ADAPTIVE WAVELET FILTER DESIGN FOR OPTIMIZED IMAGE SOURCE ENCODING

by

ROOPESH KUMAR, B.E.

A THESIS

IN

ELECTRICAL ENGINEERING

Submitted to the Graduate Faculty
of Texas Tech University in
Partial Fulfillment of
the Requirements for
the Degree of

MASTER OF SCIENCE

IN

ELECTRICAL ENGINEERING

Approved

_____
Chairperson of the Committee

_____

_____

_____

Accepted

_____
Dean of the Graduate School

December, 2002

ACKNOWLEDGEMENTS

I bow to Lord Narayana, who blesses every being, and who is the bestower of all wisdom and dreams.

I would like to take this opportunity to give my heartfelt thanks to my advisor, Dr. Sunanda Mitra, for her guidance, patience and mother-like support. Working with her was a pleasure and a great experience overall. I also express my sincere thanks to all the committee members, Dr. Tanya Karp, Dr. Thomas Krile, and Dr. Brian Nutter, for their help and interest shown in my thesis. Also grateful have been all my friends, whom I love a lot.

My appreciation for the unconditional love, my parents and my wife provide, is beyond words. It is by their support, encouragement and hope, that I regain my energy and vision from. My life is indeed indebted to them.

Lastly, I would like to thank one person, Dr. Ramesh Krishnan, who has changed my life for the best. He is the one who has given me constant motivation, education, and has shown me that there is so much more in life worth accomplishing. I ardently wish that every person has the privilege of such a mentorship in his life.

TABLE OF CONTENTS

ABSTRACT

Despite intensive research being conducted on the topic of adaptive filter design in general, adaptive filter design in the discrete wavelet transform (DWT) domain with specific constraints is still an active research area. The present work investigates the advantages and limitations of the design of a 2-chanel perfect-reconstruction wavelet filter which is adapted and optimized under minimum energy constraints in a specific band. Such a filter can be used with a quantizer and entropy encoder of a wavelet based image encoder to give optimum performance.

An optimal 2-channel conjugate quadrature filter (CQF) bank has been designed and optimized using Sequential Quadratic Programming methods. The filter bank problem is solved using recently developed optimization techniques for general non-linear, non-convex functions. The results indicate an improved performance for this method compared to the earlier-used Interior-Point optimization method.

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1
INTRODUCTION

The importance of visual communications has increased significantly in recent years with applications in several areas such as image/video databases, digital television transmission, and video conferencing. Uncompressed image and video data require considerable storage capacity and high bandwidth networks for their transmission. For example, an uncompressed color image of size 512x512 pixels with 24 bit color would require about 0.8 Mbytes, or a gray level lumbar medical image of size 2000x2500 would require 5Mbytes of space. The problem is very acute in cases where it is required to handle thousands of such images. Image compression in such a case is the only solution to decrease the amount of memory required to store the image and to reduce the bandwidth required to transmit the image. Efficient data compression techniques are highly essential for superior results.

1.1 Discrete Wavelet Transform

In recent years, the discrete wavelet transform (DWT) has become a standard technique in image data compression. DWT has high correlation and energy compaction efficiency. It has certain other distinct advantages, over other popular transforms, like the elimination of blocking artifacts and its suitability to the human visual system characteristics.

A wide variety of wavelet-based schemes are available ranging from simple scalar-based coding to more complex techniques such as vector quantization and tree encoding. A wavelet based image encoder can be classified into three components: decomposition of the image using wavelets, quantization of the wavelet coefficients and entropy encoding of quantized coefficients. Although in every component there is freedom to choose from different options available, they collectively affect the coder performance, and the choice of the appropriate wavelet or the filter bank is very crucial. This is because if the performance of the wavelet filter is poor, the subsequent steps

would be expected to perform poorly too. At the same time, if the most appropriate wavelet is found, then the overall performance of the coder would be optimum. Most of the well known image coding algorithms use a particular filter bank chosen from a variety of filters designed and developed over the years. In such algorithms, the same filter is used for coding and decoding all kinds of input images, whether it is a natural, synthetic or a medical image, which all have very different characteristics. This approach may not always give the best overall performance of the coder. Hence it is of significant interest to find the best or the optimum wavelet filters in any given case. This work does exactly that and it concentrates on the first coder step of designing the wavelet or the filter bank. An effective method is presented to find the optimum filter bank adaptively for the given input of the image source which depends on the statistics of the input. This work is founded on earlier similar work [1] done on adaptive wavelet filter banks.

Also, designing a filter bank for any application requires that a lot of variables be taken into account. The filters can be either finite impulse response (FIR) or infinite impulse response (IIR). They can be either orthogonal or biorthogonal and perfect reconstructing or near perfect reconstructing. Since the application is intended for images, the filters have to be compactly supported in time, so that there is time-localization and hence FIR filters are the obvious choice. The filter bank type that is chosen is a 2-channel conjugate quadrature filter (CQF) bank. A CQF bank is very similar to the more popularly known quadrature mirror filter (QMF) bank. It has been found to have slightly better distortion cancellation and frequency domain characteristics [2].

1.2 Optimization Technique

Since a CQF bank design for a given filter order can have many possible solutions, the search for the best filter is an optimization problem. The problem is formulated by making use of an objective function which is dependent upon the input image statistics and various other criteria that imply perfect reconstruction of the image.

This optimization is performed using the well-known Sequential Quadratic Programming (SQP) algorithm.

SQP is a class of algorithms that solves a given problem by transforming it into a quadratic optimization and solves it using standard methods like the quasi-Newton step finding method and line-search. The results obtained are subsequently presented.
The results are based on a comparison of using such optimum filter banks as opposed to using standard filter banks like the famous Daubechies filters. The results are also compared with the interior point algorithm method used in the earlier work [1].

1.3 Outline of the Thesis

Chapter 2 deals with the wavelet and filter bank theory. A brief note about the importance of the application of wavelets to image source encoding is presented. The concepts of the CQF bank are noted and the filter design constraints using the perfect reconstruction criteria are derived. The complete optimization problem is formulated.

Chapter 3 deals with the optimization concepts and the classification the various optimization techniques, with special emphasis on the Interior-Point and SQP methods. The problem, formulated in chapter 2, is classified in terms of convexity and non-convexity, and then solved using the SQP algorithm.

Chapter 4 presents the application of the filter bank to image source encoding and it includes the results of the tests performed on various images and comparison with the earlier-used Interior-Point method.

CHAPTER 2

ADAPTIVE WAVELET FILTER DESIGN

2.1 Wavelet Basics and Equations

Wavelets are functions defined over a finite interval and have an average value of zero [1]. The basic idea of the wavelet transform is to represent any arbitrary function $f(t)$ as a superposition of a set of such wavelets or basis functions. These basis functions are obtained from a single prototype wavelet called the mother wavelet, by dilations or contractions (scaling) and translations (shifts).

To formalize the discussion about wavelets and filter banks let us introduce some notations that will be used throughout this work. The notations that are most commonly used in the literature are used. By $x \in Z$, $x \in R$ or $x \in C$, we denote that x belongs to the set of integer, real or complex numbers, respectively. By $x \in R^{MxN}$ and $x \in C^{MxN}$ we denote that x belongs to the set of real and complex objects, respectively, with M rows and N columns. The symbol [ ] denotes a discrete version of the signal of the continuous signal denoted by ( ).

The scaling and wavelet functions are given by:

$$\phi(t) = \sum_n h_0(n)\phi(2t-n), \tag{2.1}$$

$$\psi(t) = \sum_n h_1(n)\phi(2t-n), \quad t \in R. \tag{2.2}$$

Functions $\phi(t)$ and $\psi(t)$ are called as the scaling and the wavelet functions respectively. If the scaling and wavelet functions satisfy Eq.(2.1) and Eq.(2.2) they can be used as kernel functions to obtain a wavelet transform of a function $f(t)$. While these equations guarantee a multiresolution representation of the function $f(t)$, these equations are not sufficient to guarantee that the inverse transform exists, to completely recover the original function from the transform coefficients. The minimum requirement for the invertability of a transform in terms of its basis is biorthogonality of the basis functions,

in this case the scaling and wavelet functions [1]. In terms of the coefficients h_0 and h_1, the conditions for complete reconstruction from the inverse wavelet transform are:

$$\sum_n f_0(n+2m)h_0(n) = \delta(m), \qquad m \in Z, \tag{2.3}$$

$$\sum_n f_1(n+2m)h_1(n) = \delta(m), \tag{2.4}$$

$$\sum_n f_1(n+2m)h_0(n) = 0, \text{ and} \tag{2.5}$$

$$\sum_n f_0(n+2m)h_1(n) = 0. \tag{2.6}$$

Coefficients f_0 and f_1 are used in the inverse wavelet transformation. In order for the functions $\phi(t)$ and $\psi(t)$ to form a bi-orthogonal basis, Eqs.(2.3)-(2.6) have to be satisfied.

2.2 Advantages of Wavelets in Image Encoding

The wavelet transform offers a very promising method for image compression. Very high compression rates are achieved by using wavelet transforms. A brief explanation is given regarding the advantages of using wavelets for image source encoding. Wavelets offer certain advantages compared to popular transforms like the Discrete Cosine Transform (DCT) and Discrete Fourier Transform (DFT), which do a good job in localizing the energy in the frequency domain but behave relatively poorly in the space domain since they do not admit non-stationary signals [3]. Wavelets, on the other hand, admit non-stationary signals offering good localization in both the frequency and space domains. Such localization in both frequency and space is very important, especially in applications like image encoding since one must be able to analyze the image at different scales (which is spatial frequency resolution) and to be able to encode the positional locations of the image (which is spatial resolution).

Like the other transforms available, the wavelet transform reduces the entropy of the image, which means that the wavelet coefficient matrix has lower entropy than the image it is encoding. Apart from good frequency and space localization, there are also fast algorithms available for implementing wavelet transforms. They have high de-correlation and energy compaction properties. Effects like blocking artifacts and aliasing distortions can be eliminated by designing proper filters. An example of such a blocking artifact, called as checker-board artifact, which is caused due to non-smooth scaling function or the low-pass filter, is shown in Fig. 2.1. It is better suited to the human visual system since its basis functions are localized both in frequency and space. These properties make wavelets very attractive for image compression. A more detailed analysis of the wavelet subband coding and its advantages is given in a later chapter.

Fig. 2.1: An image with an Artifact

2.3 Wavelets and Filter Banks

Now let us, briefly, try to understand the relationship between wavelet theory and filter banks as applied to multiresolution analysis. Multirate filter banks are the fundamental building blocks for subband decomposition, which can be used for signal compression by splitting the input signal into multiresolution frequency bands and then coding or compressing each of the bands. Since a two-channel filter bank is used, we will

discuss such a filter bank. Fig. 2.2 shows a two-channel filter bank. Here $x[n]$ is the

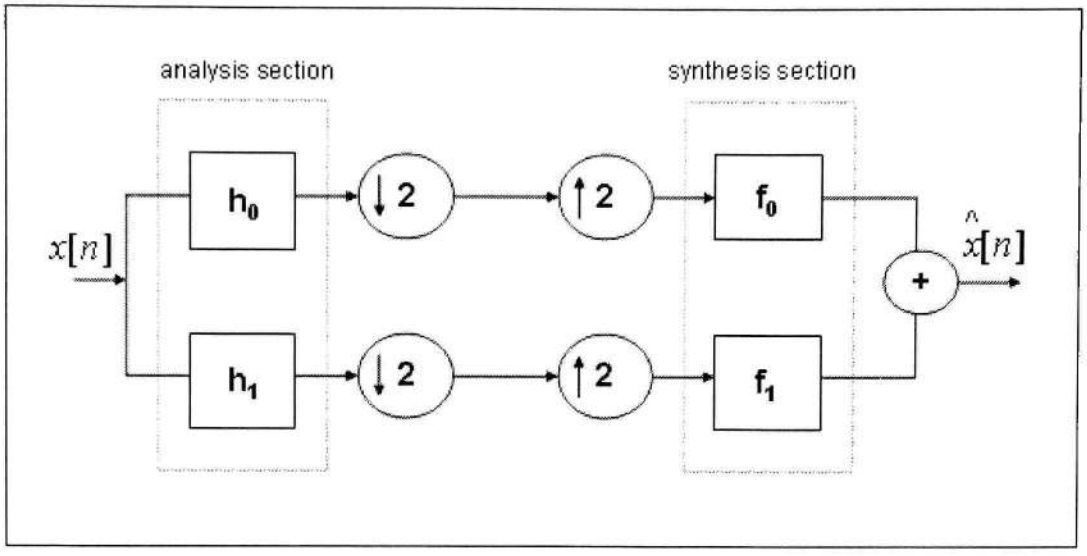


Fig. 2.2: A 2-channel Filter bank

input signal and $h_o, h_1, f_0, \mathrm{f}_1$ are the analysis/decomposition and synthesis/reconstruction low-pass and high-pass FIR filters, respectively, of length N. It can be shown that the perfect reconstruction conditions for the output signal, $\hat{x}[n]$, in such a case are

$$\sum_n h_0[n+2m]f_0[n] = \delta[m], \tag{2.7}$$

$$\sum_n h_1[n+2m]f_1[n] = \delta[m], \tag{2.8}$$

$$\sum_n h_1[n+2m]f_0[n] = 0, \tag{2.9}$$

$$\sum_n h_0[n+2m]f_1[n] = 0. \tag{2.10}$$

Eqs.(2.3)-(2.6) and (2.7)-(2.10) are identical and thus the wavelet definitions given by Eqs.(2.3)-(2.7) also form a perfect reconstruction filter bank. Thus Fig. 2.2

represents an implementation of the forward and inverse discrete wavelet transform.

This work considers perfect reconstruction Conjugate Quadrature Filter (CQF) banks since they have been the popular choice recently for sub-band coding of speech and image signals for efficient wavelet compression techniques. A CQF is an orthogonal filter bank, similar to a Quadrature Mirror Filter (QMF) bank, whose low-pass and high-pass filters form a time reversed pair. The relationships among the filters for a CQF bank are given by the following simple time domain relationships [2]:

$$h_1[n] = (-1)^n h_0[N-1-n], \tag{2.11}$$

$$f_1[n] = -(-1)^n h_0[n], \tag{2.12}$$

$$f_0[n] = h_0[N-1-n]. \tag{2.13}$$

Thus the filter bank design reduces to designing either the analysis high-pass or low-pass filter alone. Let us introduce the concept of the Nyquist(M) condition. A sequence $c(n)$ is said to be Nyquist(M) if $c(nM) = \delta(n)$. The sequence $h_1(n)$ is said to satisfy the Nyquist(2) criteria if

$$\sum_n h_1(n)h_1(n+2m) = \delta(m). \tag{2.14}$$

It can be shown that in a two-channel filter bank, if one of the filters $h_1(n)$ or $h_0(n)$ satisfies the Nyquist(2) condition and the other filter is found using Eq.(2.11), then the other filter also satisfies the Nyquist(2) and they form an orthogonal, and therefore a perfect reconstruction (PR) filter bank. Thus the CQF bank design reduces to just solving the Nyquist(2) criteria for $h_1(n)$ or $h_0(n)$ as in Eq.(2.14).

2.4 Adaptive Wavelet Filter Design

Wavelet filter banks have the ability to perform multiresolution analysis, wherein signals at different frequencies are analyzed at different resolutions. They are designed to

give good time/space resolution and poor frequency resolution at high frequencies and good frequency resolution and poor time/space resolution at lower frequencies [4]. This property is very useful since most of the signals (including images) encountered in practical applications have high frequency components over small areas and low frequency components over larger areas. In images, the low frequency spectrum represents the smooth and the coarse details which constitute most of the image, whereas the high frequency spectrum represents the edges and the finer details which constitute a minor portion of most images. The filter bank in Fig. 2.2 splits the input into high frequency and low frequency components, depending on the design of the high-pass and low-pass filters. The low-pass output very closely resembles the input and it gives the essence whereas the high-pass output gives the detail of the input as shown in the Fig. 2.3. The low frequency and the high frequency wavelet coefficients are equal in number and approximately half the size of the input signal. Thus compression can be achieved by reconstructing the input signal by just using the approximate coefficients and discarding the detail information, which constitutes only a minor part of the original input. This is the basic idea in the compression of signals and images using wavelet filter banks.

Since Eq.(2.14) has an infinite number of solutions, there are as many ways by which orthogonal filters can be designed depending upon the application on hand. The knowledge that most of the signals are low-pass processes, that is most of the energy of the signals is concentrated in the low frequencies, is the central aspect used in designing the desired CQF filter bank. The low-pass output of the CQF bank provides a coarser resolution approximation of the input signal and the high-pass output is the finer resolution approximation. Due to the nature of the CQF bank, these two outputs are orthogonal to each other and that filter bank which provides the maximum energy compaction in the low-pass channel, and thus the minimum energy in the high-pass channel, results in the minimum approximation error. Since the input can be any random 2-dimensional image signal, the output energy is considered as the qualifying criteria instead of the output signal.

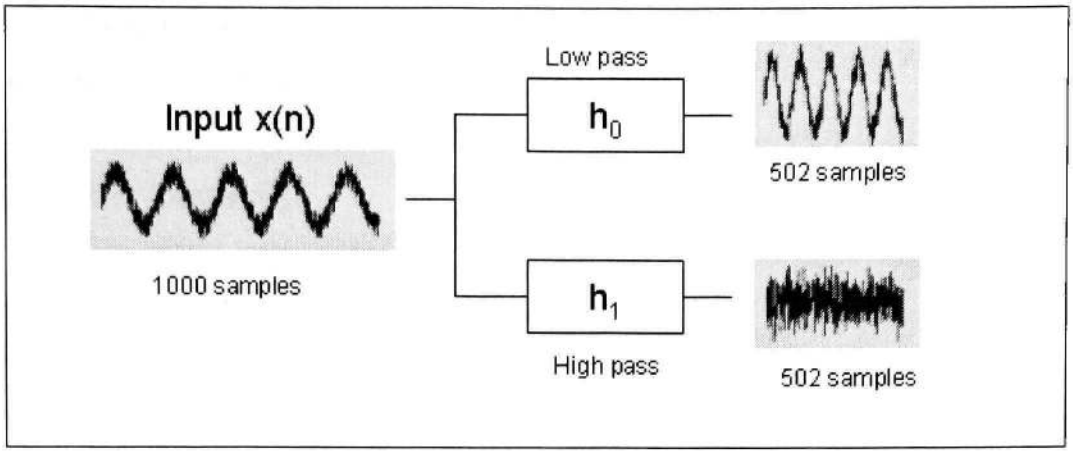


Fig. 2.3: Analysis filter section low-pass and high-pass outputs for the sine wave input with high frequency noise added

This filter bank selection involves the statistics of the input signal, hence making the design adaptive, and the filter that produces the minimum error for that input is the optimum filter bank for the same input. The detailed analysis of this concept is explained in the later chapters. This is the principle for the proposed method for image size reduction using an optimum wavelet filter design and execution time reduction, when implemented in a Digital Signal Processor [1].

Much of the earlier research has been carried using PR-QMF banks in the frequency domain. This work is a continuation of the previous work, using an alternative approach in which the filter is analyzed in the spatial domain [1], wherein the CQF bank is instead used. The space-domain formulation describes the complete set of exact reconstruction conditions based on a simple set of matrix operations and avoids any notion of frequency selective filtering operations. Such a design methodology provides an environment in which many design optimization parameters like energy compaction, orthonormal PR requirement, aliasing distortion and input statistics can be simultaneously optimized using a single design procedure. The design problem is solved using a non-linear optimization technique of Sequential Quadratic Programming and Newton's method, which are explained in the next chapter.

Now we introduce the parameters that have been considered in the optimization of the 2-channel CQF bank design.

2.5 Energy Compaction

The energy term of the high-pass filter forms the objective function of the optimization problem and is the main criterion for the filter bank design. As seen earlier, the wavelet distributes the energy of the input signal into low and high frequency components and the input is then approximated using the low frequency components. Also, as will be seen later in detail, it is possible to redistribute the energy in the low-pass and the high-pass channel, so that one channel receives more energy than the other. As said earlier, most of the signals and images encountered can be considered to be low-pass in nature, and the optimum filter bank for the given input is one that allows maximum energy through the low-pass channel and subsequently minimum energy through the high-pass channel, since the channels are orthogonal. This ensures minimum distortion as the low-pass output approximates the input signal very closely.

The design is a dual optimization problem, that is, either the low-pass channel energy can be maximized or high-pass channel energy can be minimized. Perfect reconstruction without aliasing takes place only when the output energy of the high-pass filter is zero. When this energy is not zero, there is some loss of information and consequently there is some aliasing. The high-pass energy is zero in the ideal case and hence in practice it is a lossy type of compression and the optimization tries to minimize the error.

From stochastic process theory, it can be shown that the total high-pass channel output energy, $\sigma_t^{\,2}$, is given by [1] as,

$$\sigma_t^{\,2} = \frac{1}{2\pi} \int_{-\pi}^{\pi} |H_1(\omega)|^2 S_{xx}(\omega) |F_1(\omega)|^2 \, d\omega . \tag{2.15}$$

In the equation above, $S_{xx}(\omega)$ is the power spectral density of the input signal. This formulation is in the frequency domain. Since the design is in the space-domain, the equivalent equations are written as shown next.

$$\sigma_t^2 = h_1^T R_{xx} h_1, \qquad h_1 \in R^N. \tag{2.16}$$

R_{xx} is the NxN symmetric Toeplitz input autocorrelation matrix. $S_{xx}(\omega)$ is the Fourier equivalent of the autocorrelation matrix R_{xx}. Eq.(2.16) is the space-domain equivalent of the energy in the amplitude spectrum of the high-pass filter. This autocorrelation matrix is the input statistic that the optimization problem uses in selecting the optimum filter bank adaptively. The energy term is the objective function and the solution, h_1, is the one that minimizes this function along with satisfying the constraints for perfect reconstruction. Other desired properties are also analyzed.

2.6 Orthonormal PR Requirement

This requirement is included to obtain the perfect reconstruction condition. This condition is particularly critical in signal coding applications [5]. The PR condition for an orthonormal 2-channel PR-CQF bank is nothing but the implementation of the Nyquist(2) criterion

$$\sum_n h_1(n)h_1(n+2m) = \delta(m). \qquad \begin{matrix} n = 0,1...N-1 \\ m = 0,1.... \end{matrix} \tag{2.17}$$

Eq.(2.17) can be expressed in matrix form as

$$h_1^T T_1 h_1 = 1, \tag{2.18}$$

$$h_1^T T_i h_1 = 0, \qquad i = 2...\frac{N}{2}. \tag{2.19}$$

T_1 is an identity matrix of order N. T_i, $i = 2...\frac{N}{2}$, are a set of symmetric toeplitz matrices given by the first row, a_i, whose coefficients, $t_i(n)$, are

$$\begin{aligned} t_i(n) &= 0.5 \quad \text{if } n = 2i-1 \\ &= 0 \qquad \text{otherwise.} \end{aligned} \tag{2.20}$$

2.7 Zero DC High-Pass Filter

Most of the signals have a significant amount of their energy around the DC component. In order to represent this energy in the low-pass function for minimum error, the mean of the high-pass function is constrained to be zero. This restricts the function to have a zero DC value. This is an additional constraint and is important since the absence of this can give rise to artifacts such as the checker board that was discussed earlier.

$$\sum_n h_1(n) = 0. \tag{2.21}$$

The complete optimization problem can be stated as shown

$$\text{Minimize} \quad h_1^T R_{xx} h_1 \tag{2.22}$$

subject to

$$h_1^T h_1 = 1,$$

$$h_1^T T_i h_1 = 0, \qquad i = 2 \ldots \frac{N}{2}$$

$$\sum_n h_1(n) = 0.$$

T_i are as defined as in Eq.(2.20). This is a non-convex optimization problem, which is explained in detail in the next chapter. Previous work done on this problem has converted this non-convex problem into an equivalent convex one and has made use of primal-dual Interior Point optimization techniques for finding the optimum filter vector. The present work deals with the problem in its original form, which is non-linear and non-convex, using Sequential Quadratic Programming frame-work along with the Newton's method that was used in Interior-Point method. Optimization techniques are explained in detail in the next chapter.

Once the high-pass filter has been designed, the complete filter bank can be easily obtained using the Eqs.(2.11)-(2.13).

CHAPTER 3
NON-LINEAR NON-CONVEX OPTIMIZATION

3.1 Optimization Basics

Optimization problems consist of three basic components.

1. An Objective Function which we want to minimize or maximize. For instance, in a typical wavelet filter design problem, minimization of the reconstruction error is desired. Often, it is required to optimize a number of different objectives at once. For instance, in the above filter design problem, it would also be required to minimize stop band energy, minimize the pass band ripple and maximize pass band energy simultaneously. Usually, the different objectives are not mutually compatible and the variables that optimize one objective may be far from optimal for the others. In practice, such problems are formulated as single-objective problems by either forming a weighted combination of the different objectives or else replacing some of the objectives by constraints.

2. A set of Unknowns or Variables. The unknowns affect the value of the objective function. In the filter design problem above, the unknowns are the impulse response or the frequency response coefficients.

3. A set of Constraints which allow the unknowns to take on certain values but exclude others by defining boundaries in the feasible region. In the filter design, we would want to impose perfect reconstruction constraints like the Nyquist(2) criterion.

An Optimization problem is then to: Find the value of the variables that minimize or maximize the objective function while satisfying the constraints (if present). To be able to classify the optimization techniques, familiarity with some concept of functions is desirable. A set, S, is said to be affine [1] if the line through any two distinct points in S lies in S, that is, for any $x_1, x_2 \in S$ and any $\theta \in R$, we have

$$\theta x_1 + (1-\theta)x_2 \in S .$$

which implies that S contains every linear combination of any two points in S, provided that the coefficients of the combination sum to one. It can be shown that an affine set contains every affine combination of its points.

A set, C, is said to be convex [6] if the line segment between any two points in C lies in C, that is, for any $x_1, x_2 \in C$ and any θ with $0 \le \theta \le 1$, we have

$$\theta x_1 + (1-\theta)x_2 \in C. \tag{3.1}$$

It can be observed that every affine set is also convex since the convex set is a subset of the affine space. Fig. 3.1 shows some examples of such sets.

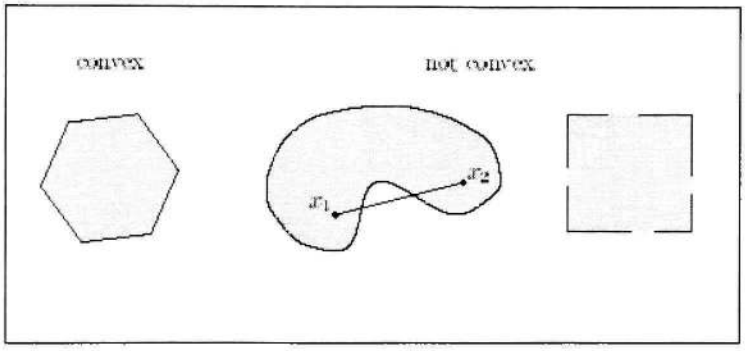


Fig. 3.1: Convex and non-convex set examples

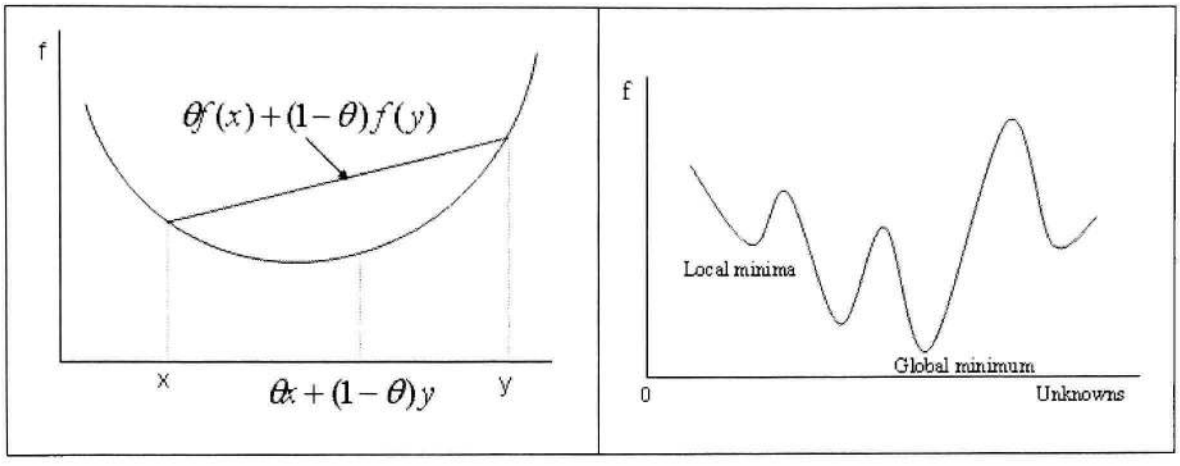
It is seen in Fig. 3.1 that points between the line segment (x_1, x_2) do not lie in the set, and the set is hence non-convex.

A function, $f(x)$, is said to be convex, if the domain (dom) f is a convex set and if for all $x, y \in f$, and θ with $0 \le \theta \le 1$, we have

$$f(\theta x + (1-\theta)y) \le \theta f(x) + (1-\theta)f(y). \tag{3.2}$$

Geometrically it means that the line segment between $(x, f(x))$ and $(y, f(y))$ lies above the graph of f, as shown in Fig. 3.2(a).

If the function $-f$ is convex, then f is said to be concave. Functions that do not satisfy Eq.(3.2) are said to be non-convex. Fig. 3.2(b) gives an example of such a non-convex function. It can be observed that the convex function has a single minimum, called the global minimum, whereas the non-convex function has many local minima and a single global minimum.



(a) A convex function (b) A non-convex function

Fig. 3.2: Example of a convex and a non-convex function

3.2 Classification of Optimization Techniques

The various classifications of optimization problems are as follows:

1. Unconstrained or Constrained:

Unconstrained problems have only an objective function and no constraints. These problems are harder to solve since all possible values have to be tested. Constrained problems are the more common type of optimization problems encountered, wherein the feasible region of the objective solutions is restricted by the accompanying constraints. Such problems have a well defined solution boundary and are a little easier to solve compared to unconstrained problems.

2. Linear or Non-linear:

Linear optimization is used to solve a linear objective subject to linear constraints.

A Non-linear optimization problem is one which has a non-linear objective function or constraints or both. These problems are a little harder to solve compared to linear problems. The wavelet filter design problem that has been formulated in the previous chapter is non-linear.

3. Convex or Non-convex:

A problem is a convex optimization problem if

a) The objective is a convex function;

b) The inequality constraints are convex functions; and

c) The equality constraints are affine functions.

If any of these conditions fail, then the problem fails to qualify as convex and the problem is a non-convex optimization. In a convex problem any solution is the global solution, whereas the same is not true in a non-convex problem. Non-convex problems are harder to solve and most optimization techniques do not guarantee a global or optimum solution.

3.3 Introduction to Interior-Point Methods

An introduction to the popular method of Interior-Point optimization is given here. The earlier work [1] used this method, and some introduction to it would prove to be helpful. As opposed to the simplex method [7] which generates a sequence of feasible iterates by repeatedly moving from one vertex of the feasible set to an adjacent vertex with a lower value of the objective function, Interior-Point methods start with a point that lies inside the set of feasible solutions and the iterate path strictly remains in the feasible region all the time. The most significant breakthrough in the Linear Programming field was made by Karmakar's algorithm [8] in 1984, giving rise to the Interior-Point methods.

Using the standard form notation for the linear problem [7], we define the feasible set

$$P = \{x \in R^n \mid A \cdot x = b; x \geq 0\}$$

and the associated set P^+ to be the subset of P satisfying strict non-negativity constraints

$$P^+ = \{x \in R^n \mid A \cdot x = b; x > 0\}$$

P^+ is called the strictly feasible set and its elements are called strictly feasible points.

Interior-Point methods are iterative methods that compute a sequence of iterates belonging to P^+ and converging to an optimal solution. That is completely different from the simplex method, where an exact solution is obtained after a finite number of steps. Interior-Point iterates tend to an optimal solution but never really attains it, since the optimal solutions do not belong to P^+ but to the region P minus P^+. This apparent drawback is not really serious since

Most of the time an approximate solution (with e.g., 10^{-8} relative accuracy) is sufficient for most purposes

-A rounding procedure can convert a nearly optimal interior point into an exact optimal vertex solution.

The type of algorithm used in the optimization is the main difference between the various interior point methods. Although not fully standardized, they are commonly distinguished as Path-Following algorithms, Affine-Scaling algorithms, and Potential-Reduction algorithms [9]. Of these methods, Path-following interior methods are the most popular and offer excellent results both in theory and practice. This method encompasses the primal-dual method and is explained in detail in [9,10].

The affine-scaling variant of Karmarkar's algorithm is explained below. This algorithm gives a good look at the concept of interior point methods. Details such as the initial trial solution are skipped as they do not affect the algorithm.

The basic idea can be summarized as follows:

Concept 1: Shoot through the interior of the feasible region toward an optimal solution

Concept 2: Move in a direction that improves the objective function value at the fastest rate.

Concept 3: Transform the feasible region to place the current trial solution near its center, thereby enabling a large improvement when concept 2 is implemented.

The affine algorithm is explained [11] using a standard problem defined by the set of Eqs.(3.3) and later supplemented with an example.

The optimization problem is: to find a vector x as to,

$$\text{Minimize } c^T x \tag{3.3}$$

subject to

$$Ax = b,$$

$$x \geq 0,$$

where $A = R^{(mxn)}$ is a real $(m \quad x \quad n)$ matrix and $b \in R^m$.

In the affine-scaling algorithm, the current iterate x^k is used to determine the next iterate x^{k+1} as follows: The variable x is changed as,

$$\hat{x} = D^{-1}x,$$

where

$$D = D^k = diagonal(x^k)$$

is the diagonal matrix whose (i, j) entry is x_i^k , the i^{th} component of the current iterate x^k. The constraints in Eqs.(3.3) become

$$\hat{A}\ \hat{x} = b \text{ and } \hat{x} \geq 0$$

where

$$\hat{A} = AD,$$

and the objective becomes

$$\hat{c}^T \hat{x},$$

where

$$\hat{c} = Dc.$$

The current iterate $\hat{x}$ transforms into a vector of ones (denoted as e),

$$\hat{x}^k := D^{-1}x^k = e.$$

The steepest descent direction is given by

$$\hat{p} = -P_{\hat{A}} \hat{c},$$

where $P_{\hat{A}}$ denotes projection onto the orthogonal complement of the row space of the matrix $\hat{A}$. A step in this direction is given as

$$\hat{x}^{k+1} = \hat{x}^{k} + \hat{a}\,\hat{p},$$

which translates into the original variables as,

$$x^{k+1} = x^{k} + a\,p = x^{k} - \alpha D P_{AD} D c,$$

where

$\alpha = \hat{\alpha} > 0$ is small enough so that $x^{k+1} > 0$, so that all the components of x remain positive, which is a property of the algorithm.

α is usually chosen in the range $\left[0.9\alpha^{\max}, 0.99\alpha^{\max}\right]$, where

$$\alpha^{\max} = \max\{\tau > 0 : x^{k} + p\tau \geq o\}.$$

This is the complete algorithm. The following example [12] is solved using the above algorithm. Consider

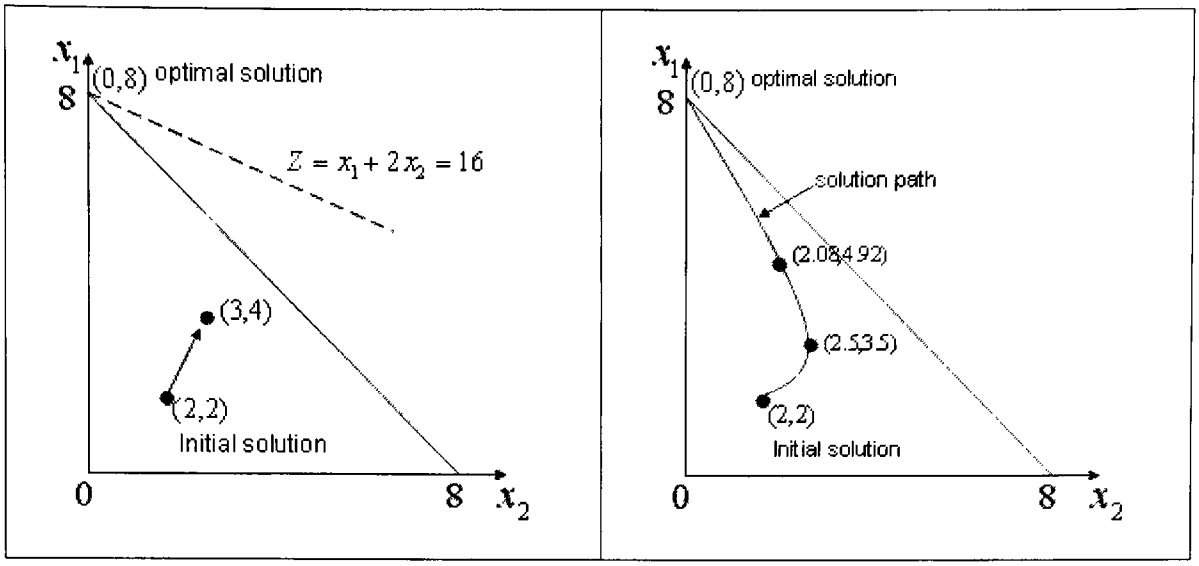
Maximize $Z = x_1 + 2x_2$

subject to

$$x_1 + x_2 \leq 8,$$
$$x_1, x_2 \geq 0.$$

The solution to this simple problem can be analytically found as Z = 16 at (0,8). The algorithm is now applied to this problem with this knowledge, with an initial trial solution chosen as (2, 2) which strictly lies in the feasible region, as shown in Fig. 3.3 (a), which also shows the initial descent direction towards the point (3, 4). Fig. 3.3 (b) shows the final solution trajectory that the algorithm takes for the example. It can be observed that the iterates lie strictly in the feasible region during all the time.

The primal-dual varies in the sense that the objective is transformed into its dual using the Lagrangian [13] of the primal problem and then the dual system is solved simultaneously. This approach promises faster convergence and is the most popular



(a) Initial descent direction (b) Trajectory of the algorithm

Fig. 3.3: Direction and path of the Interior-Point algorithm

among the Interior-Point methods.

3.4 Analysis of the Optimization Problem

The problem defined by the set of Eqs.(2.22) is analyzed and classified in terms of convex or non-convex problem. With the optimization theory in place, the original optimization problem is again shown below and is analyzed in detail.

Minimize $f = h_1^T R_{xx} h_1$ (3.4)

subject to

$$h_1^T h_1 = 0, \tag{3.5}$$

$$h_1^T T_i h_1 = 0, \qquad i = 2 \ldots \frac{N}{2} \tag{3.6}$$

$$\sum_n h_1(n) = 0. \tag{3.7}$$

This problem is shown below to be non-convex. The proof is given by showing that the problem fails to be convex as defined above.

The second order condition definition for a convex function is as stated below: If the function f is twice differentiable, that is, its Hessian $\nabla^2 f$ exists at each point in the domain f, then f is strictly convex if and only if dom f is convex and its Hessian is positive definite [6],

$$\nabla^2 f \geq 0. \tag{3.8}$$

In this case, we observe that the objective is twice differentiable and hence its Hessian exists. Since

$$\nabla^2 f = R_{xx} \tag{3.9}$$

for all values of h, the problem is convex if and only if R_{xx} is a positive definite matrix, that is,

$$R_{xx} \geq 0 \tag{3.10}$$

A matrix, R_{xx}, is defined to be positive definite, if for any non-negative vector x, we have,

$$x^T R_{xx} x \geq 0 \tag{3.11}$$

Eq.(3.11) is equivalent to the objective function in Eq.(3.4).

The objective in Eq.(3.4) is clearly a non-linear function due to the presence of the h^2 term in the form of $h^T h$. By the property of an autocorrelation matrix, the smallest eigenvalue is positive and thus the matrix R_{xx} is positive definite [1]. The objective function is convex by definition in Eq.(3.10). Though the objective is convex, the problem on the whole is non-convex. The non-convexity comes in due to the constraints and is proved below.

The feasible region for the problem is defined by the constraints in Eqs.(3.5)-(3.7). These constraints give the possible solution range for $h[n]$, and it has to satisfy these constraints. By including all the constraints in the objective, the complete problem

is considered. Now the proof to show that convexity fails is simply to show that the plot of the objective with the constraints included is non-convex having more than a single minimum.

The proof is given for a general filter of length $N = 6$. The input is assumed to be a 2-D signal, most likely an image of dimension m x n. Let the optimized filter solution be,

$$h[n] = [a, b, c, d, e, f] \tag{3.12}$$

Let the autocorrelation matrix, R_{xx}, be defined as

$$R_{xx} = \begin{bmatrix} r_{11} & r_{12} & r_{13} & r_{14} & r_{15} & r_{16} \\ r_{21} & r_{22} & r_{23} & r_{24} & r_{25} & r_{26} \\ r_{31} & r_{32} & r_{33} & r_{34} & r_{35} & r_{36} \\ r_{41} & r_{42} & r_{43} & r_{44} & r_{45} & r_{46} \\ r_{51} & r_{52} & r_{53} & r_{54} & r_{55} & r_{56} \\ r_{61} & r_{62} & r_{63} & r_{64} & r_{65} & r_{66} \end{bmatrix}. \tag{3.13}$$

As will be seen later, this matrix is a normalized symmetric Toeplitz (NxN) matrix in image processing applications. This autocorrelation matrix is derived using a 1-D vector, which is calculated either in the horizontal or the vertical directions of the image. The autocorrelation vector in the horizontal direction is defined as

$$R[i] = \sum_{p=1}^{m} \sum_{k=1}^{n-i+1} I(p,k) * I(p,k+i-1), \tag{3.14}$$

where $I(p,k)$ is a two-dimensional input, which in this case is a digital image.

The vector $R[i]$ is normalized as

$$R[i] = R[i] / \max(R). \tag{3.15}$$

The autocorrelation matrix then simplifies as shown below.

$$R_{xx} = \begin{bmatrix} 1 & r_{12} & r_{13} & r_{14} & r_{15} & r_{15} \\ r_{12} & 1 & r_{12} & r_{13} & r_{14} & r_{15} \\ r_{13} & r_{12} & 1 & r_{12} & r_{13} & r_{14} \\ r_{14} & r_{13} & r_{12} & 1 & r_{12} & r_{13} \\ r_{15} & r_{14} & r_{13} & r_{12} & 1 & r_{12} \\ r_{15} & r_{15} & r_{14} & r_{13} & r_{12} & 1 \end{bmatrix}. \tag{3.16}$$

Also due to the nature of the image coefficient values (0-255 for an 8-bit image), the coefficients have the property,

$$r_{16} < r_{15} < r_{14} < r_{13} < r_{12}. \tag{3.17}$$

The normalized vector coefficients also have the additional property, that

$$r_{12}, r_{13} .... r_{16} \in [0,1].$$

A very typical example of such an autocorrelation matrix of order $N = 6$, as calculated for the 'Lena Image' is as shown

$$R_{xx} = \begin{bmatrix} 1.0000 & 0.9965 & 0.9907 & 0.9852 & 0.9801 & 0.9754 \\ 0.9965 & 1.0000 & 0.9965 & 0.9907 & 0.9852 & 0.9801 \\ 0.9907 & 0.9965 & 1.0000 & 0.9965 & 0.9907 & 0.9852 \\ 0.9852 & 0.9907 & 0.9965 & 1.0000 & 0.9965 & 0.9907 \\ 0.9801 & 0.9852 & 0.9907 & 0.9965 & 1.0000 & 0.9965 \\ 0.9754 & 0.9801 & 0.9852 & 0.9907 & 0.9965 & 1.0000 \end{bmatrix}.$$

In other similar work [5], an autoregressive, order(1), AR(1) source model with the correlation coefficient , $\rho = 0.95$,which is a crude approximation to still images is used. The correlation sequence of this source is expressed as,

$$R[m] = \rho^{|m|} \qquad m = 0, \pm 1, \pm 2, ...$$

The corresponding toeplitz autocorrelation matrix for the AR(1) approximation is,

$$R_{xx} = \begin{bmatrix} 1.0000 & 0.9500 & 0.9025 & 0.8574 & 0.8145 & 0.7738 \\ 0.9500 & 1.0000 & 0.9500 & 0.9025 & 0.8574 & 0.8145 \\ 0.9025 & 0.9500 & 1.0000 & 0.9500 & 0.9025 & 0.8574 \\ 0.8574 & 0.9025 & 0.9500 & 1.0000 & 0.9500 & 0.9025 \\ 0.8145 & 0.8574 & 0.9025 & 0.9500 & 1.0000 & 0.9500 \\ 0.7738 & 0.8145 & 0.8574 & 0.9025 & 0.9500 & 1.0000 \end{bmatrix}. \tag{3.18}$$

Now a brief explanation of the various constraints is given in terms of the feasible solution that it defines. From the constraint in Eq.(3.5) we have,

$$a^2 + b^2 + c^2 + d^2 + e^2 + f^2 = 1. \tag{3.19}$$

The Toeplitz matrices T_2 and T_3 are given by Eq.(2.20) as,

$$T_2 = \begin{bmatrix} 0 & 0 & 0.5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.5 & 0 & 0 \\ 0.5 & 0 & 0 & 0 & 0.5 & 0 \\ 0 & 0.5 & 0 & 0 & 0 & 0.5 \\ 0 & 0 & 0.5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.5 & 0 & 0 \end{bmatrix} \quad \text{for } i = 2, \tag{3.20}$$

$$T_3 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0.5 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.5 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0.5 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.5 & 0 & 0 & 0 & 0 \end{bmatrix} \quad \text{for } i = 3. \tag{3.21}$$

The set of constraints in Eq.(3.6), using Eqs.(3.20)-(3.21) are given as

$$ac + bd + ce + df = 0, \tag{3.22}$$

$$ae + bf = 0. \tag{3.23}$$

The constraint in Eq.(3.7) ensures a zero DC-mean for the high pass filter. Thus,

$$a + b + c + d + e + f = 0. \tag{3.24}$$

It is to be noted that the minimum filter length, N, for an optimized solution is $N = 6$. The only solution for the case $N = 2$ is the Haar wavelet filter. Thus filters only with lengths $N >= 4$ are considered in this work.

The problem is now proved to be non-convex as explained earlier. Applying Eq.(3.11) to the problem, we have for convexity,

$$\begin{bmatrix} a & b & c & d & e & f \end{bmatrix} \begin{bmatrix} 1 & r_{12} & r_{13} & r_{14} & r_{15} & r_{15} \\ r_{12} & 1 & r_{12} & r_{13} & r_{14} & r_{15} \\ r_{13} & r_{12} & 1 & r_{12} & r_{13} & r_{14} \\ r_{14} & r_{13} & r_{12} & 1 & r_{12} & r_{13} \\ r_{15} & r_{14} & r_{13} & r_{12} & 1 & r_{12} \\ r_{15} & r_{15} & r_{14} & r_{13} & r_{12} & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \\ e \\ f \end{bmatrix} \geq 0. \tag{3.25}$$

Eq.(3.25) simplifies as

$$\begin{aligned} &(a^2 + b^2 + c^2 + d^2 + e^2 + f^2) + 2r_{16}[af] + 2r_{15}[ae + bf] + 2r_{14}[ad + be + cf] \\ &+ 2r_{13}[ac + bd + ce + df] + 2r_{12}[ab + bc + cd + de + ef] \geq 0 \end{aligned}. \tag{3.26}$$

Let us denote $f(a,b,c,d,e,f)$ as

$$\begin{aligned} f(a,b,c,d,e,f) &= (a^2 + b^2 + c^2 + d^2 + e^2 + f^2) + 2r_{16}[af] + 2r_{15}[ae + bf] \\ &+ 2r_{14}[ad + be + cf] + 2r_{13}[ac + bd + ce + df] + 2r_{12}[ab + bc + cd + de + ef] \end{aligned}. \tag{3.27}$$

This function does not consider the convexity condition in Eq.(3.25). This function is further simplified by using the constraints in Eqs.(3.19)-(3.24) and given as,

$$f(a,b,c,d,e,f) = (1 - r_{12}) - 2(r_{12} - r_{14})(ad + be + cf) - 2(r_{12} - r_{16})af \ . \tag{3.28}$$

A complete proof of the simplification is given in Appendix A. This is the simplified objective function with the constraints included for $N = 6$. This function is plotted with possible coefficient values which satisfy all the constraints.

Also from Eq.(3.19), it can be inferred that,

$$-1 \le ad, be, cf, af \le 1. \tag{3.29}$$

This inference is made since the maximum possible value of the square of the filter coefficients is unity.

Fig. 3.4 is a 3-D plot of the function in Eq.(3.28) over the possible filter coefficient values. It can be seen in Fig.3.4 that the function has more than a single minimum, showing that the problem is highly non-convex for the case of $N = 6$. The figures show the many local minima solutions to the adaptive filter design. To simplify plotting, Fig. 3.4 is a plot of the objective function, f, against only the first and the second coefficients of the high-pass filters $h[1]$ and $h[2]$, respectively. The various minima are observed to be close to each other and the optimum solution is very close to these local solutions. Fig. 3.5 shows the 2-D plot of the same graph above. The objective function is plotted against the index values of its vector.

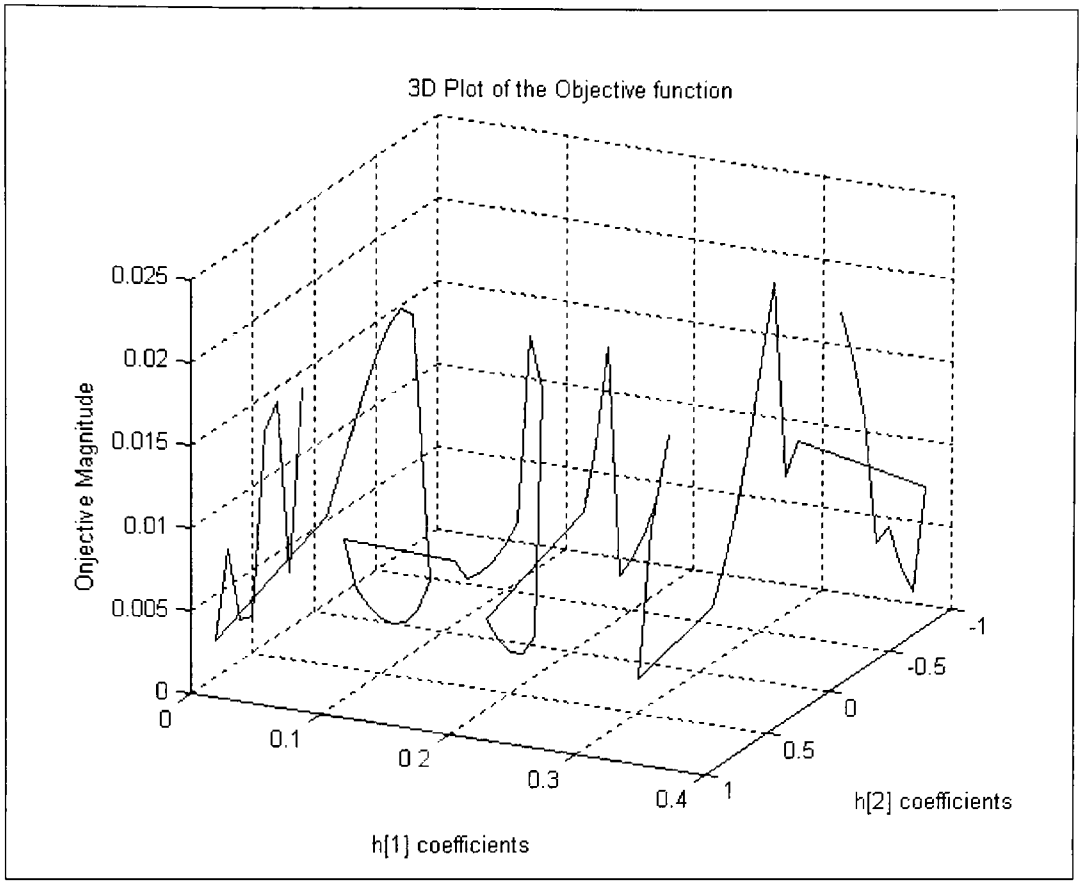


Fig. 3.4: 3-D Objective Plot with $h[1]$ and $h[2]$ coefficients

The earlier work [1] transformed the original non-convex problem to convex and made use of a primal-dual Interior-Point algorithm. In general, constrained convex problems are easier to solve than constrained non-convex problems, since only one minimum has to be found, which also happens to be the global solution. However, the globally constrained non-convex optimization problem can be solved by finding all local minima and then determining the global minimum, although the computational cost will be higher in this case. Until recently, Interior-Point methods and Sequential Quadratic Programming methods were available only for convex problems and the few other

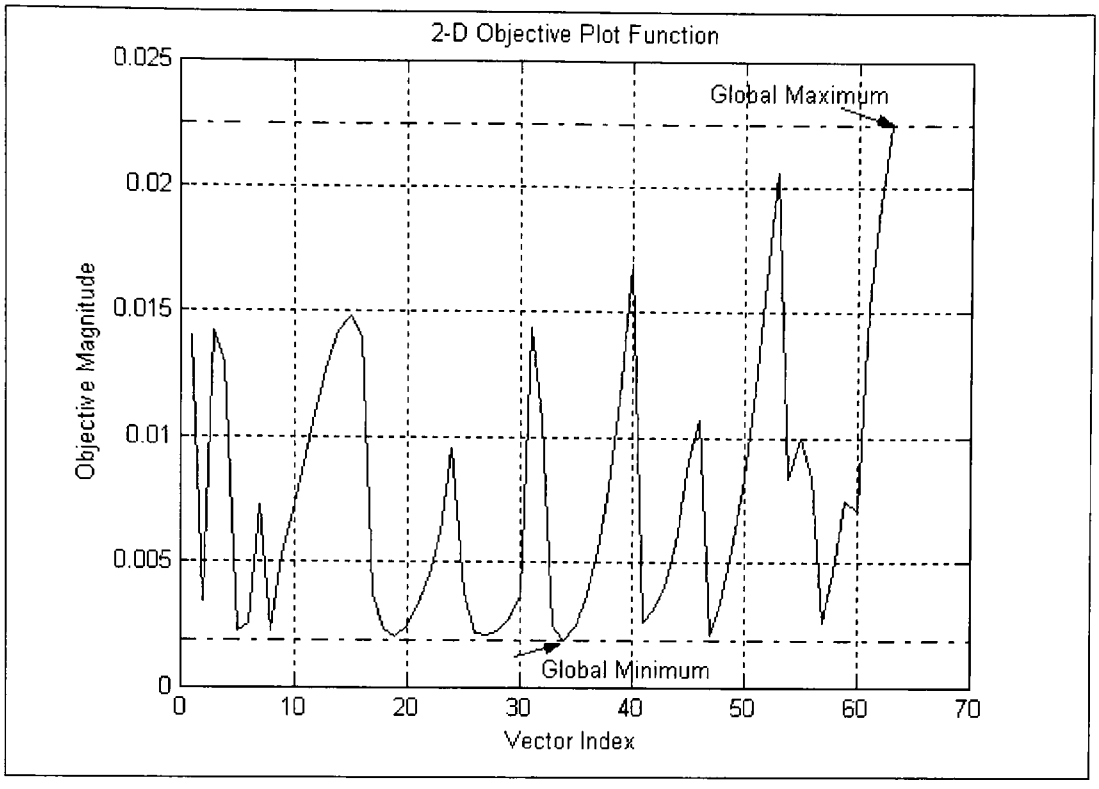


Fig. 3.5: 2-D plot of the objective function

methods available for non-convex problems did not guarantee global convergence and these methods were not comparable in speed to algorithms for convex problems. Only recently have general non-convex problems been the subject of research [13], and many advances have been made which give a very feasible computational performance and also almost guarantee global solutions. Such an algorithm is used in solving the problem in its original form.

3.5 SQP Optimization Implementation

The optimization technique used to solve this problem is discussed in detail. The optimization problem is a non-linear non-convex problem. The problem is solved using Sequential Quadratic Programming (SQP) [13] optimization, and it is shown to outperform other tested methods in terms of efficiency and accuracy over a large number of problems [14].

The proposed method also makes use of the Interior-Point optimization principle in that the iterate solution path is restricted to lie in the feasible region. This promises faster convergence to the optimal solution compared to other methods. This SQP method for constrained optimization gives the equivalent of Newton's method for unconstrained optimization problems. The analysis of the problem is simplified by considering the problem as [13],

$$\begin{aligned} &\text{Minimize } f(h) = h^T R_{xx} h \\ &\text{subject to } \quad g(h) = 0, \end{aligned} \tag{3.30}$$

where functions f and g are known to be smooth and h refers to the high-pass filter h_1

$$g(h) = \begin{bmatrix} h^T h \ -1 = 0 \\ h^T T_i h = 0 \\ \cdots \\ \sum h\ [n] = 0 \end{bmatrix} \quad i = 2 \ldots \frac{N}{2}. \tag{3.31}$$

The Lagrangian of the problem in Eq.(3.30) is,

$$L(h\ , y) = f(h) - \sum_{i=1}^{m} y_i g_i(h), \qquad i = 2 \ldots \frac{N}{2} \tag{3.32}$$

where y is the Lagrangian multiplier vector of the constraints.

The optimization routine makes an approximation of the Hessian of the Lagrangian in Eq.(3.32) at every iteration using a quasi-Newton method [14]. The original problem is then transformed into an unconstrained Quadratic Problem (QP) using a merit function.

For such problems reliable solution procedures like Newton descent and line search are readily available. The approximate Hessian is used in determining the descent conditions.

Of the various methods available for updating the Hessian, the formula given by Broyden, Fletcher, Goldfarb, and Shanno (BFGS) [15,16,17,18] which is popularly used in SQP implementation is used here. A positive definite quasi-Newton approximation of

the Hessian, H, of the Lagrangian in Eq.(3.32) is calculated using the BFGS method given as,

$$H_{k+1} = H_K + \frac{q_k q_k^T}{q_k^T s_k} - \frac{H_k^T H_k}{s_k^T H_k s_k}, \tag{3.33}$$

where $s_k = h_{k+1} - h_k$,

$$q_k = \nabla f(h_{k+1}) + \sum_{i=1}^{n} y_i . \nabla g_i(h_{k+1}) - \left(\nabla f(h_k) + \sum_{i=1}^{n} y_i . \nabla g_i(h_k) \right).$$

A positive definite Hessian is maintained, providing the term $q_k^T s_k$ is positive at each update and that H is initialized with a positive definite matrix. If $q_k^T s_k$ is not positive, then q_k is modified in steps until $q_k^T s_k$ is positive. In the initial phase, the most negative element of $q_k .* s_k$ is repeatedly halved until $q_k^T s_k$ is greater than a tolerance. If the term is not yet positive, then q_k is modified by adding a vector v multiplied by a scalar w as,

$$q_k = q_k + wv, \tag{3.34}$$

where

$$v_i = \nabla g_i(h_{k+1}) . g_i(h_{k+1}) - \nabla g_i(h_k) . g_i(h_k)$$

$$\text{if } (q_k)_i \, w \, \pi \, 0 \quad \text{and} \quad (q_k)_i . (s_k)_i \, \pi \, 0, \qquad i = (1,...m)$$

$$v_i = 0 \text{ otherwise,}$$

and also w is increased systematically until $q_k^T s_k$ is positive.

Here $m = [(N/2) - 1]$ in all cases. The first-order necessary condition for a local minimizer, h_*, of the problem in Eq.(3.30) is

$$\nabla (h_*^T R_{xx} h_*) - \nabla g(h_*)^T y = 0, \tag{3.35}$$

$$g(h_*) = 0.$$

∇g denotes the Jacobian matrix of the equality constraints. The solutions (h_k, y_k) of the above problem lie in the interior of the feasible region, and form the so-called central path. For a non-linear problem, the first-order conditions gives a local solution but do not

guarantee a global solution and hence a merit function is incorporated in the objective [13] as

$$\phi(h) = h^T R_{xx} h - g(h)^T \, y^{\sim}(h) + \frac{1}{2} c \left\| g(h) \right\|^2, \tag{3.36}$$

where

$$y^{\sim}(h) = [\nabla g(h)^T \nabla g(h)]^{-1} \nabla g(h)^T h^T R_{xx} h \,. \tag{3.37}$$

$c \geq 0$ is the penalty parameter. The merit function guides the iterates of the algorithm to a minimum solution. This is achieved by ensuring that the merit function is reduced by the penalty parameter at each iterate.

Newton's method is used of to solve the optimality conditions in Eqs.(3.35). At the k^{th} iteration, the system of linear equations to determine the Newton directions is,

$$\begin{bmatrix} H_k & -\nabla g(h_k)^T \\ \nabla g(h_k) & 0 \end{bmatrix} \begin{bmatrix} \Delta h_k \\ \Delta y_k \end{bmatrix} = -\begin{bmatrix} -\nabla L(h_k, y_k) \\ g(h_k) \end{bmatrix} = -F(h_k, y_k) \,. \tag{3.38}$$

The Newton direction, $\Delta w_k = (\Delta h_k, \Delta y_k)$, is then used to find the next iterate,

$$w_{k+1} = w_k + A_k \Delta w_k \tag{3.39}$$

$$\text{where } A_k = diag(\alpha_k I_N, \chi_k I_m) \,.$$

Here α_k and χ_k are the step-size in the interval [0,1] for the iterates h_k and y_k, respectively. I_N and I_m are identity matrices or order N and m, respectively. Furthermore

$$\alpha_k^{\max} = \gamma \max_{1 \leq j \leq N} \left\{ -\frac{h_k^{(j)}}{\Delta h_k^{(j)}} : \Delta h_k^{(j)} < 0 \right\} \tag{3.40}$$

represents the maximum allowable step-size, which guarantees that the iterates h_k always remain strictly feasible, for some chosen constant $\gamma \in (0,1)$.

The distance of the current solution from the central path is measured by the Euclidean norm, $\left\| F(h_k, y_k) \right\|$. Once this measure is less than a certain threshold value, the

iteration can be terminated.

The merit function in Eq.(3.36) provides a measure of progress towards an optimum solution of the problem in Eq.(3.30). A line-search procedure for adjusting the step-lengths of the variables is used in order to guarantee that the function decreases at each iterate. The merit function is modified in order to mimimize the computational effort of calculating $\nabla\phi(h_k)$ in a standard line search approach. The line-search formulation is,

$$\phi_k(\alpha_{k,i}) \le \phi_k(0) + \rho\alpha_{k,i}\phi_k'(0) \tag{3.41}$$

where

$$\phi_k'(0) = \nabla(h^T R_{xx} h)^T \Delta h_k - \frac{1}{\alpha_{k,i}}[y^{\sim}(h_k + \alpha_{k,i}\Delta h_k) - y_k^{\sim}]^T g_k - y_k^{\sim T} g_k \Delta h_k - c_{k,i}\|g_k\|^2 . \tag{3.42}$$

Also

$$\phi_k(0) = \phi(h_k),$$

$$\phi_k(\alpha_{k,i}) = \phi(h_k + \alpha_{k,i}\Delta h_k),$$

$$y_k^{\sim} = y^{\sim}(h_k).$$

Since $\phi_k'(0)$ does not require the calculation of the second derivatives of the objective and the constraints, the approximate merit function ϕ_k is preferable to the initial merit function ϕ .

Now we describe the mechanism which ensures that Δh_k is a descent direction which guarantees the reduction of the merit function and the algorithm is converging towards the central solution. As long as the Hessian of the Lagrangian, H_K , is positive definite, the direction is assured to be descent. Eq.(3.33) and Eq.(3.34) ensure that H_K is positive definite throughout the algorithm. In order that the function ϕ_k is reduced at each iterate, its first derivative $\phi_k'(0)$ should be negative. From Eq.(3.42) it can be noted that

this would happen if $c_{k,i}$ large enough. Hence we select the value of $c_{k,i}$ such that the descent condition is satisfied,

$$\phi_k^{'}(0) \le -\frac{1}{2}\left[\Delta h_k^T H_k \Delta h_k + c_{k,i}\|g_k\|^2\right] \le -\frac{1}{4} c_{k,i}\|g_k\|^2 \le 0. \tag{3.43}$$

The index i represents the number of times the descent condition is checked at the k^{th} iteration. If Eq.(3.43) is satisfied, the penalty parameter does not increase, that is, $c_{k,i+1} = c_{k,i}$, otherwise the penalty parameter is determined as,

$$c_{k,i} = Max\Big\{\ 2c_{k,i}, \frac{-2\Delta h_k^T H_k \Delta h_k}{\|g_k\|^2}, \\ \frac{2}{\|g_k\|^2}\left[\frac{1}{2}\Delta h_k^T H_k \Delta h_k + \nabla g_k \Delta h_k - \frac{1}{\alpha_{k,i}}\left[\tilde{y}(h_k + \alpha_{k,i}\Delta h_k) - \tilde{y_k}\right]^T g_k - \tilde{y_k}^T \nabla g_k \Delta h_k\right]\Big\} \tag{3.44}$$

The new value $c_{k,i+1}$ of the penalty parameter guarantees that $\phi_k^{'}(0) \le 0$. Then the line-search condition given by Eq.(3.41) is checked for the current values $c_{k,i+1}$ and step-size $\alpha_{k,i}$. If it is not satisfied, then the step-size is reduced by choosing a new value $\alpha_{k,i+1}$ from the interval $[\beta_1\alpha_{k,i}, \beta_2\alpha_{k,i}]$, where $\beta_1, \beta_2 \in (0,1)$ and $\beta_1 \le \beta_2$. This process is repeated until Eq.(3.41) is satisfied.

If Eq.(3.43) is satisfied instead, we set the new iterates as,

$$h_{k+1} = h_k + \alpha_{k,i}\Delta h_k \tag{3.45}$$

$$c_{k+1,0} = c_{k,i+1}$$

This is carried until the norm, $\|F(h_k, y_k)\|$ reaches the desired threshold.

3.6 SQP Algorithm

The complete algorithm is given below:

Initially, at the beginning of the k^{th} iteration, (h_k, y_k) is available.

Set $c_{k,0} \ge 0$; $\rho > 0$; $\beta_1, \beta_2 \in [0,1]$ with $\beta_1 \le \beta_2$;

Set $\chi_k = 1$ and the tolerance threshold, τ, as $0 \le \tau \le 1e-3$

Repeat Until: $\|F(h_k, y_k)\| \le \tau$.

Compute H_K using Eq.(3.33) and Eq.(3.34).

Compute Newton direction $(\Delta h_k, \Delta y_k)$ using Eq.(3.38).

Set $i = 0$ and $\alpha_{k,i} = \alpha_k^{\max}$ as in Eq.(3.40).

If Eq.(3.43) is satisfied

Then set $c_{k,i+1} = c_{k,i}$.

Else

Compute $c_{k,i+1}$ from Eq.(3.44).

Repeat Until: Eq.(3.41) is satisfied.

Choose $\alpha_{k,i+1} = [\beta_1 \alpha_{k,i}, \beta_2 \alpha_{k,i}]$.

Set $i = i+1$.

End

Set $i_k = 1$; $h_{k+1} = h_h + \alpha_{k,ik} \Delta h_k$; $c_{k+1,0} = c_{k,ik}$; $y_{k+1} = y_k + \chi_k \Delta y_k$; $k = k+1$.

End

Sample values for the variables are: $c_{o,o} = 0; \beta_1 = 0.05; \beta_2 = 0.5; \rho = 1e-3$.

This gives the complete algorithm used for the optimization. The next chapter gives the results for the image source encoding application.

CHAPTER 4
APPLICATION TO IMAGE SOURCE ENCODING

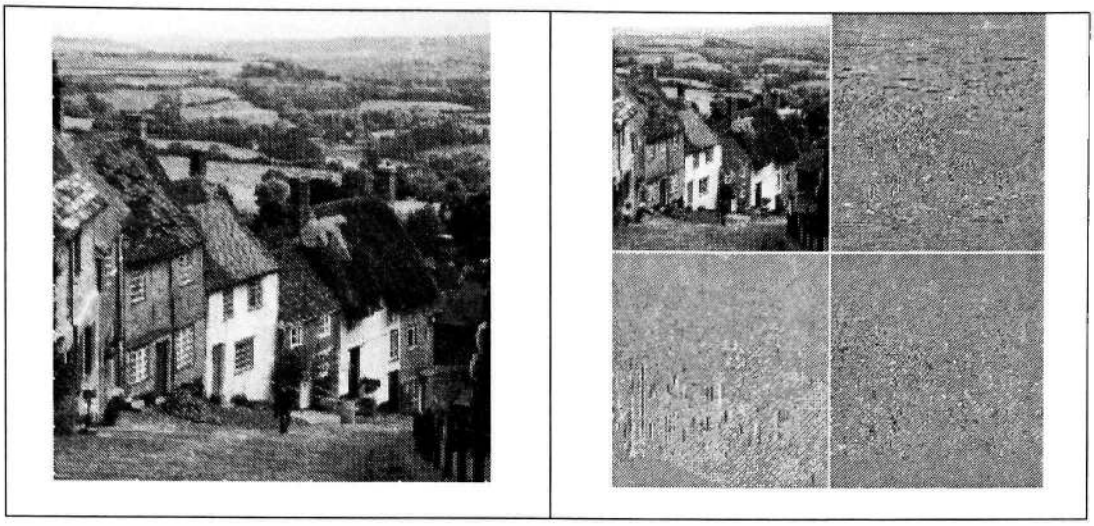
This chapter deals with the application of the filter bank optimization algorithm to image source input. As shown in the previous chapters, it is possible to redistribute the energy of the image among its sub-bands, so that one sub-band receives most of the energy. Since most of the natural images have a majority of low frequency spatial components (below $\pi/2$), they are considered to be low-pass inputs.

Since the wavelet transform is reversible, the algorithm given in the previous chapters can be easily extended for the 2D case, as in an image, by computing the discrete wavelet transform along each dimension. This 2D DWT of the image results in an average image and three directionally sensitive detail images. These newly formed images are called subbands as shown in Fig. 4.1(b).

4.1 DWT Analysis of an Image

Fig. 4.2 shows the wavelet analysis of an image's subbands. In the figure Low-Low denotes the low-pass sub-band in both the horizontal and the vertical directions of the 2-D image. High-High denotes the high-pass sub-band in both the directions. Similarly Low-High and High-Low denote corresponding sub-bands in the two directions. This gives a single resolution wavelet analysis of the image. For a multi-resolution representation, just the Low-Low band is considered and split iteratively into its respective sub-bands. The present work concentrates on single level resolution analysis only, though it can be easily applied to more levels. This kind of analysis is exploited in many wavelet-based lossy compression techniques like EZW [19], LBG-VQ [20], and HMVQ [21,22,23,24].

A brief explanation about the statistical analysis of the subimages obtained after the wavelet transform is given. It also highlights the advantage of using such a transform



(a) Goldhill Image (b) Wavelet decomposed subbands

Fig. 4.1: Example of wavelet decomposition of an image

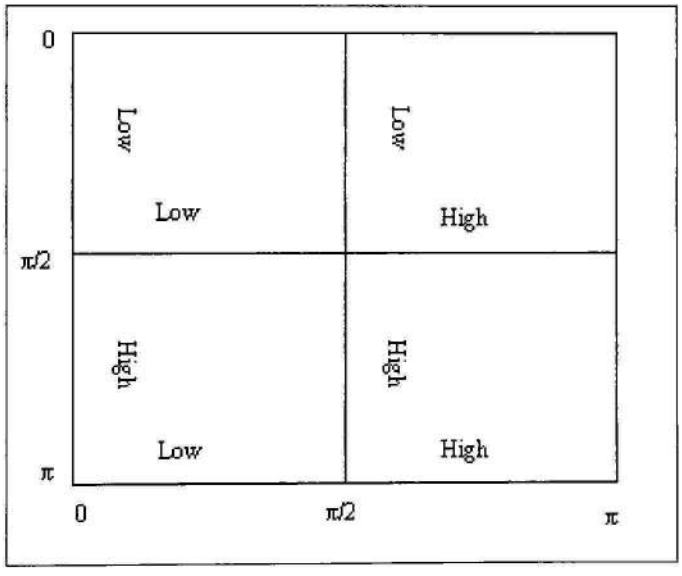
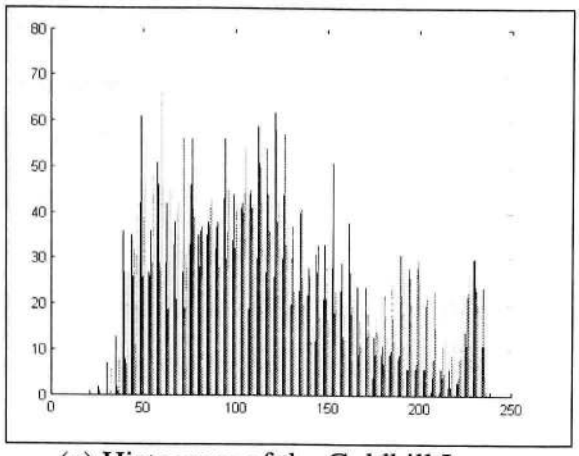
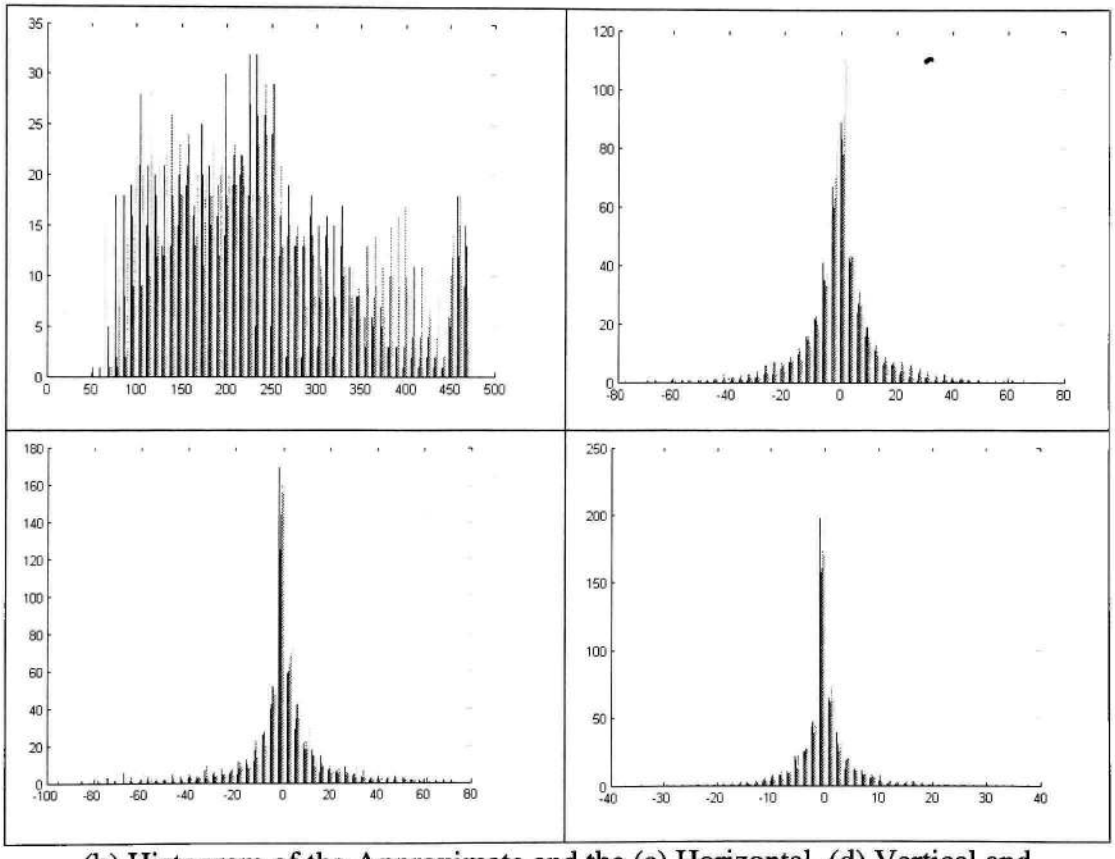


Fig. 4.2: 2-D Wavelet analysis of an image

in image source encoding applications. Figs. 4.3 (a)-(e) shows the histogram of the original image and the subimages respectively, which were plotted from the dyadic wavelet transform on the Goldhill image of Fig. 4.1, with Daub 4 filters. The



(a) Histogram of the Goldhill Image



(b) Histogram of the Approximate and the (c) Horizontal, (d) Vertical and (e) Diagonal Detail subimages (clockwise from top left corner)

Fig. 4.3: Histograms of an image and its subimages

normalized histogram, or pdf, of a subimage provides such statistical information about the coefficients of the subimage as the mean and variance of its distribution [25]. The histograms of the subimages shown in Figs. 4.3 (b)-(e) are very typical and can be observed with various other similar images. The histograms of the subimages in Fig. 4.3 (b)-(e) correspond to the respective subbands in Fig. 4.2. We note that the approximate coefficient distribution in the Low-Low subband follows the histogram of the original image very closely, and thus contains most of the image information. The detail subimage histograms on the other hand are very narrow and have fewer coefficients. Compared to the pdf of the original image, which is normally multimode and also difficult to analyse, the detail distributions are single mode and quite narrow. Its variance is low and its mean is zero indicating that there are many wavelet coefficients with low values. Thus by discarding the detail coefficients, about $3/4^{th}$ of the image size is reduced and at the same time, most of the energy is retained by the approximate coefficients, which results in minimum distortion. This Low-Low subband containing almost the same information as the original, can now be subjected to vector quantization techniques with high-fidelity reconstruction ability as in [21].

As pointed out in the earlier chapters, the coding performance of wavelet based image coding depends on the choice of the wavelets [26]. The optimal wavelet basis for the given image, which depends on the image statistics, would give the maximum coding gain (which is explained later) and the best decomposition in the Low-Low subband. Such a subband system depends upon the frequency selectivity of the filter bank that is designed. In order to obtain perfect decorrelation of interband coefficients of the subbands to have perfect reconstruction, the filter banks have to be ideal with infinite duration basis functions. Such an infinitely long impulse response would provide high frequency localization. At the same time it is very important, in visual signal processing, to have spatial localization, which implies that the impulse response of the filter cannot be very long. Various other properties that determine the performance of the wavelet filter bank are discussed.

4.2 Image Dependent Statistics

A large sampling of images having varying activity levels has been tested and the results of the output image and the filter bank have been tabulated. One of the ways to determine the activity level is to calculate the Spectral Flatness Measure (SFM) [27] of the given image. This quantity measures the redundancy as observed in the structure and the shape of the given power spectral density (PSD) function. This measure gives the overall activity of the image and is given as the ratio of the geometric mean and the arithmetic mean of the samples of the PSD. For two dimensional digital images, it is given as [24]

$$SFM = \frac{\left[\prod_{i=0}^{R-1}\prod_{j=0}^{C-1}|F(i,j)|^2\right]^{1/RC}}{\frac{1}{RC}\sum_{i=0}^{R-1}\sum_{j=0}^{C-1}|F(i,j)|^2} \tag{4.1}$$

where $F(i,j)$ is the $(i,j)^{th}$ Fourier Transform of the RxC dimensional image I. The SFM has a dynamic range of [0,1]. The inverse of the SFM gives the predictability measure of the image and has a range of [1,∞]. A totally flat spectrum has a SFM value of 1 and is related to a completely unpredictable image. On the other hand, a SFM value of 0 is related to a completely predictable image, that is, an image having all pixels except one with the same value A lower SFM value indicates a more predictable image, which means that the energy of the image is concentrated in fewer wavelet coefficients and hence gives better performance. Active images (with SFM close to 1) are difficult to compress as they contain a large number of details and lower spatial redundancy. A much easier way to comprehend this measure is to use the spectral activity measure (SAM), which is the inverse of the SFM.

$$SAM = \frac{1}{SFM}$$

The SAM is used more often than the SFM. The SAM value indicates proportionally the predictability of a given image. Images which have a SAM value of

close to 1 are highly unpredictable compared to those having values towards infinity.

Another measure for an image statistic is the Spatial Frequency (SF) [28], which gives the SFM equivalent in the spatial domain. It gives the overall activity in the spatial domain and is given in Appendix B. Higher values of SF mean that the image contains relatively larger components in the high frequency area.

It is important to consider the order of the wavelet filter used. A higher order filter implies good frequency localization and thus higher energy compaction. Also a higher order increases the regularity of the wavelet. Regularity of a wavelet filter represents the amount of smoothness that can be imposed. At the same time, a lower filter order would have better time localization and thus preserve edge information better, which can be crucial. A proper intermediate choice has to be made to give the best results. A measure of this is given by the Time-Frequency Localization [26] as described in Appendix B.

Another popular measure of performance is the Transform Coding Gain (TCG) [3]. It is a measure of energy compaction of the transform. It is given as,

$$TCG = \frac{1/2(\sigma_L^2 + \sigma_H^2)}{\left[\sigma_L^2 \sigma_H^2\right]^{1/2}}. \tag{4.2}$$

This term is the ratio of the arithmetic mean of the energy, σ_j^2, in each subband to the geometric mean.

4.3 Results

The algorithm was applied to a large sample of images, and the results follow. CQF filters can be designed for any order to satisfy the Nyquist (2) criteria. Also there is an unique filter bank for order $N = 2$, corresponding to the Haar wavelet transform, hence $N = 4$ is the shortest case for which there are many solutions. The results are compared with Daubechies filters, also referred to as Max Flat filters, since the low-pass filter has the maximum number of zeros at $\omega = \pi$ corresponding to $Z = -1$ in the Z-plane for the low-pass filter or maximum number of zeros at $\omega = 0$ corresponding to $Z = 1$ for the high-pass filter [29]. As was shown in [30], Daubechies Max Flat filters of length $N = 4$ have the optimum energy compaction and aliasing distortion under frequently met assumptions

on the autocorrelation sequence. All the results are compared with Daubechies (Daub) filters only. Also, since the filter length has to be neither too short nor too long to have a satisfactory frequency and time localization, tests were limited to the longest length of $N = 24$.

Fig. 4.4 shows the transform coding gain comparison between the adaptive and optimized filter and Daub filter for various filter lengths. It can be seen that for filter order of $N = 4$, the adaptive filter is the same as the Daubechies filter. This can also be confirmed from the tabulated results in Table 4.1 for all the tested images. This observation confirms two results.

1. For filter order $N = 4$, there is only a single solution for the perfect reconstruction criteria.

2. Since Daubechies filters give the optimum energy compaction for length $N = 2$, the optimizing algorithm successfully finds better filters than the Daubechies filters. This can further be observed in the Table 4.1, which shows that the adaptive filter performs better than the Daubechies in more than 90% of the cases. It can also be noted that the gain does not increase considerably beyond the filter order $N = 10$.

This can also be observed in the Fig. 4.5, which is the plot of the PSNR values of the reconstructed Mandrill image from the approximation wavelet coefficients for various filter lengths. It is seen that the performance of the adaptive filter is the same as the Daubechies filter for $N = 4$. Beyond that length, the adaptive filter gives a better PSNR than the Daubechies filter, thus finding better solutions. Since the problem is non-convex, further tests would be required to determine if the obtained solutions are indeed global. The Mandrill image performs poorer when compared to the other images having a lower spatial frequency value, as seen in Table 4.3, since it has a high SF value. This performance does not come as a surprise and is as expected. Similarly, the Straw image performs poorly, as it has a lot of edge or high frequency information compared to the other tested images.

This observation is further strengthened by the plot in Fig. 4.6 of the max PSNR values of the 25 tested images against the respective SF values. Except for a very few

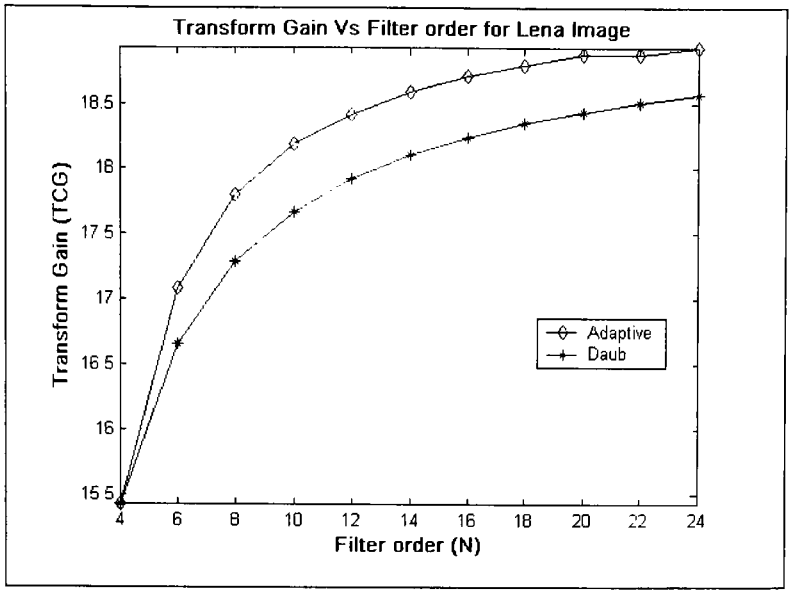


Fig. 4.4: Plot of TCG versus Filter order for Lena Image

cases, the general trend is decreasing PSNR values as the SF value increases. Since a higher SF value indicates high edge information, a lot more reconstruction aliasing takes place as more of the high frequency information is discarded by the low-pass filter.

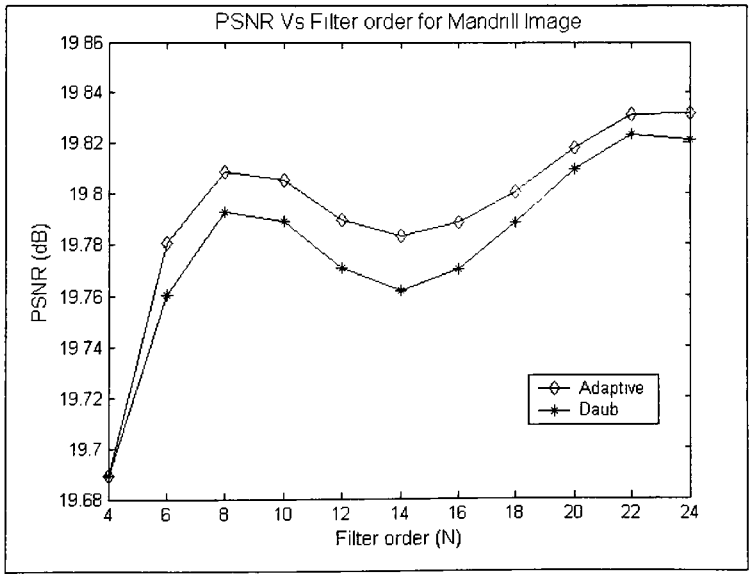


Fig. 4.5: Plot of PSNR versus Filter order for Mandrill Image

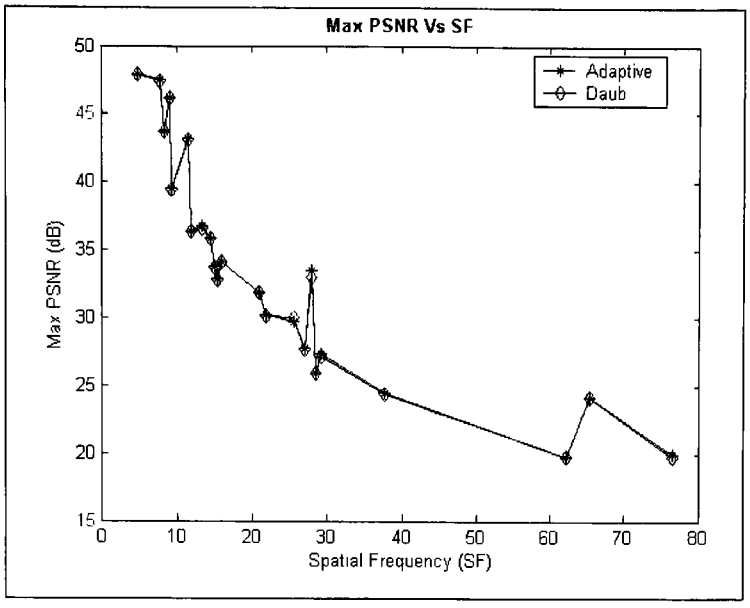


Fig. 4.6: Plot of PSNR versus SF of test images

The result of the SQP optimization is also compared to similar work done earlier [1] using Interior-Point optimization, and is shown in Table 4.1. It can be seen that the current SQP method performed better in terms of both the MSE and the PSNR values for the four standard images. Another improvement is in the computational time of the algorithm, which averages about 13 seconds using the Interior-Point method and just 2.5 seconds using the SQP method (using Daubechies filter as the starting point) described. All the tests were performed using a 1.54 GHz AMD Athlon and 480 MB RAM computer.

The present method performed better for all images considered except the Mandrill image. The performance improvement has been significant in a few cases and marginal in a few. Though exact comparison data in terms of the computational time is not available, the SQP method seems to perform better compared to the Interior point method in speed too. Hence, an overall improvement has been achieved in the present case compared to the earlier method.

Table 4.1: Comparison of results for Interior point and SQP methods

Image	Filter Order	Interior Point Method		SQP Method	
		Optimized MSE	Optimized PSNR	Optimized MSE	Optimized PSNR
Lena	N=4	27.80	33.69	22.82	34.55
	6	22.61	34.59	17.65	35.66
	8	20.88	34.93	16.13	36.05
Goldhill	4	54.80	30.74	41.13	32.0
	6	53.31	30.86	36.40	32.52
	8	48.64	31.26	35.26	32.66
Mandrill	4	275.6	23.72	698.4	19.69
	6	271.9	23.79	683.8	19.79
	8	264.8	23.90	679.5	19.81
Peppers	4	53.89	30.82	35.86	32.59
	6	48.18	31.30	21.67	35.06
	8	38.75	32.24	17.73	35.65

An explanation of the use of the autocorrelation R_{xx}, used in the objective function, is given here. For images, the autocorrelation vector is the sequence created by shifting either the columns or the rows, depending on whether it is the horizontal or the vertical autocorrelation desired, and then an averaging is performed [1]. Eqs.(3.14)-(3.15) give the mathematical form of the operations described above for the horizontal sequence. The vertical autocorrelation sequence is formed the same way, but by shifting rows instead of columns. Strictly speaking, the filter that is optimized using the horizontal autocorrelation sequence is the optimal filter in only the horizontal direction, and the filter optimized using the vertical autocorrelation sequence is optimal in the vertical direction only. This is not a problem, since we can design separable filters, which means we construct an optimal filter bank consisting of eight (four filters in each

dimension) instead of four filters. Frequently, the horizontal and vertical autocorrelation sequences are significantly correlated, and the filters optimized using only the horizontal or vertical autocorrelation sequence will perform well in both dimensions for almost all the images. This was the observation and is also shown in Table 4.2, which gives the results of the algorithm when the horizontal and vertical directions were used alternatively for the autocorrelation sequence. At N=4, the difference was zero and was very small for other cases and can be considered negligible. Though this test was performed for only four images, the observation can be extended to almost all the images. This takes care of the direction to be considered for the sequence.

Another area of interest from the point of the optimization method is the starting point of the algorithm. It is known [1] that the solution of the problem $h_1^T R_{xx} h_1 \rightarrow \min$ for any symmetric matrix is the eigenvector of R_{xx} that corresponds to the smallest eigenvalue of R_{xx}. Since $h_1^T R_{xx} h_1$ is the output energy of the signal that passes through the filter $h_1(n)$, and $h_1^T R_{xx} h_1 > 0$, R_{xx} is positive definite, i.e., its smallest eigenvalue (equal to the value of the output energy) is greater than zero. This implies that the optimal solution is the eigenvector corresponding to the smallest eigenvalue, or in the vicinity of it. With this knowledge, the starting point was chosen to be the above eigenvector to obtain the optimum filter.

To be able to qualify the results obtained by the above method, a comparison was made with a different starting point. The obvious choice was using the Daubechies filters as the starting point, since they are known to be maximally flat filters and having the maximum number of zeros at $\omega = \pi$ for a given length of the low-pass filter. The results of the comparison are given in Table 4.3. The results come as little surprise, as the Daubechies filters almost always perform better in finding the optimum solution when compared to the eigenvector starting points. The performance gain is as high as 54% for the Daubechies filters in some cases. The Daubechies filters give a better and also a more consistent output for various filter lengths. The Daubechies filters thus become the obvious choice for the starting point and all the tests were performed with them.

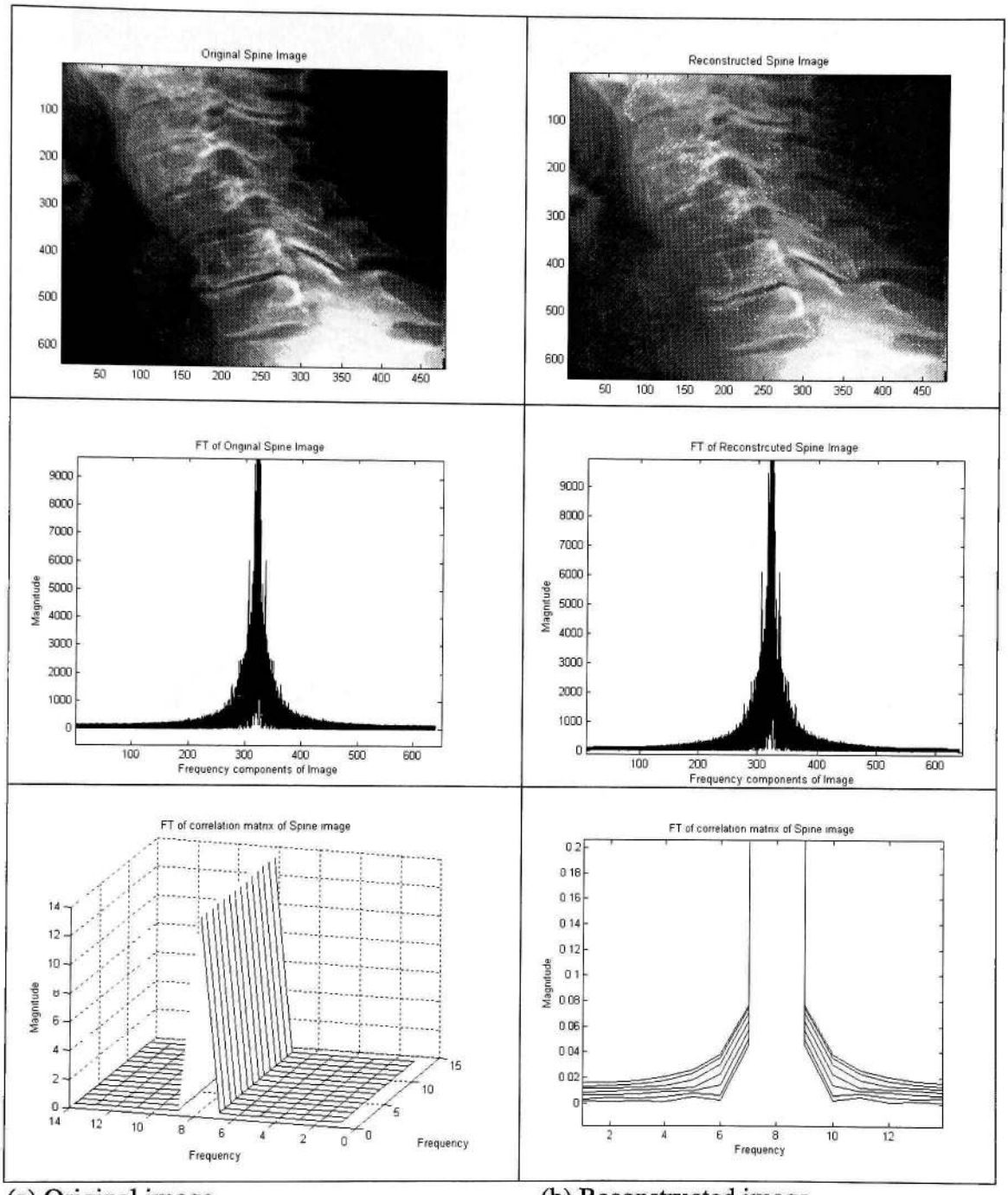
Table 4.2: Comparison of the results by using either the horizontal or the vertical autocorrelation of the image

Image	Filter order	Horizontal		Vertical correlation		% change
		MSE	PSNR(dB)	MSE	PSNR(dB)	
Lena	N=4	22.8209	34.5475	22.8209	34.5475	0
	6	17.6597	35.6610	17.6537	35.6625	0.0042
	8	16.1357	36.0529	16.1488	36.0494	0.0097
	10	15.4870	36.2311	15.7735	36.1515	0.2197
	12	15.0266	36.3622	16.0518	36.0756	0.7882
Goldhill	N=4	41.1269	31.9895	41.1269	31.9895	0
	6	36.3962	32.5202	36.3054	32.5311	0.0335
	8	35.2614	32.6578	35.2378	32.6607	0.0089
	10	34.6027	32.7397	34.5988	32.7402	0.0015
	12	33.9859	32.8178	34.0159	32.8140	0.0116
Mandrill	N=4	698.4679	19.6893	698.4679	19.6893	0
	6	683.8821	19.7810	683.8636	19.7811	0.0005
	8	679.5712	19.8085	679.5439	19.8086	0.0005
	10	680.0534	19.8054	680.0227	19.8056	0.0010
	12	682.4979	19.7898	682.9625	19.7868	0.0152
Peppers	N=4	35.8618	32.5845	35.8618	32.5845	0
	6	20.2854	35.0590	20.2490	35.0668	0.0222
	8	17.7156	35.6472	17.8995	35.6024	0.1257
	10	20.9640	34.9161	21.2096	34.8655	0.1449
	12	24.9685	34.1569	24.8443	34.1785	0.0632

Table 4.3: Comparison of the results by using either the Daub filters or the eigenvector corresponding to minimum eigenvalue as starting point for the optimization algorithm.

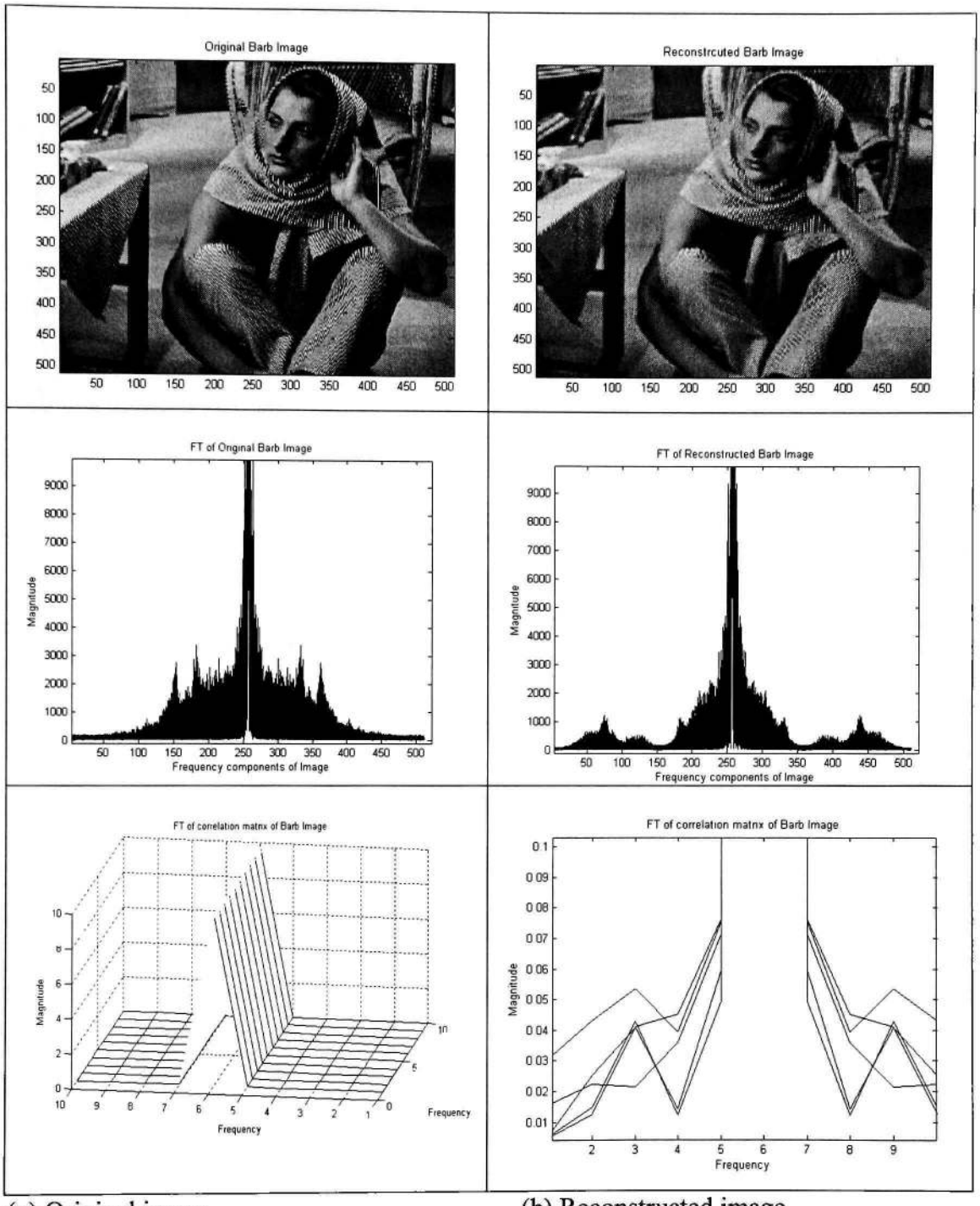
Image	Filter order	Daub filters as starting point		Min. Eigen vector as starting point		% change in PSNR
		MSE	PSNR(dB)	MSE	PSNR(dB)	
Lena	N=4	22.8209	34.5475	41.5349	31.9467	8.1411
	6	17.6597	35.6610	22.7218	34.5664	3.1667
	8	16.1357	36.0529	16.1508	36.0489	0.0111
	10	15.4870	36.2311	34.4626	32.7573	10.604
	12	15.0266	36.3622	281.834	23.6309	53.875
Goldhill	N=4	41.1269	31.9895	41.1287	31.9894	0.0003
	6	36.3962	32.5202	164.292	25.9746	25.200
	8	35.2614	32.6578	241.863	24.2951	34.421
	10	34.6027	32.7397	34.7742	32.7182	0.0657
	12	33.9859	32.8178	65.0287	29.9998	9.3934
Mandrill	N=4	698.4679	19.6893	714.504	19.5908	0.5028
	6	683.8821	19.7810	687.086	19.7607	0.1027
	8	679.5712	19.8085	742.750	19.4224	1.9879
	10	680.0534	19.8054	971.479	18.2565	8.4841
	12	682.4979	19.7898	780.961	19.2045	3.0477
Peppers	N=4	35.8618	32.5845	41.0345	31.9993	1.8288
	6	20.2854	35.0590	22.3782	34.6325	1.2315
	8	17.7156	35.6472	19.4359	35.2447	1.1420
	10	20.9640	34.9161	32.4916	33.0131	5.7644
	12	24.9685	34.1569	20.2369	35.0694	-2.602

Some results are now given in Figs. (4.7)-(4.9) showing the frequency spectrums of the input image and the Low-Low subband reconstructed image and the relationship between the activity level of the image and the performance. Images with varying spatial activity level were chosen and their respective spectra and the autocorrelation matrices are plotted. The spatial frequencies and the spectral activity measures of the three images considered are (4.76) and (215.48) for the Spine image, (28.47) and (379.8) for the Barb image and (76.53) and (18.65) for the Straw image, respectively. As mentioned earlier, higher the SF value, the more high frequency information in the form of edges is present in the spatial domain of the image and the lesser its predictability, which can be easily seen in the images shown. The Straw image has the most edge information compared to the other two images. Higher the SAM value, higher the high frequency information in the image's Fourier domain, which is seen in the amplitude spectra of the autocorrelation matrices of the images. The spectrum of the Barb image consists of the highest edge information as compared to the Straw and the Spine image. The Straw image has the smoothest spectrum of the three images since it has the lowest SAM value. The SF value determines the reconstruction error in terms of the PSNR values, as is seen in Table 4.4. The Spine image gives the highest PSNR compared to the Barb and the Straw images. The optimized PSNR of the three reconstructed images are in the range of 45, 25 and 19 dB respectively.



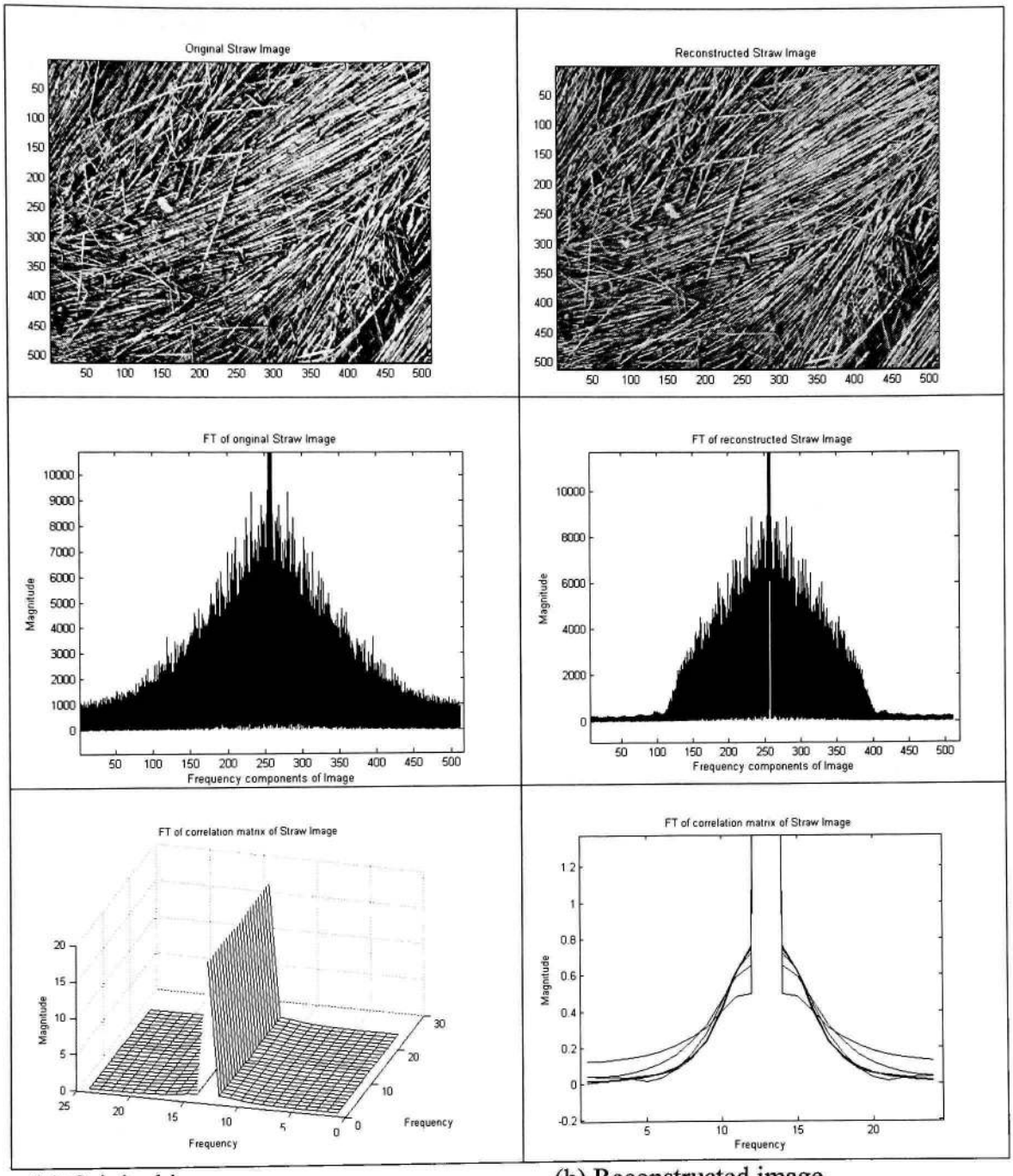
(a) Original image
(b) Reconstructed image
(c) FFT of the original image
(d) FFT of the reconstructed image
(e) 3-D FFT of the autocorrelation matrix
(f) 2-D FFT of the autocorrelation matrix
(in order of appearance)

Fig 4.7: Analysis of the Spine image



(a) Original image
(b) Reconstructed image
(c) FFT of the original image
(d) FFT of the reconstructed image
(e) 3-D FFT of the autocorrelation matrix
(f) 2-D FFT of the autocorrelation matrix
(in order of appearance)

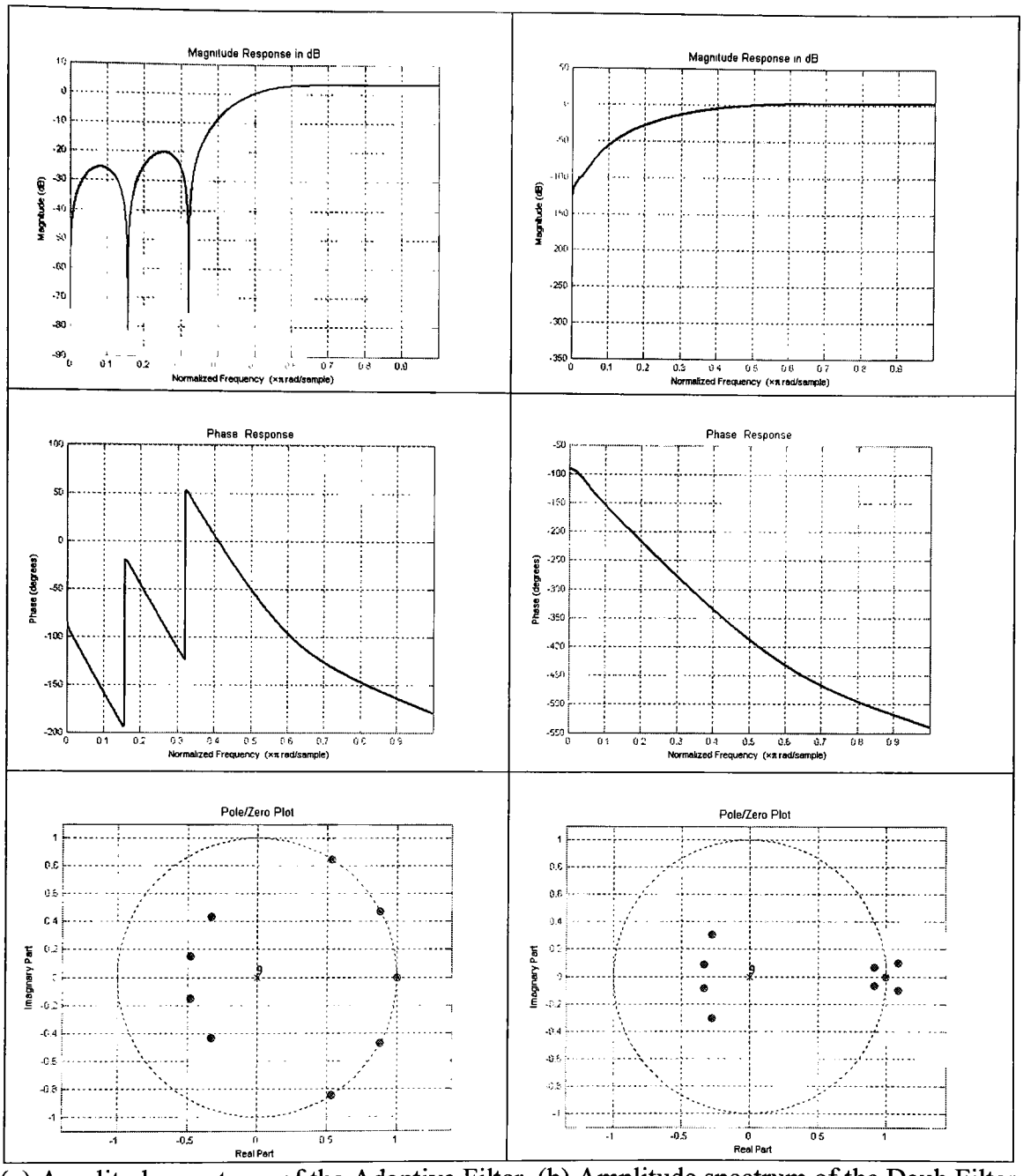
Fig 4.8: Analysis of the Barb image



(a) Original image
(b) Reconstructed image
(c) FFT of the original image
(d) FFT of the reconstructed image
(e) 3-D FFT of the autocorrelation matrix
(f) 2-D FFT of the autocorrelation matrix
(in order of appearance)

Fig 4.9: Analysis of the Straw image

Now a comparison of the filters obtained with the standard Daubechies filters is made. Fig. 4.10 (a)-(f) shows the results of the comparison made between the optimized high pass filter for the Straw image for $N = 10$ and the Daubechies high pass filter. As mentioned earlier, Daubechies filters are known to have the maximum number of zeros at $Z = 1$ for the high-pass filter. It can be seen from the Fig. 4.10(f) that there is one zero at $Z = 1$ in the pole-zero plot of the Daubechies filter. Similarly a zero is also present in the pole-zero plot of the adaptive filter. But Table 4.4 confirms that the adaptive filter performs better than the Daub filters for all filter lengths. The amplitude spectrums of the filters show that the adaptive filter has a better cut-off than the Daubechies filter. This improvement in the performance of the Adaptive filter is expected since it redistributes the energy of the input image according to the image statistics, whereas the Daubechies filter has a fixed set of filters for any input. Table 4.4 shows the results of all the tested images for various filter lengths.



(a) Amplitude spectrum of the Adaptive Filter (b) Amplitude spectrum of the Daub Filter
(c) Phase response of the Adaptive filter (d) Phase response of the Daub filter
(e) Pole-zero plot of the Adaptive filter (f) Pole-zero plot of the Daub filter
(in order of appearance)

Fig. 4.10: Comparison of the Adaptive filter and the Daub high-pass filters for the Straw image

Table 4.4: Results for various filter lengths applied to different images

Image	SF	SAM $= \frac{1}{SFM}$	Filter	N = 4			N = 6			N = 8		
				MSE	PSNR dB	TCG	MSE	PSNR dB	TCG	MSE	PSNR dB	TCG
Spine	4.76	215.48	Adap	2.00	45.11	25.59	2.02	45.06	27.77	2.16	44.78	28.16
			Daub	2.00	45.11	25.59	2.26	44.59	27.32	2.16	44.77	28.14
House	9.14	1170.22	Adap	5.93	40.40	37.82	3.18	43.10	43.02	3.01	43.34	45.36
			Daub	5.93	40.40	37.82	3.22	43.06	42.95	3.03	43.32	45.33
Lena	13.27	924.28	Adap	22.82	34.54	15.42	17.66	35.67	17.08	16.13	36.06	17.80
			Daub	22.82	34.54	15.42	18.91	35.36	16.65	17.31	35.74	17.28
Peppers	14.41	263.62	Adap	35.86	32.59	14.26	20.28	35.06	15.16	17.71	35.65	15.50
			Daub	35.86	32.59	14.26	21.68	34.77	14.97	17.73	35.64	15.29
Couple	15.14	347.90	Adap	40.0	32.10	12.90	33.62	32.87	14.05	30.72	33.26	14.48
			Daub	40.0	32.10	12.90	35.77	32.59	13.64	32.71	32.98	14.06
Goldhill	15.42	195.77	Adap	41.12	31.99	12.43	36.39	32.53	13.13	35.26	32.66	13.37
			Daub	41.12	31.99	12.43	37.69	32.36	12.94	36.33	32.52	13.18
Aerial	16.75	257.80	Adap	21.52	34.80	23.40	12.71	38.39	33.82	9.80	39.48	40.39
			Daub	21.52	34.80	23.40	12.72	37.08	30.30	9.80	38.21	35.36
Barb	28.47	379.80	Adap	168.6	25.86	4.89	199.6	25.14	5.56	190.2	25.34	5.58
			Daub	168.6	25.86	4.89	168.4	25.86	4.84	168.8	25.85	4.81
Camera -man	29.30	120.0	Adap	147.0	26.46	6.50	135.3	26.82	6.82	133.1	26.89	6.93
			Daub	147.0	26.46	6.50	140.2	26.66	6.71	138.6	26.71	6.82
Mandrill	62.26	26.99	Adap	698.4	19.69	3.38	683.8	19.79	3.40	679.5	19.81	3.41
			Daub	698.4	19.69	3.38	687.2	19.76	3.39	682.0	19.79	3.39
Straw	76.53	18.65	Adap	954.8	18.34	3.18	772.4	19.26	3.49	724.2	19.54	3.58
			Daub	954.8	18.34	3.18	842.0	18.87	3.35	785.3	19.18	3.45

(NOTE: Adap:- Adaptive Filter; Daub:- Daubechies Filter)

Table 4.4: (continued)

Image	N = 10			N = 12			N = 14			N = 16		
	MSE	PSNR dB	TCG	MSE	PSNR dB	TCG	MSE	PSNR dB	TCG	MSE	PSNR dB	TCG
Spine	1.76	45.67	28.65	1.31	46.95	28.97	1.05	47.90	29.19	1.09	47.73	29.36
	1.76	45.66	28.64	1.31	46.95	28.95	1.05	47.89	29.18	1.09	47.72	29.35
House	3.59	42.57	46.68	3.95	42.16	47.52	3.67	42.48	48.08	2.93	43.46	48.48
	3.60	42.56	46.66	3.96	42.15	47.50	3.68	42.47	48.06	2.94	43.45	48.46
Lena	15.48	36.23	18.20	15.02	36.36	18.41	14.65	36.47	18.59	14.48	36.53	18.71
	16.55	35.98	18.19	16.07	36.06	17.92	15.64	36.18	18.10	15.28	36.29	18.24
Peppers	20.96	34.92	15.62	24.96	34.16	15.79	26.03	33.98	15.84	27.40	33.76	15.72
	21.08	34.89	15.46	25.81	34.01	15.58	28.55	33.57	15.65	27.44	33.74	15.71
Couple	28.78	33.54	14.74	27.91	33.68	14.94	27.85	33.68	15.04	28.03	33.66	15.16
	30.42	33.30	14.32	29.27	33.46	14.51	29.19	33.47	14.64	29.63	33.41	14.75
Goldhill	34.60	32.74	13.49	33.99	32.82	13.57	33.69	32.86	13.60	33.30	32.91	13.64
	35.60	32.61	13.32	35.01	32.68	13.40	34.46	32.75	13.46	34.02	32.81	13.51
Aerial	7.21	39.55	44.68	7.44	39.42	46.58	6.80	39.81	48.0	5.30	40.89	49.9
	9.10	38.53	39.09	8.85	38.66	41.86	7.86	39.17	43.95	6.24	40.17	45.56
Barb	173.2	25.74	5.61	182.2	25.52	5.62	214.2	24.83	5.74	206.4	24.99	5.77
	169.3	25.84	4.79	169.7	25.83	4.77	170.1	25.82	4.76	170.5	25.81	4.75
Camera -man	130.8	26.97	7.0	129.8	27.0	7.0	127.6	27.10	7.01	125.3	27.15	7.01
	136.9	26.76	6.89	133.8	26.86	6.93	130.4	26397	6.97	127.9	27.06	6.99
Mandrill	680.0	19.81	682.4	3.41	19.79	683.5	3.41	19.78	682.6	3.42	19.79	3.42
	682.6	19.78	3.40	685.5	19.77	3.40	686.8	19.76	3.40	685.6	19.77	3.41
Straw	689.5	19.75	3.65	680.4	19.81	3.68	681.5	19.80	3.69	669.5	19.88	3.72
	754.7	19.35	3.51	737.0	19.45	3.56	723.8	19.53	3.59	712.5	19.60	3.61

Table 4.4: (continued)

Image	N = 18			N = 20			N = 22			N = 24		
	MSE	PSNR dB	TCG	MSE	PSNR dB	TCG	MSE	PSNR dB	TCG	MSE	PSNR dB	TCG
Spine	1.35	46.83	29.49	1.61	46.03	29.60	1.71	45.78	29.66	1.64	46.03	29.75
	1.35	46.82	29.48	1.61	46.03	29.58	1.72	45.77	29.67	1.60	46.07	29.74
House	2.09	44.93	48.77	1.55	46.23	49.0	1.63	46.0	49.18	2.22	44.66	49.32
	2.09	44.92	48.76	1.55	46.21	48.99	1.63	45.99	49.17	2.22	44.64	49.31
Lena	14.33	36.57	18.80	14.17	36.62	18.88	14.09	36.64	18.88	13.95	36.69	18.92
	15.01	36.36	18.35	14.84	36.41	18.43	14.75	36.44	18.50	14.69	36.45	18.56
Peppers	22.81	34.54	15.77	18.25	35.51	15.80	17.24	35.77	15.83	19.85	35.16	15.86
	22.83	34.54	15.77	18.26	35.51	15.80	17.24	35.77	15.83	19.86	35.15	15.86
Couple	28.10	33.65	15.24	27.79	33.70	15.31	27.30	33.77	15.35	36.85	33.85	15.35
	29.94	33.37	14.83	29.68	33.40	14.90	28.89	33.52	14.94	27.97	33.66	15.0
Goldhill	33.28	32.91	13.64	33.77	32.85	16.58	33.87	32.84	13.60	33.92	32.83	13.61
	33.79	32.84	13.54	33.79	32.84	13.56	33.89	32.83	13.59	33.93	32.82	13.60
Aerial	4.78	41.34	46.91	4.16	41.94	47.87	4.43	41.67	48.65	5.20	40.97	49.28
	4.80	41.31	46.80	4.18	41.91	47.79	4.44	41.65	48.58	5.21	40.95	49.22
Barb	187.5	25.40	5.61	180.5	25.57	5.67	192.1	25.30	5.81	197.1	25.19	5.87
	171.0	25.79	4.74	171.6	25.78	4.73	172.2	25.77	4.72	172.6	25.75	4.71
Camera -man	123.3	27.22	7.10	122.4	27.26	7.12	122.9	27.24	7.13	124.0	27.20	7.14
	126.6	27.10	7.01	126.5	27.10	7.02	127.4	27.07	7.04	128.3	27.04	7.05
Mandrill	680.7	19.81	3.42	678.1	19.82	3.42	676.0	19.84	3.42	675.9	19.84	3.42
	682.6	19.78	3.41	679.3	19.81	3.41	677.2	19.82	3.41	677.5	19.82	3.41
Straw	660.2	19.94	3.73	652.1	19.99	3.73	647.8	20.01	3.74	645.8	20.02	3.75
	701.9	19.66	3.63	691.6	19.73	3.64	684.0	19.78	3.66	681.2	19.79	3.67

Summarizing the test results for the reconstruction from the Low-Low subband, it can be stated that optimized filters are found that produce better MSE and PSNR than the Daubechies Maxflat filters. However, there is a chance that they will perform worse than the Maxflat filters in some cases (one out of twelve test cases). The absolute values of PSNR, in general, are in excess of 30 dB (except for very high SF measure images like the Mandrill image). This makes the adaptively designed filters acceptable to be used in a complete encoder scheme.

Here is a summary of the main points of this chapter. An adaptive two-channel PR-CQF bank design problem was considered. PR-CQF constraints were introduced to keep aliasing at a low level while minimizing the energy in the higher frequency subbands. The optimization technique that would solve the given problem was implemented. The tests conducted have shown that the minimization of energy in the higher subbands leads to minimization of aliasing. Various relationships between the performance of the images and the statistics of the images are developed and shown. The tests have shown that the minimum of the energy in the higher subband under CQF constraints is sufficient to provide acceptable distortion of the images reconstructed from the Low-Low subband only. The results are compared with the well-known Daubechies filters and an improved performance is observed. It is of special interest to note that in a few cases, even though the optimization always find a better solution compared to the starting point, the Daubechies filter give a better PSNR value than the adaptive filter. This inference can be confirmed by the values of the energy term of the high-pass filter, tabulated for a few images in Table 4.5. The table gives the values of the energy and the corresponding PSNR for 4 images as shown. It can be noted, in the case of the Barb image, that even though the PSNR values of the Daubechies filter are better, the energy values of the Adaptive filter are yet lower. The energy values of the Adaptive filter are better in all of the cases, thus confirming that the optimization indeed finds a better solution. It is also to be noted that the PSNR values may not be the best way to determine the quality of the results obtained, as seen in the case of the Barb image. Nevertheless the algorithm finds a better solution in terms of minimization

Table 4.5: Comparison of the high-pass channel energy values of the Daubechies and the Adaptive filters for various images

Image	Filter order	Daub Filter		Adaptive Filter		% improvement in energy value
		$h_1^T R_{xx} h_1$	PSNR (dB)	$h_1^T R_{xx} h_1$	PSNR (dB)	
Lena	N=4	0.0021	34.5475	0.0021	34.5475	0
	6	0.0018	35.3650	0.0017	35.6610	5.5556
	8	0.0017	35.7481	0.0016	36.0529	5.8824
	10	0.0016	35.9406	0.0015	36.2311	6.2500
	12	0.0016	36.0698	0.0015	36.3622	6.2500
Goldhill	N=4	0.0032	31.9895	0.0032	31.9895	0
	6	0.0030	32.3686	0.0029	32.5202	3.3333
	8	0.0029	32.5285	0.0028	32.6578	3.4483
	10	0.0028	32.6166	0.0028	32.7397	0
	12	0.0028	32.6885	0.0027	32.8178	3.5714
Barb	N=4	0.0212	25.8613	0.0212	25.8613	0
	6	0.0215	25.8690	0.0164	25.1303	23.7209
	8	0.0218	25.8579	0.0162	25.3398	25.6881
	10	0.0220	25.8433	0.0160	25.7444	27.2727
	12	0.0222	25.8318	0.0159	25.5231	28.3784
Mandrill	N=4	0.0449	19.6893	0.0449	19.6893	0
	6	0.0444	19.7602	0.0443	19.7810	0.2252
	8	0.0443	19.7930	0.0441	19.8085	0.4515
	10	0.0441	19.7891	0.0440	19.8054	0.2268
	12	0.0441	19.7706	0.0439	19.7898	0.4535

of the output high-pass channel energy.

CHAPTER 5
CONCLUSIONS

The present work has contributed to the adaptive filter bank design for image processing applications. The design methods for optimal perfect reconstruction two-channel filter banks for image size reduction have been investigated. The perfect reconstruction conditions are derived from the Nyquist(2) criteria. When the aliasing energy component, which is the energy through the high-pass filter channel, was kept at an acceptably low value by introducing perfect reconstruction constraints into the optimization problem adaptively, the filter banks provide lower distortion than non-adaptive filters, like the Daubechies filters, when the images were reconstructed from their LL subbands. This work has primarily focused on investigating the advantages and limitations of the adaptive wavelet filter when compared to standard filters. Various measures for the image statistics are discussed and the results analyzed.

Summarizing the results obtained, we conclude that:

1. Adaptive filters were successfully found using various starting points.
2. For $N = 2$, the only solution to the optimization problem is that of the Haar filter.
3. For $N = 4$, the adaptive filter was the same as the Daubechies filter. For this case the algorithm did not find a better solution and it converged to the starting point.
4. For $N > 4$, the optimization performed better than the Daubechies filter in more than 90% of cases when the PSNR values were compared, it gave a lower energy value of the high-pass channel in all 100% of cases compared to the Daubechies filter.

Future scope in improving and applying this work could be:

1. A few other optimality criteria can also be included such as constrained non-linear phase response of the filter and uncorrelatedness of subband signals. These constraints can improve the performance of the optimization. Although this work has been done in the area of image encoding, the adaptive filter design algorithms

presented in this work can find application in other general areas of image/signal processing.

2. The optimization technique can be used for a wider range of applications, like possibly in Edge detection by concentrating the energy in the High-High subbands

REFERENCES

[1] Vadim Kustov, "Adaptive Wavelet Filter Design for Digital Signal Processing Systems," Ph.D. Dissertation, Texas Tech University, December 2000.

[2] A N Akansu, M J T Smith, *Subband and Wavelet Transforms: Design and Application*, Kluwer Academic Publications, Norwell, MA, 1996.

[3] F Z Qureshi, "Image Compression Using Wavelet Transform," *Computational Vision Project*, University of Toronto, http://citeseer.nj.nec.com/425270.html.

[4] R Polikar, "The Wavelet Tutorial," Rowan University http://engineering.rowan.edu/~polikar/WAVELETS/WTtutorial.html.

[5] H Caglar, Y Liu, A N Akansu, "Statistically Optimized PR-QMF Design," *Visual Communications and Image Processing*, Vol. 1605, SPIE, 1991.

[6] Stephen Boyd, Lieven Vandenberghe, *Convex Optimization*, December 2001. http://www.ee.ucla.edu/ee236b/ee236b.html#book.

[7] R J Vanderbei, *Linear Programming: Foundations and Extensions*, Kluwer Academic Publishers, II Ed., Norwell, MA, June 2001.

[8] N Karmarkar, "A New Polynomial-time Algortihm for Linear Programming," *Combinatorica*, pp.373-395, 4 (1984).

[9] P A Jensen, J F Bard, *Interior Point Methods*, Operations Research Models and Methods, John Wiley and Sons, New York, 2003.

[10] S J Wright, *Primal-Dual Interior Point Methods*, SIAM, 1997.

[11] D M Gay, "Pictures of Karmarkar's Linear Programming Algorithm," Computing Science Technical Report, No.136, AT&T Bell Labs, 1987.

[12] F S Hillier, *Introduction to Mathematical Programming*, Mc-Graw Hill, II Edition, New York, 1995.

[13] I Akrotirianakis and B Rustem, "A Primal-Dual Interior Point Algorithm with an Exact And Differentiable Merit Function for General Non-Linear Programming Problems," Technical Report 98-09, http://citeseer.nj.nec.com/akrotirianakis99primaldual.html.

[14] T F Coleman, M A Branch, A Grace, "Matlab Optimization Toolbox User's Guide," Version 2, MathWorks Inc., Natick, MA.

[15] Broyden, C.G., "The Convergence of a Class of Double-rank Minimization Algorithms," *J. Inst. Maths. Applics.*, Vol. 6, pp 76-90, 1970.

[16] Fletcher, R, "A New Approach to Variable Metric Algorithms," *Computer Journal*, Vol. 13, pp 317-322, 1970.

[17] Goldfarb, D, "A Family of Variable Metric Updates Derived by Variational Means," *Mathematics of Computing*, Vol. 24, pp 23-26, 1970.

[18] Shanno, D F, "Conditioning of Quasi-Newton Methods for Function Minimization," *Mathematics of Computing*, Vol. 24, pp 647-656, 1970.

[19] Shapiro J, "Embedded Image Coding Using Zerotrees of Wavelet Coefficients," *IEEE Transactions on Signal Processing,* Vol. 41, No. 12, December 1993.

[20] M Antonini, M Barlaud, P Mathieu, I Daubechies, "Image Coding Using Wavelet Transform," *IEEE Transactions on Image Processing*, 1(2), pp 205-220 (1992).

[21] Shuyu Yang, "Performance Analysis from Rate Distortion Theory of Wavelet Domain Vector Quantization Encoding" Ph.D. Dissertation, Texas Tech University, May 2002.

[22] Shuyu Yang, Sunanda Mitra, Enrique Corona, Brian Nutter, and D. J. Lee, "Multi-level Wavelet Feature Statistics for Efficient Retrieval, Transmission, and Display of Medical Images by Hybrid Encoding," *EUROSIP Journal in Applied Signal Processing*, 2002.

[23] Sunanda Mitra, Shuyu Yang, Roopesh Kumar, and Brian Nutter, "An Optimized Hybrid Vector Quantization for Efficient Source Encoding", Invited paper at the *45th MWSCAS*, 4-7 August, Tulsa, Oklahoma, 2002.

[24] S. Mitra, and Shu-Yu Yang, "High Fidelity Adaptive Vector Quantization at Very Low Bit Rates for Progressive Transmission of Radiographic Images," *Journal of Electronic Imaging*, Vol.8, No. 1, pp 23-35, 1999.

[25] M Antonini, T Gaidon, P Mathieu, M Barlaud, "Wavelet Transform and Image Coding," *Wavelets in Image Communication*, Advances in Image Communication, Vol. 5, Elsevier, 1994.

[26] M K Mandal, S Panchanathan and T Aboulnasr, "Choice of Wavelets for Image Compression," Lecture Notes in Computer Science, Vol. 1133, Springer Verlag, New York, 1996.

[27] N S Jayant, P Noll, *Digital Coding of Waveforms: Principles and Applications to Speech and Video*, Prentice Hall, Englewood Cliffs, NJ, 1984.

[28] Mislav Grgic, Nikola Sprljan, Branka Zovko-Cihlar, "Image Analyser-Educational Tool," University of Zagreb, Croatia.

[29] A Kumar, J Kusuma and K Ramachandran, "Lecture 10: Wavelet Packet Trees," http://www-inst.eecs.berkeley.edu/~ee225a/spring2002/lectures/ee225a_lec9.ps, Feb 2002, University of Berkeley.

[30] P Moulin and M Mihcak, "Theory and Design of Signal-Adapted FIR Paraunitary Filter Banks," *IEEE Transactions on Signal Processing,* Vol. 46, No. 4, August 1998.

APPENDIX A

OBJECTIVE FUNCTION DERIVATION

Proof of Eq.(3.27)

Eq.(3.26) is

$$f(a,b,c,d,e,f) = (a^2 + b^2 + c^2 + d^2 + e^2 + f^2) + 2r_{16}[af] + 2r_{15}[ae + bf] + 2r_{14}[ad + be + cf] + 2r_{13}[ac + bd + ce + df] + 2r_{12}[ab + bc + cd + de + ef] \quad (A.1)$$

From Eqs.(3.18),(3.21),(3.22),(3.23) we have

$$f(a,b,c,d,e,f) = 1 + 2r_{16}[af] + 2r_{14}[ad + be + cf] + 2r_{12}[ab + bc + cd + de + ef] \quad (A.2)$$

Now consider,

$$(a+b+c+d+e+f)^2 = (a^2 + b^2 + c^2 + d^2 + e^2 + f^2) + 2[ac + ae + bd + bf + ce + df] + 2[ab + ad + af + bc + be + cd + cf + de + ef] = 0 \quad (A.3)$$

since $(a+b+c+d+e+f) = 0$ from Eq.(3.21)

Eq.(A.3) simplifies as

$$ab + bc + cd + de + ef = -1/2 - ad - af - be - cf \quad (A.4)$$

Substituting Eq.(A.4) in (A.2), we get Eq.(3.27)

$$f(a,b,c,d,e,f) = (1 - r_{12}) - 2(r_{12} - r_{14})(ad + be + cf) - 2(r_{12} - r_{16})af \quad (A.5)$$

APPENDIX B

DEFINITION OF SOME IMAGE DEPENDENT STATISTICS

Spatial Frequecny (SF) is defined as [28],

$$SF = \sqrt{P^2 + Q^2} \tag{B.1}$$

where, $P = \sqrt{\frac{1}{RC}\sum_{i=1}^{R}\sum_{j=2}^{C}[I(i,j) - I(i,j-1)]^2}$

$$Q = \sqrt{\frac{1}{RC}\sum_{j=1}^{C}\sum_{i=2}^{R}[I(i,j) - I(i-1,j)]^2}$$

The Joint Time-Frequency Localization (TFL) is defined as [26],

$$TFL = \sigma_n \sigma_\omega \tag{B.2}$$

where, $\sigma_n^2 = \frac{1}{E}\sum_n (n - \bar{n})^2 |h[n]|^2$

$$\sigma_\omega = \frac{1}{E}\frac{1}{2\pi}\int_{-\infty}^{\infty}(\omega - \bar{\omega})^2 |H(\omega)|^2 d\omega$$

where,

$\bar{n} = \frac{1}{E}\sum_n n|h[n]|^2$, the centre of mass of the sequence

$\bar{\omega} = \frac{1}{2\pi E}\int_{-\infty}^{\infty}\omega|H(e^{j\omega})|^2 d\omega$, the centre of mass in frequency domain

$E = \sum_n |h[n]|^2 = \frac{1}{2\pi}\int_{-\infty}^{\infty}|H(\omega)|^2 d\omega$, is the total energy of the sequence

The TFL of any wavelet is lower bounded by 0.5.

The Transform Coding Gain (TCG) for an orthogonal transform is defined as the ratio of the arithmetic mean of the energy, σ_j^2, in each subband to the geometric mean [3],

$$TCG = \frac{\frac{1}{N}\sum_{j=0}^{N-1}\sigma_j^2}{\left[\prod_{j=1}^{N-1}\sigma_j^2\right]^{1/N}} \tag{B.3}$$

For a 2-channel PR-QMF Filter bank, Eq.(B.3) reduces to

$$TCG = \frac{\sigma_x^2}{\left[\sigma_L^2\sigma_H^2\right]^{1/2}}, \tag{B.4}$$

where σ_x^2, σ_L^2 and σ_H^2 are the variances of the input, the low-pass band and the high-pass band respectively.

The input variance in such a case is expressed as,

$$\sigma_x^2 = 1/2(\sigma_L^2 + \sigma_H^2). \tag{B.5}$$

The output energy of the low-pass and the high-pass filters are defined as,

$$\sigma_L^2 = h_0 R_{xx} h_0^T, \text{ and}$$

$$\sigma_H^2 = h_1 R_{xx} h_1^T.$$

PERMISSION TO COPY

In presenting this thesis in partial fulfillment of the requirements for a master's degree at Texas Tech University or Texas Tech University Health Sciences Center, I agree that the Library and my major department shall make it freely available for research purposes. Permission to copy this thesis for scholarly purposes may be granted by the Director of the Library or my major professor. It is understood that any copying or publication of this thesis for financial gain shall not be allowed without my further written permission and that any user may be liable for copyright infringement.

Agree (Permission is granted.)

_____ _____

Student Signature Date

Disagree (Permission is not granted.)

_____ _____

Student Signature Date