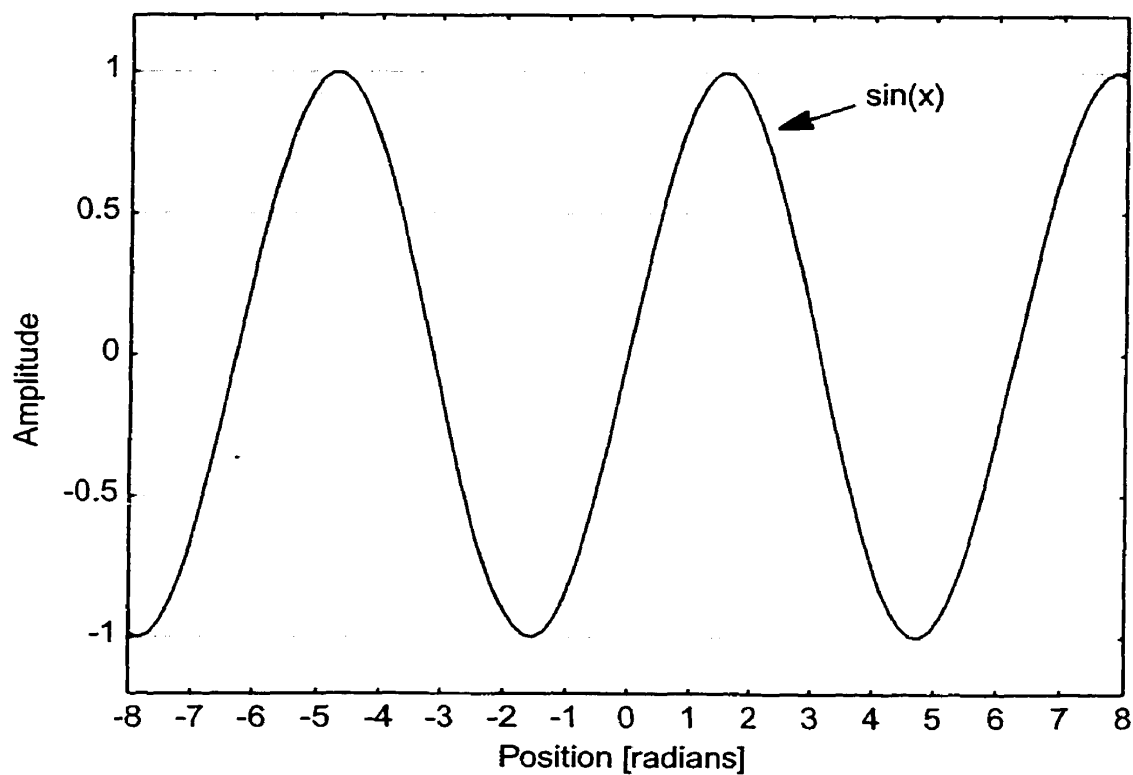
Fig. 4.2. Example basis function in Fourier analysis.

While not a wavelet transform, the Fourier transform serves to outline some of the

key issues, advantages, and disadvantages with this class of mathematical transforms. The

Fourier transform is an analysis and representation tool which models a function/signal as

the sum of scaled and dilated trigonometric functions (*i.e.* sine and cosine functions). A sample basis function in Fourier analysis, the sine wave, is plotted in Fig. 4.2.
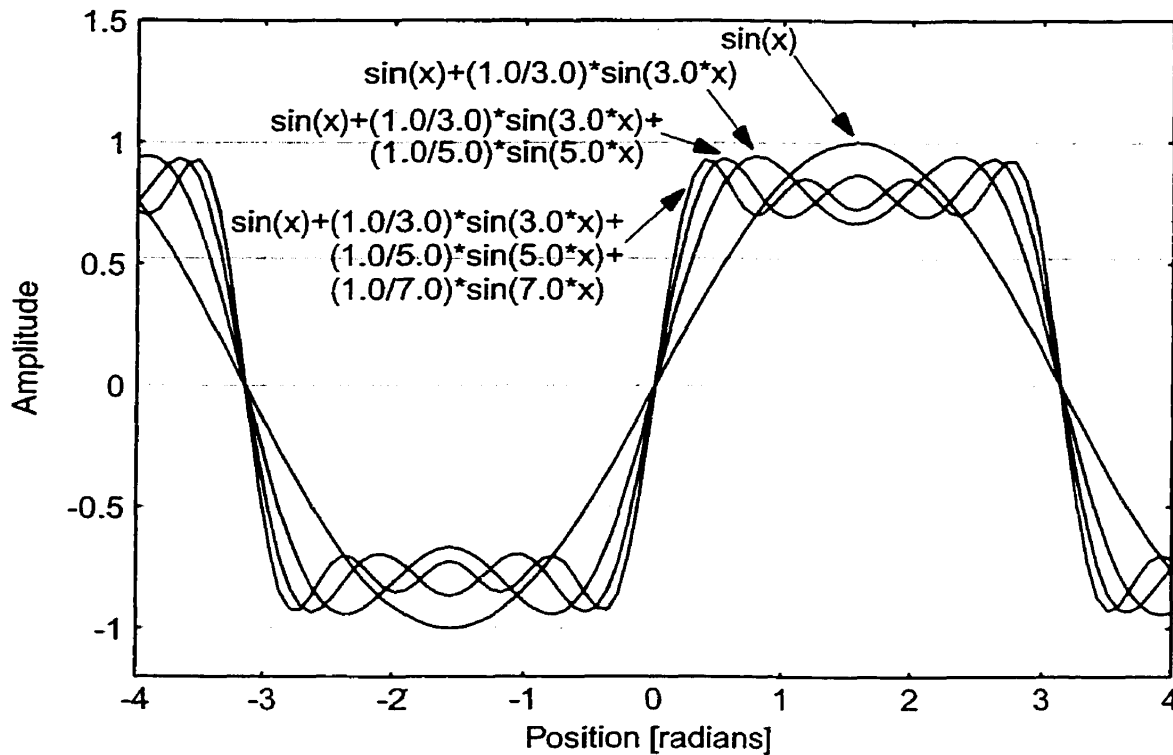
Fig. 4.3. Approximations of a square wave using Fourier decomposition
with the first 4 lowest frequency components.

As mentioned, the power behind Fourier analysis and the Fourier transform is the ability to decompose a signal into scaled and phase shifted versions of the basis function. To illustrate this, Fig. 4.3 and Fig. 4.4 demonstrate how a square wave is approximated by sinusoids. Illustrated in Fig. 4.3 is the first four approximations of a square wave using Fourier analysis. The first wave is a sinusoid, in this case $\sin x$. The second wave is a superposition of two sinusoids, namely $\sin x + \frac{1}{3}\sin 3x$. The third and fourth approximations to this square wave are $\sin x + \frac{1}{3}\sin 3x + \frac{1}{5}\sin 5x$ and

$\sin x + \frac{1}{3}\sin 3x + \frac{1}{5}\sin 5x + \frac{1}{7}\sin x$, respectively. With the first 50 of these sinusoids, as shown in Fig. 4.4, the approximation to the square wave improves even further. If this process is continued to infinity, then a perfect square wave is obtained.
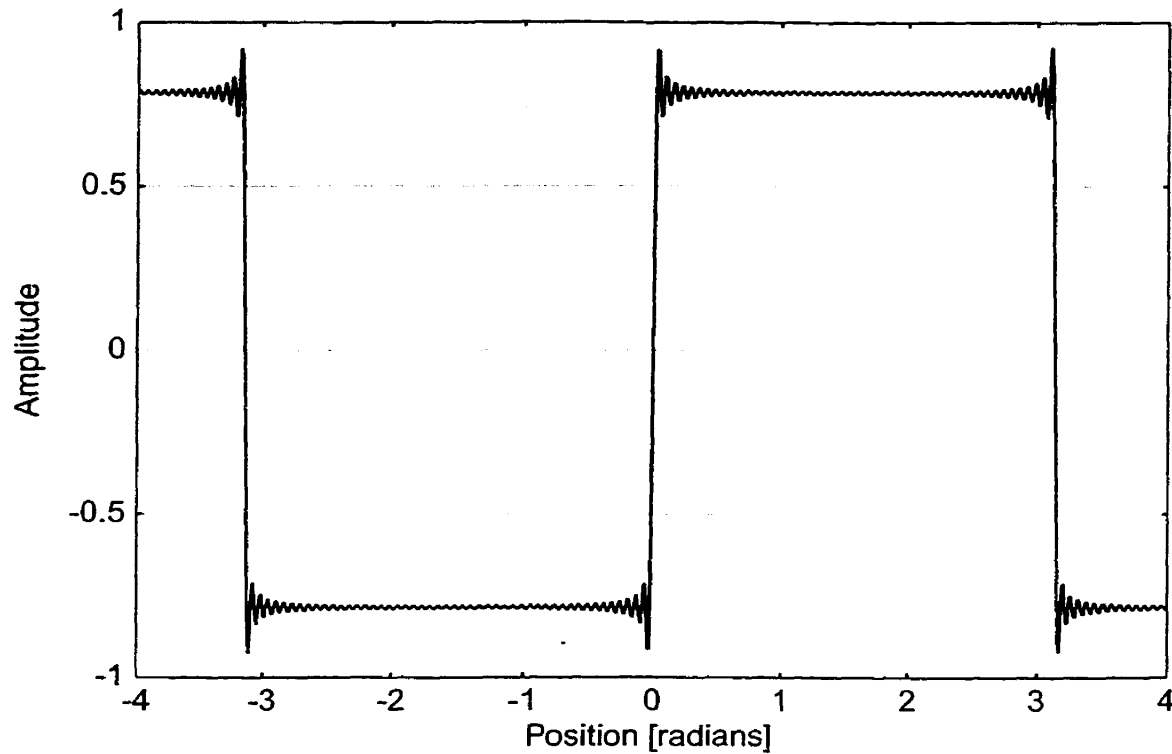
Fig. 4.4. Approximation of a square wave using Fourier decomposition
with the first 50 lowest frequency components.

One of the problems with using the Fourier transform is that the sinusoids used to decompose the signal span the entire spatial domain (*i.e.* from $-\infty$ to ∞). In image representation this basis span results in no localization of image features in the spatial domain even though localization is precisely achieved in the frequency domain. This limitation is seen by the infinite number of sinusoids required to represent a signal as simple as a square pulse, while only one Fourier basis function is required to represent a sine wave.

The question then is whether basis functions other than sinusoids can be used in the decomposition of signals. The next subsection delves into the wavelet transform which uses a class of basis functions that do allow for good spatial localization in addition to frequency localization.

4.3   Wavelets and the Wavelet Transform

A number of different approaches exist for image compression and progressive image compression. Some of the most promising approaches incorporate the wavelet transform at the principal signal processing level. This section provides a brief introduction to some of the ideas behind wavelets and the wavelet transform, as well as some of the reasons that wavelets are of interest to signal compression researchers. This section also outlines some of the ideas behind multiresolution analysis and how wavelets can decompose a signal for multiresolution analysis.

4.3.1   The Wavelet Transform

Similar to the Fourier transform, the wavelet transform is a mathematical transformation or decomposition of a signal with a set of orthogonal basis functions. First explicitly introduced by Grossman and Morlet [GrMo84], the wavelet transform decomposes a signal using mother wavelet functions as the analyzing function instead of sinusoids as with the Fourier transform. These mother wavelet functions will be denoted as $\psi(t)$. To

form the set of wavelet basis functions, the mother wavelet $\psi(t) \in L^2(\mathbf{C})$ is scaled and translated as follows

$$\psi_{u,s}(t) = \frac{1}{\sqrt{s}}\psi\left(\frac{t-u}{s}\right) \tag{4.17}$$

where $s > 0$ is a scaling parameter and u is a translation parameter. With this scaling and translation of the mother wavelet $\psi(t)$, the following definition for the continuous wavelet transform can be made.

Definition 4.11: Let $\psi(t) \in L^2(\mathbf{C})$ be a mother wavelet function, $s > 0$ be a scaling parameter for $\psi(t)$, u be a translation parameter for $\psi(t)$, and $\psi_{u,s}(t)$ be the scaled and translated version of $\psi(t)$. The *continuous wavelet transform* (CWT) of a function $f(t) \in L^2(\mathbf{R})$ is defined as

$$W_\psi f(u, s) = \langle f, \psi_{u,s} \rangle = \int_{-\infty}^{\infty} f(t)\overline{\psi_{u,s}(t)}dt = \frac{1}{\sqrt{|s|}}\int_{-\infty}^{\infty} f(t)\psi\left(\frac{t-u}{s}\right)dt \tag{4.18}$$

Also, let $W[\cdot]$ be an operator denoting the continuous wavelet transform. $\qquad\square$

Note that the CWT in Eq. 4.18 is an inner product of the form of Eq. 4.4. Therefore, the set of scaled and translated mother wavelets $\{\psi_{u,s}(t)\} \in L^2(\mathbf{C})$ forms the signal basis in the CWT. The inverse CWT is defined as follows [Mall99].

Definition 4.12: Let $\psi(t) \in L^2(\mathbf{C})$ be a mother wavelet function, $s > 0$ be a scaling parameter for $\psi(t)$, u be a translation parameter for $\psi(t)$, and $\psi_{u,s}(t)$ be the scaled and

translated version of $\psi(t)$. If $f \in L^2(\mathbf{R})$, then the *inverse continuous wavelet transform* (ICWT) is defined as

$$f(t) = \frac{2}{C_\psi} \text{Real} \left[\int_0^\infty \int_{-\infty}^\infty \frac{1}{s^2} W_\psi f(u, s) \psi_{u, s}(t) du ds \right] \tag{4.19}$$

where C_ψ is a constant which is defined as

$$C_\psi = \int_{-\infty}^\infty \frac{|F_\omega[\psi(t)]|^2}{|\omega|} d\omega = \int_{-\infty}^\infty \frac{|\Psi(\omega)|^2}{|\omega|} d\omega \tag{4.20}$$

with $F_\omega[\psi(t)] = \Psi(\omega)$ the Fourier transform of $\psi(t)$. Also, let $W^{-1}[\cdot]$ be an operator denoting the inverse continuous wavelet transform.   ❏

In looking at the forward and inverse continuous wavelet transforms, some important properties behind $\psi(t)$ can be noted. One of the most important properties for $\psi(t)$ follows from the C_ψ factor in Eq. 4.19. For a perfect reconstruction of $f(t)$ to occur, the C_ψ factor must be a constant and also finite such that

$$C_\psi = \int_{-\infty}^\infty \frac{|\Psi(\omega)|^2}{|\omega|} d\omega < \infty \tag{4.21}$$

Therefore, if reconstruction of the signal is desired after decomposition, $\psi(t)$ must be restricted to satisfy Eq. 4.21 which is referred to as the *admissibility condition*. This equation indicates that the square of the Fourier transform $\Psi(\omega)$ of $\psi(t)$ must decay faster than the factor $|\omega|$ at $\pm\infty$ for the inequality to hold. This result indicates that the mother wavelet $\psi(t)$ must be *bandlimited* in frequency. For purposes of signal analysis, a bandlimited signal is useful when analyzing the signal in the frequency domain. In signal anal-

ysis this means that a wavelet transform can decompose a signal into constituent

frequency components from the bandlimited mother wavelet.

Another important point about the admissibility condition is that $\Psi(\omega)$ must be

zero when $\omega = 0$, otherwise the inequality in Eq. 4.21 does not hold. Expanding this

condition with the Fourier transform of $\psi(t)$ itself, the following equation can be stated

$$\Psi(\omega = 0) = F[\psi(t)]|_{\omega = 0} = \int_{-\infty}^{\infty} \psi(t)e^{-j(0)t}dt = \int_{-\infty}^{\infty} \psi(t)dt = 0 \qquad (4.22)$$

which shows that the following must also be true of $\psi(t)$ for the admissibility condition to

hold

$$\int_{-\infty}^{\infty} \psi(t)dt = 0 \qquad (4.23)$$

The significance of Eq. 4.23 is that it indicates that $\psi(t)$ must be a *zero mean* function.

Another important property that results from the admissibility condition is that the

wavelet function $\psi(t)$ must have finite energy for reconstruction, so following Lemma 4.1

and Defn. 4.1, the following must be true

$$\int_{-\infty}^{\infty} |\psi(t)|^2 dt < \infty \qquad (4.24)$$

Combining the properties from Eq. 4.23 and Eq. 4.24, it is noticed that $\psi(t)$ has

finite span. This characteristic is unlike sinusoids in Fourier analysis since sinusoids have

infinite span and are not square integrable, so Eq. 4.24 does not hold. This property of

$\psi(t)$ means that wavelet analysis leads to spatial localization unlike Fourier analysis

which has no spatial localization. Combining the two localization features of wavelets, namely frequency localization and spatial localization, it should become clear that wavelets have advantages of representing non-stationary signals over what Fourier analysis can achieve. Specific non-stationary portions of a signal can be analyzed apart from the rest of the signal. From these frequency and spatial localization properties of $\psi(t)$, as well as $\psi(t)$ having zero mean and finite energy, it should be realized that $\psi(t)$ is likely a small oscillating curve dampened at $\pm\infty$. Hence, $\psi(t)$ is a small wave or a *wavelet* as it has become known.

Wavelet functions are also often designed with specific properties depending on the task. One important property is the number of vanishing moments that a wavelet possesses, where vanishing moments are defined as follows.

Definition 4.13: Let $\psi(t) \in L^2(\mathbf{C})$ be a wavelet function. Then $\psi(t)$ is said to have *m vanishing moments* if

$$\int_{-\infty}^{\infty} t^p \psi(t) dt = 0 \qquad 0 \le p < m. \tag{4.25}$$

❏

An important result of Eq. 4.25 is that a wavelet with m vanishing moments is orthogonal to any polynomial of degree $m - 1$ and is a multiscale differential operator of order m [Mall99].

A final useful property for $\psi(t)$ to achieve *good* spatial and frequency analysis is that $\psi(t)$ should be well localized in the spatial domain and $\Psi(\omega)$ should be well local-

ized in the frequency domain. Following this idea of localization, Daubechies developed a number of *compact support* wavelet functions [Daub90], [Daub92]. A compact support wavelet is one that minimizes the span of the wavelet function in both spatial and frequency domains simultaneously. By doing this minimization, signals can be decomposed into sparse sets of coefficients which helps in approximating a non-stationary signal by fewer coefficients; the primary goal of image compression. An example of a wavelet, the well known *Daubechies 4-tap wavelet* (DAUB4) [Daub90], [Daub92] is plotted in Fig. 4.5.
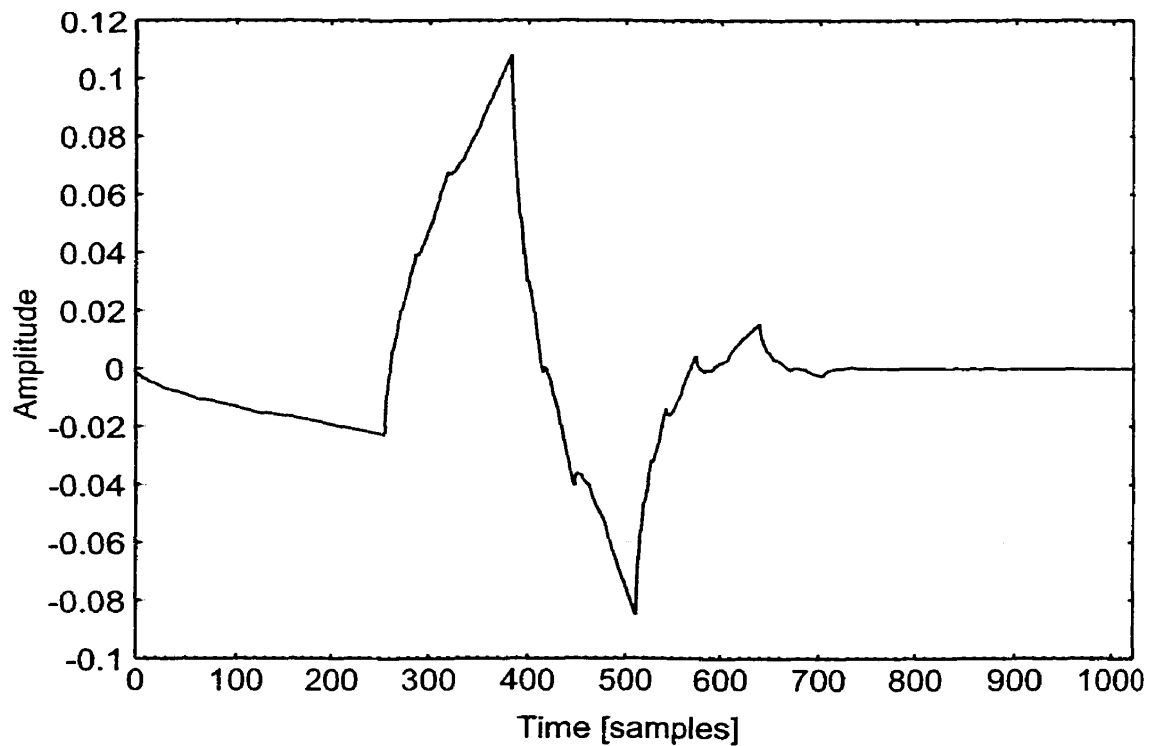
Fig. 4.5. DAUB4 wavelet for a unit vector in the 5th component of a vector of length 1024 (i.e. DAUB4 e_5).

As with all wavelets, the DAUB4 wavelet can serve as the basis function for decomposing signals. As a small example, Fig. 4.6a illustrates two scalings/translations of the DAUB4 wavelet. The first is the DAUB4 e_5 wavelet, also given in Fig. 4.5, over a vector with 1024 samples. When a wavelet transform using the DAUB4 basis function is applied to this signal it produces a vector with 1024 samples which are all zero except for the 5th component which has a unit impulse (hence the e_5). Similarly, the second wavelet which is the DAUB4 e_{40} wavelet over a vector with 1024 samples produces a unit impulse in the 40th component when transformed. When these two signals are added in the spatial domain, as shown in Fig. 4.6b, the resulting signal has a wavelet transform with unit vectors in the 5th and 40th components.

This discussion demonstrates how wavelets work as basis functions for signal decomposition through the scaling and translation of the wavelets. Seemingly complex looking signals can often be decomposed into a sparse set of wavelet coefficients when the right basis function $\psi(t)$ is selected. The whole process is very similar to the scaling and translation of sinusoid functions as with the Fourier transform. Unlike the Fourier transform, the wavelet transform is not restricted to one basis function. There are actually an infinite number of basis functions as long as they meet the properties outlined. Selection of the mother wavelet is entirely dependent on the application and results will vary depending on which mother wavelet is chosen.
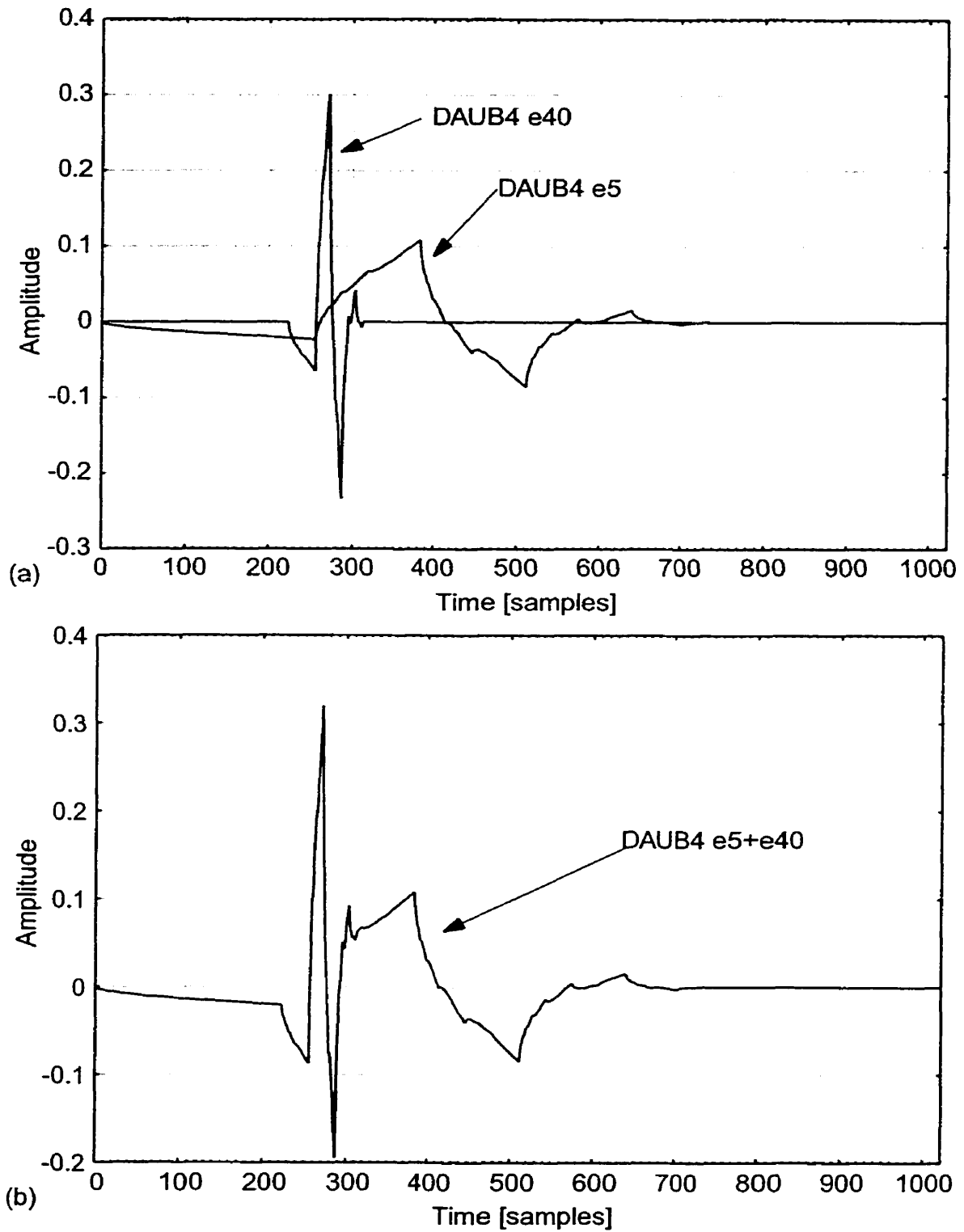
Fig. 4.6. Sample wavelet representation.
(a) two wavelets that make up time domain signal and (b) time domain signal.

4.3.2   Discrete Wavelet Transform Through Filter Banks

With the discrete nature of the array image representation discussed in Chapter 2, it is convenient to have a discrete version of the CWT presented in the previous section. Mallat describes a simple algorithm for doing a *discrete wavelet transform* (DWT) through multiresolution analysis using filter banks [Mall89].

Before getting into the details of the DWT, note the following definitions for downsampling and upsampling of a signal. These terms are used when developing approximation and detail signals in the DWT.

Definition 4.14: (after [GoCh99]) Let $\{x(n)\}$ be an input sequence. Then an M-point *decimation* of $\{x(n)\}$ is given by

$$y(n) = x(nM) \qquad \text{for } n \in \mathbf{Z} \tag{4.26}$$

This concept is also referred to as *downsampling* and can be represented diagramatically as $\boxed{\downarrow M}$. ∎

Definition 4.15: (after [GoCh99]) Let $\{y(n)\}$ be an input sequence. Then an M-point *interpolation* of $\{y(n)\}$ is given by

$$x'(n) = \begin{cases} y\left(\dfrac{n}{M}\right) & \text{for } n = kM, k \in \mathbf{Z} \\ 0 & \text{otherwise} \end{cases} \tag{4.27}$$

This concept is also referred to as *upsampling* and can be represented diagramatically as $\boxed{\uparrow M}$. ∎

It is clear from Defn. 4.14 that an input sequence is taken and every Mth element is extracted to perform an M-point decimation. With Defn. 4.15 an input sequence is taken and expanded by adding $M-1$ zeros between the M elements to perform the interpolation.

The general procedure behind Mallat's DWT algorithm [Mall89] is to decompose the discrete signal into an approximation signal $\mathbf{A}_s$ and a detail signal $\mathbf{W}_s$, where s is the scale in the multiresolution analysis. The approximation signal $\mathbf{A}_s$, or lowpass signal, is formed by the projection of the original signal onto the space formed by the basis $\{\phi_{k,s}:2^{s/2}\phi(2^s t - k);k \in \mathbf{Z}\}$ where $\phi(t)$ is a scaling function. The detail signal $\mathbf{W}_s$, or highpass signal, is similarly formed with the basis $\{\psi_{k,s}:2^{s/2}\psi(2^s t - k);k \in \mathbf{Z}\}$ where $\psi(t)$ is a mother wavelet. Both of these bases are effectively scaled versions of $\phi(t)$ and $\psi(t)$ that have been downsampled by a 2^s-point decimation to a lower resolution.

A few properties of $\phi(t)$ and $\psi(t)$ exist. Firstly, $\psi(t)$ must follow all of the previously discussed properties of a mother wavelet for this to be a DWT. An important property of $\phi(t)$ is that it must be selected so that approximations $\mathbf{A}_s$ are subsets of higher resolutions approximations $\mathbf{A}_{s+k}$, $(k>0) \in \mathbf{N}$, as follows

$$\{0\} \leftarrow ... \leftarrow \mathbf{A}_{-1} \subset \mathbf{A}_0 \subset \mathbf{A}_1 \subset ... \subset L^2 \tag{4.28}$$

More importantly, an approximation at one scale should be made up of the lower level approximation and detail signals as follows

$$\mathbf{A}_{s+1} = \mathbf{A}_s \otimes \mathbf{W}_s \qquad s \in \mathbf{Z} \tag{4.29}$$

where $\otimes$ is an operator that combines the approximation and detail signal through the inverse DWT. In addition, the approximation and detail signals at a specific resolution should not contain overlapping components such that

$$A_s \cap W_s = \{0\} \qquad s \in Z \tag{4.30}$$

It should also be noticed that if A_s is the original signal, then the decomposition into wavelet detail signals at various resolutions is expressed as follows

$$A_s = W_{s-1} \otimes A_{s-1} = W_{s-1} \otimes (W_{s-2} \otimes A_{s-2})$$

$$= W_{s-1} \otimes W_{s-2} \otimes W_{s-3} \otimes \cdots \tag{4.31}$$

The algorithm that Mallat proposed actually uses filter banks G and H to filter the signal as $\phi(t)$ and $\psi(t)$ would form A_s and W_s at each scale s. To get the required relationships as in Eq. 4.29, Eq. 4.30, and Eq. 4.31, Mallat showed that when a filter bank for $\psi(t)$ is developed that the following impulse response relationship between the two filter banks holds

$$g(n) = (-1)^{1-n} h(1-n) \tag{4.32}$$

where G is the mirror of H and the pair are known as *quadrature mirror* filters [Mall89]. Hence, knowing the mother wavelet $\psi(t)$, then the scaling function $\phi(t)$ can be developed. With both filter banks in place then the DWT algorithm as depicted in Fig. 4.7 results. Following Fig. 4.7, the original discrete signal is convolved with the quadrature mirror filter bank pairs H and G, and then decimated by a factor of 2 to get the 2^s-point decimation. This procedure results in two discrete signals, A_{s-1} and W_{s-1}, which each have half the number of samples as A_s. This process is repeated on A_{s-1} to form A_{s-2} and W_{s-2}, and then again on the following approximation signals until the entire signal

has been decomposed to the wavelet terms. Decomposition can be stopped at any chosen

level, though will be limited to when the resulting approximation and detail signals have

been decomposed down to one sample each.
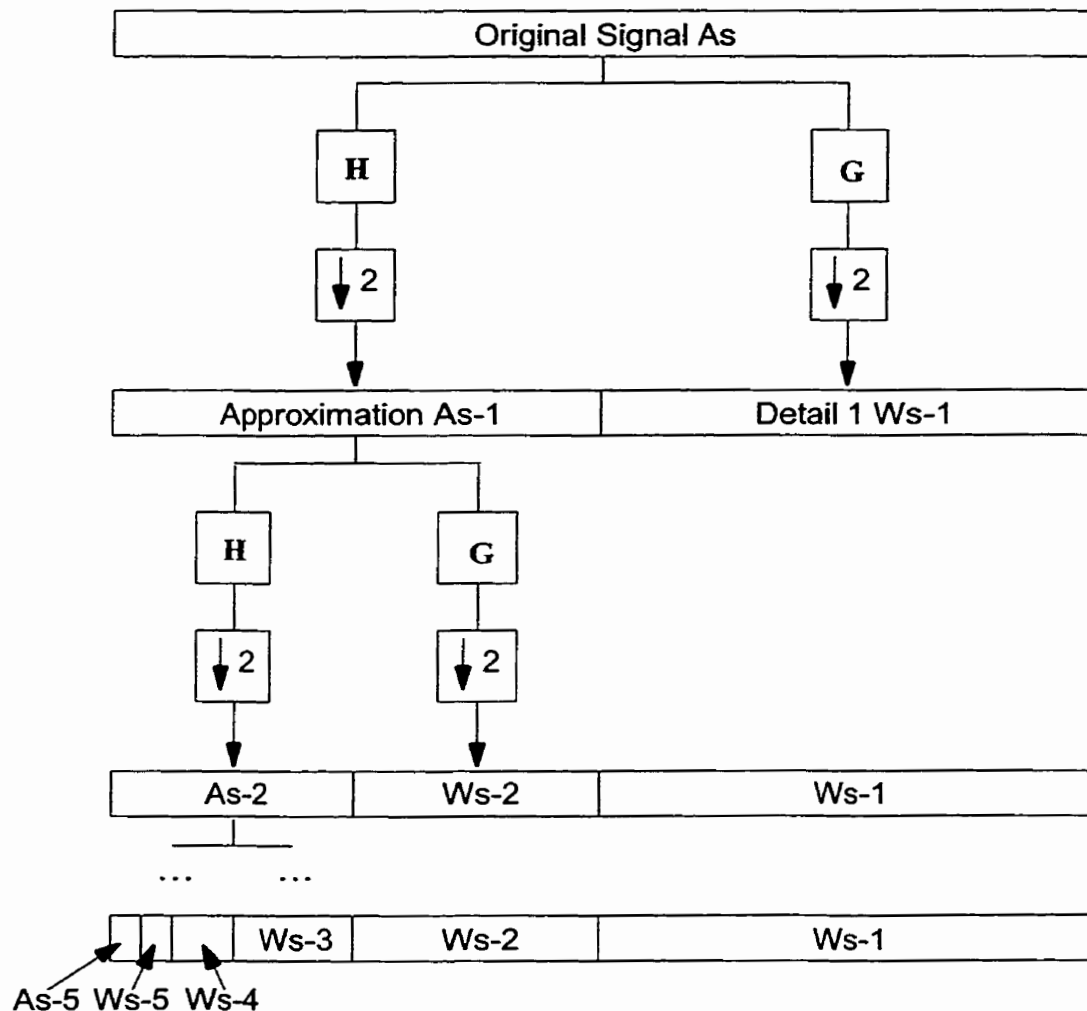
Fig. 4.7. Discrete wavelet transform using filter banks (after [Mall89]).

Reconstructing of the signal is done by reversing the process depicted in Fig. 4.7.

The approximation signals A_s and W_s are upsampled with a 2 -point interpolation. Then

the upsampled versions of A_s and W_s are convolved again with the **H** and **G** filter

banks, respectively and the resulting vectors added together to form the approximation at the next higher resolution A_{s+1}.

4.3.3   Two Dimensional Discrete Wavelet Transform

With image compression and progressive image transmission being the focus of this thesis, a two dimensional DWT algorithm is in order. For developing a 2D DWT algorithm, an important thing to note is that the 2D wavelet transform is separable [Mall99]. If $\{\psi_{u,s}(x)\}$ is an orthogonal basis then $\{\psi_{(u,s)_1}(x)\psi_{(u,s)_2}(y)\}$ is an orthogonal basis for a 2D space. Given a 2D function $f(x,y) \in L^2(\mathbf{R}^2)$, the function can be decomposed, as usual, through the following inner product

$$\langle f, \psi_{(u,s)_1}\psi_{(u,s)_2}\rangle = \int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty} f(x,y)\overline{\psi_{(u,s)_1}(x)}\,\overline{\psi_{(u,s)_2}(y)}dxdy \tag{4.33}$$

which can be rearranged to the following

$$\langle f, \psi_{(u,s)_1}\psi_{(u,s)_2}\rangle = \int\limits_{-\infty}^{\infty}\overline{\psi_{(u,s)_2}(y)}\int\limits_{-\infty}^{\infty} f(x,y)\overline{\psi_{(u,s)_1}(x)}dxdy \tag{4.34}$$

The rearrangement of the inner product in Eq. 4.34 shows that the inner product in the wavelet transform is *separable*. Using the separability of the inner product, a simple algorithm for performing the DWT on images can be developed using only the 1D DWT. This 2D transform is done by applying the DWT filter banks first on each of the rows of an image and then afterwards applying the DWT filter banks on each of the columns. This procedure decomposes the image into an approximation image, as well as horizontal, vertical, and diagonal details. This process is then repeated with the approximation portion of

the decomposition. The first three passes of the 2D DWT decomposition are illustrated in Fig. 4.8.

a - Approximation of signal
v, h, d - Detail from signal

				aa	ha			a	h		
a		h						v	d		
				va	da						
v		d									

Original Image   First Pass   Second Pass   Third Pass
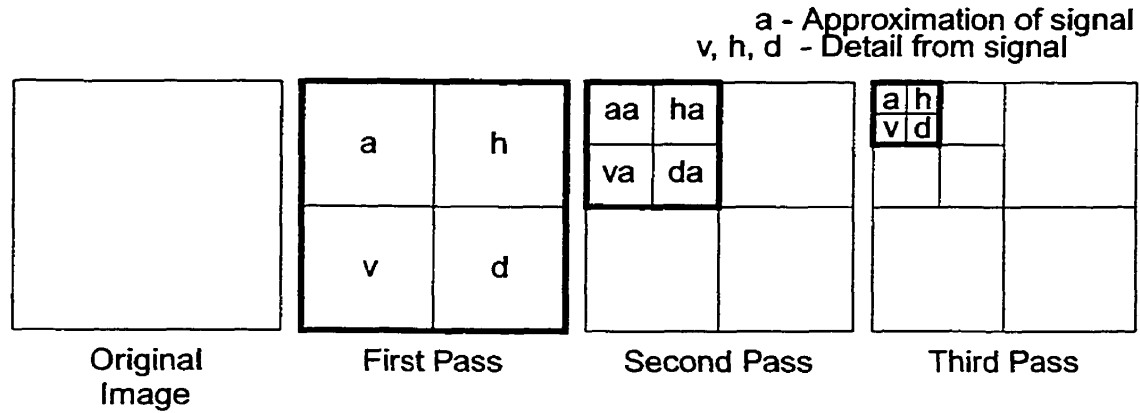
Fig. 4.8. Two dimensional wavelet transform on an image.

As Fig. 4.9 illustrates, there is a hierarchy in the 2D DWT of an image. Each detail wavelet coefficient at the highest level (at the top-left of Fig. 4.9) is the parent of four wavelet coefficients down one level. Continuing on, these four wavelet coefficients are the parents of the sixteen wavelet coefficients at the next detail level. With this form of hierarchy, a higher level view of the signal can be made by using only some of the more global detail values in the top-left quadrants of this 2D DWT. More details can be incorporated into an analysis by going down through the hierarchy. Effectively, this structure allows for an analysis of a signal at different resolutions or scales, depending what level of lower details are included. This multiresolution structure will come in useful when considering progressive image transmission techniques since approximations of the image can be transmitted first and then various details subsequently transmitted.
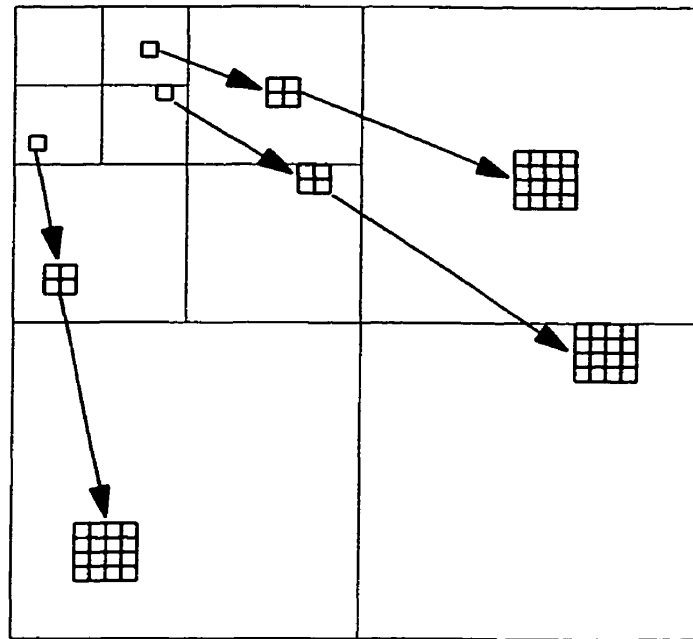
Fig. 4.9. Multiresolution analysis from the structure of the
two dimensional discrete wavelet transform.

. 4.4   Wavelet Packets and the Wavelet Packet Transform

As noticed in the previous section, the wavelet transform consists of decomposing

a signal using a mother wavelet $\psi(t)$. When performing the DWT described previously,

the mother wavelet is used to decompose a signal into consecutive octave bands in the fre-

quency domain [GoCh99]. It is sometimes desirable to obtain finer resolutions within spe-

cific octave bands, so further splitting of these bands can be done by again performing the

DWT within the band. For the DWT, this effectively results in recursively decomposing

the detail signals in addition to the already decomposed approximation signal [Wick94],

[Mall99]. This category of transformation is a generalization of the DWT and is referred

to as a wavelet packet decomposition or the wavelet packet transform, due to the further combining of extra wavelet decompositions within octave bands.

4.4.1   Discrete Wavelet Packet Transform Through Filter Banks

The *wavelet packet transform* (WPT) is very similar to the wavelet transform. Using filter banks, the *discrete wavelet packet transform* (DWPT) on a 1D signal is performed as illustrated in Fig. 4.10. This figure is similar to that of the DWT shown in Fig. 4.7 with the exception that at every pass, both the approximation signal A_s and the detail signal W_s are decomposed instead of only the approximation signal A_s as in the DWT. The packet transform works by taking the original signal, passing it through the quadrature mirror filter bank pair H and G, and then decimating these filtered signals by a factor of 2 to form the various approximation $a_{section}^{pass}$ and detail $d_{section}^{pass}$ signals for each pass and for each section of the signal. This procedure is recursively repeated for each resulting approximation $a_{section}^{pass}$ and detail $d_{section}^{pass}$ signal to form the full wavelet packet decomposition. In image compression, the important thing to realize with the wavelet packet transform is that the bulk of the energy in the wavelet coefficients might be redistributed in fewer coefficients depending on the signal. This feature may allow for better approximations of an image to be made with fewer coefficients.
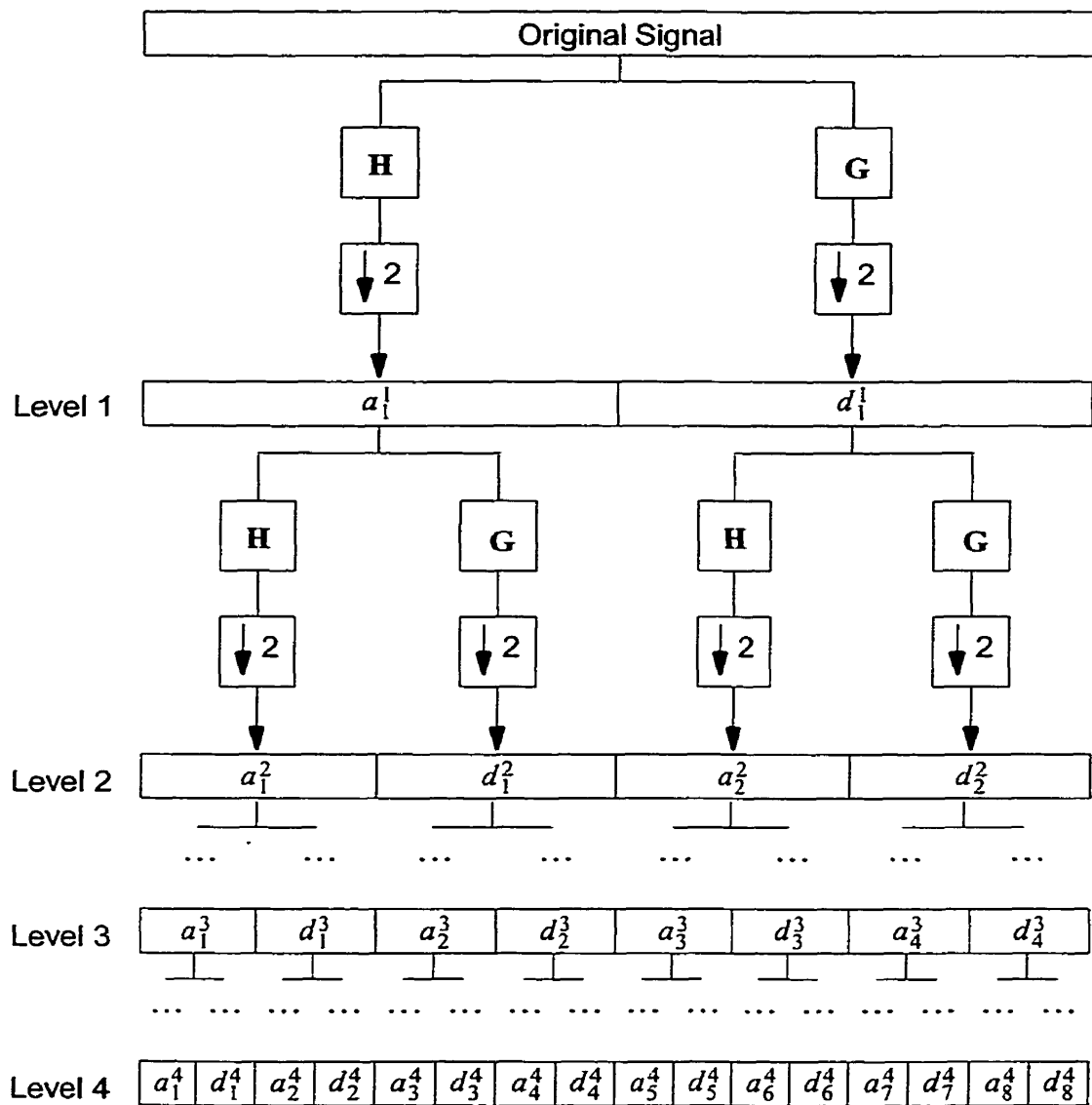
Fig. 4.10. Discrete wavelet packet transform using filter banks (after [Mall99]).

4.4.2   Adaptive Selection of Wavelet Packet Bases

While it is apparent from Fig. 4.10 that the wavelet packet transform decomposes signals more fully than the wavelet transform alone, another concept often used with

wavelet packets is the adaptive selection of the wavelet packet basis. The decomposition

shown in Fig. 4.10 shows a *full* wavelet packet transform. Often, a partial, or selective,

WPT is more useful in specific applications and is generally dependent on the signal char-

acteristics.

Full Wavelet
Packet Transform

(a)

Selective Wavelet
Packet Transform

(b)

Wavelet
Transform

(c)
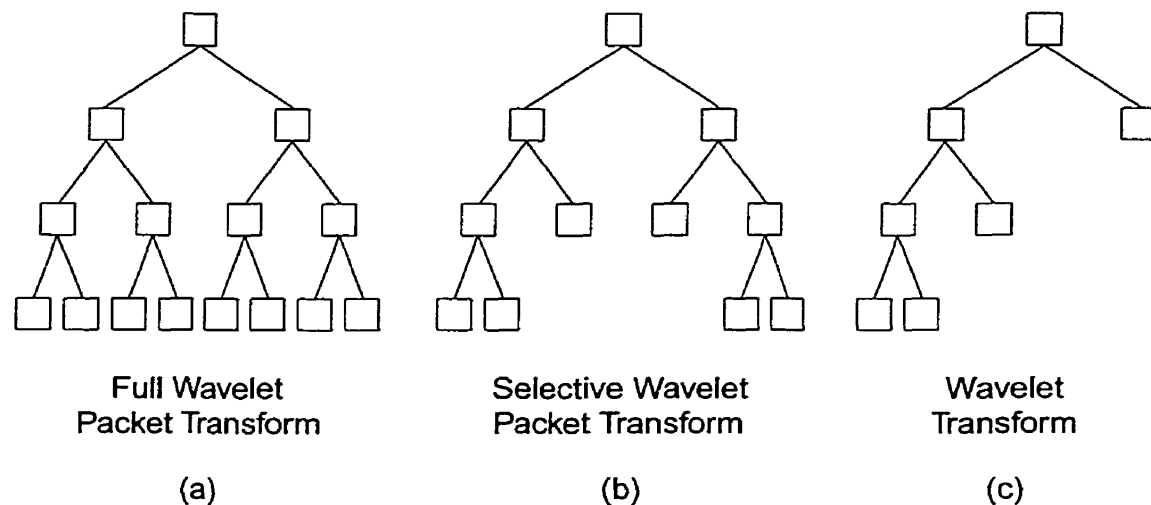
Fig. 4.11. One dimensional wavelet and wavelet packet transform decompositions.
(a) Discrete wavelet packet transform,
(b) selective discrete wavelet packet transform, and (c) discrete wavelet transform.

For instance, Fig. 4.11a illustrates the decomposition of a 1D signal using the full

DWPT as described in the previous subsection (*i.e.* a simplified illustration of Fig. 4.10).

When selecting the function basis using the DWPT, it may be advantageous to only

decompose certain approximation and detail signals to a certain level. This selective

decomposition is illustrated in Fig. 4.11b where some approximation and detail decompo-

sitions are no longer decomposed further at later levels. The stopping criteria for deciding

where to end further decomposition depends on the signal and the application, but can

include factors such as entropy measures or mean squared error measures after quantiza-

tion [CoWi92], [Wick94]. In Chapter 7 of this thesis, multifractal measures that are used

to evaluate image quality are then used to help in determining the stopping criteria for the

wavelet packet decomposition. It should also be noted that the standard DWT is actually a

specific case of the selective wavelet packet decomposition, as illustrated in Fig. 4.11c.

4.4.3   Two Dimensional Discrete Wavelet Packet Transform

The 2D DWPT used for signals such as images is very similar to that of the 2D

DWT. As illustrated in Fig. 4.12, the full wavelet packet decomposition takes the image

and in the first pass decomposes it both horizontally and vertically to obtain an approxima-

tion image, a horizontal detail image, a vertical detail image, and a diagonal detail image.

The DWPT decomposition is recursively repeated for these four smaller sections at each

subsequent pass. This process can be compared to the 2D DWT decomposition illustrated

in Fig. 4.8 where only the approximation portion of the signal is further decomposed.
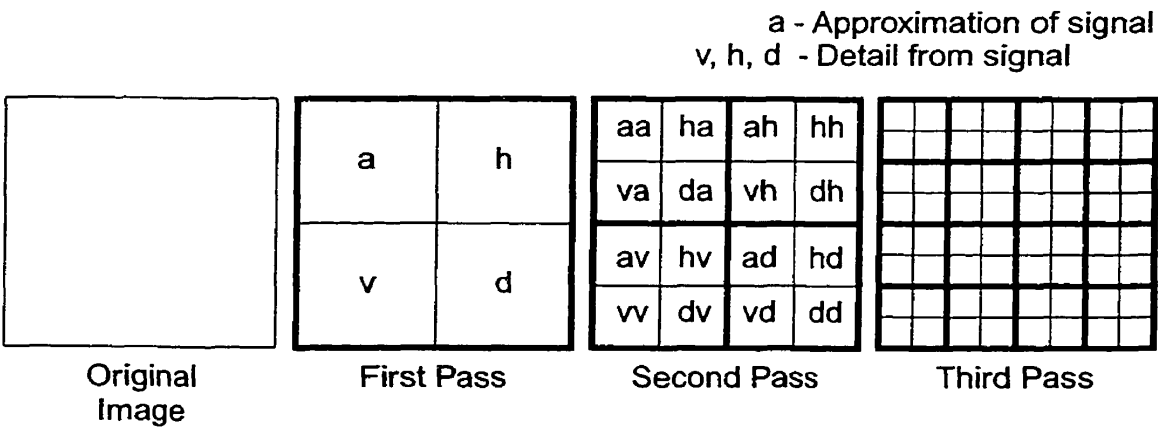
Fig. 4.12. Full 2D wavelet packet decomposition.

a - Approximation of signal
v, h, d - Detail from signal

aa	ha	ah	hh
va	da	vh	dh
		ad	hd
		vd	dd

a   h

v   d

Original Image   First Pass   Second Pass   Third Pass
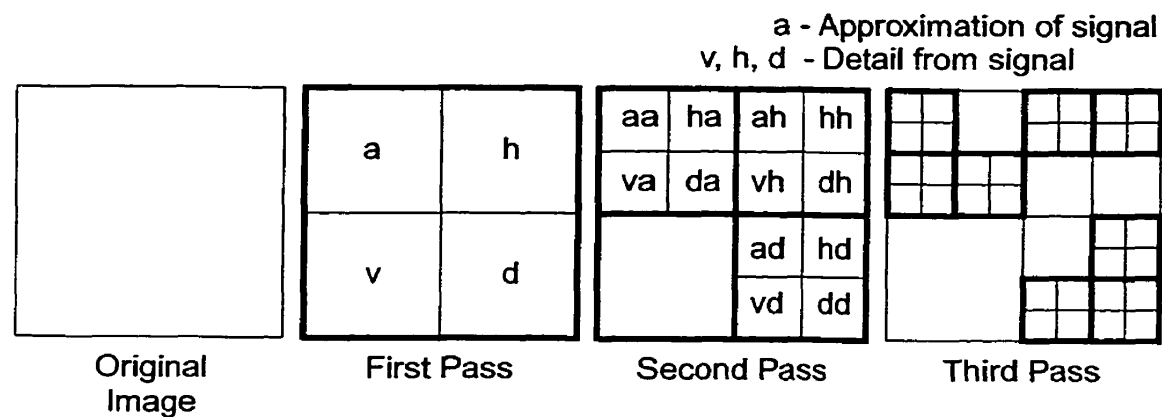
Fig. 4.13. Example selective 2D wavelet packet transform decomposition.

The process is similar for the selective 2D DWPT for a 2D signal such as an image. As illustrated in the example decomposition in Fig. 4.13, an image is taken and the DPWT is applied, yielding the approximation image, horizontal detail image, vertical detail image, and diagonal detail image. At this point it must be decided which of these four smaller sections of the signal would most benefit our application with further decomposition using the stopping criteria for the current signal. Once selected, the chosen smaller signal portions are recursively decomposed with the DWPT and this process repeated at each pass. Once completed, some portions of the decomposed image will have had more wavelet packet transforms performed than other portions of the image.

4.5   Summary

This chapter has provided the background behind wavelets and wavelet packets. This discussion included an overview of the properties behind wavelets as well as how to decompose 1D and 2D signals using wavelets and wavelet packets. The following chap-

ters describe the progressive image transmission techniques developed, as well as the mul-

tifractal image complexity measures developed. This explanation includes progressive

image transmission techniques based on fractal surface segmentation and interpolation,

with wavelet and wavelet packet decomposition. The ideas of adaptive wavelet packet

decomposition will come into play when the developed multifractal measures are used to

help in the selection of the wavelet packet bases to be used for a particular image.

CHAPTER V

PROGRESSIVE FRACTAL IMAGE COMPRESSION USING FRACTAL SURFACE SEGMENTATION AND INTERPOLATION

This chapter introduces a new and novel approach to image compression [DaKi96], [DaKi97] and progressive image transmission [DaKi98a] that has been developed for this thesis. This method is based on an interpolation scheme using fractal surfaces and will be referred to as *fractal surface segmentation and interpolation* (FSSI). Fractal surface segmentation and interpolation allows for the representation of certain texture features within an image using the fractal dimension of the texture. These texture features can then be reproduced in a fractal sense to synthesize natural looking images. Hence, the goal is not necessarily to reproduce the textures in an image from a pixel by pixel basis, but, instead to synthesize textures with similar fractal characteristics that appear psychovisually similar to the original textures. Some of the findings and realizations from this work with FSSI are used in Chapter 6 for the development of progressive image quality measures based on multifractal complexity measures. It must be noted that the FSSI techniques described in this chapter are not currently bit rate competitive with other state of the art techniques such as EZW [Shap93] or SPIHT [SaPe96], but have potential for being combined with other techniques to improve compression.

Of particular interest for this thesis is the development of progressive image transmission techniques. An important approach to performing progressive image transmission is to represent the image in a multiresolution form so that increasing levels of image

detail can be transmitted and reconstructed in the image, as discussed in Sec. 2.2.1. Of course, many methods exist for representing an image in a multiresolution form. This chapter focuses on methods which use region segmentation of an image and then interpolation in the reconstruction process to estimate unknown values located between known values. To perform this estimation of the unknown values, the simplest approach is to do a linear interpolation between the unknown values. Other attempts have been made using non-linear interpolation functions such as using B-splines [Wata97]. This chapter presents a method of using fBm in the interpolation process as an alternative with advantages for representing some classes of textures, particularly random-like textures or noise-like textures.

5.1   Interpolation and Fractal Interpolation

Interpolation is the process of estimating or predicting unknown values within a signal that are between other known values in the signal. For instance, if given the signal illustrated in Fig. 5.1 then approximations of the signal can be formed by sampling the signal at regular intervals and performing a linear interpolation as shown in Fig. 5.2.
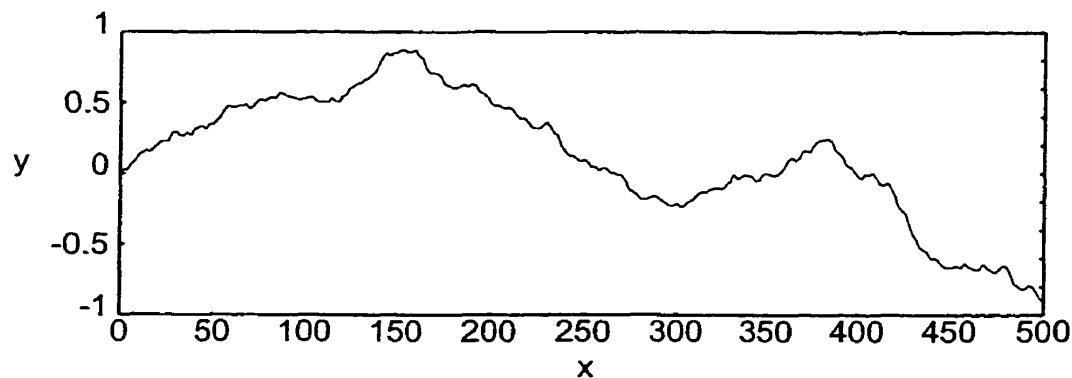
Fig. 5.1. Example signal to perform successive interpolations.
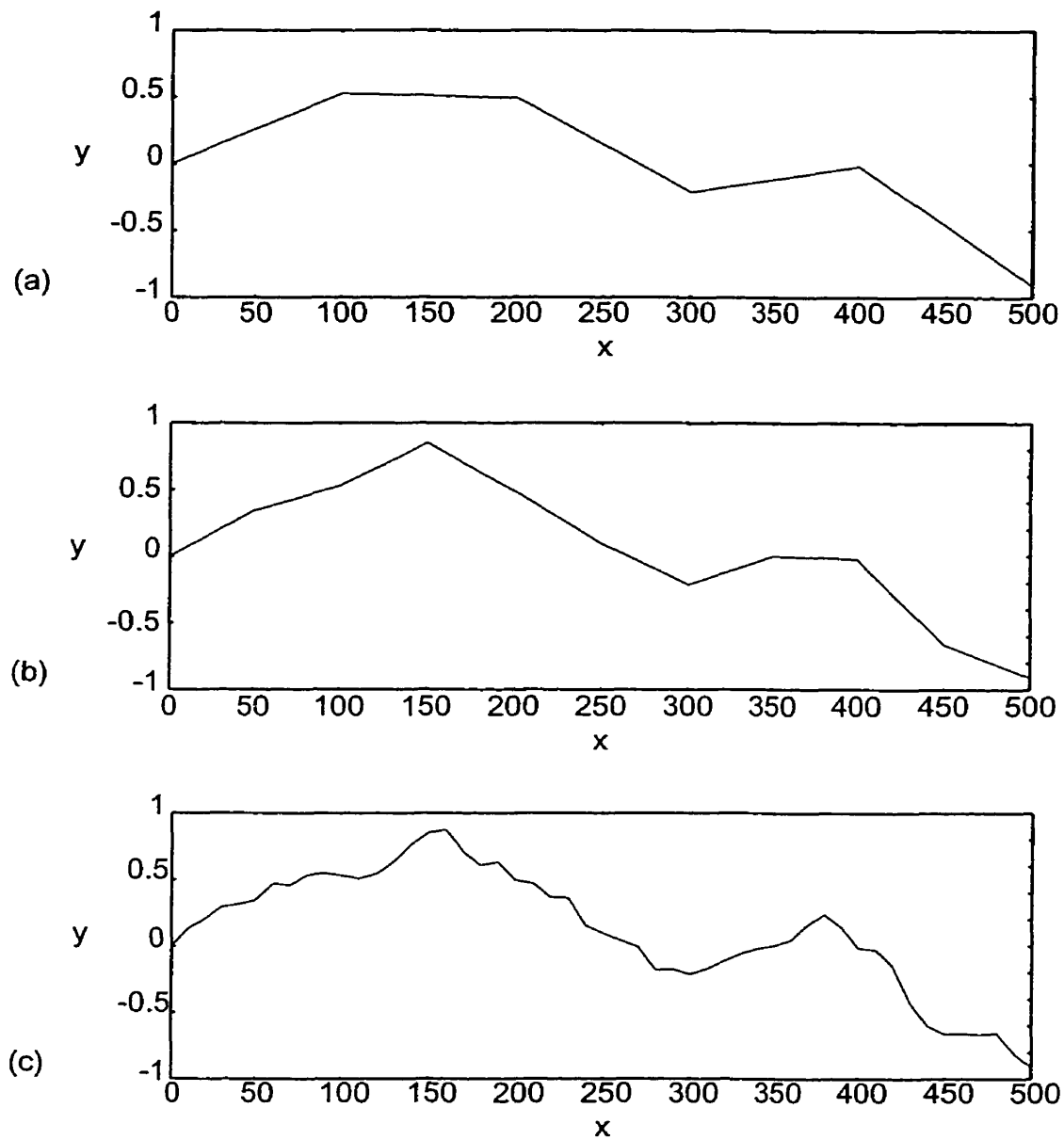
Fig. 5.2. Linear interpolation of example signal where sampling is at every:
(a) 100 points, (b) 50 points, and (c) 10 points.

From Fig. 5.2 it is seen that linear interpolation does produce increasingly better

approximations of the original signal in Fig. 5.1 as the sampling interval is decreased. A

drawback of this interpolation is perceptual accuracy is poor with only a few samples.
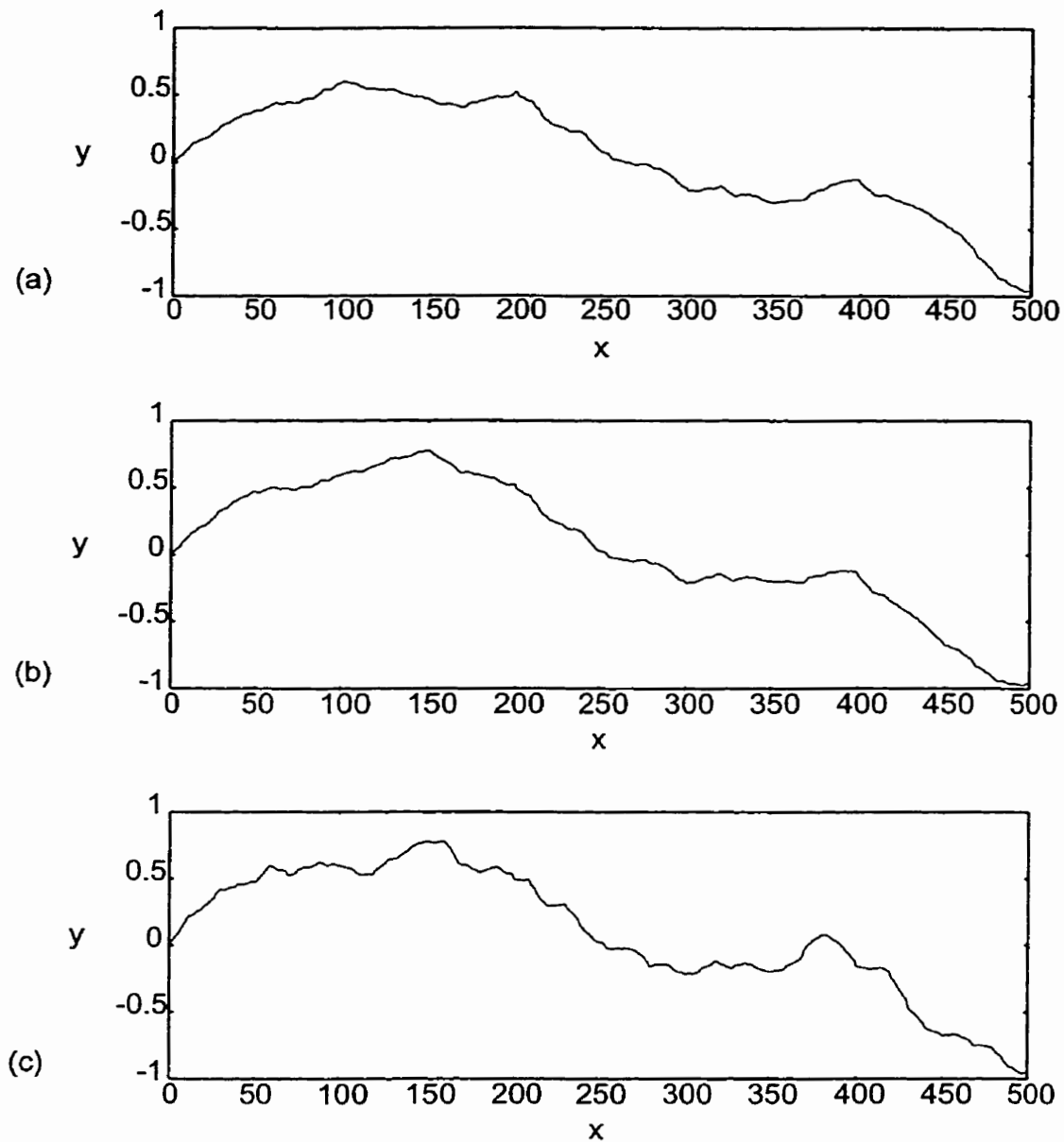
Fig. 5.3. Fractal interpolation of example signal where sampling is at every:
(a) 100 points, (b) 50 points, and (c) 10 points.

Since there are more characteristic features to the original signal in Fig. 5.1 other

than straight lines, it seems logical to extend the form of interpolation to incorporate other

features in the signal. The feature of interest for this thesis is the fractal dimension of the

signal. By calculating the fractal dimension of the signal in Fig. 5.1 and using this to gen-

erate fBm, then a different form of interpolation can be developed, namely fractal interpo-

lation. An example of this fractal interpolation is illustrated in Fig. 5.3 and can be

compared with the linear interpolation illustrated in Fig. 5.2. From a perceptual view-

point, the fractal interpolations in Fig. 5.3 better represent the original signal than the lin-

ear interpolations in Fig. 5.2. It should be noted that fractal interpolation is not limited to

signals that are coarser and rougher such as Fig. 5.1. Smoother signals with fractal dimen-

sions closer to 1 will result in fractal interpolations composed of straighter line segments

formed through the fBm. Signals closer to white noise would be interpolated with fBm

having a higher fractal dimension. Therefore, the fractal interpolation can be considered a

generalization of linear interpolation.
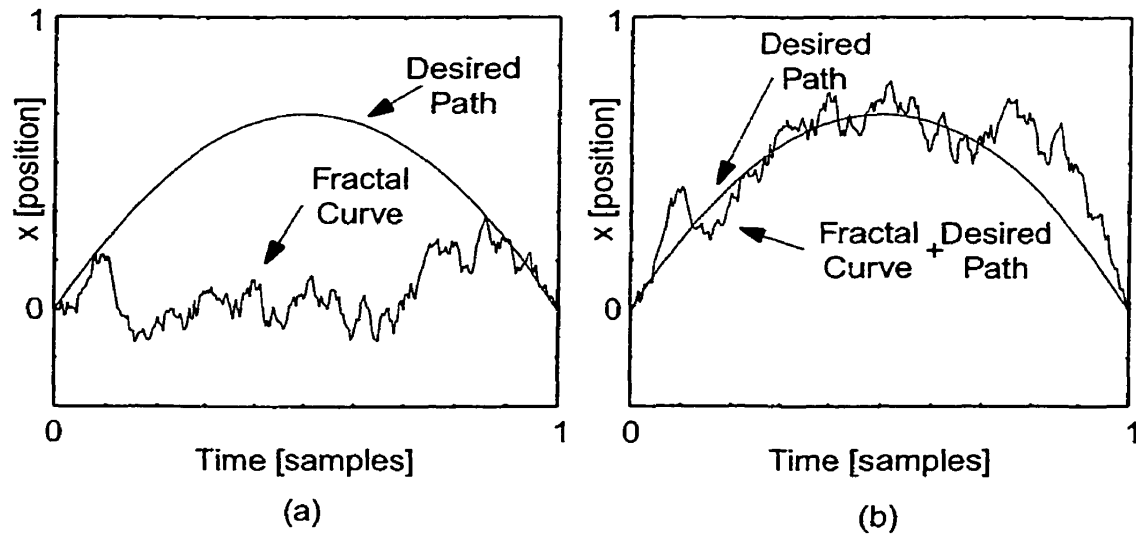
Fig. 5.4. Fractional Brownian motion interpolation along a path.
a) not along the desired path, b) superposition with the desired path.

Of course there are other interpolation functions that can be used such as B-splines, Bezier curves, and higher order polynomial curves. These types of interpolation functions can still be used in conjunction with fractal interpolation. This combined interpolation is done by using one of the interpolation curves, say B-splines, to get an approximation of the signal's global path and then superposing a fractal curve with the fractal dimension of the original signal. For instance, Fig. 5.4a illustrates two curves: a self-affine fBm fractal curve, and a desired global path for the curve. It is clear that the fractal curve, which is generated between the two endpoints of the curve, does not follow the desired path. By performing a superposition of the two curves Fig. 5.4b, is obtained where the fractal curve follows the desired path but still retains its fractal characteristics. An important feature to note is that this superposition does not alter the fractal characteristics of the fBm curve as long as the fractal dimension of the desired path does not exceed that of the fBm. This property can be seen by considering again the Hausdorff-Besicovitch fractal dimension in Eq. 3.11 from Defn. 3.1. From the Hausdorff-Besicovitch fractal dimension it should be noted that as $s \rightarrow 0$ then N_s will be dominated by the vel covering needed for the object with the highest fractal dimension. Therefore when measuring an object that is composed of more then one internal object, the fractal dimension is dominated by the internal object with the highest fractal dimension. This is the case when using any of the curves generated through B-spline interpolation, Bezier interpolation, and any high order polynomial style interpolation technique since the fractal dimension of these simpler curves is 1. Of course one problem with these calculations is that having $s \rightarrow 0$ means that the signal being measured must be of infinite resolution for the higher dimension fractal to completely dominate. With the finite resolution of the images being

measured, it was noticed that two quantitative measures can be extracted, namely the *textural fractal dimension* and the *structural fractal dimension* [CrIv89]. The textural fractal dimension measures the fBm curve in Fig. 5.4 while the structural fractal dimension measures the inverted parabola in Fig. 5.4. In a finite sense these two measures do interact, but, the textures will still dominate although not completely suppress the structure of the curve [SaCh92]. Therefore, the fractal dimension of the curve resulting from the superposition of simpler curves with the fBm curve is close to the fractal dimension for the fBm curve itself.

5.2   Generating Fractional Brownian Motion

For the fractal interpolation technique discussed in Sec. 5.1 to be useful a method of generating fractal curves is necessary. Fractional Brownian motion as discussed in Sec. 3.4.2 can be used to generate these self-affine fractal curves for fractal interpolation. Unfortunately, the main formulation of fBm given in Defn. 3.8 with Eq. 3.70 does not give a clear way of efficiently generating fBm with an arbitrary Hurst exponent H^*, and particularly for generating surfaces as is desired for images. Dealing with the stochastic integral in Eq. 3.70 is a bit cumbersome and does not immediately offer a simple approach to generating fBm curves and surfaces. To this end, a number of different techniques for generating fBm have been developed over the years. These methods include the random midpoint displacement algorithm [Carp80], [FoFC82], [Saup88]; successive random additions [Voss85]; fast fractional Gaussian noise [Mand71]; independent or random cuts algorithm [Mand82]; Fourier spectral synthesis [Voss85]; wavelet synthesis [Mall89], [Flan92]; multidimensional recursive filters [BrBa94]; and chaotic techniques [MuBa90].

In this thesis the *random midpoint displacement* (RMPD) algorithm is used for generating the desired fBm fractal curves and surfaces. Some other studies have shown that the "random midpoint displacement algorithm is in practice much simpler to implement, faster to generate, and results in a comparable accuracy" for generating fBm [StLN95].

5.2.1   Random Midpoint Displacement Algorithm in One Dimension

The procedure behind the RMPD algorithm [Carp80], [FoFC82], [Kins95a] for generating fBm embedded one dimension is illustrated in Fig. 5.5. The approach that the RMPD algorithm takes is to start with a coarse, low resolution curve and through a refinement process develop a high resolution curve containing the desired fractal characteristic. This refinement process must be carefully performed if the self-affine property of the resulting fBm curve is to be maintained over all scales of the fBm curve.

The algorithm begins, as illustrated in Fig. 5.5a, by starting with point **A** and **B'** which are located on the abscissa for the fBm fractal being generated. The point **B** is found by offsetting point **B'** perpendicular to the x-axis into the dimension that this fractal will span. This offset is determined by a Gaussian random variable, $n(x; \mu = 0, \sigma)$, with zero mean and a variance of σ^2. The line segment formed by **AB** now completes the construction of the coarsest level for the fBm.

From this point in the fBm construction, the line segment **AB** is broken into smaller pieces and then iteratively refined. The first step of this refinement is illustrated in Fig. 5.5b. As the name suggests for the random midpoint displacement algorithm, the
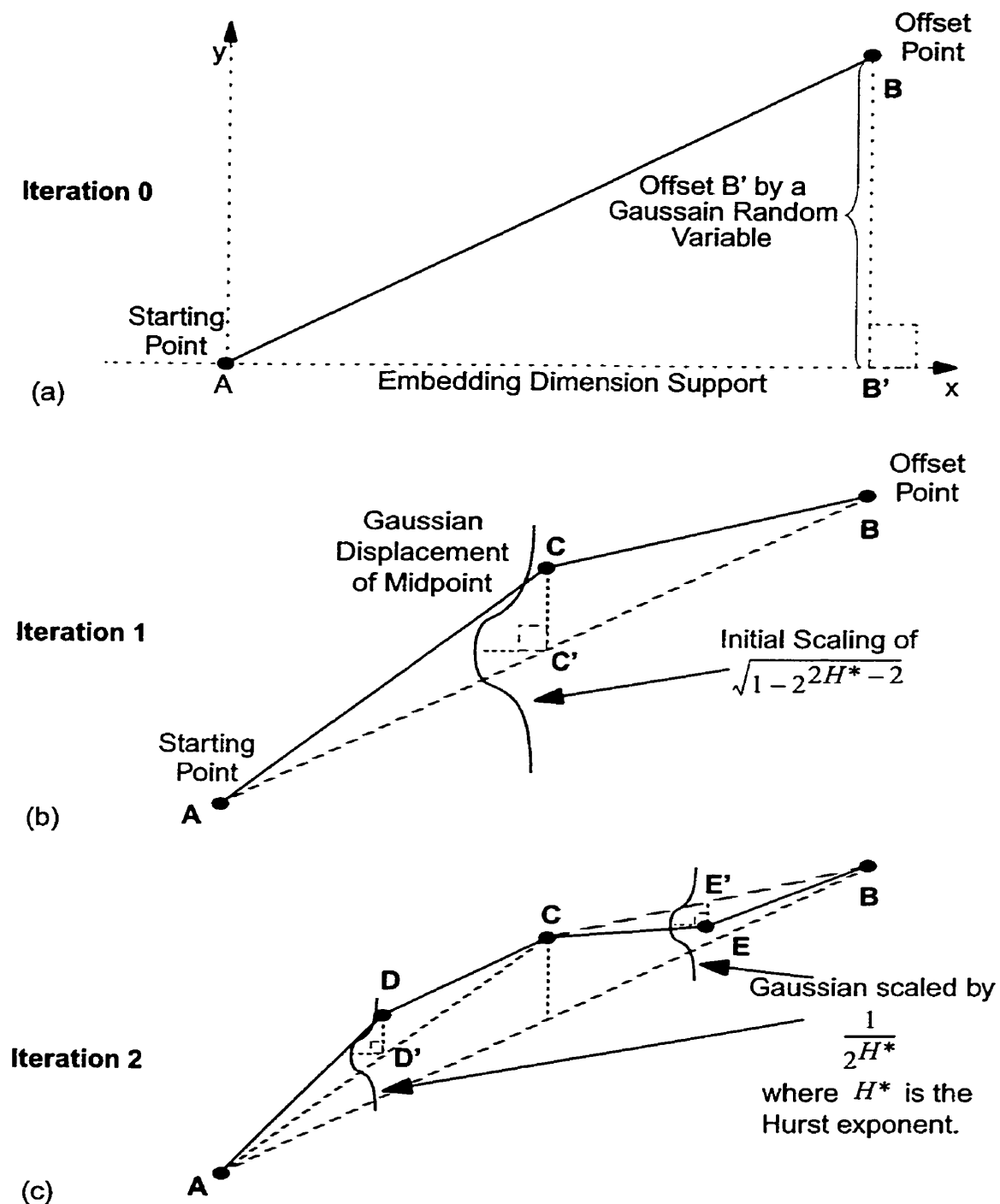
Fig. 5.5. Midpoint displacement algorithm for generating fractional Brownian motion.
(a) initial setup, (b) refinement iteration 1, and (c) refinement iteration 2. (after [Kins95a])

midpoint **C'** along the line segment **AB** is found and then this point is displaced with a random Gaussian variable to get the new point **C**. For the purposes of this thesis, this displacement will always occur perpendicular to the abscissa as illustrated in Fig. 5.5b. Fournier *et al.* suggest that this displacement can also occur perpendicular to the current line segment under consideration [FoFC82]. To achieve the appropriate scaling for fBm the random Gaussian variable must be scaled as shown in Eq. 5.1 for this first iteration with $i = 1$ [Voss85]

$$displacement_{i=1} = \frac{n(x; \mu = 0, \sigma)}{\sqrt{1 - 2^{2H^* - 2}}} \tag{5.1}$$

As can be seen in Eq. 5.1, the scaling is dependent on the Hurst exponent H^* so that different levels of fBm can generated. Recall that the Hurst exponent is related to the fractal dimension and the spectral exponent according to Eq. 3.78. In the case of the fBm curves being generated, $E = 1$ for Eq. 3.78.

Note that the scaling factor in Eq. 5.1 is only valid for the first stage of the RMPD algorithm. This equation arises because Mandelbrot and van Ness [MaVa68], [Mand85] showed that any interval t in the fBm has a variance related to the following equation.

$$\langle |B_H(t) - B_H(0)|^2 \rangle = \sigma^2 t^{2H^*} \tag{5.2}$$

When splitting up the fBm curve recursively in half, this gives a scaling of the Gaussian random variable with variance σ^2 at the $(i - 1)th$ stage to be [Voss85]

$$\sigma_i^2 = \frac{\sigma^2}{(2^i)^{2H^*}} - \frac{\sigma^2}{4} = \frac{\sigma^2}{(2^i)^{2H^*}}[1 - 2^{2H^* - 2}] \tag{5.3}$$

So, at the first stage the scaling is done by the standard deviation of $\sqrt{1 - 2^{2H^* - 2}}$. The first iteration of the fBm refinement process is now complete.

After the first iteration, subsequent refinement iterations follow a similar procedure of recursively finding the midpoint of each line segment and performing the midpoint displacement. The second iteration of this refinement is illustrated in Fig. 5.5c. In this figure, line segment **AC** and **CB** have their midpoints **D'** and **E'**, respectively, displaced perpendicular to the abscissa by a Gaussian random variable. For this second iteration and all subsequent iterations the scaling factor has changed. In this case the displacement for all subsequent iterations i is expressed as

$$displacement_{i \geq 1} = \frac{n(x; \mu = 0, \sigma)}{(2^{H^*})^i} \tag{5.4}$$

This equation immediately follows from the standard deviation by taking the square root of additional stages in Eq. 5.3. This scaling of the Gaussian random variable allows for proper scaling so that the fBm generated remains with the same Hurst exponent throughout all scales [Voss85], [PeJS92, p. 495] and hence the fBm has the proper self-affine fractal characteristics. This refinement process is continued until the desired resolution is obtained.

The description of the midpoint displacement algorithm for fBm curves with $E = 1$ in the $E + 1$ Euclidean dimension is now completed. The next section describes extensions needed to form a two dimensional midpoint displacement so that surfaces can be generated with fBm characteristics as opposed to only curves.

5.2.2   Random Midpoint Displacement Algorithm in Two Dimensions

Generating fBm surfaces using the RPMD is similar to that for generating fBm curves. Figure 5.6 illustrates one method for generating fBm surfaces based on triangle partitioning [FoFC82], [Saup88]. Quadrilaterals, which are sometimes convenient when working with pixel arrays, can also be considered, but their partitioning is a little more complex than that for triangle partitioning [FoFC82].

| (a) Iteration 0 | (b) Iteration 1 | (c) Iteration 2 |
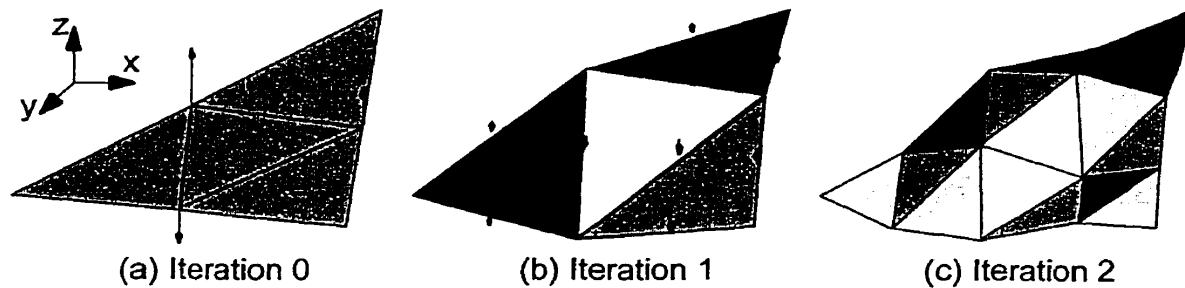
Fig. 5.6. Fractional Brownian motion surfaces through midpoint displacement.
(a) initial setup, (b) refinement iteration 1, and (c) refinement iteration 2.

At the initial setup of the RMPD for surfaces, as illustrated in Fig. 5.6a, a triangle is established in the embedding dimension similar to how the initial line segment was set up in Fig. 5.5 for fBm curves. The initial surface starts in the xy-plane in this three dimensional space (note that $E = 2$ for the $E + 1$ Euclidean space of Eq. 3.78).

The midpoints along each of the three triangle edge line segments are found and then displaced along the z-axis by the scaled Gaussian random variable using Eq. 5.1. This process segments the initial triangle into four smaller triangles as illustrated in

Fig. 5.6b. It should be noted that the shading of the triangles in Fig. 5.6 is solely to help illustrate the different triangles in the surfaces.

Figure 5.6c illustrates the second recursive iteration of the RMPD on each of the four resulting triangles in Fig. 5.6b. The displacement of the midpoints in the four triangles is again done along the z-axis, but, this time using the scaling factor in Eq. 5.4 as was done in the case for fBm curves. Again note that the displacement could be made perpendicular to the triangle surface itself, but, this would cause confusion since triangle surfaces adjoin with each other and then some decision would have to be made as to how to combine the two surfaces' normals. Also, displacement according to surface perpendiculars would cause the surface to no longer be a function in $z = f(x, y)$. Limiting the displacement in the z-axis direction only eliminates these two problems. Similar to fBm curves, this process of midpoint displacement is recursively repeated with each of the subsurfaces until the desired surface resolution is obtained.

(a)   (b)

Gap between
surfaces
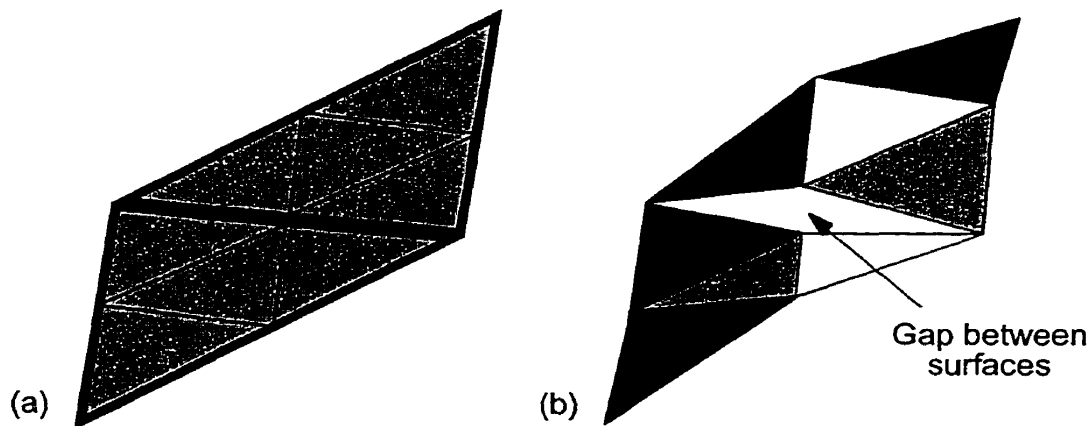
Fig. 5.7. Example of creasing problem with two dimensional
random midpoint displacement.
(a) two adjoining surfaces, and (b) disjoint surfaces after midpoint displacement.

A special note must be made about implementing the RMPD algorithm for surfaces. Treating each of the generated surfaces separately, as through a recursive process, results in disjoint or creased fractal surfaces as demonstrated in Fig. 5.7. Figure 5.7a shows two adjoining triangle surfaces that have become disjoint through the RMPD algorithm in Fig. 5.7b. When these surfaces are used as image greyscale intensities, as is done later in this chapter, this disjointness or creasing manifests itself as triangle artifacts throughout the image as demonstrated in Fig. 5.8a. This creasing and disjointness in the surfaces occur when the RMPD algorithm is applied separately to each subsurface without considering neighbouring surfaces. To eliminate these artifacts, the RMPD algorithm must allow neighbouring surfaces to properly share a common edge when performing the midpoint displacement. Using this strategy with the same parameters as in Fig. 5.8a, Fig. 5.8b was produced which does not contain the same triangle artifacts. Figure 5.8b is not as visually disturbing as Fig. 5.8a and the fractal dimension desired for the corresponding Hurst exponent, $H^* = 0.5$ for this example, is properly maintained.

(a)   (b)

Fig. 5.8. Example images generated using 2D RMPD with $H^* = 0.5$ (Brownian motion) and average greylevel value half of full scale.
(a) recursive RMPD producing creasing, and (b) RMPD with creasing eliminated.

To give a sense of some of the types of surfaces and textures possible with the fBm surface interpolation, an example set of fBm surfaces is illustrated in Fig. 5.9. Shown, from left to right and top to bottom, are 25 sample fBm surfaces with a mean pixel intensity of half scale and a Hurst exponent H^* range from 0.04 to 1.00 in increments of 0.04. It is surfaces similar to these that will be used for the FSSI image compression technique. One difference will be that the surfaces will be of varying sizes. The surfaces will act as interpolating surfaces similar to what was done in Fig. 5.4 with fractal interpolating curves. Hence, the surfaces will follow the contours of the image being represented.
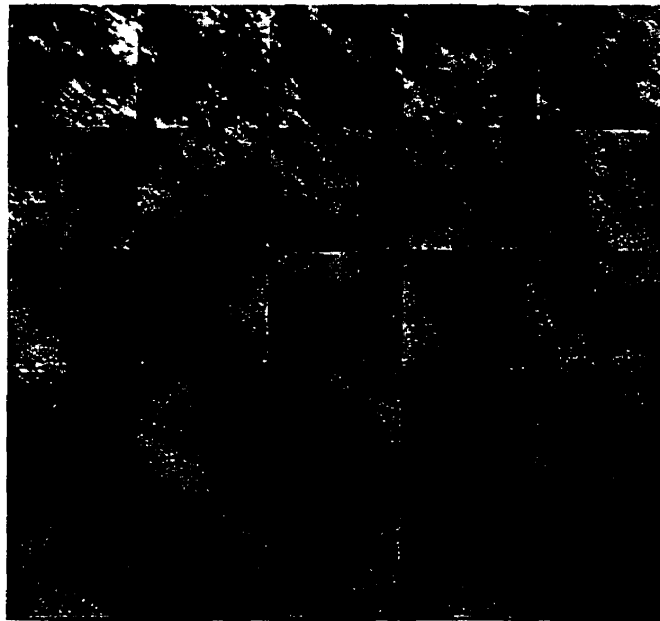
Fig. 5.9. Example surface tiling showing a set of 25 possible fBm surface with a Hurst exponent ranging from 0.04 to 1.00 in 0.04 increments (from left to right and top to bottom).

5.3   Measuring Fractal Dimension Locally in Images for FSSI

Having an interpolation scheme such as the fractal interpolation scheme presented in Sec. 5.1 is only useful if the fractal dimension of the signal can be measured. Before modeling images and textures with fractal surfaces, their fractal nature must be determined. This information is found by calculating the fractal dimension of localized regions within the image. Many algorithmic approaches exist for calculating the fractal dimension locally within an image including those from [PNHA84], [DaKe85], [KeCC89], [CrIv89], [SaCh92], [ChSa95]. This thesis takes the approach of using the Rényi dimension spectrum D_q to calculate the fractal dimension in local regions of an image. For the purposes of the FSSI technique developed in this chapter, the texture measurements are limited to single fractals. Namely, the Rényi dimension spectrum is used with $q = 1$. Rewriting Eq. 3.45 for the Rényi dimension spectrum, the following expression is obtained for $D_{q=1}$.

$$D_{q=1} = D_I = \lim_{s \to 0} \frac{\dfrac{-\sum_{x \in \chi} p_s(x) \log p_s(x)}{\sum_{x \in \chi} p_s(x)}}{\log\left(\dfrac{1}{s}\right)} = \lim_{s \to 0} \frac{H(X)}{\log\left(\dfrac{1}{s}\right)} \tag{5.5}$$

When $q = 1$ for the Rényi dimension spectrum, the resulting fractal dimension is generally referred to as the information dimension because of the use of Shannon entropy $H(X)$ from information theory. It must be noted that the probabilities $p_s(x)$ in Eq. 5.5 are dependent on the current scale of measurement. For the developed FSSI techniques, which are described in the following section, the scale of measurement is limited to 5×5

blocks of pixels to get the local dimension calculations. For each 5×5 block being measured in an image, the probability $p_s(x)$ is calculated as

$$p_s(x) = \frac{\text{pixel}(x) - \min_{\forall i}[\text{pixel}(i)]}{\sum_{\forall i}\left[\text{pixel}(i) - \min_{\forall i}[\text{pixel}(i)]\right]} \tag{5.6}$$

where pixel(x) is the centre pixel in the 5×5 block and the summation goes through each of the 25 pixels in the 5×5 block and adds up the greyscale intensity values. The pixel greyscale values in Eq. 5.6 also have the minimum pixel value in that 5×5 subtracted. This subtraction will partially normalize the texture so that similar textures that are very dark versus the same texture that is very bright results in the same fractal dimension. Note that this formulation of $p_s(x)$ ensures that $\sum_{x \in \chi} p_s(x) = 1$ so that Eq. 5.5 can be rewritten using a complete probability distribution as follows

$$D_{q=1} = D_I = \lim_{s \to 0} \frac{-\sum_{x \in \chi} p_s(x) \log p_s(x)}{\log\left(\frac{1}{s}\right)} \tag{5.7}$$

The following section describes the FSSI techniques developed using the idea of fractal surface interpolation and this form of local fractal dimension calculation.

5.4   Image Compression and Progressive Image Transmission using Fractal Surface Segmentation and Interpolation

The goal behind fractal interpolation and fractal surface interpolation is to represent a curve or surface faithfully from a perceptual point of view using fractal curves and

surfaces. For image representation specifically, the goal is to represent the surfaces of an image projection in a perceptually faithful manner. For instance, the two images displayed in Fig. 5.10 are of the original image of lena [Lena99], [USC99] in Fig. 5.10a and its three dimensional projection as a surface plot in Fig. 5.10b. The goal therefore is to find a method of using fractal surface segmentation and interpolation to represent the this surface projection using fBm surfaces that model the fractal nature of the surface projection. This segmentation can follow a process similar to the feature extraction as used by Ferens and Kinsner [FeKi95] on surface plots. The interpolation can follow a process similar to the natural scene description done by Pentland [Pent84], with modifications so that explicit surface contours can be generated that model specific image surface plots. The next two subsections present two techniques developed in this thesis for this image segmentation and interpolating representation using fractal dimensions and fBm.

(a)   (b)
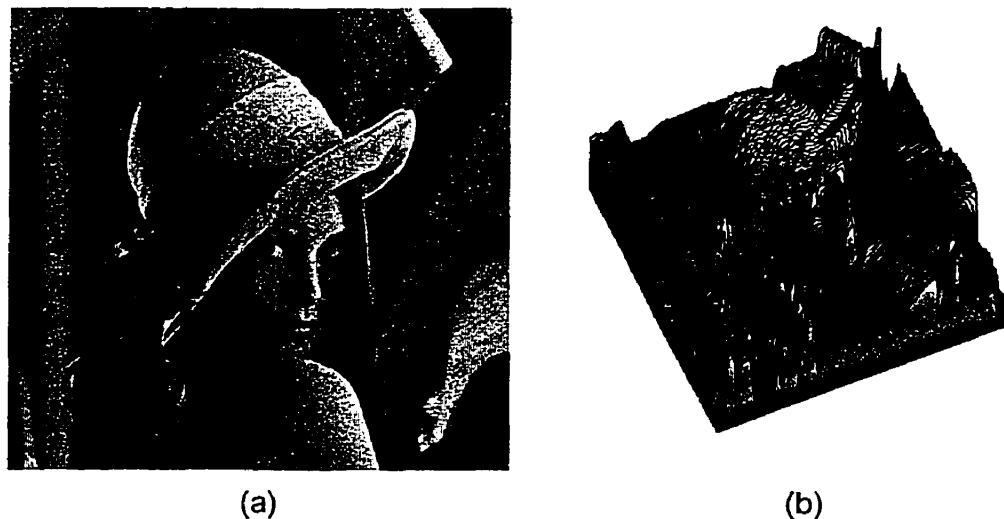
Fig. 5.10. (a) Image of lena [Lena99], [USC99] and
(b) a three dimensional surface plot of lena.

5.4.1   FSSI Image Representation/Compression Technique 1

With the tools in place for generating fractal surfaces of a desired dimension and a method for measuring the fractal dimension of local regions in an image, a scheme is now developed for image representation and compression which was first described by Dansereau and Kinsner [DaKi96]. The high-level approach taken in this image compression scheme is to segment an image surface plot, such as in Fig. 5.10b, into regions according to areas of uniform fractal dimension. The surfaces in these regions of uniform fractality can then be represented as fBm surfaces with interpolation being performed between the edges of the regions.

Different approaches can be taken to implement this high-level view of this first FSSI image compression scheme. A first step is to calculate the local fractal dimensions throughout the image. This is done, as partially described in Sec. 5.3, by taking a 5×5 block mask of pixels and calculating the fractal dimension, D_f, of that block according to Eq. 5.7 in combination with Eq. 5.6. The calculated fractal dimension is recorded for that block and then the block is shifted by one pixel to get a complete coverage of the image being analyzed. Using a 5×5 sliding window block to perform the fractal dimension calculations on an $N \times M$ image results in an $(N-4) \times (M-4)$ matrix of fractal information dimension values for local calculations throughout the image.

With a matrix of local fractal dimensions now calculated, the next step is to decide on how to segment the image into regions of uniform fractality. Different techniques can be used such as edge detection, region growing, or split-merge algorithms, but, for simplicity and for proof of concept the segmentation of the image is done using a triangle tes-

sellation. That is, the matrix is split into a set of triangular regions of varying sizes. This method will later simplify the generation of fBm surfaces since the RMPD algorithm described in Sec. 5.2.2 is based on triangle surface partitioning.
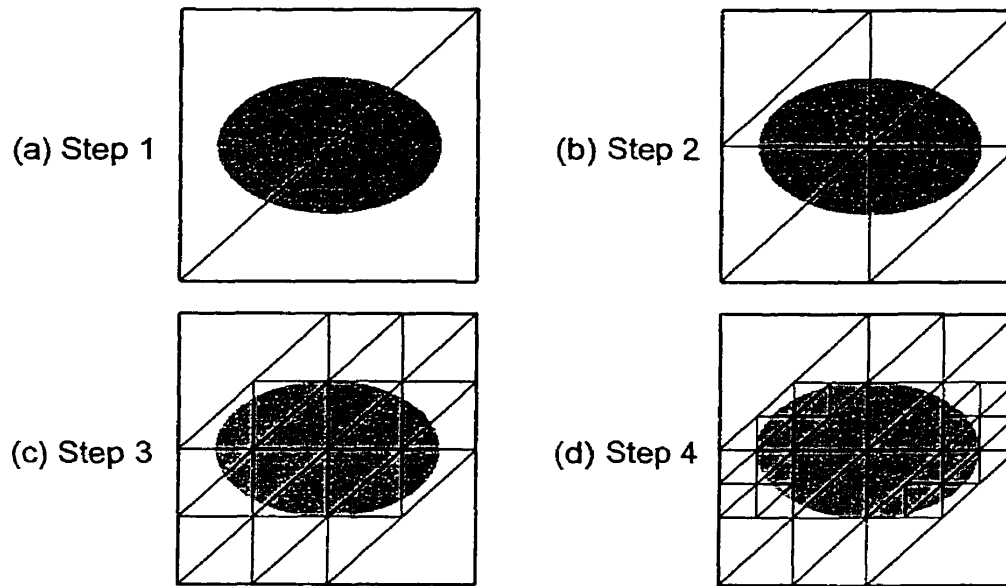
Fig. 5.11. Example tessellation of an oval.
(a) Step 1, (b) Step 2, (c) Step 3, and (d) Step 4.

Representation of this triangle tessellation is of paramount importance to achieve compression. The approach taken for this FSSI technique is illustrated in Fig. 5.11. Shown in Fig. 5.11a is an example fractal dimension matrix which happens to contain a region of uniform fractality in the shape of an oval versus the background which has a different uniform fractality. The tessellation scheme used first splits the matrix into the two initial triangles as shown. The algorithm then continues by checking each triangle and deciding whether the region within the triangle is of uniform fractal dimension. If the region within the triangle does not have uniform fractal dimension then the triangle is par-

titioned into four equal sized triangles, as illustrated in Fig. 5.11b. If the triangle region is already uniform then no further partitioning is required. This process is recursively repeated with each triangle. A sample image tessellation of lena, shown in Fig. 5.12, illustrates how an image can be subdivided using this triangle tessellation scheme.
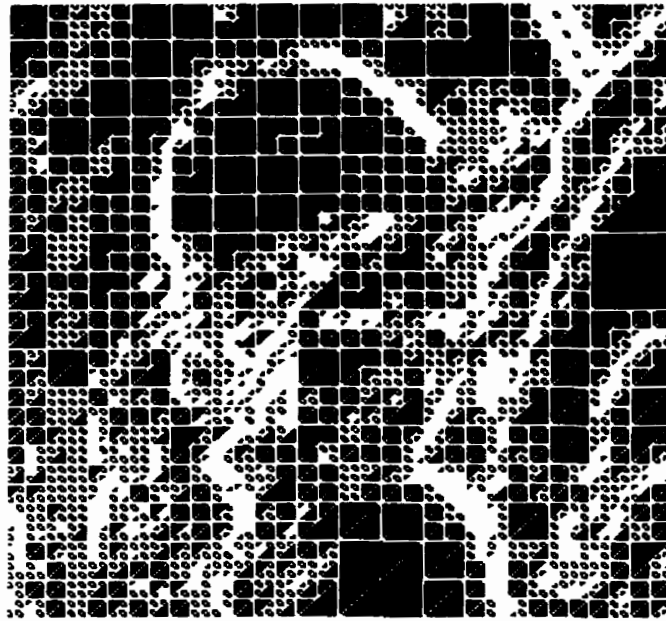
Fig. 5.12. Sample tessellation of lena with triangles.

Efficient encoding of the triangle partitioning is done by representing the partitioning as a quad-tree, that is, a tree structure where each node has four children nodes, with each node representing one triangle. Only 1 bit is then required to determine whether a specific triangle is a leaf node in the quad-tree or if the triangle is further partitioned with four children triangles. A slightly more complicated scheme can be used where the triangle is partitioned into unequal sized triangles. This approach can produce a more adaptive matrix tessellation with fewer triangles depending on the layout of the fractal dimensions,

but, would increase the bit rate since the sizes of the triangles in each partitioning would somehow have to be represented (*i.e.* more than 1 bit).

With the fractal dimension matrix now tessellated with triangles, the next step required for the fractal surface interpolation is to determine sets of points to interpolate between. The tessellation in the fractal dimension matrix alone only gives uniform regions of fractal complexity in each triangle. So, points must be chosen from the original image so that the a fractal surface with the corresponding fractal dimension can be generated to interpolate between these points. The easiest approach to this requirement is to encode the greyscale pixels that correspond to the corners of all of the triangles in the fractal dimension matrix tessellation. Considering Fig. 5.6a, this is equivalent to setting the three corners of the triangle to a height equal to the corresponding greyscale pixel values from the original image. Hence, there are three points forming a triangular region with which to perform the fractal surface interpolation. With the three interpolation points for the triangle and the fractal dimension for the surface, the RMPD algorithm can be performed to synthesize a surface that interpolates the three points. This fractal surface should be perceptually better than if only a smooth surface was interpolated between the three points. Doing this surface interpolation for the entire image will reconstruct an approximation to the original image.

A data flow chart for the entire encoding and decoding phase of this FSSI technique is illustrated in Fig. 5.13. The process starts with the original image as input. Localized fractal dimension measures are calculated for the image and in this case stored as the Hurst exponent H^* since this is what the RMPD from Sec. 5.2.2 uses (recall that

the fractal dimensions can be related to H^* through Eq. 3.78). The fractal dimension matrix is then tessellated into uniform regions of fractality, with the inclusion of an extra stopping criteria when a triangle partitioning reaches a desired resolution. The tessellation information, Hurst exponent for each tessellating triangle, and interpolation points from the corners of each tessellating triangle are then encoded into the output data stream. Further lossless encoding can also be performed on the output data stream.
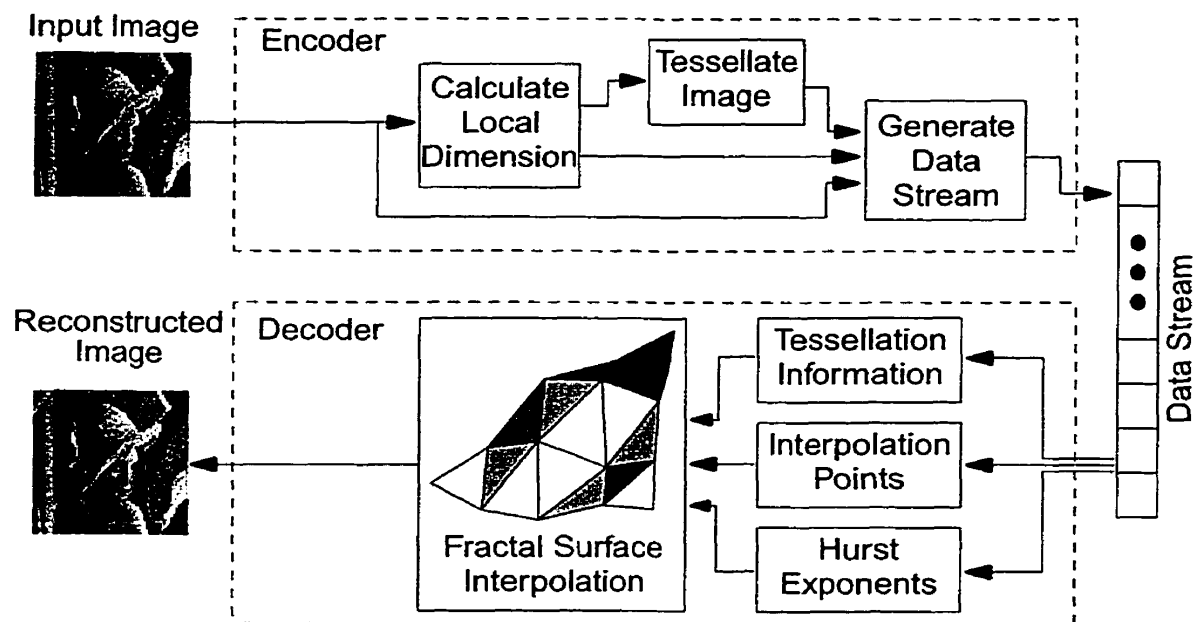
Fig. 5.13. Data flowchart of the first FSSI image compression/reconstruction scheme.

For the image reconstruction, the tessellation information, interpolation points, and Hurst exponent are extracted from the data stream for each of the tessellating triangles. The tessellation quad-tree binary decisions, encoded as 1 bit yes/no answers, are used to determine the triangle layout in the fractal dimension matrix. To the triangles the interpolation point intensities are overlaid at the corners of the triangles and the Hurst exponents

associated back to the corresponding triangles. Using this decoded data stream, the RMPD algorithm can be performed to synthesize the fractal surfaces between the interpolation points to reconstruct the image.

5.4.2   FSSI Image Representation/Compression Technique 2

A second approach to performing the fractal surface interpolation, first proposed by Dansereau and Kinsner [DaKi97], is presented in this section. This technique has many similarities to the technique presented in Sec. 5.4.1, but, in addition the fractal surface interpolation is superposed over a wavelet based approximation of the image, similar to what was done with the curve superposition in Fig. 5.4 of Sec. 5.1.

The data flow for this second FSSI technique is illustrated in Fig. 5.14. The compression, or encoding stage, starts by taking the input image and passing it through two separate decomposition processes. The first decomposition process is to calculate the local fractal dimensions within the image in the same manner as is done in Sec. 5.4.1. The result is a two dimensional matrix containing local Hurst exponent H^* measurements that are later used in the FSSI reconstruction phase. This Hurst exponent matrix is passed through a DWT, as described in Sec. 4.3.3. The discrete wavelet coefficients from this transformation then have a hard threshold applied which zeroes out any wavelet coefficients with a magnitude smaller than a user chosen threshold value. This hard threshold of the coefficients forms an approximation of the matrix in the form of the non-linear approximation discussed for Eq. 4.12. The result is then efficiently encoded using an algorithm developed for the DWT by Shapiro [Shap93] known as embedded *zerotree coding* (ZT), which is discussed in more detail in Chapter 7. For now, just consider the zerotree coder

to be an efficient method of representing discrete wavelet coefficients for progressive

transmission. Finally, the output of the zerotree coder is passed through a lossless entropy

coder to remove any further redundancy. In the experiments performed in Sec. 5.5.2 using

this FSSI technique, an adaptive arithmetic coder, as described by Witten, Neal and Cleary

[WiNC87], is used as this final stage of lossless entropy coding.
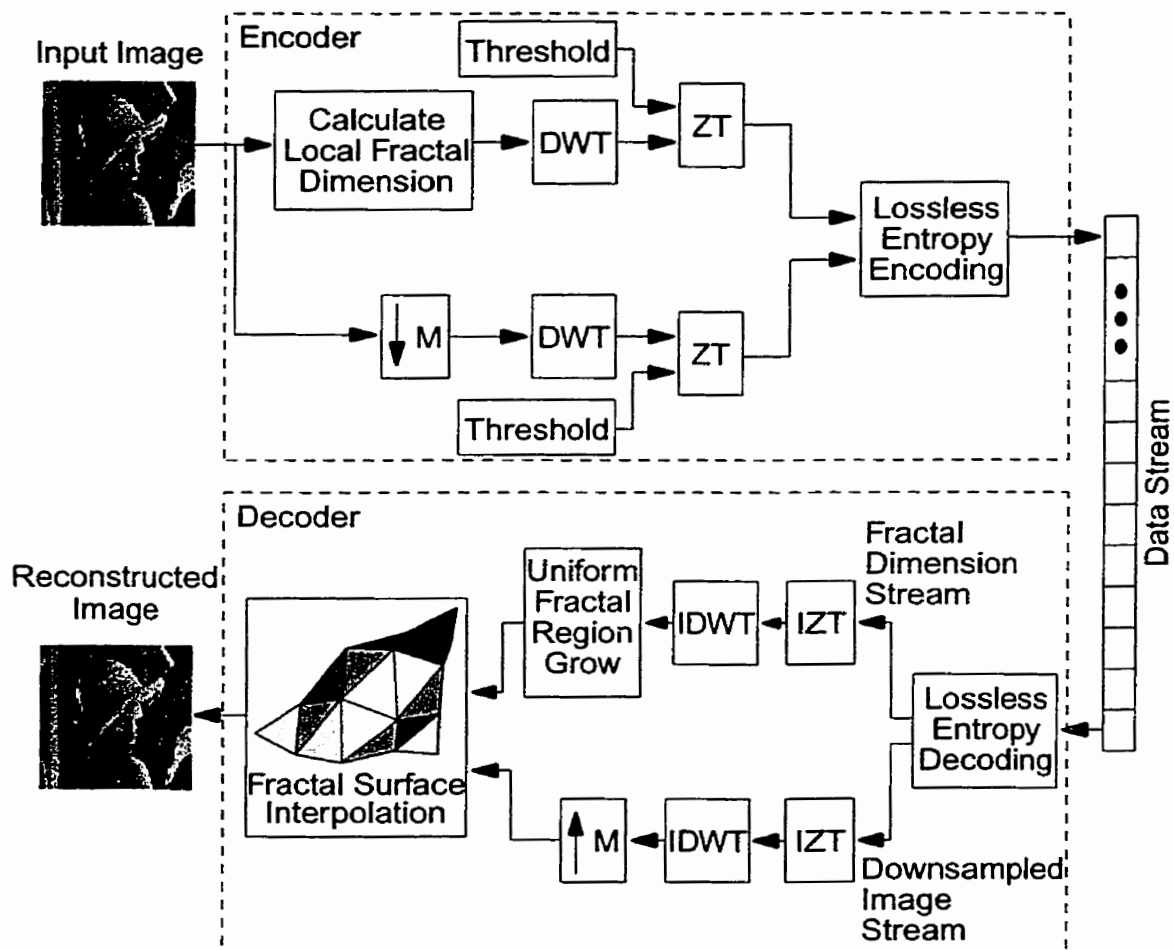
Fig. 5.14. Data flowchart of the second FSSI image compression/reconstruction
scheme with wavelet support.

The second half of the encoding process is shown in the bottom branch of the encoder stage. This stage of the encoding starts by downsampling the input image by performing an M-point decimation (refer to Defn. 4.14) on all rows and columns of the image. The level M of downsampling affects the compression ratio and reconstruction quality of the image, and is chosen by the user. Similar to the Hurst exponent matrix, the downsampled image is passed through a DWT, hard thresholded, zerotree encoded, and then lossless entropy encoded using the adaptive arithmetic coder. The compression side of the image encoding process is now complete.

Image reconstruction reverses this process and performs the FSSI as illustrated in the bottom half of Fig. 5.14. The compressed data stream is decoded through the lossless entropy decoder. Then the downsampled image components and the Hurst exponent components are separated and each passed through the *inverse zerotree* (IZT) decoder and the *inverse discrete wavelet transform* (IDWT). The downsampled image data is upsampled (refer to Defn. 4.15) back to its original size. This upsampled image now contains the required interpolation points for the fractal surface interpolation.

Regions of uniform fractal dimension in the Hurst exponent matrix are then identified. With the interpolation points and regions with uniform Hurst exponents identified, fractal surface interpolation is then performed superposing the results on the approximation of the image that had been downsampled in the encoding phase. Effectively, as with Fig. 5.4, the discrete wavelet approximation of the image is used as the overall path desired for the surface and the Hurst exponents are used to add in extra fractal characteris-

tics according to the original image. The reconstruction phase of the image is now completed.

5.4.3   Progressive Image Transmission with FSSI Techniques

The two FSSI techniques presented in Sec. 5.4.1 and Sec. 5.4.2 can easily be adapted to allow for progressive image transmission as is described by Dansereau and Kinsner [DaKi98a]. For the technique presented in Sec. 5.4.1, the quad-tree partitioning into a triangle tessellation can inherently be sent progressively because of the partition process. To make the rest of the algorithm progressive, the interpolation points at the corners of triangles must be transmitted as the partitioning occurs as well as the Hurst exponent for each triangle. To improve encoding efficiency, the full Hurst exponent does not need to be sent, but, instead the change in the Hurst exponent. That is, the difference between a child triangle's Hurst exponent and the parent triangle's Hurst exponent expressed as

$$\Delta H^* = H^*_{child} - H^*_{parent} \tag{5.8}$$

is transmitted instead of H^*_{child}.

For the FSSI technique presented in Sec. 5.4.2, the zerotree coding technique developed by Shapiro [Shap93] is progressive by design. Therefore, reconstruction of the image can continuously be done as more and more data is received at the decoding stage.

5.5   Experimental Results with the FSSI Techniques

A number of experiments were performed with the FSSI techniques described in Sec. 5.4 to determine any important properties of these techniques. The following two subsections describe these experiments and present results from the experiments.

5.5.1   Experimental Results for FSSI Technique 1

Three images, lena, peppers and baboon, as shown in Appendix A, were used to test the first FSSI image compression scheme. These images are 512×512 8-bit grey-scale images. To start, a number of test image compressions were performed to check the basic perceptual quality of the image reconstructions. To illustrate the quality of the image reconstructions, a sample of these results for lena are provided in Fig. 5.15. Shown are reconstructions of the lena image with 25.34 dB PSNR at 0.450 bpp, 27.09 dB PSNR at 0.665 bpp, and 29.74 dB at 1.303 bpp. These samples show that some triangle and smudging artifacts appear at smaller PSNR values in the images but that none of the serious surface creasing as demonstrated in Fig. 5.8a occurs.

(a) Original image   (b) 25.34 dB PSNR at 0.450 bpp

(c) 27.09 dB PSNR at 0.665 bpp   (d) 29.74 dB PSNR at 1.303 bpp
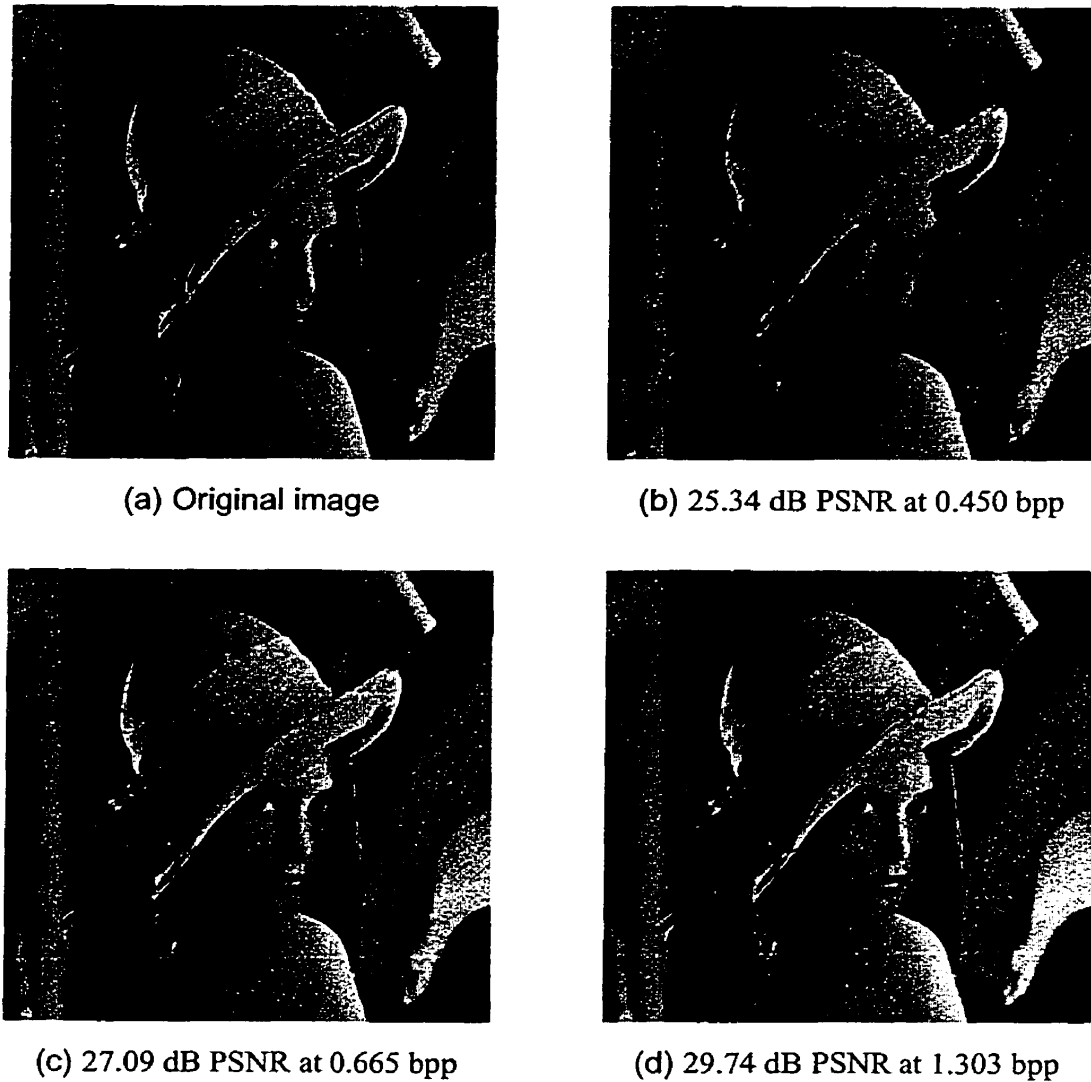
Fig. 5.15. Sample image reconstructions of lena. (a) original image
(b) 25.34 dB PSNR at 0.450 bpp, (c) 27.09 dB PSNR at 0.665 bpp, and
(d) 29.74 dB PSNR at 1.303 bpp.

With the basic image reconstruction verified for the FSSI image compression

scheme, a number of image compressions and reconstructions were performed on the test

images to evaluate the compression performance. These experiments were done by vary-

ing the sensitivity of the image tessellation to the following: (1) the relative uniformity of

local fractal dimensions calculations and (2) the minimum allowable triangle size. Both of these are factors in the tessellation of an image being compressed. The relative uniformity of local fractal dimensions is important since some criteria needs to be in place to determine whether a triangle needs to be further subdivided. This decision can be made by ensuring that all fractal dimensions in a region fall within a certain small range of values. The minimum allowable triangle size affects the resolution in the final image compression and hence will affect the bit rate for the compression.

Results of the peak signal-to-noise ratio versus bits per pixel for these experiments are given in Fig. 5.16. Lines have been fitted to the data points to show the trend in PSNR as the rate in bpp increases. As expected, for each of the test images the PSNR improves as the bpp increases. From the three plots, lena and peppers have similar PSNR values over most of the range of bpp values unlike the plot for baboon which has much lower PSNR values than lena and peppers. It is also noted that below approximately the 0.3 bpp mark, the PSNR drops significantly for all three images using this FSSI technique. This PSNR drop is primarily due to there not being enough data to reconstruct much of the image.
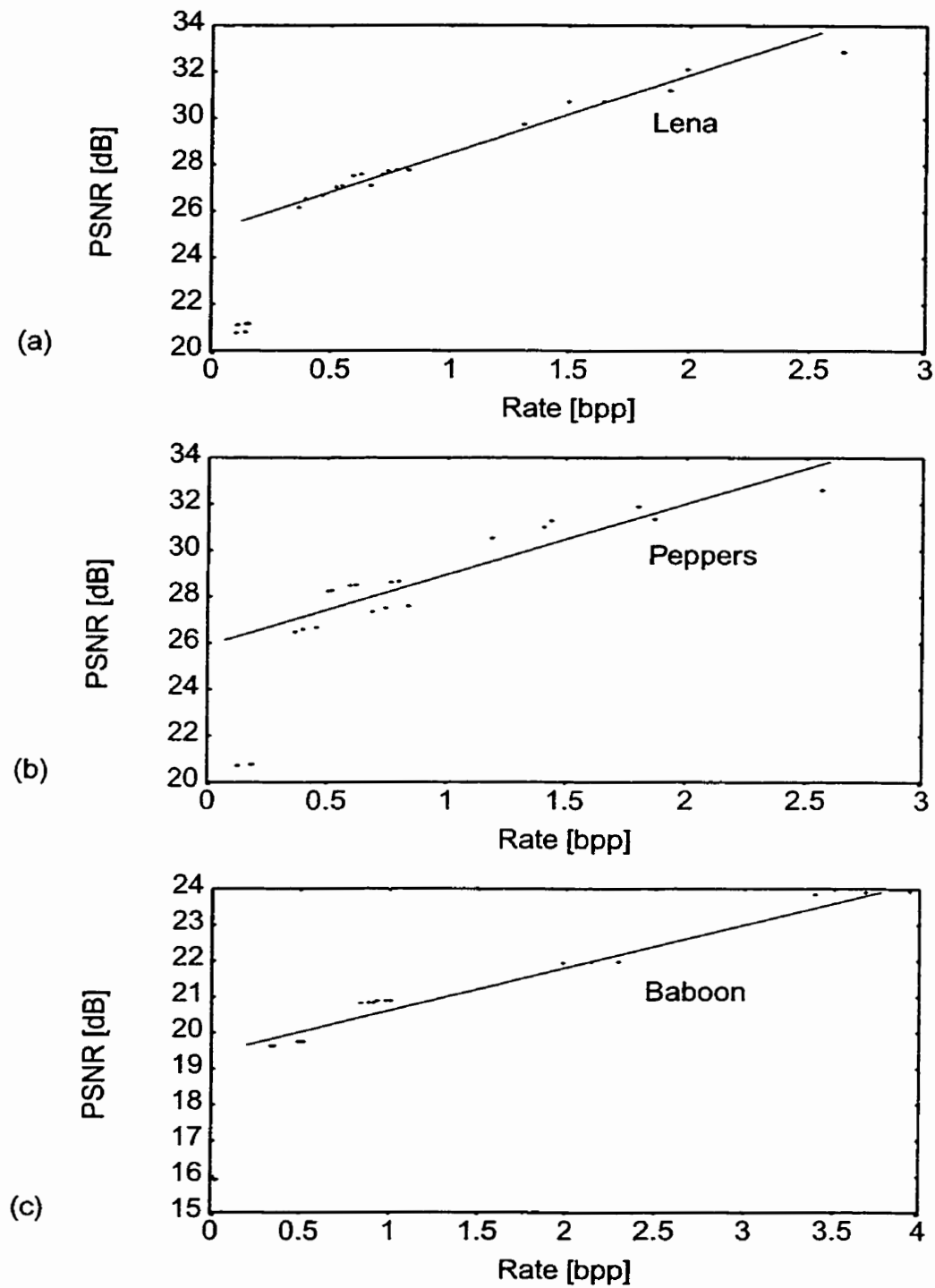
Fig. 5.16. Plots of PSNR vs. bpp experimental results with
FSSI technique 1 on the images
(a) lena, (b) peppers, and (c) baboon.

The baboon image results in PSNR measures consistently 5-6 dB below that of lena and peppers. This difference can be explained by the relative fractal complexity of the baboon's hair in the image (refer to Appendix A). As indicated in Chapter 3, smoother objects will typically have a smaller fractal dimension where rougher objects will typically have a higher fractal dimension. The baboon's hair in the image is much closer to a white noise type surface then say a smoother black noise type surface. This observation leads to the question as to why the PSNR for the baboon image is consistently lower then that of lena and peppers, which are composed much more of smoother surfaces. Is fractal surface interpolation limited only to smoother surfaces versus the rougher surfaces as in baboon's hair? Considering that the FSSI scheme should be able to statistically model the baboon's hair just as well as lena and peppers from a fractal point of view, this drop in PSNR image quality seems out of place.

The problem lies in what PSNR is actually calculating. Consider the equation for PSNR in Eq. 2.11. It should be realized that PSNR effectively measures the pixel energy differences throughout the image reconstruction on a *pixel by pixel basis*. This factor is key to understanding why rougher surfaces like the baboon's hair have a much lower PSNR using this FSSI algorithm. The fractal surface interpolation attempts to synthesis a surface between the interpolation points with the same fractal characteristics as in the original image. The hope is that this surface will be perceptually similar to the original.

(a) Original image of Lena   (b) Image of Lena shifted by one pixel
PSNR = ∞   PSNR = 25.4385 dB
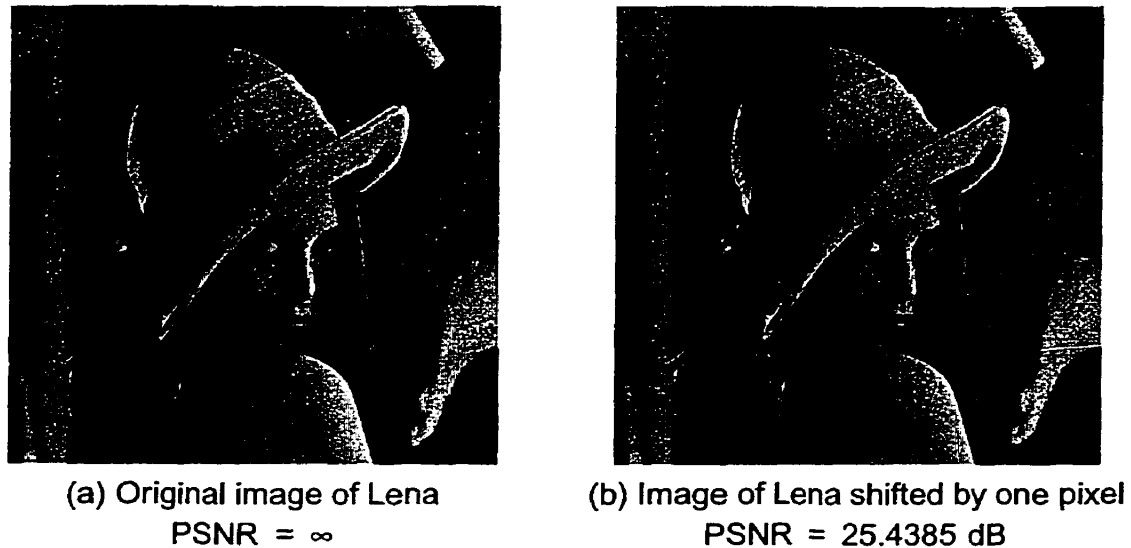
Fig. 5.17. (a) Original image of Lena and (b) image of Lena translated
by one pixel diagonally down to the right.

An important realization is that this interpolation does not attempt to make specific

pixels a specific value. In a rough surface, one pixel may be very bright while the neigh-

bouring pixel may be very dark. The fractal surface interpolation will recreate the rough

surface, but, having specific pixels as bright or dark is not guaranteed even if the percep-

tual nature of the surface is similar to the original. In other words, the fBm generation has

a tendency to create perceptually similar surfaces but not necessarily exact pixel intensi-

ties. Therefore, in terms of the FSSI technique developed, the PSNR measure is not suit-

able as an image quality measure. This statement can be emphasized by considering what

happens with the PSNR measure when translation of the image occurs, which is in essence

what occurs with the FSSI technique on a local scale. For instance, consider the two

images in Fig. 5.17 which are the original image of Lena and the image of Lena shifted by

one pixel diagonally down and to the right. Since there is no noise for the original image,

the PSNR is infinite. For the image that has been translated by one pixel, the PSNR is 25.4385 dB. From the general rule of thumb use of PSNR, this PSNR suggests an image that is of perceptually poor quality compared to the original image. This is obviously not the true when the perceptual quality of the two images are compared. From the realization that PSNR is an unreliable indicator of image quality for the FSSI technique, Chapter 6 explores other measures more suitable for compression techniques such as FSSI.

As an additional experiment, the Shannon entropy (Eq. 3.19) of the output data streams was measured and compared with the compression rates obtained. Depending on the data compression scheme used, an average boost of 1:1.25-1:1.40 extra compression over the original output data stream would be obtained. This would bring the bit rate down from, for example, 0.5 bpp to 0.35 bpp.

5.5.2   Experimental Results for FSSI Image Compression Technique 2

A number of experiments were performed with the second FSSI image compression scheme on the 512×512 8-bit greyscale image of lena, which is shown in Appendix A. Of interest is the compression rate in bpp versus the reconstruction PSNR, as the input parameters of the compression are changed.

(a) Original image   (b) 26.033 dB PSNR at 0.2528 bpp

(c) 26.639 dB PSNR at 0.5497 bpp   (d) 26.9889 dB PSNR at 0.7737 bpp

Fig. 5.18. Sample image reconstructions with FSSI technique 2 on image of lena.
(a) original, (b) 26.033 dB PSNR at 0.2528 bpp,
(c) 26.639 dB PSNR at 0.5497 bpp, and (d) 26.9889 dB PSNR at 0.7737 bpp.

To demonstrate the image reconstruction with this FSSI technique, Fig. 5.18 pre-

sents three different image reconstructions of lena at different bits per pixel. The first

reconstruction shown in Fig. 5.18b is lena at 0.2528 bpp with a PSNR of 26.033 dB. This

is one of the higher compression rates achieved where the image is still recognizable. The textures are smudged throughout the image and some artifacts arise near edges at region borders where the Hurst exponent is different. The second image shown in Fig. 5.18c is lena at 0.5497 bpp with a PSNR of 26.639 dB. This image has fewer artifacts compared to Fig. 5.18b since some extra textures are visible such as in the hat and the edge of the scarf around the base of the hat. Also, some of the artifacts from Fig. 5.18b have disappeared such as those along the white bar running down the left side of the image. More definition exists in the feathers and the curved texture of the shoulder is sharper. Finally, Fig. 5.18d shows lena at 0.7737 bpp with a PSNR of 26.9889 dB. Again, the textures improve while some of the artifacts disappear. In this case a gain of less than 1 dB of PSNR is obtained by tripling the bit rate. This result shows that this technique in its current state still requires careful selection of parameters to get good image reconstructions at a low bit rate.

Since a large increase in the number of bits required versus the increase in PSNR is noticed with the sample reconstructions in Fig. 5.18, experiments were designed to test how the different input parameters affect the image reconstructions. For the experimentation, there are three parameters that affect image compression. The first parameter is the level the input image is downsampled before being passed through the DWT. The two other parameters are the threshold values used at both zerotree coding stages. Along with these parameters at the compression stage, there is also one hidden reconstruction parameter for the region growing process. This variable controls the region growing process for deciding what is considered a uniform region in the Hurst exponent matrix.
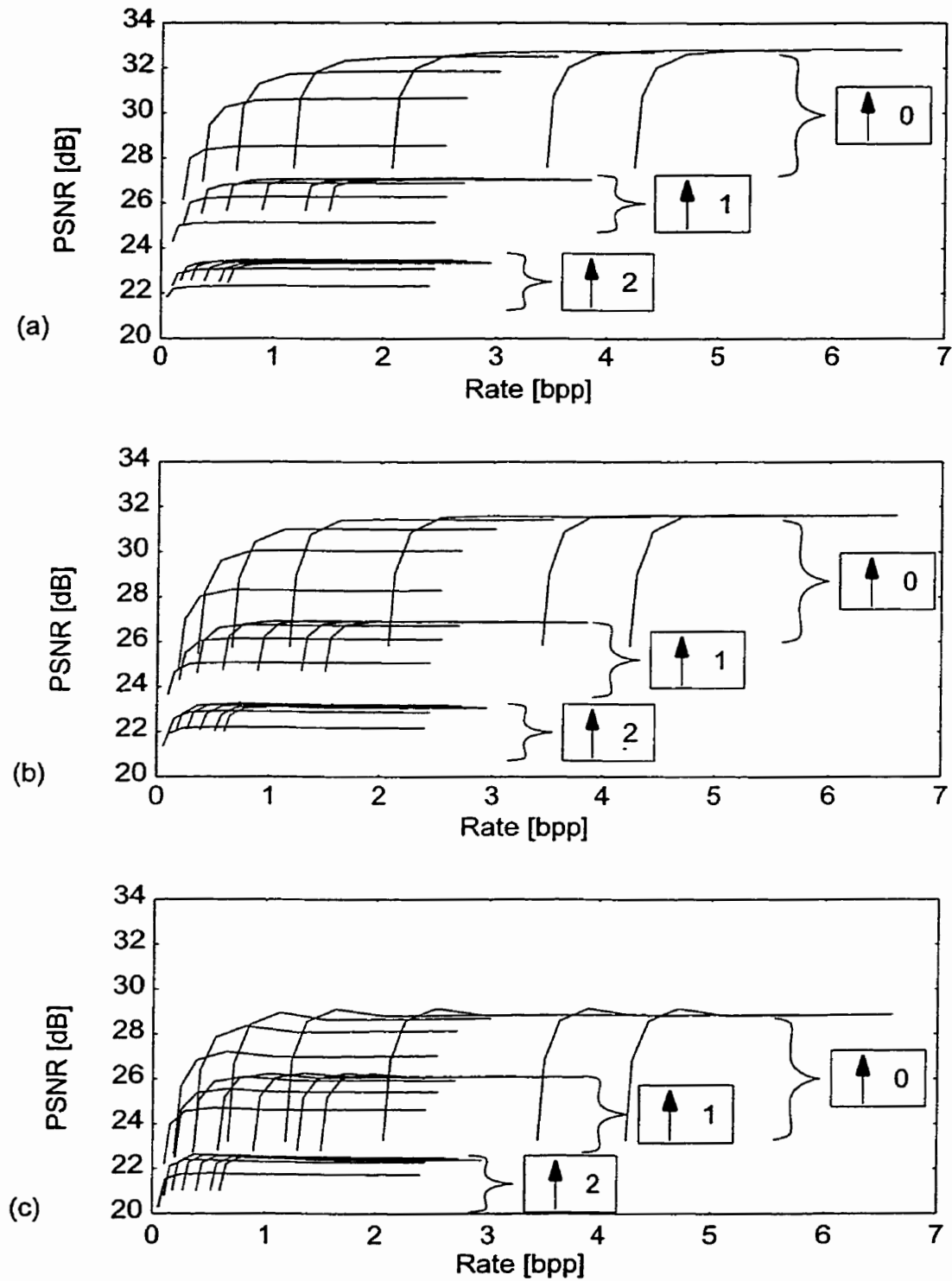
Fig. 5.19. Experimental results for FSSI technique 2 with
Hurst exponent quantization = (a) 0.05, (b) 0.1, and (c) 0.2.

The main set of experimental results are shown in the three plots in Fig. 5.19. These plots show the PSNR of the image reconstruction versus the compression rate in bpp. Focusing on Fig. 5.19a, it is evident that the plot has three separate data regions. These regions are labelled ↑0, ↑1, and ↑2 and indicate the level of downsampling/upsampling done with the input image at the compression and reconstruction phases. Each of these regions contain seven curves which represent, from left to right, decreasing zerotree threshold values for the downsampled image (64 to 1 by powers of 2). Within each curve, from the bottom of the curve to the top of the curve, are the values for decreasing zerotree threshold values for the Hurst exponent matrix (64 to 1 by powers of 2). Each of the three plots in Fig. 5.19 have these characteristics except that the Hurst exponent quantization level at the region growing stage during reconstruction is 0.05, 0.1, and 0.2 for Fig. 5.19a, Fig. 5.19b, and Fig. 5.19c, respectively. This Hurst exponent quantization dictates the range of Hurst exponent values to pass as a uniform region of fractal dimension measurements.

As the plots in Fig. 5.19 show, there is a 3-6 dB difference in PSNR going from no downsampling, ↑0, to a dyadic downsampling, ↑1, and 3-4 dB difference going from a dyadic downsampling, ↑1, to a quadratic downsampling, ↑2. This variation is understandable since downsampling the image removes information so the reconstruction quality will degrade.

When the zerotree threshold for the downsampled image is decreased, the characteristic curve moves up and to the right (*i.e.* quality gets better, but the coding rate increases). Also, when the zerotree threshold for the Hurst exponent matrix is decreased,

the quality of the reconstruction increases, while the rate increases. These results are expected since higher resolutions obtained from the zerotree coders would improve the image reconstructions while also increasing the number of bits required. An interesting point to note is that there is a significant knee in each of these curves. This knee is a transition point between the two main processes in the reconstruction of the image. Below this knee point the major contributing factor in the reconstruction quality is the Hurst exponent matrix. Above this knee point the region growing quantization level in the Hurst exponent matrix starts to work against the improved accuracy in the downsampled image. This influence causes the reconstruction quality to begin to plateau. Therefore, the optimal operating point for the best reconstruction results versus compression rate is at the knee points in these curves.

Another interesting observation from the three plots in Fig. 5.19 is that the PSNR value decreases for all of the curves as the quantization range for the Hurst exponent matrix increases. Again this is expected since a coarser Hurst exponent representation results in a less defined fractal surface reconstruction. A balance must be achieved to keep the quantization level high enough so that the enhanced texture benefits from FSSI are realized, but, not so high as to lose the detail within the image through the random process of FSSI. Additionally, the quantization must not be so low as to rely heavily on the downsampled image for interpolation points.
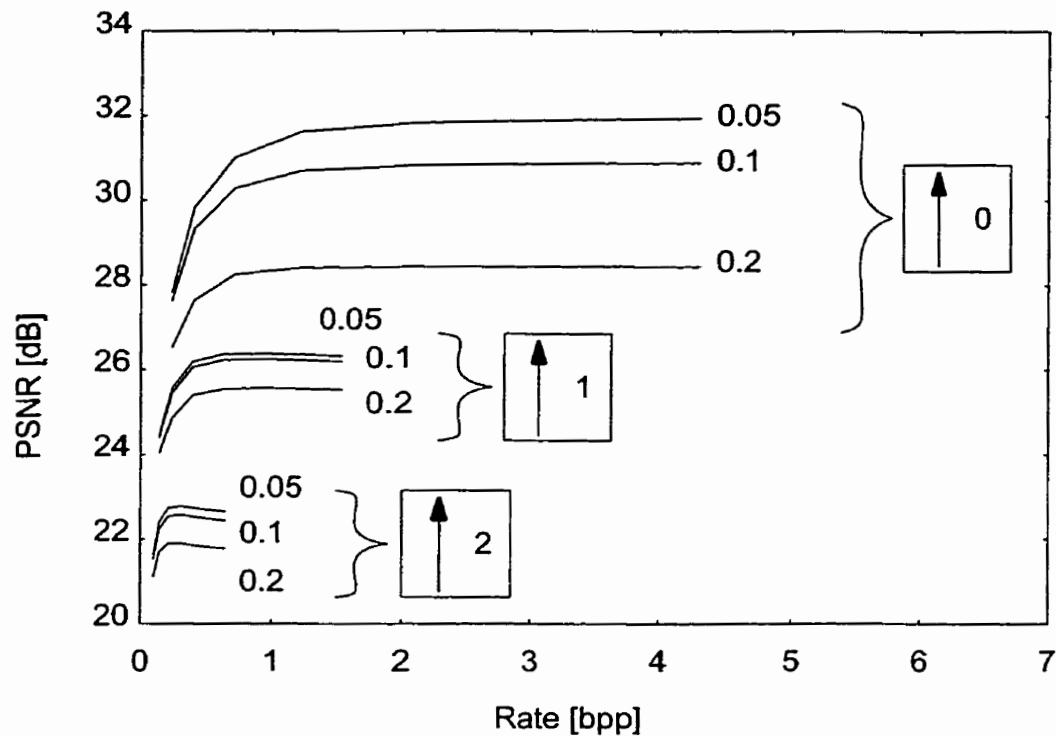
Fig. 5.20. Knee points for different Hurst exponent quantization levels and
downsampling levels for FSSI technique 2.

The plot in Fig. 5.20 gives a more clear relationship between the optimal operation

point versus Hurst exponent quantization. In this figure, the knee points for each of the

curves from Fig. 5.20 have been extracted and plotted, forming curves with other knee

points having the same Hurst exponent quantization (0.05, 0.1, and 0.2). These curves are

grouped according to the downsampling/upsampling level (↑0, ↑1, and ↓2) used in the

experiment. It is again clear that the quantization level of the Hurst exponent matrix plays

an important role since the smaller the quantization, the better the reconstruction. It

should also be noted that there is actually a slight decrease in reconstruction quality for

downsampling levels ↑1 and ↓2 resulting in a slight hump in the curve. This feature is also

apparent in Fig. 5.19 for these downsampling levels. These humps are likely due to an

optimal point of operation for the implemented FSSI technique. The implemented FSSI

operates with a minimum resolution area for the surfaces it generates, primarily to avoid

triangles that are long and thin.

5.6   Summary

A main focus of this chapter is the presentation of a new form of image compres-

sion and progressive image transmission based on FSSI done with fractional Brownian

motion. Two progressive image transmission FSSI techniques were presented with some

experimental results looking at the PSNR versus compression rate. One important obser-

vation made is that the PSNR metric is a poor measure when this type of random surface

generation is performed to reconstruct an image. This issue is considered further in the

following chapter which develops new image quality measures based on multifractal com-

plexity measures that do not have some of the shortcomings of the PSNR measure. The

presented FSSI techniques are unique since images are decomposed and represented by

fractional noise generated by fBm. This approach is different from the majority of image

compression techniques since here an image is comsidered as a collection of fractional

noises, instead of the normal view where an image is a pure signal that is possibly contam-

inated by noise.

CHAPTER VI

MULTIFRACTAL MEASURES OF PROGRESSIVE IMAGE TRANSMISSION SCHEMES

An important concern to those researching lossy image compression and progressive image transmission techniques is how to evaluate the quality of the image reconstructions. Unfortunately, with the lack of a precise model for human perception, the only current arguably accurate approach to evaluating image quality is through a *mean opinion score* (MOS) perceptual quality rating of images. This MOS is done by having a group of human observers rate the perceptual quality of test images, with the results statistically analyzed to get a population average. Since a human observer is the final intended recipient for the class of images considered in this thesis, the MOS has, arguably, the final say since the subjective perceptual quality of an imperfect image is otherwise difficult to characterize in an objective manner. The primary problem with the MOS is that it is time consuming to set up the experiments, find people to participate in rating the images, and then compile the results. Actually, few perceptual studies have been done for progressive image transmission specifically. One study that was conducted to evaluate the perceptual quality of progressive image transmissions was done by Cen *et al.* [CPSC97] to evaluate the differences between progressive JPEG [JPEG99] and SPIHT [SaPe96], as well as between *embedded zerotree wavelet* (EZW) coding [Shap93] and SPIHT [SaPe96]. This study had 25 observers view 118 images in a controlled environment and took roughly 45 minutes for each observer. The setup of such an experiment is time consuming and then requires the statistical analysis of the results to draw any conclusions. Obviously, the

MOS is not feasible for adaptive image compression schemes where the selection of image compression and progressive image transmission parameters are chosen and adapted "on the fly" for each image. This chapter attempts to develop a new objective measure based on multifractal complexity that can be used in such scenarios.

6.1   Introduction

The most commonly used objective measure in the image compression literature for evaluating image reconstruction quality is the peak signal-to-noise ratio distortion measure [CPSC97] as expressed in Eq. 2.10 (or more specifically Eq. 2.11 for 8 bpp images). While PSNR and other similar mean squared error distortion metrics are useful, they are limited in terms of the psychovisual information measured. Considering Eq. 2.10 closer, PSNR is limited to calculating energy differences between corresponding pairs of pixels in the original and reconstructed images. These energy differences are averaged and then formed into a ratio with the largest possible pixel intensity acting as a normalizing factor. The PSNR is limited as a psychovisual measure since

- only energy differences between individual reconstructed pixels are considered

- local and global neighbourhoods of pixel trends are not considered

- the perceptual nature of edges is ignored

- the perceptual nature of textures is completely ignored

Attempts to model the *human visual system* (HVS) a little closer have resulted in more effective quantitative measures. Watson introduced the Cortex transform [Wats87]

as a means of filtering the signal which can then be applied for JND evaluations for image evaluation. Daly's visual difference predictor [Daly93], and Lubin's work on the human vision discrimination model [Lubi93] extend on some of these ideas for JND measuring of image quality. Lu *et al.* [LuAE95] introduced the *picture quality scale* (PQS) as another attempt at modeling the HVS. Work by Jayant *et. al* consider techniques for signal compression based on human perception [JaSJ93] and the HVS for images.

The focus of this chapter is to develop a new quality measure that can analyze image content from global, regional, local, and structural viewpoints through multifractal complexity measures. If this can be done, then this approach will yield a subjective measure that correlates well with psychovisual perception as it relates to MOS experiments.

In the development of an image quality measure, multifractal dimension measures will be considered as the primary approach, since multifractals can characterize the complexity within a signal, such as an image. Multifractal dimension measures perform this characterization at multiple scales/resolutions, so, global, regional, and local neighbourhoods of pixels are considered. With these features, multifractals will allow for an analysis of textures in an image as well. The development of these image quality measures using multifractal complexity measures is presented in the following section.

6.2   Multifractal Dimension Complexity Measures
as Progressive Image Transmission Quality Measures

Through the description and discussion of multifractal dimension measures in Chapter 3, a number of characteristic features should be noticed about these measures that

lead to a new class of objective image evaluation measures. This section considers some of the characteristics of multifractals in the design of image complexity measures as well as performs experiments to determine the feasibility of multifractals as an image quality measure. This work extends some of the research by Ferens and Kinsner [FeKi95], [Fere95] on feature extraction from signals using multifractals where, instead, characterization of the entire image is done to form the image complexity measure. In some sense, the work in this thesis also extends the work in image texture segmentation and classification with the fractal measures of Chaudhuri and Sarkar [ChSa95], [SaCh92] and the pseudo-multifractal measure of Kaplan [Kapl99]. These techniques can be considered as preliminary to an objective measure using multifractals.

Under consideration in this thesis for image complexity measures is the Rényi generalized entropy, the Rényi dimension spectrum, the Mandelbrot spectrum, and finally a generalization of the Kullback-Leibler distance and Rényi information forming a new multifractal measure which will be referred to as the relative Rényi dimension spectrum. The formulation of these measures are presented in Sec. 6.3, Sec. 6.4, Sec. 6.5, and Sec. 6.6, respectively.

When designing an image quality measure, one approach is to develop a measure that characterizes image content, apply this measure to both the original and approximated images, and then compare the results. A similar, but not identical, approach is to find the distortion or difference between the approximated and original images, forming a residual image. Then the image quality measure is applied to this residual image to determine the level of distortion between the two images. The MSE, SNR, and PSNR metrics from

Sec. 2.1.5 are examples of distortion measures since these metrics find the residual between the reconstructed image and the original image. For these three metrics, the energy difference between the two images is found and aggregated on a pixel by pixel basis. A final approach to developing an image quality measure is to form a relative measure that calculates the image quality using a form of ratio or proportional measure between the reconstructed and original images. This calculation may be done, for instance, by forming a rational expression between the reconstructed image and the original image. All of these approaches are investigated to various degrees in the multifractal techniques presented in the rest of this chapter.

6.3   Rényi Generalized Entropy as a Progressive

Image Transmission Quality Measure

This section investigates using the Rényi generalized entropy as an objective measure for use in progressive image transmission techniques. Some of this work was initially presented by Dansereau and Kinsner [DaKi99a]. The Rényi generalized entropy $H_q(X)$, as given in Defn. 3.4 with Eq. 3.23, is a generalization of Shannon entropy evaluated over the moment order q of a probability distribution function $p(x)$. It is known that Shannon entropy $H(X)$, as given in Defn. 3.3, is a measure of the average uncertainty in the random variable X that forms the probability distribution $p(x)$ [CoTh91]. From this description, the Rényi generalized entropy measures the average uncertainty in the random variable X that forms $p(x)$, but over the different orders q.

The uncertainty or "randomness" in a symbol set $\{X\}$ can therefore be measured with the Rényi generalized entropy. This method is the first approach investigated as an objective measure to characterize the level of uncertainty within the image representation. This characterization gives the beginning of a complexity measure for image quality. From the image, an objective measure will be developed by forming a probability distribution $p(x)$ using the pixel values in the image as the symbols in the sequence $\{X\}$.

To test out the properties of this approach to measuring image quality with the Rényi generalized entropy, a number of experiments were performed using a series of eight snapshots from a progressive image transmission. These images are 512×512 8-bit greyscale images and are formed from the two original images of lena and urban, as given in Appendix A. The original images and the first five progressive snapshots for each of these images are displayed in Fig. 6.1 and Fig. 6.2 for lena and urban, respectively. The other three snapshots are not shown since the differences from the original image are nearly imperceptible in the printed form of this thesis.

(a) 344 bytes at 0.131% (8)

(b) 911 bytes at 0.348% (7)

(c) 2,300 bytes at 0.877% (6)

(d) 5,068 bytes at 1.933% (5)

(e) 11,018 bytes at 4.203% (4)

(f) Original image
262,144 bytes (100%)

Fig. 6.1. Progressive transmission of lena. (a) 344 bytes at 0.131%,
(b) 911 bytes at 0.348%, (c) 2,300 bytes at 0.877%, (d) 5,068 bytes at 1.933%,
(e) 11,018 bytes at 4.203%, and (f) original image (100%) with 262,144 bytes.
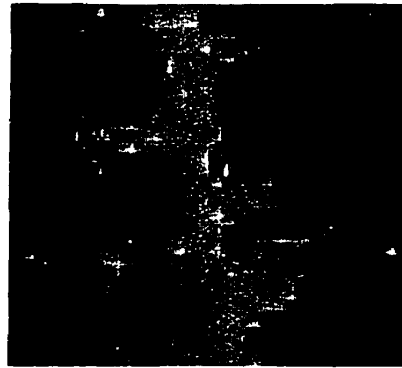
(a) 365 bytes at 0.139% (8)   (b) 886 bytes at 0.338% (7)
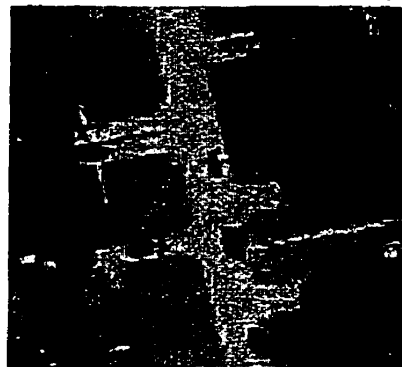
(c) 2,421 bytes at 0.924% (6)   (d) 5,552 bytes at 2.118% (5)
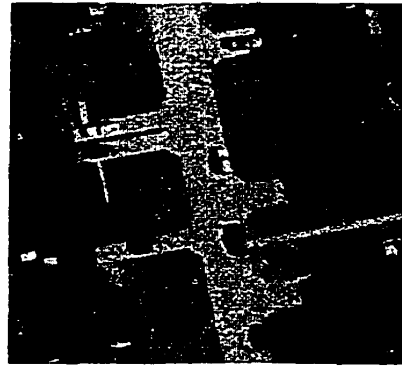
(e) 12,080 bytes at 4.608% (4)   (f) Original image
                                       262,144 bytes (100%)
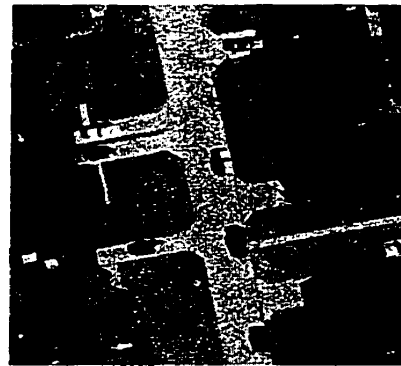
Fig. 6.2. Progressive transmission of urban. (a) 365 bytes at 0.139%,
(b) 886 bytes at 0.338%, (c) 2,421 bytes at 0.924%, (d) 5,552 bytes at 2.118%,
(e) 12,080 bytes at 4.608%, and (f) original image (100%) with 262,144 bytes.

These two images are chosen for the experiments since they compress relatively little from an information theoretic viewpoint. Lena and urban have Shannon entropies of 7.45 bpp and 7.65 bpp, respectively, which limits their first-order lossless compression to only 1.074:1 and 1.046:1, respectively. Hence, these images can roughly be classified as tough to compress. These images, therefore, serve as good examples of difficult images to compress that would benefit from a progressive form of compression, transmission, and reconstruction. The progressive image snapshots in Fig. 6.1 and Fig. 6.2 were generated using the embedded zero-tree coding algorithm of Shapiro [Shap93] with a final lossless compression stage using Witten, Neal, and Cleary's adaptive arithmetic coding [WiNC87]. At this point the exact workings of embedded zero-tree coding and adaptive arithmetic coding are not needed since the current goal is to develop objective measures using this progression of images solely as sample test images.

The first set of experiments consists of calculating the Rényi generalized entropy on the original image as well as each of the image snapshots extracted from the progressive wavelet image transmission. This approach falls in line with an objective measure applied separately on the reconstructed and original images as discussed in Sec. 6.2. The main difficulty to answer when using the Rényi generalized entropy $H_q(X)$ is how exactly to form the probability distribution $p(x)$. The approach tried is to form a histogram for the image and normalize it to form $p(x)$. Thus, $p(x)$ is formed by calculating the relative frequency for each pixel greyscale value in the image.

The results of the Rényi generalized entropy $H_q(X)$ measures using this $p(x)$ formulation are plotted for lena and urban in Fig. 6.3a and Fig. 6.3b, respectively. Plotted in

Fig. 6.3 is the Rényi generalized entropy $H_q(X)$ versus moment order $q \in [-20.0, 20.0]$ for the original images and each of the eight progressive image snapshots. It should be noted from the definition of $H_q(X)$ with Eq. 3.23 that $H_q(X)$ is a continuous function in q. Therefore, the plots in Fig. 6.3 are continuous. Also, since the pixel values are integral, the only errors in this measurement are numerical round-off errors due to double floating point precision math, which are insignificant for these plots.

An important observation about the plots in Fig. 6.3 is that when $q = 1$ then $H_{q=1}(X) = 7.45$ bpp for the original lena image and $H_{q=1}(X) = 7.65$ bpp for the original urban image. These are precisely the values calculated for Shannon entropy as is expected. This agreement helps verify that the $H_q(X)$ calculations are correct in light of Lemma 3.1. It is also observed that all of the curves are monotonically non-increasing as Lemma 3.4 requires.

From the Rényi generalized entropy plot for lena in Fig. 6.3.a, it is apparent that there is no significant changes in $H_q(X)$ across the eight reconstructed images compared to the original image during the progressive transmission. This observation likely means that the histograms of the different image reconstructions do not differ greatly from the histogram of the original image. Unfortunately, this property does not assure that the histograms of the reconstructed image matches that of the original image. There can be a shift in the entire histogram for a reconstructed image that would result in the same $H_q(X)$ versus q curves since $p(x)$ would effectively be the same from the viewpoint of $H_q(X)$.
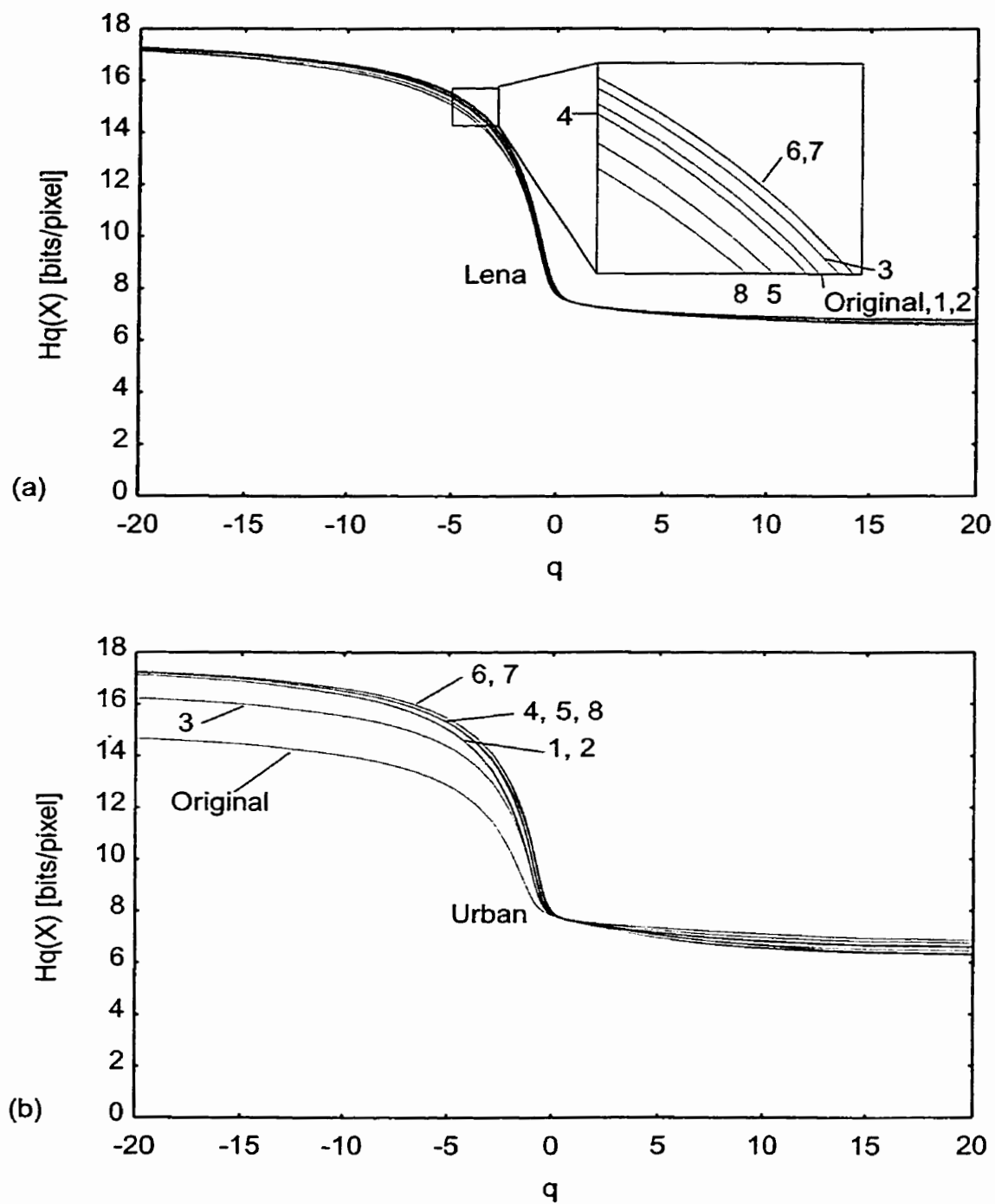
Fig. 6.3. Rényi generalized entropy on the progressive image reconstructions.
(a) lena and (b) urban.

This problem of a shifted histogram occurs with the reconstructions of the urban image. As Fig. 6.3b shows, the $H_q(X)$ versus q curves for the reconstructed images are not well aligned with the $H_q(X)$ versus q curve for the original image. Considering Fig. 6.3.b, when $q < 0$ it is evident that the progressive model of the image does not follow the original urban image's $H_q(X)$ versus q curve as closely as it did for lena in Fig. 6.3a. For the original image of urban, the curve when $q < 0$ is well below that for any of the reconstructed versions of the urban image. An increase in the $H_q(X)$ versus q curve such as this for the reconstructed images suggests that the probability distributions are flattened for the reconstructions, or the probability distributions for the original and reconstructed images are not lining up as well, or that the probability distributions for the original and reconstructed images are quite different. This difference is not directly verifiable with only the Rényi generalized entropy calculations, but, when put in conjunction with image histogram analysis and considering the perceptual quality of the images, then these three scenarios can be checked. A histogram of the original image of urban along with the last five image reconstruction steps from the progressive image transmission is produced and is plotted in Fig. 6.4. The histogram for the original image of urban is stretched and shifted to the right compared with the histograms of the other reconstructed images. This discrepancy confirms that this progressive image transmission model for approximating images does not work as well as desired for the image of urban as it appears to work for the image of lena. It also indicates that the lack of clustering of the curves in Fig. 6.3b does not result from a problem with the Rényi generalized entropy as an objective measure.
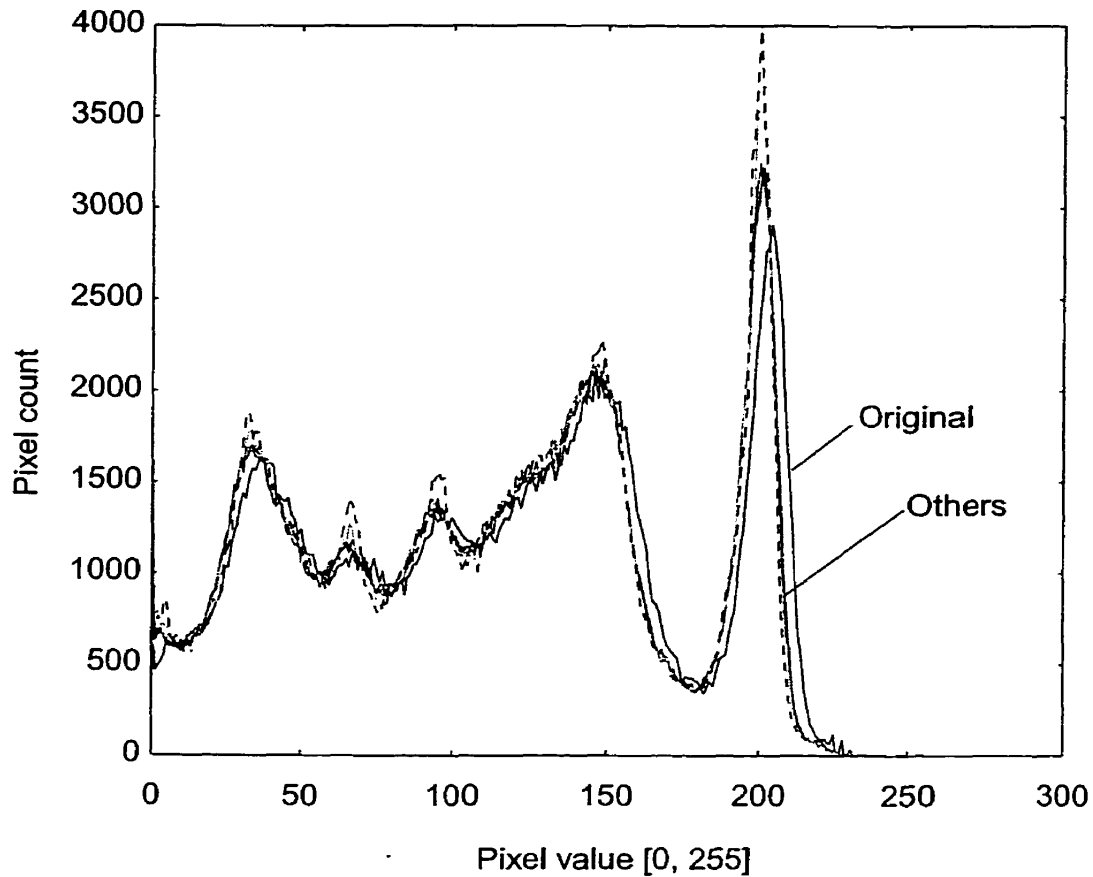
Fig. 6.4. Histogram for the image of urban and five approximations.

The next set of experiments is to perform the Rényi generalized entropy analysis on residual images formed from reconstructed and original images as is suggested in Sec. 6.2. For these experiments, the absolute differences on a pixel by pixel basis between the progressively reconstructed images and the original image are found. The Rényi generalized entropy $H_q(X)$ is then computed on these residual images. The probability distribution $p(x)$ is again set as before from the normalized histogram, but now on the residual image. The $H_q(X)$ experimental results using these residual images for the lena and urban images are plotted in Fig. 6.5a and Fig. 6.5b, respectively.

The plots in Fig. 6.5 show a larger spread in the $H_q(X)$ versus q curves for this residual approach compared with those in Fig. 6.3. An interesting observation for the curves in Fig. 6.5a and Fig. 6.5b is that for $q \geq 0$ the curves are in descending order according to the respective progressive reconstructed image. This observation can be used as a criterion for an objective measure for progressive image transmission since it would generally occur as the range of values in the residual image decreases so as to decrease $|\chi|$ from Theorem 3.4. This decrease would then limit the maximum value of $H_q(X)$ as the residual image has fewer unique values. It must be noted that this observed decrease in $H_q(X)$ for Fig. 6.5a and Fig. 6.5b is not followed for lena or urban when $q < 0$. In this case, the $H_q(X)$ curves cross each other at different values of q. This fact may also be useful in analyzing the quality of a progressive image transmission.
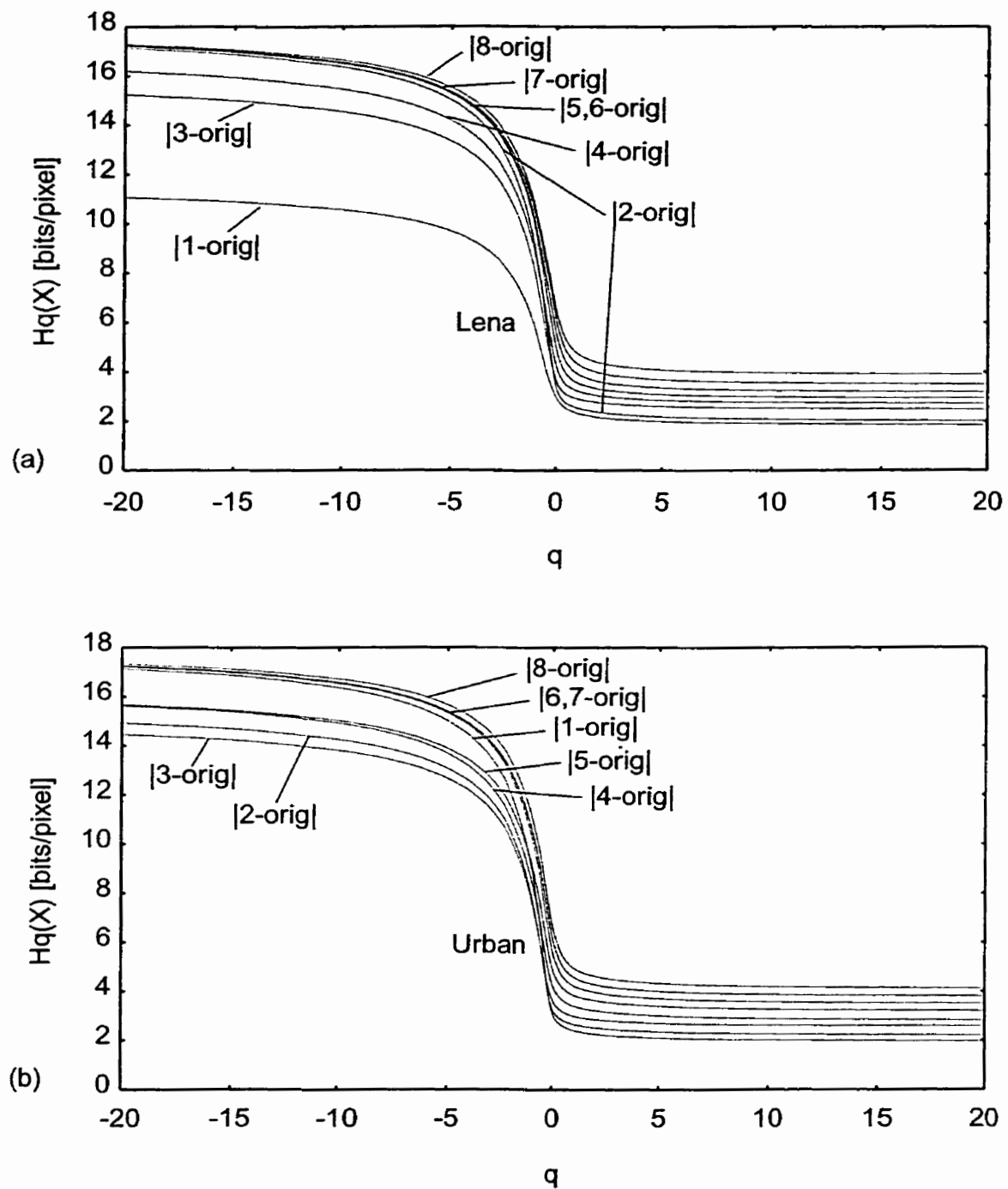
Fig. 6.5. Rényi generalized entropy on residual images from the progressively recon-
structed images and the original image for (a) lena and (b) urban.

The last set of experiments using the Rényi generalized entropy is similar to the previous, but, instead the residuals are calculated between one snapshot of the approximated image to the next successive snapshot in the progression of the approximated image. This procedure allows the image reconstruction to be viewed from the perspective of a stage to stage addition of image information throughout the progressive image transmission. The same Rényi generalized entropy calculations were performed with these new residual images for the lena and urban approximations, and the results are plotted in Fig. 6.6a and Fig. 6.6b, respectively.

It is observed from the two plots in Fig. 6.6 that the $H_q(X)$ versus q curves corresponding to the earlier progressive image snapshots tend to cluster together. Also, as the progression continues, the residual images tend to decrease the resulting $H_q(X)$ value. The clustering, or higher values of $H_q(X)$, is mostly due to the greater number of unique values in the residual image resulting in a larger $|\chi|$. Again, according to Theorem 3.4 this range allows for a high value of $H_q(X)$. The clustering at the beginning of the progression also indicates that the improvements to the image reconstructions produce similar residual images from step to step. As the progression continues, the refinement turns out to be more fine tuning as the number of unique values in the residuals decreases, causing the possible maximum of $H_q(X)$ to decrease.
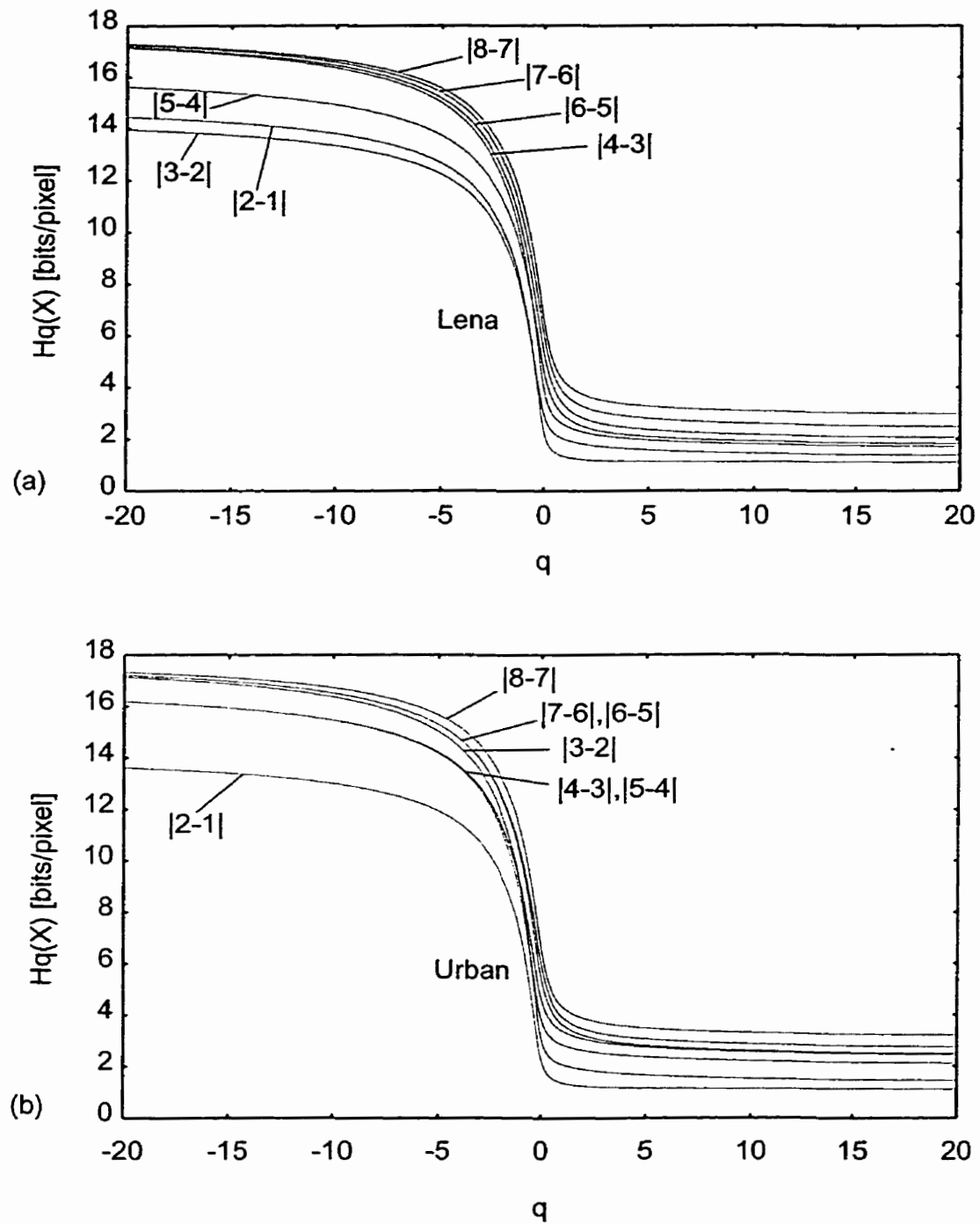
Fig. 6.6. Rényi generalized entropy on residual images between successive
image reconstructions for (a) lena and (b) urban.

From these three sets of experiments, it is concluded that the Rényi generalized entropy contains some indicators as to the progression in quality for the image transmissions. The results shown, unfortunately, do not give an obvious lead into developing an encompassing objective measure, since convergence or movement of the curves is not consistent on either side of $q = 0$. The main limitation with using the Rényi generalized entropy in this manner is that the images are analyzed only from the viewpoint of a normalized histogram. Many completely different images can form the same normalized histogram. So, the image quality information retained within the normalized histogram is too limited and not really enough for the Rényi generalized entropy to form a proper model of the image approximations. Also, the histogram approach eliminates any pixel neighbourhood or spatial correlations, which is a key factor in the human visual system. Therefore, while some interesting results are obtained with the Rényi generalized entropy, it can only serve as an indicator to some image quality characteristics and, hence, has limited use as a complete image quality measure for progressive image transmission.

The next subsection continues the experiments, now with the Rényi dimension spectrum which overcomes some of the limitations of the Rényi generalized entropy. These benefits are achieved by looking at the image at multiple scales and with probability distributions suited more to measuring the image texture complexity.

6.4   Rényi Dimension Spectrum as a Progressive

Image Transmission Quality Measure

This section investigates how the Rényi dimension spectrum $D_q(X)$ can be used as a progressive image transmission quality measure and follows some of the initial work done by Dansereau and Kinsner [DaKi98b], [DaKi99b]. The Rényi dimension spectrum $D_q(X)$, as given with Eq. 3.45 in Defn. 3.5, affords many analytical advantages over those of the Rényi generalized entropy alone. As is described in Sec. 3.3.4, the Rényi dimension spectrum is a multifractal measure that measures the fractality of an object $p_s(x)$ resolved at scale s at specified orders of q. For each order q, the fractality of the object at that scale is measured and the critical exponent D_q is found satisfying Eq. 3.44. Therefore, the rate of change in the object's complexity at different scales is measured. This idea can be applied as a measure of the complexity within an object such as an image to determine the image's complexity over different scales. This concept is different from the Rényi generalized entropy $H_q(X)$ since, instead of only local calculations within the image, the measurement is done in a multiresolution fashion so that local, neighbourhood, and global measurements of the image are all performed, resulting in the final value for $D_q(X)$. In addition, varying the order q in the $D_q(X)$ calculation changes the extent of the contribution of different inhomogeneous fractal complexities within the image, as discussed in Sec. 3.3.4. Using these characteristics, this section addresses the use of the Rényi dimension spectrum $D_q(X)$ as a progressive image transmission quality measure.

The first issue to address is how the probability distribution $p_s(x)$ in Eq. 3.45 for $D_q(X)$ is calculated for an image over the different scales s. In the case of the Rényi gen-

eralized entropy $H_q(X)$ from Sec. 6.3, the probabilities $p(x)$ are formed from the normalized histogram of the image. This normalized histogram does not readily incorporate the desired property of measuring the image over multiple scales so that the critical exponent D_q can be determined over the range of scales. For this section, the probability distribution $p_s(x)$ at scale s is calculated by summing the greyscale intensity values in a particular vel as a proportion of all greyscale intensity values in the image as follows

$$p_s(x) = \frac{\sum\limits_{i \in B_x} (\text{pixel}(i) + 1)}{\sum\limits_{i \in \forall B} (\text{pixel}(i) + 1)} \tag{6.1}$$

where $\text{pixel}(i)$ is a pixel in vel B_x of the image covering and B is the set of all vels. Note that the summations in Eq. 6.1 include the extra term of 1 to bring the greyscale range of an n-bit unsigned integer pixel depth from $[0, 2^n - 1]$ to $[1, 2^n]$ or in the case of the 8-bit greyscale images used in this thesis from $[0, 255]$ to $[1, 256]$. This addition is done to ensure that no vel results in a probability of $p_s(x) = 0$. While there is no mathematical reason that $p_s(x)$ cannot be zero, it is useful to think of the multifractal measure being applied to a surface in a three dimensional space (recall the surface projection from Fig. 5.10b). If $p_s(x) = 0$ is allowed then the surface can be thought of as having holes where $p_s(x)$ is zero. The addition of 1 to each greyscale pixel value ensures that no holes

exist in the surface. For $q = 0$, having no holes in the surface results in the Rényi dimension spectrum being

$$D_{q=0}(X) = \lim_{s \to \infty} \frac{\log\left(\dfrac{\sum\limits_{x \in \chi_s} p_s^0(x)}{\sum\limits_{x \in \chi_s} p_s(x)}\right)^{(1)^{-1}}}{\log\left(\dfrac{1}{s}\right)} = \lim_{s \to \infty} \frac{\log\left(\dfrac{\sum\limits_{x \in \chi_s} 1}{\sum\limits_{x \in \chi_s} p_s(x)}\right)}{\log\left(\dfrac{1}{s}\right)} = 2 \qquad (6.2)$$

since the rate of growth of the number of vels is 4 times the rate of growth of the scale, and hence 2 times when the logarithms are taken, such as with measuring the square in Fig. 3.6. If holes are allowed in the surface, then $D_{q=0}(X)$ would be less than 2 and would make the objective measure a harder to interpret in the event of $p_s(x) = 0$.

The experiments to follow are performed on the progressive snapshots of the images of lena and urban as described in Sec. 6.3. The first set of experiments is to calculate the Rényi dimension spectrum $D_q(X)$ for $q \in [-20.0, 20.0]$ directly on the progressive image reconstructions of the two test images. These results of applying $D_q(X)$ on the original image and eight progressive reconstructed images are plotted in Fig. 6.7a and Fig. 6.7b for lena and urban, respectively.
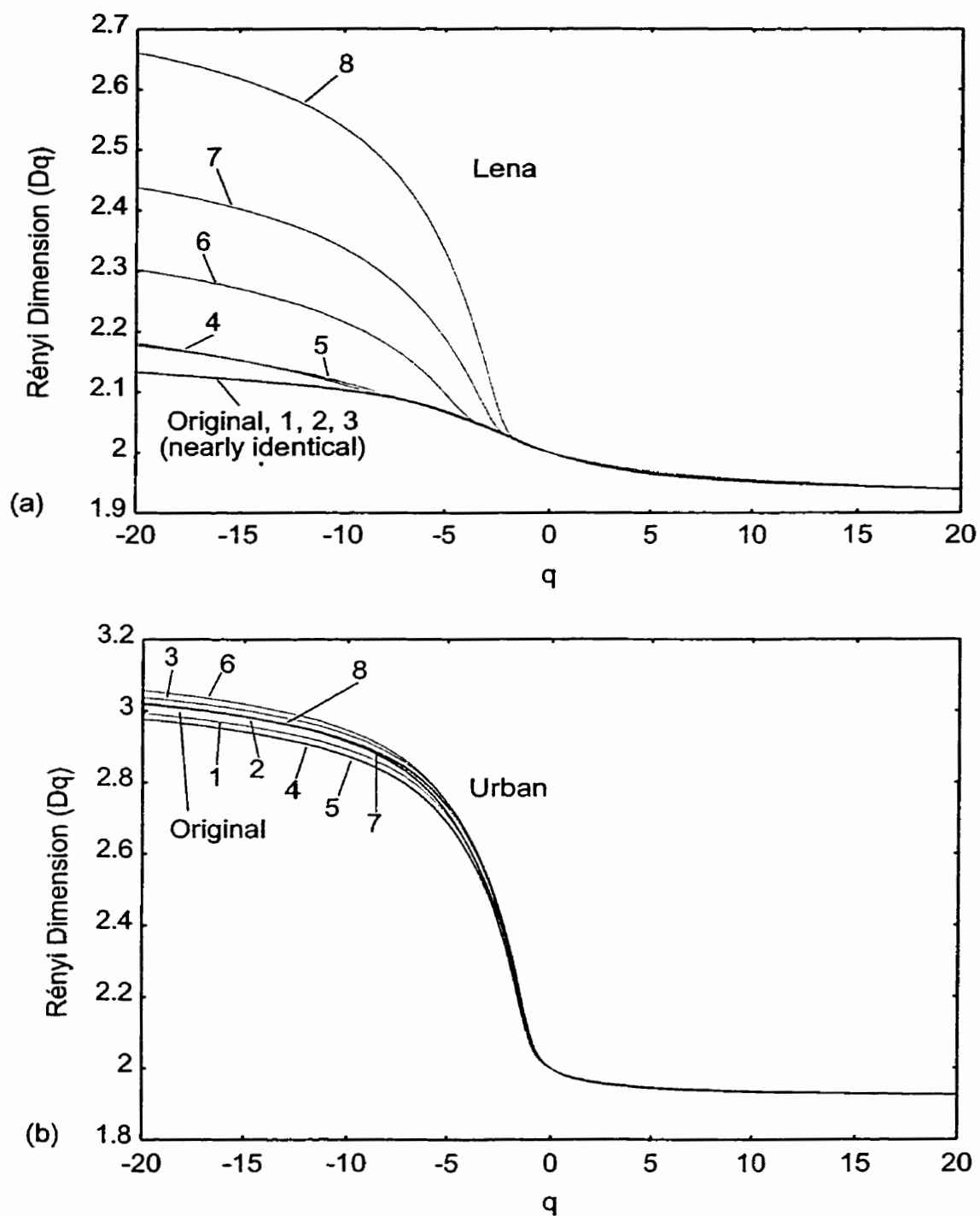
Fig. 6.7. Rényi dimension spectrum experimental results at progressive
steps for the image of (a) lena and (b) urban.

Looking at the results for lena in Fig. 6.7.a, it is noticed that there is a larger spread among the $D_q(X)$ curves for $q < 0$. This variation suggests that this particular progressive image transmission is not as representative of the multifractal complexity in the original image at the earlier stages of the transmission. It is interesting to notice that as the image transmission progresses, the $D_q(X)$ versus q curves do converge to the $D_q(X)$ versus q curve of the original image. This result is reasonable since as the progressive image transmission occurs, the reproductions of the image improve. Therefore, the measure clearly indicates how the complexity within the image is represented better as the transmission occurs.

Looking at the results for the image of urban in Fig. 6.7b, it is noticed that the $D_q(X)$ versus q curves are clustered together. This clustering means, from a multifractal perspective, that the progressive reproductions of the image during transmission have similar complexity. This result follows the goal of maintaining that same multifractal complexity throughout the image reconstruction. Unfortunately, from an objective measure viewpoint, Fig. 6.7b does not offer a view into the improvement in quality of the image reconstructions.
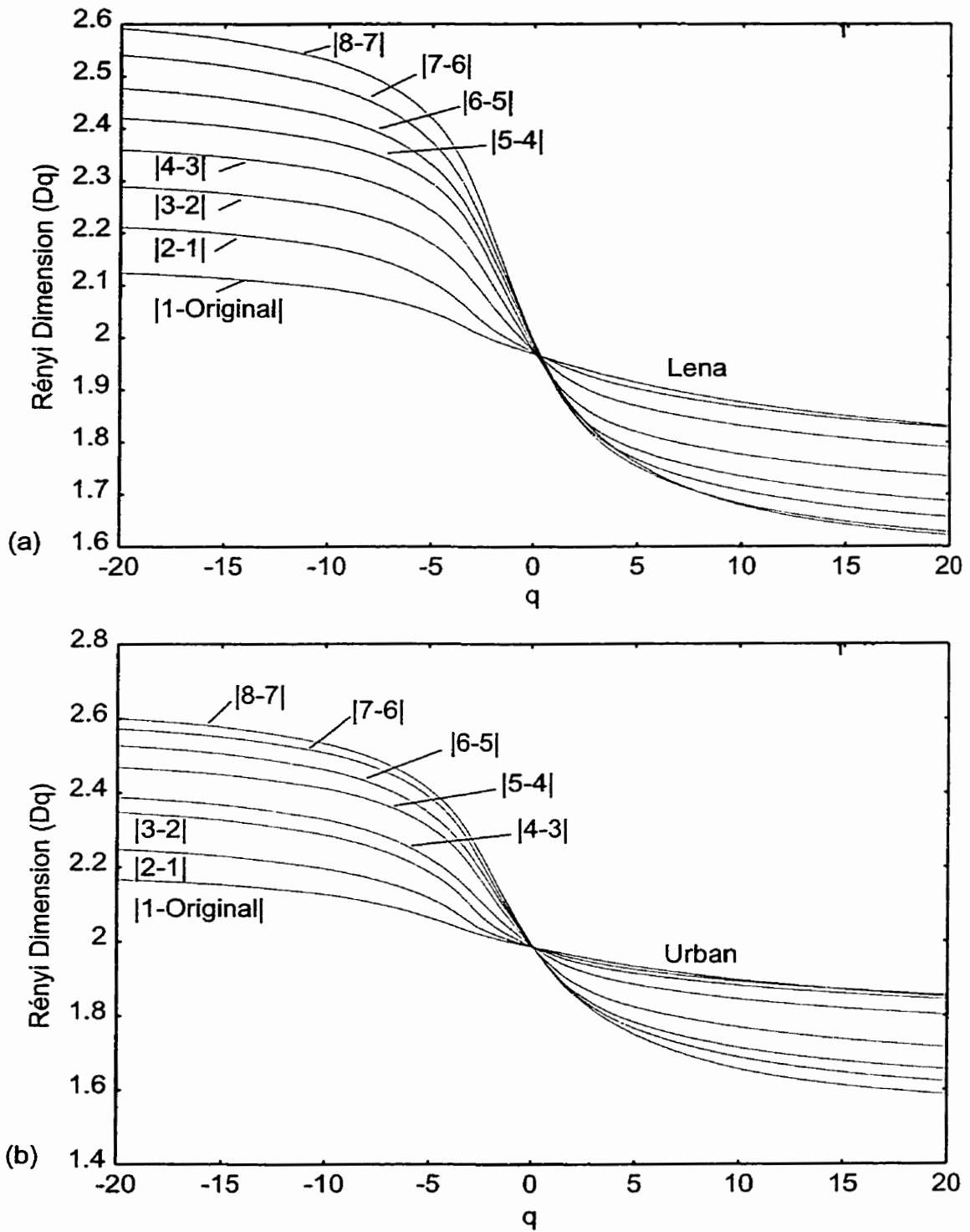
Fig. 6.8. Rényi dimension spectrum calculations on residuals of successive images in the progressive image transmission for the image of (a) lena and (b) urban.

The next set of experiments conducted investigates what type of complexity characteristic is present in the absolute residual images between successive image reconstructions. This calculation is done for successive image reproductions with both lena and urban using the Rényi dimension spectrum $D_q(X)$, and the results are plotted in Fig. 6.8a and Fig. 6.8b, respectively. The most important point to notice about both plots in Fig. 6.8 is that as the progressive image transmission occurs, the $D_q(X)$ versus q curves for the residual images become flatter. This flattening of the $D_q(X)$ versus q means there is a reduction in the multifractal complexity between successive reconstructed images, since the curve approaches the topological dimension of 2. A reduction in the multifractal complexity between successive reconstructed images means that there are fewer different features between successive reconstructed images. Hence, the overall image quality improves as more data is reconstructed into the image. These results give a first glimpse into what could be used as an objective measure, where the goal is to flatten out the $D_q(X)$ versus q curve to the topological dimension of 2 as quickly as possible for the residual images.

When considering what is desired, an optimal progressive image transmission would transmit the main multifractal complexity within the image very early in the transmission process. Therefore, the expected result would be that the $D_q(X)$ versus q curves in the residual images would flatten as more image information is incorporated into the image reconstruction. This flattening of the $D_q(X)$ versus q means that $D_{-\infty}(X)$ and $D_{\infty}(X)$ in the residual images are converging as the progressive image transmission continues and hence using Lemma 3.5 and Lemma 3.6, the probability range for $\sup(p_s(x))$ and $\inf(p_s(x))$ is decreasing. This phenomenon is witnessed in both plots in Fig. 6.8 for

the images **lena** and **urban**. This result suggests not only that the Rényi dimension spectrum $D_q(X)$ can be used as an objective measure by following the rate at which the $D_q(X)$ versus q curves flatten, but, also the rate at which $D_{-\infty}(X)$ and $D_\infty(X)$ converge versus the image data rate in the progressive transmission.

With the Rényi dimension spectrum $D_q(X)$ as a promising objective measure, the question now is how to encapsulate the $D_q(X)$ versus q curves into a single value. This single value will help indicate the image quality of the reconstruction at any phase of the progressive image transmission. Different approaches can be taken, but one simple approach is to calculate the *root mean squared error* (RMSE) of the Rényi dimension spectrum for the original image and the reconstructed image as follows

$$\text{RMSE}(D_{q_{orig}}(X), D_{q_{rec}}(X)) = \left[\lim_{T \to \infty} \frac{1}{T} \int_{\frac{-T}{2}}^{\frac{T}{2}} |D_{q_{orig}}(X) - D_{q_{rec}}(X)|^2 dq \right]^{1/2} \qquad (6.3)$$

If $D_{q_{rec}}(X)$ and $D_{q_{orig}}(X)$ are calculated for a total of n specific values of q then Eq. 6.3 can be approximated by

$$\text{RMSE}(D_{q_{orig}}(X), D_{q_{rec}}(X)) = \left[\frac{1}{n} \sum_{i=1}^{n} |D_{q_{orig}(i)}(X) - D_{q_{rec}(i)}(X)|^2 \right]^{1/2} \qquad (6.4)$$

Using Eq. 6.4, a single value results for the objective measure of a specific image reconstruction compared with the original image. To test this objective measure, a number of experiments can be performed by applying this measure over a set of image reconstructions. Since Eq. 6.4 returns only one value per image reconstruction, a greater range of image reconstructions can be analyzed and visualized above that of the eight image reconstructions previously used from Fig. 6.1 and Fig. 6.2. For ease of duplication of these

experiments, the approach taken for generating the image reconstructions is to take the original image, perform the DWT using the Daubechies 4-tap mother wavelet [Daub92] (see Sec. 4.3.3), and then apply a hard-threshold in the wavelet domain to form approximations of the image. The hard-threshold is applied in the wavelet domain by choosing I_k in Eq. 4.12 of Sec. 4.1.3 as the set of wavelet coefficient indices for those wavelet coefficients that have a magnitude greater than the hard-threshold value. In essence, this sets any wavelet coefficient with a magnitude less than or equal to the hard-threshold to zero, which results in an approximation of the image in the wavelet domain using Eq. 4.12. For the experiments performed in this section, the range of hard-threshold values chosen are the integers in the range $[1, 250]$ which results in a total of 250 progressively reconstructed image approximations.

Before presenting the experimental results, a few notes should be made about the accuracy of the $D_q(X)$ calculations that were noticed during this process. The calculation of $D_q(X)$ is done by finding the critical exponent in the proportionality given in Eq. 3.44. This critical exponent is calculated by finding the limiting slope in a log-log plot using the numerator $H_q(X)$ and the denominator $\log(1/s)$ from Eq. 3.45 for the ordinate and abscissa axes, respectively. The difficulty with this process is finding the slope within this log-log plot and turns out to be the source of error in the $D_q(X)$ calculations. Initial implementations for finding the limiting slope in the log-log plot used the least-squares line fitting algorithm. It is found that for the purposes of calculating $D_q(X)$, that the least-squares line fitting algorithm is too sensitive to small deviations by outliers at coarser scales of s. To combat this sensitivity, the line fitting algorithm was switched to a more robust technique as described in Numerical Recipes for C for fitting a line by minimizing

absolute deviation [PTVF92]. This alternative method greatly improves the line fitting results obtained and helps remove small anomalies that are present because of the inaccurate line fitting by the least-squares algorithm. Another detail to help improve the line fitting results is to increase the number of points with which to perform the line fitting. The initial implementation of the $D_q(X)$ calculations went over the range of scales s in a dyadic   form   (i.e.   for   a 512×512 image   the   dyadic   set $\{1, 2, 4, 8, 16, 32, 64, 128, 256, 512\}$ for s). This was changed so that any scale within the entire size of the image could be selected. The experiments to follow used the set of scales $\{1, 2, 4, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 32, 36, 40, 44, 48, 52, 56, 60, 64,$ $80, 96, 112, 128, 192, 256, 512\}$ for the value of s. Of course, more scale measurements can be used but this set seems sufficient for the $D_q(X)$ calculations since the points fall in a relatively straight line and there are enough to counter-balance a few outliers when using the robust line fitting technique.

Using these improvements to the $D_q(X)$ calculations and with the set of 250 progressive image reconstructions for each image, Fig. 6.9, Fig. 6.10, Fig. 6.11, Fig. 6.12, and Fig. 6.13 were produced for the images of lena, urban, baboon, peppers, and farm, respectively (refer to Appendix A for the original images). These plots in Fig. 6.9 through Fig. 6.13 show the RMSE calculations from Eq. 6.4 versus the hard-threshold value used in the wavelet domain to approximate the image from a Daubechies 4-tap DWT. The purpose of these figures is to investigate the relationship between the RMSE of the $D_q(X)$ versus q curves and the image reconstructions from coarser approximations to more complete approximations according to Eq. 4.12. The plots in Fig. 6.9 through Fig. 6.13 have been combined in Fig. 6.14 to see the range of all of the curves together.
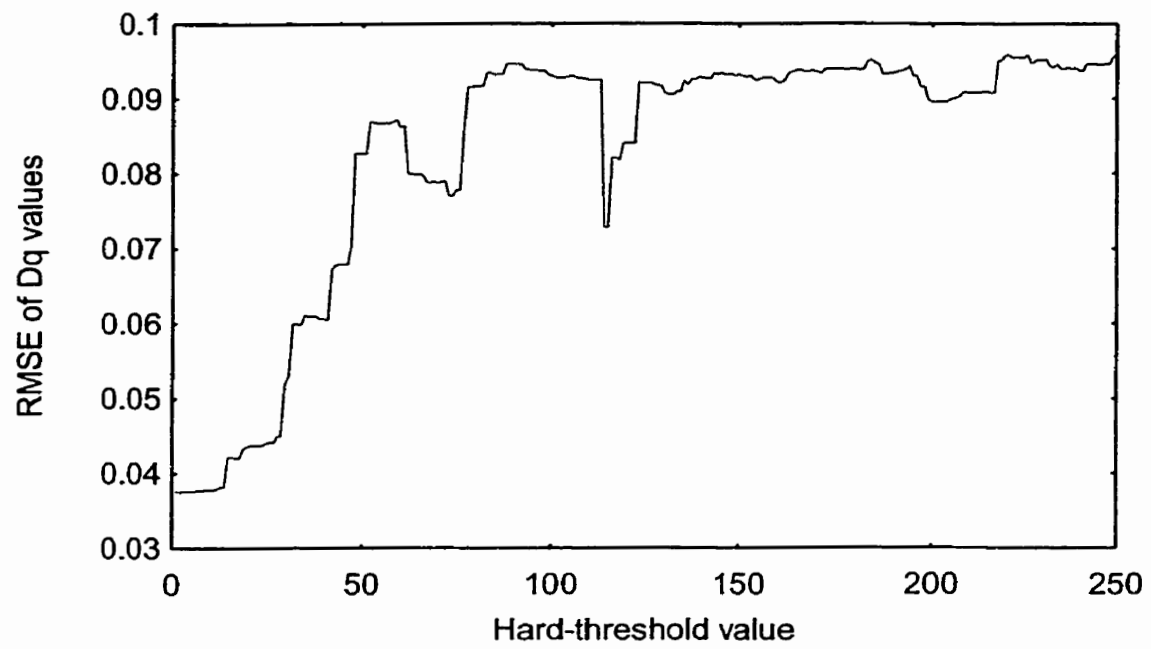
Fig. 6.9. RMSE tracking of Dq for lena (Daub4).
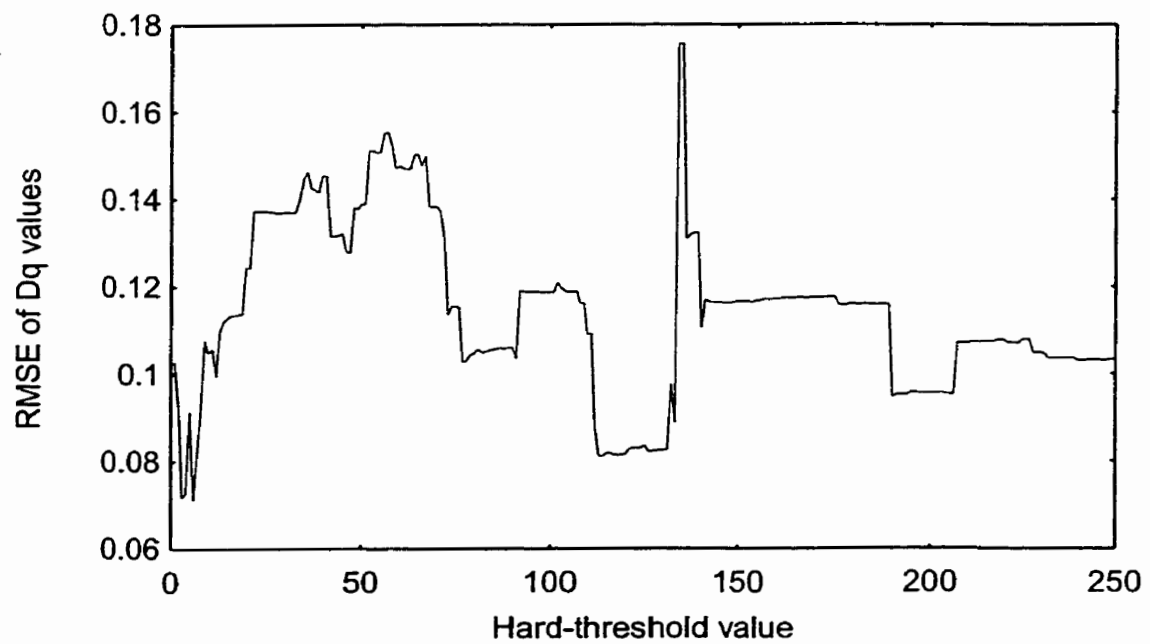
Fig. 6.10. RMSE tracking of Dq for urban (Daub4).
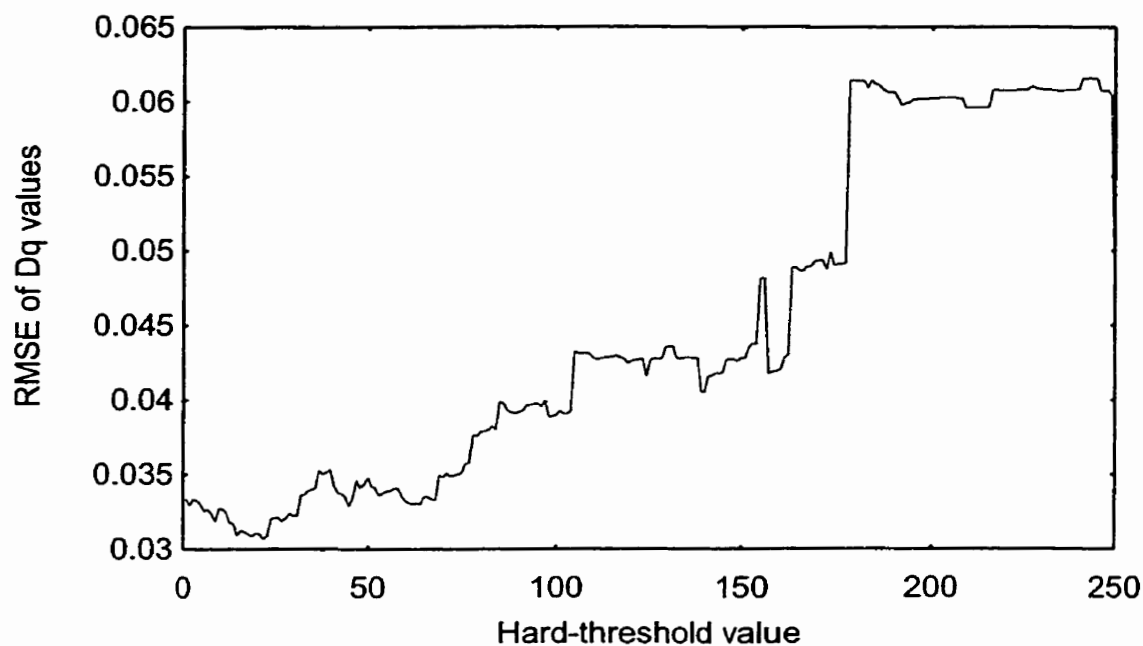
Fig. 6.11. RMSE tracking of Dq for baboon (Daub4).
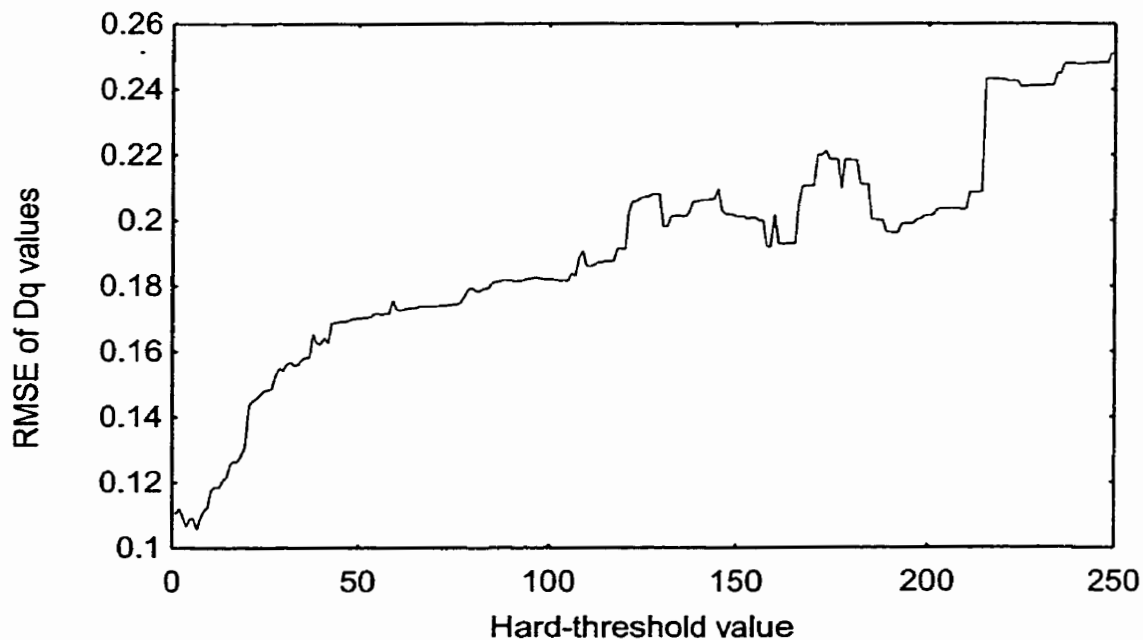
Fig. 6.12. RMSE tracking of Dq for peppers (Daub4).
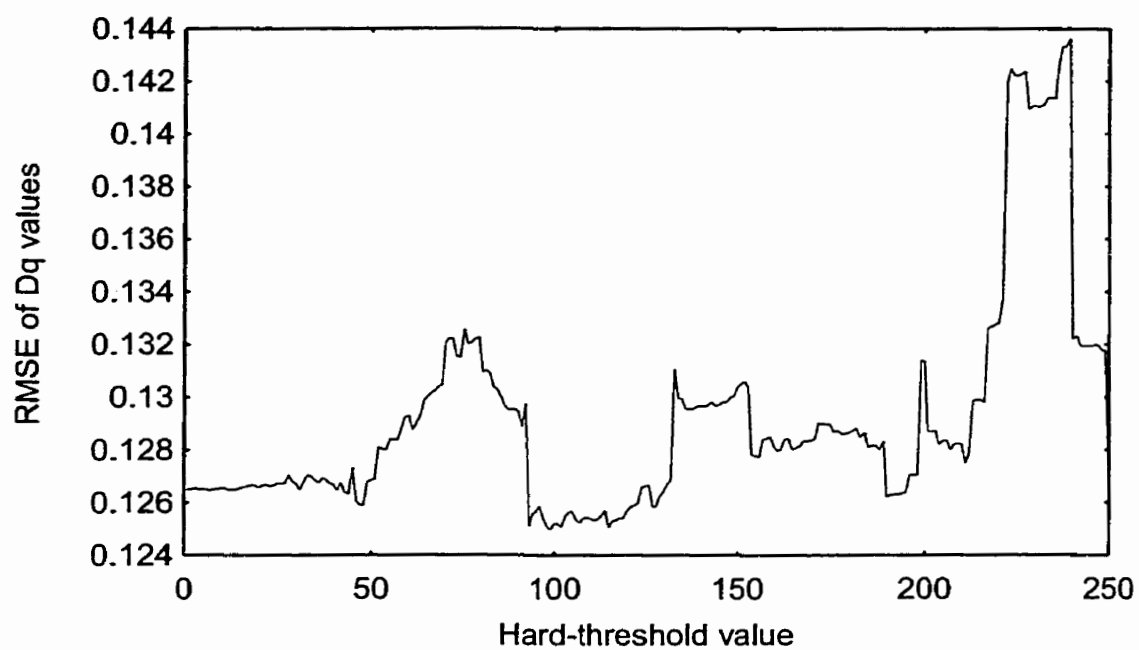
Fig. 6.13. RMSE tracking of Dq for farm (Daub4).
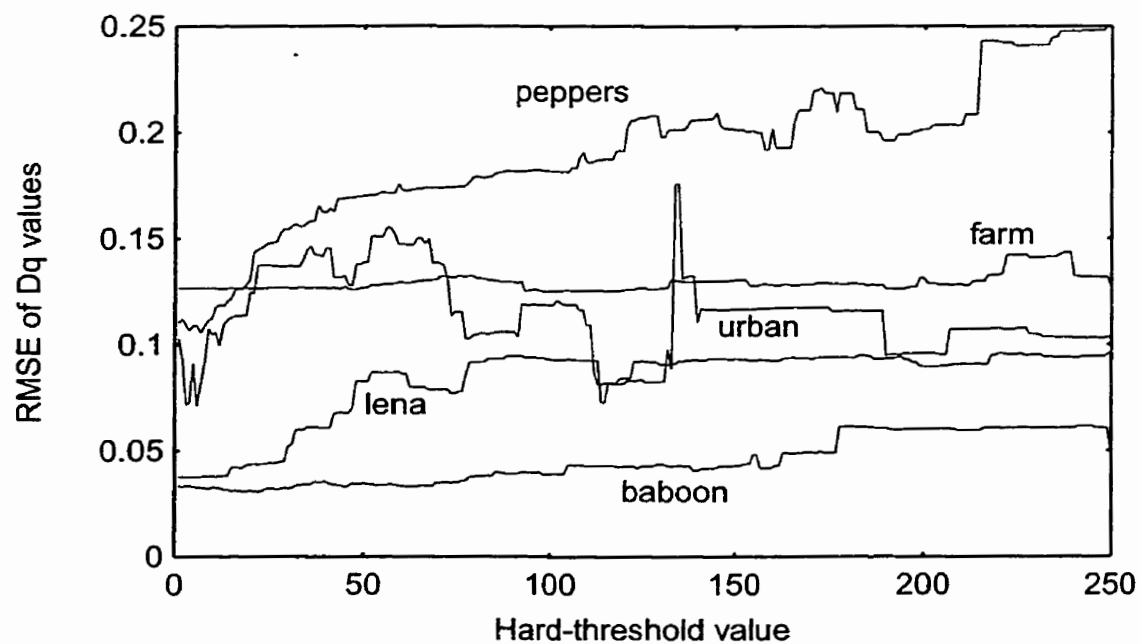
Fig. 6.14. RMSE tracking of Dq for all five test images (Daub4).

Looking at the RMSE results for lena, baboon, and peppers in Fig. 6.9, Fig. 6.11, and Fig. 6.12, respectively, it is noticed that the error in the $D_q(X)$ values approach smaller values as the hard-threshold value decreases. This decrease in $D_q(X)$ error with the hard-threshold value is desired since this correlates with what is intuitively expected for an image quality measure, as more image information is added for the image reconstructions. This decrease in $D_q(X)$ error indicates that the multifractal complexity of the reconstructed images approaches the multifractal complexity of the original image. Unfortunately, this decrease in $D_q(X)$ error values is not monotonic and there are definite sharp transitions and spikes in the changing $D_q(X)$ error values. For instance, in Fig. 6.9 there is a sharp change in the $D_q(X)$ error value at a hard-threshold of 114.

This same tendency for $D_q(X)$ error to decrease with the hard-threshold is not apparent for the images of urban or farm in Fig. 6.10 and Fig. 6.13, respectively. Figure 6.10 shows a multifractal complexity error that tends to change more haphazardly. Figure 6.13 shows a complexity error that remains roughly constant throughout the image reconstruction after it hits roughly the 225 hard-threshold mark.

A question arising from the plots in Fig. 6.9 through Fig. 6.13 is whether there is any perceptual significance to the sharp changes and spikes in the experimental results. To address this matter, an informal mean opinion score experiment was conducted to determine whether any perceptual correlations can be made with the Rényi dimension spectrum $D_q(X)$ for the experiments run on the five images.

The MOS experiment followed the protocol described in Sec. B.1 of Appendix B. This protocol defines how the experiment is directed in which an original image is shown

to the participants, followed by an approximation of the image. The participants then rank the perceived level of distortion in the approximated image compared to the original image using the 5 point scale given in Sec. B.2 of Appendix B. This MOS experiment was held over four sessions with a total of 32 participants and a 50 set of original image/ approximated image pairs were presented.

As a baseline, 5 sets of image originals/approximations were repeated once at random throughout the MOS experiment. The purpose for repeating some images is to determine the range of MOS values obtained when the same approximated image is presented. On the 5 point scale, an average of a 0.135 difference in MOS value was found for identical images that were presented. This result means that differences in MOS values on single tests are likely statistically similar if they fall within this range of difference.

The first point of interest from the MOS experiments is whether any noticeable perceptual difference between images exists when a large change in the RMSE for $D_q(X)$ values occur for close hard-threshold values. For instance, the sharp spike in Fig. 6.9 at a hard-threshold of 114, as previously discussed, was tested. The approximation of lena with a hard-threshold of 113 and a hard-threshold of 114 were presented to the MOS experiment participants. A difference in MOS value of 0.25 was found in this case. This value is greater than the 0.135 average difference found for identical images, but, not a large step in terms of the 5 point MOS range. Other pairs of hard-threshold values for the other image approximations were also tested in the MOS experiment and these results are listed in Table 6.1.

**Table 6.1 MOS value differences for hard-threshold approximations
of the original image.**

Original Image	Two Hard-Threshold Values Tested	MOS Obtained for Hard-Thresholds	MOS Value Difference
baboon	156, 163	2.00, 2.02	-0.02
baboon	178, 250	1.72, 1.61	0.11
farm	92, 93	2.40, 2.48	-0.08
farm	132, 133	2.08, 1.94	0.14
farm	201, 221	1.23, 1.25	-0.02
farm	221, 222	1.25, 1.17	0.08
lena	76, 77	3.21, 3.21	0.00
lena	113, 114	2.61, 2.36	0.25
peppers	214, 215	2.16, 1.82	0.34
urban	21, 22	4.87, 4.78	0.09
urban	133, 190	2.11, 1.84	0.27

The results in Table 6.1 show that most of the sharp change in the Rényi dimension spectrum have a difference on the MOS values of less than the 0.135 average noticed for identical images. This result suggests that from a perceptual point of view, the sharp changes in the Rényi dimension spectrum $D_q(X)$ error values versus hard-threshold value are nearly imperceptible. Therefore, it can be concluded that small changes in the hard-threshold value has only a small effect on the result of the MOS value for perceptual distortion. This fact is important to note since at many of the hard-threshold values listed in Table 6.1 there are sudden jumps in the respective RSME plots for that image. It must also be noted from Table 6.1 that results typically show that the image approximation with more image information (lower hard-threshold) has a slightly higher MOS value than the

approximations with less image information as indicated by the positive MOS value differences. All the negative differences in MOS value fall below the 0.135 MOS value average for identical images presented to the participants.
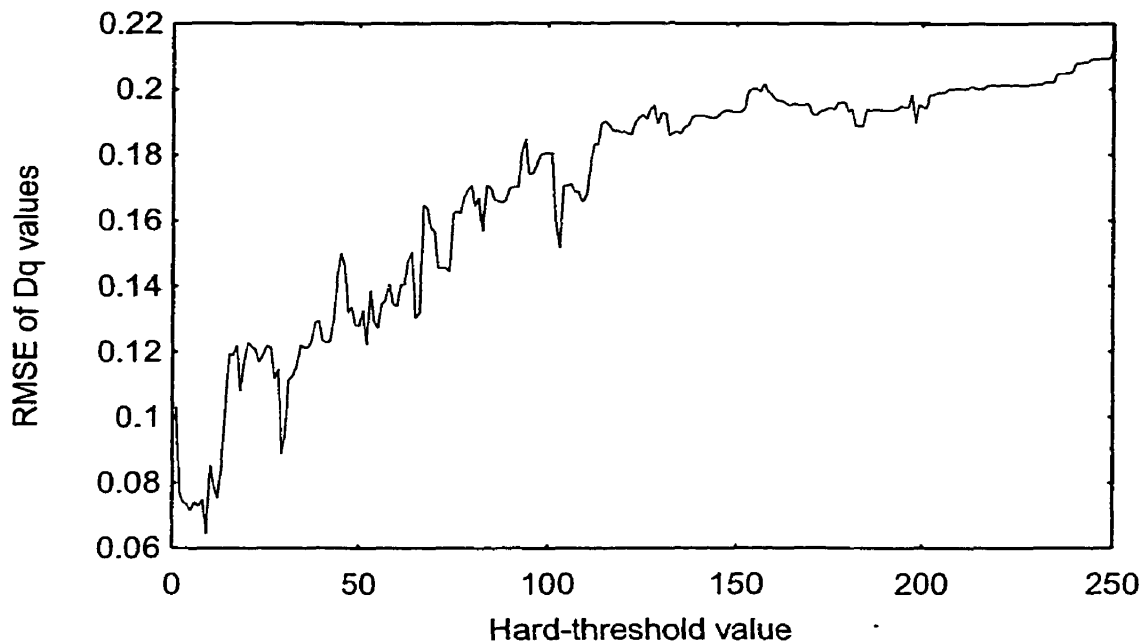
Fig. 6.15. RMSE tracking of Dq for urban (Daub20).

As noted previously, for three of the five test images the resulting RMSE plots tended to have decreasing $D_q(X)$ error values for decreasing hard-threshold while the other two images have results that are more difficult to characterize. One factor in forming the approximations is the selection of the mother wavelet. The Daubechies 4-tap mother wavelet [Daub92] was used, but this is likely not an optimal mother wavelet for all classes of images. Some RMSE experiments were performed using other mother wavelets in place of the Daubechies 4-tap mother wavelet. Figure 6.15 shows the RMSE results when using the Daubechies 20-tap mother wavelet [Daub92] to form the hard-threshold approx-

imations of the urban image. These results show a clearer tendency for the $D_q(X)$ error values to decrease along with the hard-threshold value compared with those results in Fig. 6.10. Therefore, choosing the right mother wavelet is also important in being able to model the multifractal complexity within the image properly.

6.5   Mandelbrot Spectrum as a Progressive Image Transmission Quality Measure

The next multifractal dimension measure to investigate as an objective measure for progressive image transmission is the Mandelbrot spectrum as described in Sec. 3.3.5. Some study of the Mandelbrot spectrum as an objective measure has been described by Dansereau and Kinsner [DaKi99b]. From Defn. 3.6, the information carried by the Mandelbrot spectrum is expected to be similar to that of the Rényi dimension spectrum since it is a transformation of $D_q(X)$ and q to form f_q and α_q. Since this is the case, only a few experiments are performed to show the differences between the Mandelbrot spectrum and the Rényi dimension spectrum discussed in Sec. 6.4.

The experiments done in this section are with the lena and urban images from Fig. 6.1 and Fig. 6.2, respectively. The Rényi dimension spectrum $D_q(X)$ results from Fig. 6.7 are taken and Eq. 3.54 and Eq. 3.55 applied for $f_q(\alpha)$ and α_q, respectively, to obtain the Mandelbrot spectra plotted in Fig. 6.16. A similar type of spread is noticed in the Mandelbrot spectra in Fig. 6.16a for the image of lena as compared to the Rényi dimension spectra $D_q(X)$ in Fig. 6.7a. Similarly, the clustering of the $D_q(X)$ curves in Fig. 6.7b for the image of urban is reflected in the $f_q(\alpha)$ curves in Fig. 6.16b. One

important point to notice is that the inflection points of the $D_q(X)$ curves from Fig. 6.7 at $q = 0$ have been transformed into the maxima points in the $f_q(\alpha)$ curves in Fig. 6.16. In addition, D_∞ and $D_{-\infty}$ are now contained in the plot of the Mandelbrot spectra, which is an analytical advantage in having the entire curve bounded in the plot.

In developing an objective measure based on the Mandelbrot spectrum, similar to those looked at for the Rényi dimension spectrum $D_q(X)$ in Sec. 6.4, the problem is turned into finding a convergence in the concave functions represented by the $f_q(\alpha)$ versus α curves. Even though the Mandelbrot spectrum has been shown to be useful in choosing parameters for other wavelet packet based image compression schemes [Jang97], [JaKi97], this thesis will not address the Mandelbrot spectrum fully. The reason for this brevity is that the information is available in the Rényi dimension spectrum $D_q(X)$ and extra computational errors are introduced when calculating the Mandelbrot spectrum using numerical differentiation for the $\dfrac{dD_q}{dq}$ term in Eq. 3.55. Therefore, the · focus in this thesis will be with the Rényi dimension spectrum $D_q(X)$ instead of the Mandelbrot spectrum. Also, as noticed in Fig. 6.16, there are artifacts in the $f_q(\alpha)$ versus α curves for larger values of α, which also correspond to larger magnitude negative values of q. These same artifacts were noticed by Jang [Jang97] and Chen [Chen97] and result from sensitivities in the calculation to the large values of q.
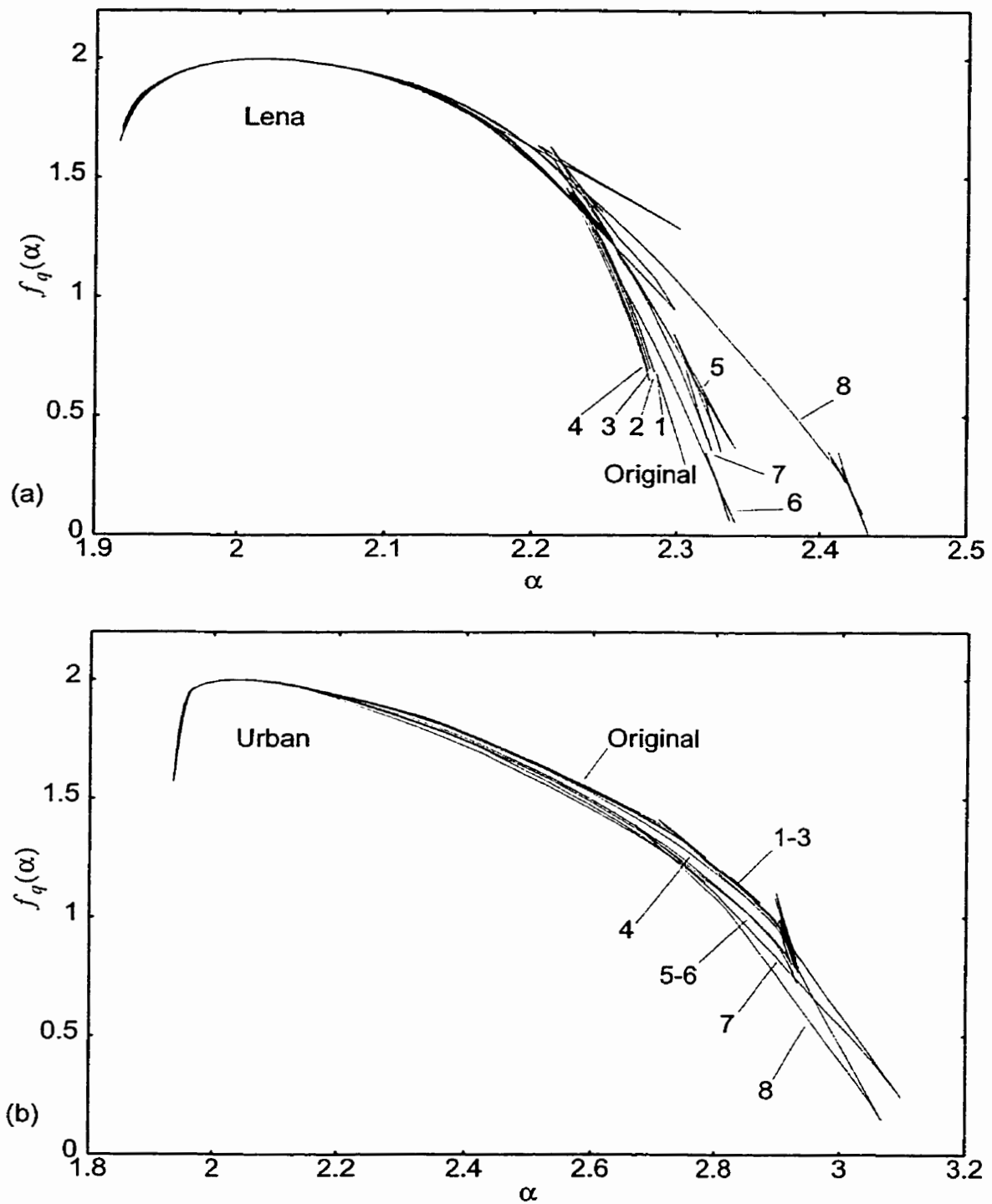
Fig. 6.16. Mandelbrot spectrum experimental results at progressive steps for the image of (a) lena and (b) urban.

6.6   Rényi Information and Relative Multifractal Measures as a
Progressive Image Transmission Quality Measure

From a theoretical standpoint, the idea of measuring the complexity within an image at multiple resolutions is a reasonable approach to forming an objective measure. This technique allows for the characterization of the complexity within an image from local to global image features. The approaches given in Sec. 6.4 and Sec. 6.5 using the Rényi dimension spectrum or Mandelbrot spectrum as the complexity measures fill the need for measuring the complexity within an image over these multiple resolutions. Unfortunately, when actually applied as described, the results obtained are difficult to interpret in an automated way as an objective measure and are sensitive to the whole measuring process, in particular the line fitting of the log-log plots. The problems in interpretation stem from the sharp changes in the RMSE calculations of Rényi dimension spectrum curves in Fig. 6.9 through Fig. 6.13. An improved approach is needed for performing these multifractal measures to help remove some of these sensitivities and help stabilize the measure into one that can more appropriately be used as an objective measure.

Stepping back and considering at the problem at hand, the goal behind lossy image compression and progressive image transmission is to construct a model that approximates the original image which is refined over time. In view of the multifractal measures of interest, instead of only constructing a model that approximates the original image, the key interest is to construct a model having a multifractal complexity that approximates the multifractal complexity in the original image. Therefore, the measure should consider the

multifractal complexity models of both the original and reconstructed images to form the objective measure, and preferably at the same time.

The models formed so far with the Rényi dimension spectrum and the Mandelbrot spectrum use the probability distribution function $p_s(x)$ at scale s as given in Eq. 6.1. Currently, $p_s(x)$ for the original image and the reconstructed image are characterized separately with the multifractal measures. A different approach combining the characterization into one multifractal measure for both probability distributions could be beneficial. Lanterman *et al.* [LaSM99] looked at using the Kullback-Leibler distance [KuLe51] measure as an approach to comparing two different models. The Kullback-Leibler distance is defined as follows.

Definition 6.1: Let $u(x)$ and $v(x)$ be complete probability distributions defined over the alphabet χ. The *relative entropy* or *Kullback-Leibler distance* $D(u \| v)$ between $u(x)$ and $v(x)$ is defined as

$$D(u \| v) = \sum_{x \in \chi} u(x) \log \frac{u(x)}{v(x)}. \tag{6.5}$$

❑

It should be noted that $D(u \| v) \neq D(v \| u)$ in general. Also, the Kullback-Leibler distance is not a true distance metric since it does not, in general, satisfy the triangle inequality rule of $D(u \| v) + D(v \| w) \geq D(u \| w)$, given the probability distributions $u(x)$, $v(x)$, and $w(x)$.

In the approach by Lanterman *et al.* [LaSM99], the probability distribution of the true system, $p^{\text{true}}(x)$, and the probability distribution from the model, $p^{\text{model}}(x)$, are defined for their system. Given these two probability distributions, the following expression is stated using the Kullback-Leibler distance

$$0 = D(p^{\text{true}} \| p^{\text{true}}) \le D(p^{\text{true}} \| p^{\text{model}}) \tag{6.6}$$

This indicates that the distance of the true system with itself is zero while the distance between the true system and any model approximating the system will be greater than zero. More importantly, if given two approximations of the system then $p^{\text{model}(1)}(x)$ can be considered better than $p^{\text{model}(2)}(x)$ if

$$D(p^{\text{true}} \| p^{\text{model}(1)}) \le D(p^{\text{true}} \| p^{\text{model}(2)}) \tag{6.7}$$

Equation 6.7 meets the goal of the objective measure desired since multiple approximation models can be generated for an image and then this form of comparison used to decide which model is of better quality.

Two main limitations and advantages exist between the Kullback-Leibler distance $D(u \| v)$ and the Rényi generalized entropy $H_q(X)$. The first limitation is that the Kullback-Leibler distance cannot inherently perform multifractal analysis on a model since there is no way of weighting the probability distributions to extract other inhomogeneous characteristics from the probability distribution set. This ability is an advantage of the Rényi generalized entropy since the use of the moment order q allows for biasing of different probabilities so that inhomogeneous characteristics can be extracted. The second limitation is with the Rényi generalized entropy since it only considers a single probability distribution in its calculation. This limitation does not exist with the Kullback-Leibler dis-

tance since it forms a relative measure between two different probability distributions and therefore makes inequalities such as Eq. 6.7 possible.

The goal then is to combine the advantages of the Rényi generalized entropy and the Kullback-Leibler distance into a single measure. Two additional criteria are placed in the design of a measure. The first criterion is that the measure must always be greater than or equal to zero with equality only when the two probability distributions are identical. The second criterion is that when $q = 1$ that the measure reduces to the Kullback-Leibler distance. The measure developed following these criteria is defined as follows.

Definition 6.2: Let $u(x)$ and $v(x)$ be probability distributions and let q be the moment order. The *generalized relative entropy* $RH_q(u \parallel v)$ or *generalized Kullback-Leibler distance* is then defined as

$$RH_q(u \parallel v) = \frac{1}{q-1} \log \frac{\sum_{x \in \chi} v(x) \left(\frac{u(x)}{v(x)} \right)^q}{\sum_{x \in \chi} u(x)} . \tag{6.8}$$

❑

It turns out that Eq. 6.8 is the Rényi information $I_q(u \parallel v)$ of the two probability distributions as developed by Rényi in [Rény60] which means that

$$RH_q(u \parallel v) = I_q(u \parallel v) \tag{6.9}$$

As can be seen, Eq. 6.8 incorporates the different characteristics outlined for this measure. The first is that it incorporates the moment order power q in a manner similar to the Rényi generalized entropy in Eq. 3.23. Another important characteristic is that Eq. 6.8 allows for two probability distributions to be included in the measure. One difference that

is obvious with Eq. 6.8 is that negative values are possible. This is will not be a problem in our application since the absolute value of the measure could be taken to give a similar result as the Kullback-Leibler distance.

The following theorem shows that the measure is zero when the two probability distributions are identical.

Theorem 6.1: Let $u(x)$ and $v(x)$ be probability distributions where $u(x) = v(x)$. The generalized relative entropy $RH_q(u \| v)$ gives

$$RH_q(u \| v) = 0 \qquad \text{for } u = v \tag{6.10}$$

Proof:

$$RH_q(u \| v) = RH_q(u \| u) = I_q(u \| u) \tag{6.11}$$

$$= \frac{1}{q-1} \log \frac{\sum_{x \in \chi} u(x) \left(\frac{u(x)}{u(x)} \right)^q}{\sum_{x \in \chi} u(x)} \tag{6.12}$$

$$= \frac{1}{q-1} \log \frac{\sum_{x \in \chi} u(x) \left(\frac{1}{1} \right)^q}{\sum_{x \in \chi} u(x)} \tag{6.13}$$

$$= \frac{1}{q-1} \log \frac{\sum_{x \in \chi} u(x)}{\sum_{x \in \chi} u(x)} \tag{6.14}$$

$$= \frac{1}{q-1} \log 1 \tag{6.15}$$

$$= 0. \tag{6.16}$$

$\square$

The second criterion in the design of the measure in Eq. 6.8 is that when $q = 1$ that the measure reduces to the Kullback-Leibler distance. This property is shown with the following theorem.

Theorem 6.2: Let $u(x)$ and $v(x)$ be complete probability distributions. The generalized relative entropy $RH_q(u \| v)$ is precisely the Kullback-Leibler distance $D(u \| v)$ for $q = 1$ as follows

$$RH_{q = 1}(u \| v) = D(u \| v).$$
(6.17)

Proof: Finding the limit as $q \to 1$, we have

$$RH_{q = 1}(u \| v) = \lim_{q \to 1} \frac{1}{q - 1} \log \frac{\sum_{x \in \chi} v(x) \left(\frac{u(x)}{v(x)} \right)^q}{\sum_{x \in \chi} u(x)}$$
(6.18)

Since

$$\lim_{q \to 1} \log \frac{\sum_{x \in \chi} v(x) \left(\frac{u(x)}{v(x)} \right)^q}{\sum_{x \in \chi} u(x)} = \log \frac{\sum_{x \in \chi} v(x) \frac{u(x)}{v(x)}}{\sum_{x \in \chi} u(x)}$$
(6.19)

$$= \log \frac{\sum_{x \in \chi} u(x)}{\sum_{x \in \chi} u(x)}$$
(6.20)

$$= \log 1$$
(6.21)

$$= 0$$
(6.22)

and

$$\lim_{q \to 1} q - 1 = 0$$
(6.23)

then l'Hôpital's rule can be used on Eq. 6.18 to find the limit when $q \to 1$ as follows

$$RH_{q=1}(u \| v) = \lim_{q \to 1} \frac{1}{q-1} \log \frac{\sum_{x \in \chi} v(x)\left(\frac{u(x)}{v(x)}\right)^q}{\sum_{x \in \chi} u(x)} \tag{6.24}$$

$$= \lim_{q \to 1} \frac{\frac{d}{dq} \log \frac{\sum_{x \in \chi} v(x)\left(\frac{u(x)}{v(x)}\right)^q}{\sum_{x \in \chi} u(x)}}{\frac{d}{dq}(q-1)} \tag{6.25}$$

$$= \lim_{q \to 1} \frac{\frac{d}{dq} \log \sum_{x \in \chi} v(x)\left(\frac{u(x)}{v(x)}\right)^q - \frac{d}{dq} \log \sum_{x \in \chi} u(x)}{\frac{d}{dq}(q-1)} \tag{6.26}$$

$$= \lim_{q \to 1} \frac{d}{dq} \log \sum_{x \in \chi} v(x)\left(\frac{u(x)}{v(x)}\right)^q \tag{6.27}$$

$$= \lim_{q \to 1} \frac{\frac{d}{dq} \sum_{x \in \chi} v(x)\left(\frac{u(x)}{v(x)}\right)^q}{\sum_{x \in \chi} v(x)\left(\frac{u(x)}{v(x)}\right)^q} \tag{6.28}$$

$$= \lim_{q \to 1} \frac{\sum_{x \in \chi} v(x)\frac{d}{dq}\left(\frac{u(x)}{v(x)}\right)^q}{\sum_{x \in \chi} v(x)\left(\frac{u(x)}{v(x)}\right)^q} \tag{6.29}$$

$$= \lim_{q \to 1} \frac{\sum_{x \in \chi} v(x)\left(\frac{u(x)}{v(x)}\right)^q \log \frac{u(x)}{v(x)}}{\sum_{x \in \chi} v(x)\left(\frac{u(x)}{v(x)}\right)^q} \tag{6.30}$$

$$= \frac{\displaystyle\sum_{x \in \chi} v(x)\frac{u(x)}{v(x)}\log\frac{u(x)}{v(x)}}{\displaystyle\sum_{x \in \chi} v(x)\frac{u(x)}{v(x)}} \tag{6.31}$$

$$= \frac{\displaystyle\sum_{x \in \chi} u(x)\log\frac{u(x)}{v(x)}}{\displaystyle\sum_{x \in \chi} u(x)} \tag{6.32}$$

Since the probability distributions for the Kullback-Leibler distance $D(u \| v)$ are defined

as complete, then $\displaystyle\sum_{x \in \chi} u(x) = 1$. Therefore

$$RH_{q=1}(u \| v) = \sum_{x \in \chi} u(x)\log\frac{u(x)}{v(x)} = D(u \| v) \tag{6.33}$$

which is the Kullback-Leibler distance. $\square$

These theorems show that the measure given in Defn. 6.2 follow the criteria out-

lined for the desired measure. The measure presented is not yet a multifractal measure

since it does not consider the self-similarity of the object over multiple scales to build up

the appropriate power law relationship similar to Eq. 3.5. This power law relationship is

necessary for finding the critical exponent that balances out the measured value over the

multiple scales. The following newly introduced definition transforms the relative Rényi

generalized entropy $RH_q(u \| v)$ from Eq. 6.8 into the form of a multifractal measure sim-

ilar to that of Eq. 3.45 for the definition of the Rényi dimension spectrum $D_q(X)$.

Definition 6.3: Let s be the scale of measurement, $u_s(x)$ and $v_s(x)$ be probability distri-

bution resolved at scale s, and q be the moment order. The *relative Rényi dimension*

spectrum $RD_q(u \| v)$ or *relative multifractal dimension spectrum* with order q between

$u_s(x)$ and $v_s(x)$ is defined as

$$RD_q(u \| v) = \lim_{s \to \infty} \frac{RH_q(u \| v)}{\log \frac{1}{s}} = \lim_{s \to \infty} \frac{\frac{1}{q-1} \log \frac{\sum_{x \in \chi_s} v_s(x) \left(\frac{u_s(x)}{v_s(x)} \right)^q}{\sum_{x \in \chi_s} u_s(x)}}{\log \frac{1}{s}} . \tag{6.34}$$

$\square$

Definition 6.3 presents a new class of multifractal measures that, to the best of our

knowledge, has not been seen in literature. This definition opens up a new approach to

performing multifractal measures where two probability distributions are compared in a

relative manner, similar to that of relative entropy in information theory. One interesting

property of the relative Rényi dimension spectrum is given in the following theorem.

Theorem 6.3: Let $u_s(x)$ and $v_s(x)$ be complete probability distributions, s be the scale

of measurement, and let the moment order $q = 0$. The relative Rényi dimension spec-

trum is then

$$RD_{q=0}(u \| v) = 0. \tag{6.35}$$

Proof:

$$RD_q(u \| v) = \lim_{s \to \infty} \frac{\frac{1}{(0)-1} \log \frac{\sum\limits_{x \in \chi_s} v_s(x) \left(\frac{u_s(x)}{v_s(x)}\right)^{(0)}}{\sum\limits_{x \in \chi_s} u_s(x)}}{\log \frac{1}{s}} \tag{6.36}$$

$$= \lim_{s \to \infty} \frac{-\log \frac{\sum\limits_{x \in \chi_s} v_s(x)}{\sum\limits_{x \in \chi_s} u_s(x)}}{\log \frac{1}{s}} \tag{6.37}$$

$$= \lim_{s \to \infty} \frac{\log \sum\limits_{x \in \chi_s} u_s(x) - \log \sum\limits_{x \in \chi_s} v_s(x)}{\log \frac{1}{s}} \tag{6.38}$$

With $u(x)$ and $v(x)$ as complete probability distributions, then Eq. 6.38 reduces to

$$RD_q(u \| v) = \lim_{s \to \infty} \frac{\log 1 - \log 1}{\log \frac{1}{s}} = \lim_{s \to \infty} \frac{0}{\log \frac{1}{s}} = 0 \tag{6.39}$$

$\square$

Theorem 6.3 states that regardless of the probability distributions $u_s(x)$ and $v_s(x)$, when $q = 0$ the relative Rényi dimension spectrum between $u_s(x)$ and $v_s(x)$ is zero. This property is interesting since it shows that when $q = 0$, comparing two proba-

bility distributions over different scales has constant growth. This can be shown by the following power law relationship expressing Eq. 6.34 when $q = 0$

$$\left(\frac{\sum\limits_{x \in \chi_r} v_s(x)\left(\frac{u_s(x)}{v_s(x)}\right)^q}{\sum\limits_{x \in \chi_r} u_s(x)}\right)^{\frac{1}{q-1}}\Bigg|_{q=0} \propto \left(\frac{1}{s}\right)^{RD_{q-0}(u \parallel v)} \tag{6.40}$$

This proportionality can be reduced as follows since $RD_{q=0}(u \parallel v) = 0$

$$\left(\frac{\sum\limits_{x \in \chi_r} u_s(x)}{\sum\limits_{x \in \chi_r} v_s(x)}\right) \propto 1 \tag{6.41}$$

which shows that the growth is constant.

Using the relative Rényi dimension spectrum, a number of experiments were performed to see the relationship between $RD_q(u \parallel v)$ and q. For these experiments, the hard-threshold approximations of lena as described in Sec. 6.4 are taken and the relative Rényi dimension spectrum applied over a range of values for q. For reasons of space and to avoid cluttering the plots, only the approximations of lena with a hard-threshold of 2^n for $0 \leq n \leq 10 \in \mathbf{N}$ are chosen for the experiments. The values of $RD_q(u \parallel v)$ are calculated for $q \in [-40.0, 40.0]$ in intervals of 0.25. Using these parameters, Fig. 6.17 was produced.

Fig. 6.17. Plot of relative Rényi dimension spectrum versus q for lena approximated from a Daub4 DWT using a hard-threshold of 2^n for $n = 0$ to $n = 10$.

The first thing to note about Fig. 6.17 is that all of the curves pass through the origin when $q = 0$ as Theorem 6.3 requires. Actually, this is not entirely accurate since the curves actually pass through the point 6.20542214e-09 in the calculations. This result is likely due to some numerical rounding and imprecision issues with the actual calculations. It should also be noticed from Fig. 6.17 that as the quality of the image improves with decreasing hard-threshold, that the curves converge to the abscissa. This observation is promising for using the relative Rényi dimension spectrum as an objective measure, particularly when looking at how the hard-thresholds from $2^{10} = 1024$ down to $2^6 = 64$ converge.

Unfortunately, some of the greater complexity within the curves of Fig. 6.17 cannot be clearly seen due to the scale. The curves for $2^0 = 1$ through $2^7 = 128$ are replot-

ted in Fig. 6.18 through Fig. 6.25, respectively, so that closer analysis can be done. The plots for a hard-threshold of $2^0 = 1$ and $2^1 = 2$ in Fig. 6.18 and Fig. 6.19, respectively, show a near straight line with a constant negative slope. The plots become more interesting in Fig. 6.20 for a hard-threshold of $2^2 = 4$ where some of the points begin to form separate line segments in the plot for $q < -15$ and $q > 30$. The segment for $-15 < q < 30$ is still roughly linear with a constant slope, though, a larger slope than in the previous two figures.

Starting with Fig. 6.21, the additional line segments on either side of $q = 0$ begin to take on a different shape. The further away from $q = 0$, the more each line segment approaches a larger positive slope. This trend continues with Fig. 6.22 and Fig. 6.23. With Fig. 6.24 and Fig. 6.25 the line segments continue approaching larger positive slopes the further from $q = 0$, but, it is clear that the line segments continue from there and change from a large positive slope to flip to a large negative slope that then begins to slope back down. In essence, there is a slight rotation in each of the line segments formed as the moment orders of q move away from zero.

At this point it is still unclear why these line segments are appearing, but, it may be due to the double floating point precision used and the large powers of q. It should also be noted that the span of the line segments are contained in a growing envelope for this range of q.
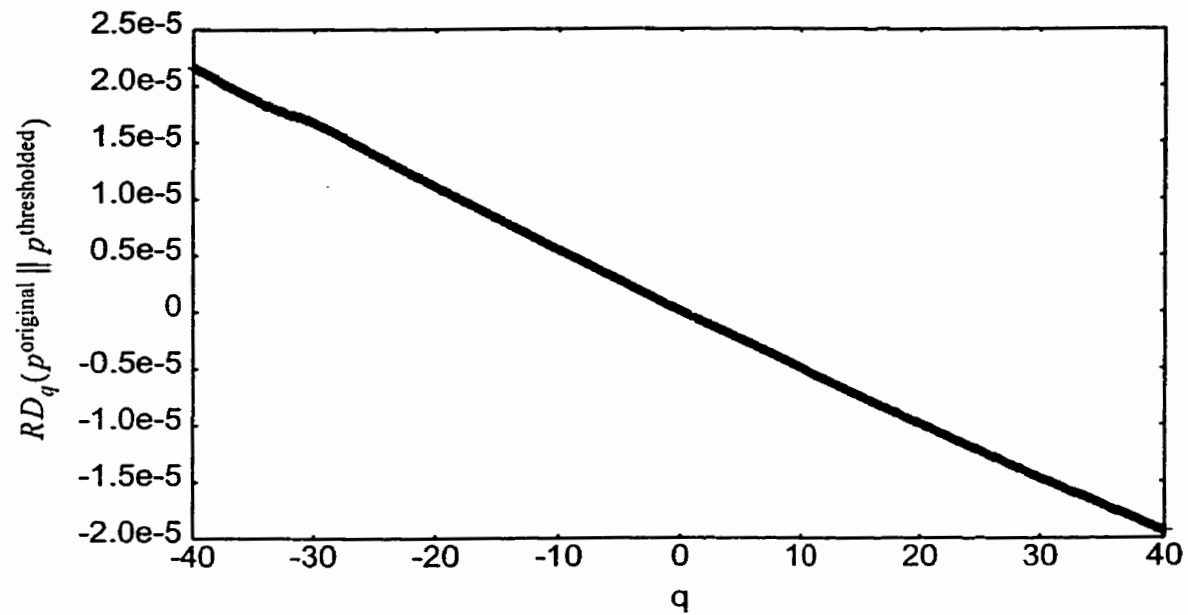
Fig. 6.18. Plot of relative Rényi dimension spectrum versus q for lena approximated from a Daub4 DWT using a hard-threshold of 1.0.
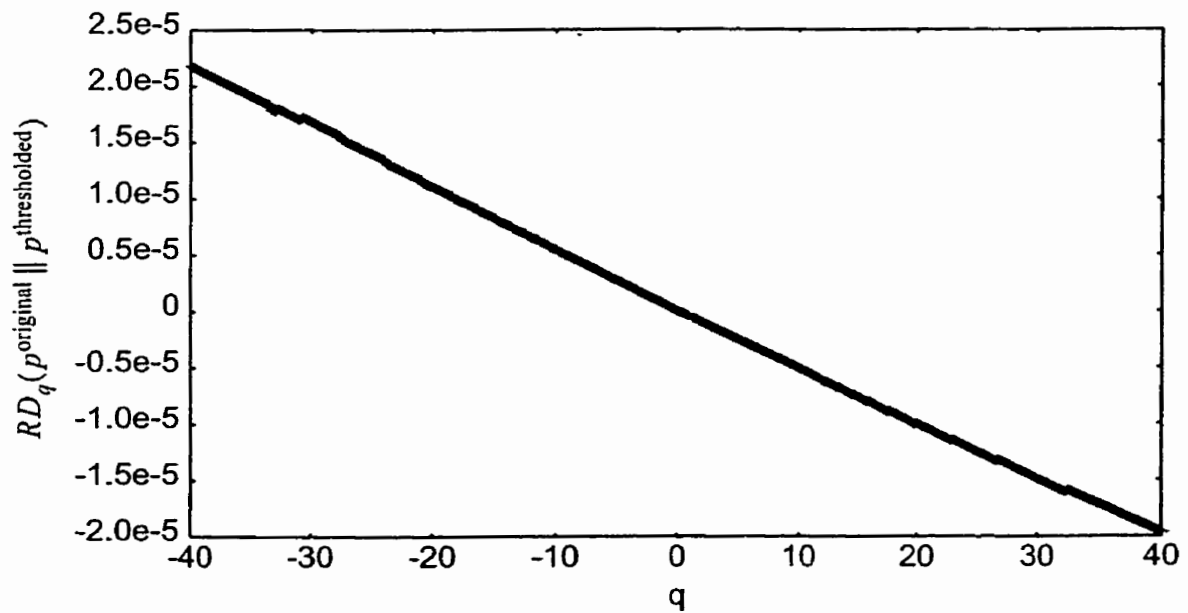
Fig. 6.19. Plot of relative Rényi dimension spectrum versus q for lena approximated from a Daub4 DWT using a hard-threshold of 2.0.
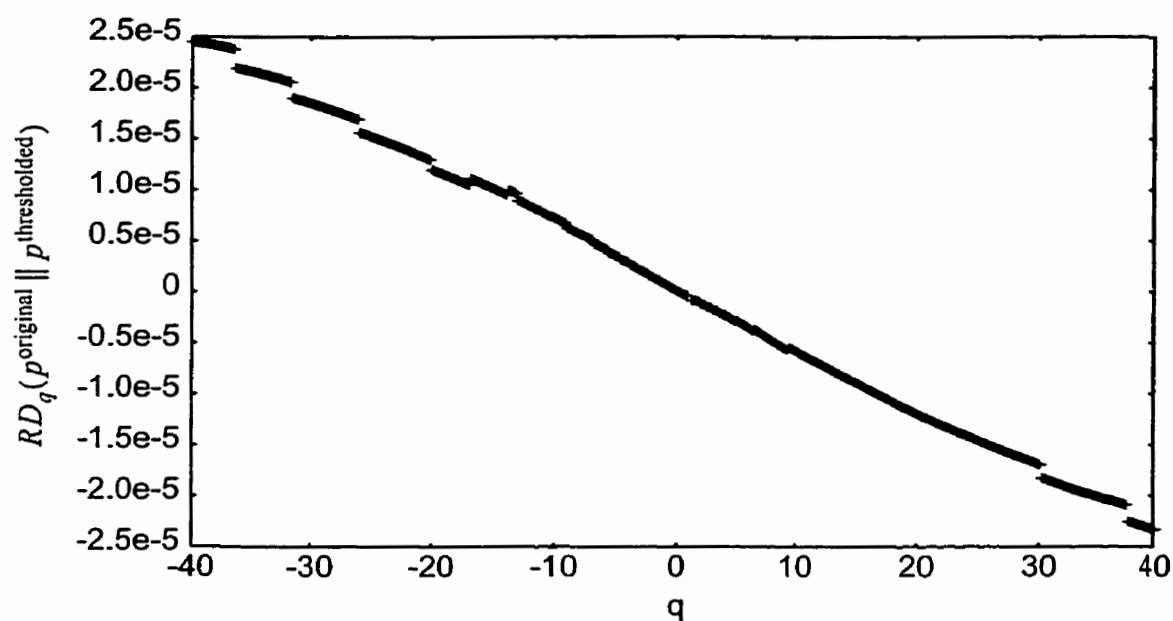
Fig. 6.20. Plot of relative Rényi dimension spectrum versus q for lena approximated from a Daub4 DWT using a hard-threshold of 4.0.
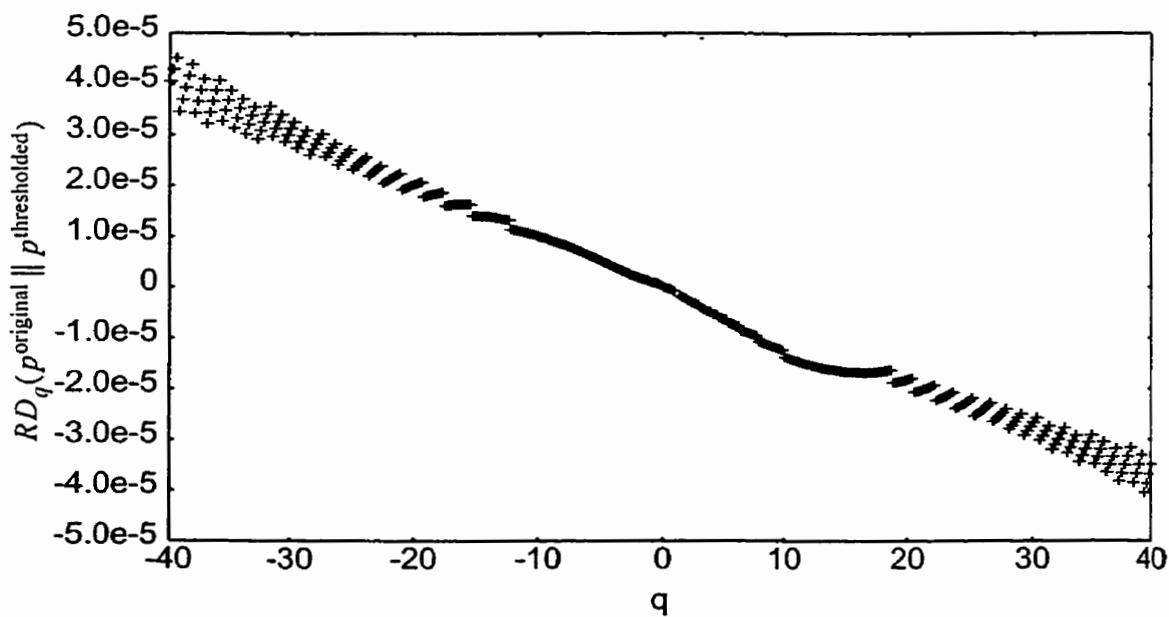
Fig. 6.21. Plot of relative Rényi dimension spectrum versus q for lena approximated from a Daub4 DWT using a hard-threshold of 8.0.
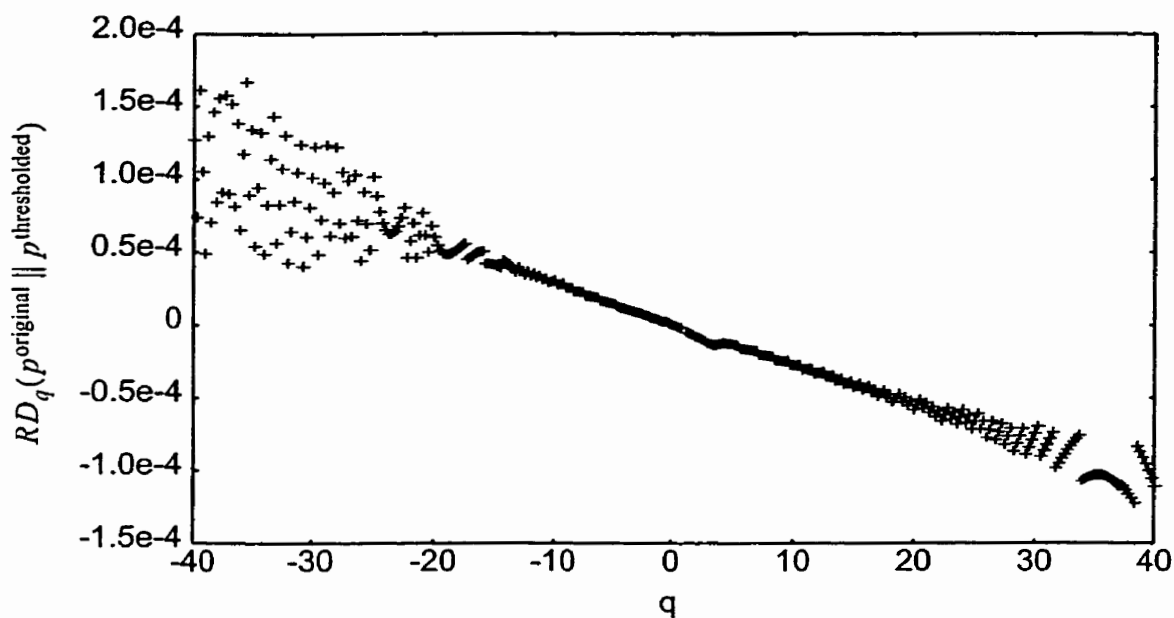
Fig. 6.22. Plot of relative Rényi dimension spectrum versus q for lena approximated from a Daub4 DWT using a hard-threshold of 16.0.
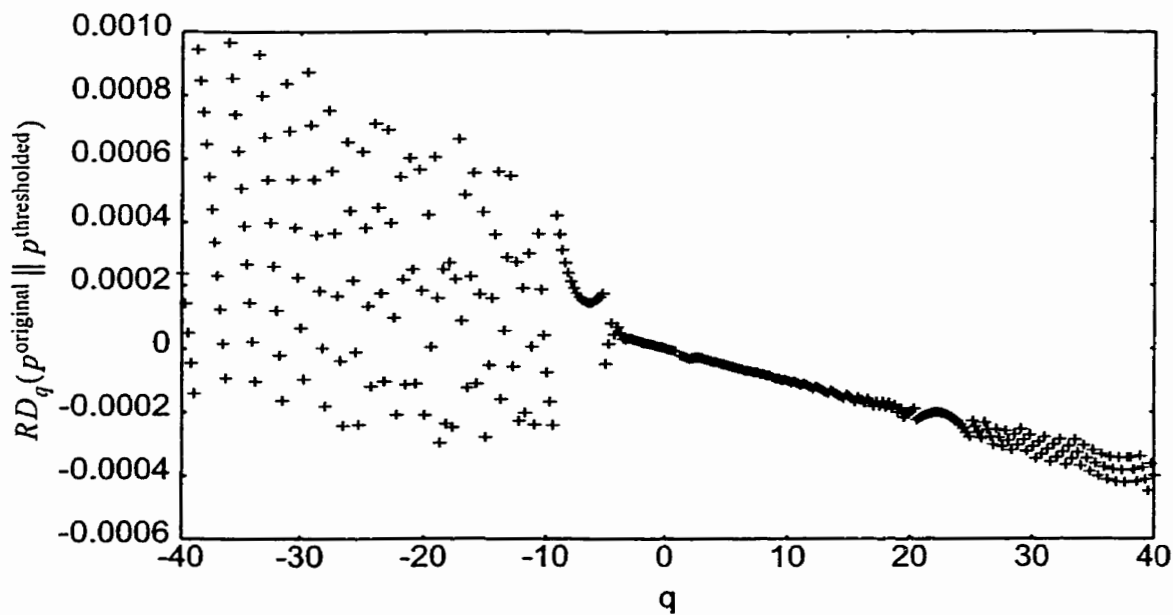
Fig. 6.23. Plot of relative Rényi dimension spectrum versus q for lena approximated from a Daub4 DWT using a hard-threshold of 32.0.
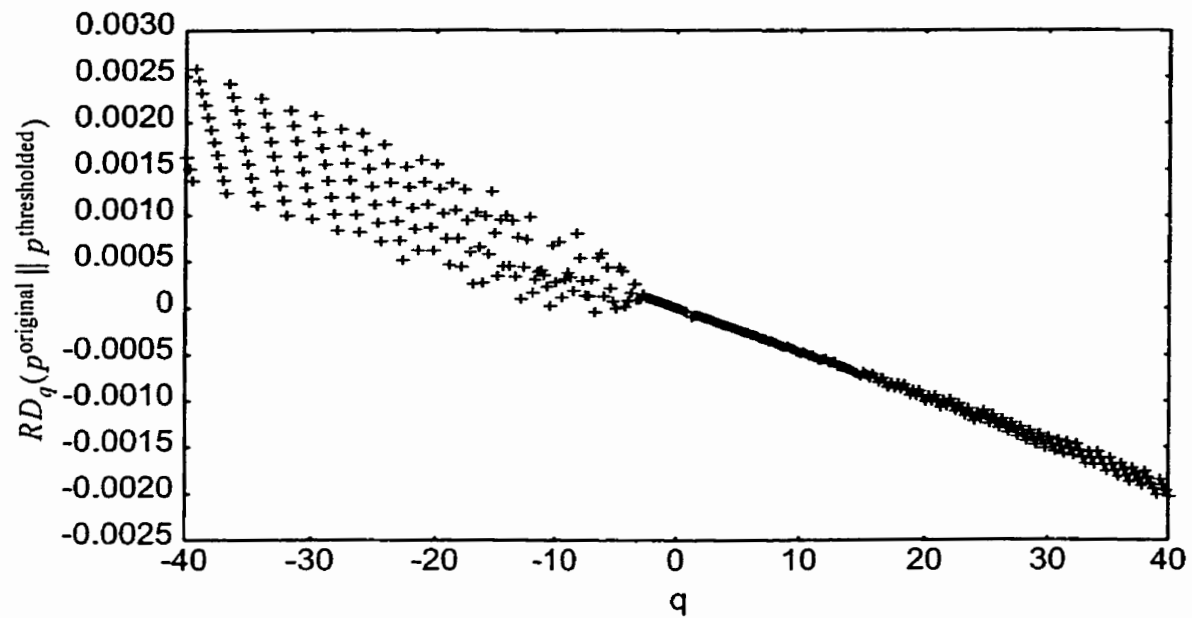
Fig. 6.24. Plot of relative Rényi dimension spectrum versus q for lena approximated from a Daub4 DWT using a hard-threshold of 64.0.
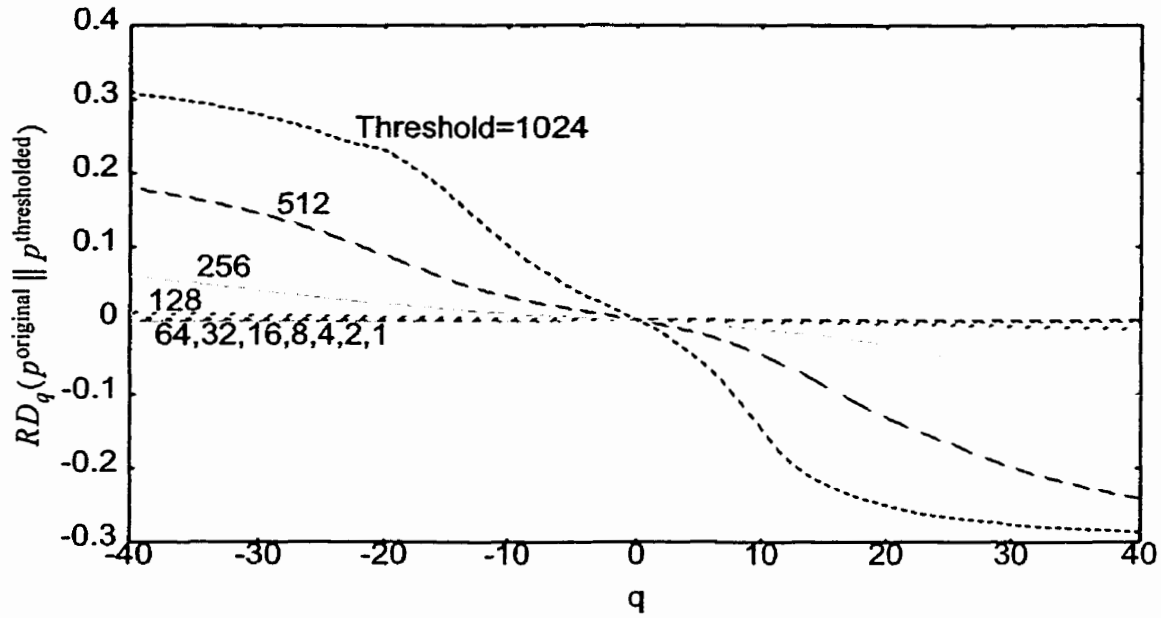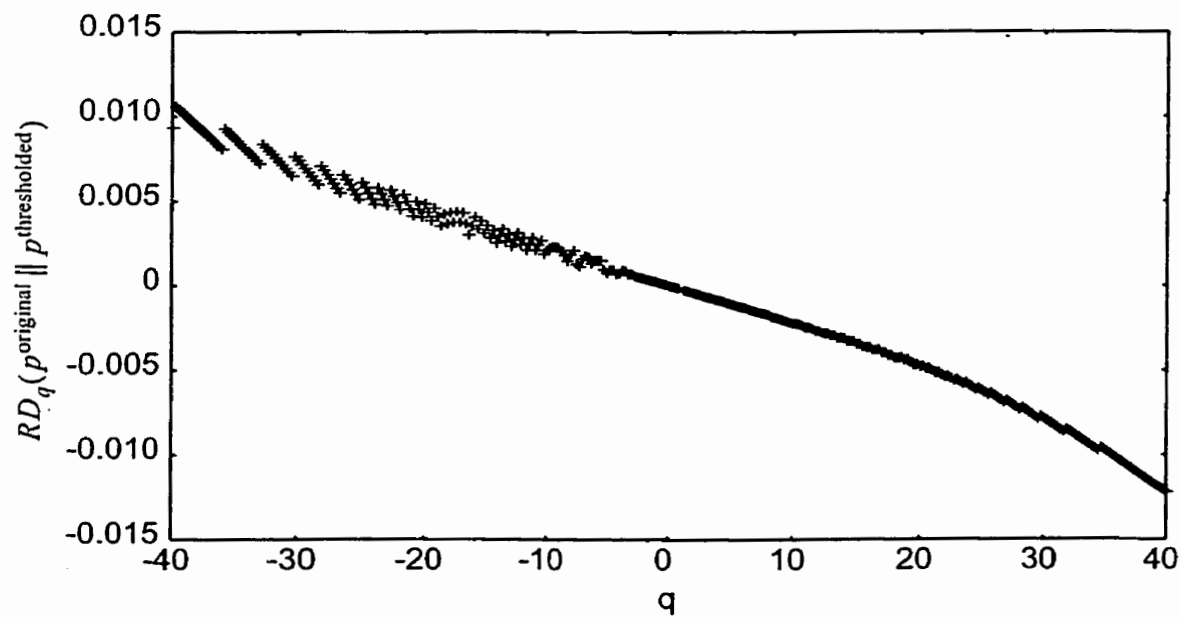

Fig. 6.25. Plot of relative Rényi dimension spectrum versus q for lena approximated from a Daub4 DWT using a hard-threshold of 128.0.

The results given in Fig. 6.18 through Fig. 6.25 do not make the relative Rényi dimension spectrum unusable as an objective measure, but, care will be needed in interpreting the results if larger absolute values of q are used. Another observation that can be made about the plots, where extra line segments appear for larger absolute values of q, is that the trend tends to follow the path made by the line segment in and around the $q = 0$ mark. Therefore, using the values further out from $q = 0$ as an average will approximate an extended version of the inner line segment.

A simple approach to using the relative Rényi dimension spectrum as an objective measure is to calculated the spectrum over a number of moments of order q and then sum these values. This objective measure will be referred to as IQM and expressed as follows

$$\text{IQM}(p^{\text{original}} \parallel p^{\text{thresholded}}) = \sum_{q \in I_q} RD_q(p^{\text{original}} \parallel p^{\text{thresholded}}) \tag{6.42}$$

where I_q is a set of values for the moment order q.

Using this objective measure, a number of experiments are performed on the sets of 250 hard-threshold images described in Sec. 6.4. The purpose of the experiments is to observe how the IQM changes for different hard-threshold values. The set of values for $q \in I_q$ of Eq. 6.42 are chosen in the range of $[-20.0, 20.0]$ in increments of 0.5. These experiments are performed for lena, urban, baboon, peppers, and farm with the results plotted in Fig. 6.26, Fig. 6.27, Fig. 6.28, Fig. 6.29, and Fig. 6.30, respectively. Note that for Fig. 6.27 and Fig. 6.29 that the RMSE curves from Fig. 6.10 and Fig. 6.12, respectively, have been overlaid in the plot for the purpose of analysis and comparison.
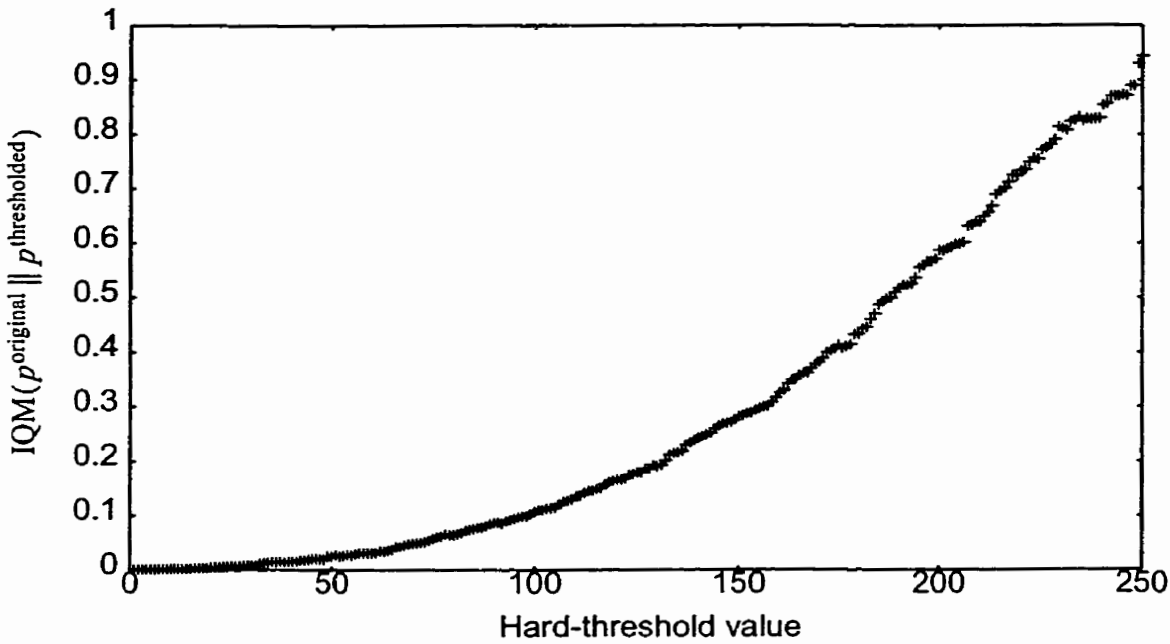
Fig. 6.26. IQM measure versus hard-threshold value for image of lena (Daub4).
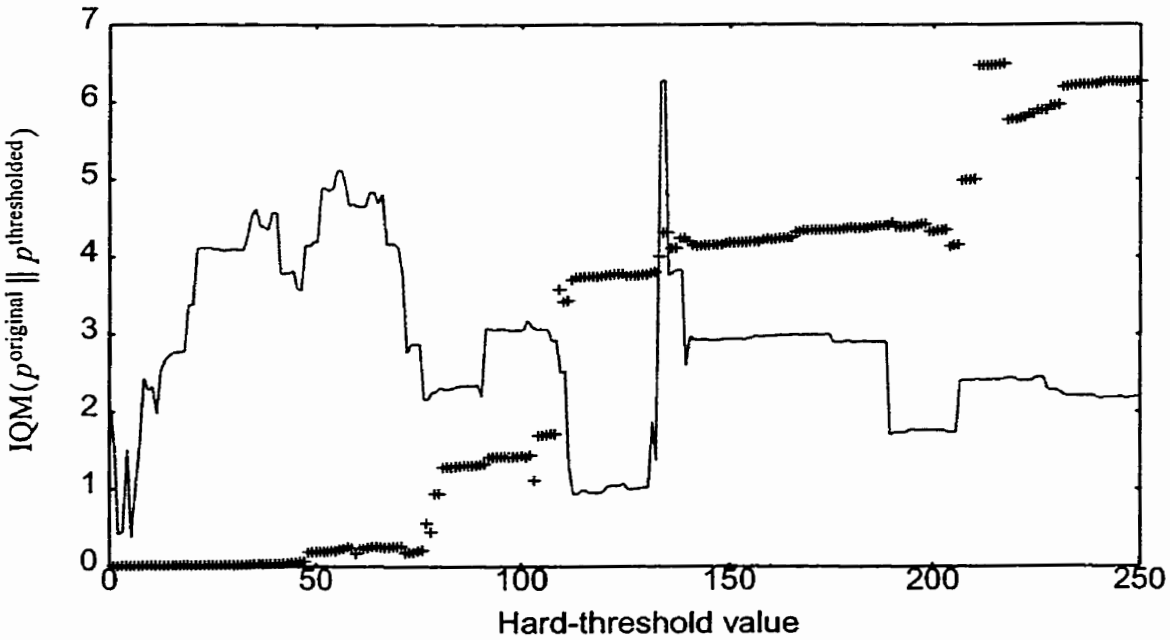
Fig. 6.27. IQM measure versus hard-threshold value for image of urban (Daub4).
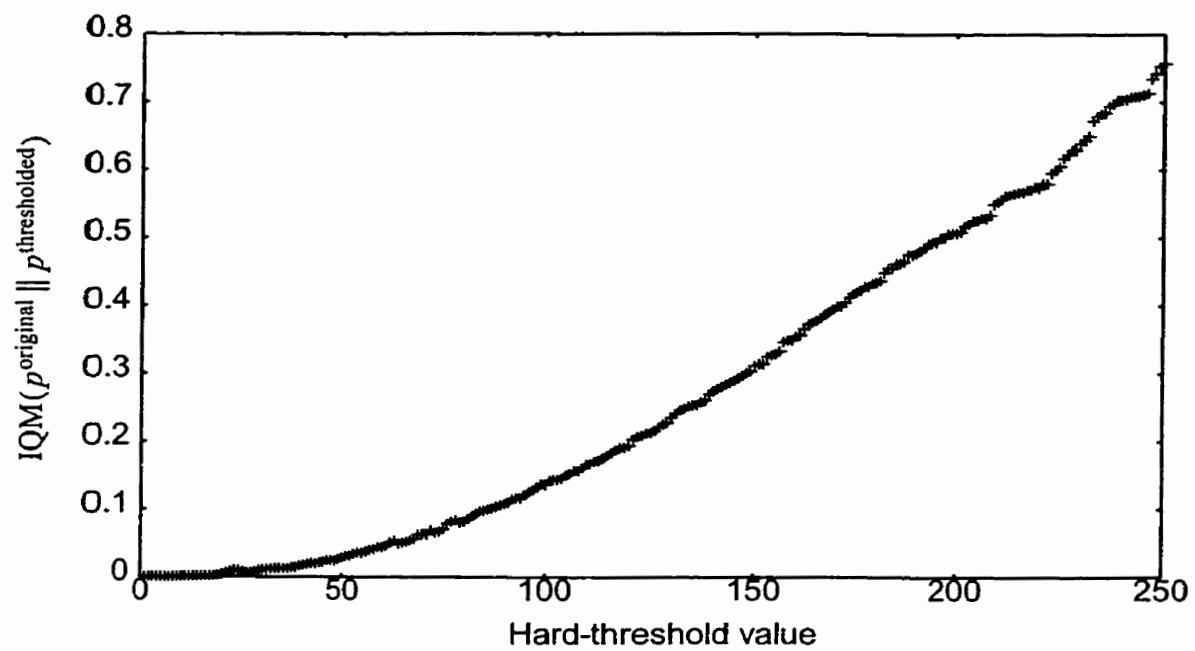
Fig. 6.28. IQM measure versus hard-threshold value for image of baboon (Daub4).
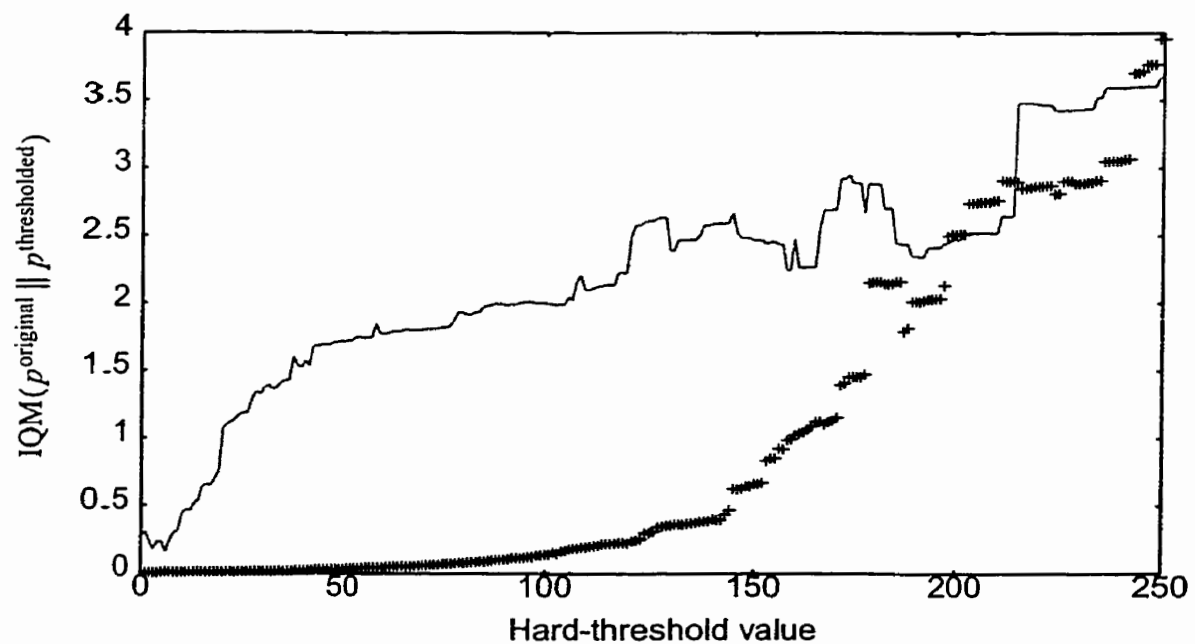
Fig. 6.29. IQM measure versus hard-threshold value for image of peppers (Daub4).
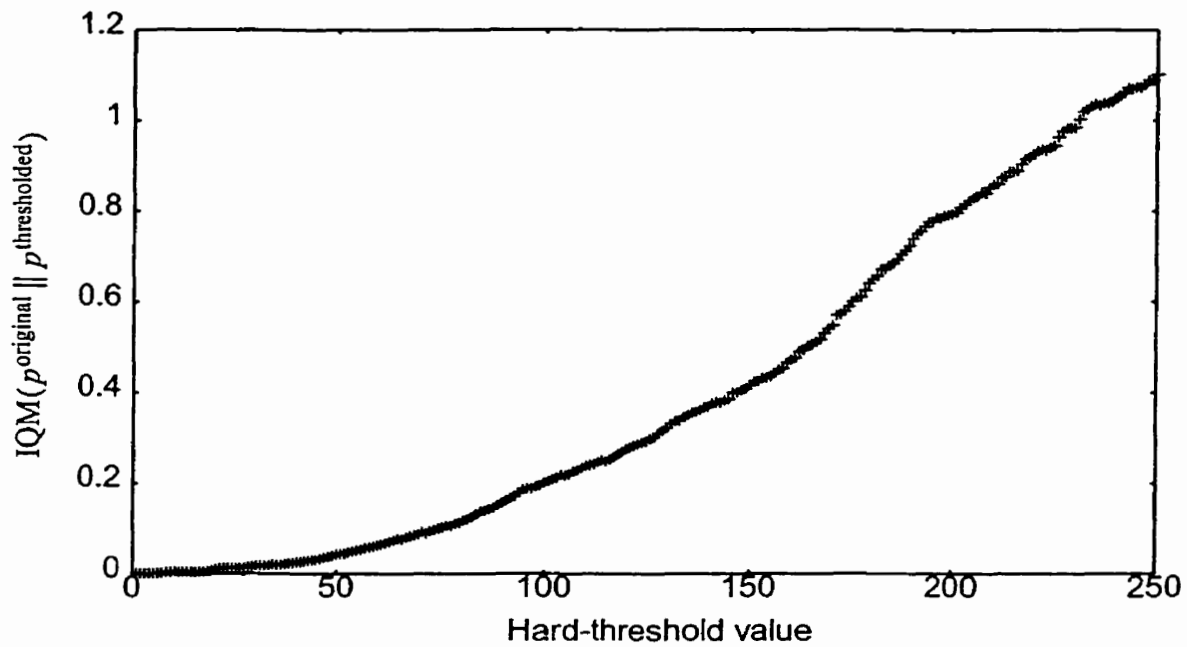
Fig. 6.30. IQM measure versus hard-threshold value for image of farm (Daub4).

The results of the IQM experiments are very encouraging, particularly for lena, baboon, and farm in Fig. 6.26, Fig. 6.28, and Fig. 6.30, respectively, which have a near exponential drop in the IQM value as the hard-threshold value decreases. This shape follows what would intuitively be expected for an objective measure of image quality, since the reconstructions have this sort of improvement in image quality as the hard-threshold level decreases and more image information is available in the reconstruction.

The IQM results for the image of peppers in Fig. 6.29 also predominantly has an exponential change in IQM versus hard-threshold level. In some ranges of the hard-threshold values, the IQM results for peppers deviate from the exponential curve such as the range from 178 to 186 where there is a step and the range from 211 to 242 where the curve plateaus. An interesting observation is that these two regions correspond to the step

and plateau in the overlaid curve from Fig. 6.12. Therefore, some correlations exist between the IQM and the Rényi dimension spectrum measures presented in Sec. 6.4. In addition, a small rise in the points around the hard-threshold value of 125 corresponds with a smaller step in the overlaid curve, and the small jump at 145 corresponds with a small peak in the overlaid curve. Not all sharp changes in the overlaid curve are readily noticed in the IQM results, but many correlate well.

The IQM results for the urban image given in Fig. 6.27 show far less of an exponential change in IQM versus hard-threshold compared with the other four sets of results. While the points, in general, show that the IQM value decreases in correspondence with a decrease in hard-threshold, there are more plateaus visible in this plot with sharper drops as compared with the other IQM results. Again, a very interesting observation is that the IQM results have many correlations with the overlaid results for the Rényi dimension spectrum from Fig. 6.10. For instance, the hard-threshold values of 48, 79, 109, 133, 207 in the IQM results have noticeable transitions that correlate well with sharp changes in the overlaid results for the Rényi dimension spectrum from Fig. 6.10.

With the promising results for IQM observed in Fig. 6.26 to Fig. 6.30, the next step is to see how the IQM measure correlates with MOS experimental results. Using the previously described MOS experiments, the correlation between the IQM measure and the MOS results are plotted in Fig. 6.31. As would be expected from a good objective measure, there is a high correlation between the plotted 1/log(IQM) and the results obtained from the MOS experiments as the fitted line indicates in Fig. 6.31. A more indepth set

MOS experiments need to be run to support these initial experiments, but correlation does

exist between the developed measure and the MOS results.
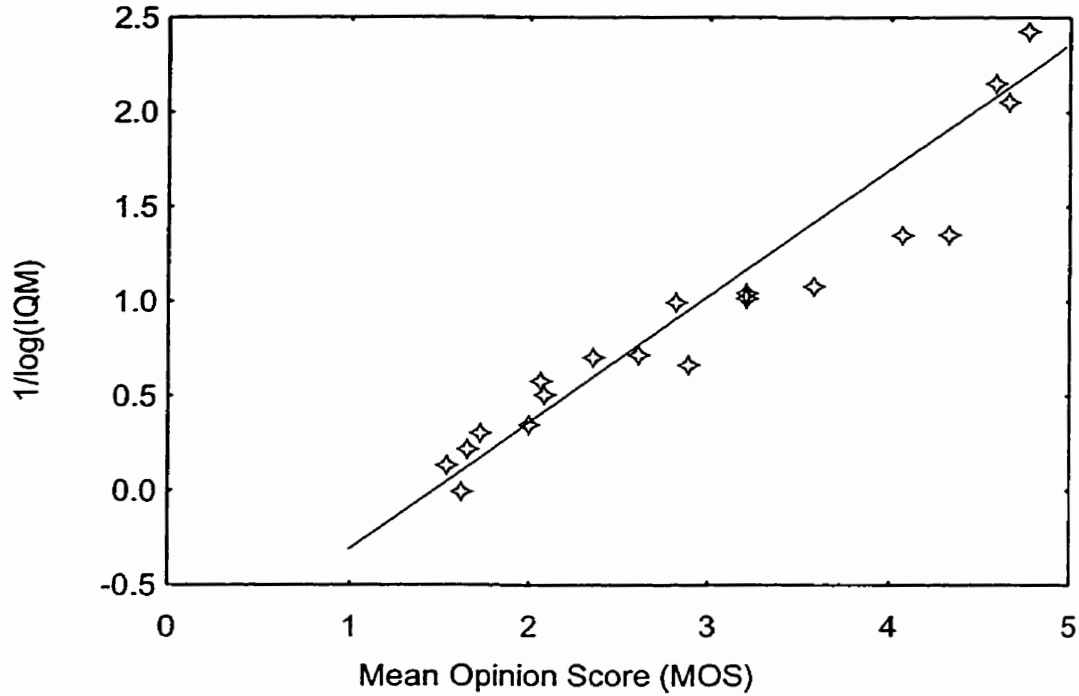
Fig. 6.31. The developed IQM objective measure versus
mean opinion score experiment correlation results.

In conclusion, when considering the relative Rényi dimension spectrum and the

IQM measure, a few important characteristics must be noted. The first is that these mea-

sures allow for two probability distributions to be compared and hence the original image

and the reconstructed image to be compared. More importantly, the combination of the

probabilities within the summation and logarithm of the relative Rényi dimension spec-

trum $RD_q(u \| v)$ in Eq. 6.34 versus looking at each probability separately as with the

$D_q(X)$ in Eq. 3.45 removes many of the numeric sensitivity in the calculation, particu-

larly the errors from the line fitting required for finding the critical exponents. Therefore,

the relative Rényi dimension spectrum $RD_q(u \parallel v)$ is better suited as an objective measure than the other multifractal measures presented.

6.7   Summary

This chapter has investigated and outlined the use of multifractal dimension measures such as Rényi generalized entropy, Rényi dimension spectrum, Mandelbrot spectrum, and the newly proposed relative Rényi dimension spectrum for analyzing the multifractal complexity of images. This analysis, in turn, makes way for a new method of objectively analyzing the quality of signals in terms of signal complexity and in particular the quality of lossy image reconstructions through differences in complexity. This approach is especially important since the multifractal analysis is done over all resolutions of the image, while many other objective measures only consider localized changes in the image, such as with the PSNR which only looks at squared pixel energy differences on a pixel by pixel basis. The proposed approach, therefore, for an objective measure is to use the IQM measure developed in Sec. 6.6 since it has some useful properties, has an encouraging exponential drop as the image quality improves, and removes some of the sensitivity to the calculation from the result.

The following chapter considers the development of wavelet and wavelet packet based progressive image transmission techniques. Of particular interest from the standpoint of this chapter, the following chapter considers the use of the developed multifractal dimension measures to assist in the selection of parameters, wavelet coefficients, and wavelet packet bases, so that the overall progressive image transmission can be improved.

Chapter VII

Progressive Wavelet Image Compression with Multifractals

This chapter explores using the measures developed in the previous chapters in the context of progressive image transmission techniques, with focus on wavelet and wavelet packet image compression techniques. One component looked at in this chapter is progressive wavelet image transmission where *regions of complexity disparity* can be autoidentified using the developed multifractal measures. Another component of this chapter is the development of a new generalized wavelet packet best basis cost function based on the Rényi generalized entropy. The following section begins this chapter with a brief introduction to progressive wavelet image transmission in the context of perceptually motivated measures.

7.1   Progressive Wavelet Image Transmission

As described in Chapter 2, progressive image transmission is the compression and transmission of image information such that the image can be roughly approximated at an early stage of reconstruction and then refined as more image information is received. As pointed out by Jayant *et al.* [JaJS93], image coding, and signal coding in general, has advanced to the point where researchers need to focus on the perceptual coding of these signals. For very low bit rates, good image coding algorithms will need perceptually motivated distortion measures. Similar perceptual coding efforts are done with other signals, such as digital audio signals as given in the review by Painter and Spanias [PaSp97].

Therefore, progressive image transmission also needs to take advantage of perceptually

motivated distortion measures.

808 bytes (0.308%)	1977 bytes (0.754%)	5262 bytes (2.01%)
20678 bytes (7.89%)	58173 bytes (22.19%)	96359 bytes (36.76%)
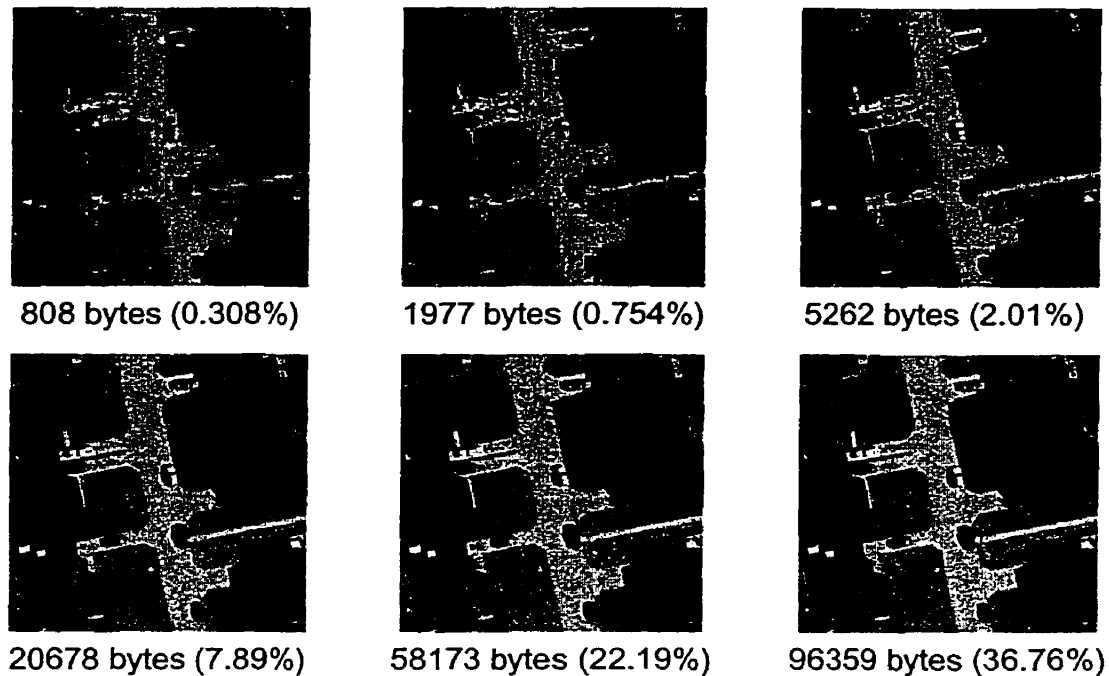
Fig. 7.1. Sample progressive transmission of the aerial ortho image
urban with size in bytes and % of original.

Current state of the art wavelet techniques include the EZW algorithm developed

by Shapiro [Shap93], [Shap93b], which was later generalized and improved by Said and

Pearlman as the SPIHT algorithm [SaPe96]. Example images from a progressive image

transmission using the EZW algorithm is shown in Fig. 7.1 for the image of urban. While

the EZW and SPIHT algorithms perform well, are computationally efficient, and mini-

mize the tractable distortion criteria such as MSE and PSNR [HöKS97], neither of these

algorithms consider any perceptually motivated distortion measures in the coding of the

images. They function by taking advantage of the structure of the wavelet decomposition and allocating bits for wavelet coefficients in an efficient manner.

Many researchers are looking at developing perceptually motivated progressive image transmission. One technique, as presented by Frajka *et al.* [FrSZ97], modifies the SPIHT algorithm by Said and Pearlman [SaPe96] so that it can focus on certain regions in the image during the embedded image coding. This focus on the regions within the image is performed by region masking, where important regions are identified and then the next few phases in the algorithm transmit bits only for this region.

Other efforts at modifying EZW and SPIHT algorithms were done by Höntsch *et al.* [HöKS97], where the wavelet coefficients are weighted by a JND factor using the Minkowsky metric. No side information is sent in this technique so it does not really adapt to specific image dynamics. The resulting wavelet coefficients are perceptually tuned to the Minkowsky metric, as opposed to the MSE, through the weighting of the wavelet coefficients.

Subband coding is another approach to wavelet image compression where each level of decomposition in the wavelet transform is considered a separate band, or subband, of dilations. This approach includes some of the first wavelet subband coding by Woods and O'Neil [WoON86], the DWT subband coding by Mallat [Mall89], and the DWT sub-band progressive image transmission by Jafarkhani and Farvardin [JaFa96]. The latter is a fast approach that can reconstruct the image as every wavelet coefficient arrives without increasing the computational cost. Some of the compression in progressive wavelet sub-band coding can be optimized by using techniques such as that given by Buccigrossi and

Simoncelli [BuSi97], which considers the joint statistics between subbands. Perceptual forms of subband coding have been introduced by Safranek and Johnson [SaJo89] with the perceptual subband image coder (PIC). Höntsch and Karam [HöKa97] present an adaptive form of the PIC algorithm that determines the JND for local wavelet coefficients to improve the image quality.

Another approach to wavelet based image compression follows the singularity detection of Mallat and Hwang [MaHw92]. This technique can represent a signal through the zero crossings in the wavelet transform, which form the edges within an image. Langi and Kinsner present one such method that encodes the singularities in a chain code representation that can reconstruct the edges in the image [LaKi95a], [LaKi95c].

Another area of research to improve image compression and progressive image transmission is image denoising. While not explicitly following perceptual models, this preprocessing step helps remove uncorrelated noise from an image, which makes the job of an image coder easier. Donoho introduced image denoising using wavelet shrinkage [Dono92], [Dono95]. Extensions to these ideas were made by Krim *et al.* [KTMD99] who looked at choosing wavelet packet best basis representations that can then have a hard-threshold applied instead of soft-thresholds [Dono92], [Dono95]. Denoising effects in wavelet based progressive image transmission techniques was looked at by Langi and Kinsner [LaKi95b], [Lang96], [Kins95b].

The next section considers the effects of enhancing regions within an image. The effects on the developed multifractal measures of enhancing regions are investigated along with how to identify these RCD using the multifractal measures.

7.2   Progressive Image Transmission with Enhanced Regions

One approach used for some progressive image transmission techniques to improve image quality is to identify regions that should be emphasized. During the encoding phase in the progressive image transmission, these regions are encoded first or encoded with greater precision initially, perhaps through a rebalancing of bit allocation in the quantization step of the coefficients. This special treatment of regions allows for the key portions of the image reconstruction to appear faster if the important regions are identified properly. An example of a region might be the head and shoulder portion of a passport image, which would generally be of more interest than the background of the image. The finer details of the background can therefore be delayed, while the head and shoulders are filled in with more details early in the progressive image transmission.

7.2.1   Encoder/Decoder Design for Enhanced Region Progressive Wavelet Image Transmission

A number of techniques exist for varying the bit allocation of regions for progressive wavelet image transmission. Shapiro presented a method for preprocessing and post-processing the wavelet coefficients to adjust the bit allocation of the EZW algorithm [Shap93b]. Frajka *et al.* [FrSZ97] present a method of extending the Said and Pearlman's SPIHT algorithm [SaPe96], which also works with Shapiro's EZW algorithm, by establishing an importance function for each of the wavelet coefficients. This importance function is used to effectively determine the priority of the bit allocation process for the wavelet coefficients in the image decomposition. Other adaptive techniques exist for pro-

gressive wavelet image transmission, but, this section focuses on the effects of enhancing regions on the IQM multifractal measure developed in Sec. 6.6.

To set up an experimental environment for this section, a simple progressive wavelet image transmission technique is needed that allows for the enhancement of regions at an early stage in the progressive image transmission. The data flow diagram in Fig. 7.2 illustrates the environment used for generating images with enhanced regions. In the progressive image transmission technique shown in Fig. 7.2, an image is first decomposed with the DWT. Afterwards, the wavelet coefficients from the DWT are taken and then the coefficients that spatially map back to specified regions are scaled by a set of multipliers. The goal is to increase the magnitude of these wavelet coefficients corresponding to the region so that the EZW encoder allocates bits to these coefficients earlier than it would otherwise. Of course, the scaling of the wavelet coefficients must be done across the different subbands in the wavelet decomposition, following the parent/child relationship of the wavelet coefficients. The coefficient scaling may also be different from subband to subband for each region. After the EZW encoding of the scaled wavelet coefficients, the data stream is further losslessly compressed with an adaptive arithmetic coder [WiNC87]. Some sideband information must also be sent in the data stream including the location of the region as well as the scaling factors used for the region. The reconstruction of the image is the reverse of the process illustrated in Fig. 7.2.
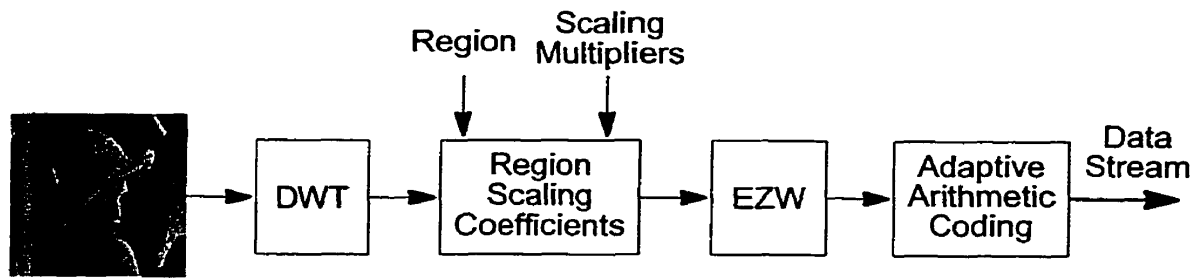
Fig. 7.2. Data flow of the test progressive wavelet image transmission coder with enhanced regions.

7.2.2   Relative Rényi Dimension Spectrum Experiments on Enhanced Regions

To show how the image compression scheme depicted in Fig. 7.2 improves the image quality of selected regions, a number of experiments are performed on the image of lena with three enhanced regions, as shown in Fig. 7.3. The experiments consider the image of lena with no enhanced regions, the image of lena with the two enhanced regions around the eyes, and the image of lena with the face as the enhanced region.
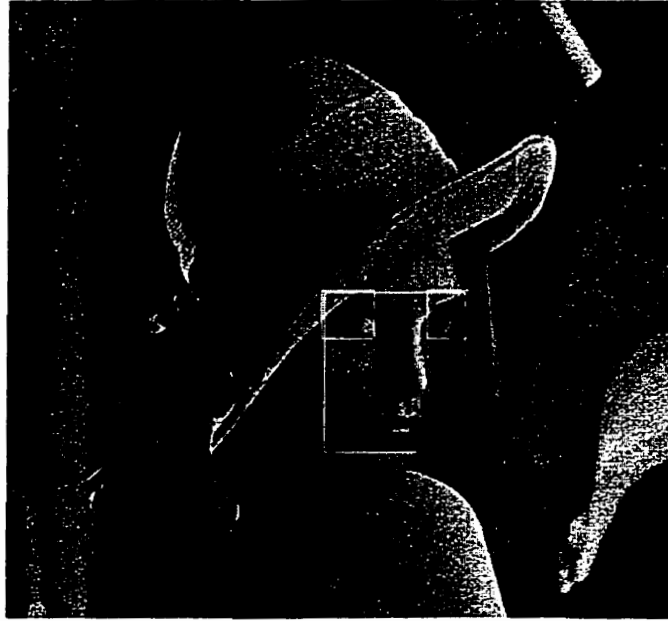
Fig. 7.3. Image of **lena** with three selected regions,
namely the two eyes and face.

Using this experimental setup, the image of **lena** was compressed and recon-
structed using the encoder/decoder described in Sec. 7.2.1. Cropped versions of the
resulting reconstructed images are given in Fig. 7.4 and Fig. 7.5 at 0.05 bpp and 0.1 bpp,
respectively. These two figures each show six cropped images of **lena** consisting of

- a. the original image of **lena**,

- b. results when using no enhanced regions,

- c. results with the eyes enhanced and the wavelet coefficients multiplied by 2.0,

- d. results with the face enhanced and the wavelet coefficients multiplied by 2.0,

- e. results with the eyes enhanced and the wavelet coefficients multiplied by 4.0,

- f. and results with the face enhanced and wavelet coefficients multiplied by 4.0.

Looking at these images, a number of observations can be made. First, looking at the image in Fig. 7.4b versus Fig. 7.4c and Fig. 7.4d, it is noticed that the eyes have more details in Fig. 7.4c then in Fig. 7.4b, and that the entire face in Fig. 7.4d has marginally more details than in Fig. 7.4b. Both Fig. 7.4c and Fig. 7.4d have some features that are not as coarse than in Fig. 7.4b, particularly around the eyes and the tip of the nose. These results follow what is expected, since multiplying the wavelet coefficients by 2.0 for the region containing the eyes and face emphasizes these features slightly at an earlier stage of the encoding. This emphasis is more prominent when the wavelet coefficients are multiplied by 4.0 as evidenced in Fig. 7.4e and Fig. 7.4f. Finer details are brought into the regions of the eyes for both reconstructions. Also, the nose, nostril, cheek, and mouth area in Fig. 7.4f have finer details than in Fig. 7.4b when no enhanced regions are used. All of these observations can similarly be made about the respective reconstructions at 0.1 bpp in Fig. 7.5. The eyes and face in Fig. 7.5 are better defined with some finer resolution components around the eyes, nose, mouth, and cheek.
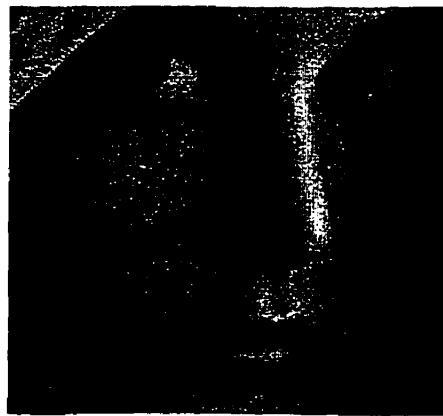
(a) Original image

(b) 0.05 bpp with no enhancement

(c) 0.05 bpp with eyes enhanced * 2.0

(d) 0.05 bpp with face enhanced * 2.0

(e) 0.05 bpp with eyes enhanced * 4.0

(f) 0.05 bpp with face enhanced * 4.0

Fig. 7.4. Cropped image reconstructions of lena with regions compressed
at 0.05 bpp. (a) Original image, (b) no enhancement, (c) eyes enhanced * 2.0,
(d) face enhanced * 2.0, (e) eyes enhanced * 4.0, and (f) face enhanced * 4.0.
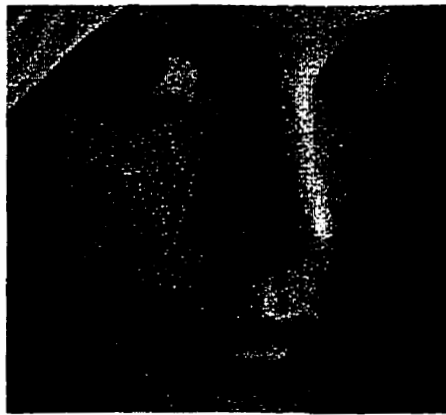
(a) Original image
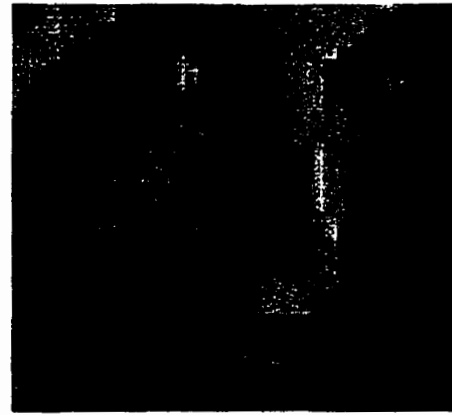
(b) 0.1 bpp with no enhancement
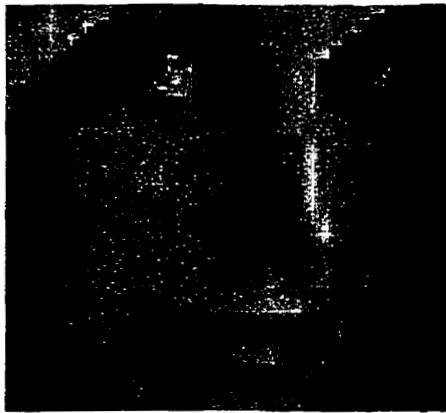
(c) 0.1 bpp with eyes enhanced * 2.0

(d) 0.1 bpp with face enhanced * 2.0
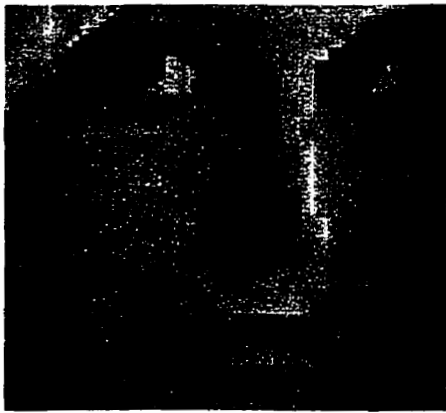
(e) 0.1 bpp with eyes enhanced * 4.0

(f) 0.1 bpp with face enhanced * 4.0

Fig. 7.5. Cropped image reconstructions of lena with regions compressed
at 0.1 bpp. (a) Original image, (b) no enhancement, (c) eyes enhanced * 2.0,
(d) face enhanced * 2.0, (e) eyes enhanced * 4.0, and (f) face enhanced * 4.0.

With the images in Fig. 7.4 and Fig. 7.5, the next set of experiments is to take the generated image reconstructions and apply the relative Rényi dimension spectrum $RD_q(u \| v)$ from Sec. 6.6. These experiments are performed in two sets. The first set of experiments applies $RD_q(u \| v)$ on the uncropped versions of the images with the reconstructed image as v and the original image as u. The second set of experiments applies $RD_q(u \| v)$ on the cropped versions of the images, as shown in Fig. 7.4 and Fig. 7.5, again with v and u as the reconstructed and original images, respectively. Performing the $RD_q(u \| v)$ analysis in this manner allows us to see how the measure is affected over the entire image versus only in the area of the enhanced region.

The results of the first set of experiments using the full images of lena in the $RD_q(u \| v)$ analysis are plotted in Fig. 7.6 and Fig. 7.7. The plots in Fig. 7.6 reflect the $RD_q(u \| v)$ analysis results using the full images from Fig. 7.4, while the plots in Fig. 7.7 reflect the $RD_q(u \| v)$ analysis results using the full images from Fig. 7.5. Analyzing Fig. 7.6 it is apparent that the image reconstruction with no region enhancement and the image reconstructions with the eye and face enhanced scaled by 2.0 have quite similar $RD_q(u \| v)$ versus q curves. For the image reconstructions with the enhanced regions scaled by 4.0 the $RD_q(u \| v)$ versus q curves move away from the abscissa. This departure indicates that the overall image reconstruction complexity is degraded when the enhanced regions are included. This result is expected since using more bits for the enhanced region means that fewer bits are available for the rest of the image. Therefore, the overall multifractal image complexity is degraded when the regions are included, even though the specific regions have improved quality. This observation corresponds to the

PSNR results obtained by Frajka *et al.* [FrSZ97] who observed the same change when enhanced regions are added to an image.

Analyzing Fig. 7.7 where the bit rate is 0.1 bpp, the $RD_q(u \parallel v)$ versus q curves are all very similar. The use of extra bits in the progressive image transmission has brought the $RD_q(u \parallel v)$ versus q curves closer together compared to Fig. 7.6, which means that the overall multifractal complexities in the five images are similar. While difficult to notice properly in the plot, for $q > 0$ all of the images with enhanced regions are marginally closer to the abscissa than the image with no enhancement, except for the image with the eyes as the enhanced region using a scaling of 4.0. This result indicates that the majority of the images with enhanced regions follow the overall multifractal complexity of the original quite well, but that the enhancement of the region has redistributed a few additional bits to actually improve the overall image complexity disparity. This effect is an interesting reversal from the plots in Fig. 7.6, since it shows that at some point in the transmission the overall image compleixty is improved by including the enhanced regions. This reversal is not noticed in the PSNR results of Frajka *et al.* [FrSZ97]. The curve that is further from the abscissa compared to the one for the image with no enhancement indicates that the balance has not yet been achieved in the number of bits allocated to the smaller regions versus the entire image. More multifractal complexity details are needed for the rest of the image at this enhancement scaling before this reversal is obtained. These same observations can be made for $q < 0$, but, it is less clear since some of the rotating line segment effect noticed in Sec. 6.6 is visible, starting roughly at the $q < -5.0$ mark.
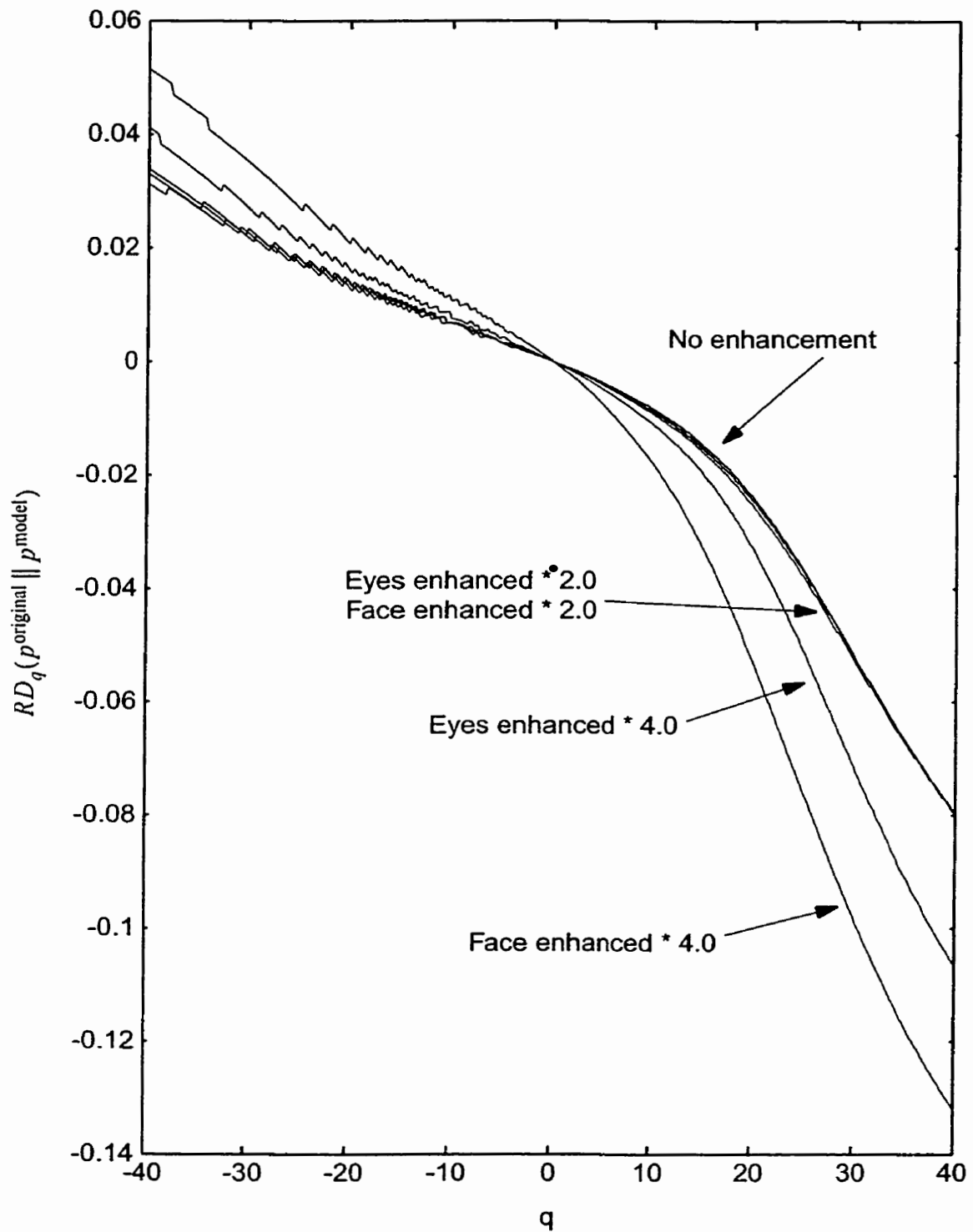
Fig. 7.6. Plot of relative Rényi dimension spectrum versus q for Daub4 EZW reconstruction of lena at 0.05 bpp with no enhancement, eyes enhanced, and face enhanced using a scaling of 2.0 and 4.0.
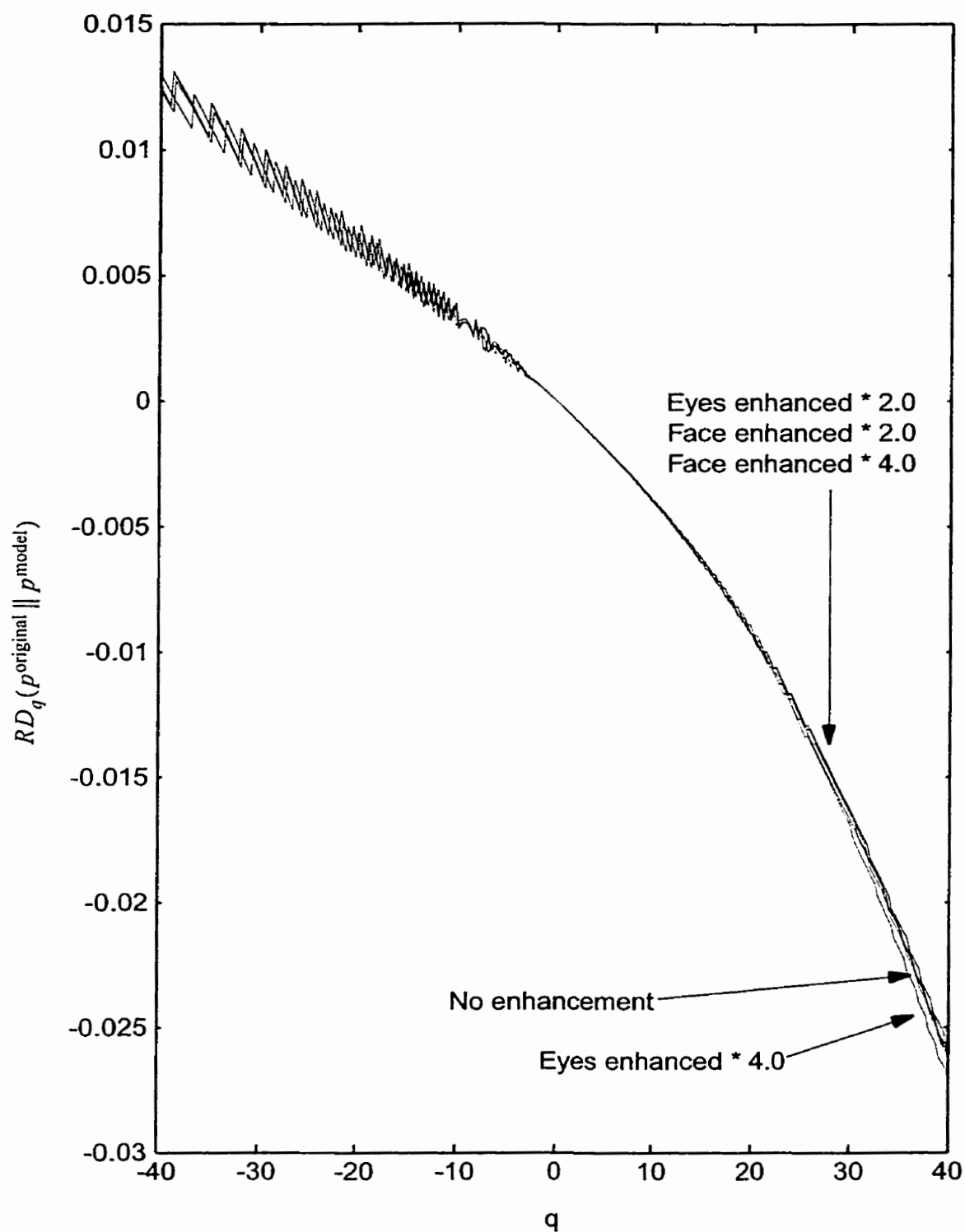
Fig. 7.7. Plot of relative Rényi dimension spectrum versus q for Daub4 EZW reconstruction of lena at 0.1 bpp with no enhancement, eyes enhanced, and face enhanced using a scaling of 2.0 and 4.0.

The next set of experiments calculate the relative Rényi dimension spectrum $RD_q(u \| v)$ of the cropped images of lena, as opposed to the full image reconstructions. The same experiments are performed as in Fig. 7.6 and Fig. 7.7 with these cropped versions of the images for u and v, respectively. The $RD_q(u \| v)$ versus q results are plotted for a bit rate of 0.05 bpp and 0.1 bpp in Fig. 7.8 and Fig. 7.9, respectively.

For the results at 0.05 bpp, Fig. 7.8 shows that three of the four curves with enhanced regions converge more towards the abscissa than the cropped reconstruction without any enhancements. Since the cropping limits the analysis to the region around the face of lena, these results are expected since this region has better quality. Here the increase in quality is due to the greater number of bits allocated to this region. Of surprise is that the curve for the reconstruction with the face as the enhanced region at a scaling of 4.0 is further from the abscissa. This result is not expected since Fig. 7.4f appears to be of better perceptual quality than all of the other image reconstructions in Fig. 7.4. Thus, the larger distance from the abscissa is possibly due to some extra edge artifacts in the image reconstruction that add an extra level of multifractal complexity which should not be present.
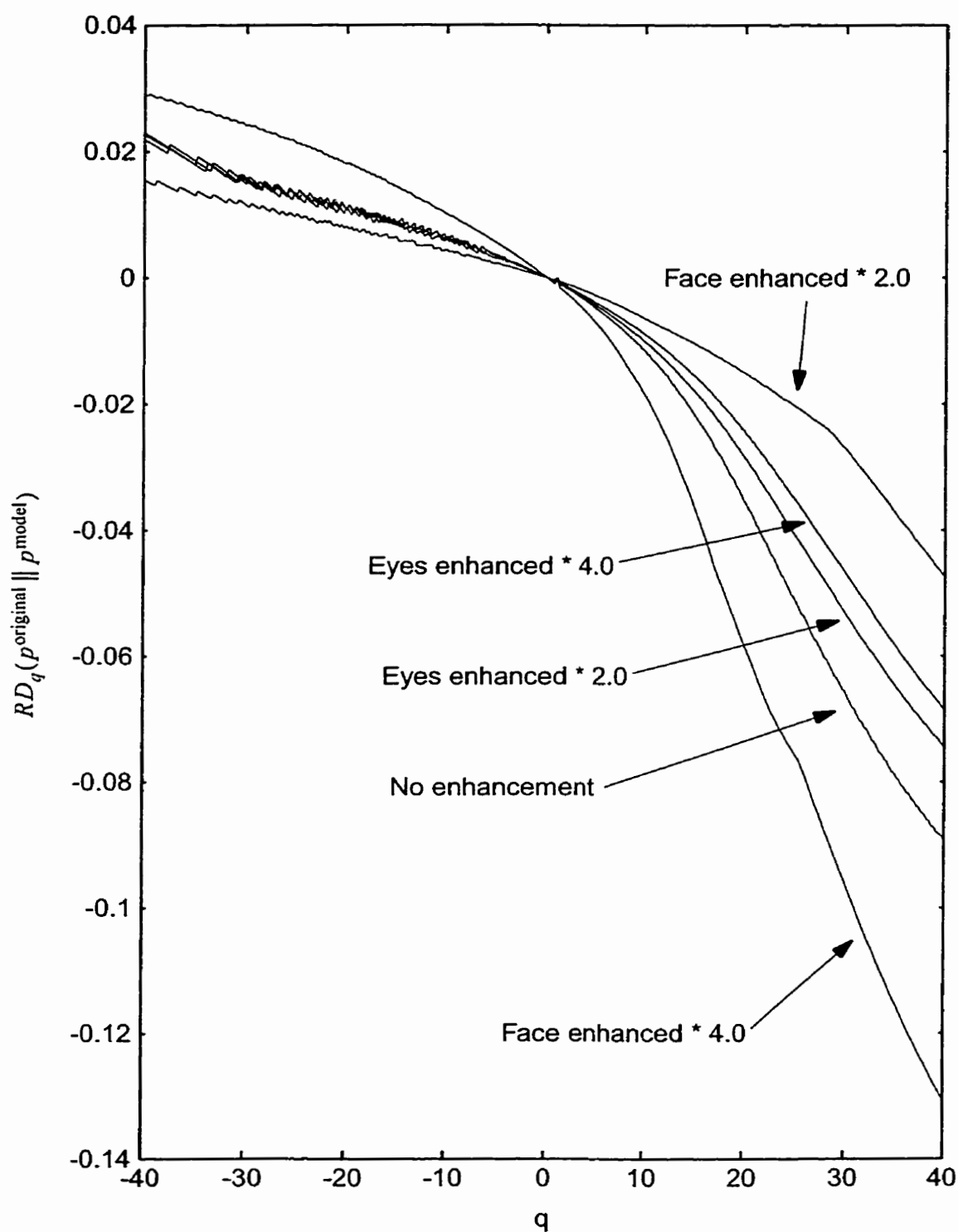
Fig. 7.8. Plot of relative Rényi dimension spectrum versus q for cropped Daub4 EZW reconstruction of lena at 0.05 bpp with no enhancement, eyes enhanced, and face enhanced using a scaling of 2.0 and 4.0.
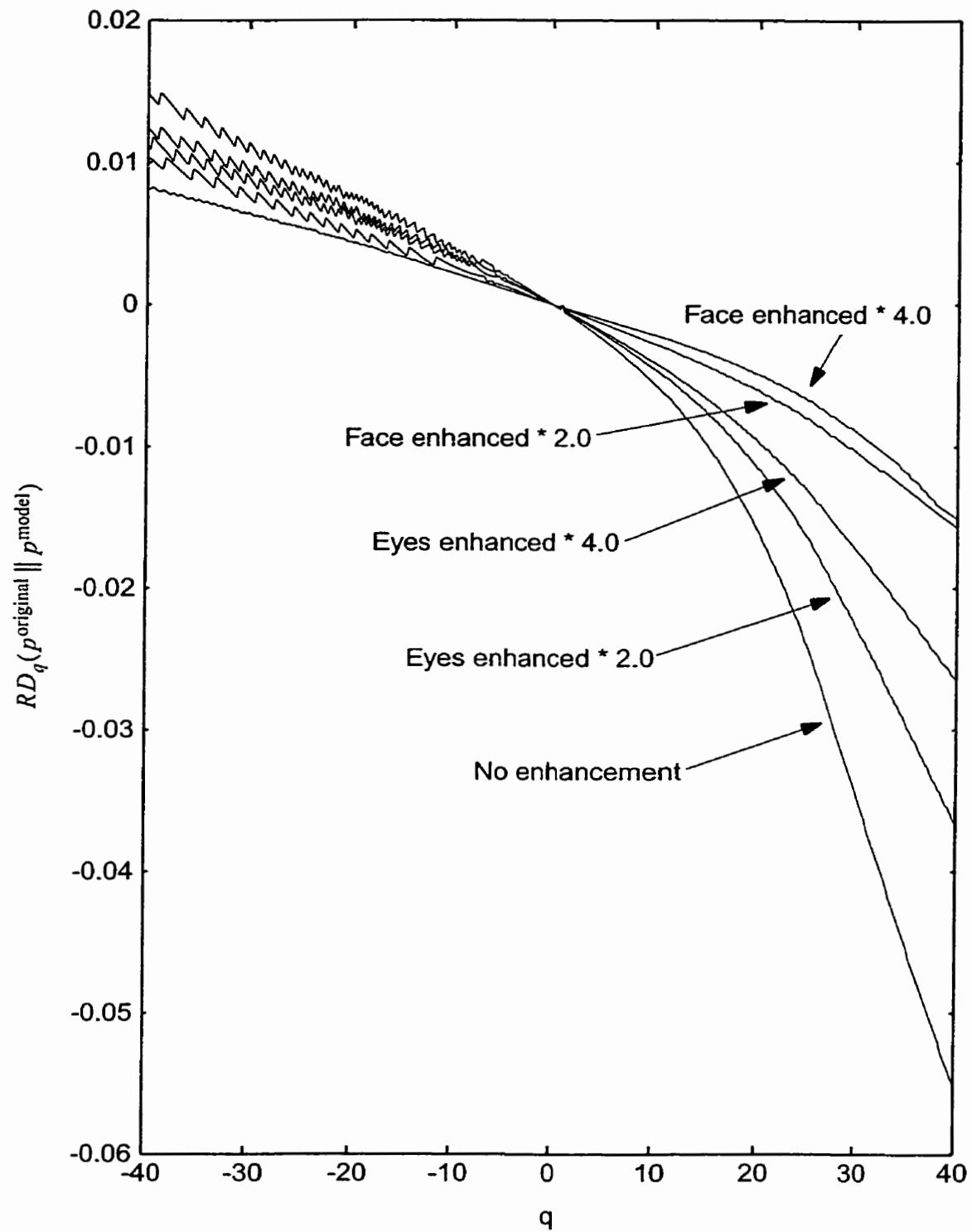
Fig. 7.9. Plot of relative Rényi dimension spectrum versus q for cropped Daub4 EZW reconstruction of lena at 0.1 bpp with no enhancement, eyes enhanced, and face enhanced using a scaling of 2.0 and 4.0.

For the results at 0.1 bpp, Fig. 7.9 shows much more of what is expected of all four images with enhanced regions. All four curves are shifted closer to the abscissa as compared to the image reconstruction with no enhancements. Another interesting observation is that the two curves for enhanced regions around the eyes and the two curves for enhanced regions around the face are grouped together with a slight improvement when going from a scaling of 2.0 to 4.0. This ordering of the face and eye curves is expected, since measuring the face when only the eyes are enhanced would only marginally change the $RD_q(u \parallel v)$ value for the cropped image.
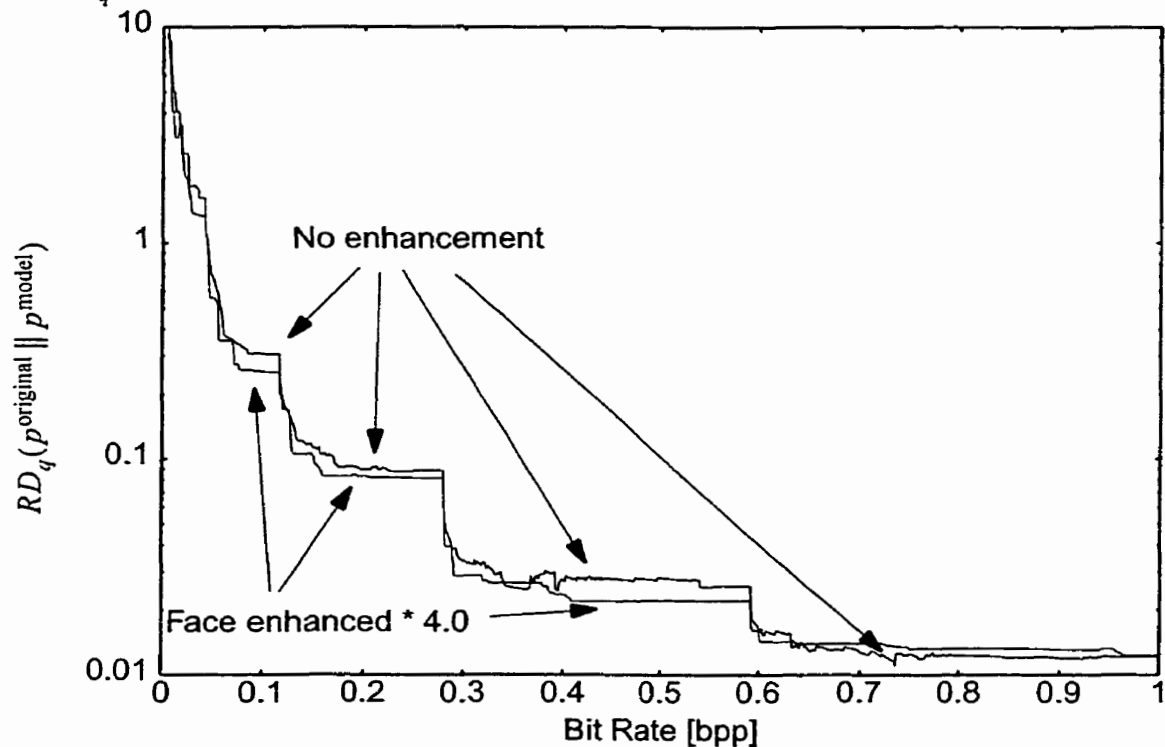
Fig. 7.10. IQM measure versus compression rate for image of lena with and without enhanced regions around face.

Another experiment conducted looks at how the IQM analysis developed in Sec. 6.6 is affected by including some enhanced regions in an image. The experiment is

conducted by applying the IQM analysis on the reconstructed image with the enhanced regions around the face included. This IQM analysis is applied on the full image as well as the cropped version of the image over a range of bit rates, with the results plotted in Fig. 7.10. This plot shows that the IQM value is lower for the cropped image at the majority of bit rates from 0.005 to 0.6 bpp. This result indicates that the enhanced image is of closer signal complexity throughout the transmission until roughly the 0.640 bpp mark, where enough bits have been allocated to the rest of the image to improve the complexity representation compared to the enhanced image. This result also shows that 0.640 bpp is the point when the enhanced image and the full image have balanced out in terms of overall image complexity from the standpoint of the IQM analysis.

(a) IQM map of lena   (b) Histogram equalized
                                          IQM map of lena
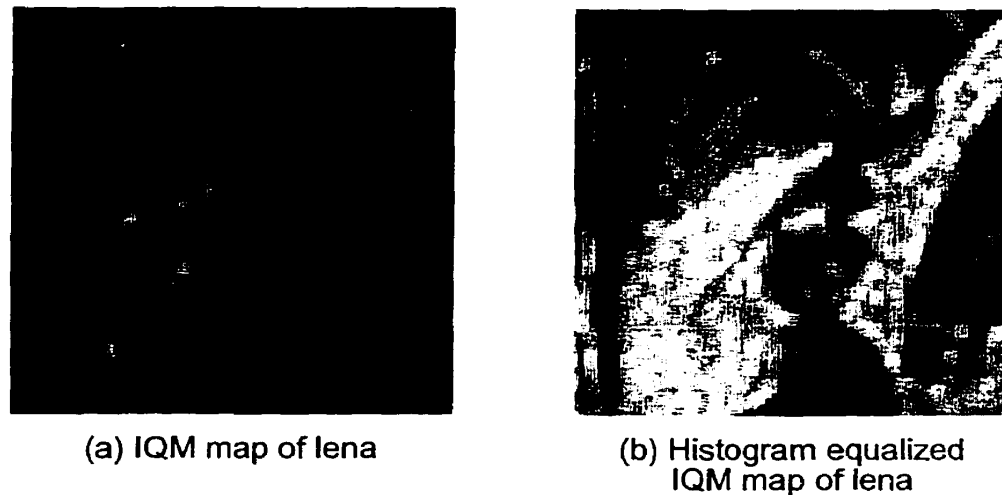
Fig. 7.11. IQM map of image of lena at 0.1 bpp using a 16x16 pixel sliding window. (a) IQM map of lena, (b) Histogram equalized IQM map of lena.

Following up further on the IQM experiments, one view of the image that is interesting is to consider how the IQM value changes over different parts of the image. An experiment was set up to map the IQM values by calculating the IQM value on a 16×16 pixel sliding window over the entire image. The resulting IQM map is given in Fig. 7.11a

for the image of lena reconstructed with no enhancement at 0.1 bpp. A histogram equalized version of Fig. 7.11a is given in Fig. 7.11b to provide more detail by using the full pixel range available. The darker regions in these two IQM maps are regions where the IQM value indicates that little difference exists between the original and reconstructed images. The brighter areas indicate those regions where finer details are required to improve the complexity disparity between the original and reconstructed images. From Fig. 7.11a, it is apparent that the feathers, eyes, and mirror trim in the image require the addition of finer details to further improve the image complexity from an IQM analysis standpoint. The histogram equalized version of the map, given in Fig. 7.11b, shows that the more textured areas of the image and regions where edges exist also require further details to be included to improve the image complexity from an IQM analysis standpoint.

7.2.3   Automatic Identification of Regions of Complexity Disparity

One problem with using enhanced regions for the improvement of overall image complexity is that these regions usually need to be chosen by hand. It would be useful to have an automated method of identifying regions where the complexity differs significantly between the two images being compared. These regions will be refered to as regions of complexity disparity (RCD). The IQM analysis introduced in Sec. 6.6 can be used to perform this identification of RCD. This identification is done by using the IQM mapping approach described for generating Fig. 7.11 in Sec. 7.2.2. With this approach, a map of the IQM values throughout the image can be established at specified resolutions, and then the regions with the highest IQM values identified as the RCD.

Using this approach of IQM mapping, the five test images of lena, urban, baboon, peppers, and farm are mapped, and the dominant RCD identified in Fig. 7.12, Fig. 7.13, Fig. 7.14, Fig. 7.15, and Fig. 7.16, respectively. Each of these figures show the RCD identification for 8, 32, and 128 windows of size 64x64, 32x32, and 16x16, respectively.

Some observations can be made about the RCD identification in Fig. 7.12 through Fig. 7.16. For lena, at a window size of 64x64 the RCD identification chooses important edges such as the shoulder, hat, and hair. At the finer resolution of 32x32, the RCD identification continues to identify edges around the hat, shoulder, cheek, and mirror as important, as well as one of the eyes. This RCD identification continues at 16x16 with more RCD around the eyes, and edges around the hat, border of the feathers, and the mirror. The IQM mapping has returned primarily edge components as the RCD, and has ignored most of the finer textures in the interior of the hat and feathers.

(a)   (b)   (c)
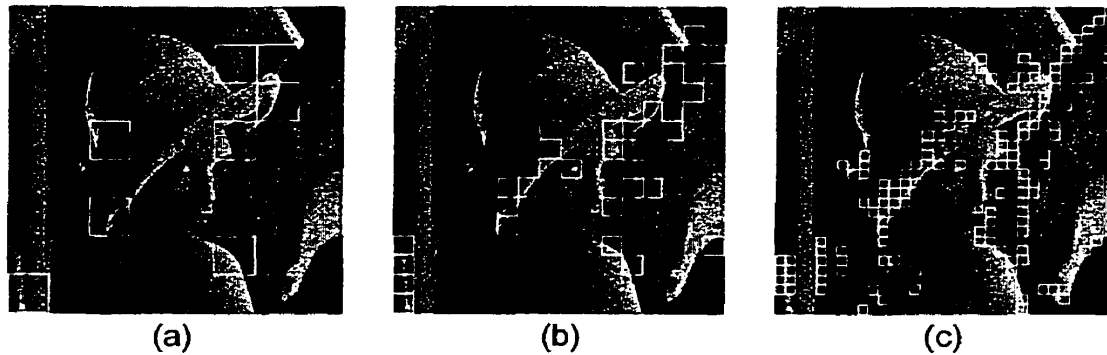
Fig. 7.12. Results of identification of regions of complexity disparity using the IQM map for the image of lena at scales of (a) 64x64, (b) 32x32, and (c) 16x16.
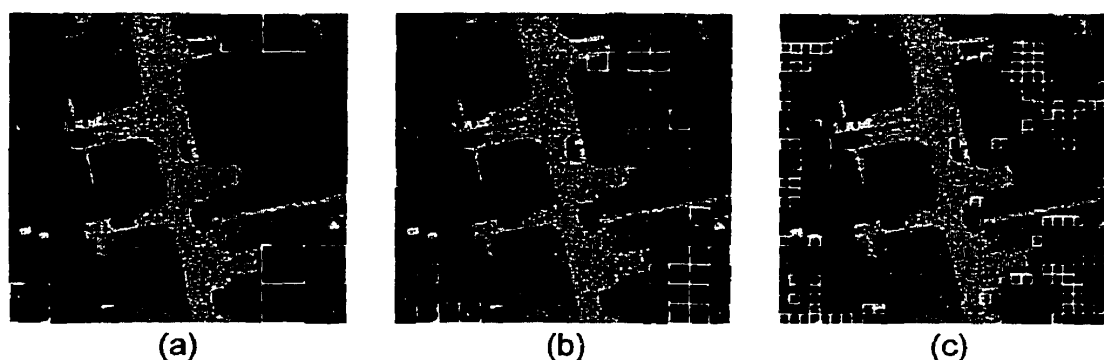
Fig. 7.13. Results of identification of regions of complexity disparity using the IQM map for the image of urban at scales of (a) 64x64, (b) 32x32, and (c) 16x16.
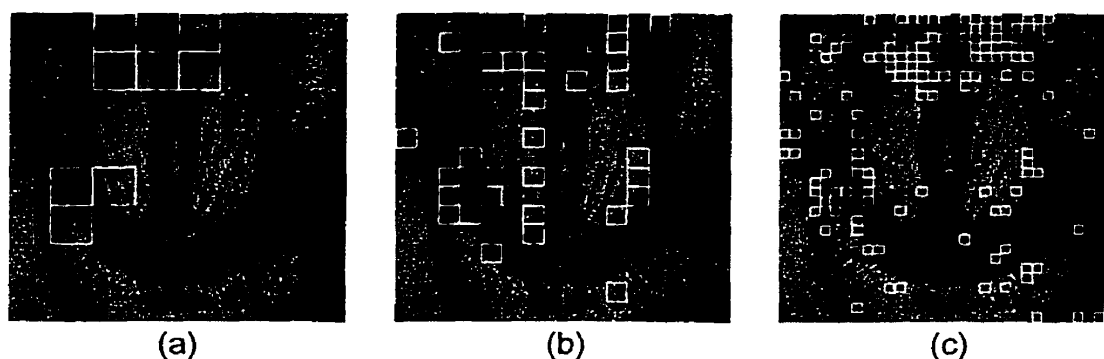
Fig. 7.14. Results of identification of regions of complexity disparity using the IQM map for the image of baboon at scales of (a) 64x64, (b) 32x32, and (c) 16x16.

The RCD identification in the image of urban tends to focus on the houses in the image. The cars only begin to be identified with a window size of 16x16 and the roads are not identified at all, though, the roads are already relatively sharp. This example may show a weakness in the IQM analysis in that, for instance, the house in the top right of the image is not significantly perceptually different from the original. The IQM analysis can discern features past the level of human JND, which suggests that some extra preprocess-

ing for JND features, perhaps through a smoothing filter, should be done with the images

before performing IQM mapping.

The RCD identification in the image of baboon shows a strong preference for the

area around the eyes as well as along the nose at the border of the hair. It is interesting to

note that few regions are identified within the hair of the baboon.

The RCD identification in the image of peppers shows a strong preference for the

edges around each of the peppers, particularly where there is a sharp transition from light

to dark. This result agrees with what is known of the importance of edges in human per-

ception [JaJS93].

(a)   (b)   (c)
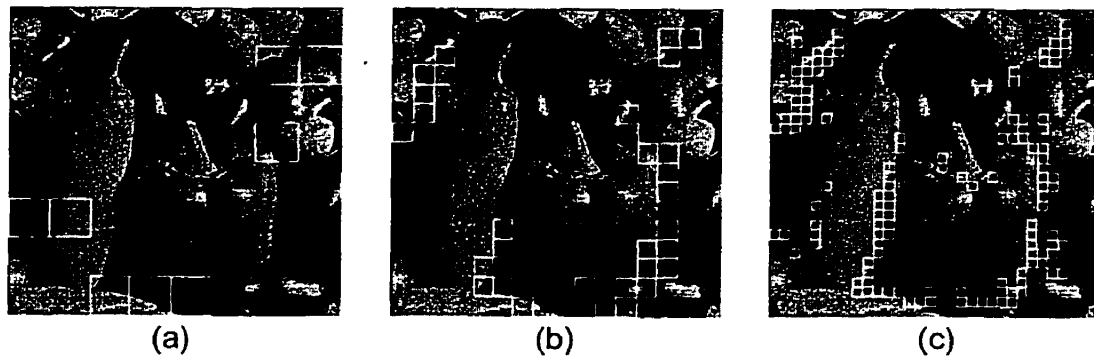
Fig. 7.15. Results of identification of regions of complexity disparity using the IQM map
for the image of peppers at scales of (a) 64x64, (b) 32x32, and (c) 16x16.
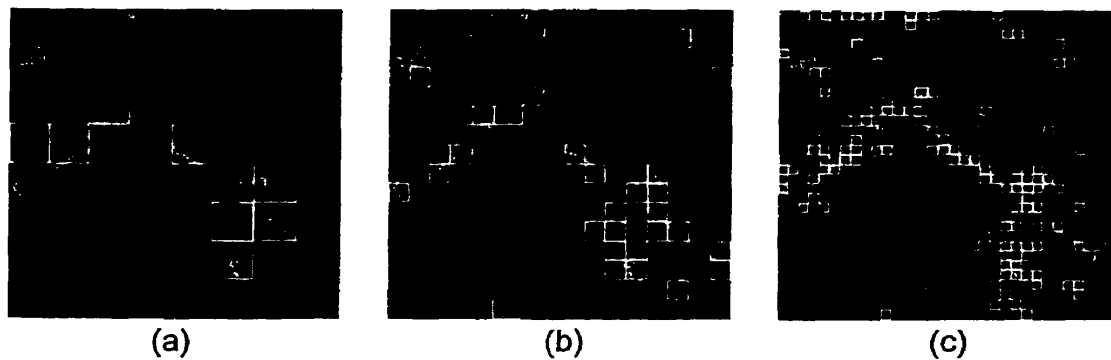
Fig. 7.16. Results of identification of regions of complexity disparity using the IQM map for the image of farm at scales of (a) 64x64, (b) 32x32, and (c) 16x16.

Finally, the image of farm shows that most of the RCD identification is in the regions with the houses and roads. The finer details of the fields are mostly ignored and treated as having a multifractal complexity much closer to that of the original image. It is the more noticeable changes in the image· that are identified as RCD.

From the experiments conducted·, IQM analysis shows that edges and areas of more irregular changes are identified in these wavelet compressed images as the RCD. The next section looks at a different problem in signal decomposition, namely that of finding the best basis wavelet packet representation of a signal.

7.3   Wavelet Packet Best Basis Search using the

Rényi Generalized Entropy

Wavelet packets, as discussed in Sec. 4.4, are a generalization of the wavelet transform such that detail portions of the signal are also decomposed in the same manner as approximation portions of the signal (refer to Fig. 4.10). The advantage behind using

wavelet packets in signal approximation is that choosing this decomposition in an adaptive manner according to the signal may yield better results than the wavelet transform itself. As indicated in Sec. 4.4.2, an issue with using wavelet packets is deciding how many levels of decomposition should be performed for each of the approximation and detail signals for a particular signal. In other words, what is the stopping criteria that indicates the level of wavelet decomposition.

A number of approaches are available for choosing a wavelet packet decomposition, or wavelet packet basis as it is often referred. A general approach is to form a dictionary of all wavelet packet decompositions, and then perform an exhaustive search for the basis representation that minimizes some criteria. For a binary tree decomposition of a time series, as in Fig. 4.10, or a quaternary tree decomposition for a 2D signal such as an image, as in Fig. 4.12, this function basis dictionary grows exponentially the deeper the tree is decomposed. With this exponential growth, an exhaustive search of all bases in the dictionary tree for the best basis is computationally prohibitive.

Coifman and Wickerhauser [CoWi92], [Wick94] proposed an efficient search algorithm for the best basis in the dictionary tree by use of additive cost measures. Other search algorithms include that by Taswell [Tasw94], [Tasw95] which finds near-best bases according to non-additive cost functions, as well as pursuit algorithms such as the greedy approach in the matching pursuit algorithm by Mallat and Zhang [MaZh93].

In this section, our interest is with extending the entropy based best basis search algorithm as proposed by Coifman and Wickerhauser [CoWi92]. The approach described first defines the idea of an additive cost function as follows.

Definition 7.1: A map $C(x_i, \phi)$ from sequences $\{x_i\}$ to $\mathbf{R}$ for the orthonormal basis ϕ

is called an additive information cost function if $C(0, \phi) = 0$ and

$C(\{x_i\}, \phi) = \sum_i C(x_i, \phi)$.

From this definition, Coifman and Wickerhauser continue by specifying that a best

basis representation of a signal is a basis from the dictionary that minimizes $C(f, \phi)$ for

the function f. If an additive information cost function is defined for an orthonormal basis

ϕ_j, taken at decomposition level j, then it follows from Defn. 7.1 that

$$C(f, \phi_j^0 \cup \phi_j^1) = C(f, \phi_j^0) + C(f, \phi_j^1) \tag{7.1}$$

where $(\phi_j^0, \phi_j^1) \subset \phi_j$. The search algorithm that follows from this additive property is to

search through the tree representation of the dictionary and determine if a parent node has

a cost function greater or less than the sum of the cost functions of the children nodes. If

the following inequality holds for the parent node and the n children nodes in a decompo-

sition

$$C(f, \phi_j^i) > C(f, \phi_{j+1}^{2i}) + \dots + C(f, \phi_{j+1}^{2i+n}) \tag{7.2}$$

then the children $\phi_{j+1}^{2i} \cup \dots \cup \phi_{j+1}^{2i+n}$ are chosen as part of the wavelet packet decomposi-

tion of that portion of the signal. On the other hand, if

$$C(f, \phi_j^i) \leq C(f, \phi_{j+1}^{2i}) + \dots + C(f, \phi_{j+1}^{2i+n}) \tag{7.3}$$

then the parent ϕ_j^i is chosen as the wavelet packet decomposition for that portion of the

signal. An example best basis for a cost function $C(f, \phi)$ is illustrated in Fig. 7.17, which

is a simplified version of Fig. 4.10. The shaded regions illustrate the selection of a wavelet

packet decomposition of the signal into coefficients at various levels of ϕ_j^i. For instance,

the comparison $C(f, \phi_3^0) > C(f, \phi_4^0) + C(f, \phi_4^1)$ results in the children nodes ϕ_4^0 and ϕ_4^1

being selected for the wavelet decomposition. Also, $C(f, \phi_2^2) \le C(f, \phi_3^4) + C(f, \phi_3^5)$ results in the parent ϕ_2^2 being selected for the wavelet packet decomposition.
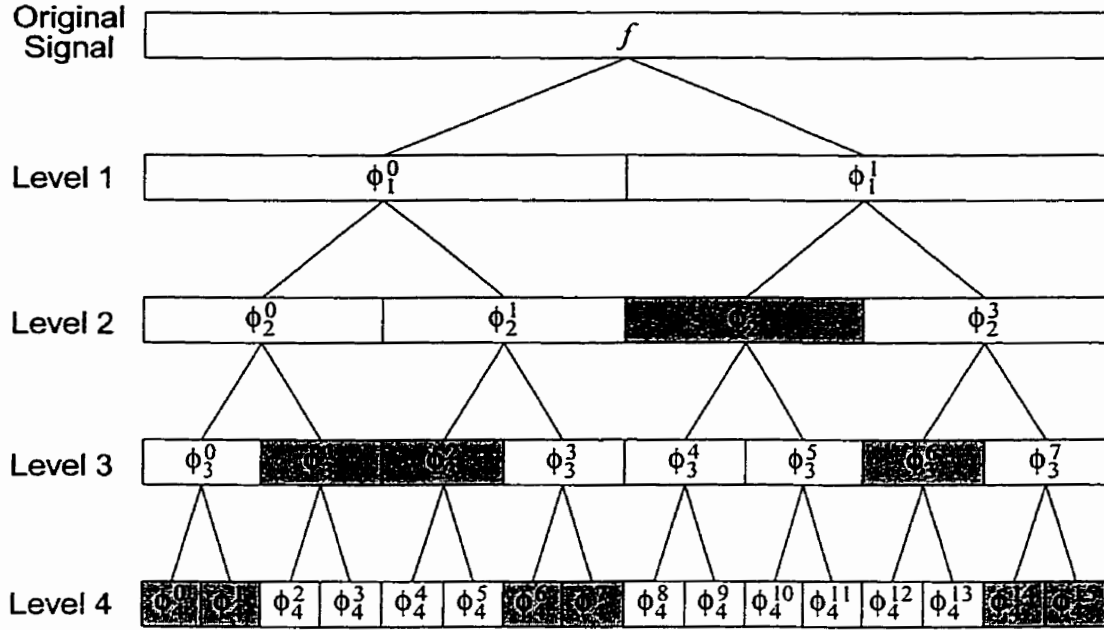
Fig. 7.17. Example best basis, shown in shaded regions, for a one dimensional wavelet packet decomposition.

The additive information cost function proposed by Coifman and Wickerhauser [CoWi92] can be expressed as follows [Mall99]

$$C(f, \psi_{u,s}) = -\sum_u \frac{|\langle f, \psi_{u,s} \rangle|^2}{\|f\|^2} \log\left(\frac{|\langle f, \psi_{u,s} \rangle|^2}{\|f\|^2}\right). \tag{7.4}$$

This is effectively the Shannon entropy from Eq. 3.19 with a probability distribution of

$$p(x) = \frac{|\langle f, \psi_{u,s} \rangle|^2}{\|f\|^2}. \tag{7.5}$$

For the additivity requirement in Eq. 7.1 to hold with Eq. 7.4, then the following must be true of the Shannon entropy for the probability distributions X and Y

$$H(X \cup Y) = H(X) + H(Y). \tag{7.6}$$

This relationship is shown to be true as follows

$$H(X \cup Y) = - \sum_{x \cup y \in X \cup Y} p(x \cup y) \log p(x \cup y) \tag{7.7}$$

$$= - \sum_{x \in X} p(x) \log p(x) - \sum_{y \in Y} p(y) \log p(y) \tag{7.8}$$

$$= H(X) + H(Y). \tag{7.9}$$

Therefore, the cost function in Eq. 7.4 is additive since it has the same form as Shannon entropy.

From this entropy based approach for an additive cost function, there is potential for using the Rényi generalized entropy in place of the Shannon entropy. The concept of additivity for entropy was generalized by Rényi [Rény55], [Rény59], [Rény60] for incomplete probability distributions such that

$$H(X \cup Y) = \frac{W(X)H(X) + W(Y)H(Y)}{W(X) + W(Y)} \tag{7.10}$$

where, as given in Eq. 3.21, $W(X) = \sum_{x \in X} p(x)$. This generalization remains consistent with the idea of additivity in Eq. 7.1 and therefore can be used for evaluating functions as possible additive cost functions for the best basis search algorithm of Coifman and Wickerhauser [CoWi92]. Another note about Eq. 7.10 is that it reformulates the additive relationship into an arithmetic mean of the weighted values of $H(X)$ and $H(Y)$.

A generalization of the arithmetic mean in Eq. 7.10 is also made by Rényi for a more general case and states that

$$H(X \cup Y) = g^{-1}\left[\frac{W(X)g(H(X)) + W(Y)g(H(Y))}{W(X) + W(Y)}\right] \qquad (7.11)$$

where $y = g(x)$ is an arbitrary strictly monotonic and continuous function and $x = g^{-1}(y)$ is its inverse [Rény60]. Instead of the arithmetic mean used for Eq. 7.10, an exponential mean in the form of

$$g_{q,b}(x) = b^{(1-q)x} \qquad (7.12)$$

where $b \in \mathbf{R}$, can also be used which maintains the relationship of Shannon entropy in Eq. 7.11. In addition, using Eq. 7.12 then the Rényi generalized entropy meets the requirement of Eq. 7.11 over all orders of q and can be expressed as

$$H_q(X \cup Y) = g_{q,b}^{-1}\left[\frac{W(X)g_{q,b}(H_q(X)) + W(Y)g_{q,b}(H_q(Y))}{W(X) + W(Y)}\right]. \qquad (7.13)$$

This relationship for the Rényi generalized entropy over all orders of q is shown to be true when using Eq. 7.12 as follows

$$H_q(X \cup Y) = g_{q,b}^{-1}\left[\frac{W(X)g_{q,b}(H_q(X)) + W(Y)g_{q,b}(H_q(Y))}{W(X) + W(Y)}\right] \qquad (7.14)$$

First calculating $g_{q,b}(H_q(X))$ gives

$$g_{q,b}(H_q(X)) = g_{q,b}\left(\frac{1}{1-q}\log_b\frac{\sum\limits_{x\in X}p^q(x)}{W(X)}\right) \tag{7.15}$$

$$= b^{(1-q)\frac{1}{1-q}\log_b\frac{\sum\limits_{x\in X}p^q(x)}{W(X)}} \tag{7.16}$$

$$= \frac{\sum\limits_{x\in X}p^q(x)}{W(X)}. \tag{7.17}$$

Replacing $g_{q,b}(H_q(X))$ and $g_{q,b}(H_q(Y))$ in Eq. 7.14 with Eq. 7.17, the following is obtained

$$H_q(X\cup Y) = g_{q,b}^{-1}\left[\frac{W(X)\dfrac{\sum\limits_{x\in X}p^q(x)}{W(X)} + W(Y)\dfrac{\sum\limits_{y\in Y}p^q(y)}{W(Y)}}{W(X)+W(Y)}\right] \tag{7.18}$$

which reduces to

$$H_q(X\cup Y) = g_{q,b}^{-1}\left[\frac{\sum\limits_{x\in X}p^q(x) + \sum\limits_{y\in Y}p^q(y)}{\sum\limits_{x\in X}p(x) + \sum\limits_{y\in Y}p(y)}\right]. \tag{7.19}$$

The inverse of Eq. 7.12 can be solved x to give $g_{q,b}^{-1}$ as follows

$$g_{q,b}^{-1}(y) = x = \frac{1}{1-q}\log_b y. \tag{7.20}$$

Substituting $g_{q,b}^{-1}$ into Eq. 7.19 gives the following

$$H_q(X \cup Y) = \frac{1}{1-q} \log_b \left[\frac{\displaystyle\sum_{x \in X} p^q(x) + \sum_{v \in Y} p^q(y)}{\displaystyle\sum_{x \in X} p(x) + \sum_{y \in Y} p(y)} \right]. \tag{7.21}$$

The summations in Eq. 7.21 can be collapsed and written as follows

$$H_q(X \cup Y) = \frac{1}{1-q} \log_b \left[\frac{\displaystyle\sum_{x \in X, v \in Y} p^q(x \cup y)}{\displaystyle\sum_{x \in X, y \in Y} p(x \cup y)} \right] \tag{7.22}$$

which is the Rényi generalized entropy $H_q(X)$ for the probability distribution $X \cup Y$.

Since the Rényi generalized entropy $H_q(X)$ is effectively an additive function in the form of the exponential function $g_{q,b}(x)$ using Eq. 7.14, then $H_q(X)$ can also be used as an additive information cost function in this new sense of additivity. Therefore, a new form of the additive cost function using the probabilities from Eq. 7.5 can be written using the Rényi generalized entropy $H_q(X)$ as follows

$$C(f, s) = \frac{1}{1-q} \log_b \frac{\displaystyle\sum_u \left(\frac{|\langle f, \psi_{u,s} \rangle|^2}{\|f\|^2} \right)^q}{\displaystyle\sum_u \frac{|\langle f, \psi_{u,s} \rangle|^2}{\|f\|^2}} \tag{7.23}$$

where additivity is in the sense of Eq. 7.14 using the $g_{q,b}(x)$ from Eq. 7.12.

7.4   Summary

This chapter has looked at progressive wavelet image transmission using region enhancement and wavelet packet best basis selection. The region enhancement presented

in Sec. 7.2 investigates how the multifractal measures development in Chapter 6 are affected when there are regions selected in an image. It was found that the enhanced regions improve the local image complexity according to the IQM analysis and slightly degrade the complexity of the rest of the image at low bit rates. Also presented in Sec. 7.2 is a method of mapping the IQM values throughout an image so that identification of RCD can be performed. The results from the experiments conducted for the identification of RCD show that major edge boundaries between different textures are identified, as well as localized objects that stand out from nearby features.

The section on wavelet packets presented a generalized entropy approach of Coifman and Wickerhauser's best basis selection [CoWi92]. It was shown mathematically that the Rényi generalized entropy can be used as an additive information cost function if the definition of additive is expanded to mean value as given by Rényi [Rény60].

The presentation of background, theory, and experimental results for this thesis is now concluded. The next chapter closes the thesis with conclusions drawn from the research and recommendations for future work.

CHAPTER VIII

CONCLUSIONS

8.1   Conclusions

This thesis has presented the development and study of progressive image transmission techniques involving fractals and wavelets, as well as progressive image quality measures based on multifractal complexity measures. One approach developed and presented is a progressive image transmission technique that synthesizes image textures using fractal surfaces. Objective measures were then formulated and studied to measure the complexity of the surfaces within an image. Finally, the image complexity measures were applied in the development of wavelet based progressive image transmission techniques, including region of complexity disparity selection and the search for a wavelet packet best basis.

It was shown in Chapter 5 that the synthesis of textures is one approach to modelling the psychovisual characteristics of the surfaces within an image. This modelling was done by measuring the fractal dimension of the image and then interpolating a fractal surface to approximate the textures in the image. The interpolation process produces textures with similar characteristics as the original surfaces, but does not attempt to reproduce specific pixel values. The model produced perceptually good results, therefore, it is concluded that the fractal dimension is useful for characterizing and synthesizing of textures.

With the development of new and novel techniques for representing images with fractal surface interpolations, observations of the texture synthesis gives credence that fractal dimensions can capture important characteristics of image textures and features. Thus, fractal dimensions are candidates for the basis of an image quality measure. Chapter 6 explored the potential of using fractal dimensions, or more specifically multi-fractal dimensions, in the formation of these objective measures. Considered was the Rényi generalized entropy, the Rényi dimension spectrum, the Mandelbrot spectrum, and the newly developed relative Rényi dimension spectrum in Sec. 6.3 through Sec. 6.6, respectively. The Rényi generalized entropy, presented in Sec. 6.3, has the shortcoming of only considering the measurement at one scale. Therefore, the Rényi generalized entropy does not consider the image from a global viewpoint and its results only reflect local cal-culations. The Rényi dimension spectrum and the Mandelbrot spectrum, presented in Sec. 6.4 and Sec. 6.5, respectively, fix this shortcoming by extending the measurement, in a limiting case, to an infinite resolution. This extension allows the measurement to con-sider the image from a global scale through to the smallest local scale and, hence, gives a better means to an objective measure. The primary problem that exists with these two multifractal measures is that they are sensitive to the approach taken for calculating the limiting case for both the original and reconstructed images. The relative Rényi dimen-sion spectrum, presented in Sec. 6.6, minimizes some of the effects of this problem by per-forming the comparison of the reconstructed image with the original image before the limiting case in the multifractal measure. Therefore, it is concluded that of the image complexity measures presented, the relative Rényi dimension spectrum produces the best results. Particularly, the IQM analysis developed in Sec. 6.6 encapsulates the results of the

relative Rényi dimension spectrum into a single value, giving a clean exponential decay in value as the image quality improves.

From the results obtained in Chapter 6 with image complexity measures, applications to wavelet based progressive image transmission were then considered in Chapter 7. The effects of including enhanced regions in an image decomposition on the developed image complexity measures were first investigated in Sec. 7.2.2. It was found that the image complexity measures can be applied successfully to these enhanced regions. When applied, the measure results show correspondence with the perceived image quality enhancements of the regions. The IQM analysis was then applied to identify regions of complexity disparity in Sec. 7.2.3. From the results obtained, the regions of complexity disparity selected correspond to many of those that the human visual system would consider perceptually important. Finally, the Rényi generalized entropy was considered in the generalization of the entropy based best basis selection for a wavelet packet decomposition. The theoretical framework was developed in Sec. 7.3 that generalizes the notion of an additive information cost function as well as generalizes the entropy based information cost function with the Rényi generalized entropy.

With the development and experimentation with progressive image transmission techniques involving fractals and wavelets, as well as progressive objective measures, it is concluded that the objectives of this thesis have been achieved. The following section lists the contributions of this thesis.

8.2   Contributions

We believe that this thesis and the research done towards its completion has provided the following contributions.

1. A new and novel approach to image compression and progressive image transmission based on fractal surface segmentation and interpolation using fractional Brownian motion generated surfaces. This technique is not bit rate competitive with state of the art techniques such as SPIHT or EZW, but does show a unique approach signal representation.

2. A study of using multifractal dimension measures as objective meausres of image quality for progressive image transmission.

3. The development of a new theoretical class of multifractal measures referred to as relative multifractal measures.

4. The development of a new measure referred to as the relative Rényi dimension spectrum $RD_q(u \parallel v)$.

5. The development of an objective measure, IQM, using the relative Rényi dimension spectrum complexity measure.

6. A study of the effects of enhancing regions in a progressive image transmission on the newly developed relative Rényi dimension spectrum.

7. The development of an approach to identify regions of complexity disparity in an image using the objective measure, IQM.

8. The generalization of the concept of additivity for additive cost functions in wavelet packet best basis searches.

9. The theoretical development of a new best basis additive cost function using the Rényi generalized entropy.

8.3   Recommendations for Future Work

Based on the work done in this thesis, the following future work to extend some of the ideas presented is recommended.

1. The extension of the multifractal measures for colour images.

2. The extension of the progressive image transmission techniques to colour images.

3. The development of more robust techniques, such as line fitting techniques, to reduce the sensitivity in the multifractal measures when calculating the limiting case.

4. A study of the effects on the image quality measures of preprocessing the image to a just noticeable difference level from the viewpoint of the human visual system, perhaps through a smoothing filter.

5. Further research should be done to explain the results of the $RD_q(u \parallel v)$ calculations when q is a large negative or positive number.

6. Experimentation with the newly formed best basis cost function developed using the Rényi generalized entropy and the new generalization of additive cost functions.

7.   Research into using the multifractal measures to select certain classes of fea-

      tures from an image when looking at regions of complexity disparity. This

      may be useful in cases where, say, roads need to be selected from an aerial

      photo.

REFERENCES

[AtSV88]   H. Atmanspacher, H. Scheingraber, and W. Voges, "Global scaling proper-
ies of a chaotic attractor reconstructed from experimental data," *Phys. Rev.
A*, vol. 37, no. 4, pp. 1314-1322, Feb. 1988.

[Barn88]   M. Barnsley, *Fractals Everywhere.* New York: Academic Press, 1988.

[Brow28]   R. Brown, "A brief account of microscopical observations made in the
months of June, July, and August, 1827, on the particles contained in the
pollen of plants; and on the general existence of active molecules in
organic and inorganic bodies," *The London and Edinburgh Philosophical
Magazine and Annals of Philosophy*, vol. 4, no. 21, pp. 161-173, 1828.

[BrBa94]   L. Bruton and N. Bartley, "Simulation of fractal multidimensional images
using multidimensional recursive filters," *IEEE Trans. on Circuits and Sys.*,
vol. 41, no. 3, pp. 181-188, Mar. 1994.

[BuSi97]   R. W. Buccigrossi and E. P. Simoncelli, "Progressive wavelet image coding
based on a conditional probability model," *Proc. of IEEE Intl. Conf. Acous-
tics, Speech, and Signal Processing - ICASSP'97*, vol. 4, pp. 2957-2960,
Apr. 1997.

[BuGe90]   W. Buskist and D. W. Gerbing, *Psychology: Boundaries and Frontiers.*
HarperCollins Publishers, 667 pp., 1990.

[Cant83]   G. Cantor, "Über unendliche, lineare Punktmannigfaltigkeiten V," *Mathe-
matische Annalen*, vol. 21, pp. 545-591, 1883.

[Carp80]   L. Carpenter, "Computer rendering of fractal curves and surfaces," *SIG-GRAPH'80 Conference Proc.*, pp. 109, 1980.

[CPSC97]   S. Cen, H. Persson, D. Schilling, P. Cosman, and C. Berry, "Human observer responses to progressively compressed images", *Proc. of the 31st Asilomar Conf. on Signals, Systems and Computers, Pacific Grove, California*, vol. 1, pp. 657-661, Nov. 1997.

[ChSa95]   B. B. Chaudhuri and N. Sarkar, "Texture segmentation using fractal dimension," *IEEE Trans. on Pat. Anal. and Mach. Int.*, vol. 17, no. 1, pp. 72-77, Jan. 1995.

[Chen97]   H. Chen, Accuracy of Fractal and Multifractal Measures for Signal Analysis, *M. Sc. Thesis, University of Manitoba*, Winnipeg, MB, Canada, 115 pp., 1997.

[CoWi92]   R. R. Coifman and M. V. Wickerhauser, "Entropy-based algorithms for best basis selection," *IEEE Trans. on Information Theory*, vol. 38, no. 2, pp. 713-718, March 1992.

[CoTh91]   T. M. Cover and J. A. Thomas, *Elements of Information Theory*, John Wiley & Sons, 524 pp., 1991.

[CrIv89]   R. Creutzberg and E. Ivanov, "Computing fractal dimension of image segments," *Proc. III Int. Conf. on Computer Analysis and of Images and Patterns*, pp. 110-116, 1989.

[Daly93]   S. Daly, "The visible difference predictor: An algorithm for the assessment of image fidelity," in *Digital Images and Human Vision*, ed. A.B. Watson, MIT Press, Cambridge, MA, pp. 179-206, 1993.

[DaKC95]   R. Dansereau, W. Kinsner, and S. Chow, "A study of codebooks for compression of head-and-shoulder images," *Proc. 12th Int. Conf. on Math. and Computer Modelling and Scientific Computing (Boston)*, vol. 6, 7 pp., July 1995.

[DaKi00]   R. Dansereau and W. Kinsner, "Psychovisual correlations with multifractal measures for wavelet and wavelet packet progressive image transmission," *Proc. of the Canadian Conf. on Elec. and Comp. Eng. (CCECE'2000)*, May 2000.

[DaKi96]   R. Dansereau and W. Kinsner, "Perceptual image compression through fractal surface interpolation," *Proc. of the Canadian Conf. on Elec. and Comp. Eng. (CCECE'96)*, pp. 899-902, 1996.

[DaKi97]   R. Dansereau and W. Kinsner, "Image compression through fractal surface interpolation and wavelet compression," *Proc. of IEEE Western Canada Conference on Communications, Computer, and Power (WESCANEX'97)*, pp. 94-99, 1997.

[DaKi98a]   R. Dansereau and W. Kinsner, "Progressive transmission of images through fractal surface interpolation," *Proc. of the Canadian Conf. on Elec. and Comp. Eng. (CCECE'98)*, pp. 277-280, May 1998.

[DaKi98b]   R. Dansereau and W. Kinsner, "Progressive transmission of images using wavelets: Evaluation using the Rényi generalized entropy," *Proc. of the Canadian Conf. on Elec. and Comp. Eng. (CCECE'98)*, pp. 273-276, May 1998.

[DaKi99a]   R. Dansereau and W. Kinsner, "Rényi generalized entropy analysis of images from a progressive wavelet image transmission," *Proc. of the Canadian Conf. on Elec. and Comp. Eng. (CCECE'99)*, 4 pp., May 1999.

[DaKi99b]   R. Dansereau and W. Kinsner, "Progressive image transmission using wavelet packets guided by metrics," special issue of *Mathematical Modelling and Scientific Computing*, vol. 10, 2000. (ISSN 1067-0688)

[Dans01]   R. Dansereau, "Supplemental Report for Source Code used in Dissertation Research", *Technical Report*: University of Manitoba, 2001.

[Daub90]   I. Daubechies, "Orthonormal bases of compactly supported wavelets," *IEEE Trans. Inform. Theory*, vol. 36, pp. 961-1005, Sept. 1990.

[Daub92]   I. Daubechis, *Ten Lectures on Wavelets*, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, 1992.

[DaKe85]   J. A. Davidson and D. J. Keller, "Measurement of image sharpness through the approaches used to describe fractals," *SMPTE Journal*, vol. 94, no. 8, pp. 802-809, Aug. 1985.

[DjVu99]   DjVu document compression software, *AT&T*, 1999.
            http://dejavu.research.att.com/

[Dono92]   D. L. Donoho, "De-noising via soft-thresholding," *Technical Report*, Dept. of Statistics, Stanford University, 37 pp., 1992. (available through the Internet from ftp://playfair.stanford.edu/pub/donoho)

[Dono95]   D. L. Donoho, "De-noising by soft-thresholding," *IEEE Trans. on Inf. Theory*, vol. 41, pp. 613-627, May 1995.

[FeLK93]   K. Ferens, W. Lehn, and W. Kinsner, "Image compression using learned vector quantization," *Proc. IEEE WESCANEX'93*, (Sk, Canada), IEEE 93CH3317-5, pp. 299-312, 1993.

[FeKi95]   K. Ferens and W. Kinsner, "A multifractal entropy measure for feature extraction of images," *Math. Modelling and Scientific Comp. (Boston)*, vol. 6, 6 pp., July 1995.

[Fere95]   K. Ferens, "Perceptual measures of signal features," *Ph. D. Thesis*, University of Manitoba, Winnipeg, MB, Canada, 1995.

[Flan92]   P. Flandrin, "Wavelet analysis and synthesis of fractional Brownian motion," *IEEE Trans. on Inform. Theory*, vol. 38, no. 2, pp. 910-917, 1992.

[FoFC82]   A. Fournier, D. Fussell, and L. Carpenter, "Computer rendering of stochastic models," *Comm. ACM*, vol. 25, no. 6, pp. 371 - 384, June 1982.

[FrSZ97]   T. Frajka, P. G. Sherwood, and K. Zeger, "Progressive image coding with spatially variable resolution," *Intern. Conf. on Image Proc. ICIP '97 (Santa Barbara, CA)*, vol. 1, pp. 53-56, Oct. 1997.

[GeGr92]   A. Gersho and R. M. Gray, *Vector Quantization and Signal Compression*, Kluwer Academic Press, 1992.

[GIF87a]   "GIF Graphics Interchange Format: A standard defining a mechanism for the storage and transmission of raster-based graphics information," report by *CompuServe Inc.*: Columbus, OH, 15 pp., June 15, 1987.

[GIF90]   "Graphics Interchange Format: Version 89a," report by *CompuServe Inc.*: Columbus, OH, 34 pp., July 1990.

[GoCh99]   J. C. Goswami and A. K. Chan, *Fundamentals of Wavelets: Theory, Algorithms, and Applications*, John Wiley & Sons, 306 pp., 1999.

[GoWo92]   R. C. Gonzalez and R. E. Woods, *Digital Image Processing.* Addison-Wesley Publishing, 716 pp., 1992.

[Gray84]   R. M. Gray, "Vector quantization," *IEEE Acoustics, Speech, and Signal Processing Magazine*, vol. 1, pp. 4-29, Apr. 1984.

[GrMo84]   A. Grossman and J. Morlet, "Decomposition of Hardy functions into square integrable wavelets of constant shape," *SIAM Journal on Mathematical Analysis*, vol. 15, no. 4, pp. 723-736, July 1984.

[HJKP86]   T. C. Halsey, M. H. Jensen, L. P. Kadanoff, I. Procaccia, and B. I. Shraiman, "Fractal measures and their singularities: The characterization of strange sets," *Phys. Rev. A*, vol. 33, no. 2, pp. 1141-1151, Feb. 1986.

[HePr83]   H. G. E. Hentschel and I. Procaccia, "The infinite number of generalized dimensions of fractals and strange attractors," *Physica D: Nonlinear Phenomena*, vol. 8D, no. 1 & 2, pp. 435-444, July 1983.

[Hilb91]   D. Hilbert, "Ueber die stetige abbildung einer linie auf ein flächenstück," *Mathematische Annalen*, vol. 38, pp. 459-460, 1891.

[HöKS97]   I. Höntsch, L. J. Karam, and R. J. Safranek, "A perceptually tuned embed-

ded zerotree image coder," *Intern. Conf. on Image Proc. ICIP '97 (Santa

Barbara, CA)*, vol. 1, pp. 41-44, Oct. 1997.

[HöKa97]   I. Höntsch and L. J. Karam, "APIC: Adaptive perceptual image coding

based on subband decomposition with locally adaptive perceptual weight-

ing," *Intern. Conf. on Image Proc. ICIP '97 (Santa Barbara, CA)*, vol. 1,

pp. 37-40, Oct. 1997.

[Huff52]   D. A. Huffman, "A method for the construction of minimum-redundancy

codes," *Proc. Inst. Electr. Radio Eng.*, vol. 40, pp. 1098-1101, Sept. 1952.

[HuBS65]   H. E. Hurst, R. P. Black, and Y. M. Simaika, *Long-Term Storage: An

Experimental Study*, Constable, London, U.K. 145 pp., 1965.

[Hurs51]   H. E. Hurst, "Long-term storage capacity of reservoirs," *Trans. of the

American Society of Civil Engineers*, vol. 116, pp. 770-799, 1951.

[Infi99]   Lightning Strike wavelet image compression software, *Infinop, Inc.*, 1999.

http://www.infinop.com/

[Iter99]   STiNG iterated function systems image compression software, *Iterated

Systems, Inc.*, 1999. http://www.iterated.com/

[Jacq90]   A. Jacquin, "A novel fractal block-coding technique for digital images,"

Proc. ICASSP, pp. 2225-2228, 1990.

[JaFa96]   H. Jafarkhani and N. Farvardin, "Fast reconstruction of subband decom-

posed signals for progressive transmission," *Proc. of IEEE Data Compres-

sion Conference (DCC '96)*, pp. 230-239, April 1996.

[JaKi97]   E. Jang and W. Kinsner, "Multifractal wavelet compression of finger-prints," *Proc. of WESCANEX'97 (Winnipeg)*, pp. 313-321, May 1997.

[Jang97]   E. Jang, *Compression of Fingerprints based on Wavelet Packet Decomposition and Fractal Singularity Measures*, Thesis: University of Manitoba, 206 pp., 1997.

[JaJS93]   N. Jayant, J. D. Johnston, and R. J. Safranek, "Signal compression based on models of human perception," *Proc. IEEE*, vol. 81, pp. 1385-1422, Oct. 1993.

[JPEG99]   "Joint Photographic Experts Group (JPEG) image compression standard," 1999. http://www.jpeg.org/

[Kapl99]   L.M. Kaplan, "Extended fractal analysis for texture classification and segmentation," *IEEE Trans. on Image Processing*, vol. 8, pp. 1572-1585, Nov. 1999.

[KeCC89]   J. M. Keller, S. Chen, and R. M. Crownover, "Texture description and segmentation through fractal geometry," *Compter Vision, Graphics, and Image Process.*, vol. 45, no. 2, pp. 150-160, Feb. 1989.

[Kins94a]   W. Kinsner, "Fractal dimensions: morphological, entropy, spectral and variance classes," *Technical Report, DEL94-4*, Univ. of Manitoba, 146 pp., May 1994.

[Kins94b]   W. Kinsner, "The Hausdorff-Besicovitch dimension formulation for fractals and multifractals," *Technical Report, DEL94-7*, Univ. of Manitoba, 7 pp., June 1994.

[Kins94c]   W. Kinsner, "Noise, power laws, and the spectral fractal dimension," *Technical Report, DEL94-1*, Univ. of Manitoba, 44 pp., Feb. 1994.

[Kins95a]   W. Kinsner, *Fractal and Chaos Engineering, 24.721 Course Notes*, University of Manitoba, 1995.

[Kins95b]   W. Kinsner, "Compression of digital images, including ORTHO-W1 and ORTHO-P1 techniques. Phase 2: Denoising of lossless ortho image compression," *Technical Report, TRLW95-1*, June 21, 1995, 13 pp.

[Koch04]   H. von Koch, "Sur une courbe continue sans tangente obtenue par une construction géometrique élémentaire," *Arkiv för Matematik, Astronomi och Fysik*, vol. 1, pp. 681-702, 1904.

[Koho89]   T. Kohonen, *Self-organization and associative memory*. Berlin, Germany: Springer-Verlag, 3rd ed., 1989.

[Koho90]   T. Kohonen, "The self-organizing map," *Proc. IEEE* 78(9), pp. 1464-1480, 1990.

[Krey89]   E. Kreyszig, *Introductory Functional Analysis with Applications*, John Wiley & Sons, 688 pp., 1989.

[KTMD99]   H. Krim, D. Tucker, S. Mallat, and D. Donoho, "On denoising and best signal representation," *IEEE Trans. on Information Theory*, vol. 45, pp. 2225-2238, Nov. 1999.

[KuLe51]   S. Kullback and R. A. Leibler, "On information and sufficiency," *Annals of Math. Stats.*, vol. 22, no. 1, pp. 79-86, 1951.

[LaKi95a]   A. Langi and W. Kinsner, "Optimal compression of large aerial ortho images," *J. Mathematical Modelling and Scientific Computing* (special issue), ISSN: 1067-0688, 1995.

[LaKi95b]   A. Langi and W. Kinsner, "Signal denoising for compression of aerial ortho images," *J. Mathematical Modelling and Scientific Computing* (special issue), ISSN: 1067-0688, 1995.

[LaKi95c]   A. Langi and W. Kinsner, "Wavelets compression for aerial ortho images," *Math. Modelling and Scientific Comp.*, vol. 6, 9 pp., July 1995.

[Lang96]   A. Langi, Wavelet and Fractal Processing and Compression of Nonstationary Signals, *Ph. D. Thesis, University of Manitoba*, Winnipeg, MB, Canada, 227 pp., 1996.

[LaSM99]   A. D. Lanterman, J. A. O'Sullivan, and M. I. Miller, "Kullback-Leibler distances for quantifying clutter and models," *Optical Engineering*, vol. 38, no. 2, pp. 2134-2146, Dec. 1999.

[Lena99]   "512 × 512 8-bit grayscale image of Lena", *NASA*, 1999.

            http://zeus.arc.nasa.gov/testimg.html

[LGI95]   "Urban and rural ortho images," *Linnet Geomatics International*. Private communication (W. Kinsner), 1995.

[LuAE95]   J. Lu, V. R. Algazi, and R. R. Estes, "Comparison of wavelet image coders using the picture quality scale (PQS)," *Proc. of SPIE Conf. on Wavelet Applications*, vol. 2491, no. 1, pp. 1119-1130, Apr. 1995.

[Lubi93]   J. Lubin, "The use of psychophysical data and models in the analysis of

                display system performance," *Digital Images and Human Vision*, A.B.

                Watson, ed., MIT Press, Cambridge, MA, pp. 163-178, 1993.

[Mall89]   S. G. Mallat, "A theory for multiresolutional signal decomposition: The

                wavelet transform," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 11,

                pp. 674-693, July 1989.

[Mall99]   S. G. Mallat, *A Wavelet Tour of Signal Processing*, 2nd ed., San Diego, CA:

                Academic Press, 637 pp., 1999.

[MaHw92]   S. G. Mallat and H. L. Hwang, "Singularity detection and processing with

                wavelets," *IEEE Trans. on Inform. Theory*, vol. 38, no. 2, pp. 617-643,

                Mar. 1992.

[MaZh93]   S. G. Mallat and Z. Zhang, "Matching pursuits with time-frequency dictio-

                naries," *IEEE Trans. on Signal Processing*, vol. 41, no. 12, pp. 3397-3415,

                Dec. 1993.

[Mand71]   B. Mandelbrot, "A fast fractional Gaussian noise generator," *Water

                Resources Research*, vol. 7, no. 3, pp. 543-553, June 1971.

[Mand75]   B. Mandelbrot, *Les Objects Fractals: Forme, hasard et dimension*, Flam-

                marion, Paris, 192 pp., 1975.

[Mand82]   B. Mandelbrot, *The Fractal Geometry of Nature*, 1982.

[Mand85]   B. Mandelbrot, "Self-affine fractals and fractal dimension," *Physica

                Scripta*, vol. 32, pp. 257-260, 1985.

[MaVa68]   B. Mandelbrot and J. van Ness, "Fractional Brownian motions, fractional noises and applications," *SIAM Review*, vol. 10, no. 4, pp. 422-437, 1968.

[MaWa69a]   B. Mandelbrot and J. R. Wallis, "Computer experiments with fractional Gaussian noises. Part 1, Averages and variances," *Water Resources Research*, vol. 5, no. 1, pp. 228-241, Feb. 1969.

[MaWa69b]   B. Mandelbrot and J. R. Wallis, "Computer experiments with fractional Gaussian noises. Part 2, Rescaled ranges and spectra," *Water Resources Research*, vol. 5, no. 1, pp. 242-259, Feb. 1969.

[MaWa69c]   B. Mandelbrot and J. R. Wallis, "Computer experiments with fractional Gaussian noises. Part 3, Mathematical appendix," *Water Resources Research*, vol. 5, no. 1, pp. 260-267, Feb. 1969.

[MuBa90]   A. R. Murch and R. H. T. Bates, "Colored noise generation through deterministic chaos," *IEEE Trans. on Circuits and Syst.*, vol. 37, no, 5, pp. 608-613, 1990.

[PaSp97]   T. Painter and A. Spanias, "A review of algorithms for perceptual coding of digital audio signals," *Proc. of 13th Intern. Conf. on Digital Signal Processing*, vol. 1, pp. 179-208, July 1997.

[PeJS92]   H.-O. Peitgen, H. Jurgens, and D. Saupe, *Chaos and Fractals: New Frontiers of Science*. Springer-Verlag, 984 pp., 1992.

[PNHA84]   S. Peleg, J. Naor, R. Hartley, D. Avnir, "Multiple resolution texture analysis and classification," *IEEE Trans. on Pattern Anal. and Machine Intell.*, vol. PAMI-6, no. 4, pp. 518-523, July 1984.

[PeVe95]   R. F. Peltier and J. Lévy Véhel, "Multifractional Brownian motion: defini-
            tion and preliminary results," *Rapport de recherche: Institut National de
            Recherche en Informatique et en Automatique (INRIA)*, 39 pp, Aug. 1995.

[PeMi92]   W. B. Pennebaker and J. L. Mitchell, *JPEG Still Image Data Compression
            Standard.* New York: Van Nostrand Reinhold, 1992.

[Pent84]   A. Pentland, "Fractal-based description of natural scenes," *IEEE Trans.
            Pattern Anal. and Machine Intell.*, no. 6, pp. 661-674, 1984.

[PTVF92]   W. H. Press, S. A. Teukolsky, W. T. Vetterling, B. P. Flannery, Numerical
            Recipes in C: The Art of Scientific Computing, 2nd ed., *Cambridge Uni-
            versity Press*, 994 pp., 1995.

[Rény55]   A. Rényi, "On a new axiomatic theory of probability," *Acta Mathematica
            Hungarica*, vol. 6, pp. 285-335, 1955.

[Rény59]   A. Rényi, "On the dimension and entropy of probability distributions,"
            Acta Mathematica (Budapest, Hungary), vol. 10, pp. 193-226, 1959.

[Rény60]   A. Rényi, "On measures of entropy and information," *Proc. Fourth Berke-
            ley Symp. Math. Statist. Probab.*, University of California Press, Berkeley,
            vol. 1, pp. 547-561, 1960.

[SaJo89]   R. J. Safranek and J. D. Johnson, "A perceptually tuned sub-band image
            coder with image dependent quantization and post-quantization data com-
            pression," *Proc. of IEEE Intl. Conf. Acoustics, Speech and Signal Process-
            ing - ICASSP '89*, pp. 1945-1948, 1989.

[SaPe96]   A. Said and W. A. Pearlman, "A new fast and efficient image codec based on set partitioning in hierarchical trees," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 6, no. 3, pp. 243-250, June 1996.

[SaCh92]   N. Sarkar and B. B. Chaudhuri, "An efficient approach to estimate fractal dimension of textural images," *Pattern Recognition*, vol. 25, pp. 1035-1041, 1992.

[Saup88]   D. Saupe, "Algorithms for random fractals," *The Science of Fractal Images*, ed. H.-O. Peitgen and D. Saupe, Springer-Verlag, pp. 71-136, 1988.

[Schr91]   M. Schroeder, *Fractals, Chaos, Power Laws: Minutes from an Infinite Paradise*, W.H. Freeman and Company: New York, 429 pp., 1991.

[Shan48a] C. E. Shannon, "A mathematical theory of communications," *Bell Syst. Tech. J.*, vol. 27, pp. 379-423, July 1948.

[Shan48b] C. E. Shannon, "A mathematical theory of communications," (continued) *Bell Syst. Tech. J.*, vol. 27, pp. 623-656, October 1948.

[Shap93]   J. M. Shapiro, "Embedded image coding using zerotrees of wavelet coefficients," *IEEE Trans. Sig. Proc.*, vol. 41, pp. 3445-3462, Dec. 1993.

[Shap93b] J. M. Shapiro, "Smart compression using the embedded zerotree wavelet (EZW) algorithms," *Proc. of 27th Asilomar Conf. on Signals, Systems & Computers*, vol. 1, pp. 486-490, Nov. 1993.

[Shil74]   G. E. Shilov, *Elementary Functional Analysis*, Dover Publications, Inc., 334 pp., 1974.

[Sier15]   W. Sierpinski, "Sur une courbe dont tout point est un point de ramifica-
tion," *Compt. Rendus Acad. Sci.: Paris*, vol. 160, pp. 302-305, 1915.

[StMe88]   H. E. Stanley and P. Meakin, "Multifractal phenomena in physics and
chemistry," *Nature*, vol. 335, pp. 405-409, Sept. 1988.

[StLN95]   M. A. Stoksik, R. G. Lane, and D. T. Nguyen, "Practical synthesis of accu-
rate fractal images," *Graphical Models and Image Processing*, vol. 57,
no. 3, pp. 206-219, May 1995.

[Tasw94]   C. Taswell, "Near-best basis selection algorithms with non-additive infor-
mation cost functions," *Proc. Symposium on Time-Frequency Time-Scale
Analysis*, vol. 94TH8007, pp. 13-16, Oct. 1994.

[Tasw95]   C. Taswell, "Image compression by parameterized-model coding of wave-
let packet near-best bases," *Proc. of SPIE Conf. on Wavelet Applications*,
vol. 2491, no. 1, pp. 153-161, Apr. 1995.

[USC99]   "Original scanned colour image of Lena", *Univ. of South California*, Signal
and Image Processing Institute, 1999.
http://sipi.usc.edu/services/database/Database.html

[Voss85]   R. F. Voss, "Random fractal forgeries," In R.A. Earnshaw, editor, *Funda-
mental Algorithms in Computer Graphics*, Springer-Verlag, pp. 805-835,
1985.

[VoCl78]   R. F. Voss and J. Clarke, ""1/f noise" in music: Music from 1/f noise," *J. of
the Acoustical Society of America*, vol. 63, no. 1, pp. 258-263, Jan. 1978.

[Wall91]   G. K. Wallace, "The JPEG still picture compression standard," *Commun.*
             ACM, pp. 31-43, Apr. 1991.

[Wall93]   L. M. Wall, "Reduced search fractal block coding using frequency sensitive
             neural networks," *M. Sc. Thesis, University of Manitoba*, Winnipeg, MB,
             Canada, x+200 pp., 1993.

[Wata97]   T. Watanabe, "Image coding making use of B-spline surfaces," *IEEE Trans.*
             on Circuits and Syst. for Video Tech., vol. 7, no. 2, Apr. 1997.

[Wats87]   A.B. Watson, "The Cortex transform: rapid computation of simulated neu-
             ral images," *Computer Vision, Graphics, and Image Processing*, vol. 39,
             pp. 311-327, 1987.

[WaGo89]   L. Wang and M. Goldberg, "Progressive image transmission using vector
             quantization on images in pyramid form," *IEEE Trans. Commun.*,
             pp. 1339-1349, Dec. 1989.

[Welc84]   T. Welch, "A technique for high performance data compression," *IEEE*
             Computer, vol. 17, no. 6, pp. 8-19, June 1984.

[Wick94]   M. V. Wickerhauser, *Adapted Wavelet Analysis from Theory to Software*,
             Wellesley, MA: A K Peters, 486 pp., 1994.

[Wisn96]   J. S. Wisniewski, "Colors of noise pseudo FAQ," version 1.3, *Newsgroup:*
             alt.sci.physics.acoustics, Oct. 7, 1996
             <http://www.msaxon.com/colors.htm>.

[WiNC87]   I. H. Witten, R. M. Neal and J. G. Cleary, "Arithmetic coding for data com-
             pression," *Commun. ACM*, vol. 30, no. 6, pp. 520-540, June 1987.

[WoJa99]   B. Wohlberg and G. de Jager, "A review of the fractal image coding litera-
            ture," *IEEE Trans. on Image Processing*, vol. 8, pp. 1716-1729, Dec. 1999.

[WoON86]   J. W. Woods and S. D. O'Neil, "Subband coding of images," *IEEE Trans.
            Acoust., Speech, and Signal Proc.*, vol. ASSP-34, pp. 1278-1288,
            Oct. 1986.

[ZiLe77]   J. Ziv and A. Lempel, "A universal algorithm for sequential data compres-
            sion," *IEEE Trans. Inform. Theory*, IT-23, pp. 337-343, May 1977.

[ZiLe78]   J. Ziv and A. Lempel, "Compression of individual sequences by variable
            rate coding," *IEEE Trans. Inform. Theory*, IT-24, pp. 530-536, Sept. 1978.

APPENDIX A

ORIGINAL IMAGES USED IN EXPERIMENTS
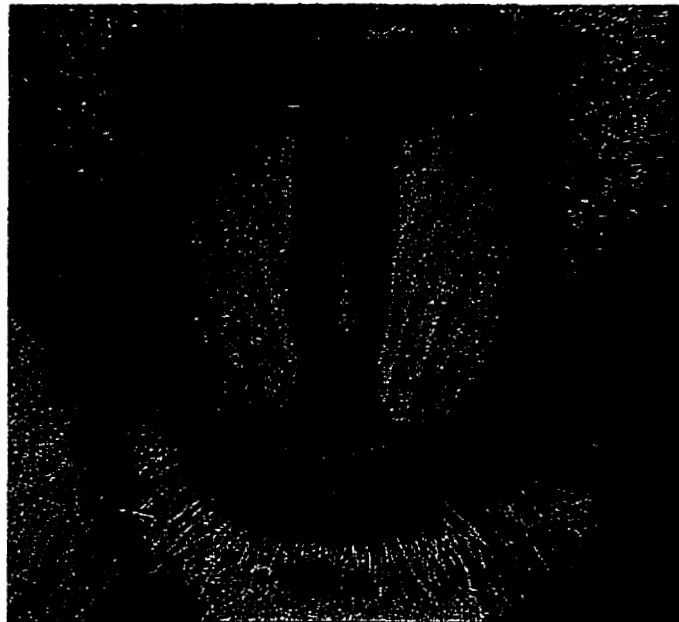

Fig. A.1 Original image of lena.

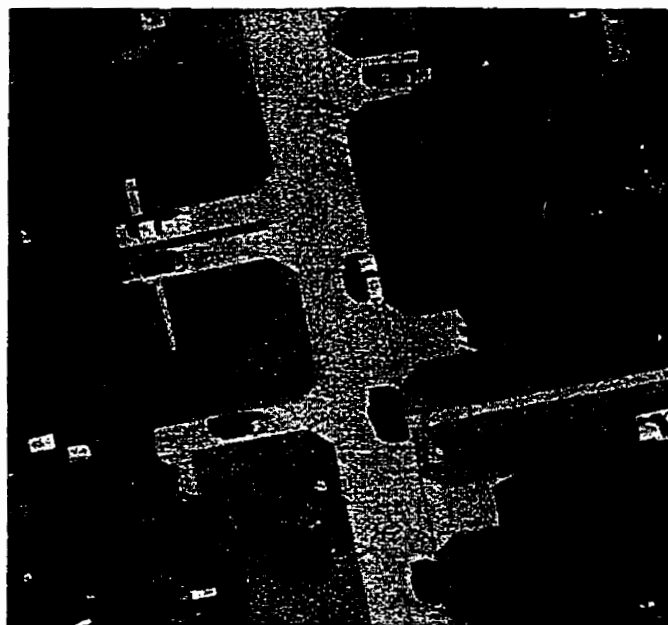Fig. A.2 Original image of baboon.

Fig. A.3 Original image of peppers.

Fig. A.4 Original image of urban.

Fig. A.5 Original image of farm.

APPENDIX B

MEAN OPINION SCORE SUPPLEMENTARY INFORMATION

B.1   Statement of Protocol for Mean Opinion Score Experiment

I would like to thank you for volunteering today to participate in the perceptual image quality project. To standardize this presentation and help remove any bias in this study, I have prepared a statement of the procedure and protocol which I will read now.

The purpose of this study is to measure the perceptual quality of images that have been distorted for purposes of digital image compression. I have asked you here today to help in assessing the perceptual quality of a number of images. The protocol used for this assessment is as follows:

1. An image will be presented for 5 seconds

2. A 2 second pause will then follow

3. A second image will be presented for 5 seconds

4. A 8 second period will be given to fill in the evaluation form with your rating

5. This protocol procedure is continued with another image pair

In addition to this protocol procedure, the following should also be noted

- The first image presented in the image pair is a reference image

- The second image in the image pair is the image to be evaluated

- The second image may or may not be the original reference image

- Image pairs are presented in a random order so do not anticipate a pattern

For the images being evaluated, you are asked to score the level of distortion that you perceive compared to the reference image. Scoring is on a 5 point scale with the following ratings:

a) 5 for a level of distortion that is imperceptible

b) 4 for a level of distortion that is perceptible but not annoying

c) 3 for a level of distortion that is slightly annoying

d) 2 for a level of distortion that is annoying

e) and 1 for a level of distortion that is very annoying

The perceptual score for an image is to be marked on the mean opinion score sheets provided with an "X" for that image pair set. If you feel that the perceived rating falls somewhere between two rating levels, you may mark it along the line proportional to your perceived rating.

This study should last no more than 30 minutes. To remove any biases in the scoring, I would like to remind you that participation in this study is on a voluntary basis and that you will not be remunerated or receive any payment for participating. Also, for purposes of anonymity please do not put your name or any other identifying marks on the mean opinion score sheets.

To start, we will go through two test image pairs so that we are all familiar with the timings and what to expect. For these image pairs, imagine what you would mark on the provided scale if you were to rate the perceptual quality. If there are any questions, please feel free to ask them now or during the experiment.

B.2   Ranking Scale Used for Mean Opinion Score Experiment

The following scale was used in the mean opinion score experiment for people to mark their perceived perceptual ranking of an image pair.

Image Pair #1

1	2	3	4	5
very annoying	annoying	slightly annoying	perceptible, but not annoying	imperceptible