A Novel Enhanced Compression Scheme for Digital Images

and its Application for Medical Imaging

Lakshmi Sugavaneswaran

A Thesis

in

The Department

of

Electrical and Computer Engineering

Presented in Partial Fulfillment of the Requirements
for the Degree of Master of Applied Science at
Concordia University,
Montreal, Quebec, Canada

July 2007

Canada

ABSTRACT

A Novel Enhanced Compression Scheme for Digital Images and its Application for Medical Imaging

Lakshmi Sugavaneswaran

With the advent of digital technology, the amount of digital medical images generated is increasing at an exponential rate. When used in interactive telemedicine applications, these raw images need a huge storage unit, thereby necessitating the use of high-bandwidth medium for transmission. In order to reduce the cost of transmission and enable archiving, especially for medical applications, image compression is performed. Most of the existing research in compression is focused on minimizing mathematically tractable, easy to measure, distortion metrics like the peak signal-to-noise ratio (PSNR) for gray scale and root mean square error (RMSE) for color coded images. These metrics often poorly correlate with the visual quality of the image as the human visual sensitivity characteristics are not incorporated in these evaluation techniques. For digital images, especially for medical images, perceptual quality of image is important along with compression efficiency since diagnosis is done based on the visual attributes of the reconstructed image.

This thesis is concerned with the problem associated with obtaining an enhancement in the quality of the image with lossy compression schemes. An enhanced image compression scheme using the nearly-orthogonal wavelet to achieve an enhancement in the perceptual quality of the image is proposed. This approach differs from the

conventional design scheme in that it incorporates the human visual system (HVS) characteristics directly in each sub-band of the wavelet decomposed image. The enhancement in the visual quality of the reconstructed image is achieved by using the newly derived contrast sensitivity function band peak-average (CSF-BPA) mask at each decomposition level during wavelet transformation phase in the compression system. The nearly-orthogonal wavelet is used to achieve improved compression performance for low and medium bit rate applications. The proposed scheme is further extended for perceptual medical image compression application for images such as x-ray, MRI and ultrasound scan. The dependency of the calculated PSNR to the observed visual quality has been addressed for some of the standard images as well as for medical images. The presence of additive white Gaussian noise (AWGN), which is the most common source for image artifacts, is considered during the evaluation of the performance of the scheme. The effectiveness of the proposed scheme is demonstrated and compared with that of the conventional scheme through extensive quantitative and qualitative assessments. Finally, a multiplier-less design for the implementation of the BNC 17/11 wavelet filter in a FPGA or ASIC hardware is proposed.

ACKNOWLEDGEMENTS

been pivotal to the success of this work. Although I may not have shown my appreciation, for their unconditional support and encouragement, I hope they are proud of me. Many thanks are due to my close family and friends, especially to Sivaramakrishna, Jagannathan, Seshanarayan, Rajeev, Karthik, Gayatri, Hamsa, Sujata, Ramaa, Karuna, Deepak and Govind for their constant help and encouragement.

I dedicate this work to my

loving family

Contents

List of Figures

List of Tables

List of Symbols

ω :   Radial frequency

t   :   Time

τ :   Translation

s :   Scale

$\gamma(\tau, s)$:   Continuous wavelet transform of a function

W_j :   Details subspace

V_j :   Signal subspace

$\phi(k)$:   Scaling function

$\psi(k)$:   Wavelet function

$\tilde{\phi}(k)$:   Dual of $\phi(k)$

$\tilde{\psi}(k)$:   Dual of $\psi(k)$

ϕ_{ij} :   Scaling function with the ith position in x-axis with a scale of 2^j

ψ_{ij} :   Wavelet function with the ith position in x-axis with a scale of 2^j

$CSF(f)$:   Contrast sensitivity function at frequency 'f' in cycles per degree

E_{ms} :   Mean square error

E_{rms} :   Root mean square error

$e(x,y)$:   Error component

SNR_{ms} :   SNR corresponding to mean square error

SNR_{rms} :   SNR corresponding to root mean square error

$PSNR$:   Peak signal to noise ratio

L	:	Luminance constant
f_x	:	Spatial frequency in cycles per degree define in the x-direction
f_y	:	Spatial frequency in cycles per degree define in the y-direction
f_c	:	Centre frequency in cycles per degree
f_s	:	Surround frequency in cycles per degree
f	:	Spatial frequency in cycles per degree
f_{max}	:	Maximum frequency
M, N	:	Number of rows and columns of an image
a_i	:	Low-pass or Approximation coefficients at level 'i'
d_i	:	High-pass or Detail coefficients at level 'i'
a_i^k	:	a_i defined at stage 'k'
d_i^k	:	d_i defined at stage 'k
L_n, H_n	:	Low and High frequency bands of an n-level DWT
x, y, i, j, k, l, n, m	:	Integers
p_n, q_n	:	Mask coefficients
a, b, K, K_0, K_1	:	Constants
δ_{kl}	:	Delta function
h_0, h_1	:	Analysis filters
g_0, g_1	:	Synthesis filters
$h[n], f[n]$	:	Impulse reponse of the low-pass filters
$g[n], j[n]$	:	Impulse reponse of the high-pass filters
$H(\omega), F(\omega)$	:	Transfer function of low-pass filters
$G(\omega), J(\omega)$	:	Transfer function of high-pass filters

$H(z), F(z)$	:	Z-transform of low-pass filters
$G(z), J(z)$	:	Z-transform of high-pass filters
$H_e(z)$	:	Even component of $H(z)$
$H_o(z)$	:	Odd component of $H(z)$
$G_e(z)$	:	Even component of $G(z)$
$G_o(z)$	:	Odd component of $G(z)$
$F_e(z)$	:	Even component of $F(z)$
$F_o(z)$	:	Odd component of $F(z)$
$J_e(z)$	:	Even component of $J(z)$
$J_o(z)$	:	Odd component of $J(z)$
$H_p(z)$	:	Analysis polyphase matrix
$F_p(z)$	:	Synthesis polyphase matrix
$X_e(z), X_o(z)$	:	Even and odd components of $X(z)$
$s_i(z), t_i(z)$	:	Update and predict lifting coefficients
$\alpha_i, \beta_i, \gamma_i, \delta_i$	:	Lifting coefficients
ζ	:	Scaling factor
P_n, U_n	:	Predict and Update sequences
$x[n], y[n]$	:	Input and output signals
In_1, In_2	:	Input lines
Out_1, Out_2	:	Output lines
C_ψ	:	Admissability constant

List of Abbreviations

AFB   :   analysis filter-bank

ASIC   :   application-specific integrated circuit

AWGN   :   additive white Gaussian noise

BNC   :   biorthogonal nearly coiflet

BPA   :   band peak-average

CDF   :   Cohen, Daubechies and Feaveau

CMYK   :   cyan, magenta, yellow and key (black)

CSF   :   contrast sensitivity function

CT   :   contrast threshold

CWT   :   continuous wavelet transform

dB   :   Decibel

DCT   :   discrete cosine transform

DWT   :   discrete wavelet transform

EBCOT   :   embedded block coding with optimized truncation

EZW   :   embedded zero-tree wavelet

FIR   :   finite impulse response

FPGA   :   field programmable gate array

HP   :   high pass

HPF   :   high pass filter

IGS   :   improved gray scale

JND   :   just-noticeable distortion

KL	:	Karhunen-Loeve
LOT	:	lapped orthogonal transform
LP	:	low pass
LPF	:	low pass filter
LWT	:	lazy wavelet transform
LZC	:	Lempel-Ziv compressor
MOS	:	mean opinion score
MRA	:	multi-resolution analysis
MRI	:	magnetic resonance imaging
MSE	:	mean square error
PIC	:	perceptual image coder
PSNR	:	peak signal-to-noise ratio
QMF	:	quadrature mirror filter
RMSE	:	root mean square error
SFB	:	synthesis filter bank
SNR	:	signal-to-noise ratio
SNR_{ms}	:	mean-square SNR
SNR_{rms}	:	root mean-square SNR
SPIHT	:	set partitioning in hierarchical trees
WPS	:	whole-point symmetry

Chapter 1

Introduction

1.1 Introduction and Motivation

Telemedicine is beneficial for populations living in remote regions. With the advent of interactive applications, it is currently being used for online consultation, monitoring and diagnosis services offered by medical experts. The implementation of such a system requires an image sending station, a transmission channel, and an image review station. Such a system needs a huge storage unit and a very high-bandwidth medium to transmit the raw medical data and hence, increases the cost of transmission. In order to provide a cost-effective solution, transmission over the low-bandwidth medium is considered a good alternative. This helps the medical images such as ultrasound, x-ray, CT scan, and fMRI scan in applications like teleradiology, and video conferencing. The main advantages of digitally represented images are that they enable electronic archiving and efficient transmission over the network, and provide useful manipulation of diagnostic information. Hence, the quality of storage and transmission of medical images has a significant potential towards patient care [1, 2]. In spite of offering perfect reconstruction, lossless compression methods provide a maximum compression ratio of only 2:1, and even if transmitted through a high speed DSL line, it is very time consuming. For instance, a 512 x 512 image takes about 7 hours to be transmitted. Table 1.1 clearly illustrates the storage space and transmission time constraints to be met for different uncompressed multimedia data types.

Table 1-1 Uncompressed storage space and transmission time requirements for different multimedia data types [1].

Data	Duration/ Size	Uncompressed size (B for bytes)	Transmission time
A page of text	11" x 8.5"	4-8 KB	1.1 - 2.2 sec
Speech	10 sec	80 KB	22.2 sec
Grayscale image	512 x 512	262 KB	1 min 13 sec
RGB image	512 x 512	786 KB	3 min 39 sec
Medical image	2048 x 1680	5.16 MB	23 min 54 sec
Video	640 x 480, 1 min (30 frames/sec)	1.66 GB	5 days 8 hrs

Hence, in order to achieve an improvement in the compression performance, the application of certain high quality lossy compression methods would be acceptable, if nearly-lossless image reconstruction without much loss in the diagnostic content is ensured [3]. Consequently, most of the modern image compression strategies exploit the fact that the human visual system (HVS), as the ultimate consumer of an image, is an imperfect sensor. For these compression schemes, an exact bit-for-bit representation is unnecessary and the data can be coded in a non-invertible or lossy fashion. Based on the HVS characteristics, the lossy image compression can be classified as

    a) visually lossless, in which the reconstructed image is visually indistinguishable from the original, and

    b) visually lossy, in which the reconstructed image contains visible artifacts or distortions.

To facilitate this, most of the state-of-the-art image compression algorithms employ the discrete wavelet transform (DWT). The DWT not only approximately decorrelates the image and provides good energy compaction, but also establishes a multi-resolution representation of the signal input. This type of representation attempts to mimic the decomposition performed by the cortical basis, with a relatively low complexity. The wavelet-based algorithms also incorporate psychophysical results based on traditional sub-band distortion, white-noise distortion and/or contrast masking based on the value of luminance rather than the image pixel value [4].

1.1.1 Compression Principle

A common characteristic of most images is that the neighboring pixels are correlated and therefore contain redundant information. This redundant information available in the image data is reduced for achieving a less correlated representation or compressed version of the image. The two fundamental features of any compression process are redundancy and irrelevancy reduction. *Redundancy reduction* aims at removing duplication from the signal source (image/video). *Irrelevancy reduction* omits parts of the signal that will not be noticed by the signal receiver, namely the HVS. In general, three types of redundancy can be identified [5].

- *Spatial redundancy,* where correlation exists between the neighboring pixel values.

- *Spectral redundancy,* where correlation exists between different spectral bands.

- *Temporal redundancy,* where correlation exists between the adjacent frames in a sequence of images (in video applications).

Research in image compression aims at reducing the number of bits needed to represent an image by removing the spatial and spectral redundancies as much as possible. Since our focus is on still image compression, the effect of temporal redundancy is not considered in this thesis.

1.1.2 Compression techniques

A brief description of the classification of compression techniques is provided in this sub-section. Based on the quality of the retrieved image, any compression technique can be broadly classified as lossless and lossy.

(a) Lossless compression: Lossless compression attempts to reconstruct a compressed image so that its quality matches the original image [6]. Although lossless compression sounds ideal, it does not provide much compression and the file size remains quite large. For this reason, lossless compression is used where detail information is extremely important, for example, when planning to make large prints. Lossless compression is offered in the form of TIFF and RAW file formats for use in digital cameras.

(b) Lossy compression: Lossless compression is not a practical solution for many applications. Hence, most of the recent digital cameras offer lossy compression in addition to or in place of lossless compression, and this may result in some image degradation. In many situations, such as posting images on the web or making small to medium sized prints, the image degradation is not that obvious.

In general, a picture can be compressed much more than a text, since it allows a higher degree of information loss and there is more redundancy in a picture than in a text. A 5:1 compression ratio is considered good when a lossless compression scheme is employed. Some of the lossless compression schemes in use are statistical compression methods

such as Huffman-coding, entropy-coding, arithmetic coding, predictive coding methods (can be either lossy also), discrete cosinus-transform (DCT) (used in JPEG algorithm) and CCITT group 1 to 4. For achieving a higher compression ratio of upto 100:1, lossy compression methods like fractales, splines and wavelet-based coders (EZW, EBCOT, LZC, SPIHT, etc) are employed.

1.2 Scope of the thesis

Medical advancements, taking advantage of the digital technology, have paved the way for interactive teleradiology and many new medical imaging techniques. Each medical image generated in a hospital has about 1 million bytes on an average. The size and number of such images are large, and compression offers a means to reduce the cost of storage and increased speed of transmission. There is a precipitous fall in the cost of storage with the increase in the capacity and the cost of transmission, resulting in a strong demand for medical image compression. The quality of the final reconstructed digital image is of great importance for these images, since diagnosis is done based on the visual attributes of the reconstructed image. Hence, for images where need for smooth quality is more critical, lossless compression schemes are employed. These schemes provide a maximum compression ratio of 4:1. To achieve higher compression ratios, certain lossy compression algorithms are used provided the HVS characteristics are considered. The integration of the HVS model in image compression for better perceptual quality has spawned a wide range of research. The complex perceptual-based compression schemes proposed in [1, 7, 8, 9] are costly to develop, implement and deploy. In this dissertation, we have attempted to develop a scheme for the enhancement of the visual quality of the

reconstructed compressed images with reduced complexity. A contrast sensitivity function (CSF) masking method is proposed here for use with the nearly-orthogonal wavelet transform, and the overall effect analyzed. An extensive evaluation of the system performance in the noisy environment is conducted for a variety of standard test and digital medical images. Also, a multiplier-less design for the BNC 17/11 wavelet filter is proposed for implementation in a FPGA or ASIC hardware.

1.3 Thesis Organization

The dissertation consists of five chapters. Since most of the compression algorithms are applied to graphical images, the basic concept of graphical image storage has been introduced in this chapter. *Chapter 2* provides a brief overview about multi-resolution analysis (MRA) and wavelets. An introduction to orthogonal and biorthogonal wavelet families is also provided. The symmetry conditions are described along with a brief discussion about BNC wavelets. A description of perceptual modeling of the HVS is given. A brief overview about the perceptual modeling of the HVS using the CSF is provided. A concise background is given for the CSF model and the different quality assessment metrics. A state-of-the-art overview pertinent to the CSF models and perceptual image compression approaches are also described. *Chapter 3* presents our contribution towards enhanced digital image compression. A novel CSF band peak-average (CSF-BPA) sub-band specific mask is derived using the Mannos and Sakrison's model. Further, a scheme incorporating this newly-deduced CSF masking and using the recently explored BNC wavelets is proposed in an attempt to achieve an enhancement in the perceptual quality of the reconstructed medical image. A multiplier-less design for the implementation of the BNC 17/11 filter in a FPGA or ASIC hardware is also proposed.

6

Chapter 4 is dedicated to the analysis of the compression results. A performance comparison is made between the most commonly used JPEG-standard CDF 9/7 biorthogonal wavelet system and two of the recently explored nearly-orthogonal or biorthogonal nearly coiflet (BNC) wavelets, namely, the symmetrical BNC 17/11 wavelet and the asymmetrical BNC 22/14 wavelet, for a variety of standard test natural and digital medical images. These include the objective and subjective qualities of the reconstructed images for different wavelet types, for a specific range of compression ratio, with and without the application of the CSF masking. *Chapter 5* draws the main conclusions from the thesis dissertation and highlights the contributions. Some recommendations for future work are also provided.

Chapter 2

Multi-resolution analysis and its extension to wavelets

This chapter presents an overview of the multi-resolution analysis, its extension to wavelet study and perceptual modeling of the human visual system. The various assessment techniques for testing the image quality of the reconstructed image are also discussed in brief.

2.1 Multi-Resolution Analysis (MRA)

A signal or an image is composed of a smooth background and fluctuations, usually referred to as details. The distinction between the background and the fluctuations is determined by the resolution or the scale below which the details cannot be detected. At a given resolution, a signal is approximated by ignoring all fluctuations below that scale. With an increase in the resolution, finer details are added, providing a better approximation of the signal. At infinite resolution, the exact signal can be recovered. In analytical representation, consider a function $f(t)$ and let j be the resolution level. At each resolution the function is composed of background $f_j(t)$ and fluctuation $d_j(t)$ information specific to j. The function at the next level of resolution $j+1$, is given as, $f_{j+1}(t) = f_j(t) + d_j(t)$. The original function $f(t)$ is recovered when the resolution is infinite.

Mathematically,

$$f(t) = f_j(t) + \sum_{k=j}^{\infty} d_k(t)$$

where the time instant, $t = 1, 2,.., \infty$. This type of analysis of a signal, where there is simultaneous presence of different resolutions is referred to as multiresolution analysis. The above equation represents one way of analyzing a multi-resolution signal. Similarly, one can view the space of functions that are square integrable, $L^2(R)$, as composed of a sequence of signal (approximation) and detail (fluctuations) subspaces V_j and W_j. At each level 'j', the sum of smooth and detail signals combine into a multiresolution of the input image or signal at a finer level $j+1$ as shown in Fig. 2.1. In short, the given image or signal is divided into different scales of resolution. To summarize, a multiresolution analysis of $L^2(R)$ is a nested sequence of subspaces $\{V_j\}_j \in Z$ (Z is the set of all integers) such that the following conditions are met.

- Two detail subspaces (W_j, W_{j+1}) of different resolution are orthogonal and the detail subspace (W_j) at a given resolution j is orthogonal to the approximation subspace (V_j) in the same resolution.

- All square integrable functions are included in the finest resolution level and the zero function is included in the coarsest level. As the resolution becomes coarser i.e. when j tends to infinity, all the details will be removed and only a zero function will remain.

- The third condition is scale invariance. This states that all the spaces (V_j) are scaled versions of the core space (V_0). If $f(t)$ is in V_j, i.e., there is no detail information for scales lower than $1/2^j$, then $f(2t)$ is obtained by squeezing $f(t)$ by a

9

factor of 2. Here $f(2t)$ does not have any detail information for scales lower than $1/2^{j+1}$.

- The subspaces V_j should be shift invariant.

Signal at level 'j'   Details at level 'j'

(Local averages)   (Local differences)

$+$

Signal at level j+1

$L^2 \rightarrow$ Completeness of space

V_0 V_j – Signal subspace (contained in next subspace)

V_1

V_2 W_j – Details subspace

W_0

W_1

$V_{j+1} = V_j (+) W_j$
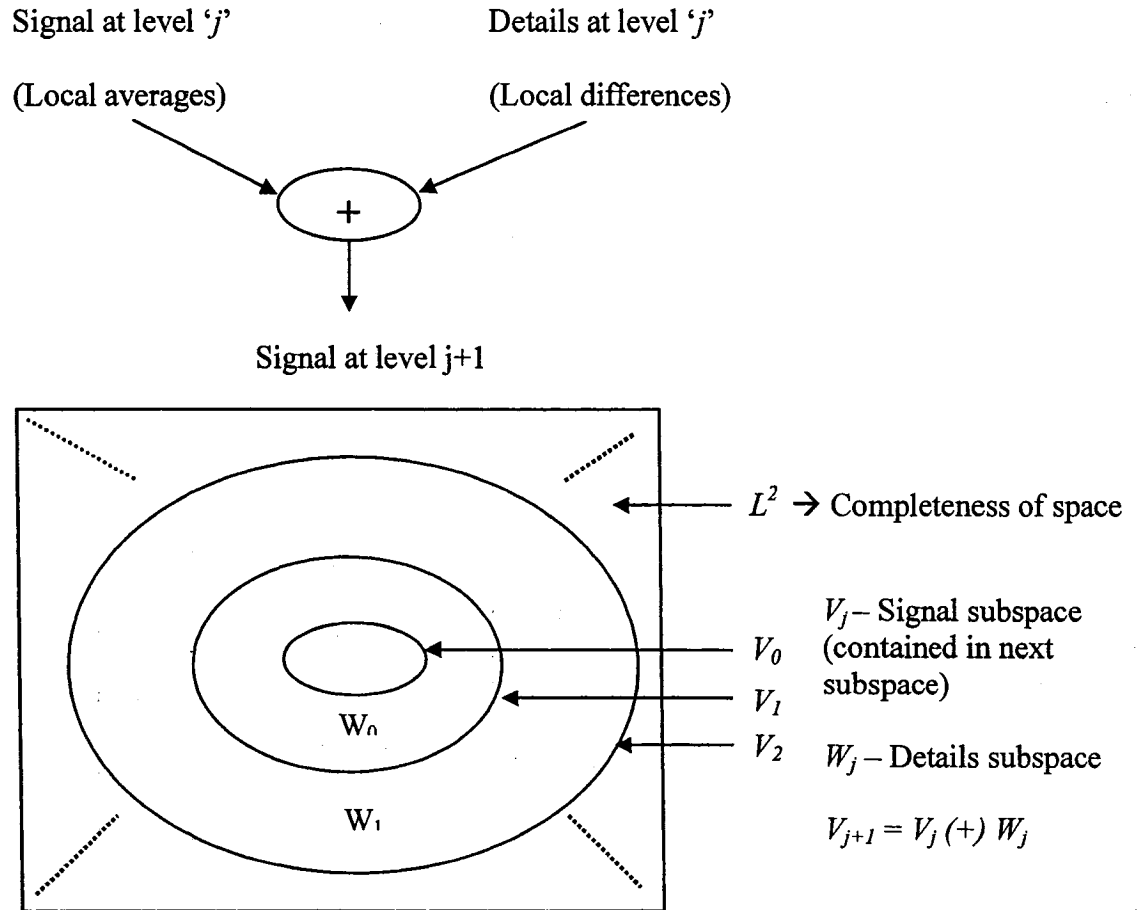
Figure 2-1 Representation of multi-resolution analysis

Wavelets constitute a special class of basis functions for $L^2(R)$. At a given resolution of an image or a signal, scaling functions $\Phi(2^j t - k)$ form the basis function for a set of approximate coefficients (smooth signals), and the wavelet functions $w(2^j t - k)$ generate

the set of detail coefficients. A single wavelet generates a whole basis and the functions

$\Phi(2^j t - k)$ & $w(2^j t - k)$ are obtained by dilation and translation (for all j & k) [10, 11, 12].

A detailed description about the wavelet transform is provided in the following section.

2.1.1 Wavelet basics

Wavelets are functions defined over a finite interval of time, having an average value of zero [13]. The basic idea behind the wavelet transform is to represent any arbitrary function $f(t)$ as a superposition of a set of wavelets or basis functions. These basis functions or baby wavelets are obtained from a single prototype wavelet called the mother wavelet by dilations or contractions (scaling) and translations (shifts). Wavelets are broadly classified as continuous and discrete. There are two kinds of transforms defined for these wavelet types, namely, the continuous wavelet transform (CWT) and the discrete wavelet transform (DWT). The CWT of a function $f(t)$ is defined by

$$\gamma(\tau,s) = \int_{-\infty}^{+\infty} f(t) \frac{1}{\sqrt{|s|}} \psi^* \left(\frac{t-\tau}{s} \right) dt \qquad (2.1)$$

where t represents time, τ represents translation, s represents scale, $\psi(t)$ is the *mother wavelet and* $\psi^*(t)$ is the complex conjugate of ψ. The original function $f(t)$ can be reconstructed using the following inverse relation given as follows.

$$f(t) = \frac{1}{C_\psi} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \gamma(\tau,s) \psi \left(\frac{t-\tau}{s} \right) d\tau \frac{ds}{|s|^2} \qquad (2.2)$$

where $C_\psi = \int_{-\infty}^{+\infty} \frac{|\psi(\varsigma)|^2}{|\varsigma|} d\varsigma$ is called the admissibility constant.

The DWT of a digital signal x[n] is calculated by passing it through a series of low pass and high pass filter banks (one such filter bank is shown in Fig. 2.2 where g[n] and h[n] represent the impulse response of the low pass and high pass filter respectively). The filtered low pass and high pass coefficients are then interpolated or decimated, and the process is repeated till the required decomposition is achieved before further processing. The DWT is defined by the following equation:

$$y[n] = x[n] * g[n] = \sum_{k=-\infty}^{\infty} x[k].g[n-k] \qquad (2.3)$$
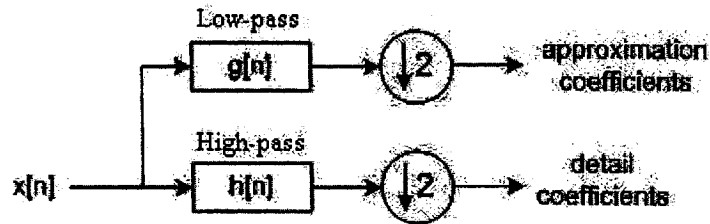
Figure 2-2 Block diagram of a filter bank

The discrete wavelet transform (DWT) of a finite length signal x(n) having N components is expressed by an N x N matrix. The DWT transforms a signal from the time domain to the joint time-scale domain. This means that the wavelet coefficients are two-dimensional, i.e., to compress the transformed signal not only are the coefficient values coded but also their position in time. For an image signal, the position in time is expressed as its position in space. The wavelet transformed image is represented using trees (root-descendant relation between the different spatial frequency regions) because of the sub-sampling that is performed in the transform. Each of the coefficients in the lower sub-band has four descendants in the next higher sub-band.

For a single level wavelet transform, the signal is passed through a pair of filters, one being a low-pass and the other a high-pass (see Fig 2.2). The low and high pass filter banks are designed with some imposed orthogonality or biorthogonality criterion. The theory behind orthogonal and biorthogonal wavelets is provided in Section 2.2. The low pass (LP) filtered output values or the approximation coefficients are obtained through an averaging of the neighboring pixel values. The high pass (HP) filtered output values or the detail coefficients are obtained through differencing the neighboring pixel values. At each level as shown in Fig 2.3, the signal is decomposed into low and high frequencies. Due to the decomposition process the input signal must be a multiple of 2^n, where n is the number of levels as shown in Fig. 2.4. The pair of LP and HP filters at the input end is referred to as the analysis filter banks (AFB) and at the output end those used for performing the inverse transform are usually referred to as the synthesis filter banks (SFB).
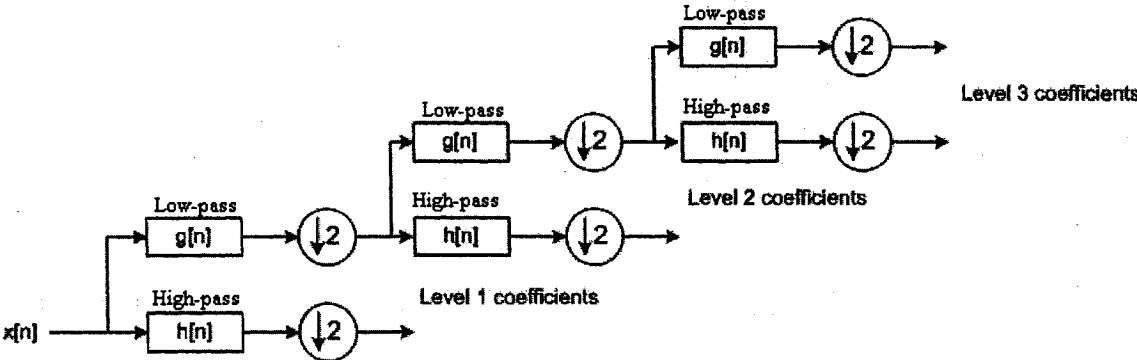
Figure 2-3   Representation of a 3-level filter bank

For a 1-D DWT, the input array (row or column matrix) is passed through the AFB and decomposed into 'n' (user-defined) levels. The decomposed signal is further processed

13

and transmitted. At the receiver end, the received signal is reconstructed back to the original size array by passing it through the SFB. The 2-D DWT can be visualized as 1-D DWT applied individually to the row and column matrices of the input [14]. The wavelet filter transformed coefficients are obtained by calculating the average and difference between adjacent pixels.
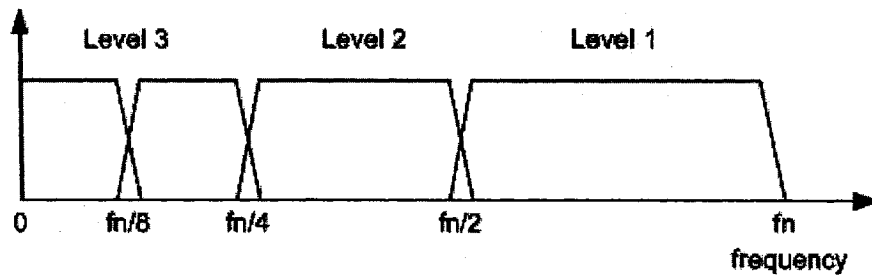
Figure 2-4   Frequency spectrum of a 3-level DWT space

2.1.2 Significance of wavelet-based compression

Despite the advantages of the JPEG compression scheme (based on DCT) such as simplicity, satisfactory performance and availability of special purpose hardware for implementation, there exist certain shortcomings. The input image needs to be divided into blocks and the correlation across the block boundaries is not eliminated. This results in noticeable and annoying "blocking artifacts", particularly noticeable at low bit rates as shown in Fig. 2.5. Lapped orthogonal transforms (LOT) attempt to solve this problem by using the overlapping blocks smoothly. Although blocking effects are reduced in LOT compressed images, increased computational complexity of such algorithms do not justify wide replacement of DCT by LOT.
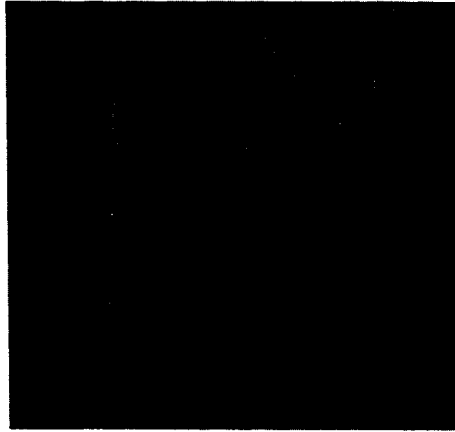
(a)
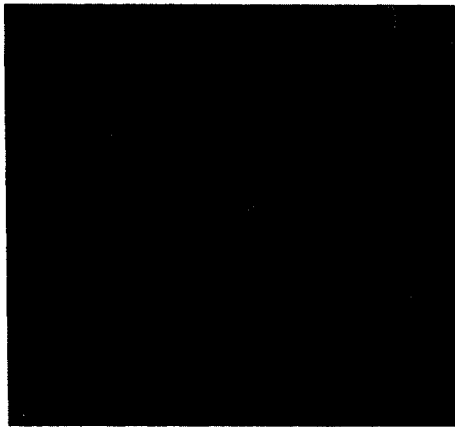
(b)

Figure 2-5 (a) Original and (b) reconstructed Lena (with DC component only) using

DCT to show blocking artifacts

The performances of the wavelet-based codecs are found to be comparable with the

lossless schemes. These codecs have been designed based on the following observations.

1. A natural image has a low pass spectrum. The energy of the sub-bands decreases with

decreasing scale (i.e. high resolution) when a wavelet transformation is performed.

2. Large wavelet coefficients are more important than small wavelet coefficients. This is due to the fact that just the large coefficients are sufficient to define the image and the HF or small coefficients add only the fine details of the image.

Over the years, the wavelet transform has gained widespread acceptance in signal processing, especially in image compression. In many applications, wavelet-based schemes (also referred to as sub-band coding) outperform other coding schemes. Wavelet coding schemes with higher compression avoid blocking artifacts, since their basis functions have variable length and there is no need to divide the input image into blocks. Wavelet-based coding is more robust for transmission and for decoding errors, and also facilitates progressive transmission of images. In addition, they are better matched to the HVS characteristics. In view of their inherent multi-resolution nature [5], wavelet coding schemes are especially suitable for applications where scalability and tolerable degradation are important. All of the above mentioned advantages are considered for employing the wavelet transform for this study.

2.1.3 Relation between filterbank and wavelet transform

Daubechies [15] was the first to discover that the discrete-time filters or quadrature mirror-filters (QMFs) can be iterated and that under certain regularity conditions will lead to continuous-time wavelets. This has proved to be a very practical and extremely useful wavelet decomposition scheme, since FIR discrete-time filters can be used for implementation. The latter follows from the fact that the orthonormal bases correspond to a subband coding scheme with exact reconstruction property using the same finite impulse response (FIR) filters for decomposition as well as for reconstruction. There are several ways in which wavelet transforms decompose a signal into various subbands.

These include uniform decomposition, octave-band decomposition and adaptive or wavelet-packet decomposition. Out of these, octave-band decomposition is the most widely used one. This is a non-uniform band splitting method that decomposes the lower frequency part into narrower bands, and the high-pass output at each level is left without any further decomposition. Fig. 2.6 shows the various subband images of a 3-level octave-band-decomposed Barbara image using the popular CDF-9/7 biorthogonal wavelet [5, 10, 11]. Like JPEG, wavelets can compress an image into a very small file without a noticeable degradation. Also, wavelets offer the following additional advantages for image signal analysis:

- *CMYK image processing* [16]: All commercial implementations of JPEG (discrete cosine transform and Huffman coding form the underlying algorithm.) are restricted to RGB and gray-scale images. But the ISO standard defines only the RGB procedure, so any file containing non-RGB data would be deemed proprietary format and would not be portable. In contrast, all of the wavelet implementations can handle CMYK color model and multi-channel images.

- *Lossless or lossy:* Wavelets offer a theoretically lossless compression. That is, an image that has been compressed and decompressed can be bit-for-bit identical with the original image. For the lossless compression case with a specified compression ratio requirement, wavelets promise to yield a smaller file compared to PDF's ZIP (flate) algorithm.

- *Artifacts:* For moderate compression, wavelet artifacts are less objectionable than JPEG artifacts. Wavelet artifacts are smooth blurs that can often blend into the surrounding imagery, while JPEG artifacts surround each detail with clouds of

17

visual "echoes". The difference is most pronounced with text over an image background.

LLL, LLL	LLL, LLH	LL, LH	L, H
LLH, LLL	LLH, LLH		
LH, LL		LH, LH	
H, L		H, H	

(a)
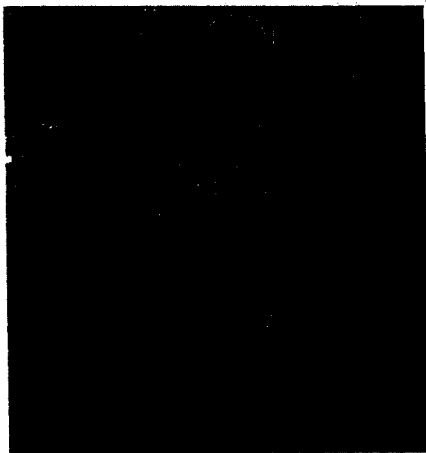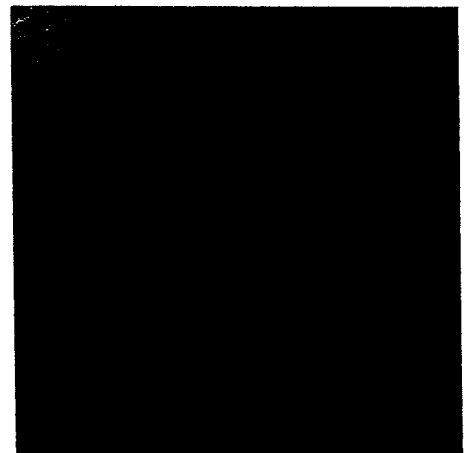
(b)   (c)

Figure 2-6   (a) Spectral Decomposition (b) Original *Barbara*

(c) Octave-band representation for *Barbara*.

18

2.2 Wavelet Classification

2.2.1 Orthogonal wavelets

The scaling function $\{\phi(t-k)\}$ satisfies orthonormality. Thus we have, in the frequency domain (ω is the radial frequency),

$$\sum_{k=-\infty}^{\infty} |\phi(\omega + 2\pi l)|^2 = 1 \qquad (2.4)$$

under conditions

$$\langle \phi_{jk}, \phi_{jl} \rangle = \delta_{kl} \qquad (2.5a)$$

$$\langle \psi_{jk}, \psi_{jl} \rangle = \delta_{ij}\delta_{kl} \qquad (2.5b)$$

$$\langle \phi_{jk}, \psi_{jl} \rangle = 0 \qquad (2.5c)$$

where (1) ϕ_{ij} is the scaling function with the ith position along the x-axis for a scale of 2^j

(2) ψ_{ij} is the wavelet function with the ith position along the x-axis for a scale of 2^j

(3) j, k and l are integers

(4) Delta function, $\delta_{ij} = \begin{cases} 1 & i = j \\ 0 & otherwise \end{cases}$

If a set of basis functions $\{\phi(t-k)\}$ are not orthonormal, it is difficult to orthogonalize them in the Fourier space. The scaling $\phi(t-k)$ and wavelet $\psi(t-k)$ functions are said to be orthogonal if

$$H(\omega)\overline{G(\omega)} + H(\omega+\pi)\overline{G(\omega+\pi)} = 0 \qquad (2.6a)$$

where $H(\omega)$ and $G(\omega)$ are the transfer functions of the low pass ($h[n]$) and high filter ($g[n]$) respectively.

19

This equation relates the coefficients appearing in the wavelet equation to those appearing in the dilation equation. The orthogonal wavelet (satisfy orthogonality criteria) filters $g[n]$ and $h[n]$ are chosen such that, $g[n] = (-1)^n h[1-n]$. Here, a single set of low-pass and high-pass filter is used. These wavelets are obtained using spectral factorization techniques.

2.2.2 Biorthogonal wavelets

A basis that spans a space need not necessarily be orthogonal. In order to gain a greater flexibility in the construction of the wavelet bases, non-orthogonal wavelets are employed. For example, it is well-known that the Haar wavelet is the only known wavelet that is compactly supported, orthogonal and symmetric. In many applications, the symmetry of the filter coefficients is also desirable, since it results in linear phase of the transfer function. For the construction of the biorthogonal wavelet systems that are compactly supported symmetric wavelets, the orthogonality conditions are relaxed.

Here, in contrast to the construction of orthogonal wavelets, the spectral factorization is not needed. Instead of one set of low-pass and high-pass filters for the orthogonal case, two sets are considered for the biorthogonal case, one for the decomposition or analysis and the other for the reconstruction or synthesis. There exist a dual for each of the scaling ($\phi(t), \tilde{\phi}(t)$) and wavelet ($\psi(t), \tilde{\psi}(t)$) functions. In other words, the inner product of the scaling function and its dual is the delta function δ_{ij}; a similar result is true for the wavelet function, that is,

$$\left\langle \tilde{\phi}_i, \phi_j \right\rangle = \delta_{ij}$$

and (2.6b)

$$\langle \widetilde{\psi}_i, \psi_j \rangle = \delta_{ij}$$

2.2.3 Symmetry conditions

For audio, image and many other DSP applications, one of the highly desirable filter characteristics is the linear phase, which corresponds to the symmetry of the associated wavelet system. There are two types of symmetry defined for any kind of a wavelet system, namely, the whole-point symmetry (WPS) and the half-point symmetry (HPS). Table 2.1 summarizes the two types of symmetry for wavelet systems. For a WPS (HPS) wavelet system, all the four filters have odd (even) length [17]. Unfortunately, there does not exist any non-trivial WPS or HPS two-channel orthogonal wavelet system [18, 19], i.e., for any non-trivial two-channel wavelet system, symmetry and orthogonality cannot be achieved simultaneously.

2.2.4 Biorthogonal Nearly Coiflets (BNCs)

Wavelets are classified as belonging to biorthogonal and orthogonal family based on the type of the scaling and wavelet functions used. The biorthogonal wavelets are constructed such that they have temporal symmetry as compared to the orthogonal wavelets. For most image compression applications, the standard CDF 9/7 wavelet filter is used because of its linear phase response.

To achieve higher compression ratios, coiflet wavelets [17] can be used. Coiflets are high-performance orthonormal wavelet systems, which emphasize the vanishing moments for both the scaling and wavelet coefficients. These coiflet wavelets when used

21

for image compression provide a better compression performance. These coiflet wavelets, though orthogonal, are found to be closer to a symmetric wavelet.

Table 2-1   Types of symmetry for wavelet systems

Whole-point symmetry (WPS)	Half-point symmetry (HPS)
$\tilde{h}[n] = \tilde{h}[-n]$	$\tilde{h}[n] = \tilde{h}[1-n]$
$\tilde{g}[n] = \tilde{g}[-n]$	$\tilde{g}[n] = -\tilde{g}[1-n]$
$h[n] = h[-n]$	$h[n] = h[1-n]$
$g[n] = g[-n]$	$g[n] = -g[1-n]$
$\tilde{\phi}(t) = \tilde{\phi}(-t)$	$\tilde{\phi}(t) = \tilde{\phi}(1-t)$
$\tilde{\psi}(t) = \tilde{\psi}(-t)$	$\tilde{\psi}(t) = -\tilde{\psi}(1-t)$
$\phi(t) = \phi(-t)$	$\phi(t) = \phi(1-t)$
$\psi(t) = \psi(-t)$	$\psi(t) = -\psi(1-t)$

The orthogonal wavelets, irrespective of the fact that they provide better compression ratios in comparison with the biorthogonal wavelets, seldom satisfy symmetry requirements for achieving a linear phase response. The biorthogonal wavelets satisfy both the symmetry and linear phase requirements. BNC wavelets [4, 19, 20] were developed to overcome the problem of having symmetrical orthogonal wavelet systems. BNC wavelets are the recently explored biorthogonal wavelet systems, which are considered to be the biorthogonal counterparts of certain coiflet families. These filters are realized when the orthogonality constraints are relaxed for the coiflet wavelet and the vanishing moments are equally distributed for the scaling and wavelet functions of a

regular biorthogonal wavelet system. The resulting wavelets have interpolating scaling functions, linear phase filter banks and dyadic fractional filter coefficients which can be easily implemented using a multiplier-less architecture.

These new class of biorthogonal wavelets have been developed both for the even-ordered and odd-ordered conditions, and their comparative performances have been studied. In this category, the BNC 22/14 (asymmetric realization) [20] and BNC 17/11 [19, 21] are claimed to have the best performance among all the BNCs. The filter coefficients for BNC 17/11 and BNC 22/14 are given in Table 2.2 and Table 2.3, respectively [19]. The frequency response of the analysis and synthesis filters for the BNC 17/11 wavelet have been compared with the standard CDF 9/7 response in [19]. Fig 2.7 gives the analysis and synthesis frequency response characteristics for not only the CDF 9/7 and BNC 17/11, but also for the BNC 22/14 wavelet filter. From the figure, it is seen that these BNCs give a response comparable to the 9/7 wavelet [19]. The response of the BNC 17/11 wavelet filter is found to be closer to the standard filter response. Also, these BNC wavelets possess maximum number of vanishing moments. In applications such as signal compression and denoising, it is desirable for the analysis function to have vanishing moments and regularity, and for the synthesis function to possess smoothness. The scaling function $\phi(x)$ and the wavelet function $\psi(x)$ and their corresponding duals $\tilde{\phi}(x)$ and $\tilde{\psi}(x)$ for a given set of analysis (h_0, h_1) and synthesis (g_0, g_1) filters are defined by the following equations:

$$\phi(x) = \sqrt{2} \sum_n h_0(n)\phi(2x - n) \tag{2.7}$$

$$\tilde{\phi}(x) = \sqrt{2} \sum_n g_0(n)\tilde{\phi}(2x - n) \tag{2.8}$$

23

$$\psi(x) = \sqrt{2} \sum_n h_1(n)\phi(2x - n) \qquad (2.9)$$

$$\tilde{\psi}(x) = \sqrt{2} \sum_n g_1(n)\phi(2x - n) \qquad (2.10)$$

Table 2-2 Filter coefficients for BNC 17/11 wavelet [19]

	BNC 17/11 – Symmetric wavelet
Analysis Filter	0.8402696692 0.4090630083 -.1073757602 -.0621741791 0.0533641923 0.0073357876 -.0135767155 -.0006712263 0.0010068394
Synthesis Filter	0.7568252267 0.4226067872 -.0331456304 -.0814830079 0.0082864076 0.0124296114

Table 2-3 Filter coefficients for BNC 22/14 wavelet [19]

	BNC 22/14 – Asymmetric wavelet
Analysis Filter	0.51620125 0.05573021 -.10097515 0.01279669 0.02604553 -.00659508 -.00465364 0.00085361 0.00068975 -.00005047 -.00004270
Synthesis Filter	0.45822144 0.11455536 -.06873322 -.01963806 0.01527405 0.00208282 -.00176239

Equations (2.7) and (2.8) converge to compactly supported basis functions for regular filters if both the following assertions hold [13]:

$$\sum_n h_0(n) = \sqrt{2}$$

and   (2.11)

$$\sum_n h_1(n) = \sqrt{2}$$

24
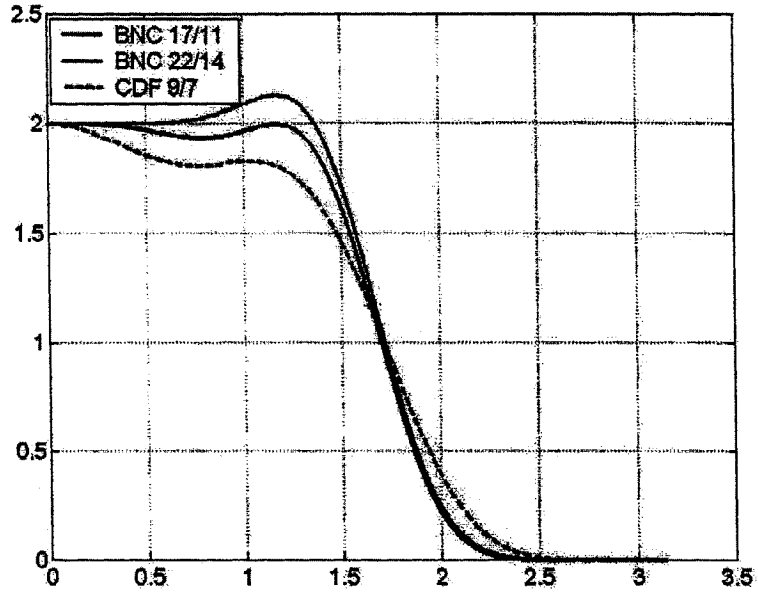
(a)
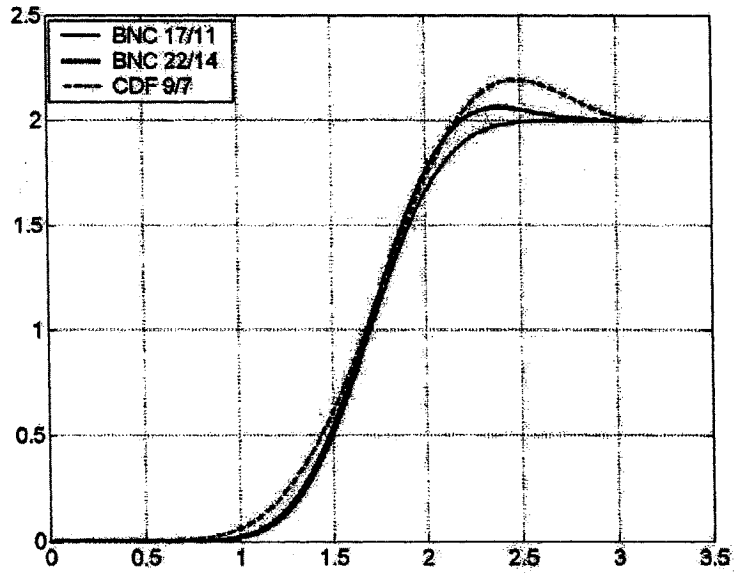
(b)

Figure 2-7 Low-pass frequency responses of analysis and synthesis CDF9/7, BNC 17/11 and BNC 22/14 wavelet filters. (a) Analysis LPF. (b) Synthesis LPF.

These biorthogonal nearly coiflet wavelets are hence found to be appropriate for representing discrete-domain data. The significant properties of these wavelets are listed as follows.

- Sparse representations for smooth signals

- Interpolating scaling functions

- Linear phase filterbanks

- Dyadic fractional filter coefficients.

These properties promise superior rate-distortion performance, better perceptual quality, and lower computational complexity in the field of image coding [17].

2.3 Perceptual models

Perceptual based techniques promise performance gains such that an image interpreted by the human eye provides perceptual accuracy. Due to the non-linear nature of the HVS, the degree of effort required to attain an acceptable perceptual accuracy varies relative to the image. The rendering algorithms based on perception of the HVS take advantage of this phenomenon and attempt to expend only "just-necessary" effort over the image to meet the perceptual accuracy criteria. To accomplish this, certain computational models are proposed for the visual system [22]. These vision models [2, 22, 23] predict the variations in the sensitivity of the human vision with respect to the background illumination levels, the spatial frequency and scene contrast features.

2.3.1 Contrast Sensitivity Measurement

Distortion is a signal superimposed on the original image by the coding process and leads to a degradation of the image quality. This is not perceived by a human observer and

perceptually lossless compression can be obtained as long as the magnitude of the distortion is below the level of visibility threshold. The visual acuity or visual threshold depends mainly on the spatial frequency, the luminance value and the contrast value of the background. More information about these factors can be found in [24].

The contrast sensitivity function can be considered as a "spatial frequency response" of the human visual system at low to moderate light levels. It is an input-output function of the overall visual system, in the same sense that a computer vision system has a spatial frequency response. At low to moderate input levels, the visual system can be viewed as a linear system. But at higher levels, the system becomes nonlinear. In terms of visual perception, the CSF is defined by the sensitivity plotted as a function of the spatial frequency. Here, the sensitivity is measured as the reciprocal of the minimum visible contrast of sinusoidal grating stimuli. A sinusoidal grating is simply a two-dimensional (2-D) sinusoid plus a bias or background to maintain positivity. The sensitivity is normally measured for a few frequencies to define a CSF curve that is plotted on a logarithmic scale. This procedure is carried out based on the assumption that the actual CSF curve is quite smooth, even though the measurements may be noisy, so that a few observations (sensitivity-frequency pairs) can be used to define the shape of the curve. This same assumption leads to the conclusion that the CSF measurements can be used to estimate the parameters in some functional (mathematical) representation of the CSF function for a given subject or group of subjects. In other words, the measured sensitivity data can be fit to some mathematical model, and in the fitting process, a small number of parameters in the model can be determined from the measured data. This is analogous to

a simple linear regression, in which noisy data can be fit by a straight line with two parameters.

Given the appropriate model, the estimated parameters represent the entire CSF curve or equivalently, a small number of estimated parameters represent the CSF for the experimental subject or group of subjects. If a model adequately represents the CSF from individual subjects, then the CSF differences between subjects or groups should be encoded in the different model parameters. Therefore, the estimated parameters quantify the differences, and they form the basis for further psychophysics studies of the individual differences. In robotic vision, the CSF curve can be viewed as a one-dimensional (1-D) input-output function representing human vision at low to moderate light levels and in steady-state. Various threshold contrast and supra-threshold contrast measurements have been made, and various models have been developed based on the contrast threshold of the HVS [22, 25, 26, 27, 28, 29]. Some of the available CSF models for the HVS are discussed in the following sub-section.

2.3.2 Review of CSF models

Several models and approximations for the CSF have been proposed in the literature [22, 25, 26, 27, 28, 29]. Some are based on fundamental theory, but most have been formulated by curve fitting procedures. In each case, the general form of the model equation has been assumed, and the parameters of the model equation adjusted empirically to achieve a good fit to an experimental CSF curve. Some of the candidate models are discussed in this section.

(a) Sakrison and Algazi's CSF

Sakrison and Algazi [30] suggested a model of the form

$$CSF(f) = K_0 [1 + K_1 . f^2] e^{(-af^2)} \qquad (2.12)$$

where f is the radial spatial frequency.

Sakrison and Algazi's CSF model is based on the rate-distortion theory analysis for optimal imaging encoding. Their optimal frequency weighting, which was then compared to a CSF, is given as

$$CSF(f) = [1 + 0.05f^2] e^{-0.0004f^2} \qquad (2.13)$$

Equation (2.12) is a generalization of their weighting function, with fitting parameters K_0, K_1 and a. Here the number of parameters is $P = 3$. These three parameters will be adjusted to achieve the optimal fit to the CSF data.

The peak sensitivity occurs at the frequency

$$f_{max} = \sqrt{((K_1 - a)/a.K_1)} \qquad (2.14)$$

The sensitivity is obtained by evaluating (2.13) at the frequency $f = f_{max}$. Here, the sensitivity at very low frequencies does not go towards zero. Instead, the sensitivity approaches K_0, which was unity in the original model.

(b) Mannos and Sakrison's CSF

Mannos and Sakrison [25] continued the image encoding studies based on the rate-distortion theory and empirically arrived at the frequency weighting function defined by

$$CSF(f) = 2.6(0.0192 + 0.144.f) e^{-(0.144.f)^{1.1}} \qquad (2.15)$$

where the radial spatial frequency is given by

$$f = \sqrt{f_x^2 + f_y^2}$$ (2.16)

f_x and f_y being the spatial frequencies in the x and y-axis respectively. Like the Sakrison and Algazi expression in (2.12), the Mannos and Sakrison frequency weighting function has an amplitude (sensitivity) which approaches K_0 at very low frequencies. The peak sensitivity in the Mannos and Sakrison equation occurs at the central spatial frequency. This model is found to correlate more accurately with the band pass behavior of the HVS characteristics [25].

(c) Modified Dooley CSF

Dooley [31], as noted in Levine [32], suggested an alternative expression for representing the contrast sensitivity function at each spatial frequency f, given by

$$CSF(f) = 5.05[1 - e^{-0.1f}]e^{-0.138f}$$ (2.17)

(d) Pelli, Legge and Rubin's CSF

Pelli et al [26] suggested that a simple parabolic fit to CSF would be fast, can easily be implemented in a spreadsheet and is sufficiently accurate for a narrow range of frequencies near the peak sensitivity. Their equation for sensitivity is

$$[S - S_0] = -K[f - f_0]^2$$ (2.18)

which is a parabola on log-sensitivity (S) versus log-frequency (f), with a peak at the coordinates (f_0, S_0). The parameter K is a "shape factor" which controls the width of the parabola. The number of parameters is $P = 3$, but Pelli and colleagues set $K = 1.15$ to reduce the number of parameters to $P = 2$. In terms of sensitivity CSF and frequency f, their more general equation is

$$\log CSF - \log CSF_0 = -K[\log f - \log f_0]^2 \qquad (2.19)$$

(e) Ahumuda's CSF Filter

Ahumuda [33] modeled the CSF characteristics to be the balanced difference of two Gaussians with the centre and surround frequencies represented by f_c and f_s with a_c and a_s as the corresponding amplitudes. The CSF is given as

$$CSF(f) = a_c.e^{(f/f_c)^2} - a_s.e^{(f/f_s)^2} \qquad (2.20)$$

Based on Ahumuda's test results the default values are set as, $a_c = 1$, $a_s = 0.685$, $f_c = 97.3227$ and $f_s = 12.1653$.

(f) Daly's CSF

Daly [29] modeled the CSF of the HVS such that it describes sensitivity variations as a function of the radial spatial frequency, light adaptation level (luminance level L), image size and lens accommodation due to distance, orientation and eccentricity. This model possesses a band pass characteristic and anisotropic behavior caused by a lack of sensitivity in the region of 45 degrees. This model corresponds well with the oblique effect. Daly's CSF is given by

$$CSF(f) = \left(\frac{0.008}{f^3} + 1\right)^{-0.2} 1.42.f.e^{(-0.3f)}\sqrt{1 + 0.06e^{0.3f}} \qquad (2.21)$$

(g) Movshon and Kiorpes CSF

This is the simplest representation of CSF using a three parameter exponential function, first described by Movshon and Kiorpes [27]. It is given by

$$CSF(f) = a.f.e^{-b.f} \qquad (2.22)$$

where f is the spatial frequency in cpd (cycles per degree), and a and b are constants. In general, the values of a and b are chosen as 75 and 0.2 respectively. This model possesses a band-pass shape for the achromatic CSF. For color images, a 2-D form of the equation is used. The main limitation of this model is that it is the same for all viewing conditions such as luminance level and viewing distance.

(h) Barten's CSF

Barten [28] proposed a slightly complicated model for square root integral (SQRI) image quality model. This contrast sensitivity model begins with the optical modulation transfer function (MTF) of the human eye, which is expressed as a Gaussian function. The MTF is then modified with models of photon and neural noise to obtain an isotropic band-pass shape. The modified MTF is now a function of the pupil diameter, image size and luminance level. This model provides slightly greater flexibility to varying viewing conditions and luminance (L). Barten's CSF is defined as

$$CSF(f) = a.f.e^{(-b.f)}\sqrt{1 + 0.06.e^{(b.f)}} \qquad (2.23)$$

where a and b are constant terms defined by

$$a = 440.\left(1 + \frac{0.7}{L}\right)^{-0.2}$$

$$b = 0.3\left(1 + \frac{100}{L}\right)^{-0.15}$$

Daly, Barten, Movshon and Kiorpe's CSF models presumed a low pass characteristic for the visual acuity function. The modified Dooley model [31] is considered the best based on the calculated value of the rms error at the N frequencies used in the least-squares. However, this model provides a good fit only at few points and the resulting function

32

does not provide a good fit to the HVS characteristics at lower and higher frequencies. Mannos and Sakrison's model [25] matches the human visual characteristics more closely by assigning a band pass model for the contrast sensitivity function.

2.4 Image Quality Measurement

To evaluate the effectiveness of digital image processing techniques, one often compares the quality of the image before and after a digital image processing procedure. In compression applications, to compare two compression strategies, the image quality of the final reconstructed image generated by each method is compared. If the quality of the image is found to be similar, then the one that requires fewer bits for coding is considered to be superior in performance. Surprisingly, the measurement of the image quality is not a straightforward procedure as it sounds to be, and usually two types of image quality assessment, namely, objective assessment and subjective assessment, are to be employed [34]. Each of these assessment techniques along with their merits and drawbacks are discussed in the following section.

2.4.1 Objective assessment

The objective quality assessment of an image or a signal is made using certain criteria, such as the root mean square error (RMSE) and peak signal-to-noise ratio (PSNR) values. Let $f(x,y)$ be the input image (of size M x N) to a processing system such as a filter, subsampling system, transformation stage, or a compression sub-system and $g(x,y)$ represent the output of the system. Then, the mean square error (MSE) E_{ms} of the system is defined as

$$E_{ms} = \frac{1}{M.N} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} e(x,y)^2 \qquad (2.24)$$

where $e(x,y)$ is the error component obtained by calculating the difference between the original and reconstructed image or signal.

The root mean square error (RMSE) is defined as

$$E_{rms} = \sqrt{E_{ms}} \qquad (2.25)$$

The signal-to-noise ratio (SNR) is widely used as a metric in evaluating the objective quality of the image. It is usually referred to as SNR_{ms} or SNR_{rms} depending on whether the mean square error or the root mean square error value is considered. SNR_{rms} is defined by

$$SNR_{ms} = 10 \, log_{10} \left[\frac{\sum_{x=0}^{M-1} \sum_{y=0}^{N-1} g(x,y)^2}{MN \cdot E_{ms}} \right] \qquad (2.26)$$

and

$$SNR_{rms} = \sqrt{SNR_{ms}} \qquad (2.27)$$

The peak signal-to-noise ratio (PSNR) is a modified version of SNR_{ms} and is the ratio between the maximum possible power of a signal and the power of corrupting noise that affects the fidelity of its representation. It is defined as follows

$$PSNR = 10 \, log_{10} \left(\frac{255^2}{E_{ms}} \right) \qquad (2.28)$$

In this assessment technique, a larger (smaller) value of the PSNR corresponds to a better (worse) image quality, which implies that the output image corresponds well (poorly) to the original image. But, this interpretation does not consider the characteristics of the

visual stimuli. The perception of the HVS is not simple and generally needs higher level processing to understand its non-linear response. Thus, the SNR does not always provide a very reliable assessment of the image quality. As an example, IGS quantization technique which is used to achieve a higher compression performance, (only 4 bits instead of the usual 8 bits are used for quantization) produces a higher subjective quality, but because of the low-frequency quantization noise and added high-frequency random noise, it results in a lower SNR value.

The objective assessment scheme can be modeled easily and provides the ease for hardware implementation. The primary advantage of mathematical metrics is that they do not require any information about the viewing conditions, need not necessarily adapt to local image content and the computations are simple. These measures are good indicators of distortion for random errors. Since image pixels are correlated, these distortion indicators cannot provide reasonable approximation for these correlated errors. One main argument used against mathematical metrics is that they typically provide a single number for the entire image, and thus cannot reflect the spatial variations in the image quality.

2.4.2 Subjective assessment

This assessment technique takes into account that, in estimating the visual quality of a signal (image or video), the human observer will be the ultimate decision maker. Hence, such an evaluation plays an important role in visual communication applications. Any technique for a subjective assessment is imperfect. Each method may have incorrect assumptions built in, and plausible alternatives invariably exist for any interpretation of the data [35]. Hence, observers are invited to subjectively evaluate the perceptual quality

of the reconstructed images, and are asked to rate the pictures based on a mean opinion score (MOS) scale, which is usually a five-scale rating system. This rating system is used by Bell Laboratories and has also been adopted to be the standard scale in CCIR Recommendation 500-3 (CCIR 1986). As per CCIR recommendations [36], the 1-5 rating is defined by the following factors.

i. Artifacts are not noticeable.

ii. Artifacts are just noticeable.

iii. Artifacts are definitely noticeable, but not objectionable.

iv. Artifacts are objectionable.

v. Artifacts are extremely objectionable.

The subjective analysis is usually performed using the blind test procedure. Blinding is a basic tool to prevent conscious and unconscious bias in research. Because of the complex response of the HVS, typical measurements of the image system quality such as the detective quantum efficiency, mean transfer function and SNR cannot always be used to determine the conditions for optimal perceptual image quality. The *Just Noticeable Difference* (JND) is determined by a group of random observers and also their mean opinion score analysis is recorded. This can be done using single-blind testing, double-blind testing or triple-blind testing.

(a) Single-Blind Trial

In a single-blind trial, the observers are not aware that they are the so-called "test" subjects for the experiment. This trial is used by the experimenter either when he knows all the facts about the experiment, or when it is assured that no bias will be induced by him/her. However, the catch here is that the observers are influenced by interaction with

the researchers, which is usually referred to as the "experimenter effect". Single-blind trials are risky in certain research domains like psychology and social science, where the experimenter may consciously or unconsciously influence the behavior of the subject.

(b) Double-Blind Trial

Double-blind describes a stringent way of conducting the quality assessment experiment, usually on human subjects, in an attempt to eliminate any subjective bias on the part of both the observers and the researchers. The double blind method is an important part of the scientific method, used to prevent research outcomes from being 'influenced' by the placebo effect [37] or observer bias. These experiments are usually carried out to achieve a higher standard of scientific rigor. In a double-blind experiment, neither the individuals nor the researchers are aware of the outcome. Double-Blind research is an important and widely used tool in many fields of research such as medicine, psychology, social sciences and forensics.

(c) Triple-Blind Trial

Triple-blind trials are double-blind trials in which the statistician interpreting the results also does not know which intervention has been given. Sometimes 'triple-blind' is used to mean that multiple investigators are all blinded to the protocol (such as the clinician giving the treatment and a radiologist or pathologist who interprets the results) [37]. This kind of assessment is motivated by the fact that pixel-by-pixel based distortion measures/metrics fail to give a correct assessment. The error metric accounts for several mechanisms of HVS.

2.5 HVS-based image compression: A brief review

Safranek and Johnston [9] proposed a perceptual image coder (PIC) in which each subband is quantized using a uniform quantizer with a fixed step size. The step size is determined by the JND threshold for uniform noise by the most sensitive coefficient in the subband. Different approaches to perceptually motivated bit allocation have been proposed by Malo et al [38] and O'Rourke et al [39] using the sine wave CSF [7] which does not correlate well with the actual HVS structure. Pappas et al [40] have studied different coding algorithms and compared their performance for supra-threshold image compression. Their study included coding schemes like JPEG, Safranek's perceptual JPEG [43], Safranek and Johnston's PIC [9] and the Said-Pearlman's SPIHT algorithm [41], which is based on Shapiro's embedded zero-tree wavelet algorithm [42]. They have shown that at low bit rates, the performance of SPIHT and the PIC coders perform best, and that they provide the best correlation with subjective quality evaluation. Also the authors have reported that scaling of the threshold quantizer step sizes in the PIC produced the best performance for visually lossy compression, and not for visually lossless compression contradicting Safranek's [43] conclusion.

All these methods exploit the frequency and orientation sensitivity of the HVS to some extent, but do not fully exploit the locally varying masking thresholds, which differ for each transform coefficient based on the image content under consideration. This limitation may result in allocating too many bits to less sensitive coefficients, resulting in over-coding of some image components or in unnecessary visible artifacts.

Hontsch and Karam [7] proposed a locally adaptive perceptual image coding which eliminates the necessity of transmitting the side information for perceptually enhanced

coding. Here, the per-coefficient quantizer step size is produced for each subband. It uses a simple perceptual model and exploits the low pass character of the perceptual masking for natural images in estimating the local amount of available masking. But this method does not fully adapt to the local changes in the available masking. Also, the cost of the locally adaptive quantization without side information is a constraint on the order of transmitting the transform coefficients. This method has shown an improvement of about 20% - 30% in comparison with PIC. Some methods have used the HVS properties in wavelet-based image compression. Antonini et al [44] have used vector quantization followed by HVS-based allocation of bits to each subband. The number of bits assigned to each subband is computed via a weighted mean-squared-error (MSE) distortion criterion in which the weights are determined based on Campbell and Robson's sine wave CSF [45]. This approach has been reported to have a performance comparable to that of the previous non-HVS based methods of higher complexity; but, the CSF model assumes a low pass structure for the HVS model.

Lai and Kuo [46] have introduced a scheme in which the contrast sensitivity and visual masking adjustments are performed within the context of a wavelet-based algorithm that uses the Haar filters. Detection thresholds are measured using a low pass representation of the CSF. Watson et al [47] have measured luminance thresholds for both individual wavelet basis function and simulated wavelet subband quantization distortions in an unmasked detection experiment.

Albanesi [48] has proposed a method for incorporating the HVS characteristics directly into the transfer stage of a wavelet-based coder using the HVS model characterized by Mannos and Sakrison [25]. This approach uses the wavelet 9/7 filters, and region-of-

interest coding [7]. Nadenau and Reichel [49] have developed a HVS scheme that uses the entire W subspace of the CSF DWT decomposition as filters for color image compression using the biorthogonal 9/7 wavelet system. A performance improvement, based on the value of the PSNR, of 13% to 49% has been reported in comparison with Albanesi's method. Nadenau et al [50] have also reintroduced the CSF filtering scheme along with a CSF masking scheme. They have incorporated the HVS properties into a wavelet-based coding algorithm via a noise-shaping filtering stage which precedes quantization. The main objective of this approach is to shape the frequency spectrum of the quantization noise to match the inverse of the CSF. This proposed filtering of the transform coefficients via a subband-specific "HVS filter" has been found to increase the compression ratio performance by approximately 30% over baseline JPEG 2000 for images containing near-threshold luminance and color distortions. However, only results for the CSF filtering and not CSF masking have been reported in [50].

Several perceptual quantization approaches have been proposed [4, 9, 33, 38, 43, 47] and the main problem associated with these approaches is the need for true perceptual quantization, which requires the computing and making use of image-dependent, locally varying, masking thresholds. Thus, locally-varying masking thresholds are needed both for encoding and decoding. This, in turn, adds to the storage needed as this side information is sent to the decoder end, thus affecting the bit rate significantly.

A CSF masking scheme, which is based on image content, has been reported by Ramos and Hemami [51]. In this scheme, the grayscale images are segmented into smooth, edge and detail regions and the mask is applied to each of these three regions. The main drawback of this scheme is that the weighting mask values are determined by trial and

error and does not directly incorporate the CSF. Also, the results indicate a decrease in the PSNR value when a mask is applied, and a slight perceptual improvement is achieved only at low bit rates. Two similar approaches have been proposed by Beegan et al [52]. A "CSF mask" is used to adjust the transform coefficients prior to quantization. Beegan et al designed and evaluated two CSF masking methods. Their experimental results show that the 11-weight DWT-CSF mask provides a better PSNR performance, and the band-average CSF mask subjectively outperforms the DWT-CSF mask for all the image types under consideration [52].

To summarize, the previous research to achieve better perceptual image compression has attempted to integrate the HVS model during the quantizer design, coder design or in contrast and luminance masking methods. However, the computational complexity involved in the design of the coder and quantizer is relatively high when compared to the incorporation of masking methods in a compression system. Also, most of the approaches proposed earlier have used the conventional biorthogonal 9/7 wavelet filter, and the evaluation is done solely for noise-free images. This wavelet has irrational filter coefficients. Some of the recently explored wavelet filters [19, 20, 53] promise superior performance compared to the standard 9/7 wavelet for compression application and also possess rational filter coefficients. This thesis work is concerned with proposing a new scheme for low-complexity medical image compression incorporating the HVS characteristics for enhanced perceptual quality of the reconstructed image.

2.6 Summary

This chapter has provided a brief introduction about the multi-resolution analysis and the wavelets. In order to facilitate the understanding of the HVS, the perceptual modeling of the HVS using contrast sensitivity function has been discussed in brief. The state-of-the-art overview pertinent to CSF models and perceptual image compression approaches has also been presented in this chapter.

Chapter 3

Proposed approach for image compression

3.1 Introduction

In a given image, the neighboring pixels are correlated and contain redundant information. If this redundant information can be removed or reduced, a less correlated representation of the image is obtained. Image compression is achieved by reducing the number of correlated pixels. To facilitate the compression process, there exists a variety of transforms. Image compression using discrete wavelet transform (DWT) has proven to be an efficient tool for achieving higher compression ratio [5]. The main advantage of using this transform is that it retains both frequency and spatial information in contrast to existing transform like discrete cosine transform (DCT), where only the frequency information is retained.

In a DWT, the image is decomposed into low frequency or approximation coefficients and high frequency or detail coefficients. The approximation coefficients contain maximum information about the image and hence, they are considered significant for the reconstruction. The detail coefficients are highly correlated and can be removed to a certain extent. Most of the current research is focused towards the development of compression schemes that provide a higher compression ratio without much loss of uncorrelated information.

3.1.1 Conventional Wavelet-based Compression

A conventional wavelet-based image compression scheme is shown in Fig 3.1. Wavelet transform is a filtering process that uses an orthogonal or a biorthogonal wavelet filter bank. The input image pixels are first transformed into spatial coordinates using the 2-D DWT. Here, a large proportion of the low frequency information is converted into fewer coefficients. This minimizes the amount of information to be coded and transmitted. The transformed wavelet coefficients are then quantized with a given level of accuracy. The step size of the quantizer is selected to meet the desired distortion limit and the bit rate considerations. The quantized bits are coded using lossy or lossless coding techniques, for transmission through a channel. At the receiver end, these compressed bits are decoded and the transmitted image reconstructed using the inverse-DWT (IDWT).

It is evident from the literature survey presented in Chapter 2 that most of the existing compression schemes aims to achieve a better peak-signal-to-noise ratio (PSNR) and mean-square-error (MSE) metrics for a given compression ratio. Usually, for the performance assessment of the compression scheme, only objective metrics are used. However, the quality of the reconstructed image is also subjective in nature. Most of the existing quality assessment schemes are objective in nature. These schemes do not provide a meaningful assessment for applications like remote sensing and telemedicine, where the visual quality of the reconstructed image is of paramount importance.

In the following sub-section, we present a compression scheme, to address the shortcomings of the conventional wavelet-based compression scheme, in order to achieve a good visual quality and an acceptable objective measure.
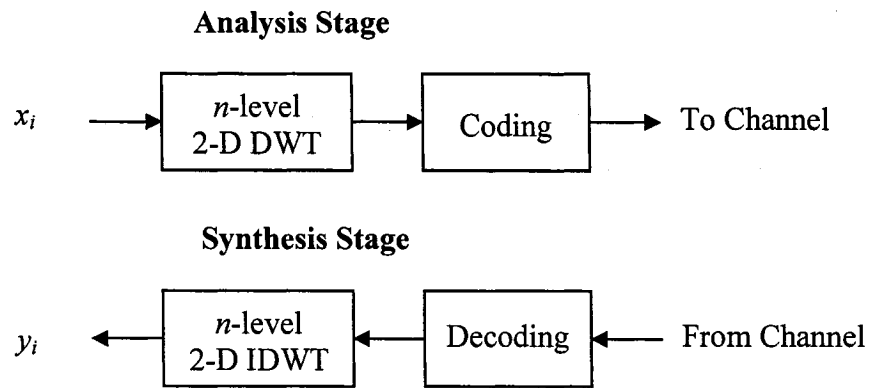
Analysis Stage

Synthesis Stage

Figure 3-1 A conventional wavelet based image compression scheme, where x_i represents the original image and y_i represents the reconstructed image

3.2 Proposed Work

In this section, we discuss the desired characteristics of each of the functional blocks in the compression scheme and the focus of this thesis work. The proposed nearly-orthogonal wavelet-based compression scheme using a novel band peak-average mask (BPA mask) is explained in the later sections.

3.2.1 Focus

In general, the performance of a compression scheme depends on the following factors.

- Selection of the filter (application-specific)

- Selection of the codec (lossy or lossless)

- Accuracy of the quantizer (number of bits used)

The visual characteristics influence the way in which the reconstructed image is perceived by a human observer. Hence, incorporating modules that account for the human eye sensitivity (HVS) characteristics in a compression scheme will result in an enhancement of the visual quality in the final image. This implementation needs to be done in a way so as to obtain an optimal combination of objective and subjective

45

performance. The HVS characteristics can be imposed in one of the following operations, namely, wavelet transform, quantization and coding. The design of the codec and quantizer sub-systems are complex in nature compared to wavelet transform sub-system. Considering the complexity involved in the design and implementation of the HVS-based codec and quantizer modules, the HVS characteristics are incorporated in the wavelet transform to achieve the desired enhancement in the perceptual quality of the image in this thesis work. An investigation of the existing algorithms also shows a good potential for achieving an improvement in the visual quality of the image, when an optimal masking is used combined with an appropriate wavelet filter. The masking module needs to be designed in such a way that its response shares a good correlation with the HVS which is, in fact, an important issue in the research work of this thesis. Such a module, if designed, can provide a cost-effective solution to the problem discussed in section 3.1.1 associated with the perceptual quality of the image. Furthermore, this enhancement has to be achieved without an increase in the overall complexity specified in terms of the computational procedure and the transmission of the coded image information.

In this thesis, the focus is mainly on two issues namely, the design of a new masking method and the selection of a suitable filter. The design of the masking method is discussed in Section 3.2.4. The selection of a suitable filter influences the perceptual quality of the reconstructed image for low and medium bit-rate applications up to a certain extent [19, 21, 44, 46]. This is in view of the fact that the regularity and smoothness of the filter effects the quality of the image. The wavelet filter is selected based on the compression requirements and the application.

Most of the in-use compression systems employ the standard biorthogonal CDF 9/7 wavelet; but, its filter coefficients are irrational and hence, need an infinite precision implementation resulting in a high computation complexity [19]. In particular, it rules out the possibility of using multiplication-free processing for hardware implementation.

A new class of nearly-orthogonal wavelets [19], usually referred to as biorthogonal nearly coiflet (BNC) wavelets, has been developed in order to improve the objective performance and to facilitate the implementation. Some of these wavelet filters possess rational coefficients, and provide higher compression ratios because of the presence of additional vanishing moments in their wavelets and the scaling functions. The BNC wavelets are obtained by sacrificing one vanishing moment in the analysis low frequency stage. Their frequency response is then tailored such that they have an increased width for their pass band. This results in a reduction of the ringing oscillations in the first side lobe. This feature aids in the retention of relatively more high frequency components or detail energy. Considering these features, it has been inferred that these wavelets have the potential for providing a superior rate-distortion performance, better perceptual quality and lower computational complexity, compared to the conventional wavelets in the field of image coding.

In our investigation, we have considered the conventional CDF 9/7 and the biorthogonal nearly coiflet wavelet filters. The proposed scheme for image compression, as discussed in the next sub-section, includes a new masking module and uses the BNC filtering for the wavelet transform.

3.2.2 Proposed scheme for image compression

In this thesis, a digital image compression scheme is proposed to improve the perceptual quality of the final reconstructed image. This scheme comprises of two stages, namely, the analysis and synthesis stages, as shown in Figure 3.2 [61]. The analysis stage consists of the following three functional blocks:

(a) wavelet transform employing the BNC wavelet,

(b) masking module, and

(c) coder.
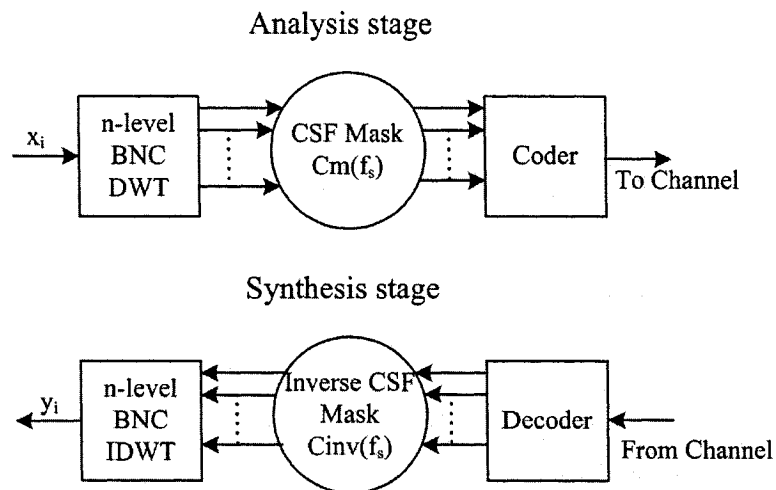
Analysis stage

Synthesis stage

Figure 3-2 Proposed scheme for enhancement of perceptual quality reconstructed compressed grayscale images, where x_i represents the original image and y_i represents the reconstructed image

The synthesis stage also consists of three functional blocks each of which performs the inverse operation corresponding to each of the analysis stage blocks. The masking module plays a significant role towards achieving an enhancement in the perceptual

quality. The characteristics of the masking module need to be determined in accordance with the human visual sensitivity characteristics. In particular, some of the high frequency components usually get attenuated during the coding stage, resulting in a loss of the edge information, and the masking module should reinforce these high frequency components with reference to the perception of human eyes. Also, the use of such a module results in a uniform distribution of the visual errors, thereby converting them into symmetric error components. Such errors are usually less perceived by the human eyes. In other words, by uniformly distributing the visual errors, the masking module aids in achieving higher compression ratios, as it tends to make the artifacts less pronounced at higher compression rates. The masking module thereby improves the visual quality of the reconstructed image.

In addition to the masking considerations, the selection of an appropriate filter bank is crucial for the image quality and the system design. In general, a good filter is needed to have a perfect reconstruction of the image. The perfect reconstruction property assures that the input energy is completely transferred through the filter. The filter should also be symmetric, have linear phase characteristics and should possess good filter regularity allowing acceptable number of vanishing moments [11]. Use of symmetric filters can minimize the detection of false edges while the vanishing moments in a wavelet correspond to the degree of approximation that can be achieved at each decomposition level. Considering all the above mentioned criteria, a symmetric orthogonal wavelet filter can be used for the compression application. Hence, the recently explored BNC wavelet filters [4, 19, 20] are chosen in this thesis work. Further, research [10] has shown that a short-length low pass filter (LPF) in the synthesis stage accumulates fewer ringing

artifacts than a long filter, resulting in a reduction of the most significant perceptual distortion in wavelet transform-coded images. The BNC wavelet filters are chosen with long-length LPF in the analysis stage so as to retain maximum information, and with a relatively short LPF in the synthesis stage.

The masked-wavelet coefficients are coded for transmission. In order to facilitate the coding and decoding processes, the standard zero-tree wavelet-based set partitioning in hierarchical trees (SPIHT) codec [41] is used. This codec is chosen as it is considered as the premier state-of-the-art algorithm in image compression, and is proven to have excellent coding performance. Also, SPIHT produces an optimal embedded bit stream, i.e., the bit stream can be truncated at any instant with a guarantee to yield the best possible reconstruction [41].

Compared to the conventional compression system (see Fig 3.1), the proposed scheme shown in Fig 3.2 has the following two differences:

1) The DWT and the IDWT blocks of the proposed scheme employ the BNC wavelet filters for the n-level decomposition and reconstruction operations so as to obtain an improved rate-distortion performance for higher compression ratios with an ease of computation.

2) The contrast sensitivity function (CSF) masking modules, using the mask values obtained by the proposed masking scheme, are included in the analysis and synthesis stages between the DWT and the Coder, and the Decoder and the IDWT stages, respectively, in order to improve the perceptual quality of the final reconstructed image.

In the proposed scheme, the masking modules are represented by the CSF mask and the inverse CSF mask units in Fig 3.2. The input x_i denotes the original image and y_i the reconstructed image. The input digital image is split into desired number of frequency and orientation selective levels, using the DWT employing the BNC wavelet filters, instead of the conventional biorthogonal CDF 9/7 wavelet. The number of decomposition levels used, represented by n, depends on the input image size. For a given $M \times N$ image, n is chosen to be less than or equal to $2^{\min(M,N)}$. A higher value of n leads to a higher compression ratio, but may lead to a removal of certain essential bits other than the redundant ones and this in turn, can result in the presence of visible artifacts in the reconstructed image.

In the analysis stage of the proposed scheme, a CSF masking module is used to address the issue of the loss of high frequency components. This is achieved by applying a mask $(C_m(f_s))$ to each sub-band of the wavelet coefficients. The mask values are obtained using the proposed band peak-average (BPA) mask method [61] described in Sections 3.2.3 and 3.2.4.

The masked wavelet coefficients are then coded using the zero-tree based SPIHT coder and transmitted through a channel. At the receiver end of the channel, these bits are decoded using the SPIHT decoder. The decoded wavelet coefficients are passed through the inverse masking filter block before reconstruction as image pixels using the inverse BNC-DWT block. The proposed scheme is also extended for the enhancement of the reconstructed medical images, such as the x-ray and ultrasound scan in the presence of additive noise information.

3.2.3 CSF masking model

In this section, we will discuss about the selection of a suitable CSF model for use in the generation of the mask values. The two important parameters of the HVS [54] for any given CSF model are the *contrast threshold* and the *contrast sensitivity*. *Contrast threshold* of the HVS is defined as the limit of visibility of a low contrast pattern before it becomes indistinguishable from the background field for a fixed value for luminance. The inverse of this parameter denotes the *contrast sensitivity*. The basic structure of the CSF model is bell-shaped or Gaussian in nature. The location of the peak in the curve depends on the viewing distance. For human perceivable signals like image and speech, contrast threshold (for images) or the limit of audibility (for speech signal) parameter needs to be considered during the development of perceptual-based processing sub-systems. This is due to the fact that the final processed signal is ultimately analyzed by a human observer. For the computation of the contrast threshold, the widely-used Mannos and Sakrison's CSF model [25] is used. This model defines a band-pass structure for the HVS and its analytical representation corresponds well with the human vision model at various spatial frequencies. Fig 3.3 shows the Mannos and Sakrison's CSF model defined by (2.15).

CSF masking is defined as the process of desensitizing the visual system to certain stimulus, using the values of contrast threshold. This is performed by providing weights based on desired perceptual importance to the wavelet-transformed coefficients, and subsequently inverting them after the decoding stage. A simplified representation of the proposed scheme, incorporating the CSF masking and inverse masking blocks, is shown in Fig 3.4. From the figure it can be seen that the CSF masking module is visualized as a

sub-band specific multiplier. The wavelet transformed image coefficients obtained from the BNC-DWT stage are masked or weighted relative to their perceptual importance in accordance with the visual sensitivity characteristics. This is obtained by multiplying the weights computed from the CSF curve with the wavelet coefficients specific to their
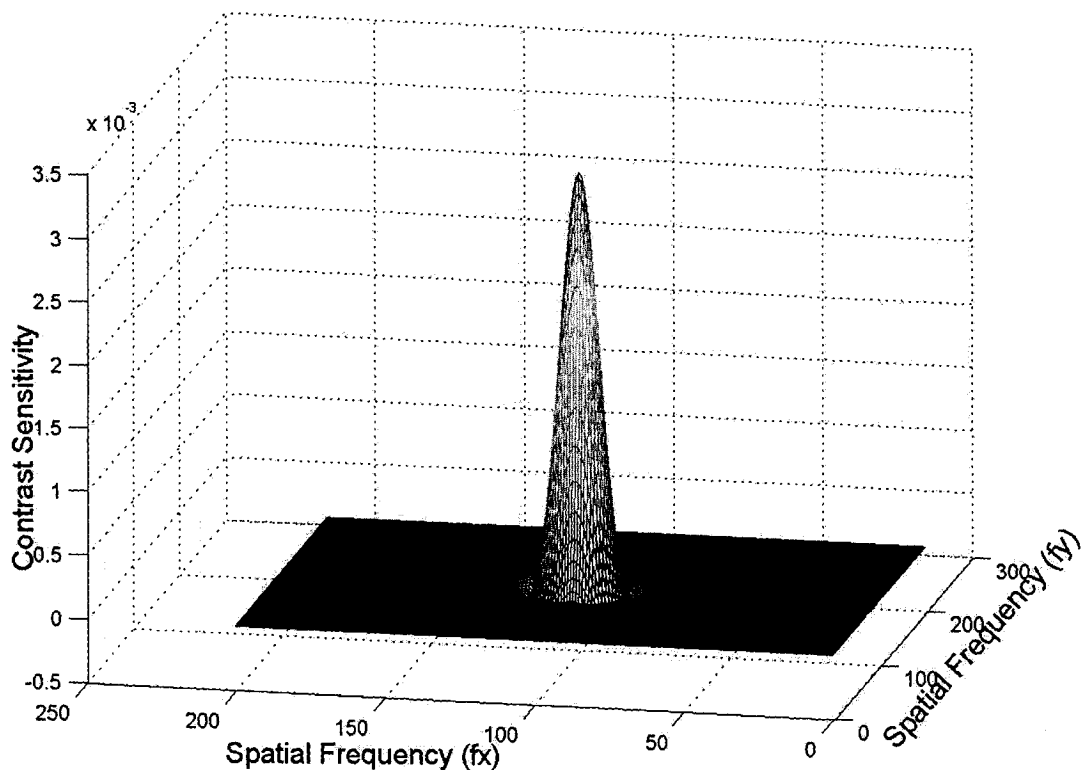
Figure 3-3   Mannos & Sakrison's CSF model curve

spatial sub-band. These coefficients are then coded and transmitted into the channel. At the receiver end, the received coded information is decoded and an inverted CSF mask is applied. The inverse mask filter is used so as to achieve an enhancement both in terms of the contrast and the perceptual quality of the reconstructed image. Moreover, the masking and inverse masking filters are chosen in such a way that they do not introduce any additional distortion in the compression system. The mask factors are determined directly

from the CSF curve using the band peak-average mask algorithm given in the following section. This technique is simple to implement and a luminance-independent processing is achieved.

3.2.4 Proposed band peak-average CSF mask (CSF-BPA mask)

We now describe a technique for obtaining the CSF weighting coefficients defined as the CSF band peak-average (CSF-BPA) mask. The CSF plotted for a viewing distance of 3 feet ($\simeq$ 1.5 x screen height) and background luminance of 100 cd/m^2 defined over a maximum spatial frequency of 64 pixels/degree is shown in Fig 3.3. In this masking method, the CSF curve is normalized and broken into m different spatial frequency bands. The value of m is equal to $n + 1$, where n is the number of decomposition levels used in the DWT-BNC stage. For a 5-level wavelet decomposition of the image, $n = 5$ and $m = 6$. The weight of the CSF mask for each sub-band is computed as the peak-average value of the curve specific to that sub-band as given in the following mask calculation algorithm. Finally, the HF weight factor is normalized to unity.
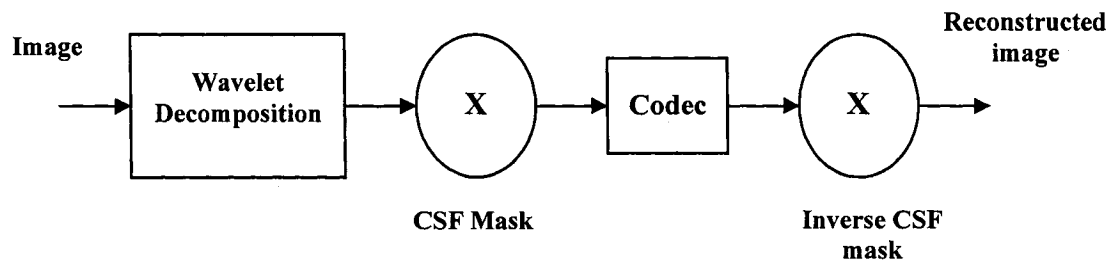
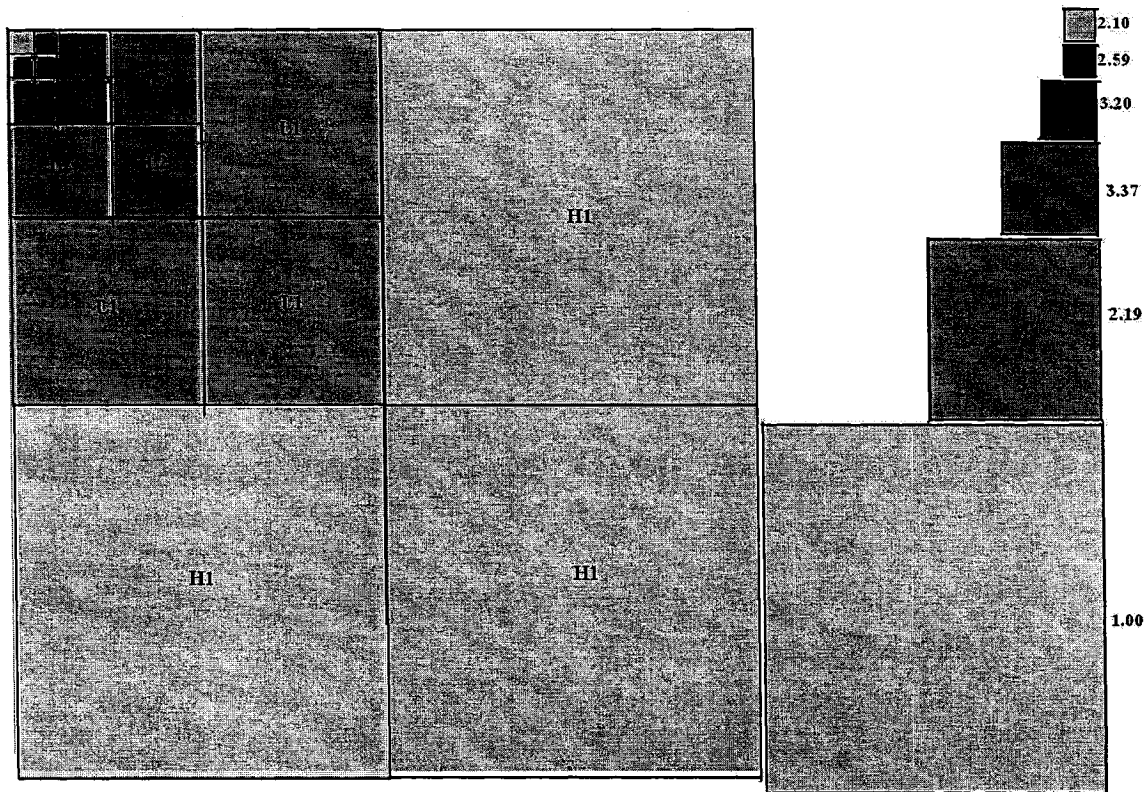Figure 3-4 Block diagram of CSF masking block for compression applications

54

Figure 3-5 6-weight band peak-average CSF mask model for a 5-level wavelet decomposition, where Ln represents the low frequency sub-band of an *n*-level DWT and Hn represents the high frequency sub-band of an *n*-level DWT

Algorithm: *Mask calculation*

1. Compute the CSF curve using the Mannos and Sakrison's transfer function [26] normalized for a luminance value of 100 cd/m^2 and a viewing distance of 3 feet.

2. Divide the curve into *m* different spatial frequency regions. Now, each region represents the part of the CSF within each of the approximation and detail sub-band.

3. Divide each of the *m* sub-bands further into *n* sub-levels.

4. Calculate the peak average of n values in each sub-band and represent the band-peak average values as p_n, p_{n-1}, ..., p_2, p_1 and q_1 respectively (starting from the origin of the curve).

5. The weights for the LH^1, HL^1, and HH^1 sub-bands are given by p_n.

6. Similarly the weights for the LH^2, HL^2, and HH^2 sub-bands are given by p_{n-1}. In each subsequent level, this scheme of weighting is followed for each of the band pass sub-bands.

7. At the final level the weight of the final low frequency band, LL^n is given by q_1.

8. The lowest weight, p_n, is taken as unity (since the HF is normalized to unity).

The 6-weight CSF-BPA mask values obtained for a 5-level image decomposition structure are shown in Fig 3.5. In the following section we discuss the proposed multiplierless design structure for the FPGA implementation of the lifted BNC 17/11 wavelet filter coefficients.

3.3 Proposed design for BNC 17/11 filter

In a wavelet-based image compression scheme, the selection of the filter bank is often crucial in determining the performance of the scheme, and a careful selection of the filter band enables superior quality for a broad class of images. For compression applications, a wavelet filter should possess finite filter length (FIR), enable perfect reconstruction of the input and meet the regularity constraints. An FIR filter is needed for proper hardware implementation. The approximation and the detail signals generated at each level should contain the necessary information for perfect reconstruction of the approximation signal at the next higher resolution level. For wavelet-based image compression schemes,

regularity is another important constraint as it influences the number of vanishing moments in the filter and aids in the reduction of oscillatory distortion. BNC wavelet filters [19, 20] provide high filter regularity and smoothness in the pass-band. These filters also have a linear phase response, which helps in the preservation of perceptually important image edges.

In the next sub-section, the convolution structures are briefly described following which the proposed design for the realization of BNC 17/11 wavelet filter [19] in the FPGA hardware is discussed in detail. The subscripts '*e*' and '*o*' refer to the corresponding even and odd components.

3.3.1 Realization of convolution structures

A convolution structure can be realized either in the form of non-polyphase or a polyphase structure, as shown in Fig 3.6.

(a) Non-polyphase structure

In a non-polyphase structure, in the analysis stage, decimation (or down-sampling) is performed after filtering and this discards half of the samples computed by the filters. Fig. 3.6 (a) shows the analysis side of a filter bank, where a filter $h[n] = (h_0, h_1, h_2, h_3 \ldots) \leftrightarrow H(z)$ is followed by a downsampling operation. This necessitates the filters to operate at double the input rate. In the synthesis phase, interpolation (or up-sampling) is performed before filtering, which means half of the filter multiplications are performed through zero operations. Thus, half of the mathematical operations are wasted. Hence, a non-polyphase structure is inefficient in terms of throughput, regardless of the direct form or cascade form implementation of the filters.
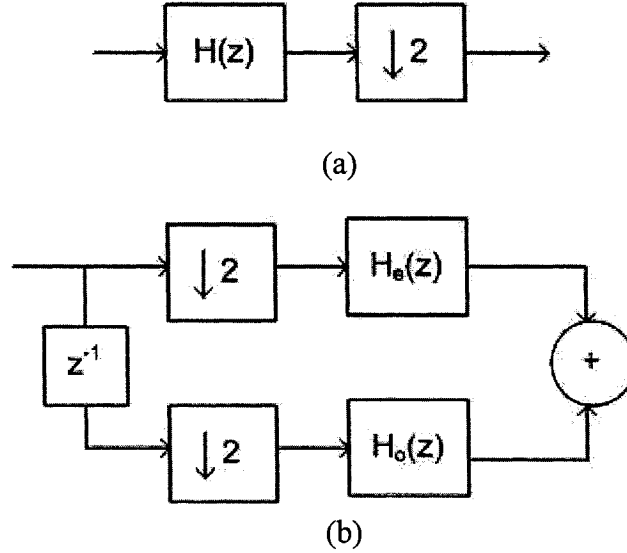
Figure 3-6 Direct form representation of analysis filters. (a) Non-polyphase structure.
(b) Polyphase structure.

(b) Polyphase structure

Polyphase structures were developed to overcome the unnecessary overhead faced with the use of the non-polyphase structures. In a polyphase structure (see Fig 3.6(b)), the filter is separated into odd and even powers as follows.

$$H(z) = h_0 + h_1 z^{-1} + h_2 z^{-2} + h_3 z^{-3} + \dots$$

$$= \left(h_0 + h_2 z^{-2} + h_4 z^{-4} + \dots \right) + z^{-1} (h_1 + h_3 z^{-2} + h_5 z^{-4} + \dots)$$

$$= H_e(z^2) + z^{-1} H_o(z^2) \tag{3.1}$$

Here,

$$H_e(z) \leftrightarrow \left(h_0, h_2, h_4, \dots \right) \quad \text{(even components } h_{2i})$$

$$H_o(z) \leftrightarrow \left(h_1, h_3, h_5, \dots \right) \quad \text{(odd components, } h_{2i+1})$$

where $i = 1, 2, 3, \dots, N/2 -1$. Once the even and odd components are separated, the order of the downsampling and the filter is exchanged, according to the first Noble identity [45].

58

The structure so obtained is computationally equivalent to the original, but is more efficient, since double throughput is obtained as the unnecessary *multiply-by-zero* operations are avoided. A similar transformation is applied on the synthesis side of the filter bank. Theoretically, such a structure provides twice the throughput in comparison with the original structure. In an FPGA implementation, the actual increase in throughput depends on many hardware timing parameters, and may not always equal the theoretical throughput. More detail about these two realization structures can be found in [14]. In the following section, the polyphase structure for the realization of the BNC wavelet filter is provided.

3.3.2 Proposed structure for BNC 17/11 filter using lifting coefficients

Figure 3.7 shows the non-polyphase structure for a two-channel filter biorthogonal bank. The analysis low pass and high pass filters are defined by the transfer functions $H(z)$ and $G(z)$ respectively. The low pass and high pass filters on the synthesis side are defined by $F(z)$ and $J(z)$.

Fig 3.8. shows the polyphase structure for a two-channel filter bank. Here, the input signal $X(z)$ is split into even and odd phases and is inputted into the corresponding polyphase form of the analysis filter. The analysis filters defined using polyphase representation are defined as follows.

$$H(z) = H_e(z) + z^{-1} H_o(z) \tag{3.2}$$

$$G(z) = G_e(z) + z^{-1} G_o(z) \tag{3.3}$$
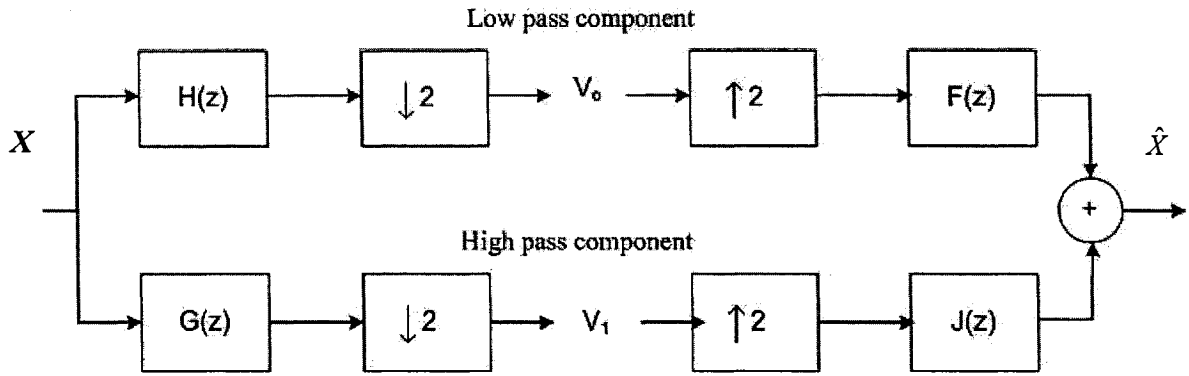
Figure 3-7 A non-polyphase structure for a two-channel biorthogonal filter bank

The output at the analysis stage is given by the values of $V_0(z)$ and $V_1(z)$.

$$\begin{bmatrix} V_0(z) \\ V_1(z) \end{bmatrix} = \begin{bmatrix} H_e(z) & H_o(z) \\ G_e(z) & G_o(z) \end{bmatrix} \begin{bmatrix} X_e(z) \\ z^{-1}X_o(z) \end{bmatrix} \qquad (3.4)$$

The reconstructed signal $\hat{X}(z)$ is obtained at the output of the synthesis stage.

$$\hat{X}(z) = [z^{-1} \quad 1] \begin{bmatrix} F_0(z^2) & J_o(z^2) \\ F_e(z^2) & J_e(z^2) \end{bmatrix} \begin{bmatrix} V_0(z^2) \\ V_1(z^2) \end{bmatrix} \qquad (3.5)$$

To obtain perfect reconstruction [14],

$$F_p(z).H_p(z) = z^{-1} \begin{bmatrix} 0 & 1 \\ z^{-1} & 0 \end{bmatrix} \qquad (3.6)$$

where

$$H_p(z) = \begin{bmatrix} H_e(z) & H_o(z) \\ G_e(z) & G_o(z) \end{bmatrix} \qquad (3.7)$$

and

$$F_p(z) = \begin{bmatrix} F_o(z) & J_o(z) \\ F_e(z) & J_e(z) \end{bmatrix} \qquad (3.8)$$

60

are the analysis and the synthesis polyphase matrices respectively, of the two-channel filter bank. When the polyphase matrices satisfy the perfect reconstruction condition (3.6), the reconstructed signal is a delayed version of the original signal.

$$x[n] = x[n - n_0]$$

For a BNC 17/11 wavelet, the analysis and synthesis filters are given by the following equations.

$$
\begin{aligned}
H(z) = h_0 &+ h_1(z + z^{-1}) + h_2(z^2 + z^{-2}) + h_3(z^3 + z^{-3}) + h_4(z^4 + z^{-4}) \\
&+ h_5(z^5 + z^{-5}) + h_6(z^6 + z^{-6}) + h_7(z^7 + z^{-7}) + h_8(z^8 + z^{-8})
\end{aligned}
\tag{3.9}
$$

$$
\begin{aligned}
G(z) = g_0 z &+ g_1(z^2 + 1) + g_2(z^3 + z^{-1}) + g_3(z^4 + z^{-2}) \\
&+ g_4(z^5 + z^{-3}) + g_5(z^6 + z^{-4})
\end{aligned}
\tag{3.10}
$$

$$
\begin{aligned}
F(z) = f_0 &+ f_1(z + z^{-1}) + f_2(z^2 + z^{-2}) + f_3(z^3 + z^{-3}) \\
&+ f_4(z^4 + z^{-4}) + f_5(z^5 + z^{-5})
\end{aligned}
\tag{3.11}
$$

$$
\begin{aligned}
J(z) = j_0 z &+ j_1(1 + z^{-2}) + j_2(z + z^{-3}) + j_3(z^2 + z^{-4}) + j_4(z^3 + z^{-5}) \\
&+ j_5(z^4 + z^{-6}) + j_6(z^5 + z^{-7}) + j_7(z^6 + z^{-8}) + j_8(z^7 + z^{-9})
\end{aligned}
\tag{3.12}
$$
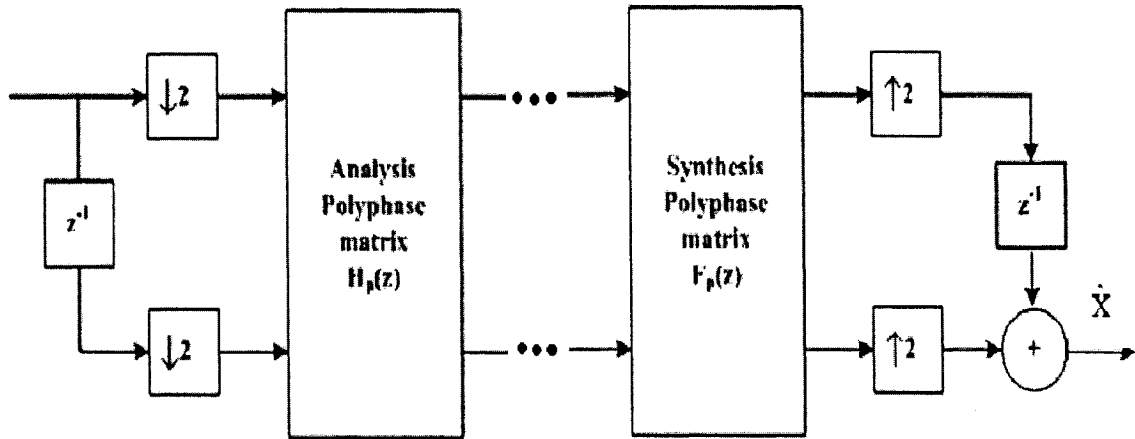
Figure 3-8 A polyphase structure for a two-channel filter bank

The values of h[n] and g[n] are provided in Table 2.2 and the synthesis filters are defined as follows.

$$j[n] = (-1)^n h[n]$$

$$f[n] = (-1)^n g[n]$$

Substituting the values of $H(z)$, $G(z)$, $J(z)$ and $F(z)$ in (3.7) and (3.8),

$$H_p(z) = \begin{bmatrix} \{h_8(z^4 + z^{-4}) + h_6(z^3 + z^{-3}) + & \{h_7(z^4 + z^{-3}) + h_5(z^3 + z^{-2}) + \\ h_4(z^2 + z^{-2}) + h_2(z + z^{-1}) + h_0\} & h_3(z^2 + z^{-1}) + h_1(z+1)\} \\ \{g_5(z^3 + z^{-2}) + g_3(z^2 + z^{-1}) & \{g_4(z^3 + z^{-1}) + g_2(z^2 + 1) \\ + g_1(1+z)\} & + g_0 z\} \end{bmatrix} \quad (3.13)$$

$$F_p(z) = \begin{bmatrix} \{f_1(1+z) + f_3(z^{-1} + z^2) & \{j_0 + j_2(z^{-1} + z) + j_4(z^{-2} + z^2) + \\ + f_5(z^{-2} + z^3)\} & j_6(z^{-3} + z^3) + j_8(z^{-4} + z^4)\} \\ \{f_0 + f_2(z + z^{-1}) & \{j_1(1 + z^{-1}) + j_3(z + z^{-2}) \\ + f_4(z^2 + z^{-2})\} & + j_5(z^2 + z^{-3})\} \end{bmatrix} \quad (3.14)$$

The factorized filter matrix $H_p(z)^T$ is obtained by decomposing the filters $\{H(z), G(z)\}$ using the algorithm proposed in [59].

$$H_p(z)^T = \prod_{i=1}^{m} \begin{bmatrix} 1 & s_i(z) \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ t_i(z) & 1 \end{bmatrix} \begin{bmatrix} k & 0 \\ 0 & 1/k \end{bmatrix} \quad (3.15)$$

This factorization is not unique, and a number of $\{s_i(z), t_i(z)\}$ pairs that lead to the same analysis and synthesis polyphase matrices can be obtained. We have factorized the BNC 17/11 polyphase matrix into eight two-tap "predict" and "update" filters with a scaling step at the end.

62

$$H_p(z)^T = \begin{bmatrix} 1 & \alpha_1(1+z^{-1}) \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 \\ \alpha_2(1+z) & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & \beta_1(1+z^{-1}) \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 \\ \beta_2(1+z) & 1 \end{bmatrix} \cdot$$

$$\begin{bmatrix} 1 & \gamma_1(1+z^{-1}) \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 \\ \gamma_2(1+z) & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & \delta_1(1+z^{-1}) \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 \\ \delta_2(1+z) & 1 \end{bmatrix} \cdot \begin{bmatrix} \varsigma & 0 \\ 1 & 1/\varsigma \end{bmatrix}$$

(3.16)

where $\alpha_1, \alpha_2, \beta_1, \beta_2, \gamma_1, \gamma_2, \delta_1, \delta_2$ and ς are the lifting [see Appendix I] coefficients

obtained during the factorization and their values are listed in Table 3.1. From (3.16), the

scheme for the realization of the BNC 17/11 forward wavelet transform has four lifting

steps (four pairs of predict P_n and update U_n sequences) as shown in Fig 3.9. The z-

transforms of the predict operators, P_1, P_2, P_3 and P_4 and the update operators, U_1, U_2, U_3

and U_4 are listed below.

$$P_1 = \alpha_1 * (1+z)$$

$$U_1 = \alpha_2 * (1+z^{-1})$$

(3.17a)

$$P_2 = \beta_1 * (1+z)$$

$$U_2 = \beta_2 * (1+z^{-1})$$

(3.17b)

$$P_3 = \gamma_1 * (1+z)$$

$$U_3 = \gamma_2 * (1+z^{-1})$$

(3.17c)

$$P_4 = \delta_1 * (1+z)$$

$$U_4 = \delta_2 * (1+z^{-1})$$

(3.17d)
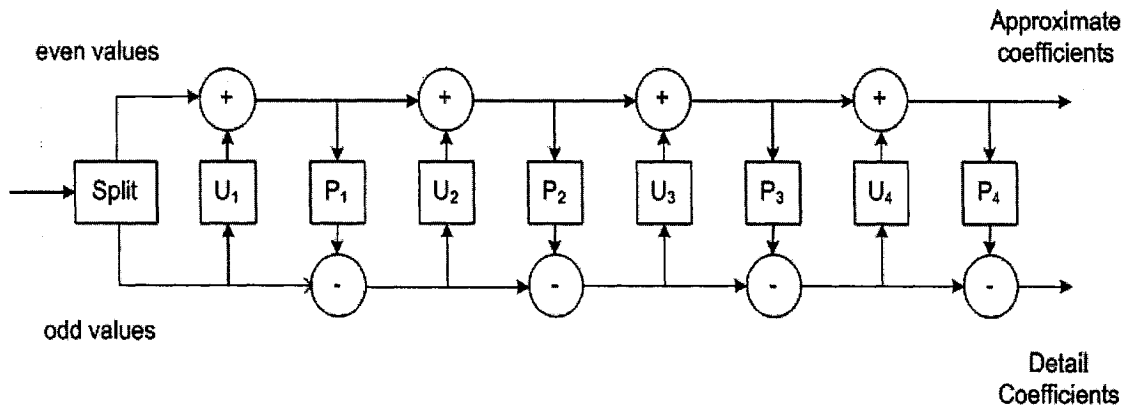
Figure 3-9 A single lifting step for BNC 17/11 wavelet filter

Table 3-1   Quantized and un-quantized lifted coefficients for BNC 17/11 wavelet filter

Symbol	Lifted filter coefficients	
	Un-quantized representation	Quantized representation
$\alpha 1$	- 1.5	-3/2 = -1.5
$\alpha 2$	0.187499	3/16 = 0.1875
$\beta 1$	-0.36673	-3/8 = -0.375
$\beta 2$	-0.12843	-1/8 = -0.125
$\gamma 1$	0.7301155	3/4 = 0.75
$\gamma 2$	-0.440046	-7/16 = -0.4375
$\delta 1$	0.2181397	7/32 = 0.21875
$\delta 2$	0.91063233	29/32 = 0.90625
ζ	1.19	5/4 = 1.25

The design structure for the BNC 17/11 analysis filter bank using the lifting scheme, i.e.,

implementation for the matrix $H_p(z)$, is as shown in Fig. 3.10. For an input signal x_i let x_{2i}

64

(x_{2i+1}) denote the even (odd) indexed samples. The approximation and detail coefficients generated during each of the predict (update) stage at the analysis phase of the filter bank is represented as a_i^m (d_i^m) representing the low-pass (high-pass) coefficients, where $m = 0$, 1, 2, 3, 4. The equations for obtaining these values are given as follows.

1) Splitting phase:

$$d_i^0 = x_{2i+1} \tag{3.18}$$

$$a_i^0 = x_{2i} \tag{3.19}$$

2) Predict & Update phase:

$$d_i^1 = d_i^0 - \alpha_1 * (a_i^0 + a_{i+1}^0) \tag{3.20a}$$

$$a_i^1 = a_i^0 - \alpha_2 * (d_{i-1}^1 + d_i^1) \tag{3.20b}$$

$$d_i^2 = d_i^1 - \beta_1 * (a_i^1 + a_{i+1}^1) \tag{3.21a}$$

$$a_i^2 = a_i^1 - \beta_2 * (d_{i-1}^2 + d_i^2) \tag{3.21b}$$

$$d_i^3 = d_i^2 - \gamma_1 * (a_i^2 + a_{i+1}^2) \tag{3.22a}$$

$$a_i^3 = a_i^2 - \gamma_2 * (d_{i-1}^2 + d_i^2) \tag{3.22b}$$

$$d_i^4 = d_i^3 - \delta_1 * (a_i^3 + a_{i+1}^3) \tag{3.23a}$$

$$a_i^4 = a_i^3 - \delta_2 * (d_{i-1}^3 + d_i^3) \tag{3.23b}$$

3) Scaling phase:

$$d_i = d_i^2 / \varsigma \tag{3.24a}$$

$$a_i = \varsigma * a_i^2 \tag{3.24b}$$
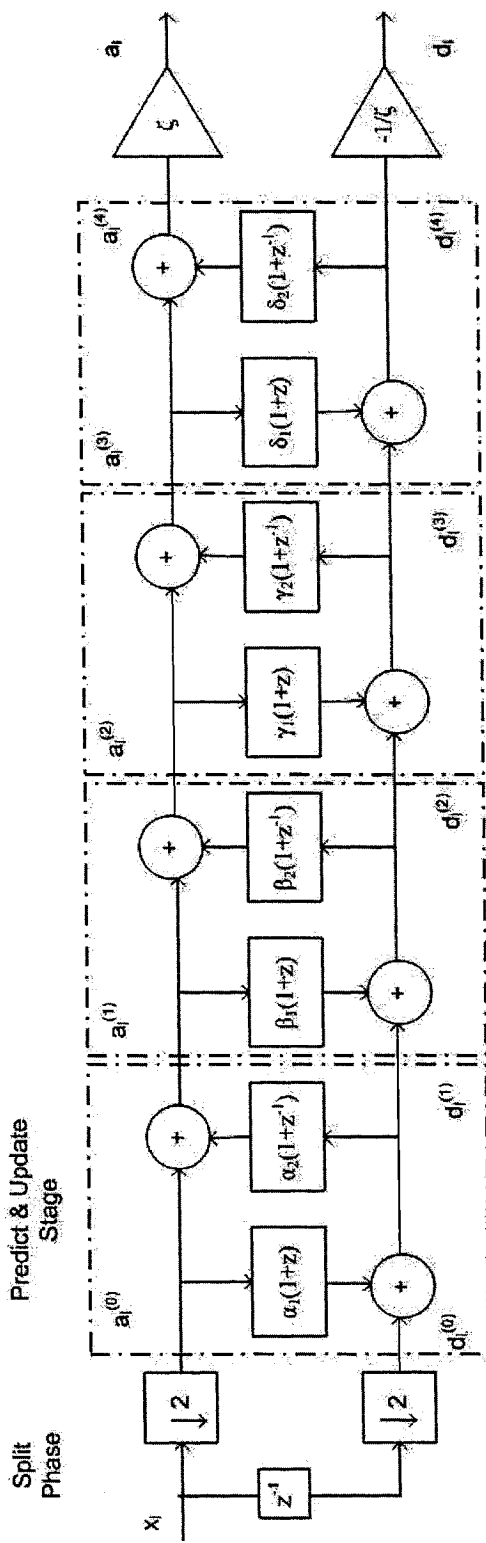
Figure 3-10   Analysis structure of lifted BNC 17/11 wavelet filter

The synthesis stage of the filter bank is obtained by inverting the scaling coefficients and by reversing the sequence of the predict and update stages as shown in Fig. 3.11. From the figure, it can be seen that if the scaling factors ς and $1/\varsigma$ are ignored and the synthesis stage immediately follows the analysis stage (i.e., no compression), and the output of the analysis stage is exactly inverted regardless of the values of $\alpha_1, \alpha_2, \beta_1, \beta_2, \gamma_1, \gamma_2, \delta_1$ and δ_2. For implementation in fixed-point hardware the lifted coefficients are approximated. This approximation is done in such a fashion that a multiplierless architecture can be realized. The approximated lifting coefficients are then used for the realization of the filter structure. Jer Min Jou [56] proposed a hardware architecture for the implementation of biorthogonal wavelets. This structure incorporates symmetry property and uses polyphase decomposition. Considering the advantages offered in terms of the computational efficiency and the ease for realization offered by Jer Min Jou's architecture, we have extended this architecture for the realization of the odd length nearly-orthogonal filters (BNC 17/11 filter to be specific). In this thesis, an architecture for the realization of the BNC 17/11 wavelet filter using a pipelining approach is proposed. The lifting approach for the BNC 17/11 wavelet filter has been discussed in Section 3.3.2.

The complete structure for the BNC 17/11 filter is shown in Fig 3.12. The main advantage of this scheme is that the same structure can be used for both the forward and reverse DWT operations in view of the addition of the multiplexer unit. Further, the use of pipelining facilitates the reduction in computation time. The direction of operation (forward or reverse) is determined by the multiplexer unit using a select input signal. When the input to the select line is '0', a forward DWT operation is performed and the
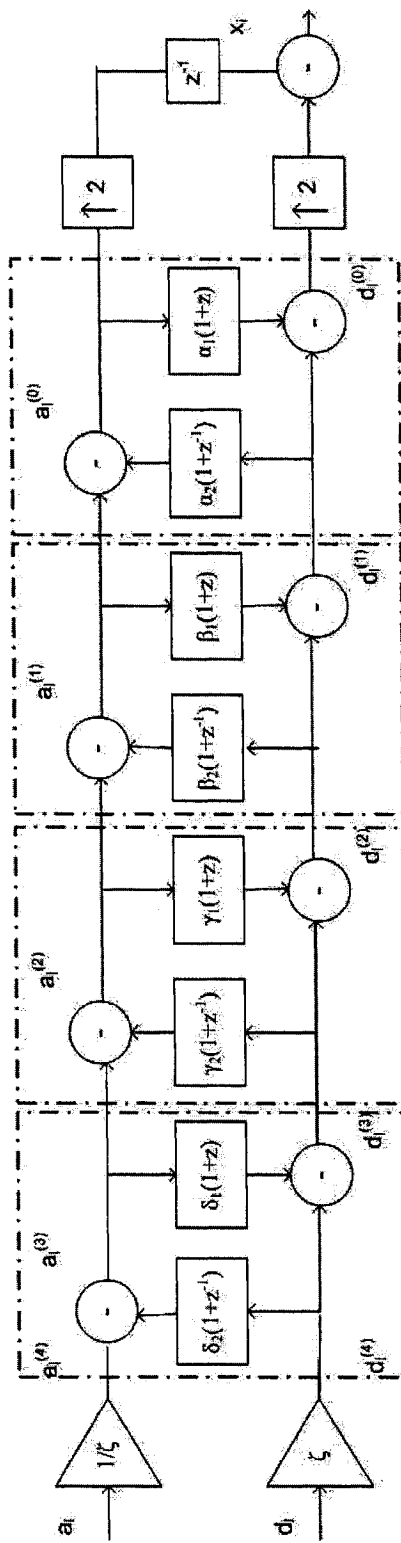
Figure 3-11   Synthesis structure of lifted BNC 17/11 wavelet filter
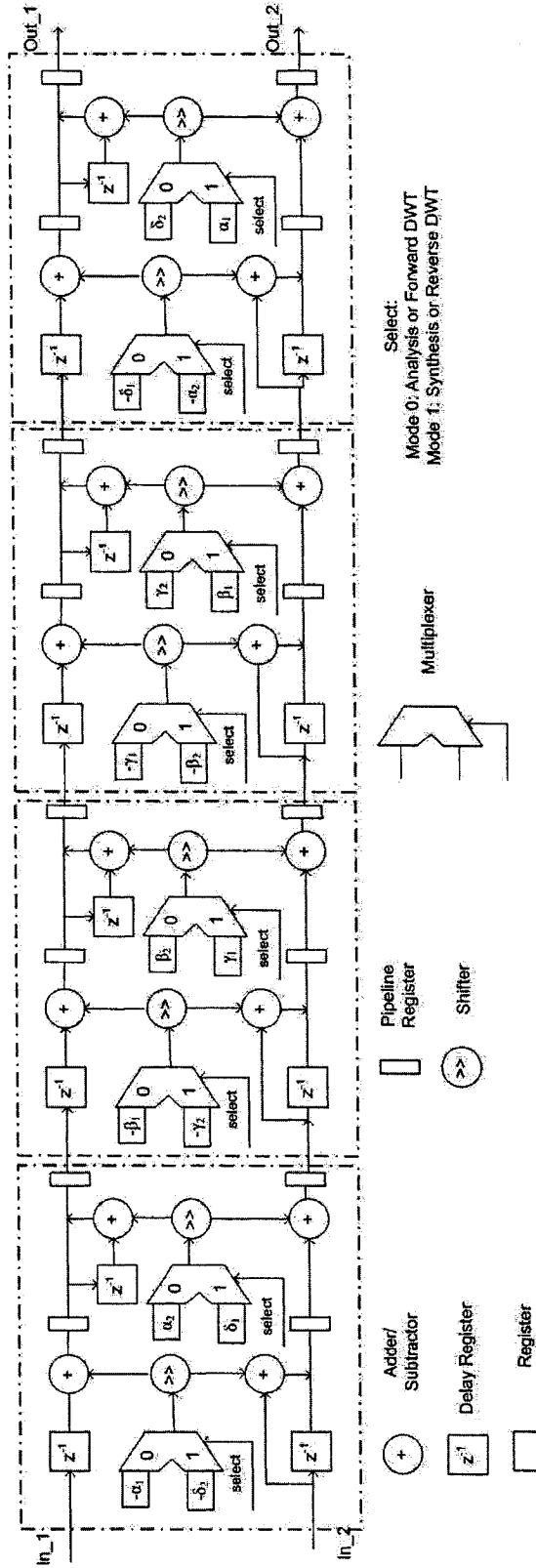
Figure 3-12 Architecture for implementation of lifted BNC 17/11 wavelet filter

approximate and detail coefficients are obtained across the output lines Out_1 and Out_2 respectively. For a select line input of '1', the lifted coefficients responsible for the reverse DWT operation are selected. At each clock cycle, there are two input samples entering the system, In_1 and In_2. While performing the DWT, the even and odd samples are given across the input lines, In_1 and In_2, and for the case of the inverse DWT, the approximate and detail coefficient values are used at the input. Once the computation is performed, the final results (wavelet coefficients) are obtained as the output through the data lines, namely, out_1 and out_2.

This design can be implemented either on a FPGA or ASIC chip. In this design, a total of 8 shifters/ multipliers, 16 adder units, 36 registers (8 for coefficients, 16 for pipelining and 1 for delay) and 8 multiplexers are employed.

3.4 Potential applications of the proposed scheme

Since the mask values are calculated in advance, no side information needs to be transmitted between the source and the destination. This, in turn, avoids the possibility of bit overhead and facilitates a low entropy representation of the image. Potential areas of application of the proposed scheme are as follows.

➢ Medical applications such as storage and transmission of medical data such as X-ray, MRI and ECG

➢ Biosignal processing

➢ Internet multimedia, audio and speech compression applications

➢ Video compression and perceptual watermarking

70

3.5 Summary

In this chapter, we have first studied the limitations of the conventional compression system. Then, in order to overcome some of these limitations and to obtain an enhanced compression performance as well as to facilitate a multiplication-free hardware realization, we have developed a compression scheme for digital images. In place of the conventional wavelets, the nearly-orthogonal wavelet filters have been used in this scheme in order to obtain an improvement in the compression efficiency. A CSF masking method (CSF-BPA mask) has been proposed for achieving an enhancement in the perceptual visual quality of the reconstructed image. This masking module has been used in conjunction with the nearly-orthogonal discrete wavelet transform in the proposed scheme. Further, this scheme has been extended for perceptually-enhanced reconstruction of digital medical images. Finally, a multiplierless design structure for fixed-point realization of the BNC 17/11 wavelet filter in a FPGA or ASIC hardware has also been proposed.

Chapter 4

Simulation Results

In Chapter 3, a new scheme for the enhancement of the perceptual quality of the reconstructed compressed grayscale images has been proposed. The proposed scheme uses the nearly-orthogonal discrete wavelet transform (DWT) integrated with the contrast sensitivity function (CSF) masking stage. The mask values have been derived from the Mannos and Sakrison's CSF model [26] curve using the proposed band peak-average (BPA) mask method. The nearly-orthogonal DWT uses either of the recently explored biorthogonal nearly-coiflet (BNC) wavelets, namely, the symmetric BNC 17/11 wavelet [19] or the asymmetric BNC 22/14 wavelet filter [20].

In this chapter, the performance of the proposed scheme is studied and compared with that of the previously used CSF-DWT mask [52] and the conventional scheme employing the standard biorthogonal 9/7 wavelet filter [18]. In order to facilitate this, the proposed scheme is tested on a variety of standard test and medical images. The characteristics of the chosen test images are quite distinct in terms of their high frequency and low frequency content. The images considered for compression analysis fall into the following broad categories.

(a) Natural images - *Lena*, *Barbara* and *Baboon*

(b) High-detail images - *Pot*, *Cambridge*, *Boat* and *Text*

(c) Medical scan images - *MRI*, *Ultrasound scan*, *X-Ray* and *Diatoms*.

The original test images can be found in Appendix - II for reference. The input image is first corrupted with additive white Gaussian noise (AWGN) before processing. The noise-added image is then wavelet decomposed and coded using the SPIHT coder [41]. At the receiver end, the transmitted bits are decoded with the use of a SPIHT decoder and the image is reconstructed using the inverse DWT. Simulation is carried out using MATLAB and the wavelet toolbox. The results are collected by reconstructing the image using each of the three wavelet filters, namely, CDF 9/7, BNC 17/11 and BNC 22/14 while switching between the three CSF modes (CSF off, CSF-DWT mask on & CSF-BPA mask on) under different noise variances and compression ratios. The performance of the scheme is studied using both objective (quantitative) and subjective (qualitative) assessment metrics. The calculated value of the peak signal to noise ratio (PSNR) is used for the objective assessment of the quality of the reconstructed image. The subjective assessment is conducted by carrying out a double blind test procedure on the reconstructed images. Further, a no-reference quality index [57] is calculated to verify the correlation between the objective and subjective test results. In addition to the above mentioned assessments, the principal component or the eigen vector analysis is performed by calculating and plotting the values of eigen vectors for the original and the reconstructed images.

4.1 Quantitative Assessment

The PSNR, measured in decibels (dB), is used as the metric for assessment of the quantitative or objective performance of the proposed scheme incorporating the nearly-orthogonal wavelet transform and the deduced CSF-BPA mask. The PSNR value is calculated with reference to the original test image. Tables 4-1 and 4-2 give the

quantitative results of the simulation indicating the PSNR and the computation time of the reconstructed *Ultrasound scan* and a zoomed portion of noise-free *MRI* images respectively without the application of the CSF mask. From these tables, it is seen that, for low and medium bit rates, the BNC wavelets, in addition to providing a higher PSNR, take less computation time in comparison with the standard CDF 9/7 wavelet. For a compression ratio of 0.5 bpp, a maximum PSNR improvement of 0.13 dB is achieved for the reconstruction of the *Ultrasound scan* image and 0.26 dB for the *MRI* image using the BNC 22/14 wavelet. However, it is evident that, among the BNC wavelets, the BNC 17/11 wavelet facilitates a faster processing of the image for a variety of low bit rates as compared to the BNC 22/14 wavelet.

Tables 4-3, 4-4 and 4-5 give the PSNR results for the reconstructed set of eleven test images compressed at different rates and corrupted with noise variances 8, 13 and 26, respectively. The values in bold font indicate the highest values of PSNR and the corresponding filters for each of the images. In this assessment metric, the performance is estimated based on the calculated PSNR value. In other words, higher the PSNR better is the performance.

From Table 4-3, corresponding to an added noise variance of 8, it is seen that, for three of the eleven reconstructed images under consideration (*Barbara*, *Diatoms* and *Text*), the application of the mask has resulted in a reduction in the overall PSNR, with a maximum of 2 dB reduction for the *Diatoms* image. For the other eight images, the application of the mask shows a significant improvement in the PSNR value. A maximum PSNR improvement of 2 dB is achieved for the case of *Lena* with the application of CSF-BPA mask integrated with the BNC wavelet transform. Further, it is noticed that the CSF-BPA

mask outperforms the previously used CSF-DWT mask by achieving an increase in PSNR of 1 dB recorded for the *Lena* test image.

From Table 4-4, corresponding to an added noise variance of 13, a pattern similar to that of the case when the images are corrupted with a noise variance of 8, is observed. Here also, the application of the mask has resulted in a reduction in the overall PSNR for three

Table 4-1 PSNR and computation time results for noise-free *Ultrasound scan* image reconstructed without the application of the CSF mask.

bpp	CDF 9/7		BNC 17/11		BNC 22/14	
	PSNR (dB)	Time (sec)	PSNR (dB)	Time (sec)	PSNR (dB)	Time (sec)
0.125	74.1149	8	74.1196	8	74.2007	8
0.250	74.5788	28	74.6621	30	74.6660	34
0.500	74.6020	124	75.7173	105	74.7300	109

Table 4-2 PSNR and computation time results for *MRI* image reconstructed without the application of the CSF mask.

Bpp	CDF 9/7		BNC 17/11		BNC 22/14	
	PSNR (dB)	Time (sec)	PSNR (dB)	Time (sec)	PSNR (dB)	Time (sec)
0.125	74.2818	9	74.3357	9	74.3530	8
0.250	74.9573	23	75.0679	22	75.2126	21
0.500	75.0950	95	75.1681	91	75.2628	99

Table 4-3 PSNR results for reconstructed set of standard test and medical images corrupted with zero-mean AWGN of variance 8.

bpp	No mask			CSF-DWT mask			CSF-BPA mask		
	9/7	17/11	22/14	9/7	17/11	22/14	9/7	17/11	22/14
Barbara									
0.20	69.6499	69.7685	69.7149	70.1895	70.1067	70.2625	70.6087	70.6970	**70.8241**
0.40	68.8360	68.9458	68.8582	70.3336	70.3722	70.4608	70.7108	70.7941	**70.9602**
0.60	67.4849	67.5481	67.6053	70.3117	70.3565	70.3532	70.8170	70.8864	**71.0352**
Lena									
0.20	71.1971	71.3219	71.2999	72.4732	72.4754	72.5367	73.3083	**73.6664**	73.5721
0.40	69.4504	69.5244	69.4529	72.2369	72.3386	72.4833	73.1469	73.2993	**73.4696**
0.60	67.6766	67.7643	67.7649	72.0879	72.154	72.1667	73.1653	**73.5499**	73.4140
Baboon									
0.20	69.7623	69.9090	**69.9791**	68.8520	68.9650	69.0251	68.1919	68.2803	68.0014
0.40	68.5789	**68.7038**	68.7019	67.3087	67.3340	67.3722	67.3776	67.4219	67.4670
0.60	66.7994	66.9113	**66.9636**	65.3169	65.3946	65.4596	65.3164	65.3764	65.3769
Boat									
0.20	70.5011	70.6857	70.5160	71.2553	71.2110	71.2355	71.6405	**71.7300**	71.6238
0.40	69.1516	69.2801	69.2496	71.3266	71.3962	71.2966	71.6885	**71.8175**	71.7626
0.60	67.5613	67.6521	67.6948	71.2175	71.2874	71.1747	71.7205	**71.8086**	71.7596
Diatoms									
0.20	69.1957	**69.2753**	69.2047	67.3439	67.4527	67.3994	67.1219	67.2046	67.0369
0.40	68.9303	69.0516	**69.0688**	67.6710	67.7213	67.6294	67.3727	67.4570	67.2889
0.60	67.7633	**67.8235**	67.7855	66.6632	66.7275	66.6313	66.4007	66.4797	66.3137
Pot									
0.20	70.8925	**71.0772**	71.0460	70.5006	70.5653	70.5906	70.7802	70.8303	70.7497
0.40	69.5769	69.7066	69.5955	70.6620	70.8017	70.7050	70.8332	70.9429	**70.9564**
0.60	67.9309	67.9868	67.9002	70.5575	70.6482	70.5925	70.8551	70.9350	**70.9420**
Text									
0.20	67.0543	**66.9852**	66.9488	65.1699	65.1935	65.1330	65.0673	65.0525	65.0062
0.40	69.0999	**68.9733**	68.8836	65.6008	65.5287	65.4447	65.3886	65.3533	65.3184
0.60	69.0224	**68.9787**	68.8989	65.7062	65.5866	65.6590	65.5536	65.5014	65.4581
Ultrasound scan									
0.20	72.1091	72.2155	72.2563	72.5150	72.6316	72.7228	73.1207	73.2320	**73.3112**
0.40	69.8032	69.9422	70.0640	72.1897	72.3060	72.5538	73.0914	73.1956	**73.2983**
0.60	68.2471	68.3011	68.2386	72.2685	72.3541	72.4448	73.1169	73.2011	**73.3321**
MRI									
0.20	72.4972	72.5141	72.3810	73.0109	73.0314	73.0631	73.2720	**73.3497**	73.0491
0.40	70.4430	70.4171	70.2446	72.5425	72.4886	72.5558	73.1571	**73.2048**	73.1065
0.60	68.4846	68.4603	68.3761	72.4315	72.4501	72.4790	73.0491	**73.1897**	73.1081
X-Ray									
0.20	**70.8177**	70.7750	70.5652	69.6789	69.7256	69.8312	69.9760	69.9162	69.9341
0.40	69.2126	69.1553	68.9642	69.5811	69.5587	69.7292	69.9132	69.9192	**69.9922**
0.60	67.5927	67.5511	67.4631	69.5902	69.6205	69.7579	69.8805	69.9235	**69.9954**
Cambridge									
0.20	68.4995	68.5916	68.7418	69.6933	69.7898	69.8988	69.7941	**69.8155**	69.8003
0.40	68.2339	68.3192	68.3127	69.8959	69.9667	69.9706	70.0369	**70.1207**	70.1142
0.60	67.2406	67.2997	67.2873	69.7651	69.8175	69.8091	70.0804	70.1321	**70.1547**

Table 4-4 PSNR results for reconstructed set of standard test and medical images corrupted with zero-mean AWGN of variance 13.

bpp	No mask			CSF-DWT mask			CSF-BPA mask		
	9/7	17/11	22/14	9/7	17/11	22/14	9/7	17/11	22/14
Barbara									
0.20	68.8529	**68.8591**	68.8499	67.6738	67.7776	67.9041	67.6291	67.6301	67.4943
0.40	67.6094	**67.7407**	67.6680	66.9432	66.9517	67.1094	66.5253	66.6152	66.6043
0.60	65.4431	**65.5259**	65.4611	64.3204	64.3336	64.4534	64.0118	64.0842	64.2394
Lena									
0.20	70.4063	70.4266	70.2626	71.6622	71.7812	71.9662	72.2373	**72.7405**	72.6305
0.40	66.7354	66.8164	66.6859	71.1647	71.1974	71.3654	72.2826	**72.9975**	72.5566
0.60	65.6221	65.6900	65.6477	70.7208	70.7534	70.9146	71.9785	**72.2641**	72.2370
Baboon									
0.20	67.5443	67.6225	67.5628	68.5125	68.5661	68.5887	68.5939	**68.6982**	68.6733
0.40	65.1723	65.2114	65.2034	68.5688	68.6475	68.6366	68.8879	68.9432	**68.9643**
0.60	64.7510	64.8127	64.8180	68.5242	68.6120	68.6764	68.9646	69.0189	**69.0577**
Boat									
0.20	69.3532	69.4347	69.3775	70.5083	70.5979	70.6634	70.7193	**70.8781**	70.8554
0.40	66.4671	66.5549	66.4620	70.1981	70.2589	70.2335	70.8549	**70.9539**	70.9074
0.60	65.4756	65.5488	65.5383	69.8971	69.9701	69.9548	70.6576	**70.7380**	70.7374
Diatoms									
0.20	67.5203	67.6830	**67.7851**	65.9254	66.0827	66.0393	65.7227	65.7942	65.6142
0.40	66.7332	**66.9437**	66.9054	66.0844	66.1348	66.1137	65.9017	65.9723	65.7788
0.60	65.9257	**66.0781**	66.0769	66.0393	66.0144	66.0049	65.8938	65.9654	65.7695
Pot									
0.20	69.3812	69.4757	69.3651	69.9958	70.0241	69.9980	69.9748	**70.1052**	70.1011
0.40	66.9786	67.0463	67.0371	69.7980	69.8461	69.7108	70.1271	**70.1984**	70.1785
0.60	65.9344	66.0095	66.0158	69.4680	69.5394	69.4559	69.9390	**70.0183**	70.0041
Text									
0.20	**65.8542**	65.8155	65.8185	64.7545	64.7511	64.7144	64.6299	64.6304	64.6154
0.40	**66.9045**	66.8405	66.8047	64.7782	64.7923	64.7920	64.8305	64.7883	64.7507
0.60	**66.9184**	66.8733	66.8212	65.0326	65.0310	65.0461	64.9985	64.9945	64.9781
Ultrasound scan									
0.20	70.2306	70.1842	69.9735	71.0797	71.2717	71.3933	71.9767	**72.0410**	71.8980
0.40	67.3978	67.4803	67.4665	70.9158	71.0046	71.0431	71.8186	71.9439	**71.9568**
0.60	66.0672	66.1622	66.1752	70.5618	70.6458	70.7873	71.7547	71.7858	**71.7873**
MRI									
0.20	69.3799	69.4599	69.4567	70.8536	70.8884	70.9067	71.2567	**71.2897**	71.2544
0.40	67.9778	68.0340	68.0138	70.4717	70.5376	70.6123	71.0296	71.0588	**71.0594**
0.60	66.5577	66.6546	66.6023	70.3622	70.4276	70.4297	71.0135	**71.0439**	71.0343
X-Ray									
0.20	67.7960	67.8178	67.7701	68.4290	68.5015	68.5020	68.6232	68.6713	**68.7254**
0.40	66.6671	66.6781	66.7221	68.2761	68.3652	68.4089	68.5542	68.5946	**68.6844**
0.60	65.7016	65.6911	65.6826	68.2064	68.2490	68.3643	68.5620	68.5940	**68.6310**
Cambridge									
0.20	68.4783	68.5438	68.5055	69.2817	69.4043	69.4019	69.3689	69.4525	**69.4937**
0.40	65.5895	65.5756	65.5107	69.1309	69.1861	69.1667	69.4992	**69.5422**	69.5383
0.60	64.9310	64.9570	64.9448	68.9463	68.9855	68.9683	69.4442	69.4969	**69.5291**

Table 4-5 PSNR results for reconstructed set of standard test and medical images corrupted with zero-mean AWGN of variance 26.

bpp	No mask			CSF-DWT mask			CSF-BPA mask		
	9/7	17/11	22/14	9/7	17/11	22/14	9/7	17/11	22/14
Barbara									
0.20	59.6835	**59.7797**	59.7506	59.7185	59.7306	59.7370	59.7017	59.7062	59.7125
0.40	60.0206	**60.0972**	60.0884	60.0733	60.0972	60.0905	59.9544	59.9625	59.9699
0.60	60.6056	**60.6856**	60.6790	60.2698	60.2744	60.2807	60.1157	60.1160	60.1233
Lena									
0.20	59.6390	59.6695	59.6887	59.6852	59.6891	**59.7076**	59.6715	59.6772	59.6774
0.40	59.9146	60.0066	60.0253	60.0311	**60.0536**	60.0516	59.9166	59.9200	59.9282
0.60	60.5190	60.5770	**60.5800**	60.2262	60.2288	60.2372	60.0739	60.0711	60.0813
Baboon									
0.20	59.3032	59.3394	59.3490	59.3465	59.3549	**59.3651**	59.3092	59.3118	59.3157
0.40	59.6484	59.6969	59.7222	59.7181	59.7432	**59.7368**	59.5773	59.5865	59.6080
0.60	60.1765	60.2384	**60.2402**	59.9207	59.9255	59.9233	59.7385	59.7437	59.7464
Boat									
0.20	59.7652	59.8001	**59.8232**	59.8023	59.8042	59.8094	59.7745	59.7839	59.7768
0.40	60.1060	60.1684	**60.1738**	60.1473	60.1644	60.1693	60.0241	60.0292	60.0326
0.60	60.6961	**60.7662**	**60.7662**	60.3481	60.3493	60.3477	60.1793	60.1842	60.1864
Diatoms									
0.20	60.6290	60.6707	**60.6856**	60.4077	60.4261	60.4226	60.3511	60.3714	60.3652
0.40	61.0075	61.0988	**61.1256**	60.7428	60.7608	60.7916	60.6011	60.6101	60.6074
0.60	61.7205	61.7855	**61.7886**	60.9492	60.9663	60.9754	60.7791	60.7789	60.7636
Pot									
0.20	60.0972	60.1301	**60.1455**	60.0928	60.0952	60.0966	60.0606	60.0736	60.0683
0.40	60.4092	60.4953	**60.5253**	60.4322	60.4425	60.4619	60.3085	60.3148	60.3178
0.60	61.0555	61.1304	**61.1466**	60.6411	60.6433	60.6400	60.4731	60.4727	60.4729
Text									
0.20	61.0256	61.0498	**61.0632**	60.8639	60.8812	60.8751	60.8396	60.8425	60.8407
0.40	61.4769	61.5228	**61.5510**	61.1808	61.1943	61.1871	61.0730	61.0791	61.0730
0.60	62.2123	**62.2373**	62.2272	61.3852	61.3990	61.3868	61.2907	61.2773	61.2562
Ultrasound scan									
0.20	60.2441	60.2781	60.2896	60.3052	60.3138	**60.3300**	60.3004	60.2999	60.2983
0.40	60.5857	60.6356	60.6536	60.6479	60.6599	**60.6669**	60.5369	60.5390	60.5385
0.60	61.2833	61.3531	**61.3643**	60.8523	60.8552	60.8586	60.7086	60.6997	60.6936
MRI									
0.20	61.4382	61.4742	61.5013	61.5564	61.5632	**61.5839**	61.5495	61.5507	61.5476
0.40	61.9915	62.0421	**62.0449**	61.8812	61.8832	61.8894	61.7668	61.7709	61.7774
0.60	62.5490	**62.5834**	62.5743	62.1257	62.1143	62.1080	61.9912	61.9940	61.9890
X-Ray									
0.20	61.3369	61.3663	**61.3847**	61.3322	61.3459	61.3351	61.3201	61.3244	61.3302
0.40	61.9276	61.9881	**61.9954**	61.6416	61.6535	61.6562	61.5440	61.5521	61.5537
0.60	62.5809	62.6176	**62.6187**	61.8916	61.8944	61.8725	61.7884	61.7960	61.1775
Cambridge									
0.20	59.3242	**59.3575**	59.3520	59.3454	59.3514	59.3524	59.3100	59.3104	59.3117
0.40	59.6281	59.6879	59.7076	59.7107	59.7117	**59.7188**	59.5632	59.5730	59.5721
0.60	60.1501	**60.2214**	60.2201	59.8947	59.9010	59.8873	59.7190	59.7245	59.7155

images, namely, *Barbara*, *Diatoms* and *Text*. A maximum reduction (2 dB) in PSNR is observed in the reconstruction of *Diatoms*. The CSF-BPA mask outperforms the CSF-DWT mask by providing an improvement in PSNR of 0.7 dB for the case of the *Lena* image.

Table 4-5 gives the PSNR results for the reconstructed set of images corrupted with a zero-mean AWGN with a noise variance of 26. From the table, it is observed that, with the increase in the amount of noise, there is a considerable reduction in the calculated PSNR. Here, the number of images showing a decrease in PSNR performance has increased from three to eight in comparison with the previous noise conditions. This accounts for 73% of the total images under analysis. A maximum reduction of 0.4 dB is noticed for the case of the *Diatoms* image. A maximum PSNR improvement of 0.2 dB is achieved with the use of the CSF-DWT mask.

In view of the above mentioned results, for most of the images, the BNC wavelets provide a better PSNR in comparison with the standard CDF 9/7 wavelet irrespective of the application of the CSF mask. An improvement in the range of 0.8 - 2.0 dB is obtained when the BNC wavelet is used with integrated CSF-BPA mask. Also, the CSF-BPA mask offers a PSNR improvement in the range of 0.2 - 1.0 dB over the previously used CSF-DWT mask. The application of the CSF mask reduces the overall PSNR value for *Baboon*, *Pot* and *Text* images irrespective of the amount of noise. When the noise variance is increased beyond 23, the decreasing trend of the PSNR value is more pronounced among the standard test and medical images. For example, when the variance is increased to 26, eight images show poor PSNR performance.

4.2 Qualitative Assessment

Figs. 4-1 – 4-17 present the visual performance for the various filters and mask conditions under specific compression ratios. Fig 4.1 shows the results for a noisy *Lena* image compressed at 0.6 bpp and reconstructed without the application of the CSF mask. It can be seen that the absence of masking leaves significant amount of noise in the recovered image. On a hardcopy, there is no noticeable difference among the performance of the CDF 9/7 and BNC filters. However, the reconstruction using the BNC wavelets shows reduced visibility of noise pixels at certain regions (near the eyes and the nose). Also, in terms of the computation time required by each of the filters, the BNC 17/11 wavelet outperforms the other two filters.

Figs. 4-2 and 4-3 show the results of the reconstruction of the *Lena* image using the CSF-DWT mask and the CSF-BPA mask respectively. The overall quality of the image, defined in terms of the amount of noise in the reconstructed image, seems to have improved with the application of the mask. This result can be verified by comparing these images with Fig. 4-1 (reconstruction with the absence of mask). The deduced CSF-BPA mask seems to preserve more features, namely, the lips, nose, lower eyelashes and feathers (attached to the hat) in comparison with the CSF-DWT mask, which can be verified by comparing Figs. 4-2 and 4-3. The overall contrast is also enhanced.

The *Cambridge* image reconstructed under three different masking modes is as shown in Figs. 4-4, 4-5 and 4-6 respectively. The visibility of the tomb, the side windows and the hoisted flag in the right corner of the image frame seems to have enhanced with the application of the mask. The clarity of the entrance structures and their shape has improved in the reconstructed image when the HVS masking is considered during

compression. Also, the reconstruction using the BNC wavelet appears to have improved clarity of the image pixels and reduced overall noise in comparison with the standard CDF 9/7 wavelet. BNC 22/14 wavelet shows comparable performance with that of BNC 17/11 wavelet except for the fact that being a long even-order filter with more vanishing moments, it results in additional smoothing effect in the overall picture quality.

From the reconstructed *X-Ray* results shown in Figs. 4-7, 4-8 and 4-9, it can be noticed that the use of the BNC wavelet filters integrated with CSF-BPA mask enhances the perceptual quality compared to that obtained using the standard CDF 9/7 wavelet filter. For the case of reconstructed *Barbara* image (see Fig 4-10.), which is an overly textured image, the BNC 17/11 wavelet retains more of the checkered pattern on the side of the table cloth, a behavior which is expected from longer filters [19]. The difference between the BNC 17/11 and the BNC 22/14 wavelets is difficult to demonstrate in a hardcopy. However, in a high quality reproduction, it is noticed that the BNC 17/11 wavelet (symmetric odd-length filter) exhibits less ringing than the BNC 22/14 wavelet (asymmetric even-length filter) does. In fact, a close inspection of the figures shows that the BNC 17/11 filter exhibits less visually disturbing ringing than the CDF 9/7 filter does. In effect, for still image coding, the BNC 17/11 wavelet retains high frequencies without any significant ringing artifact.

The *Boat* and *Baboon* are balanced images with sharp edges, smooth sections and textured areas. Hence, for analysis of such images these prominent features are to be considered. In the reconstructed *Boat* image, shown in Fig 4-11., the clarity of the image pattern after reconstruction has improved considerably at the cost of a reduction in the overall sharpness. The shapes of the additional structures are well preserved by the

application of the CSF-BPA mask. As expected, the images reconstructed using the BNC wavelets have an enhanced visual appeal.

In the reconstructed compressed *Baboon* test image shown in Fig 4-12., the visibility of the block boundaries is reduced and false patterns within the blocks are eliminated when BNC wavelet is used with integrated CSF masking (using CSF-BPA mask).

For the case of the reconstruction of *Text* (see Fig 4-13) and *Pot* (see Fig 4-14) test images, the BNC wavelets when used with integrated CSF masking (with CSF-BPA mask) has resulted in a reduction in the overall noise activity. Further, an enhancement in the image contrast is noticed leading to a slight improvement in the perceptual quality.

For the case of *MRI* image shown in Fig 4-15., the reconstructed image is slightly smoothened with the application of CSF mask. The decrease in the blur of the reconstructed image is achieved at the cost of smoothening of the image.

The results of the reconstruction of *Ultrasound scan* image, shown in Fig. 4-16, shows a similar perceptual quality improvement as seen for the case of *X-Ray* images. The edges of the womb are preserved slightly better without any blur with the application of masking used integrated with the BNC wavelet in place of the CDF 9/7 wavelet. In the reconstruction of *Diatoms* test image (see Fig 4-17), the overall contrast is enhanced with a reduction in the noise activity thereby an increase in the clarity of the edges. However, the overall contrast has reduced and the edges are slightly smoothened.

In an attempt to compare and judge the performance of the filters under different masking conditions, the *Lena*, *Barbara* and *X-Ray* images, reconstructed at two compression rates, namely, 0.2 bpp and 0.6 bpp, are provided in Figs. 4-18 – 4-23. It is observed from the figures that the BNC wavelets reconstruct the image with an enhanced perceptual quality
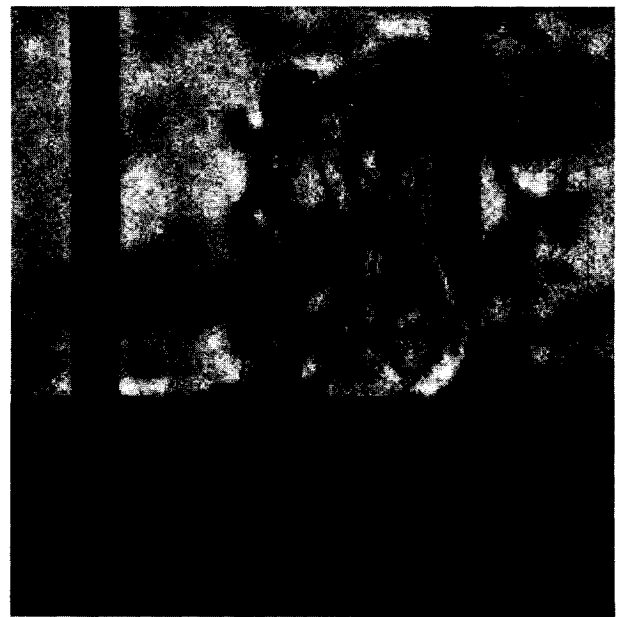
(a)
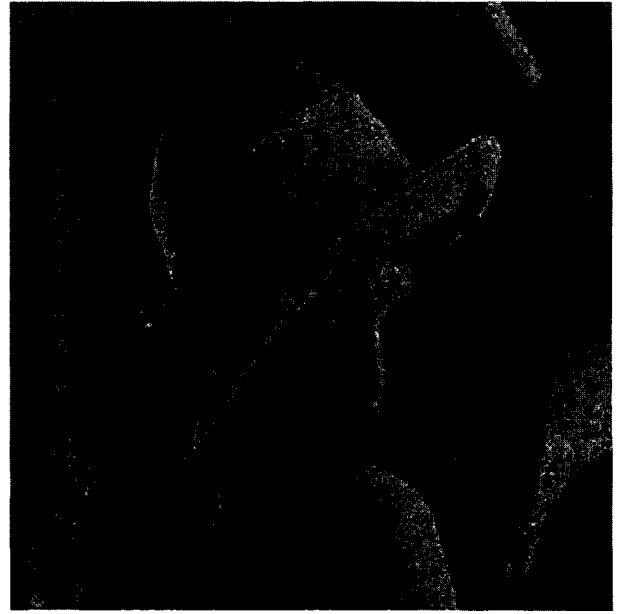
(b)

(c)
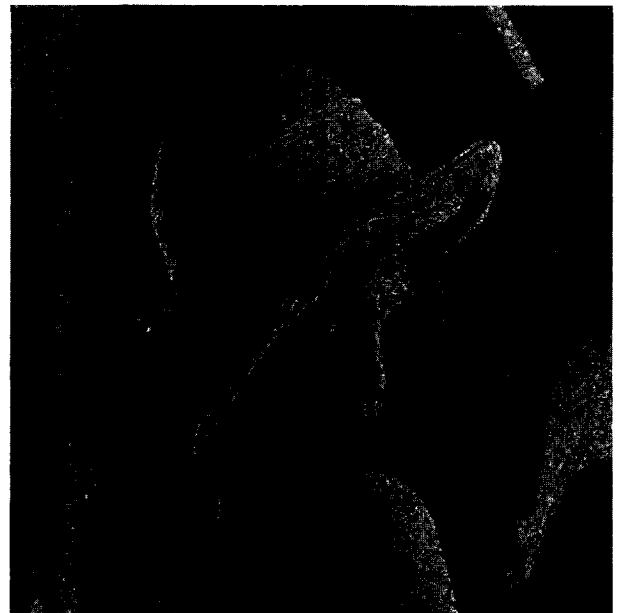
(d)

Figure 4-1 Comparison of visual image quality of the reconstructed *Lena* image compressed at 0.6bpp (with no mask). (a) Original corrupted with zero-mean AWGN of variance 26. Image reconstructed using (b) CDF 9/7 (c) BNC 17/11 (d) BNC 22/14

(a)

(b)

(c)

(d)

Figure 4-2 Comparison of visual image quality of the reconstructed *Lena* image
compressed at 0.6bpp (with CSF-DWT mask). (a) Original corrupted with zero-mean
AWGN of variance 26. Image reconstructed using (b) CDF 9/7 (c) BNC 17/11 (d) BNC
22/14

(a)

(b)

(c)

(d)

Figure 4-3 Comparison of visual image quality of the reconstructed *Lena* image compressed at 0.6bpp (with CSF-BPA mask). (a) Original corrupted with zero-mean AWGN of variance 26. Image reconstructed using (b) CDF 9/7 (c) BNC 17/11 (d) BNC 22/14
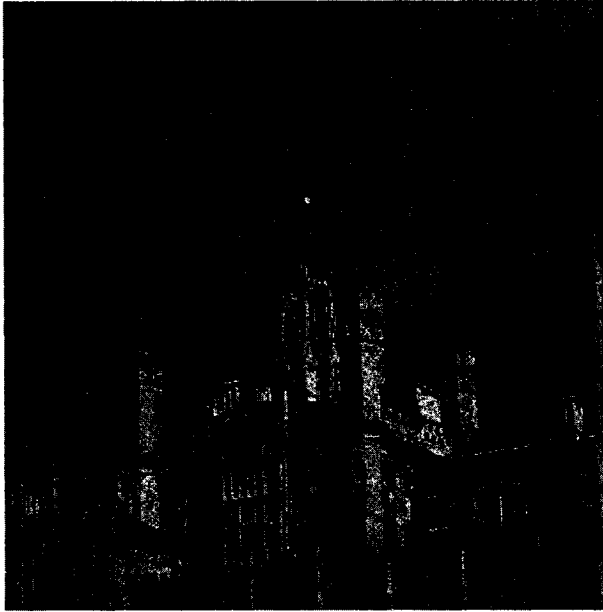
(a)
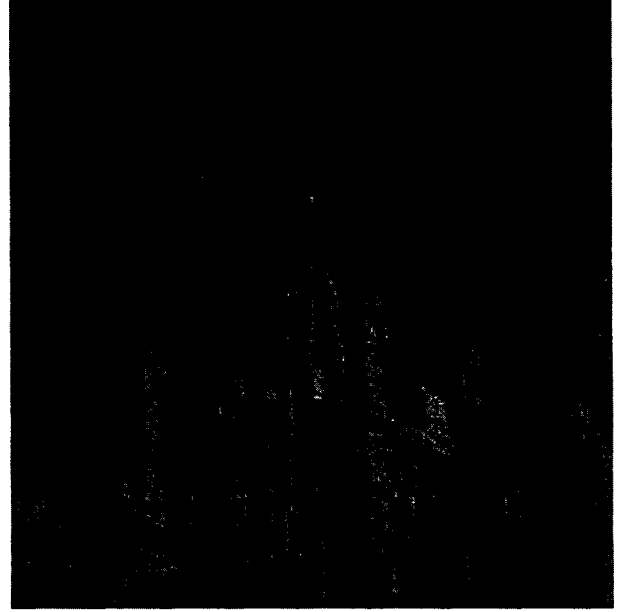
(b)
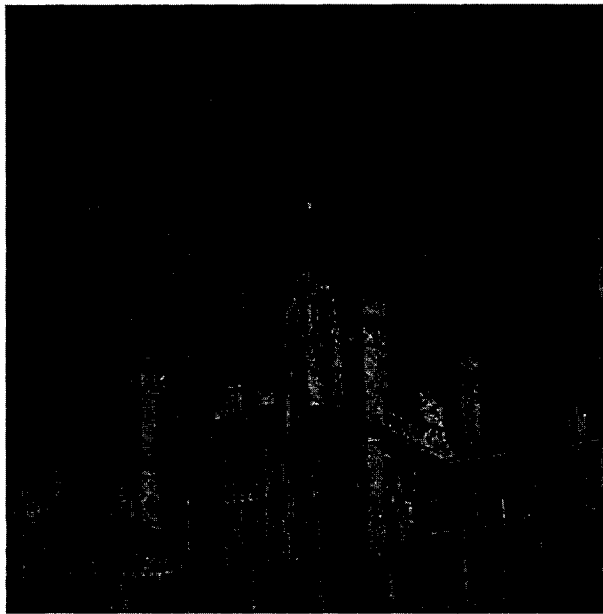
(c)
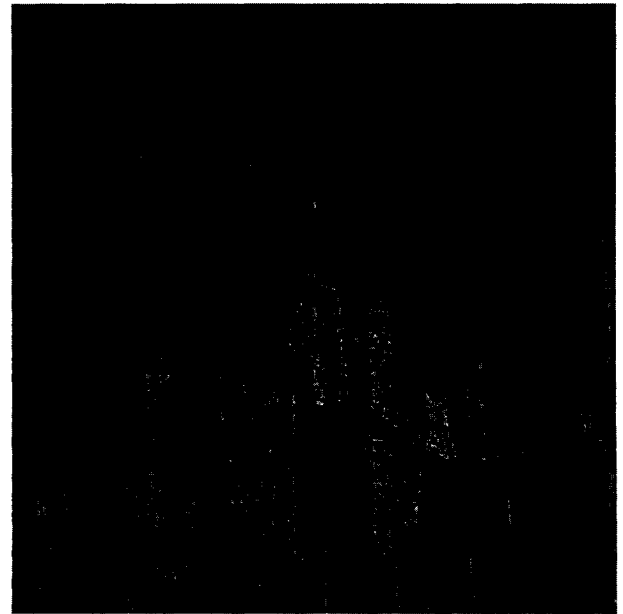
(d)

Figure 4-4 Comparison of visual image quality of the reconstructed *Cambridge* image compressed at 0.6bpp (with no mask). (a) Original corrupted with zero-mean AWGN of variance 13. Image reconstructed using (b) CDF 9/7 (c) BNC 17/11 (d) BNC 22/14
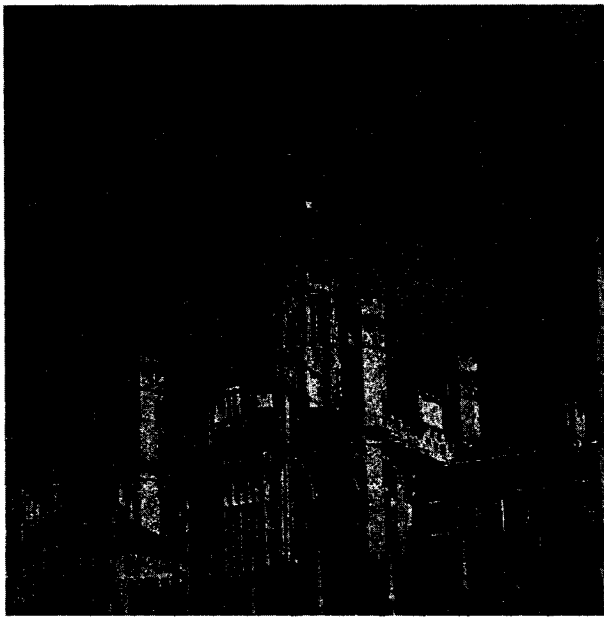
(a)

(b)
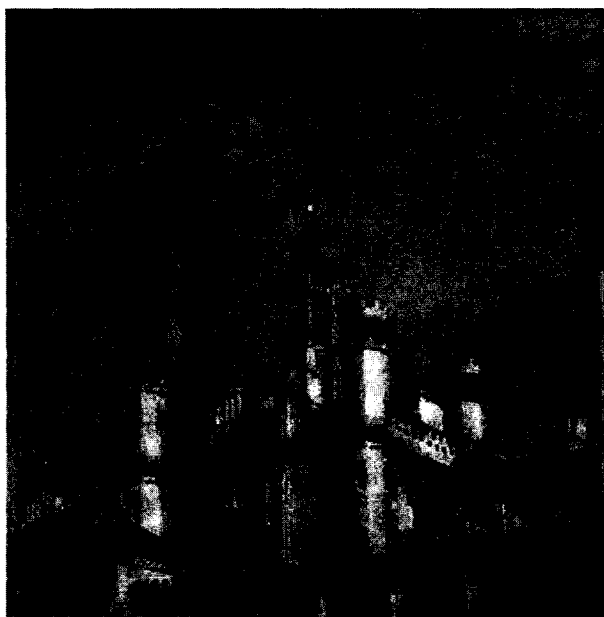
(c)
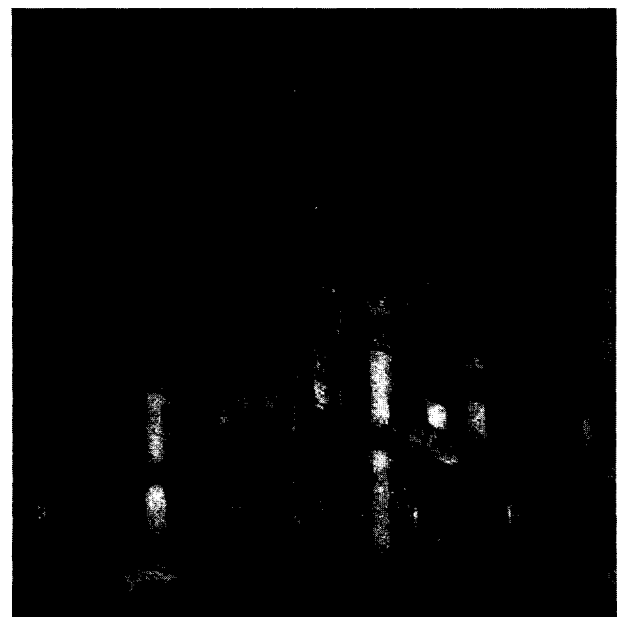
(d)

Figure 4-5 Comparison of visual image quality of the reconstructed *Cambridge* image compressed at 0.6bpp (with CSF-DWT mask). (a) Original corrupted with zero-mean AWGN of variance 13. Image reconstructed using (b) CDF 9/7 (c) BNC 17/11 (d) BNC 22/14
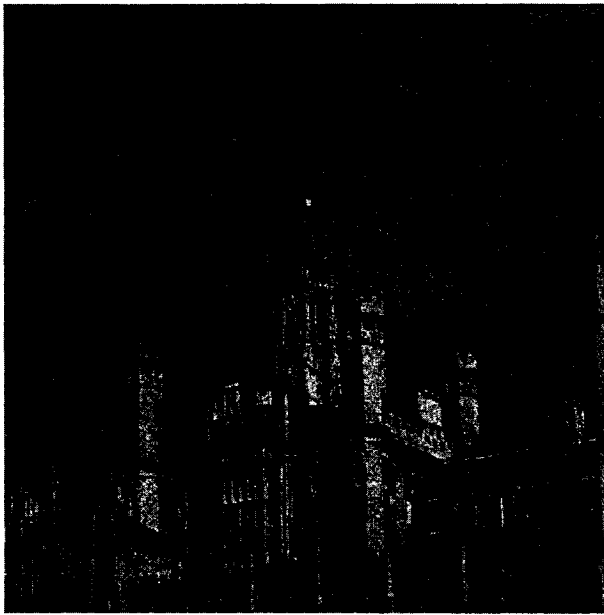
(a)

(b)
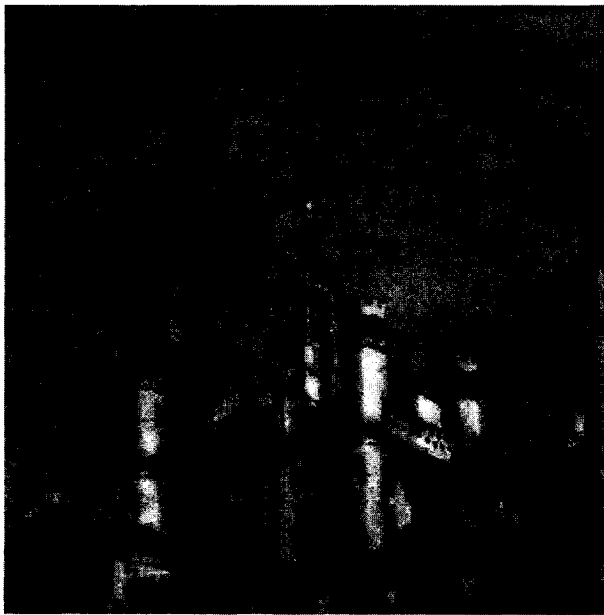
(c)
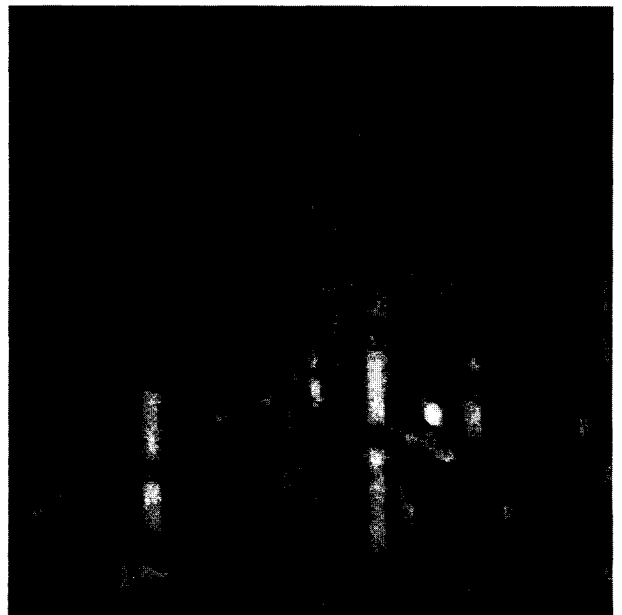
(d)

Figure 4-6 Comparison of visual image quality of the reconstructed *Cambridge* image compressed at 0.6bpp (with CSF-BPA mask). (a) Original corrupted with zero-mean AWGN of variance 13. Image reconstructed using (b) CDF 9/7 (c) BNC 17/11 (d) BNC 22/14
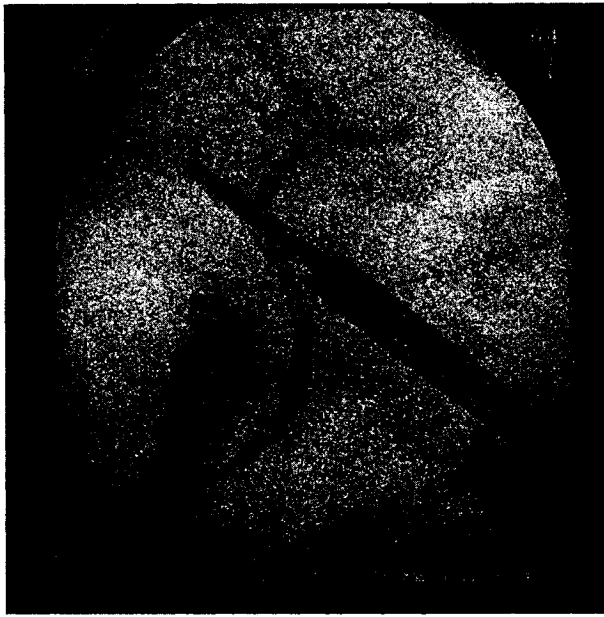
(a)
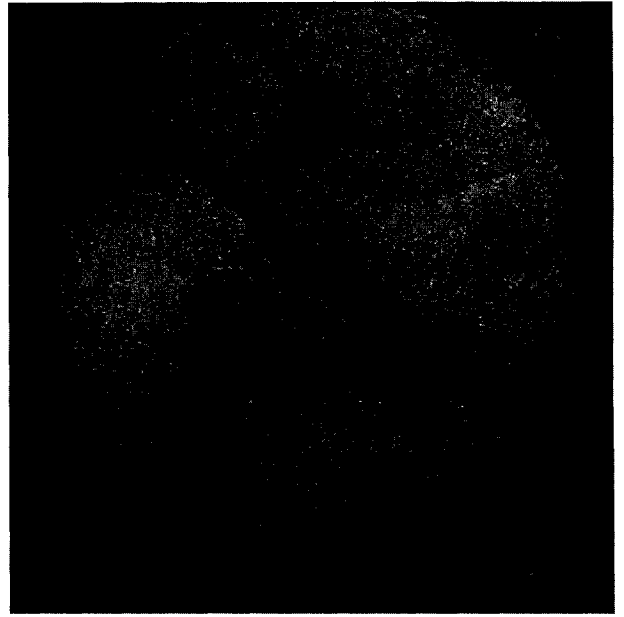
(b)
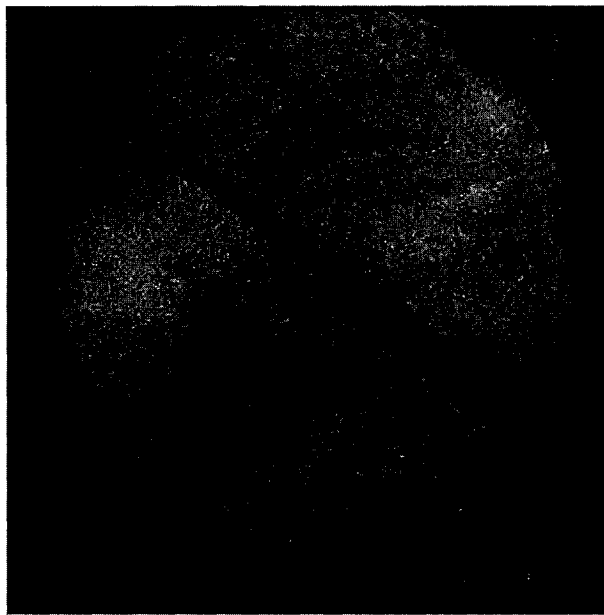
(c)
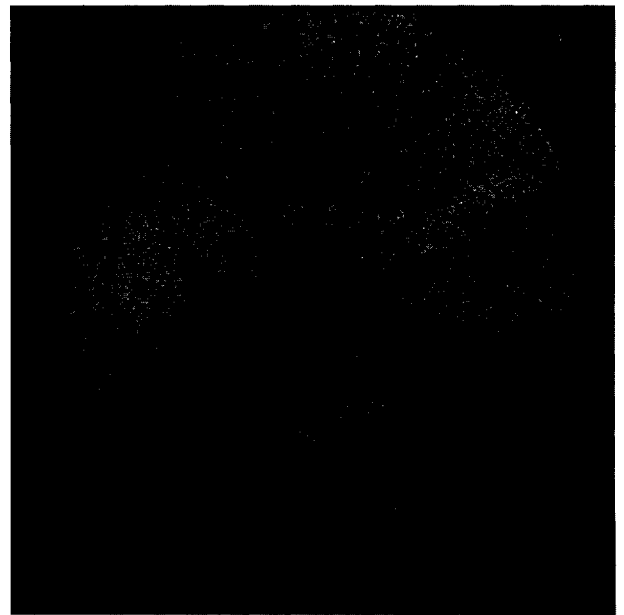
(d)

Figure 4-7 Comparison of visual image quality of the reconstructed *X-Ray* image compressed at 0.6bpp (with no mask). (a) Original corrupted with zero-mean AWGN of variance 13. Image reconstructed using (b) CDF 9/7 (c) BNC 17/11 (d) BNC 22/14
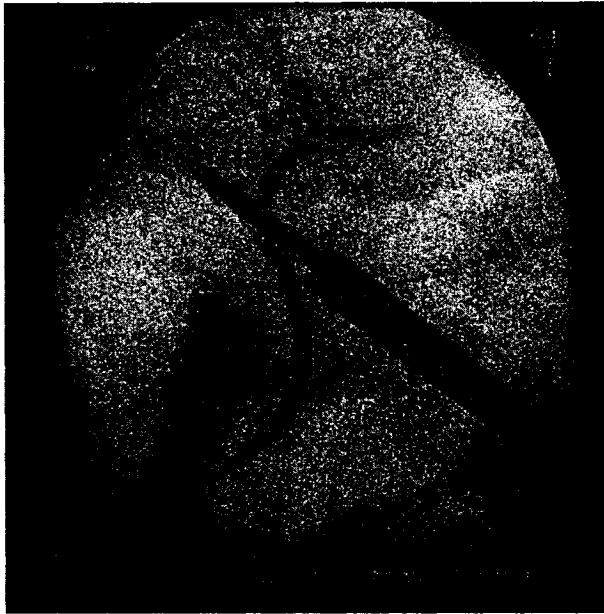
(a)
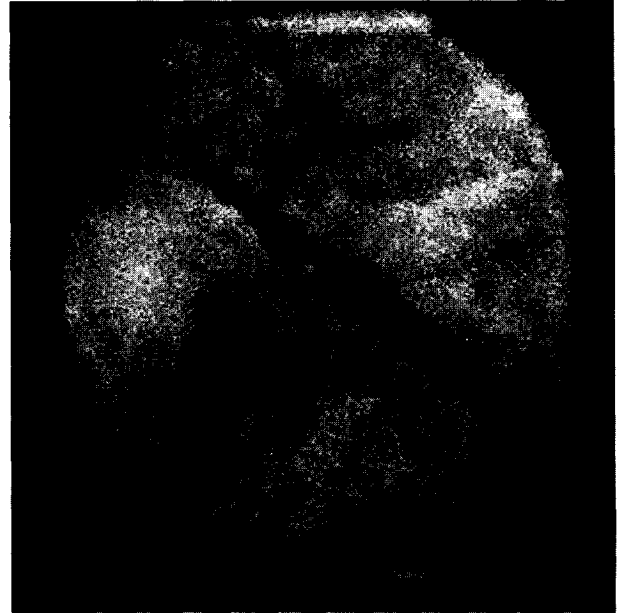
(b)
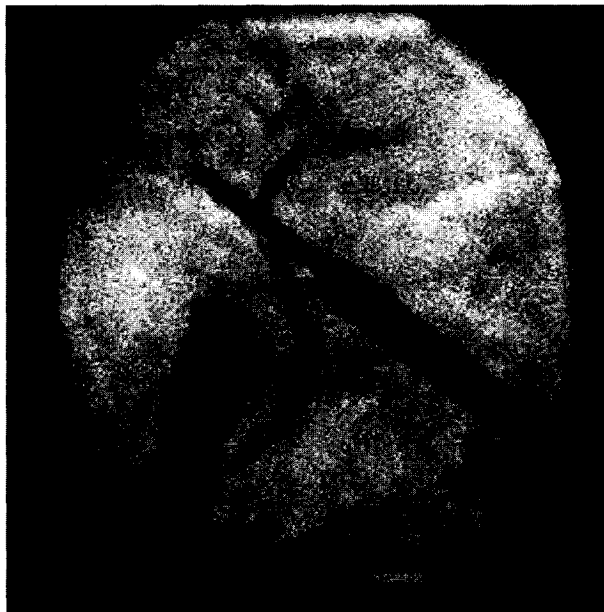
(c)
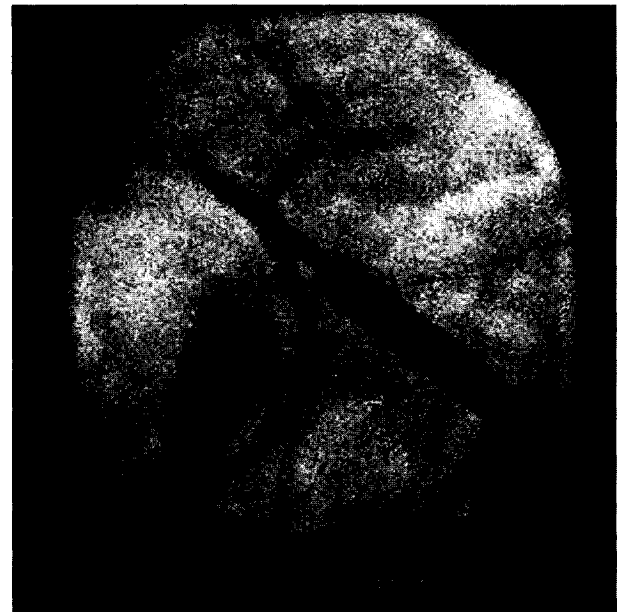
(d)

Figure 4-8 Comparison of visual image quality of the reconstructed *X-Ray* image compressed at 0.6bpp (with CSF-DWT mask). (a) Original corrupted with zero-mean AWGN of variance 13. Image reconstructed using (b) CDF 9/7 (c) BNC 17/11 (d) BNC 22/14
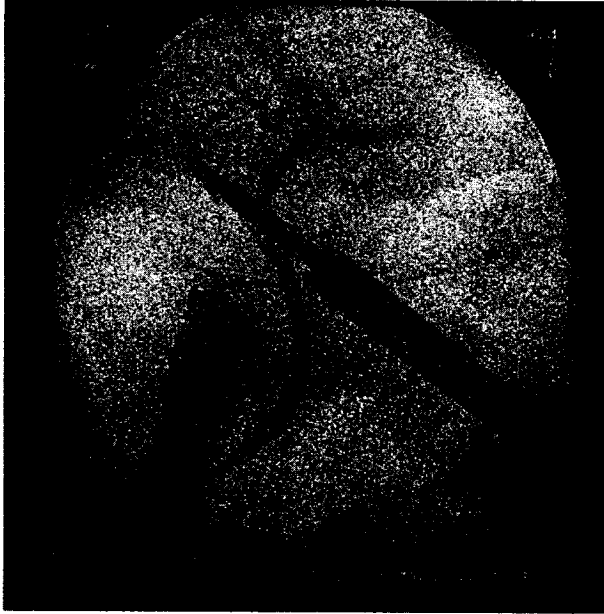
(a)

(b)
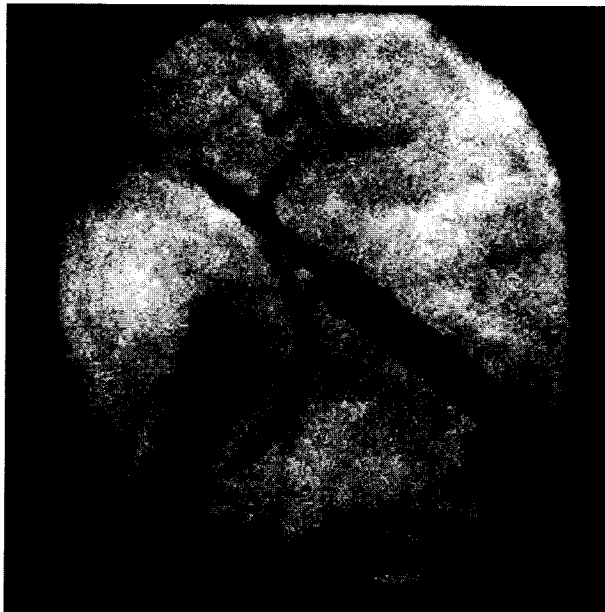
(c)

(d)

Figure 4-9 Comparison of visual image quality of the reconstructed *X-Ray* image compressed at 0.6bpp (with CSF-BPA mask). (a) Original corrupted with zero-mean AWGN of variance 13. Image reconstructed using (b) CDF 9/7 (c) BNC 17/11 (d) BNC 22/14

(a)
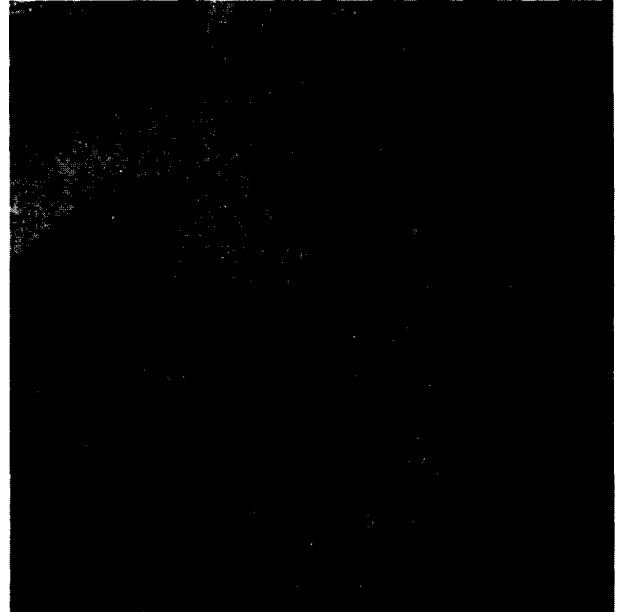
(b)

(c)

(d)

Figure 4-10 Comparison of visual image quality of the *Barbara* image compressed at 0.4bpp and reconstructed using the BNC 17/11 wavelet. (a) Original corrupted with zero-mean AWGN of variance 26. Image reconstructed with (b) no CSF mask (c) CSF-DWT mask (d) CSF-BPA mask.
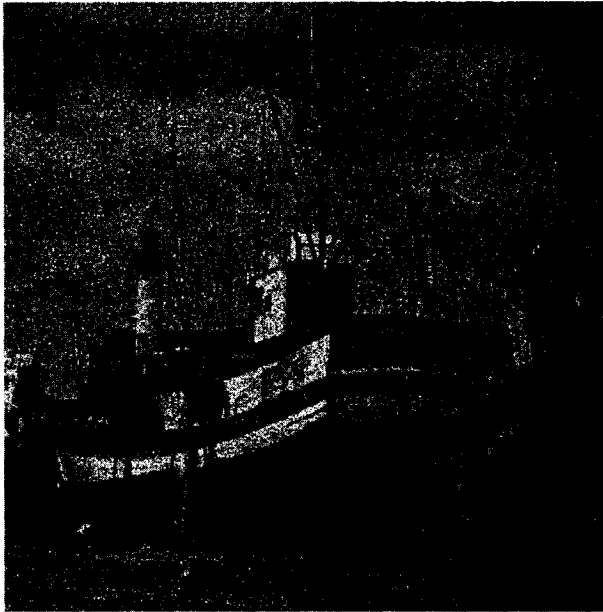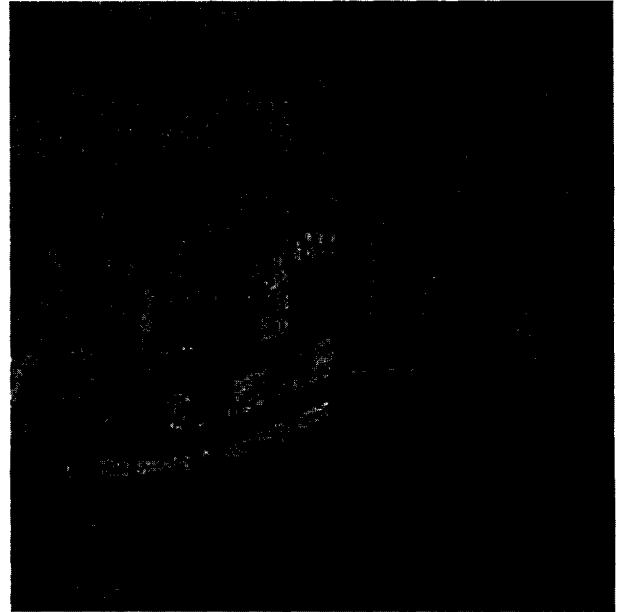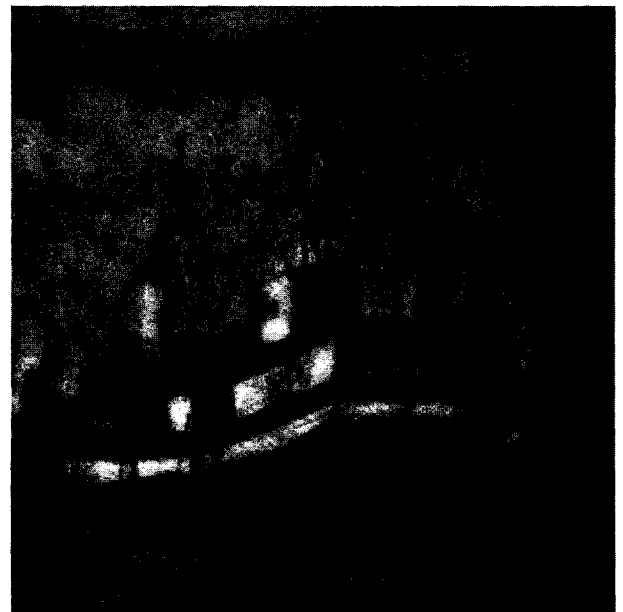
Figure 4-11 Comparison of visual image quality of the *Boat* image compressed at 0.4bpp and reconstructed using the BNC 17/11 wavelet. (a) Original corrupted with zero-mean AWGN of variance 26. Image reconstructed with (b) no CSF mask (c) CSF-DWT mask (d) CSF-BPA mask.
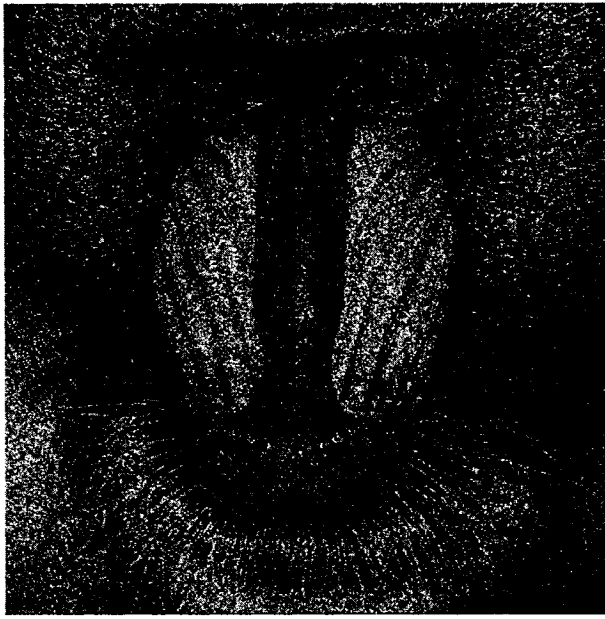
(a)
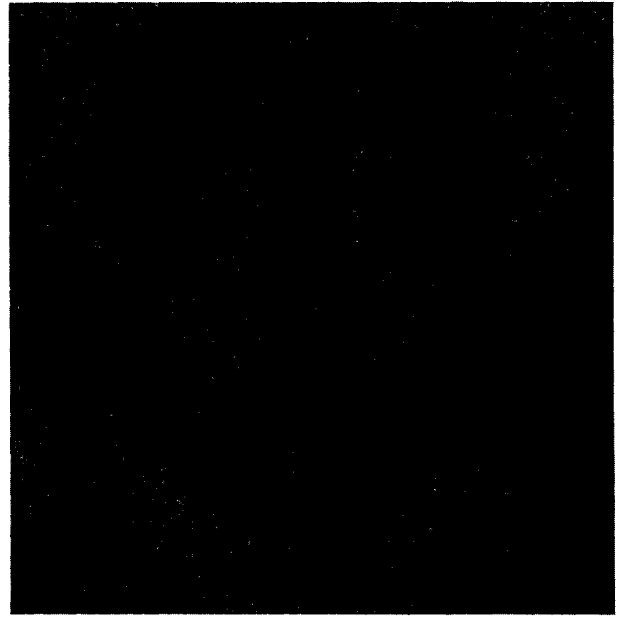
(b)

(c)

(d)

Figure 4-12 Comparison of visual image quality of the *Baboon* image compressed at 0.4bpp and reconstructed using the BNC 17/11 wavelet. (a) Original corrupted with zero-mean AWGN of variance 26. Image reconstructed with (b) no CSF mask (c) CSF-DWT mask (d) CSF-BPA mask.
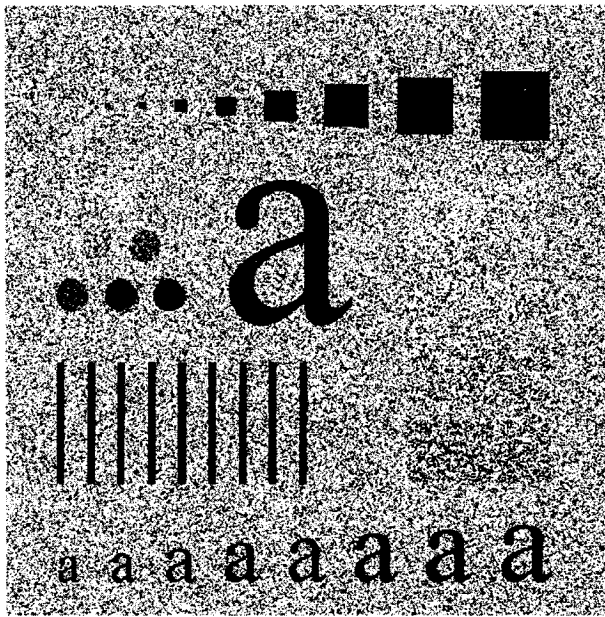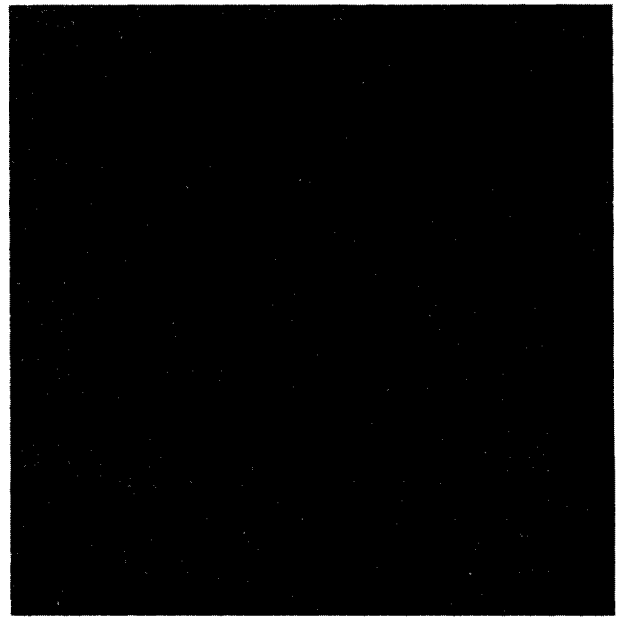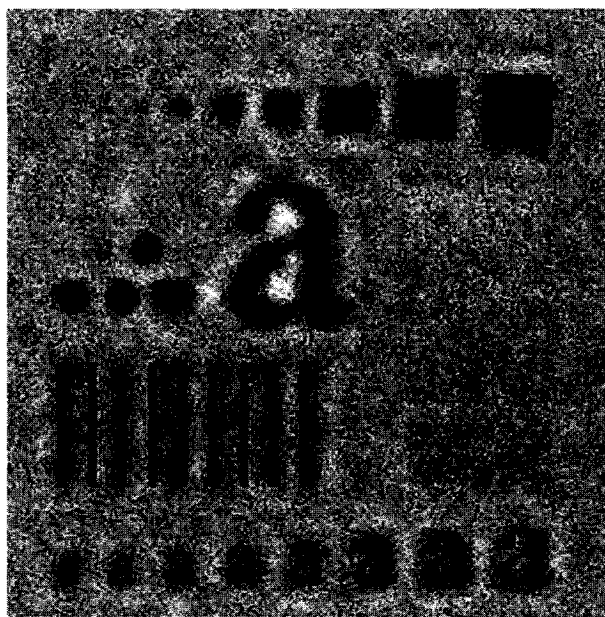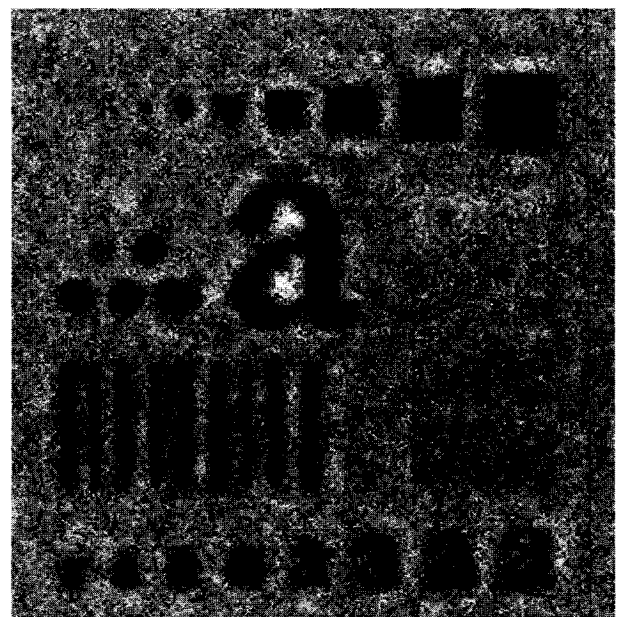
Figure 4-13 Comparison of visual image quality of the *Text* image compressed at 0.6bpp
and reconstructed using the BNC 17/11 wavelet. (a) Original corrupted with zero-mean
AWGN of variance 6. Image reconstructed with (b) no CSF mask (c) CSF-DWT mask
(d) CSF-BPA mask.

95
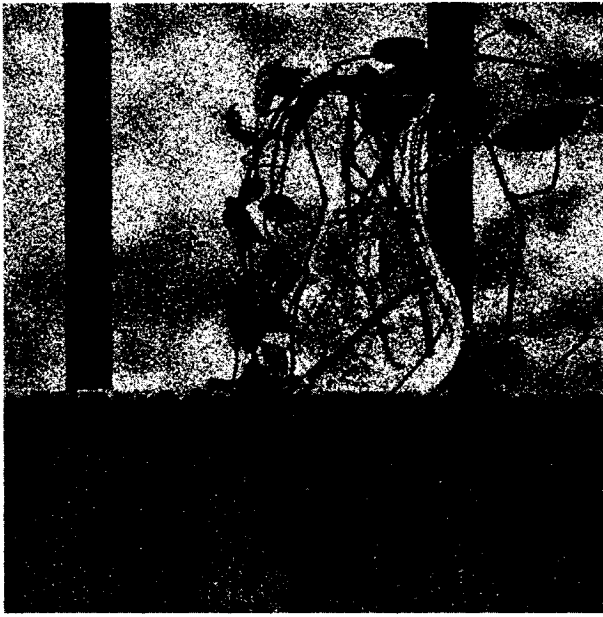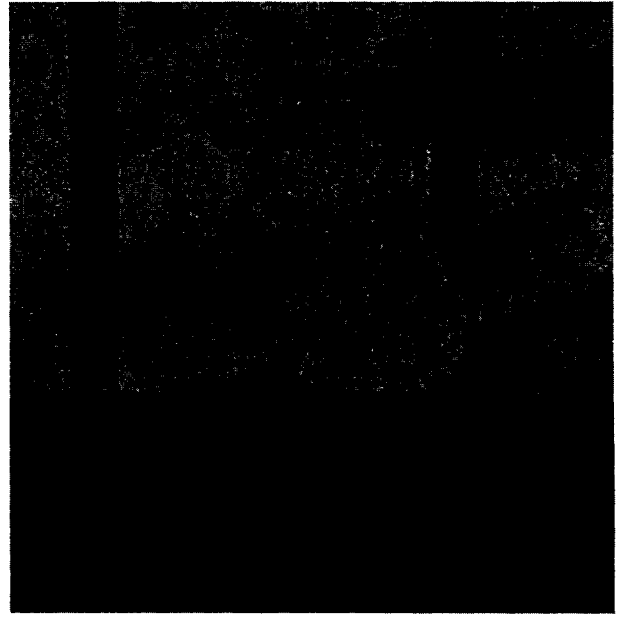
(a)   (b)

(c)   (d)

Figure 4-14 Comparison of visual image quality of the *Pot* image compressed at 0.6bpp and reconstructed using the BNC 17/11 wavelet. (a) Original corrupted with zero-mean AWGN of variance 26. Image reconstructed with (b) no CSF mask (c) CSF-DWT mask (d) CSF-BPA mask.
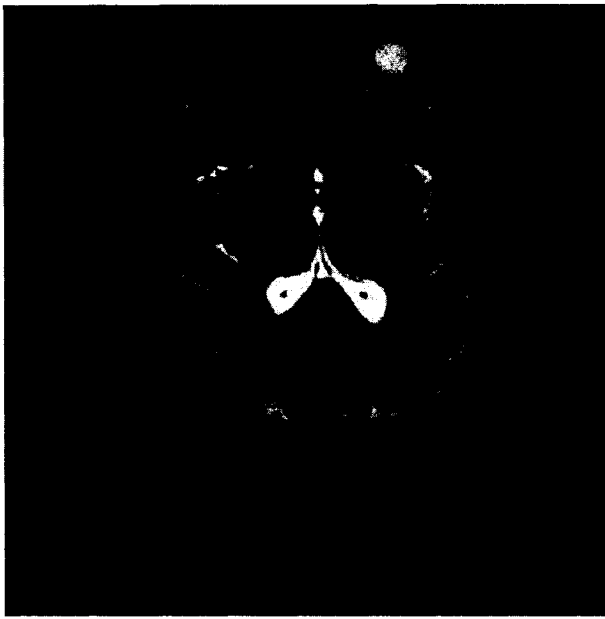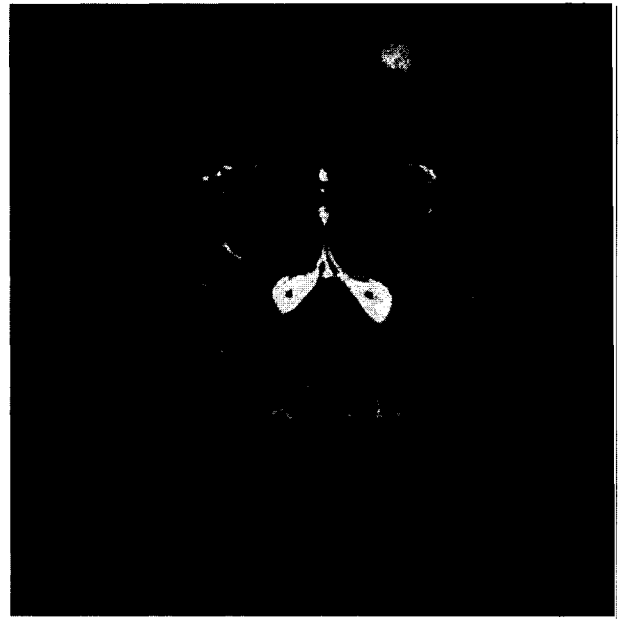
(a)   (b)
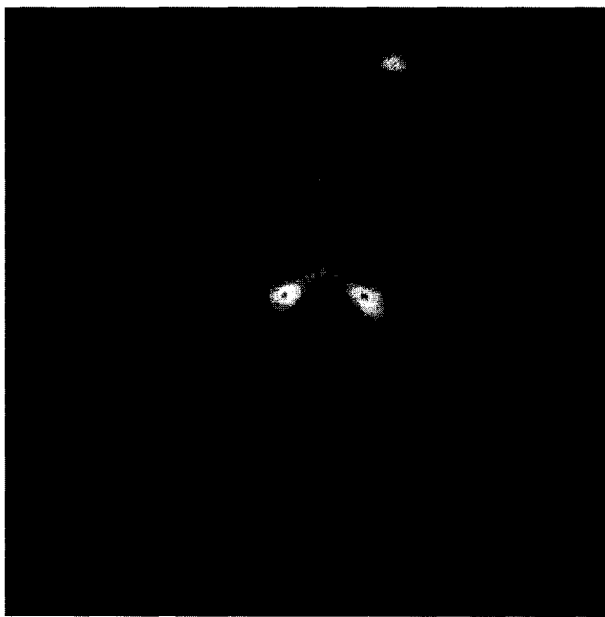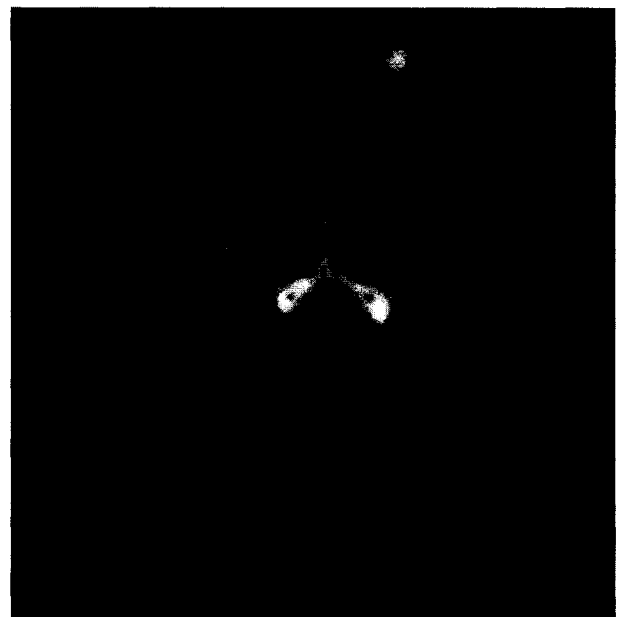
(c)   (d)

Figure 4-15 Comparison of visual image quality of the *MRI* image compressed at 0.4bpp and reconstructed using the BNC 17/11 wavelet. (a) Original corrupted with zero-mean AWGN of variance 3. Image reconstructed with (b) no CSF mask (c) CSF-DWT mask (d) CSF-BPA mask.
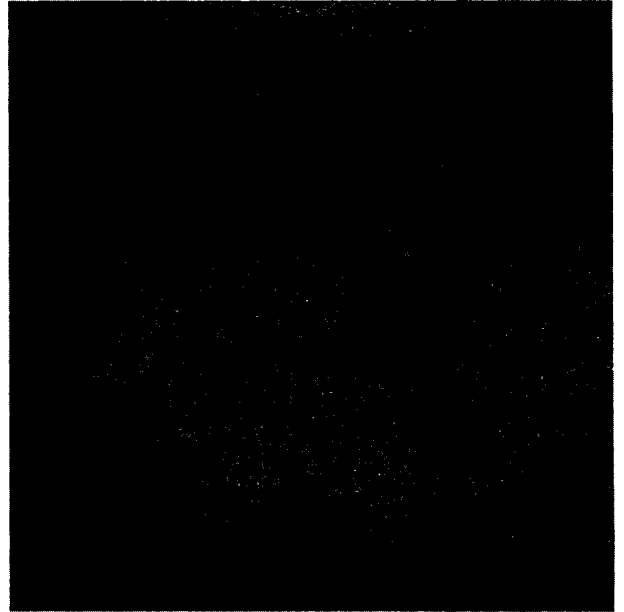
(a)　　　　　　　　　　　　　　(b)
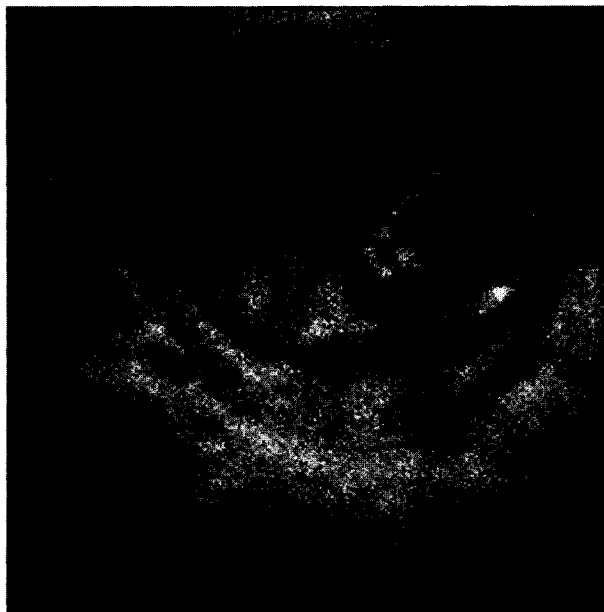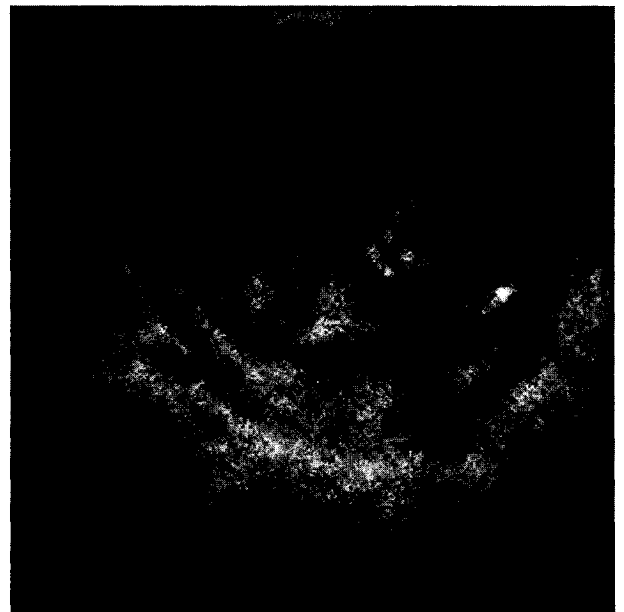
(c)　　　　　　　　　　　　　　(d)

Figure 4-16 Comparison of visual image quality of the *Ultrasound scan* image compressed at 0.4bpp and reconstructed using the BNC 17/11 wavelet. (a) Original corrupted with zero-mean AWGN of variance 26. Image reconstructed with (b) no CSF mask (c) CSF-DWT mask (d) CSF-BPA mask.
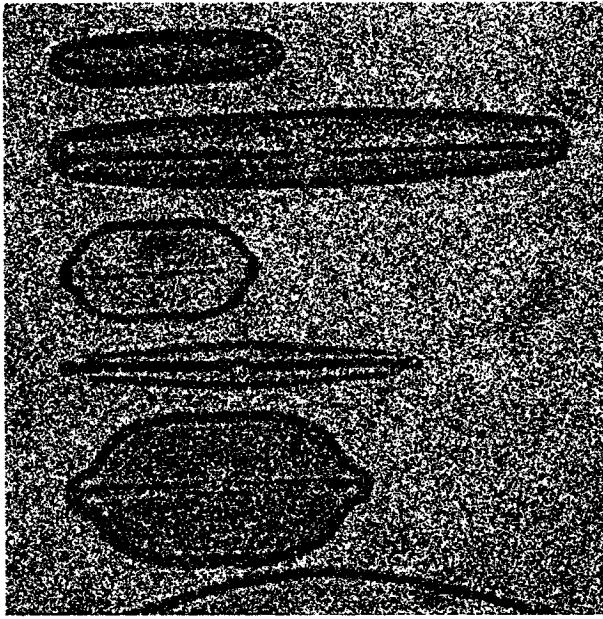
(a)
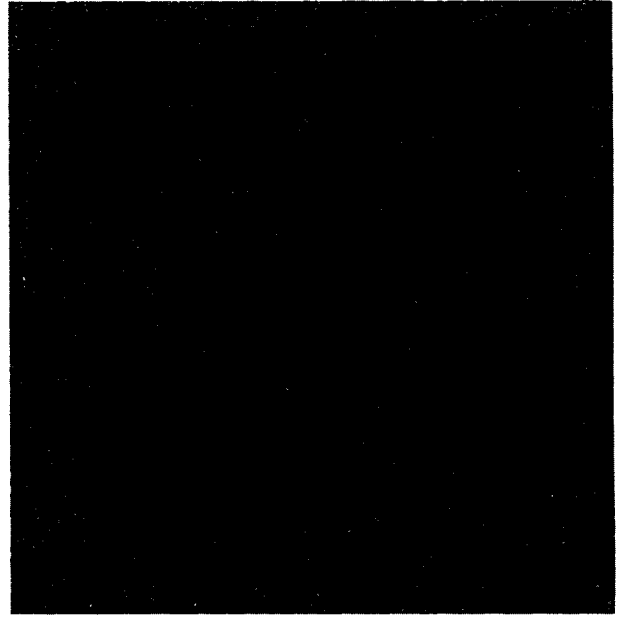
(b)

(c)
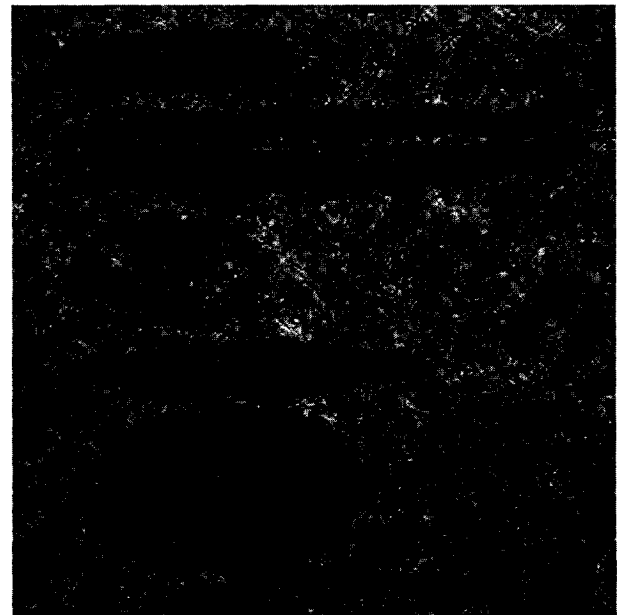
(d)

Figure 4-17 Comparison of visual image quality of the *Diatoms* image compressed at 0.4bpp and reconstructed using the BNC 17/11 wavelet. (a) Original corrupted with zero-mean AWGN of variance 13. Image reconstructed with (b) no CSF mask (c) CSF-DWT mask (d) CSF-BPA mask.

	No CSF mask	CSF-DWT mask	CSF-BPA mask
CDF 9/7			
BNC 17/11			
BNC 22/14			

Figure 4-18 Comparison of visual image quality of the reconstructed *Lena* image (corrupted with zero-mean AWGN of variance 26 compressed at 0.2 bpp) under different filtering and masking conditions. The number of resolution levels used is 5.

	No CSF mask	CSF-DWT mask	CSF-BPA mask
CDF 9/7			
BNC 17/11			
BNC 22/14			

Figure 4-19 Comparison of visual image quality of the reconstructed *Lena* image (corrupted with zero-mean AWGN of variance 26 compressed at 0.6 bpp) under different filtering and masking conditions. The number of resolution levels used is 5.
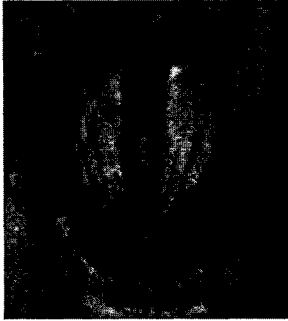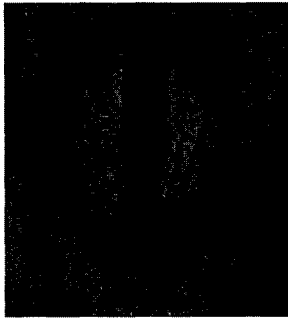
	No CSF mask	CSF-DWT mask	CSF-BPA mask
CDF 9/7			
BNC 17/11			
BNC 22/14			

Figure 4-20 Comparison of visual image quality of the reconstructed Baboon image (corrupted with zero-mean AWGN of variance 26 compressed at 0.2 bpp) under different filtering and masking conditions. The number of resolution levels used is 5.
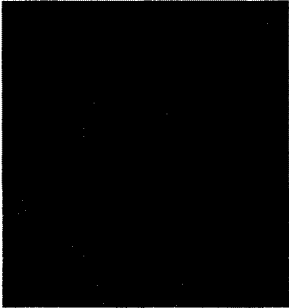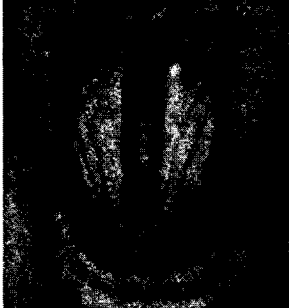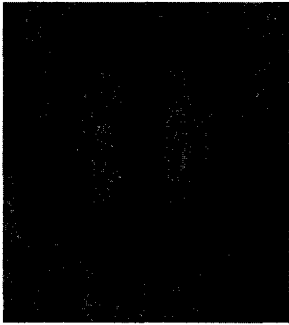
	No CSF mask	CSF-DWT mask	CSF-BPA mask
CDF 9/7			
BNC 17/11			
BNC 22/14			

Figure 4-21 Comparison of visual image quality of the reconstructed *Baboon* image (corrupted with zero-mean AWGN of variance 26 compressed at 0.6 bpp) under different filtering and masking conditions. The number of resolution levels used is 5.
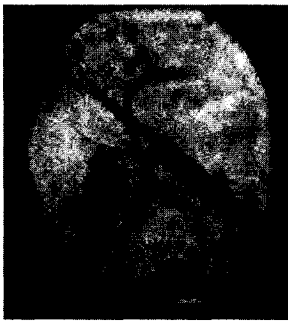
	No CSF mask	CSF-DWT mask	CSF-BPA mask
CDF 9/7			
BNC 17/11			
BNC 22/14			

Figure 4-22 Comparison of visual image quality of the reconstructed *X-Ray* image (corrupted with zero-mean AWGN of variance 13 compressed at 0.2 bpp) under different filtering and masking conditions. The number of resolution levels used is 5.
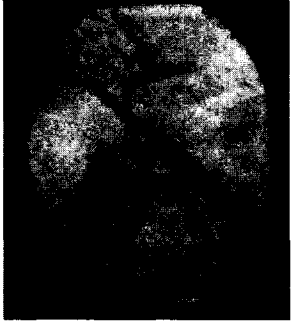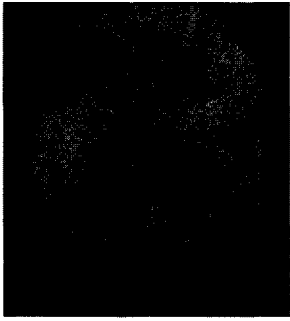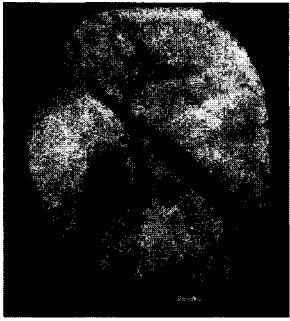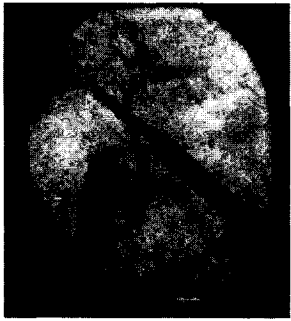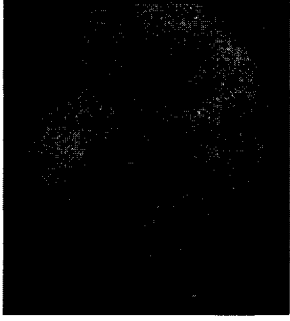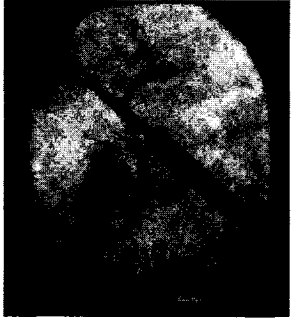
	No CSF mask	CSF-DWT mask	CSF-BPA mask
CDF 9/7			
BNC 17/11			
BNC 22/14			

Figure 4-23 Comparison of visual image quality of the reconstructed *X-Ray* image (corrupted with zero-mean AWGN of variance 13 compressed at 0.6 bpp) under different filtering and masking conditions. The number of resolution levels used is 5.
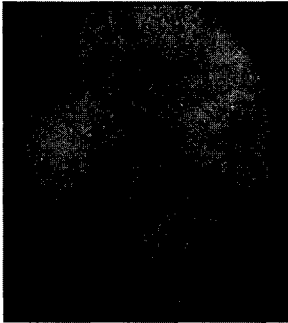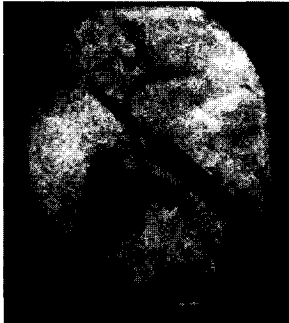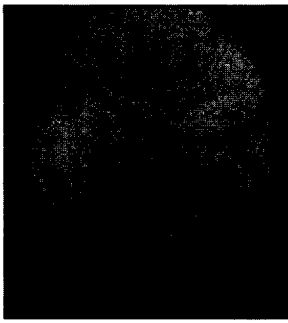
for a given compression ratio, in comparison with the standard CDF 9/7 wavelet. It is also evident that there is a significant reduction in the visible distortion of the reconstructed image. Also, a majority of the linear features corresponding to the image are successfully reconstructed independent of the image size.

4.3 Validation of qualitative assessment

The validation of the proposed scheme is performed by a subjective test, whereby a rating scale and the pair-wise comparison approaches have been used to evaluate the set of reconstructed images. Some of the high bit-rate applications make use of the *Just Noticeable Distortion* (JND) threshold [38] evaluation of the image quality, bu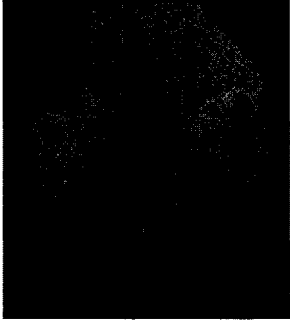t this necessitates the original and compressed images to be displayed to the observer alternatively for a short period of time, which may result in a biased decision from the observer. For all the test cases, the original image is not made available to the observer and the subjective impression of the quality of the reconstructed image is considered more important than the fidelity of the original one.

A subjective image quality testing procedure was carried out over a group of expert and non-expert observers. A double-blind test procedure was followed, where the observers were asked to rate the reconstructed image, i.e., provide their opinion score, on a 1-5 rating scale. The mean opinion score results obtained from the test are tabulated in Table 4-6. The best performance score obtained for each image is highlighted in bold. Basically, the observers were asked to judge the quality of the reconstructed images displayed on a 300 dpi color monitor screen of a personal computer. The perceived quality through the monitor is presumed to be good and none of the display related

degradation is considered. The reconstructed images were viewed from approximately two picture heights (about 3 feet).

For most of the test images, the perceptual quality is found to have improved with the application of a mask. For seven of the eleven images, the BNC 17/11 wavelet shows a better performance compared to the standard CDF 9/7 wavelet and the asymmetric BNC 22/14 wavelet. Overall, the BNC wavelets seem to provide perceptually soothing reconstruction to any casual observer (expert or non-expert) irrespective of possessing an

Table 4-6 Mean Opinion Score results determined by *Double Blind* assessment

Image	Noisy	MOS with no CSF mask applied			MOS with CSF-BPA mask applied		
		CDF 9/7	BNC 17/11	BNC 22/14	CDF 9/7	BNC 17/11	BNC 22/14
Lena	3.00	3.39	3.55	3.46	3.85	3.95	**3.96**
Barbara	2.80	3.10	3.09	3.13	3.28	3.32	**3.33**
Baboon	3.00	3.20	3.22	3.26	4.32	**4.75**	4.53
Boats	3.20	3.20	3.22	3.23	3.88	**3.99**	3.95
Cambridge	3.20	3.03	3.15	3.14	**3.48**	**3.52**	3.49
Diatoms	3.00	3.10	3.08	3.08	4.00	**4.26**	4.10
MRI	2.90	3.48	3.50	**3.51**	2.50	2.63	2.60
Pot	2.90	3.22	3.25	3.22	4.45	**4.55**	4.52
Text	2.50	3.45	3.45	3.46	**3.66**	3.65	3.65
Ultrasound scan	2.00	2.33	2.30	2.33	3.56	**3.88**	3.48
X-Ray	2.50	2.89	2.85	2.82	3.60	**3.88**	3.79

Opinion Score : 1 - very poor, image rejected
2 - poor, quality needs a lot of processing to make it perceptible
3 - fair, quality may be acceptable with some processing
4 - good, very minimum perceptual distortion
5 - very good, no visible artifacts found

increased number of vanishing moments in comparison to the standard biorthogonal 9/7 wavelet.

4.4 Eigen vector analysis

Eigen vector plays an important role in image analysis. *Eigen vector analysis* (EVA) involves a mathematical procedure that transforms a number of possibly correlated variables into a smaller number of uncorrelated variables called *principal components* or *eigen vectors* [58]. Because of this unique property, EVA is used in many areas of research like face recognition, object recognition and illumination planning. Eigen vectors are used to define the eigen regions, which are geometrical features that encompass area, location and shape properties of an image region and are used for principal component analysis or EVA. For image vectors of large dimensions, their singular value decomposition (SVD) matrix is found to determine their eigen vectors. By plotting the eigen vector of a reconstructed image, its correlation or closeness to match with the original image is determined. The only limitation of using eigen vector technique is that it is computationally complex and the accompanying cost is usually high. The simulation environment can still be used for eigen vector analysis at no additional cost except for the increase in the overall design and computation time.

Figs. 4-24 – 4-26 show the results of the eigen vector analysis on the reconstructed *Lena*, *Pot* and *Ultrasound scan* images for three masking modes. The effectiveness of retaining detail information is related to the kind of image pattern used. Fig 4-26 shows the potential suitability of the masking scheme for use in medical image compression.
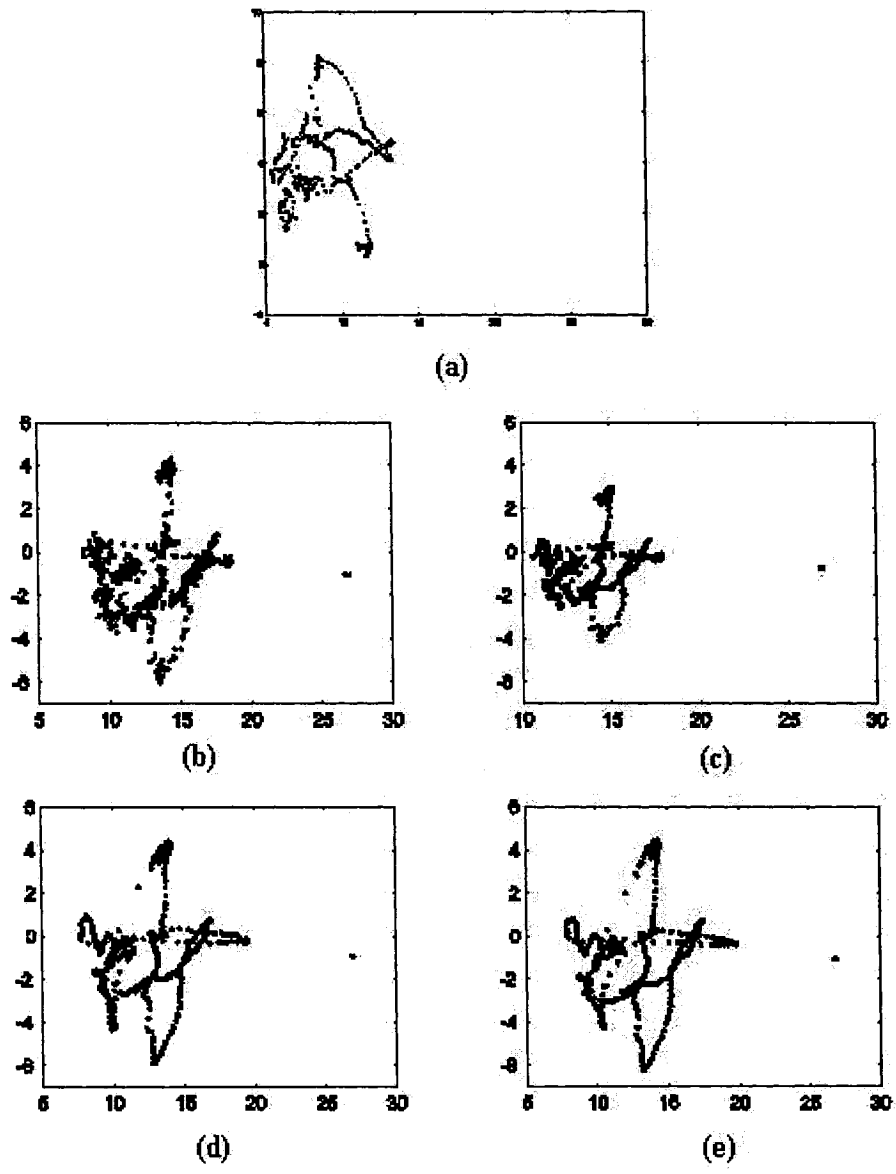
Figure 4-24 Eigen region for Lena image. (a) Original. (b) Noisy. (c) with no mask. (d) CSF-DWT mask. (e) CSF-BPA mask.
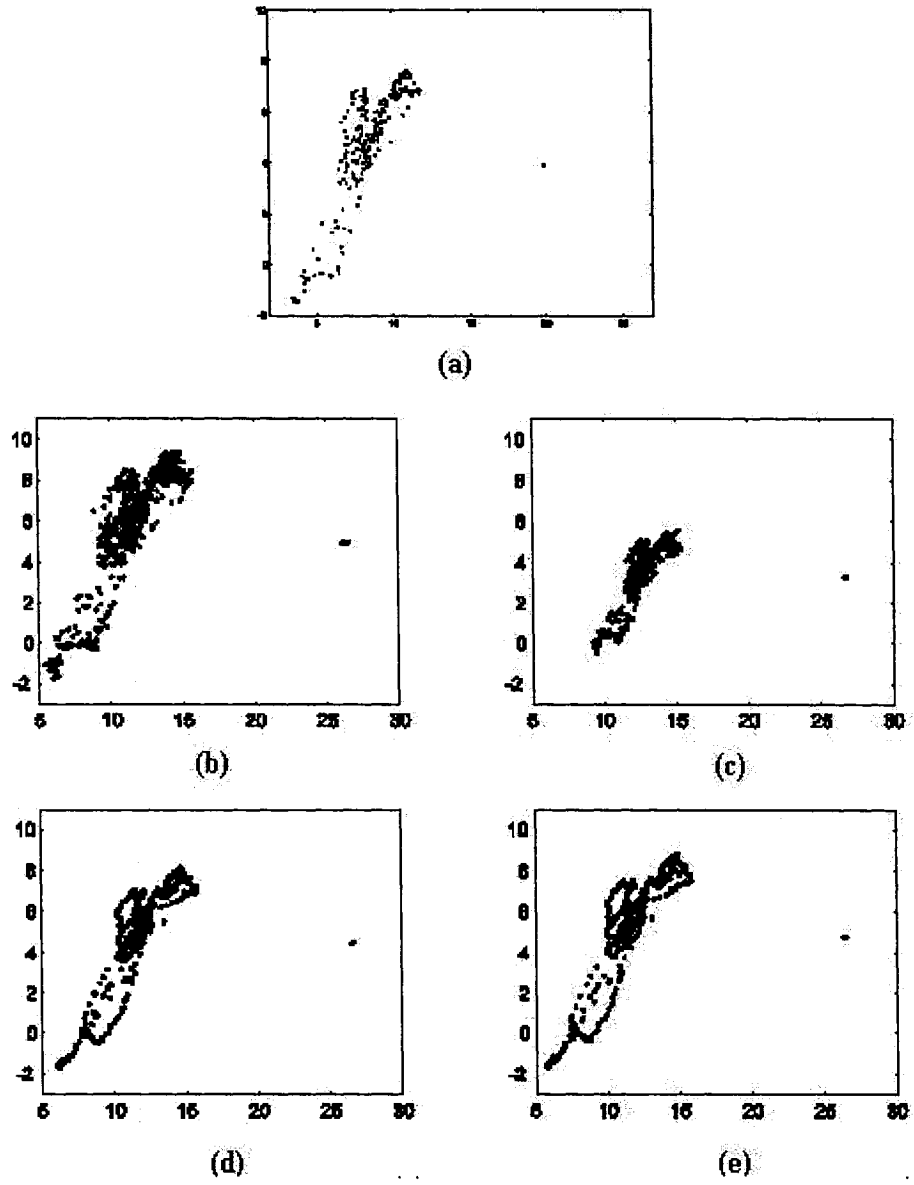
Figure 4-25 Eigen region for *Pot* image. (a) Original. (b) Noisy. (c) with no mask. (d) CSF-DWT mask. (e) CSF-BPA mask.
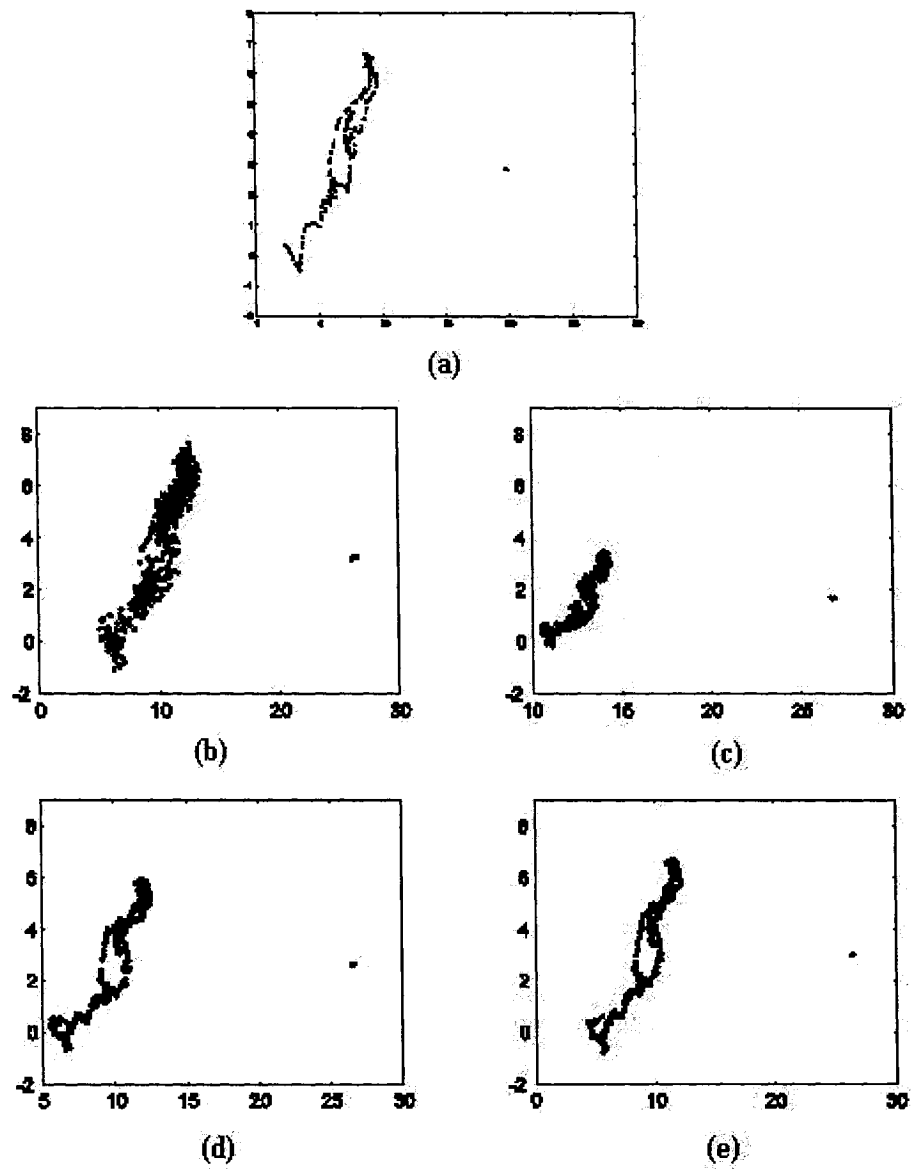
Figure 4-26 Eigen region for *Ultrasound scan* image. (a) Original. (b) Noisy. (c) with no

mask. (d) CSF-DWT mask. (e) CSF-BPA mask.

4.5 No-Reference quality index analysis

Human observers can easily assess the quality of a distorted image without examining the original image as a reference. By contrast, designing objective No-Reference (NR) quality measurement algorithms has always been a very difficult task. Currently, NR quality assessment is feasible only when prior knowledge about the types of image distortion is available. In an attempt to address this issue related to design of a no-reference quality index, Wang *et. al.* [57] established a JPEG image database and conducted a variety of subjective experiments on the images. Based on their experimental results, they came up with a computational and memory efficient NR quality assessment model for JPEG images. The NR quality assessment algorithm is an objective metric where, unlike PSNR and MSE, the quality of the reconstructed image is assessed without the necessity of having to examine the original image.

In this thesis work, we calculate the NR quality index, using the algorithm proposed in [57], to further evaluate the performance of the compression scheme incorporating the proposed masking method. Table 4-7 contains the NR quality index measurement values for all the images under consideration. These measurements have been made over a scale varying between 1 and 10, where an index value of 10 is considered equivalent to good image quality and 1 for poor visual quality. From these measurements, it can be seen that the perceived quality in about 50% of the cases, agrees with the subjective assessment results and the rest with the objective assessment. This shows the inability of the quantitative analysis metrics to predict closely the perceived quality of a reconstructed compressed image. For most of the images, under consideration, the BNC 17/11 wavelet used in conjunction with the CSF-BPA mask is found to offer a higher value for quality

index in comparison to the standard CDF 9/7 wavelet and the BNC 22/14 wavelet. The improvement in the compression performance (in terms of the final visual quality) by using the BNC wavelets with integrated CSF-BPA mask over the previously used CSF-DWT mask clearly corroborates with the previously mentioned assessment results.

Table 4-7 No-Reference Quality Index measurement results

Image	Noisy	No mask applied			CSF-BPA mask applied		
		CDF 9/7	BNC 17/11	BNC 22/14	CDF 9/7	BNC 17/11	BNC 22/14
Lena	7.55	**8.90**	8.98	9.00	8.76	8.75	8.79
Barbara	7.94	**8.78**	8.77	8.74	8.71	8.68	8.70
Baboon	7.79	8.89	8.82	8.82	9.02	**9.06**	9.01
Boats	8.68	**9.01**	**9.01**	8.90	8.75	8.78	8.77
Cambridge	8.73	8.81	8.80	8.80	9.06	**9.07**	9.03
Diatoms	6.34	**9.00**	8.96	8.93	8.93	**9.00**	8.96
MRI	8.93	8.73	**8.74**	8.72	7.48	7.52	7.50
Pot	7.68	8.95	8.85	8.91	8.93	**8.98**	8.96
Text	7.56	9.20	9.11	9.09	9.25	**9.28**	9.25
Ultrasound scan	7.11	**8.80**	8.77	**8.80**	8.75	8.73	8.75
X-Ray	8.32	**9.06**	9.01	9.01	8.52	8.52	8.58

Quality Index: Scale ranges between 1 and 10
1 - Poor visual quality
10 - High visual quality obtained

4.6 Summary

To summarize, in this chapter, the performance of the proposed system has been studied. A performance comparison has been made between the most commonly used JPEG-standard CDF 9/7 biorthogonal wavelet system and two of the recently explored nearly-orthogonal or biorthogonal nearly coiflet (BNC) wavelets, namely, the symmetrical BNC 17/11 wavelet and the asymmetrical BNC 22/14 wavelet, for a variety of standard natural images as well as digital medical images. An extensive evaluation has been carried out using both the quantitative (PSNR) and the qualitative metrics (double blind test) for different wavelet types, over a specific range of compression ratios, for various masking conditions. The presence of the AWGN, which is the most common cause for the image artifacts, is included in this evaluation. The subjective assessment procedure is followed for both the expert and non-expert observers along with a no-reference quality assessment metric.

Overall, from all the simulation results, it can be concluded that the BNC wavelets provide a better PSNR and also shows an enhancement in the perceptual quality when used in conjunction with the proposed CSF-BPA mask. The standard CDF 9/7 wavelet is found to reproduce the edge features better in comparison with the BNC 17/11 and BNC 22/14 wavelets, but the detail pixels are greatly smoothened resulting in blurring of the image. The BNC wavelets perform equally well for most of the image types. At low and medium bit rates, the BNC 17/11 wavelet preserves the high frequency pattern better than the BNC 22/14 wavelet do.

Chapter 5

Conclusions and Future work

5.1 Conclusion

Most of the existing efforts in the field of image compression have focused on developing methods to minimize mathematically tractable, easy to measure, distortion metrics. These distortion measures facilitate reasonable evaluation at high-bit rates. However, at low bit rates these metrics fail to guarantee preservation of important perceptual qualities in the reconstructed images. Integrating models of the human visual system (HVS) into the coding algorithms have the potential to facilitate increased performance, defined in terms of the image quality and bit rate. These perceptually-enhanced coding schemes attempt to distribute the visual errors uniformly over the whole image resulting in the reduction of visual artifacts in the final reconstructed image.

In this thesis work, a digital image compression scheme to achieve an enhancement in the perceptual quality of the reconstructed compressed images has been developed. The proposed system exploits certain important human visual properties like the contrast sensitivity and the contrast masking. As an initial step to the proposed scheme, the effect of contrast masking in the wavelet domain has been investigated. A band peak-average contrast sensitivity function (CSF) masking method has been developed using the Mannos and Sakrison's HVS model. The wavelet transformed coefficients were masked based on their perceptual importance using the derived CSF-BPA mask. The proposed

scheme uses a nearly-orthogonal wavelet in lieu of the biorthogonal CDF 9/7 wavelet to obtain a high compression performance.

The qualitative and quantitative performance of the proposed scheme, in enhancing the performance of the quality of the reconstructed image, has been studied under different masking conditions. The qualitative assessment has been done using the peak signal-to-noise ratio (PSNR). The mean opinion score (MOS) values were obtained from a group of observers in order to subjectively evaluate the quality of the reconstructed image. The subjective evaluation has been done only to compare the artifacts obtained from the transform themselves and not the tiling artifacts. Hence, the use of tiling was avoided whenever possible. More specifically, tiling was used for ultrasound and X-ray image due to their large size, where a tile size of 512 x 512 was employed. In order to perform the subjective testing, some critical regions from the various lossy reconstructed images were selected to be scrutinized. Many small regions of interest were considered instead of whole images because of two main reasons.

1) Some of the test images are too large to be displayed at their original resolution on a computer screen, and reducing their resolution hides most of the compression artifacts.

2) Also, if the whole image is considered, there would have been too much information for a human observer to analyze, and making an accurate and meaningful comparison would have been difficult.

The subjective testing has been restricted to lower bit rates as the human eye often cannot distinguish between a high bit rate lossy reconstruction of an image and the original. The participants involved in the subjective testing came from diverse backgrounds, a few of

116

them had any background in image coding and/or image processing while others were end users with no special technical training. In order to further evaluate the performance of the compression scheme incorporating the proposed masking method, the *No-Reference* (NR) quality assessment has been performed using the algorithm proposed in [50].

A comparison of the visual performance has been made with reference to the conventional scheme. Based on the assessment results, the proposed scheme has been found to preserve the visual quality by maintaining the semblance of the image edges. Considering the simulation results, it can be inferred that the BNC wavelets are capable of achieving state-of-the-art performance levels for lossy wavelet compression for gray scale images. Further, it has been found that longer filters give a better approximation of the transformed image and produces a greater contrast between the edges and the detail pixels after the reduction of the additive Gaussian noise.

The absence of the mask preserves the high frequency components better, but more noise appears at the output and the reconstructed image looks artificial due to the presence of the artifacts. It has been found that when the magnitude of the compression artifacts is close to the visual threshold, the subjective metrics correlate, better than the objective metrics, with the perceptual quality of the reconstructed image. One of the most significant inferences from this work is that the PSNR does not effectively take into account the human visual perception and hence, does not correlate with the mean opinion score. Finally, in order to implement the BNC 17/11 wavelet filter in a FPGA or ASIC hardware, a multiplier-less design has been proposed.

5.2 Scope For Further Research

At the present time, the only drawback for a wavelet-based compression is that it is relatively new and not yet widely supported. Every photo-editing program and web browser supports JPEG, while plugins are needed to view a wavelet-compressed image. There is no public standard for the image portability and each vendor uses his/her own file format. Below are some suggestions that can be considered for developing wavelet-based perceptually enhanced compression and watermarking systems.

1) Wavelets: The work in this thesis has shown as to how the biorthogonal counterpart of coiflet (nearly-orthogonal) wavelets aids in achieving higher compression ratios. In the field of wavelet theory, there have been some recently-explored higher-order multiwavelet systems [60]. It would be desirable to investigate these multiwavelets for image processing applications.

2) 3-D HVS modeling: The proposed work exploits the CSF characteristics of the HVS model in the 2-D domain. In reality, for a better understanding of the vision model, a 3-D model needs to be used. The distortion maps generated using a 3-D non-linear model will lead to the design of better perceptual coder systems.

3) Optimal FPGA design: In this thesis, a general design structure for implementation on a FPGA chip has been proposed. It would be worthwhile to optimize the implementation to obtain a low-power VLSI sub-system.

Appendix - I

Lifting Scheme

The lifting scheme (Fig 1), formally introduced by Sweldens [59], is a well-known method to create biorthogonal wavelet filters from the other existing wavelets. To achieve this, a lazy wavelet is used. The low-pass h and $\tilde{h}$ and the high-pass filters g and $\tilde{g}$ of a lazy wavelet are defined as follows.

$$h(z) = \tilde{h}(z) = 1$$

$$g(z) = \tilde{g}(z) = z^{-1}$$

The input signal x_0 is transformed into the approximation (x) and detail (y) signals using the lazy wavelet transform. Then, the lifting steps, namely, predict and the update, as defined by (1) and (2) are performed. In a predict step, the detail signal at next level is predicted from the current level approximation samples. During the update phase, the approximation signal at next level is updated based on the predicted values of the detail signal. The prediction P (or dual lifting) and update U (or primal lifting) steps improve the initial lazy wavelet properties.

$$y_0[n] = y[n] - P(x[n]) \tag{1}$$

$$x_0[n] = x[n] + U(y_0[n]) \tag{2}$$

Several prediction and update steps are concatenated till the desired wavelet basis is obtained. The prediction and update operators may be a linear or a non-linear combination of x and y, and by construction the lifting scheme is always reversible. Lifted wavelets are known as second generation wavelets. The main advantages of the

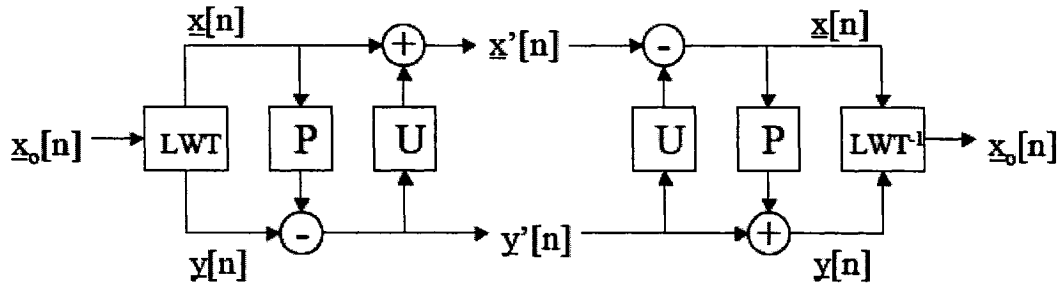lifted wavelets when compared to the new wavelets generated using classical methods are given as follows.


Figure 1: Lifting scheme

1.Inverse existence: Since by structure every lifting step is reversible, the inverse structure exists and this gives the inverse wavelet transform.

2.Critical down-sampling assured: Initial wavelet is modified with existing samples, so no additional information (redundancy) is added.

3.Computational cost reduction: Lifting reduces the computational cost involved in a standard filter implementation to one-half.

4.Memory savings: In-place lifting computation avoids auxiliary memory requirements since lifting outputs from one branch may be saved directly in the other branch.

5.FIR decomposition: Daubechies demonstrated in [13] that every wavelet transform with FIR filters can be decomposed into a finite number of lifting steps.

For images processing applications, the lifting filters are usually developed for the 1-D case and then they are extended to the separable 2-D case by a succession of a vertical and a horizontal 1-D filtering. This structure leads to a 4-band decomposition for each resolution level and the LL (the low-pass vertical and low-pass horizontal band) is successively iterated. For image compression purposes, bands with high frequency components (the HL, LH and HH bands) are not recursively filtered.
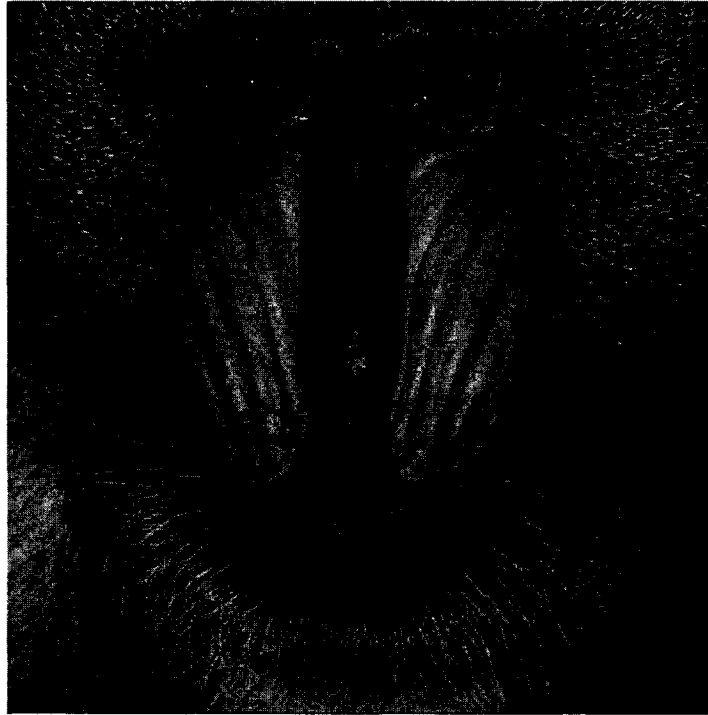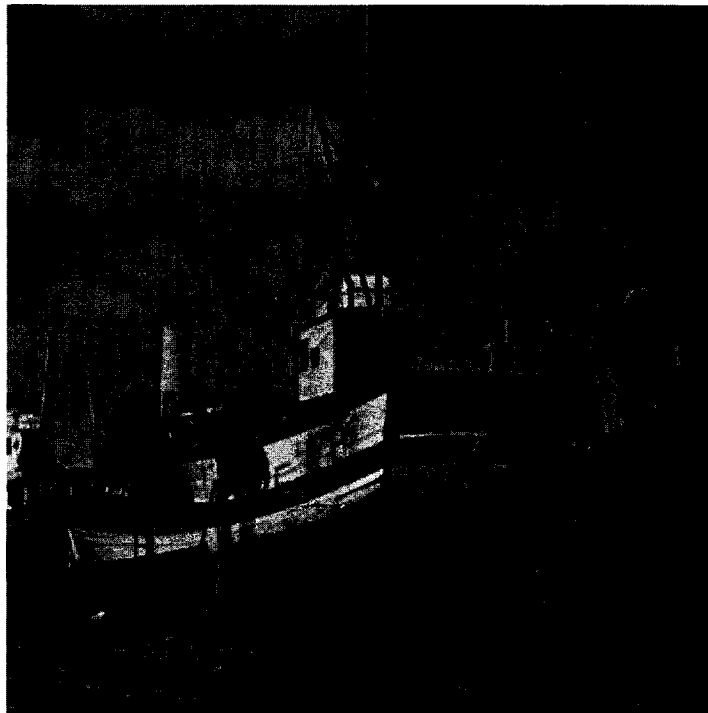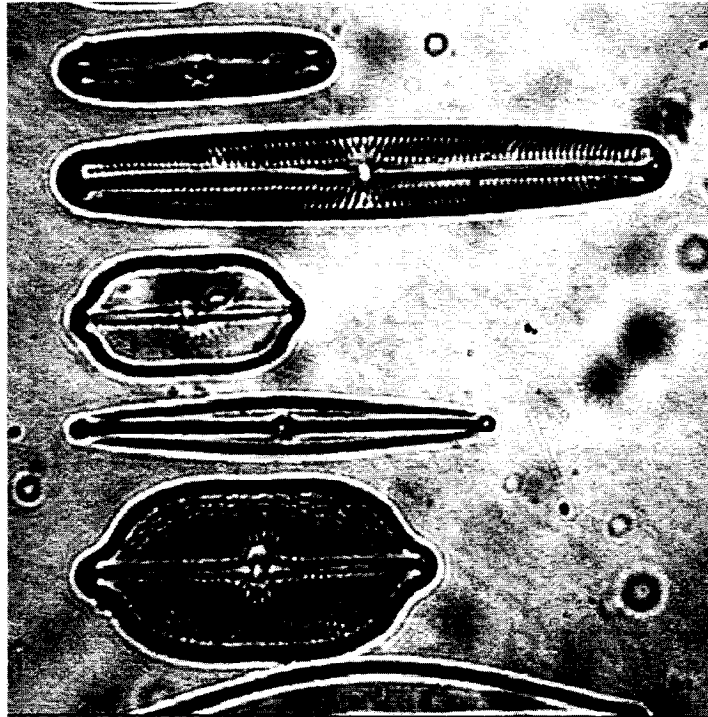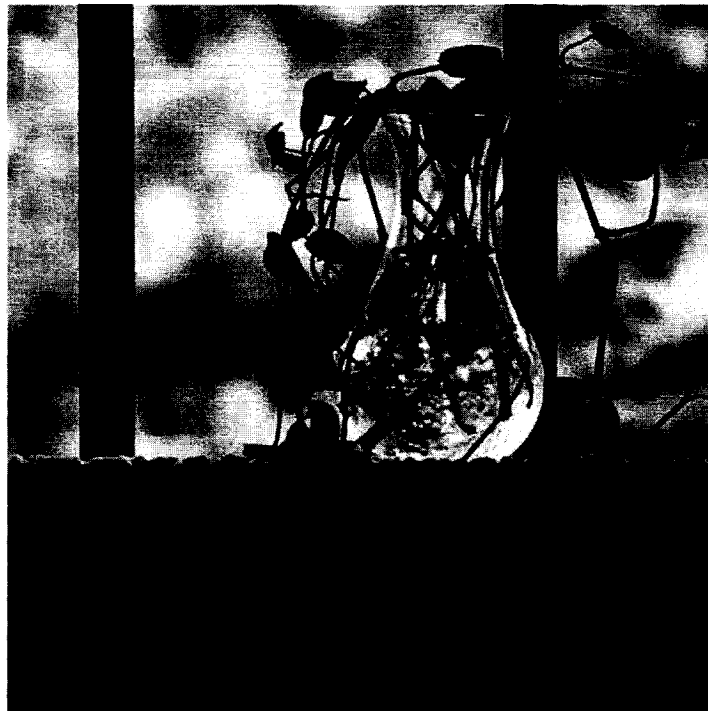
Appendix – II
Original Images

LENA
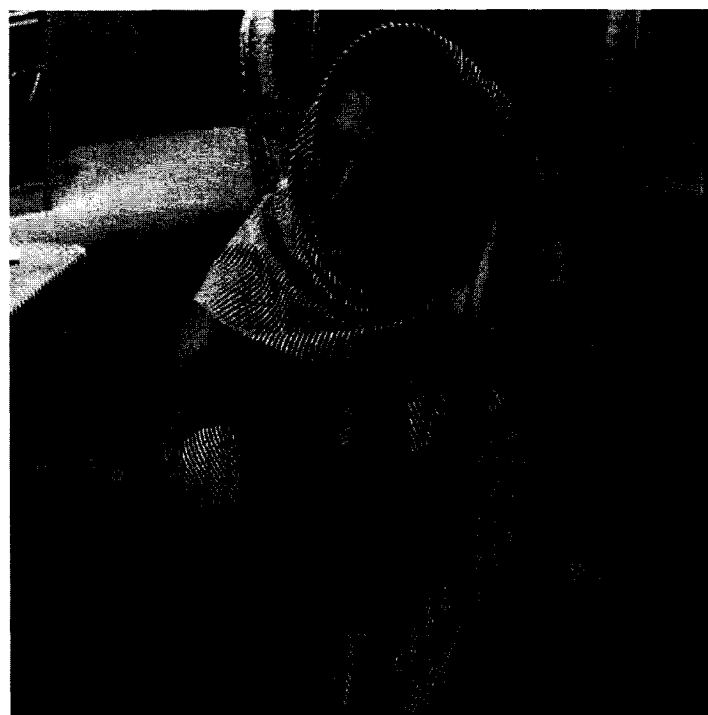
BARBARA

GOLDHILL

BOATS
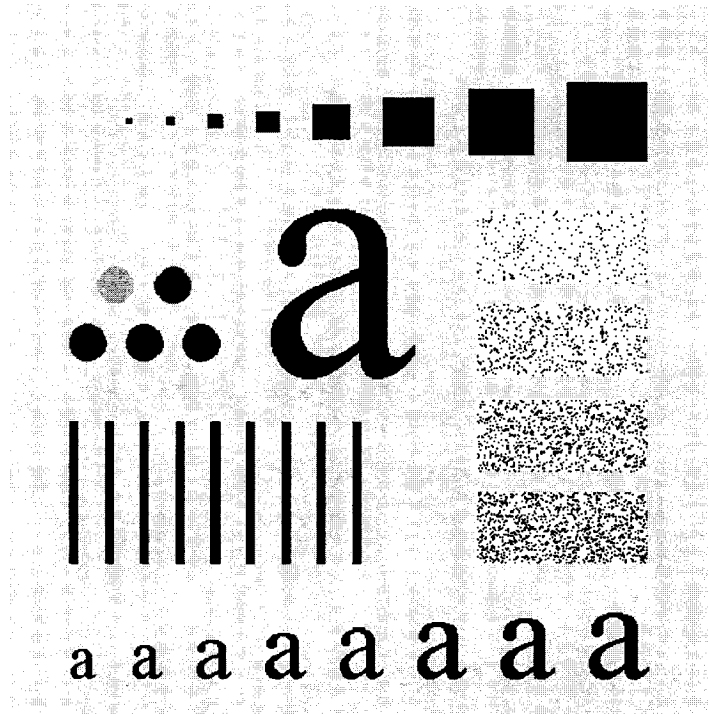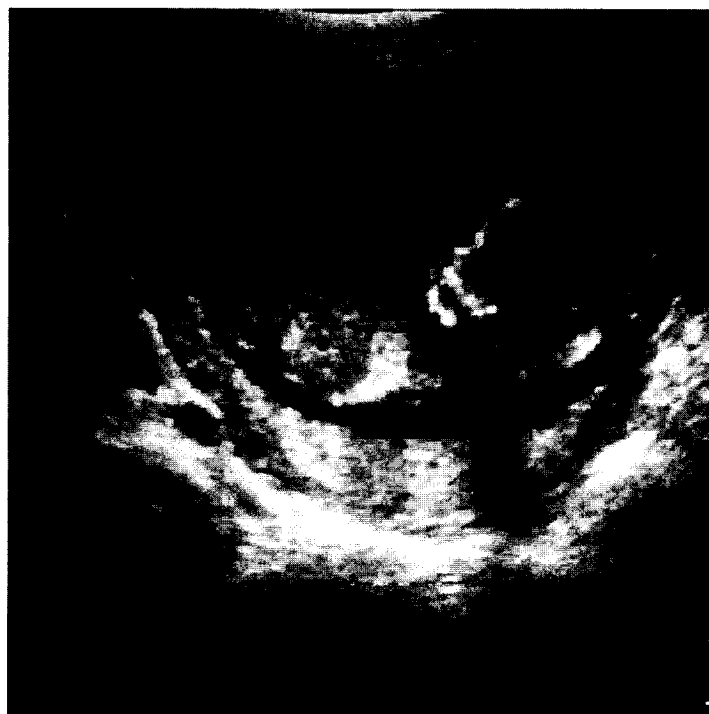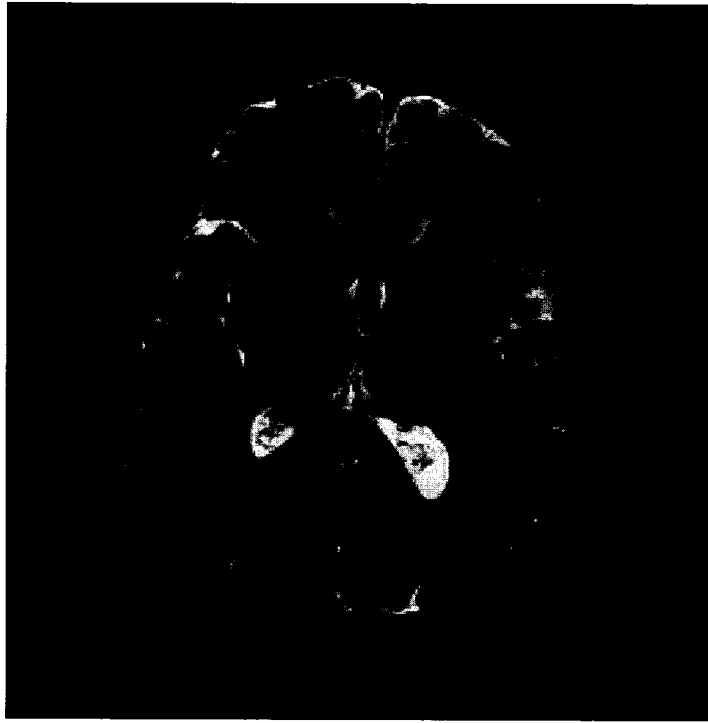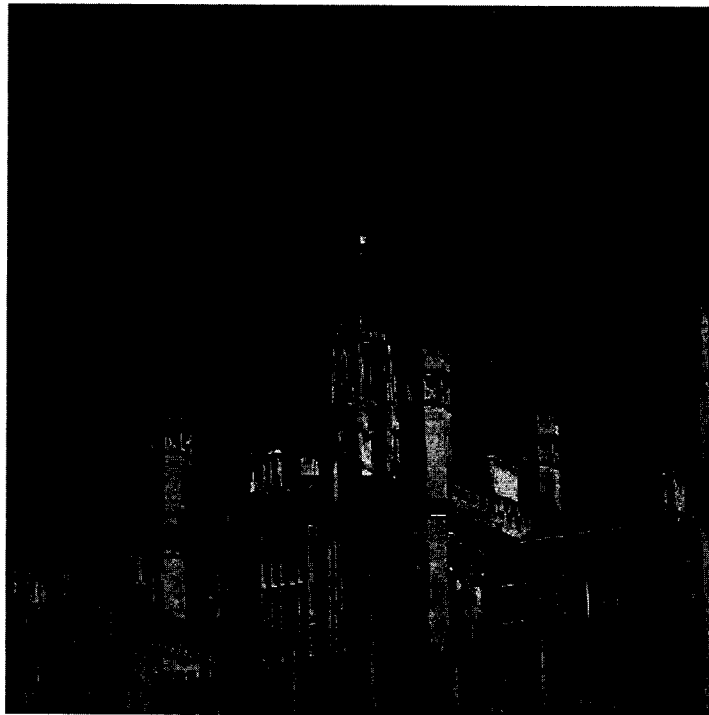
DIATOMS

POT

TEXT

ULTRASOUND

124

MRI SCAN

CAMBRIDGE

125

References

[1] M. Boliek, M. J. Gormish, E. L. Schwartz, E. L., and A. Keith, "Next Generation Image Compression and Manipulation Using CREW", in Proc. ICIP, vol. 3, pp. 567-570, 1997.

[2] B. L., Howard, P. G., and Y. Bengio, "The Z-Coder Adaptive Binary Coder", in Proc. IEEE Data Compression Conference DCC'98, pp. 13-22, Mar. 1998.

[3] J. Oh, S. I. Woolley, T. N. Arvanitis and J. N. Townend, "A Multistage Perceptual Quality Assessment for Compressed Digital Angiogram Images", IEEE Trans. Medical Imaging, vol. 20, no. 12, pp. 1352-1361, Dec. 2001.

[4] D. M. Chandler and S. S. Hemami, "Dynamic Contrast-Based Quantization for Lossy Wavelet Image Compression", IEEE Trans. Image Processing, vol. 14, no. 4, pp. 397-410, Apr. 2005.

[5] A. S. Lewis and G. Knowles, "Image Compression Using the 2-D Wavelet Transform", IEEE Trans. IP, vol. 1, no. 2, pp. 244-250, April 1992.

[6] Dennis P. Curtin, A Short Course in Choosing & Using A Digital Camera, second edition, ShortCourses Publishing Company.

[7] I. Höntsch, and L. J. Karam, "Locally Adaptive Perceptual Image Coding", IEEE Trans. Image Processing, vol. 9, no. 9, pp. 1472-1483, Sept. 2000.

[8] F. Kammoun and M. S. Bouhlel, "A Perceptual Image coding method of high compression rate", Int. Journal of Signal Processing, vol.1, no.1, pp. 79-90, 2004.

[9] R. J. Safranek and J. D. Johnston, "A perceptually tuned subband image coder with image dependent quantization and post-quantization", in Proc. IEEE ICASSP, pp.1945-1948, 1989.

[10] P. P. Vaidyanathan, Multirate Systems and Filter Banks, Prentice Hall, Englewood Cliffs, N.J., 1993.

[11] M. Vetterli and C. Herley, "Wavelets and Filter Banks: Theory and Design", IEEE Trans. Signal Processing, vol. 40, no. 9, pp. 2207-2232, Sep. 1992.

[12] S. G. Mallat, "A theory for multiresolution signal decomposition: the wavelet representation", IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 11, no. 7, pp.674-693, July 1989.

[13] I. Daubechies, Ten Lectures on Wavelets, SIAM, 1992.

[14] G. Strang and T. Nguyen, Wavelets and Filter Banks, Wellesley-Cambridge Press, Wellesley, MA, 1996,

[15] I. Daubechies, "The wavelet transform, time-frequency localization and signal analysis", IEEE Trans. Inf. Theory 36, pp. 961-1005, 1990.

[16] P. de Neve, N. K. Denecker, R. W. Philips and I. Lemahieu, "Compressing CMYK images by removing color redundancy: a perceptual-based quality evaluation", SPIE Proc., vol. 3409, p. 267-276, Sept 1998.

[17] D. Wei, Coiflet-Type Wavelets: Theory, Design and Applications, Ph. D. Thesis Dissertation, University of Texas, Austin, 1998.

[18] F. M. de Saint-Martin, P. Siohan and A. Cohen, "Biorthogonal filterbanks and energy preservation property in image compression", IEEE Trans. Image Processing, vol. 8, pp. 168-178, Feb. 1999.

[19] L. L. Winger, A. N. Venetsanopoulos, "Biorthogonal Nearly Coiflet wavelets for image compression", Signal Processing: Image Commun., vol. 16, pp. 859-869, 2001.

[20] D. Wei, H. Pai and A. C. Bovik, "Antisymmetric Biorthogonal Coiflets for image coding", in Proc. IEEE ICIP, vol. 2, pp. 282-286, Oct. 1998.

[21] T. Cooklev , A. Nishihara, M. Sablatash, "Regular Orthonormal and Biorthogonal wavelet filters", Signal Processing 57, vol. 2, pp. 121-137, Mar. 1997.

[22] C. W. Thomas, G. C. Gilmore, and F. L. Royer, "Models of Contrast Sensitivity in Human Vision", IEEE Trans. Systems, Man and Cybernetics, vol. 23, no. 3, pp. 857-864, May 1993.

[23] L. L. Winger and A. N. Venetsanopoulos, "Biorthogonal modified coiflet filters for image compression", in Proc. IEEE ICASSP, vol. 5, pp. 2681-2684, May 12-15, 1998.

[24] A. J. Ahumuda and C. H. Null, "Image Quality: A Multi-Dimensional Problem", SID'92 Digest, pp. 851-854, 1992.

[25] J. L. Mannos and D. J. Sakrison, "The effects of a visual fidelity criterion on the encoding of images", IEEE Trans. Inform. Theory, vol. 20, no. 4, pp. 525-536, July 1974.

[26] D. G. Pelli, G. S. Rubin, and G. E. Legge, "Predicting the contrast sensitivity of low-vision observers", J. Opt. SOC. Amer. A, vol. 3, no.13, pp. 56, 1986.

[27] T. Movshon and L. Kiorpes, "Analysis of the development of spatial sensitivity in monkey and human infants," J. Opt. Soc. Am. A, vol. 5, no. 12, pp. 2166-2172, Dec. 1998.

[28] P. G. J. Barten, Contrast Sensitivity of the Human Eye and Its Effects on Image Quality, SPIE Press, Bellingham, WA, 1999.

[29] S. Daly, "The visible differences predictor: An algorithm for the assessment of image fidelity", Digital Images and Human Vision, Cambridge, MA, pp. 179-206, 1993.

[30] D. J. Sakrison and V. R. Algazi, "Comparison of line-to-line and two dimensional encoding of random images", IEEE Trans. Inform. Theory, vol. 17, no. 4, pp. 386-398, July 1971.

[31] R. P. Dooley, "Predicting brightness appearance at edges using linear and nonlinear visual describing functions", SPIE Proc. Annu. Meeting, Denver, May 14, 1975.

[32] M. D. Levine, Vision in Man and Machines, New York: McGraw-Hill, 1985. Graphics and Interactive Techniques, pp. 73-82, 1999.

[33] H. A. Peterson, A. J. Ahumuda Jr. and A. B. Watson, "Improved detection model for DCT coefficient quantization", SPIE Proc., vol. 1913, pp. 191-201, Sep. 1993.

[34] R. C. Dorf, The Electrical Engineering Handbook - Circuits, Signals, and Speech and Image Processing, 3rd edition, CRC Press, 1997.

[35] F. Leavitt, (Double) Blind Faith, Psychology Department, Cal State University, Hayward, CA, USA, Natural science, June 6, 2003

[36] CCIR Recommendation 500-3, "Method for the subjective assessment of the quality of television pictures", Tech. rep., ITU, Gemeva, Switzerland.

[37] Dictionary for Clinical Trials, 2nd edition, John Wiley and Sons Ltd., April 2007.

[38] J. Malo, A. Pons, and J. M. Artigas, "Bit allocation algorithm for codebook design in vector quantization fully based on hvs non-linearities for suprathreshold contrasts", Electronics Letters, vol. 31, pp. 1222-1224, 1995.

[39] T. P. O'Rourke, R. L. Stevenson, Y. F. Huang, and D. J. Costello, "Improved decoding of compressed images received over noisy channels", in Proc. IEEE ICIP, pp. 65-68, 1995.

[40] T. N. Pappas, T. A. Michel, and R. O. Hinds, "Suprathreshold perceptual image coding", in Proc. ICIP, pp.237-240, 1996.

[41] A. Said, W. A.Pearlman, "A New Fast and Efficient Image Codec Based on Set Partitioning in Hierarchial Trees", IEEE Trans. Circuits and Systems for Video Technology, vol. 6, no. 3, pp. 243-250, June 1996.

[42] J. M. Shapiro, "Embedded Image Coding Using Zerotrees of Wavelet Coefficients", IEEE Trans. Signal Processing, vol. 41, no. 12, pp. 3445-3462, Dec. 1993.

[43] R. J. Safranek, "JPEG compliant encoder utilizing perceptually based quantization," SPIE Human Vision, Visual Processing, and Digital Display, vol. 1913, pp. 117-126, 1993.

[44] M. Antonini, M. Barlaud, P. Mathieu, and I. Daubechies, "Image coding using wavelet transforms", IEEE Trans. Image Processing, pp. 205-220, 1992.

[45] F. W. Campbell and J. G. Robson, "Application of Fourier analysis to the visibility of gratings", J. Physioi., Land., vol. 197, pp. 551-566, 1998.

[46] Y. K. Lai and C. C. J. Kuo, "Wavelet image compression with optimized perceptual quality", SPIE Proc., vol. 3460, pp. 436-447, Oct. 1998.

[47] A. B. Watson, G. Y. Yang, J. A. Solomon, and J. Villasenor, "Visual thresholds for wavelet quantization error", SPIE Proc., vol. 2657, 1996.

[48] M. G. Albanesi, "Wavelets and human visual perception in image compression", in Proc. IEEE ICPR, pp. 859-863, 1996.

[49] M. Nadenau and J. Reichel, "Compression of color images with wavelets under consideration of the HVS", SPIE Proc. Human Vision and Electronic Imaging, San Jose, CA, vol. 3644, no. 5, pp. 129-140, Jan. 24-29, 1999.

[50] M. J. Nadenau, J. Reichel and M. Kunt, "Wavelet based Color image compression: Exploiting the Contrast Sensitivity Function", IEEE Trans. Image Processing, vol.12, no.1, pp. 58-70, Jan. 2003.

[51] M. Ramos and S. Hemami, "Activity selective spiht coding", SPIE Visual Communications and Image Processing, vol. 3653, pp. 315-326, Jan. 1999.

[52] A. P. Beegan, L. R. Iyer and A. E. Bell, "Design and Evaluation of Perceptual Masks for Wavelet Image Compression", in Proc. 10th IEEE Digital Signal Processing Workshop, Pine Mountain, GA, Oct. 2002.

[53] S. Rout, A. E. Bell, "Color image compression: multiwavelets vs. scalar wavelets", in Proc. IEEE ICASSP, vol. 4, pp. 3501-3504, 2002.

[54] F. W. Campbell and D. G. Green, "Optical and retinal factors affecting visual resolution", J. Physiology, London, vol. 181, no. 3, pp. 576-593, 1965.

[55] K. K. Parhi and T. Nishitani, "VLSI architectures for discrete wavelet transforms", IEEE Trans. VLSI Systems, vol. 1, no. 2, pp. 191-202, June 1993.

[56] J. M. Jou, Y. H. Shiau and C. C. Liu, "Efficient VLSI architectures for the biorthogonal wavelet transform by filter bank and lifting scheme", in Proc. IEEE ISCAS, vol. 2, pp. 529-532, 2001.

[57] Z. Wang, H. R. Sheikh and A. C. Bovik, "No-reference perceptual quality assessment of JPEG compressed images", in IEEE ICIP, vol. 1, pp. 477-480, Sept. 2002. (http://www.cns.nyu.edu/~zwang/files/research/nr_jpeg_quality/index.html)

[58] H. Murase and M. Lindenbaum, "Partial Eigen value Decomposition of Large Images Using Spatial Temporal Adaptive Method", IEEE Trans. Image Processing, vol. 4, no. 5, pp. 620-629, May 1995.

[59] W. Sweldens, "The lifting scheme: A custom-design construction of biorthogonal wavelets", Appl. Comput. Harmon. Anal., vol. 3, no. 2, pp. 186-200, 1996.

[60] S. Rout, A. E. Bell, "Color image compression: multiwavelets vs. scalar wavelets", in Proc. IEEE ICASSP, vol. 4, pp. 3501-3504, 2002.

[61] L. Sugavaneswaran, M. N. S. Swamy, C. Wang, "Masked Nearly-Orthogonal Wavelet-Based Image Compression and its Application to Medical Imaging", in Proc. IEEE MWSCAS, pp. 285-288, Aug 2007