Development of Some Efficient Lossless and Lossy Hybrid Image Compression Schemes

Chandan Singh D Rawat

Department of Electronics & Communication Engineering
National Institute of Technology, Rourkela

Development of Some Efficient Lossless and Lossy Hybrid Image Compression Schemes

Thesis submitted to
National Institute of Technology Rourkela
for the award of the degree

of

Doctor of Philosophy

by

Chandan Singh D Rawat

under the guidance of

Prof. Sukadev Meher

Department of Electronics & Communication Engineering
National Institute of Technology Rourkela
January 2015

Dedicated

to

my family

CERTIFICATE

This is to certify that the thesis titled ***"Development of Some Efficient Lossless and Lossy Hybrid Image Compression Schemes"****, submitted to National Institute of Technology Rourkela(INDIA), by* ***Chandan Singh D Rawat****, Roll No.50709002 for the award of degree of* ***Doctor of Philosophy in Electronics and Communication Engineering****, is a bona fide record of research work carried out by him under our supervision and guidance.*
The thesis, which is based on candidate's own work, has not been submitted elsewhere for a degree/diploma. In our opinion, the thesis is of standard required for PhD in Engineering.
To the best of our knowledge, Mr. Rawat bears a good moral character and decent behaviour.

Prof. G. Haridasan
Ex-Adjunct Professor
VESIT, Mumbai
(Co-supervisor)

Prof. Sukadev Meher
Professor, EC
NIT Rourkela
(Supervisor)

Acknowledgment

I express my indebtedness and gratefulness to my supervisor **Dr.Sukadev Meher,** Professor, Department of Electronics and Communication Engineering, for his continuous guidance and support. His observations and comments helped me to remain focused in achieving my goal. I am obliged to him for his moral support through all the stages during the doctoral research work.

I am also grateful to my co-supervisor **Prof G. Haridasan,** Adjunct Professor VESIT, Mumbai, for his timely comments, guidance and support throughout the course of this work.

My special thanks go to Prof. S Hiremath, Prof. Ajit Kumar Sahoo and Prof. Manish Okade for contributing towards enhancing the quality of the work in shaping this thesis.

I would like to thank my friends Nihar Panda, Ramesh Kulkarni, Manoj Gupta, Ayaskanta Swain, for their encouragement and understanding. Their help can never be penned with words. I acknowledge the help and friendly support from the members of the department staff ans research community.

I specially thank the management and **Prof. Jayalekshmi Nair,** Principal of V.E.S. Institute of Technology, Mumbai for granting study leave during my Ph.D program.

Most importantly, none of this would have been possible without the love and patience of my family. My mother *Smt.Chandra Rawat*, my wife *Komal*, my sister *Pramila*, my son *Prashant* and my daughter *Hetal*, to whom this dissertation is dedicated to, have been a constant source of love, concern, support and strength all these years. I would like to express my heartfelt gratitude to them.

Place: N.I.T. Rourkela
Date: 11^{th} January 2015.

Chandan Singh Rawat

Contents

List of Abbreviations

bpp	Bits per pixel
CR	Compression Ratio
BDCT	Block Discrete Cosine Transform
CALIC	Context Adaptive Lossless Image Coding
CPSNR	Color peak signal to noise ratio
DCT	Discrete Cosine Transform
DCT-F	Discrete Cosine Transform-Fractals
DPCM	Differential Pulse Code Modulation
DWT	Discrete Wavelet Transform
CWT	Continious Wavelet Transform
IWT	Integer Wavelet Transform
EZW	Embedded Zero Tree Wavelet
GAP	Gradient Adjusted Predictor
JPEG	Joint Photographic Experts Group
MSE	Mean Square Error
PSNR	Peak Signal to Noise Ratio
VIFP	Visual Information Fidelity in Pixel Domain
UQI	Universal Quality Index
MSSIM	Mean Structural Similarity Index Measure
SAD	Sum of Absolute Difference
SA-DCT	Shape Adaptive Discrete Cosine Transform
LPA	Linear Polygonal Approximation
ICI	Intersection of Confidence Intervals
SPIHT	Set Partitioning in Hierarchical Trees
SOFM	Self Organizing Feature Maps

VQ	Vector Quantization
LMS	Least Mean Squares
SPIHT-VQ	Set Partitioning in Hierarchical Trees-Vector Quantization
HSPIHT-VQ	HVS based Set Partitioning in Hierarchical Trees-Vector Quantization
CALIC-SPS	Context Adaptive Lossless Image Coding-Spatial Prediction Structures

List of Symbols

x, y	Discrete spatial co-ordinates in 2-D Cartesian plane
(x, y)	Arbitrary pixel location in an image
M	Number of rows in an image
N	Number of columns in an image
$f(x, y)$	Original uncompressed image; a pixel (or, pixel value) at an arbitrary location, (x, y)
$g(x, y)$	Sub-image; image block
$f'(x, y)$	Predicted pixel value
$f''(x, y)$	Modified predicted pixel value
e	Error in prediction
ϵ	Modified prediction error
$\phi(x, y)$	Compressed image
$\tilde{f}(x, y)$	Decompressed image
$\bar{f}(x, y)$	Average (mean) of $f(x, y)$
$\bar{\tilde{f}}(x, y)$	Average (mean) of $\tilde{f}(x, y)$
σ_f^2	Variance of uncompressed image$f(x, y)$
$\sigma_{\tilde{f}}^2$	Variance of decompressed image$\tilde{f}(x, y)$
$\sigma_{f\tilde{f}}$	Covariance of original image $f(x, y)$ and decompressed image $\tilde{f}(x, y)$
η	Additive White Gaussian noise
ρ	Compression Ratio, CR (unitless)
W	Average word length / code length (unit: bpp)
P	input vector for neural network
p	Element of vector P
Q	Quality parameter for JPEG compression

$\Re$	Set of real numbers
$\mathbb{Z}$	Set of integers
t	iteration number
w	weight(LMS updation)
α	Learning parameter (LMS updation)
U_x	Adaptive shape neighborhood
γ	Hard thresholding coefficient in SA-DCT domain
T	Parameter controlling value of γ
D_{tsh}	Distance Threshold

List of Figures

List of Tables

Abstract

Digital imaging generates a large amount of data which needs to be compressed, without loss of relevant information, to economize storage space and allow speedy data transfer. Though both storage and transmission medium capacities have been continuously increasing over the last two decades, they dont match the present requirement. Many lossless and lossy image compression schemes exist for compression of images in space domain and transform domain. Employing more than one traditional image compression algorithms results in hybrid image compression techniques.

Based on the existing schemes, novel hybrid image compression schemes are developed in this doctoral research work, to compress the images effectually maintaining the quality.

Chapter 1

Introduction

1.1 Introduction to Image Compression

A picture is worth a thousand words. This expresses the essential difference among human ability to perceive linguistic information and visual information. For the same message, a visual representation tends to be perceived as being more efficient than the spoken or written words. The processing of language is inherently serial. Words and their meanings are recorded or perceived one at a time in a causal manner. Visual information, on the other hand, is processed by massively parallel interconnected networks of processing units. In the mammalian visual system, this parallelism is evident from the retina right through to the higher-order structures in the visual cortex and beyond. The efficiency of such parallel architectures over serial processing is reflected by the efficiency with which we process images over language.

In today's modern era, multimedia technology has tremendous impact on human lives. Image is one of the most important media contributing to multimedia. Information transmission is the key means to acquire and give the knowledge or data related to a particular event. For example: video conferences, medical data transfer, business data transfer and so on, require much more image data to be transmitted and stored on-line. Due to the internet, the huge information transmissions take place. The processed data required much more storage, computer processor speed and much more bandwidth for transmission. While the advancement of the computer storage technology continues at the rapid rate. The means for reducing the storage requirement of image is still needed in most of the situations. And hence it is highly desirable that the image be processed, so that efficient storage, representation and transmission of the image can be worked out. The processes involve one of the important tasks - Image Compression. Methods for digital image compression have been the subject of research over the past three decades [27].

Recently, the need for efficient image compression systems can be seen. In the rapidly growing field of Internet applications, not only still images but also small

image sequences are used to enhance the design of private and commercial web pages [28].

Meeting bandwidth requirements and maintaining acceptable image quality simultaneously are a challenge. Continuous rate scalable applications can prove valuable in scenarios where the channel is unable to provide a constant bandwidth to the application [29]. The goal of image compression is to obtain a representation that minimizes bit rate with respect to some distortion constraint. Typical compression techniques achieve bit rate reduction by exploiting correlation between pixel intensities [30]. The performance of any image compression scheme depends upon its ability to capture characteristic features from the image, such as sharp edges and fine textures, while reducing the number of parameters used for its modeling [32]. Image compression is one of the most important and successful applications of the wavelet transform [31]. Wavelets are mathematical functions that provide good quality compression at very high compression ratios, because of their ability to decompose signals into different scales or resolutions. The standard methods of image compression come in numerous ranges. Most of the well-established compression schemes use the bi-variate Discrete Wavelet Transform (DWT) [34] on wavelet-based image coding. At high compression rates, wavelet-based methods provide much better image quality in comparison with the JPEG (Joint Photogrphic Experts Group) standard, which relies on the discrete cosine transform (DCT). The good results obtained from DWT are due to multiresolution analysis, which essentially brings out information about the statistical structure of the image data. The current most popular methods rely on removing high frequency components of the image by storing only the low frequency components (e.g., DCT based algorithms). This method is used on JPEG (still images), MPEG (motion video images), H.261 (Video Telephony on ISDN lines), and H.263(Video Telephony on PSTN lines) compression algorithms.

The compression techniques can be classified as: lossless methods and lossy methods. The first class is composed of those methods which reconstruct an im-

age identical to the original; the second comprises compression methods which lose some image details after their application: the reconstruction is an approximation of the original image [36]. Well known JPEG based on DCT is lossy compression techniques with relatively high compression ratio which is done by exploiting human visual perception [62]. For the lossy compression, some irrelevant data will be thrown away during the compression. The recovered image is only an approximated version of the original image. This implies that the reconstructed image is always an approximation of the original image. Although some information loss can be tolerated in most of these applications, there is certain image processing applications that demand no pixel difference between the original and the reconstructed image [37, 38, 63].

Fractal image compression is a lossy compression method, so there will be data losses in compressed image [60]. For fractal coding, an image is represented by fractals rather than pixels. Each fractal is defined by a unique Iterated Function System (IFS) consisting of a group of affine transformations. Therefore, the key point for fractal coding is to find fractals which can best approximate the original image and then to represent them as a set of affine transformations [61].

Standard fractal coding methods rise above many other image coding techniques in the sense that it maintains high image quality after decoding but presents high compression ratios during encoding. Rather than lossy compression with relatively high compression ratio, mathematical lossless compression techniques are favored in this field [62]. A lossless scheme typically achieves a compression ratio of the order of two, but will allow exact recovery of the original image from the compressed version [63].

There is no partial reduction on data while performing the compression. The exact copy of the original image can be completely recovered. Lossless image compression algorithms are divided into *sequential* algoritms like (Fast Efficient and Lossless Image Compression System) FELICS [136], (Low Complexity Lossless Compression for Images) LOCO-I [65], (Context Adaptive Lossless Image Com-

pression) CALIC [66], some new context-based algorithms [67] and transform-based ones like lossless SPIHT (set partitioning in hierarchical trees) [57, 68]. The choice of the compression method, lossy or lossless, depends on the application [69].

1.2 Compression Metrics

The aim of image compression is to represent an image in a compact form while preserving the information content as far as possible. *Compression efficiency* is the principal parameter of a compression technique, but it is not sufficient by itself. It is simple to design a compression algorithm that achieves a low bit-rate, but the challenge is how to preserve the quality of the reconstructed image at the same time.

$$CompressionRatio\ (\rho) = \frac{Data\ Size\ of\ Input\ Image}{Data\ Size\ of\ Output\ (Compressed)\ Image} \tag{1.1}$$

$$bits\ per\ pixel\ (W) = \frac{Size\ of\ Compressed\ File\ in\ bits}{Total\ pixels\ in\ Image} \tag{1.2}$$

The two main criteria of measuring the performance of an image compression algorithm thus are *compression efficiency* and *distortion* caused by the compression algorithm. The two major compression metrics used are *compression ratio* and *bits per pixel, bpp*. The compression ratio denoted by ρ is a unitless parameter which has a minimum value of 1. Bits per pixel(bpp), denoted by W, expresses the **average code word length** required in bits to represent a pixel value in the image. For gray-scale uncompressed images, the value of bpp is **8** and for color images, the value is **24**. The standard technique to measure the compression performance (quality) using these metrics is to fix a certain compression ratio or bit-rate and then compare the *distortion* caused by the different compression techniques.

The distortion metrics, used to find the compression quality, are described in

the next section.

1.3 Distortion Metrics

Distortion measures can be divided into two categories: subjective and objective measures. A distortion measure is said to be subjective, if the quality is evaluated by human beings. Images are viewed by a number of viewers, and their evaluations are averaged [3]. The weakest point about this method is the subjectivity at the first place. It is very difficult to establish a single set of people that everyone could consult to get quality evaluation of their pictures. Moreover, the definition of distortion highly depends upon the application, that is, the best quality evaluation is not always made by people at all. In the objective measures, the distortion is calculated as the difference between the original and the reconstructed image by a predefined function. It is assumed that the original image is perfect. All changes are considered as occurrences of distortion, no matter how they appear to a human observer.

The quantitative distortion of the reconstructed image is commonly measured by the *mean absolute error* (MAE), *mean square error* (MSE), *peak signal-to-noise ratio* (PSNR) [40]. Let the original image and the decompressed image be represented by $f(x, y)$ and $\tilde{f}(x, y)$ respectively. Here x and y represent the discrete spatial coordinates of the digital image. Let the image be of size $M \times N$ pixels, i.e. $x = 1, 2, 3....M$ and $y = 1, 2, 3, ...N$. Then MAE and MSE are defined in as

$$MAE = \frac{\sum_{x=1}^{M} \sum_{y=1}^{N} \left| \tilde{f}(x, y) - f(x, y) \right|}{M \times N} \tag{1.3}$$

$$MSE = \frac{\sum_{x=1}^{M} \sum_{y=1}^{N} (\tilde{f}(x, y) - f(x, y))^2}{M \times N} \tag{1.4}$$

Higher the MAE, poorer the quality. So is the MSE. But, MAE represents noise

(error) voltage, in absolute average sense, while MSE represents noise (error) power. The PSNR is defined using logarithmic scale in dB, where higher dB value represents better quality. It is a ratio of peak signal power to noise power. Since the MSE represents the peak noise power, and the peak signal power is unit in case of normalized image signal, the image metric PSNR is defined in(1.5) as

$$PSNR = 10\log_{10}\frac{1}{MSE}, \qquad dB \tag{1.5}$$

for a normalized image. Higher the MSE, lower will be the PSNR and poorer the quality. Thus, for a high quality image the PSNR may be very high (say $40dB$ - $60dB$) whereas MSE will be very low($10^{-4} - 10^{-6}$).

For color images, the color peak signal to noise ratio(CPSNR) [43] in dB is used as performance measure. The CPSNR is defined as

$$CPSNR = 10\log_{10}(\frac{1}{3}[MSE_R + MSE_G + MSE_B]) \tag{1.6}$$

where MSE_R, MSE_B and MSE_G represent the MSE values in red, green and blue channels respectively.

Though this image metric is popularly used for evaluating the quality of the restored images and thereby the capability and efficiency of compression process, it does not give true indication of the distortion introduced by compression process. In addition to these parameters new metrics: *universal quality index* (UQI),*mean structural similarity index measure* (MSSIM) and *visual information fidelity* (VIF) [70] and a traditional metric, *sum of absolute differences* (SAD) are used as distortion measures to evaluate the distortions in the image due to compression.

These parameters are discussed below:

Universal Quality Index (UQI) [44]

The UQI is modeled considering three different factors:

(i) Loss of correlation

(ii)Luminance distortion

(iii) Contrast distortion.

It is defined by

$$UQI = \frac{\sigma_{f\tilde{f}}}{\sigma_f \sigma_{\tilde{f}}} \frac{2\bar{f}\bar{\tilde{f}}}{\bar{f}^2 + \bar{\tilde{f}}^2} \frac{2\sigma_f \sigma_{\tilde{f}}}{\sigma_f^2 + \sigma_{\tilde{f}}^2} \tag{1.7}$$

where

$$\bar{f} = \frac{1}{M \times N} \sum_{x=1}^{M} \sum_{y=1}^{N} f(x,y)$$

$$\bar{\tilde{f}} = \frac{1}{M \times N} \sum_{x=1}^{M} \sum_{y=1}^{N} \tilde{f}(x,y)$$

$$\sigma_f^2 = \frac{1}{M \times N} \sum_{x=1}^{M} \sum_{y=1}^{N} (f(x,y) - \bar{f}^2)$$

$$\sigma_{\tilde{f}}^2 = \frac{1}{M \times N} \sum_{x=1}^{M} \sum_{y=1}^{N} (\tilde{f}(x,y) - \bar{\tilde{f}}^2)$$

$$\sigma_{f\tilde{f}} = \frac{1}{M \times N} \sum_{x=1}^{M} \sum_{y=1}^{N} (f(x,y) - \tilde{f})(\bar{\tilde{f}}(x,y) - \bar{\tilde{f}})$$

The UQI, defined here, consists of three components. The first component is the correlation coefficient between the original image, f and the restored image, $\tilde{f}$ that measures the degree of linear correlation between them, and its dynamic range is [-1,1]. The second component, with a range of $[0, 1]$, measures the closeness between the average luminance of f and$\tilde{f}$. It reaches the maximum value of 1 if and only if f equals $\tilde{f}$. The standard deviations of these two images, σ_f and $\sigma_{\tilde{f}}$ are also regarded as estimates of their contrast-levels. So, the third component is

necessarily a measure of the similarity between the contrast-levels of the images. It ranges between 0 and 1, and the optimum value of 1 is achieved only when $\sigma_f = \sigma_{\tilde{f}}$.

Hence, combining the three parameters:

(i) correlation,

(ii) average luminance similarity and

(iii) contrast-level similarity, the new image metric: universal quality index (UQI) becomes a very good performance measure.

Mean Strucural Similarity Index Measure (MSSIM) [45, 71, 75]

It is a method for measuring the similarity between two images. Its value lies between $[0, 1]$. The MSSIM is designed to improve on traditional metrics like PSNR and MSE, which have proven to be inconsistent with the human eye perception. It is based on human visual system. The MSSIM is modeled considering three different factors

i) Local luminance similarity

ii)Local contrast sensitivity and

iii) Local structure similarity which are based on weighted local image statistics

$$MSSIM = \frac{2\mu_f\mu_{\tilde{f}} + C_1}{\mu_f^2 + \mu_{\tilde{f}}^2 + C_1} \frac{2\sigma_f\sigma_{\tilde{f}} + C_2}{\sigma_f^2 + \sigma_{\tilde{f}}^2 + C_2} \frac{\sigma_{f\tilde{f}} + C_3}{\sigma_f\sigma_{\tilde{f}} + C_3} \tag{1.8}$$

where μ_f and $\mu_{\tilde{f}}$ are the respective local sample means of $f(x, y)$ and $\tilde{f}(x, y)$ and σ_f and $\sigma_{\tilde{f}}$ are the respective local sample standard deviations of $f(x, y)$ and $\tilde{f}(x, y)$ and $\sigma_{f\tilde{f}}$ is the sample cross correlation of $f(x, y)$ and $\tilde{f}(x, y)$ after removing their means. The coefficients C_1, C_2 and C_3 are small positive constants that stabilize each term so that the near zero sample means, variances, or correlations do not lead to numerical instability [72].

Visual Information Fidelity

The VIF proposed by Hamid R Sheikh *et al.* is a recent image quality assessment criterion that consistently outperforms almost all other criteria. It treats image

quality assessment as an information fidelity problem [135]. At LIVE [73, 74], VIF was developed for image and video quality measurement based on natural scene statistics. Image quality assessment is done based on information fidelity, where the channel imposes fundamental limit on how much information could flow from the source(the reference image) through the channel(the image distortion process) to the receiver(human observer).

VIF is derived from a quantification of two mutual information quantities. First, the mutual information between the input and the output of HVS channel when no distortion is present (called as reference information) and mutual information between input of distortion channel and output of HVS channel for the test signal are determined. To quantify the information quantities, stochastic models for source, distortion and HVS are needed. The source or reference image is modeled by wavelet domain Gaussian Scale Mixture. The VIF distortion model assumes that the image distortion can be roughly described locally as a combination of uniform wavelet domain energy attenuation with subsequent independent additive noise as

$$d = gc + v \tag{1.9}$$

where c and d are random vectors extracted from same location in same wavelet sub-band in reference and distorted images respectively, g represents scalar deterministic gain factor, while v is independent zero mean white Gaussian noise. In the VIF receiver model, the visual distortion process is modeled as zero mean, stationary, additive white Gaussian noise process in wavelet transform domain, mainly to account for internal neural noise.

Mutual information between C and E quantifies the information the brain could extract ideally from the reference image, whereas mutual information between C and F quantifies the corresponding information that could be extracted from the test image. Thus, given the statistical models of source, channel distortion and

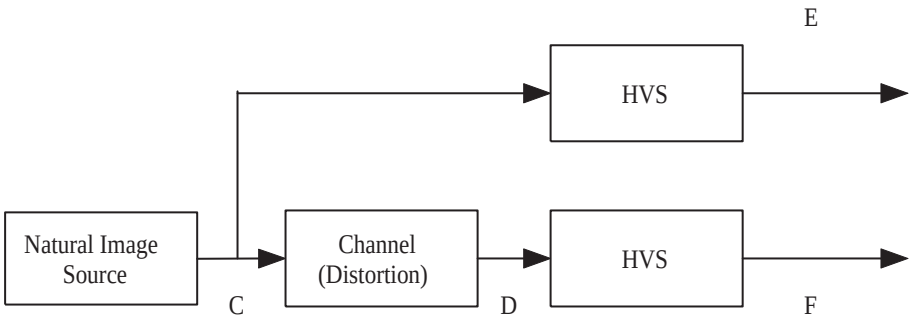


Figure 1.1: Model of VIF [5]

the receiver, the VIF is defined as

$$VIF = \frac{DistortedImageInformation}{ReferenceImageInformation} \tag{1.10}$$

$$VIF = \frac{I(C:F|z)}{I(C:E|z)} \tag{1.11}$$

The VIF has a distinction over traditional quality assessment methods, a linear contrast enhancement of the reference image that does not add noise to it will result in a VIF value larger than unity, thereby signifying that the enhanced image has a superior visual quality than the reference image. No other quality assessment algorithm has the ability to predict if the visual image quality that has been enhanced by a contrast enhancement operation. VIF is of two types: wavelet domain version and pixel domain version. The wavelet domain version is more complex [135]. In this thesis, pixel domain version VIFP [73, 74] is used as the compression performance distortion parameter.

Sum of Absolute Differences (SAD) [76]

SAD is a widely used, extremely simple metric for measuring the similarity between image blocks. It works by taking the absolute difference between each pixel in the original block and the corresponding pixel in the block being used for comparison. These differences are summed to create a simple metric of block similarity. The sum of absolute differences is used for a variety of purposes, such as the generation of disparity maps for stereo images [42], object recognition [4] and motion estimation for video compression [25].

Correlation based matching typically produces dense depth maps by calculating the disparity at each pixel within a neighborhood. This is achieved by taking a square window of certain size around the pixel of interest in the reference image and finding the homologous pixel within the window in the target image, while moving along the corresponding scan line. The goal is to find the corresponding (correlated) pixel within a certain disparity range that minimizes the associated error and maximizes the similarity.

The matching process involves computation of the similarity measure for each disparity value, followed by an aggregation and optimization step. Since these steps consume a lot of processing power, there are significant speed-performance advantages to be had in optimizing the matching algorithm. The images can be matched by taking either left image as the reference (left-to-right matching, also known as direct matching) or right image as the reference (right-to-left matching, also known as reverse matching) [14].
The SAD adds up the absolute differences between corresponding elements in the current and the reference block which is defined as

$$SAD = \sum_{i=1}^{M_1} \sum_{j=1}^{N_1} (|c_{ij} - r_{ij}|) \tag{1.12}$$

where $r_{i,j}$ are the elements of the reference block, $c_{i,j}$ are the elements of the

current block and M_1 and N_1 are the block sizes. Thus the computation of SAD is divided into three steps:

1. Compute differences between corresponding elements, $d_{ij} = c_{ij} - r_{ij}$
2. Determine the absolute value of each difference, $|d_{ij}|$
3. Add all absolute values.

1.4 Backgound and Scope

Image compression may be achieved by hybrid techniques, which would offer higher compression ratio than the other techniques, keeping the quality of reproduced image identical in both cases. Hybrid techniques combine transform and predictive coding. Fast transformations are performed on a block of data in one dimension, and the results used to predict in the other dimension to further decorrelate the image data [78].

Fractal image compression [23] can be combined with different lossless compression techniques to achieve high compression ratio and low computation time as well.

Sunil Kumar and R.C. Jain [79] have presented a fast image compression technique as well as its Progressive Image Transmission (PIT) version using fractal, which made use of a small pool of domains extracted using visually significant patterns. The affine transformations for an edge block were obtained by using its edge characteristics instead of minimum mean square error criterion. When simulated, their method was computationally simple, gave faster encoding speed and achieved good fidelity at relatively higher compression ratios than other fractal based techniques.

Vijaya Prakash *et al.* [80] have proposed a technique to enhance the data compression technique. A new DCT and Quantization (DCTQ) architecture have been designed in their work for performing image compression. Compression of

image data could be achieved by employing the DCT which is a kind of image transform. Later, compression has been achieved by performing quantization of the DCT data coefficients.
Satish Singh *et al.* [81] have discussed the pros and cons of diverse transform-based image compression models in their detailed literature survey which they have presented.

Sarantos Psycharis [82] has presented the transformation techniques that are widely used for lossy compression in order to compare them and extract conclusions for their efficiency. The images used are texture and medical images, while the methods for compression are mainly based on the DCT. For each of the above methods, the signal-to-noise ratio is computed. The methods are applied to the Vocational Training School in the Greek Education System for teaching aspects related to networks and image compression. The effectiveness of their approach was evaluated by comparing the performances of the sample students and comparing the outcomes with those of a traditional teaching approach.

Alexander Wong and William Bishop [83] have addressed the image quality issue by presenting a new algorithm that provides flexible and customizable image quality preservation by introducing an adaptive thresholding and quantization process based on content information such as edge and texture characteristics from the actual image. The algorithm is designed to improve visual quality based on the human vision system. Experimental results from the compression of various test images show noticeable improvements both quantitatively and qualitatively relative to baseline implementations as well as other adaptive techniques.

Ashutosh Dwivedi *et. al.* [84] have proposed a novel hybrid image compression technique. Their technique inherited the properties of localizing the global spatial and frequency correlation from wavelets and classification and functional approximation tasks from modified forward-only counter propagation neural network (MFOCPN) for image compression [41]. Several benchmark test images are used to investigate usefulness of the proposed technique. Results of the technique

show an enhancement in performance measures with respect to decoded picture quality and compression ratios compared to the existing wavelet and neural network based image compression techniques.

Mascher-Kampfer *et al.* [85] have related rate-distortion performance measured in PSNR to the matching scores as obtained by the recognition systems. JPEG2000 and SPIHT are correctly predicted by PSNR to be the most suitable compression algorithms to be used in fingerprint and face recognition systems. Fractal compression is identified to be least suitable to the use in the investigated recognition systems, although PSNR suggests JPEG to deliver worse recognition results in the case of face imagery. JPEG compression performs surprisingly well at high bit rates in face recognition systems, although the low PSNR performance observed.

Muhammad Azhar Iqbal *et al.* [86] have provided the implementation of a compression methodology that utilizes curvelet coefficients with SPIHT encoder. The methodology comprises three phases:

- Transformation of the stimulus image into the curvelet coefficients.
- Threshold-based selection mechanism of prominent coefficients out of different scales.
- Application of lossy SPIHT encoding technique on selected significant coefficients.

SPIHT takes advantage of the multi-scale nature of curvelet transform and eliminates the statistical and subjective redundancies. The empirical results on standard test images show higher PSNR than a few of the earlier approaches. It actually strengthens the idea of using curvelet transform in place of wavelet transform so as to attain lesser bits to represent more prominent features.

Osman Sezer *et al.* [87] propose a block-based transform optimization and associated image compression technique that exploits regularity along directional image singularities. The directionality comes out as a byproduct of the proposed

optimization rather than a built-in constraint. The work classifies image blocks and uses transforms that are optimal for each class, thereby decomposing image information into classification and transform coefficient information. The transforms are optimized using a set of training images. Their algebraic framework allows straight forward extension to non-block transforms, permitting them even to design sparse lapped transforms that exploit geometric regularity. They use an EZW/SPIHT like entropy coder to encode the transform coefficients to show that their block and lapped designs have competitive rate-distortion performance. Their work can be seen as nonlinear approximation optimized transform coding of images subject to structural constraints on transform basis functions.

Takahiro Nakayama *et al.* [88] proposed an image compression algorithm based on vector quantization technique. Adaptive resolution VQ (AR-VQ) method, which was composed of three key techniques, namely the edge detection, the resolution conversion, and the block alteration, can realize much superior compression performance than the JPEG and the JPEG-2000. In addition, they have proposed a systematic codebook design method of 4×4 and 2×2 pixel blocks for AR-VQ without using learning sequences. According to their method, the codebook applied to all kinds of images, exhibits equivalent compression performance to the specific codebooks created individually by conventional learning method using corresponding images.

Ian Berry *et al.* [89] have tested the use of common lossy and lossless compression algorithms on image file size and on the performance of the York University image analysis software by comparison of compressed Oxford images with their native, uncompressed bitmap images. This study shows that a significant 4-fold space savings (approximately) can be obtained with only a moderate effect on classification capability in biomedical images for e.g., various crystalline structures of protein images. They have investigated lossy JPEG compression as giving a potential for far greater savings that have to be offset against loss of image analysis accuracy. For their test data set (heavily biased in favor of interesting images com-

pared to our full database) a compression ratio of approximately 4:1 (95% image quality) yields a classification consistency of 75 % that may still be sufficient.

A.Kumar Kombaiya and V.Palanisamy [90] have presented an approach for an Enhanced Image Compression Method using Partial SPIHT Algorithm. Their method is based on the progressive image compression algorithm, SPIHT which is an extension of Shapiro's Embedded Zerotree Wavelet Algorithm. Their proposed Partial SPIHT Algorithm overcomes the difficulty of SPIHT that loses its efficiency in transmitting lower bit planes. In their work, they have included integer wavelet transformation and region of interest coding to Partial SPIHT and hence make it more superior to SPIHT and EZW algorithm.

K.Veeraswamy and S.Srinivas Kumar [91] have proposed an adaptive image compression algorithm based on the prediction of AC coefficients in DCT block during reconstruction of image. In the prediction phase, DC values of the nearest neighbour DCT blocks are utilized to predict the AC coefficients of centre block. Surrounding DC values of a DCT block are adaptively weighted for AC coefficients' prediction. Linear programming is used to calculate the weights with respect to the image content. Results show that their method is valid in terms of good PSNR and less blocking artifacts. In addition, an image watermarking algorithm is proposed using the DCT-AC coefficients obtained. The performance of their proposed watermarking scheme is measured in terms of PSNR and normalized cross correlation. Further, their algorithm is robust for various attacks including JPEG compression on watermarked image.

Liangbin Zhang and Lifeng Xi [92] have devised a hybrid image compression scheme using fractal-wavelet prediction where the causal similarity among blocks of different sub-bands in a wavelet decomposition of the image is exploited. The proposed coding scheme consists of predicting fractal code in one sub-band from the fractal code in lower resolution sub-band with the same orientation. By linearly adjusting the fractal code parameters in the lower resolution sub-band, an approximate forecast of the corresponding higher resolution sub-band with the

same orientation is achieved. Their experimental results show that the performance of the scheme is superior for both acceptable visual decoding image quality and an average of 20% reduction in encoding time and higher compression ratio, compared with standard Jacquin fractal coders.

Dzulkifli Mohamad *et al.* [35] have proposed a hybrid image compression technique using DCT and DWT by separating the foreground and background regions.

An efficient hybrid image vector quantization technique based on a classification in the DCT domain has been presented by Zhe-Ming Lu and Hui Pei [59]. Their algorithm combines two kinds of VQ, predictive VQ (PVQ) and DCTVQ, and also adopts a simple classifier which employs only three DCT coefficients in the 8×8 block. For each image block, the classifier switches to the PVQ coder if the block is relatively complex, and otherwise switches to the DCT-VQ coder. Experimental results have shown that their proposed algorithm can achieve higher PSNR values than ordinary VQ, PVQ, JPEG, and JPEG2000 at the same bit-rate.

A coding scheme that compresses the shape and texture of arbitrarily shaped visual objects has been presented by Martin*et al.* [94] . The presented compression scheme, Shape and Texture Set Partitioning in Hierarchical Trees (ST-SPIHT), is based on SPIHT. The ST-SPIHT utilizes the implementation of the Shape-Adaptive Discrete Wavelet Transform (SA-DWT) making use of in-place lifting, along with parallel coding of texture coefficients and shape mask pixels to attain a single embedded code that allows fine-grained rate-distortion scalability. The use of shape coding reduces the computational complexity incurred by exploiting the advantages of decomposition and spatial orientation trees employed for texture coding. Objective and subjective evaluations have been carried out to prove the superior rate-distortion performance of the ST-SPIHT scheme compared to MPEG-4 Visual Texture Coding for most bit rates.

Xiao Cheng He *et al.* [26] have considered multi-wavelets and characteristics of the Human Vision System (HVS) for image compression. To begin with, their proposed algorithm transforms a two dimensional image by selecting the BSA

(4/4)* filters. Subsequently, the HVS coefficients are applied into the sub-bands of the transformed image. The coefficients are then split into two parts: 1) the significance map and 2) the residue map. Subsequently, the significance map is encoded by making use of the proposed new modified SPIHT algorithm. The residue map is then encoded using another algorithm. Lastly, the bit stream is compressed by adopting the context-based adaptive arithmetic coding. The result proves that it is worth studying multi-wavelets and they have also compared their algorithm with other multiwavelet and JPEG2000 algorithms.

A scheme based on Self-Organizing Feature Map (SOFM) algorithm has been presented by Banu Diri and Songul Albayrak for color image compression [96]. The 1-dimensional SOFM has been utilized to map 256- color to 64−, 32− and 16− color. Subsequently, relative coding and entropy coding are performed such that there is no loss of information. The results obtained from experimentation have demonstrated the effectiveness of using SOFM for image compression.

Sharma *et al.* [97] have presented a global processing technique for training the Kohonen's network. For years now, a neural network scheme, Kohonen's Self-Organizing Map has been one of the most renowned algorithms for image compression, feature extraction and pattern recognition by association rules. The novel technique proposed for SOFM training is tested using JPEG images and a substantial decrease in the size of compressed images has been attained.

Xing-hui Zhang *et al.* [98] have made use of the shape parameters of generalized Gaussian distribution to classify the image into three different blocks. The shape parameters being very difficult to be measured by general methods, they have employed a novel associative memory neural network to determine the shape parameter. The results have demonstrated the effectiveness (accuracy) of the algorithm compared with other methods. Actually, the HVS is embedded into SPIHT algorithm so as to determine the different perceptual weights to different image blocks. The results of experimentation have illustrated that their algorithm yields significant PSNR and subjective visual quality of image after the decompression.

Tze-Yun Sung and Hsi-Chin Hsin [99] have presented a hybrid algorithm combining SPIHT and EBC (Embedded Block Coding) to encode low frequency and high frequency wavelet coefficients. The intermediary results obtained from coding of low frequency coefficients have been made use to facilitate the coding operation of high frequency coefficients. A considerably improved coding performance is attained by the hybrid SPIHT-EBC algorithm.

An extremely scalable hybrid image coding scheme (HS-HIC) has been proposed by Usama S. Mohammed [100]. It presents a hybrid coding scheme that combines a simple modification of the data in the wavelet domain and the exceptional performance of the SPIHT coding. The sub-band image data is modified based on the transformation of the high-frequency sub-band (details) in the wavelet domain. Apart from the image data in LL3, all other image data have linearly been modified based on the DFT components. The modification process outputs a new sub-band image data comprising nearly identical information as the original one but having a smaller frequency spectrum. Simulation results demonstrated that the proposed algorithm achieves better PSNR performance than that of the SPIHT test coder and some of famous image coding techniques with slightest possible addition to the computational complexity in the coding process.

A hybrid image coding scheme based on shape primitives, termed Shape Primitive Extraction and Coding (SPEC) has been presented by R.Ramya and K.Mala [101]. It is essential for a compression algorithm not only to achieve the high compression ratio, but also low complexity and high visual quality. The steps involved in the coding scheme are,

- Segmentation of the image blocks into picture and text/graphics blocks by thresholding the number of colors of each block, followed by the extraction of the shape primitives of text/graphics from picture blocks;
- Separation of small shape primitives of text/graphics from pictorial blocks using dynamic color palette that tracks recent text/graphics colors;

- Extraction of shape primitives from text/graphics blocks;
- Lossless coding of the foreground text and graphics pixels that combines shape-based and palette-based coding and
- Lossy Coding of the background pictorial pixels using Joint Photographic Expert Group (JPEG).

The efficiency and low complexity of the hybrid scheme have been demonstrated by means of the results of experimentation.

The adoption of an 8×8 DCT approach to perform DCT shrinkage, followed by modified SPIHT data organization and fidelity enhancement filter for reducing the memory needed to store a remote diagnosis and speedily transmit it has been performed by Yen-Yu Chen [102]. The unimportant DCT coefficients that correspond to the same spatial location in the high-frequency sub-bands are being reduced by a combined function proposed in association with the modified SPIHT, so as to lessen the redundancy. The quad-tree decomposition and a set of morphological filters have been utilized for artifact removal in the interim. Simulation results have illustrated that the image compression reduces the computational complexity to only a half of the wavelet based sub-band decomposition and also improves the reconstructed medical image quality both in terms of PSNR and perceptual results, close to JPEG2000 and the original SPIHT at the same bit-rate.

A color image compression algorithm making use of Kohonens self-organizing feature map has been presented by Kazuyuki Tanaka *et al* [103]. N number of neurons have been introduced for reducing a given full color image with 224 colors to an indexed color image with N colors. There are control parameters for the competitive learning between neurons in the SOFM algorithm. In the algorithm, a few of the control parameters included in a neighboring function (defined for neurons) are updated by considering the relationship among neighboring neurons. This is in contrast to Pei and Los algorithm [104], where all the control parameters are updated so as to decrease monotonically and exponentially with respect to each

iteration step. The color palette attained by the algorithm has been more robust as for control parameters than that by Pei and Los algorithm.

The images when compressed using self-organizing feature maps take longer time to converge. It is due to the fact that a given image may consist of numerous distinct gray levels with narrow differences with their neighborhood pixels. It has be determined that when the gray levels of the pixels in an image and their neighbors are mapped such that the difference in the gray levels of the neighbors with the pixel is minimum, the compression ratio as well as the convergence of the network can be improved. Thus, a high performance, in terms of compression ratio, has been achieved at the cost of a long processing time.

Durai *et al.* [106] have estimated a cumulative distribution function for the image that can be used to map the image pixels. On utilization of the mapped image pixels, the SOFM network yields a high compression ratio and it converges swiftly as well.

Kuo-Liang Chung [107] has presented a spatial as well as DCT based hybrid gray image representation approach. In the first phase, the decomposed bin tree of the input gray image has been represented using an S-tree spatial data structure (SDS), according to the bin tree decomposition principle under the specified error. Homogeneous leaves and the non-homogeneous leaves are the two types into which the constructed S-tree (SDS) leaves have been partitioned. One rectangular or square homogeneous sub-image with smooth, or, in other words, low frequency content, has been represented using the homogenous leaf; whereas, one non-homogeneous sub image with non-smooth, or, in other words, high frequency content, has been represented using a non-homogeneous leaf. The memory requirement has been reduced in the second phase by encoding, each non-homogeneous leaf by the DCT-based coding scheme.

Pandian *et al.* [108] have presented a transform domain based technique for color image compression. Vector quantization (VQ) technique has been used for compression of images and Kohonens SOFM has been used during the design of

the codebook in VQ. Special features of SOFM for generic codebook generation that permit to create the codebook only once have been exploited by their work.

K.Veeraswamy *et al.* have presented a simple approach for generation of an optimal quantization table based on HVS model [15]. This quantization table is used to quantize the Hadamard transform coefficients. This table can provide superior image compression over standard quantization tables available in the literature.

Lala Krikor*et al.*[109] have presented a technique for image encryption, which has considered certain chosen higher frequencies of DCT coefficients as the characteristic values, encrypt them according to a pseudo random bit sequence and shuffles the resulting encrypted blocks. The computational requirements of huge volumes of images have been decreased by the recent selective encryption approach.

Khalil *et al.* [110] have described and implemented a RUN-Length coder that has been made simple and more effective. Their proposed algorithm has worked on quantized coefficients of the DCT where several concurrent tokens exist. The new approach has been proved to attain competitive performance by experimental results.

Meng Meng *et al.* [111] have used DCT, VQ coding and a new proposed method that combines DCT and wavelet transform in the implementation of their proposed color image compression algorithm. This algorithm achieves high compression ratio and high efficiency.

There are many research outputs that talk about low bit rate compression, but visual quality of the output is poor. On the other hand, many contributions in literature deal with high quality decompressed image but yield lower compression ratios. Thus, there is a need of further research having high compression ratio and better image quality.

1.5 Problem Statement

The problem taken for the doctoral research work is to develop efficient hybrid image compression schemes *that yield higher quality in terms of subjective and objective evaluations at lower bit-rates.*

1.6 Chapter-wise Organization of Thesis

The chapter wise organization of the thesis is presented here.

Chapter 1 Introduction

Chapter 2 Some Basic Image Compression Schemes

Chapter 3 Development of Hybrid Image Compression Schemes using SPIHT and SOFM based Vector Quantization

Chapter 4 Development of Hybrid Image Compression Scheme using DCT and Fractal Image Compression

Chapter 5 Lossless Image Compression Scheme based on CALIC and Spatial Prediction Sctructures

Chapter 6 Conclusion

1.7 Conclusion

This chapter provides a brief introduction on image compression. Literature survey of recent work has also been presented. The fidelity criteria for evaluating the quality of decoded images are discussed. The distortion metrics used in image compression are also described. The background and scope of the work as well as the motivation and the objective of the doctoral research problem are systematically discussed. A brief chapter-wise organisation of the dissertation has also been presented.

Chapter 2

Some Basic Image Compression Schemes

Preview

Image compression research aims at reducing the number of bits needed to represent an image. In lossless compression schemes, the reconstructed image after reconstruction is numerically identical(amplitude-wise) to the original image. However, lossless compression can achieve only a modest amount of compression ratio. On the other hand, lossy schemes are capable of yielding much higher compression.

Some basic and important lossy and lossless compression schemes, available in literature, are discussed in this chapter.

2.1 Introduction

For a universal algorithm to compress images, a sequence of image pixels extracted from an image in the raster scan order is simply encoded. But, for a universal algorithm such a sequence is hard to compress. Universal algorithms are usually designed for alphabet of sizes not exceeding 2^8 and do not exploit directly the image data features [77]. As images are 2-dimensional data, intensities of neighboring pixels are highly correlated, and the images contain noise added to the image during the acquisition process. The latter feature makes dictionary compression algorithms perform worse than statistical ones for image data. Modern gray-scale image compression algorithms employ techniques used in universal statistical compression algorithms. However, prior to statistical modelling and entropy coding the image data is transformed to make it easier to compress.

To make the image data easily compressible, we use 2-dimensional image transforms, such as DCT or wavelet transform [3]. In transform algorithms, instead of pixel intensities, a matrix of transform coefficients is encoded. The transform is applied to the whole image or to an image split into fragments. Transforms can be used for both lossless and lossy compressions. Transform algorithms are more popular in lossy compression. Apart from lossy and lossless compressing and

decompressing of whole images, transform algorithms deliver many interesting features such as progressive transmission, region of interest coding, etc [17, 18]. The usages of algorithms are dependent mostly on information content of images and types of application.

Lossless compression algorithms are often predictive in nature [2, 20]. In a predictive algorithm, the predictor function is used to guess the pixel intensities and the prediction errors are calculated. The prediction errors are differences between actual and predicted pixel intensities. To calculate the predictor for a specific pixel usually intensities of a small number of already processed pixels neighbouring it is used. Next, the sequence of prediction errors, called residium, is encoded. Prediction error distribution is close to Laplacian, that is, symmetrically exponential [2, 24]. Therefore, entropy of prediction errors is significantly smaller than that of pixel values. That is why, it is easier to compress residium. In respect to the lossless compression, better results in terms of computational speed are obtained by predictive algorithms. This chapter reviews some important lossless and lossy image compression schemes employed in this thesis.

2.2 Context Adaptive Lossless Image Coding (CALIC) Algorithm

The CALIC scheme came into being in response to a call for proposal for a new lossless image compression scheme in 1994. It uses both context and prediction of the pixel values.

In an image, a given pixel generally has a value close to one of its neighbours. Which neighbour has the closest value depends on the local structure of the image. Depending on whether there is a horizontal or vertical edge in the neighbourhood of the pixel being encoded, the pixel above, or the pixel to the left, or some weighted average of neighbouring pixels may give the best prediction. How close

the prediction is to the pixel being encoded depends on the surrounding texture. In a region of the image with a great deal of variability, the prediction is likely to be farther from the pixel being encoded than in the regions with less variability. In order to take into account all these factors, the algorithm has to make a determination of the environment of the pixel to be encoded. The only information that can be used to make this determination has to be available to both encoder and decoder.

CALIC obtains higher lossless compression for the continuous-tone images than other techniques reported in the literature [21]. This high coding efficiency is accomplished with relatively low time and space complexities. CALIC puts heavy emphasis on image data modelling. A unique feature of CALIC is the use of a large number of modeling contexts to condition a non-linear predictor and make it adaptive to varying source statistics. The nonlinear predictor adapts via an error feedback mechanism. In this adaptation process, CALIC only estimates the expectation of prediction errors conditioned on a large number of contexts rather than estimating a large number of conditional error probabilities. The former estimation technique can afford a large number of modelling contexts without suffering from the sparse context problem. The low time and space complexities of CALIC are attributed to efficient techniques for forming and quantizing modeling contexts.

CALIC employs a two-step (prediction/residual) approach. In the prediction step, CALIC employs a simple new gradient based non-linear prediction scheme called gradient-adjusted predictor (GAP), which adjusts prediction coefficients based on estimates of local gradients. Predictions are then made context-sensitive and adaptive by modelling of prediction errors and feedback of the expected error conditioned on properly chosen modelling contexts. The modelling context is a combination of quantized local gradient and texture pattern; the two features that are indicative of the error behaviour. The net effect is a non-linear, context-based,

adaptive prediction scheme that can correct itself by learning from its own past mistakes under different contexts. The context-based error modelling is done at a low model cost. By estimating expected prediction errors rather than error probabilities in different modelling contexts, CALIC can afford a large number of modeling contexts without suffering from either context dilution problem or from excessive memory use. This is a key feature of CALIC that distinguishes it from existing methods.

CALIC encodes and decodes images in raster scan order with a single pass through the image. The coding process uses prediction templates that involve only the previous two scan lines of coded pixels. Consequently, the encoding and decoding algorithms require a simple double buffer that holds two rows of pixels that immediately precede the current pixel, hence facilitating sequential build-up of the image.

CALIC operates in two modes: binary and continuous tone modes. The system selects one of the two modes during the coding process, depending on the context of the current pixel. The binary mode will be triggered if all of the context pixels have less than two different gray-scale levels (not necessary to be only 0 and 1). So some flat gray-scale portions in a continuous-tone image can be coded in binary mode. To utilize the context information, an entropy codec (arithmetic coder) drives 32 different context models in binary mode. The continuous-tone mode basically has following four major components.

1. Gradient Adjusted Prediction

2. Context selection and quantization

3. Context modelling of prediction errors

4. Entropy coding of prediction errors.

CALIC uses a previous two-line buffer to construct the GAP and context modelling. It is understood that the neighbourhood pixels are known, which will

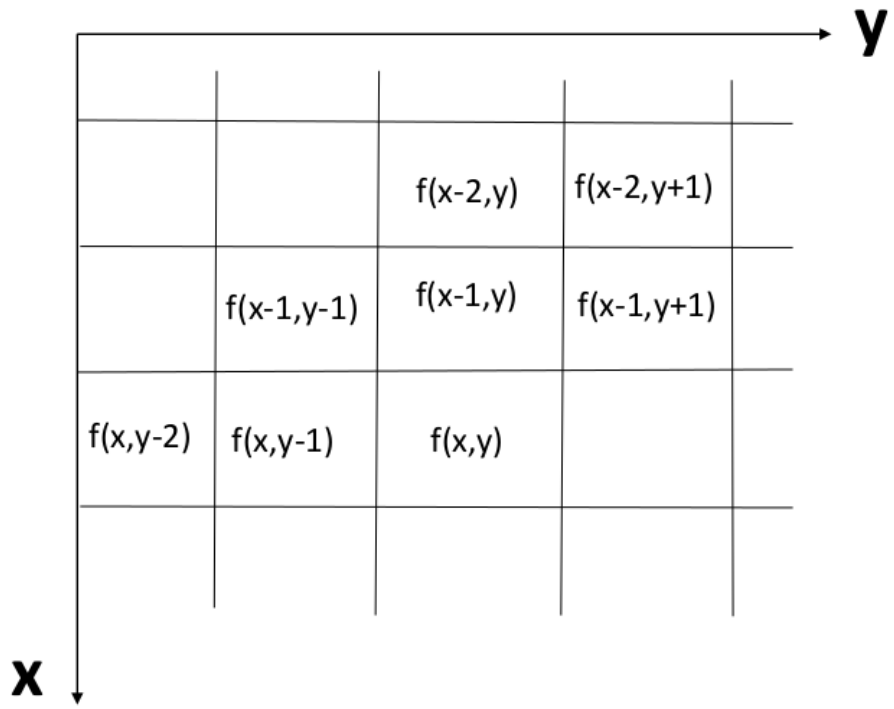


Figure 2.1: Neighborhood of pixels in CALIC

be used while encoding. Suppose we are going to encode the pixel $f(x, y)$ [Refer to Figure 2.1]. The whole algorithm will utilize totally 7 neighbourhood pixel set $f(x, y-1)$, $f(x, y-2)$, $f(x-1, y-1)$, $f(x-1, y)$, $f(x-1, y+1)$, $f(x-2, y)$ and $f(x-2, y+1)$.

The GAP predictor employed by CALIC is a simple, adaptive, nonlinear one that can adapt itself to the intensity gradients near the predicted pixel. Hence, it is more robust than the traditional DPCM-like linear predictors, particularly in areas of strong edges. The GAP differs from existing linear predictors in that it weights the neighbouring pixels according to the estimated gradients of the image. It adapts itself to the gradients of horizontal and vertical edges. The GAP tries to detect how rapidly the edge changes around the pixel, $f(x, y)$, and then by classifying the tendency of edge changing into sharp, normal and weak edge, it assigns different weights for the various neighbourhood pixels for a linear prediction of

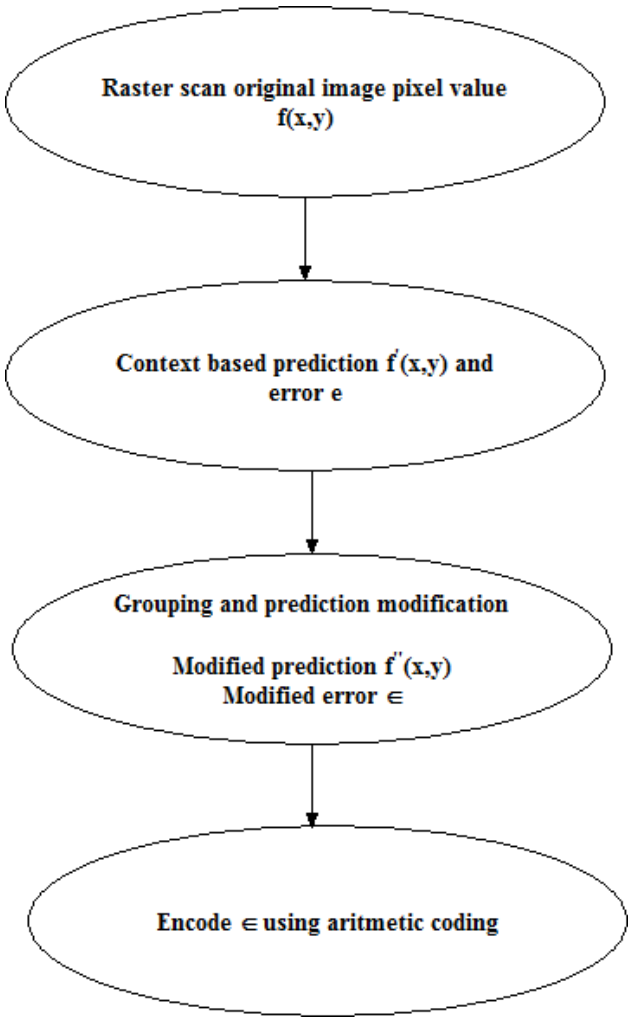


Figure 2.2: CALIC flowchart

pixel, $f(x, y)$.

The process of CALIC algorithm is summarised in Figure 2.2. The value of pixel $f(x, y)$ is predicted as $f'(x, y)$ using gradient adjusted prediction and the seven pixels in the neighborhood of pixel $f(x, y)$ as shown in Figure 2.1. The predictor coefficients and thresholds are empirically chosen. A major criterion in choosing these coefficients is ease of computation. For instance, most coefficients are powers of 2 so that multiplications/divisions can be performed by shifting.

The predicted value $f'(x, y)$ is termed as context based prediction. The error in prediction $e = f(x, y) - f'(x, y)$ is computed. Then the predicted pixel values are grouped and prediction modification is performed. The modified prediction, $f''(x, y)$, and the modified prediction error, ϵ are generated. For the scheme to be lossless, the prediction error needs to be specified along with the modified prediction. Therefore, the modified prediction error is entropy coded using arithmetic coding. The detailed description of the complete process may be found in literature [21].

2.2.1 Coding Performance of CALIC

The results tabulated in Table 2.1 indicate the compression performance of CALIC algorithm for gray-scale images of size 512×512 expressed in bpp for various test images. We observe from Table 2.1 that the CALIC algorithm yields

Table 2.1: Compression Performance of CALIC in terms of bpp

Image	bpp	Compression Ratio (ρ)
Lena	3.1509	2.5
Barbara	3.4804	2.08
Babbon	4.2813	1.87
Boat	3.5804	2.23
Goldhill	3.4663	2.31
Peppers	3.3524	2.38

a compression ratio of 1.87 (Baboon) to 2.5 (Lena) without any loss of information.

2.3 Discrete Cosine Transform based Compression

Disintegrating the images into segments is the fundamental operating principle of DCT [39]. A better signal approximation with fewer transform coefficients are provided by DCT which are real valued unlike those obtained in a Discrete

Fourier Transform [130]. In several practical image compression systems the invertible linear transform called 2-dimensional DCT is extensively used because of its compression performance and computational efficiency [81]. Data (image pixels) is converted into sets of frequencies, by DCT. The frequency sets are arranged in ascending order of frequency and descending order of significance as far as image quality is concerned. On the basis of tolerable resolution loss, the least meaningful frequencies can be discarded.

2.3.1 Global DCT versus Block DCT (BDCT)

For the purpose of image compression, DCT can be applied to the complete image (global DCT) or to an $n \times n$ block of image(BDCT). Applying DCT to entire image produces better compression but involves extremely large number of arithmetic operations thus adding to its computational complexity. Therefore, the process of compression slows down. On the other hand, applying DCT to small data units is faster but reduces the compression ratio. Moreover, in the continuous tone images, the correlations between pixels are in short range. Statistical analysis of natural images has revealed that there is little correlation between pixels more than 8 positions apart and, in fact, most of the correlations are among pixels that are within 4 positions away. The 8×8 block size is an excellent choice from both the bit-rate and the correlation-exploitation points of consideration [2, 3, 10].

2.3.2 Block DCT based Compression

Figures 2.3 and 2.4 depict the compression and decompression process based on BDCT. The image is divided into non-overlapping blocks of size 8×8 or 16×16. In standard JPEG encoding, it is divided into 8×8 blocks in the raster scan order form left to right and top to bottom. Each pixel is level shifted into signed integer by subtracting 128 from each pixel. The 8×8 block $g(x, y)$ is transformed from

the spatial domain to the frequency domain by means of the DCT transform given by

$$G(u,v) = \frac{1}{4}C(u)C(v)\sum_{x=o}^{7}\sum_{y=0}^{7} g(x,y)cos(\frac{\pi(2x+1)u}{16})cos(\frac{\pi(2y+1)v}{16}) \qquad (2.1)$$

for $u = 0, 1, ...7$ and $v = 0, 1, ...7$

where,

$$C(k) = \begin{cases} \frac{1}{\sqrt{2}}, & k = 0 \\ 1, & otherwise \end{cases}$$

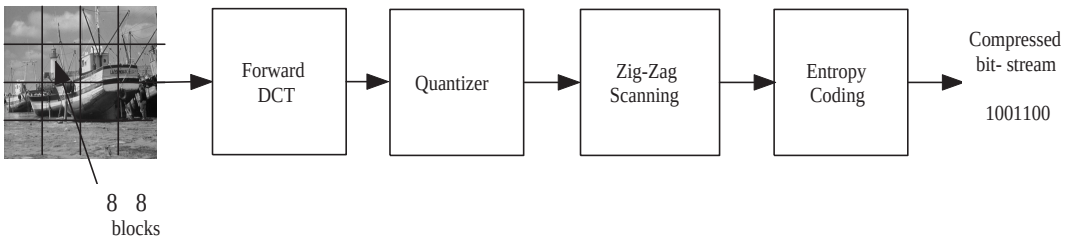


Figure 2.3: Encoder for BDCT based Compression

From the lowest (upper left corner) to the highest (lower right corner) frequencies, 64 DCT coefficients are computed for each block [2]. All DCT coefficients are encoded by using a constant number of bits. But, the importance (the ratio between an upper left corner coefficient and the one in the right bottom corner) is not the same for all the coefficients in a DCT [105]. Therefore, all coefficients

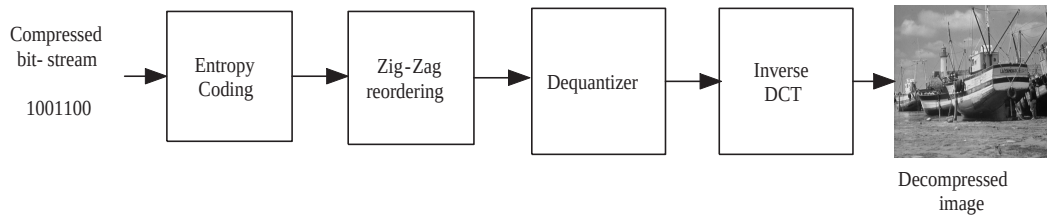


Figure 2.4: Decoder for BDCT based Compression

should not be assigned same number of bits for representation. This is achieved

by a process called quantization, which is given by:

$$G_q(u,v) = round\left(\frac{G(u,v)}{Q(u,v)}\right) \tag{2.2}$$

Here $G(u,v)$ are the DCT coefficients of the transformed image and $Q(u,v)$ is the quantization step size parameter. JPEG standard defines one quantization matrix for monochrome and for color images, two quantization matrices one each, for luminance and chrominance planes [22]. These matrices determine the visibility threshold for 2 dimensional basis functions. The matrices are best suited for natural images with 8 bit precision. The quality of the reconstructed image can be controlled by scaling these matrices. A variable called scale factor is used to control the quantization provided to the DCT coefficients and the corresponding quality of image. The scale factor varies from 1 to 50. Higher the value of scale factor more is the quantization and more the degradation in image quality.
The entire quantized coefficients are rearranged in a zigzag manner. Most of the high frequency coefficients (lower right corner) become zeros after quantization. A zig-zag scan of the matrix yielding long strings of zeros is used to exploit the number of zeros. The entropy coding used is Huffman coding.

2.3.3 Coding Performance of BDCT for Gray-scale Images

The compression performance of BDCT algorithm is tested for test images, each of 512×512 size. The scale factor for these images is kept constant at 5 and bpp is calculated for all the images. The rate distortion performance is expressed in terms of PSNR, MSSIM, VIFP and UQI. Simulation results are tabulated in Table 2.2. It is observed from Table 2.2 that the amount of compression obtained for each image is different for the same scale factor. The test images: Lena and Peppers exhibit higher compression corresponding to lower bit-rates and better quality as indicated by PSNR, VIFP, MSSIM and UQI values, whereas Barbara and Baboon images have higher bit-rates and even poorer rate-distortion per-

Table 2.2: Simulation Results of BDCT compression algorithm on various image

Image	bpp	PSNR(dB)	VIFP	MSSIM	UQI
Lena	0.4286	36.97	0.8382	0.8743	0.9998
Barbara	0.7215	25.08	0.7714	0.8113	0.9492
Baboon	0.9223	27.89	0.8092	0.7512	0.9959
Peppers	0.4399	42.15	0.8211	0.8992	0.9947

formance. Lena and Peppers give good quality because these two images have medium complexity regions. Therefore these images can be compressed further. Barbara and Baboon images have medium and high complexity regions hence the compression performance is poor.

2.4 Set Partitioning in Hierarchial Trees(SPIHT)

SPIHT is computationally fast and among the best known image compression algorithms today. The SPIHT [57] encoder works by taking advantage of the inert relationships among the wavelet coefficients across the different scales at the same spatial location in the wavelet sub-bands. SPIHT coding generally involves the coding of

- position of significant wavelet coefficients
- position of zerotrees in the wavelet subbands.

The SPIHT coder has following characteristics namely

- The greater part of an image's energy is concentrated in the low-frequency components and a decrease in variance is detected as we move from the highest to the lowest levels of the sub band pyramid.
- It has been understood that there is a spatial self-similarity amongst the sub-bands, and probably the coefficients are to be better magnitude-ordered on moving downward in the pyramid along the same spatial orientation.

A spatial orientation tree is being made use of to vividly describe the spatial relationship on the hierarchical pyramid. Figure 2.5 depicts the manner in which the spatial orientation tree is defined in a pyramid constructed with recursive four sub-band splitting. Every node in the tree represents a pixel in the image containing its corresponding pixel coordinate. The direct descendants (offspring) of a node represent the pixels of the same spatial orientation in the next finer level of the pyramid. The tree is constructed such that every node either has no offspring (the leaves) or four offsprings, which at all times form a group of 2×2 adjacent pixels. In Figure 2.5, the arrows are directed from the parent node to its four offsprings. The pixels in the highest level of the pyramid are the tree roots and are also grouped in 2×2 adjacent pixels. However, the offspring branching rule is diverse, and in each group, one of them (indicated by the black dot in LL3 band of Figure 2.5) has no descendants.

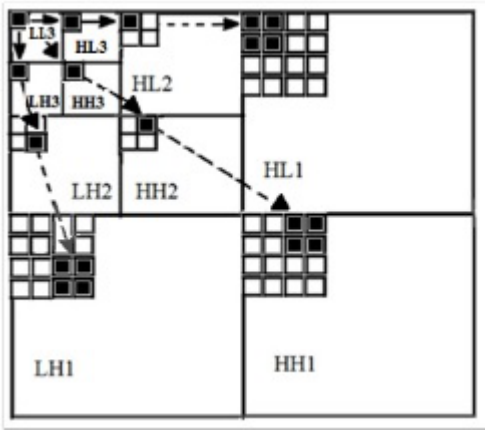


Figure 2.5: Spatial orientation tree defined in a pyramid constructed with recursive four sub-band splitting

The sets of coordinates that are being made use of to represent the coding method are as follows:

$O(i, j)$ is the set of offspring (direct descendants) of a tree node defined by pixel location (i, j).

$D(i, j)$is the set of descendants of node defined by pixel location (i, j).

$L(i,j)$ is the set defined by
$L(i,j) = D(i,j) - O(i,j)$.

Except for the highest and lowest pyramid levels, the set partitioning trees have,
$O(i,j) = [(2i,2j),(2i,2j+1),(2i+1,2j),(2i+1,2j+1)]$

The set when found significant can be split using the following rules,

- The initial partition is created with the sets (i,j) and $D(i,j)$, for all $(i,j)\varepsilon H$.
- If $D(i,j)$ is significant, then it is partitioned into $L(i,j)$ plus the four single-element sets with $(k,l)\varepsilon O(i,j)$.
- If $L(i,j)$ is significant, then it is partitioned into the four sets $D(k,l)$ with $(k,l)\varepsilon O(i,j)$.

The significant values of the wavelet coefficients contained in the spatial orientation tree are stored in three ordered lists namely,

- List of Insignificant Sets (LIS): Stores the set of wavelet coefficients that have magnitude smaller than a threshold and found in the tree structures that are insignificant. The sets prohibit the coefficients corresponding to the tree or all subtree roots, and have at least four elements. The entries in LIS are sets of the type $D(i,j)$ (type A) or type$L(i,j)$ (type B).
- List of Insignificant Pixels (LIP): Stores the individual coefficients that have magnitude smaller than the threshold.
- List of Significant Pixels (LSP): Stores the pixels that have magnitude larger than the threshold (are significant).

During the sorting pass, the pixels in the LIP that were insignificant in the previous pass are tested, and those that are found significant are moved to the LSP. Then, the sets are sequentially examined along the LIS order, and when a set is determined to be significant, it is eliminated from the list and partitioned. The new sets with more than one element are added back to LIS, while the one element sets are added to the end of LIP or LSP, according to them being significant.

Algorithm

1. Initialization: output $n = \lfloor log_2(max_{(i,j)}|c_{i,j}|)\rfloor$
   Set the LSP as an empty list, and add the coordinates $(i,j)\varepsilon H$ to the LIP, and only those with descendants also to the LIS, as type A entries.

2. Sorting Pass:
   2.1) for each entry (i,j) in LIP do
   2.1.1) output$S_n(i,j)$
   2.1.2) if $S_n(i,j) = 1$ then move (i,j) to LSP and output the sign of $c_{i,j}$

3. Refinement Pass: For each entry (i,j)in the LSP, expect those included in the last sorting pass (i.e., with same n), output the n^{th} most significant bit of $c_{i,j}$

4. Quantization-Step Update: Decrement n by 1 and go to Step 2.

The following are some of the advantages of SPIHT encoding:

- A variable bit rate and rate distortion control with provisions for progressive transmission.

- An intensive progressive capability: we can interrupt the decoding (or coding) at any time and a result of maximum possible detail can be reconstructed with one-bit precision.

- Very compact output bit stream with large bit variability; no supplementary entropy coding or scrambling has to be applied.

2.4.1 Coding Performance of SPIHT for Gray-scale Images

Selection of number of levels of decomposition

The SPIHT algorithm is tested for the gray-scale test image, Lena. The bit-rate is kept constant at 0.1 and the level of decomposition is varied from level 1 to maximum level possible. The maximum possible level depends on the size of images chosen. The images used are of size 512×512. So the maximum level of decomposition is 9. Variation of PSNR, MSSIM, VIFP and UQI are recorded in Table 2.3.

It is evident from the Table 2.3 and Figure 2.6 (a) through Figure 2.6 (d) that:

- The PSNR values vary significantly from 24 dB to 33.75 dB as decomposition levels employed change from 1 to 6 and reaches a maximum value of 33.98 dB for the 9^{th} level of decomposition. Further, increase in decomposition levels after the 6^th level does not show significant improvement in PSNR.
- At the same time, it is observed that VIFP varies from 0.0389 to 0.6316 increasing as level increases.
- MSSIM varies from 0.0468 to 0.8019, but almost attains saturation beyond level 6.
- Variation of UQI from 0.0903 to 0.9990 becomes nearly constant beyond level 6.

Though PSNR does not show significant increase in its value, visual quality of image improves with increase in levels, as indicated by the other metrics: VIFP, MSSIM and UQI. Hence, maximum level of decompositions selected for the simulation is taken as 9 to have better visual quality for a decompressed image.

Table 2.3: Simulation Results of SPIHT algorithm on Lena image at various decomposition levels.

Decomposition Level	PSNR(dB)	VIFP	MSSIM	UQI
1	24.41	0.0389	0.0468	0.0902
2	25.62	0.1426	0.2087	0.3694
3	26.58	0.1605	0.3042	0.5300
4	29.46	0.2561	0.6189	0.9818
5	33.26	0.5654	0.7782	0.9981
6	33.75	0.6065	0.7973	0.9987
7	33.93	0.6277	0.8011	0.9990
8	33.97	0.6316	0.8019	0.9990
9	33.98	0.6321	0.8021	0.9990

Table 2.4: Simulation Results of SPIHT algorithm on Lena image at various bit rates

bit-rate(bpp)	PSNR(dB)	VIFP	MSSIM	UQI
0.0156	23.13	0.1631	0.6467	0.2669
0.0313	25.17	0.2178	0.7014	0.3488
0.0625	27.35	0.2886	0.7610	0.4431
0.1250	30.19	0.3831	0.8245	0.5379
0.25	33.30	0.4865	0.8789	0.6256
0.5	36.56	0.5947	0.9181	0.7091

Table 2.5: Simulation Results of SPIHT algorithm on Barbara image at various bit rates

bit-rate	PSNR(dB)	VIFP	MSSIM	UQI
0.0156	20.67	0.1277	0.4915	0.2529
0.0313	21.88	0.1744	0.5438	0.3284
0.0625	22.92	0.2182	0.6034	0.4062
0.1250	24.20	0.2776	0.6741	0.4914
0.25	26.87	0.3572	0.7806	0.6148
0.5	30.63	0.4819	0.8793	0.7343

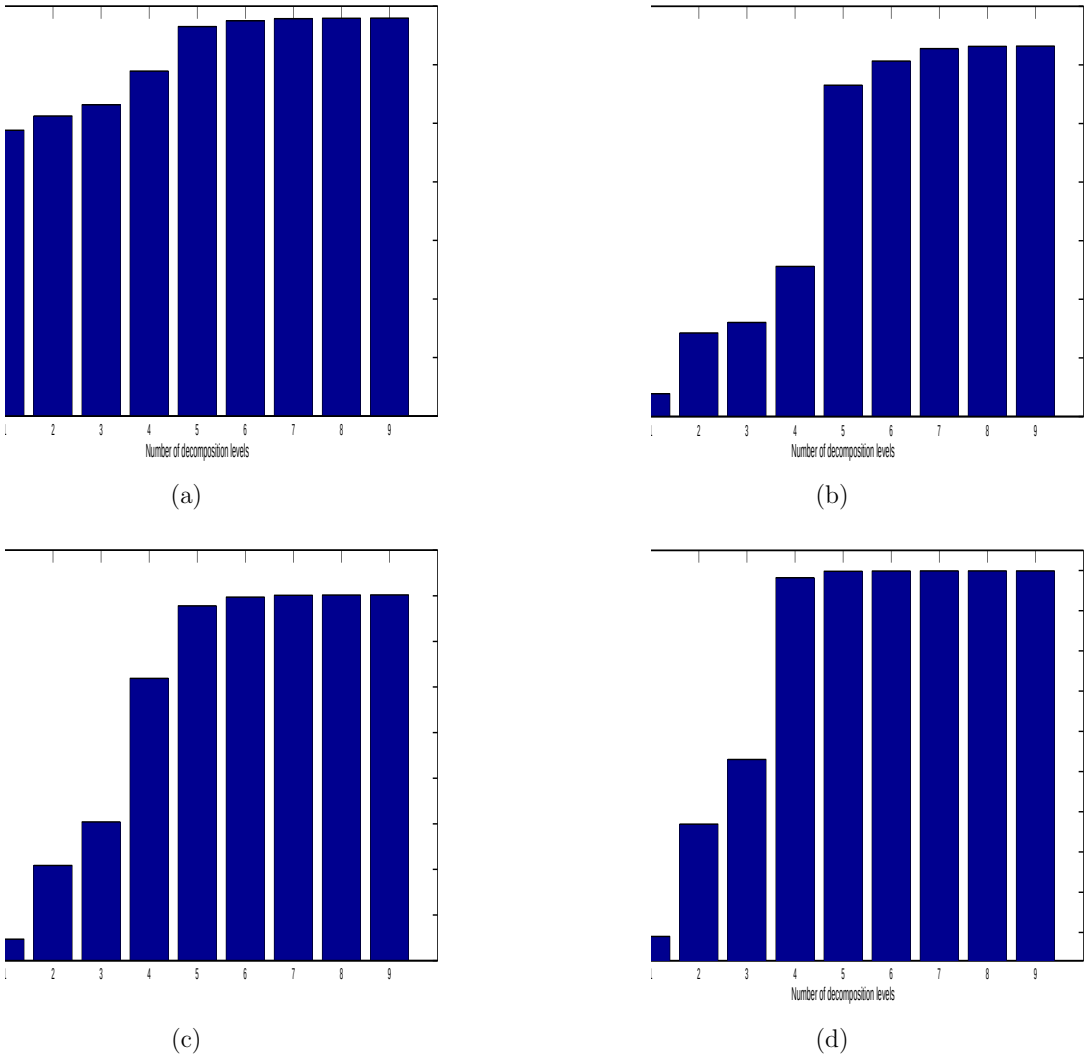


Figure 2.6: Rate distortion performance of SPIHT for different levels of decomposition in terms of(a)PSNR(dB),(b) VIFP,(c) MSSIM,(d) UQI

(a) (b)

(c) (d)

(e)

Figure 2.7: Visual quality of decompressed Lena image at different bit-rates (a)0.0313 bpp (b) 0.0625 bpp (c) 0.125 bpp (d) 0.25 bpp (e) 0.5 bpp

Table 2.6: Simulation Results of SPIHT algorithm on Peppers image at various bit rates

bit-rate	PSNR(dB)	VIFP	MSSIM	UQI
0.0156	22.41	0.1533	0.6293	0.2768
0.0313	24.44	0.2040	0.6789	0.3453
0.0625	26.84	0.2792	0.7387	0.4335
0.1250	29.79	0.3560	0.7929	0.5093
0.25	32.80	0.4546	0.8394	0.5846
0.5	35.35	0.5468	0.8753	0.6152

Table 2.7: Simulation Results of SPIHT algorithm on Baboon image at various bit rates

bit-rate	PSNR(dB)	VIFP	MSSIM	UQI
0.0156	19.49	0.0571	0.2881	0.1243
0.0313	19.88	0.0776	0.3258	0.1831
0.0625	20.45	0.0918	0.3832	0.2565
0.1250	21.39	0.1380	0.4675	0.3703
0.25	22.78	0.1865	0.6011	0.5207
0.5	25.06	0.2387	0.7207	0.6626

The tables: Table 2.4 through Table 2.7 give the coding performance for varoius test images as the bit-rate is varied from 0.0156 bpp to 0.5 bpp. These tables indicate that at lower bit-rates, PSNR values are in acceptable range for all images but other metrics numerically indicate that the visual quality is very poor. It is also observed that at lower bit rates below 0.25 bpp, the rate distortion performance of SPIHT is poor in terms of all metrics. At bit rates higher than 0.5 bpp, the performance of SPIHT is good in terms of all parameters. The test image Lena gives the best performance in terms of all the distortion metrics. The visual results of the coding performance of the test image Lena are given in Figure 2.7.

2.5 Vector Quantization (VQ)

Vector quantization [8] is a lossy compression scheme. Vector quantizers are preferred in image compression for the following two reasons.

- Most of the standard schemes need transformations from space domain to frequency domain while vector quantization operates in space domain itself, and hence its computational complexity is very less compared to transform based compression algorithms.

- Most compression schemes are symmetric for coding and decoding, whereas vector quantizers are non symmetric. Decoding involves only a lookup of the vector quantizer indices, which simplifies the decoder.

The VQ scheme is used for creating image descriptors. In VQ, the image to be compressed is partitioned into no overlapping blocks or vectors. A vector quantizer maps each input vector into a finite set of code words (codebook) using nearest neighbour rule. Once the closest codeword is found, the index of that codeword is sent to the decoder. When the decoder receives the index of the codeword, it replaces the index with the associated codeword and reconstructs the image using look-up table. Figure 2.8 illustrates this process.

A codebook that best represents the set of input vectors is difficult to design. It requires an exhaustive search for the best possible codeword in the space, and the search increases exponentially as the number of code words increases. Therefore, suboptimal codebook design schemes are resorted to and the first one that comes to mind is Linde-Buzo-Gray (LBG) algorithm [46]. It is the most common and the oldest method for vector quantization code book generation. LBG is a greedy algorithm. Therefore, its performance is sensitive to initialization. The number of code words, N or the size of the codebook is determined first. Then, N code words are selected at random, as the initial codebook. The initial code words can

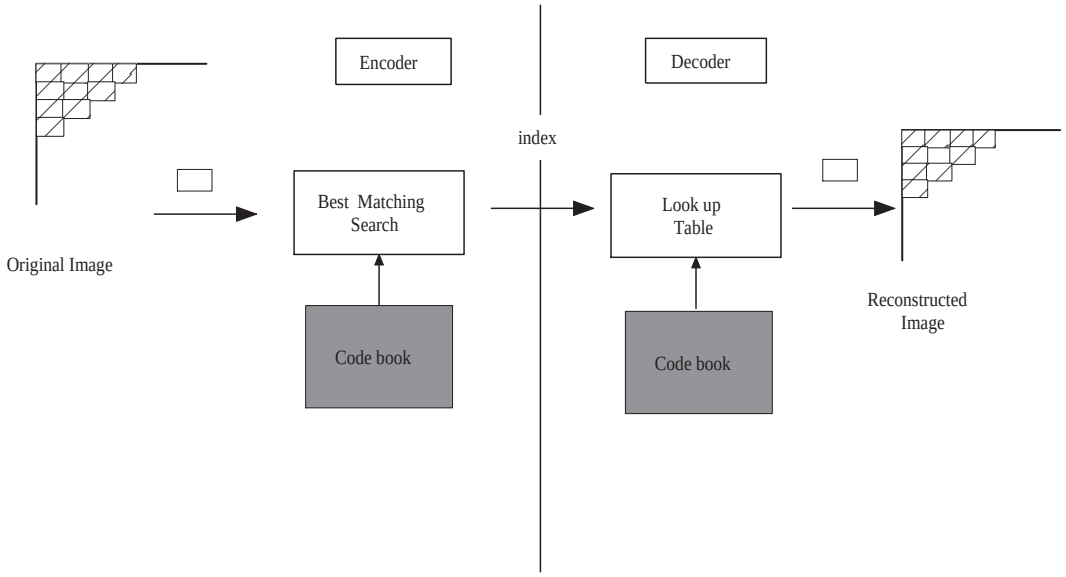


Figure 2.8: Encoder and Decoder in Vector Quantizer

randomly be chosen from the set of input vectors. Using the Euclidean distance measure, vectors around each codeword are clustered. This is done by taking each input vector and finding the Euclidean distance between it and each codeword. The input vector belongs to the cluster of the codeword that yields the minimum distance. Hence a new set of code words is computed. This is done by obtaining the average of each cluster. The component of each vector is added and divided by the number of vectors in the cluster. The process is repeated until either the code words do not change or the change in the code words become smaller than a predefined limit. This algorithm is by far the most popular one due to its simplicity, but it is very slow because input vector is compared with all the code words in the codebook for each iteration. It converges to the local minima closest to the initial point.

There are many other methods of designing the codebook, such as Pair wise Nearest Neighbor (PNN) [150], Simulated Annealing(SA) [48], Maximum Descent (MD) [49, 50] and Self Organising Feature Map SOFM [50, 52, 124].

The use of neural networks for the code book design problems has been investi-

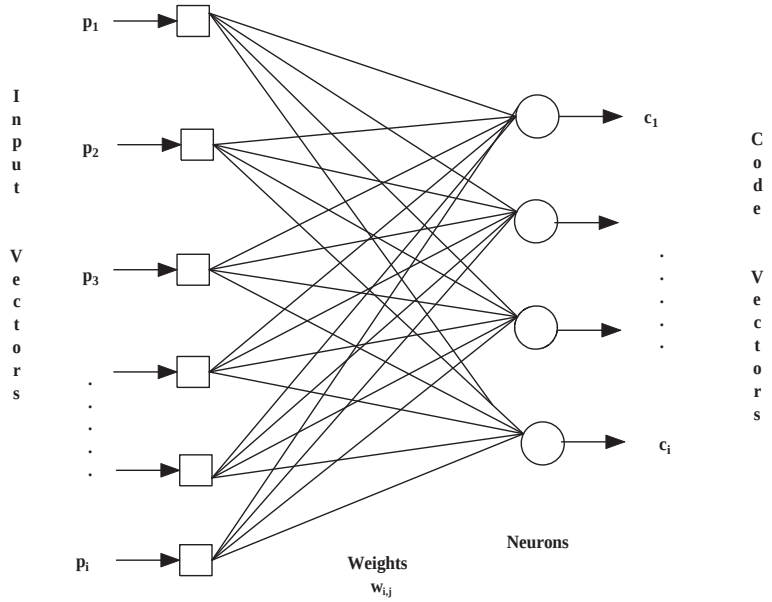


Figure 2.9: Neural network in vector quantization

gated [58]. Neural networks are attractive because of their massive parallelism, learning ability, adaptivity and fault tolerance. A self organising feature map is a neural network clustering technique having several desirable features. Consequently, it has attracted the attention of researchers in the field of vector quantization.

While designing vector quantizer, a designer faces two design issues: the dimensions of vector and the size of code book. A large vector enables the quantizer to exploit the statistical redundancy existing in the data to a greater degree. But this may reduce the reproduction error unless the codebook size is also large. The larger the code book size, the finer is the representation of input space. Since neural networks are capable of learning from input information and optimizing themselves to obtain the appropriate environment for a wide range of tasks, a family of learning algorithms has been developed for vector quantization. The input vector is constructed from a K-dimensional space. M neurons are designed to compute the vector quantization code-book in which each neuron relates to

one code-word via its coupling weights. The coupling weight, $w_{i,j}$ associated with the i^{th} neuron is eventually trained to represent the code-word in the code-book. Figure 2.9illustrates the process for one dimensional space. As the neural network is being trained, all the coupling weights will be optimized to represent the best possible partition of all the input vectors.

In order to train a network, a group of image samples known to both encoder and decoder is often designated as the training set, and the first M input vectors of the training data set are normally used to initialize all the neurons. With this general structure, various learning algorithms have been designed and developed such as Kohonen's self organizing feature mapping [127], competitive learning [19], frequency sensitive competitive learning [16, 56], fuzzy competitive learning [55], distortion equalized fuzzy competitive learning VQ [54] and predictive VQ (PVQ) neural networks [112]. Discussion of all these learning algorithms is beyond the scope of this thesis and only SOFM is discussed, in detail,in the subsequent sub-section.

2.5.1 Code Book Generation by SOFM

SOFM is realized by a two-layer network, as shown in Figure 2.10. The first layer is the input layer or fan-out layer with neurons and the second layer is the output or competitive layer. The two layers are completely connected. An input vector, when applied to the input layer, is distributed to each of the output nodes in the competitive layer. Each node in this layer is connected to all nodes in the input layer; hence, it has a weight vector prototype attached to it.

SOFM begins with a random initialization of the weight vector. Let $p \, \varepsilon \, \Re^2$ be the input to the network and let t denote the current iteration number. The neurons in second layer now compete among themselves to determine the neuron whose weight vector matches best with the input p. It finds $w_{i,t-1}$ that best matches p in the sense of minimum Euclidean distance in R^2. This neuron is called as

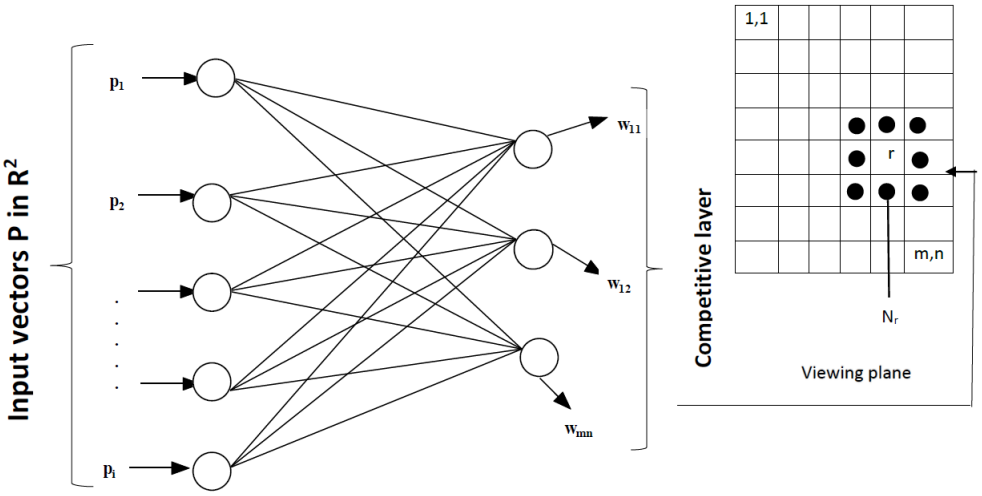


Figure 2.10: SOFM Architecture

best matching neuron(BMN). Then $w_{i,t-1}$ and the other weights in its spatial neighbourhood are updated using the the following least mean squared (LMS) updation rule [52],

$$w_{i,t} = w_{i,t-1} + \alpha_t q_t \left(p - w_{i,t-1}\right) \tag{2.3}$$

where, α_t is learning parameter and $q_t = \exp^{-dist^2(r,p)\sigma_t^2}$ is neighborhood function that depends upon lattice distance between the BMN and other neurons.

α_t and σ_t decrease with time t. The topological neighborhood also decreases with time. This scheme, when repeated long enough, preserves the spatial order, that is, the weight vectors which are metrically close in $\Re^2$ generally have visually close images in the viewing plane. Also, the distribution of the weight vectors in $\Re^2$ resembles closely the distribution of the training vectors P. So, the weight vectors approximate the distribution of the training data as well as preserve the topology of input data on the viewing plane.

In this dissertation, the vector quantizer design algorithm proposed by [52] is used for image compression. A two-dimensional SOFM is used to generate the initial codebook. The properties of SOFM are used to create and train the codebook of

the vector quantizer.

In general, in vector quantization, the codebooks are trained with test image itself. This poses a problem in practical use of such algorithm for transmission /storage of the compressed image. The compression ratio achieved is diminished to a great extent by the transmission of codebook. Hence a codebook that is generic to both transmitter and receiver is used. The reconstruction error tends to be high if the image is compressed using a codebook trained on different image. This codebook allows one time reconstruction of encoder and decoder and making code book a permanent part of it.

For the training purpose, a set of images, having varied nature in terms of details, contrasts and textures are selected. A composite image of size 768×512, shown in Figure 2.11, comprising six smaller images is used. Thus, a generic code book is constructed using some judiciously chosen images for effective compression of images having similar characteristics. Here, similar characteristics mean the images having similar distribution of gray levels over small blocks of size 8×8. Once the SOFM is trained, the codebook can easily be designed using weight vectors as reconstruction vectors. The images can be encoded by finding out, for each image vector, the code vector with least Euclidean distance.

2.5.2 Coding Performance of Vector Quantizers for Gray-scale Images

For the purpose of performance analysis, three vector quantizers using block sizes 4×4 (VQ1), 8×8 (VQ2), 4×8 (VQ3) are developed. Each VQ uses a code book of size 256 and is trained with mean removed vectors. Hence, to represent each block in the encoded image, one byte is required for index and other byte is required for block average. The average word length for VQ1 is 1 bpp; for VQ2, 0.25 bpp and for VQ3, 0.5 bpp.

Figure 2.11: The Training Image

The codeword assignment for the indices is based on the frequency distribution of the code-vectors in the encoded training image. As a strong correlation between neighbouring blocks, the absolute differences between average values of neighbouring blocks are found to have a monotonically decreasing distribution and codewords are assigned exploiting this correlation.

Table 2.8: Simulation Results of SOFM based vector quantization for VQ1

Image	PSNR(dB)	VIFP	MSSIM	UQI
Lena	35.71	0.8088	0.8444	0.9997
Barbara	32.70	0.7493	0.7951	0.9980
Babbon	30.89	0.7083	0.7832	0.9995
Peppers	36.42	0.828	0.8949	0.9945

(a)

(b)

(c)

Figure 2.12: Visual quality comparison of decoded Lena image between various vector quantizers (a) VQ1 (b) VQ2 (c) VQ3

Table 2.9: Simulation Results of SOFM based Vector Quantization for VQ2

Image	PSNR(dB)	VIFP	MSSIM	UQI
Lena	33.44	0.5341	0.7317	0.9980
Barbara	31.07	0.5030	0.6165	0.9937
Babbon	29.85	0.3758	0.4013	0.9971
Peppers	33.36	0.5319	0.7767	0.9901

The vector quantizers are tested on different test images for their compression performance. The results are summarised in the tables: Table 2.8 through Table 2.10. The comparison of numerical values of various metrics in these ta-

Table 2.10: Simulation Results of SOFM based Vector Quantization for VQ3

Image	PSNR(dB)	VIFP	MSSIM	UQI
Lena	35.50	0.8064	0.8420	0.9944
Barbara	32.70	0.7488	0.7450	0.9981
Babbon	30.86	0.6633	0.6624	0.9911
Peppers	33.61	0.7919	0.8266	0.9976

bles indicate that VQ1 exhibits better performance in terms of distortion metrics PSNR, VIFP, MSSIM and UQI than VQ2 and VQ3 for almost all images, but the compression ratio is very low(CR =8). VQ3 gives double the compression ratio with near identical distortion values when compared with VQ1. VQ2 provides compression four times that of VQ1, but with sacrifice in quality of images. Thus vector quantizer VQ1 gives better performance in terms of all the quality metrics used. Figure 2.12 shows the test image Lena, compressed using the three vector vector quantizers. The pschovisual quality of the image of Figure 2.12 (b) is poorer as indicated by blockiness in the image.

2.6 Deblocking of Gray-scale and Color Images in DCT based Compression [P8, P9, P10]

Block based discrete cosine transform (BDCT) has been widely used in image compression. In BDCT coding, an image is first divided into 8×8 non overlapping blocks. Each block is then transformed using the DCT, followed by quantization and variable length coding. At low bit rates, the coarse quantization of the DCT coefficients causes artificial discontinuities along the block boundaries in the compressed image [118]. Sometimes, the two low-frequency DCT coefficients in adjacent blocks, which are similar in value, are quantized into different quantization bins. This unpleasant visible degradation, due to the negligence of correlation among adjacent blocks, is called blocking artifact. In order to achieve high-compression rates (low bit rates), using block transform coder with visually

acceptable results, it is required to eliminate the blocking artifacts. This procedure is referred to as deblocking.

In order to reduce the annoying blocking artifacts, several deblocking algorithms have been proposed in literature. These algorithms are classified into two types: pre-processing algorithms and post-processing algorithms. Pre-processing techniques are used at the encoding end. This type of techniques requires coding schemes of its own that include transform, quantization, and bit allocation. Therefore, these approaches can hardly be applied to the commercial coding system products, such as JPEG and MPEG. Hence post-processing techniques are preferred. Alessandro Foi *et al.* [113] have proposed a post-processing method based on point-wise shape adaptive DCT (SA-DCT) for deblocking of BDCT compressed images. The algorithm is developed for denoising of images and is extended for deblocking. For deblocking purpose, the degradation by quantization noise is modelled as some additive white Gaussian noise.
The observation model is given by

$$\phi(x, y) = f(x, y) + \eta(x, y) \tag{2.4}$$

where $f(x, y)$ is the original (non-compressed) image, $\phi(x, y)$ its observation after quantization in BDCT domain, and $\eta(x, y)$ is independent Gaussian noise with variance σ^2, $\eta(.) \sim \mathcal{N}(0, \sigma^2)$.

A suitable value for the variance σ^2 is estimated directly from the quantization table $Q = [q_{i,j}]$ $i, j = 1$ *to* 8 using the empirical formula [113]

$$\sigma^2 = 0.69 \left(\frac{1}{9} \sum_{i,j=1}^{3} q_{ij} \right)^{1.3} \tag{2.5}$$

This formula uses only the mean value of the nine table entries which correspond

to the lowest-frequency DCT harmonics (including the DC-term). A higher compression obviously corresponds to a larger value for the variance. The σ^2 which is calculated by (2.5) is simply the variance of the white Gaussian noise η in the observation model of (2.4). It is the variance of some hypothetical noise which, if added to the original image $f(x, y)$, would be required in order to be removed with same level of adaptive smoothing which is necessary to suppress the artifacts generated by the BDCT quantization with the table Q.

Algorithm

- The image $f(x, y)$ is made zero-mean data by subtracting its mean from the the original image. This zero mean data is input to the denoising algorithm.
- The input to the denoising algorithm is processed upon by point-wise SA-DCT used in conjunction with anisotropic LPA-ICI (linear polygonal approximation intersection of confidence intervals). The shape adaptive transform in conjunction with LPA-ICI generates anisotropic estimates that are highly sensitive to with respect to change points in the image and allows to reveal the fine elements of images from noisy observations [114–117].
- A local estimate is obtained by performing hard thresholding in SA-DCT domain using a universal hard threshold given by $\sigma\sqrt{2ln\,|U_x| + 1}$ where U_x represents adaptive shape neighbourhood determined by anisotropic LPA-ICI. The thresholding coefficient is calculated using

$$\gamma = T \times \sigma\sqrt{2ln\,|U_x| + 1} \tag{2.6}$$

  T is usually taken as contant (typically 0.925) [113]. The current work modifies the value T so as to improve the visual quality of image

  Simulations are carried out on Lena, Barbara and Peppers images to determine the optimal value of T. The value of T is modified from 0.925 (fixed value) to a range of 0.1 to 1.0.

- All the local estimates obtained in previous step are averaged using adaptive weights that depend upon on their local variances and size of the corresponding adaptive-shape regions.
- The global estimates are used as reference to perform Wiener filtering in SA-DCT domain and local Weiner estimates are obtained.
- All the local Wiener estimates are averaged together using adaptive weights that depend on the size of the corresponding adaptive-shape regions.

2.6.1 Simulation Results for Gray-scale Images

Three quantization tables usually called **Q1**, **Q2** and **Q3** have been used by many researchers [125] in order to simulate various types of BDCT compression. These tables are given for reference.

Table 2.11: Quantization Table Q1

50	60	70	70	90	120	255	255
60	60	70	96	30	255	255	255
70	70	80	120	200	255	255	255
70	96	120	140	255	255	255	255
90	130	200	145	255	255	255	255
120	255	255	255	255	255	255	255
255	255	255	255	255	255	255	255
255	255	255	255	255	255	255	255

Table 2.12: Quantization Table Q2

86	59	54	86	129	216	255	255
64	64	75	102	140	255	255	255
75	70	86	129	216	255	255	255
75	91	118	156	255	255	255	255
97	118	199	255	255	255	255	255
129	189	255	255	255	255	255	255
255	255	255	255	255	255	255	255
255	255	255	255	255	255	255	255

Table 2.13: Quantization Table Q3

110	130	150	192	255	255	255	255
130	150	192	255	255	255	255	255
150	192	255	255	255	255	255	255
192	255	255	255	255	255	255	255
255	255	255	255	255	255	255	255
129	255	255	255	255	255	255	255
255	255	255	255	255	255	255	255
255	255	255	255	255	255	255	255

Table 2.14: Performance of deblocking algorithm for various values of threshold for quantization Table Q1 on test image Lena

T	PSNR(dB)	VIFP	MSSIM	UQI
0.1000	31.75	0.8549	1.0000	0.5734
0.2000	31.83	0.8541	1.0000	0.5759
0.3000	31.91	0.8529	1.0000	0.5780
0.4000	31.99	0.8512	1.0000	0.5791
0.5000	32.06	0.8493	1.0000	0.5789
0.6000	32.10	0.8470	1.0000	0.5776
0.7000	32.12	0.8443	1.0000	0.5756
0.8000	32.13	0.8411	1.0000	0.5732
0.9000	32.12	0.8377	1.0000	0.5705
0.9250	32.12	0.8368	1.0000	0.5698
1.0000	32.10	0.8341	1.0000	0.5677

The values of the standard deviation σ corresponding to these three tables calculated using (2.5) are 12.62, 13.21, and 22.73, respectively. In terms of image degradation, they correspond to a medium to high compression level, similar to what is obtained using JPEG with $Q = 11(Q1)$, $Q = 9(Q2)$, or $Q = 5(Q3)$.

Lena, Barbara and Peppers images of size 512×512 are used for comparison of modified point-wise SA-DCT algorithm against the original point-wise SA-DCT algorithm [113]. Table 2.14, Table 2.15 and Table 2.16 present the results for deblocking from BDCT quantization performed using these specific quantization tables for test image Lena. The value of T is varied from 0.1 to 1.0. The results

Table 2.15: Performance of deblocking algorithm for various values of T for quantization Table Q2 on test image Lena

T	PSNR(dB)	VIFP	MSSIM	UQI
0.1000	31.75	0.8496	1.0000	0.5463
0.2000	31.83	0.8488	1.0000	0.5498
0.3000	31.14	0.8474	1.0000	0.5531
0.4000	31.22	0.8457	1.0000	0.5557
0.5000	31.30	0.8437	1.0000	0.5574
0.6000	31.39	0.8413	1.0000	0.5577
0.7000	31.45	0.8385	1.0000	0.5569
0.8000	31.50	0.8353	1.0000	0.5551
0.9000	31.54	0.8318	1.0000	0.5527
0.9250	31.55	0.8308	1.0000	0.5521
1.0000	31.55	0.8341	1.0000	0.5501

Table 2.16: Performance of deblocking algorithm for various values of T for quantization Table Q3 on test image Lena

T	PSNR(dB)	VIFP	MSSIM	UQI
0.1000	28.61	0.8496	1.0000	0.7395
0.2000	28.71	0.8488	1.0000	0.7380
0.3000	28.82	0.8474	1.0000	0.7357
0.4000	28.91	0.8457	1.0000	0.7327
0.5000	28.98	0.8437	1.0000	0.7291
0.6000	29.02	0.8413	1.0000	0.7248
0.7000	29.04	0.8385	1.0000	0.7200
0.8000	29.04	0.8353	1.0000	0.7148
0.9000	29.03	0.8318	1.0000	0.7093
0.9250	29.03	0.8308	1.0000	0.7079
1.0000	26.01	0.8281	1.0000	0.7039

obtained are compared with the point-wise SA-DCT algorithm as reported in [113] which uses $T = 0.925$. A variation of around $0.37dB$ is observed in all three cases. The value of VIFP decreases as the value of T is increased, whereas the value of MSSIM remains constant. A fluctuation in UQI is observed for all three cases. For Table 2.14, the value of UQI increases as T is varied from 0.1 to 0.5, and then starts decreasing. So the value of T should be chosen between 0.1 to 0.5. Further

Table 2.17: Optimal value of T for modification of universal threshold

Lena				
Quantization Table	T	PSNR	VIFP	UQI
Q1	0.3980	31.99	0.8513	0.5791
Q2	0.5800	31.50	0.8418	0.5578
Q3	0.5201	28.99	0.7283	0.5646
Barbara				
Q1	0.4600	26.64	0.7123	0.6191
Q2	0.6101	26.36	0.6976	0.5973
Q3	0.6400	25.10	0.5514	0.4929
Peppers				
Q1	0.5	31.84	0.8739	0.5536
Q2	0.64	31.33	0.8701	0.5401
Q3	0.7000	29.07	0.7817	0.4687

investigation reveals that the optimal value of T is 0.3980. Similarly, the optimal values of T are found for other quantization tables and images and the results are reported in Table 2.17.

2.6.2 Simulation Results for Color Images

For color image compression, an image first undergoes RGB to YUV color transformation, which decomposes the image into in one luminance and two chrominance channels and then process the resulting three channels separately. According to the modeling in the previous section, it is assumed that the original image $f(x, y) = [f_R, f_B, f_G]$ in the RGB color space is represented, after BDCT quantization in YUV space as

$$\phi_C = f_C + \eta_C, \quad C = \{Y, U, V\} \tag{2.7}$$

where, f_Y, f_U, f_V are luminance and chrominance channels of $f(x, y)$ and ϕ_Y, ϕ_U, ϕ_V are the corresponding channels after quantization in BDCT domain, and $\eta =$

Table 2.18: Optimal value of T for different quality factors of JPEG compressed color images

Quality	Lena	Peppers	Baboon
4	1.9	1.9	0.985
6	1.2	1.4	0.905
8	0.985	1.4	0.9
10	0.965	1.4	0.895
15	0.94	1.4	0.935
20	0.905	1.4	0.925
25	0.925	1.4	0.928
30	0.804	1.4	0.685
40	0.784	1.4	0.625
50	0.784	1.4	0.555
75	0.784	0.925	0.355

$[\eta_Y,\ \eta_U,\ \eta_V]$ is independent Gaussian noise, $\eta_C \sim \mathcal{N}(0, \sigma_C^2)$
The estimate of the variances $\sigma_Y^2,\ \sigma_U^2$ and σ_V^2 from the corresponding quantization tables for the luminance and chrominance channels is done using (2.5). However, as the chrominance channels are down sampled, the estimated variances for the chrominances need to be further multiplied by 2, in order to take into account for the coarser sampling.

Ideally, the Y, U, and V channels are considered as independent. Therefore, the common approach is to filter the three channels separately and independently one from the other.

However, when considering natural images, the different color channels typically share some common features which are inherited from the structures and from the objects depicted in the original image. In particular, it can be observed that along the objects' boundaries all color channels of the original image usually exhibit some simultaneous discontinuities or sharp transitions.

This kind of structural correlation is exploited by imposing that the three transforms supports which are used for the filtering of the Y, U, and V channels at a particular location have the same adaptive shape. The adaptive neighborhoods

defined by the anisotropic LPA-ICI for the Y channel are used by all the three channels, because it is in the luminance that the structural information is usually better preserved after compression.

Such a constraint is imposed so that whenever some structure is detected, it is taken into account and thus preserved for the filtering of all three channels.

Table 2.19: CPSNR (dB) results for deblocking of JPEG-Compressed Color Images showing the comparison of the point-wise SA-DCT method [85] and our Proposed Modified point-wise SA-DCT method.

Quality	Lena			Peppers			Baboon		
	JPEG	SA-DCT	Modified SA-DCT	JPEG	SA-DCT	Modified SA-DCT	JPEG	SA-DCT	Modified SA-DCT
4	23.34	24.79	24.93	22.32	23.77	23.96	19.28	19.99	20.01
6	25.52	27.09	27.13	23.99	25.53	25.62	20.38	21.05	21.06
8	26.64	28.16	28.17	24.99	26.39	26.45	21.12	21.70	21.72
10	27.53	29.05	29.07	25.77	27.10	27.13	21.63	22.13	22.13
15	28.97	30.32	30.33	26.88	27.98	28.01	22.49	22.87	22.88
20	29.83	30.99	30.99	28.04	28.89	28.92	23.5	23.37	23.37
25	30.44	31.45	31.45	28.04	28.89	28.99	23.50	23.75	23.76
30	30.91	31.78	31.79	28.4	29.13	29.19	23.85	24.06	24.07
40	31.54	32.25	32.26	28.83	29.45	29.46	24.40	24.56	24.58
50	32.02	32.62	32.63	29.25	29.81	29.83	24.85	24.96	24.99
75	33.21	33.56	33.58	30.29	30.66	30.66	26.21	26.25	26.28

The point-wise SA-DCT algorithm is also used for JPEG compressed color images, from very high $Q = 4$ to very low $Q = 50$ compression levels. For the simulations in color domain, the baseline IJG JPEG implementation is used. Extensive simulations are carried out on color images of Lena, Peppers and Baboon of size 512×512 for various quality to determine the optimal value of T for universal threshold. The value of T is varied between 0.1 to1.9 for JPEG quality of $Q = 4$ to $Q = 75$. The values of T for the three test images and different quality factors are reported in Table 2.18. For different quality factors the optimal value of T is different. The results for CPSNR corresponding to the optimal T are reported in Table 2.19. It is observed that the improvement in point-wise SA-DCT is significant especially for very high and moderate compression levels as compared to JPEG. Thus, the proposed method of selecting the optimal value of T improves

the visual quality of the image.

2.7 Conclusion

In this chapter, various basic lossy and lossless image compression schemes are discussed. The compression performances of DCT based scheme, SPIHT and SOFM based vector quantization are evaluated in terms of the distortion metrics: PSNR (dB), VIFP, MSSIM and UQI. Lossless scheme, CALIC is also discussed, analysed and simulated and its compression performance is examined. These schemes will be employed in our work for development of efficient hybrid compression algorithms.

Since BDCT is usually employed in many real-time image and video compression systems (encoders) that yield unwanted blocking artifacts, we have developed a modified point-wise SA-DCT algorithm that yields promising results in smoothing and deblocking. Such schemes are expected to be used along with decoders in the receivers for obtaining better visual performance.

Chapter 3

Development of Hybrid Image Compression Scheme using SPIHT and SOFM based Vector Quantization

Preview

Many methods of embedded image compression are proposed in literature but the basic idea arises from embedded zero tree wavelet (EZW) algorithm [64]. SPIHT [57] is one of the embedded image compression algorithm that achives equal to or better performance of EZW without using arithmetic coding. The reduction in complexity from eliminating the arithmetic encoder is signicant. Therefore, SPIHT is used frequently as a benchmark for performance in evaluation of the state-of-the-art image compression algorithms. Vector quantization [8] is also known to be an efficient method for data compression.The performance of a vector quantizer depends on two factors, the quality of the codebook and the time required for codebook searching at the encoding stage. The self-organizing feature map (SOFM) introduced by Kohonen [127] can be used for constructing a good quality codebook. This chapter discusses our proposed hybrid algorithm which combines SPIHT and SOFM based vector quantization for image compression.

3.1 Introduction

Wavelets offer an elegant technique for representing the details present in an image. When an image is decomposed using wavelets, the high pass components carry less information. The possibility of elimination of high-pass components gives higher compression ratio in case of wavelet based image compression.

SPIHT is a method of coding and decoding the wavelet transform of an image. By coding and transmitting information about the wavelet coefficients, it is possible for a decoder to perform an inverse transformation on the wavelet and reconstruct the original image. The entire wavelet coefficient does not need to be transmitted in order to recover the image. Instead, when the decoder receives more information about the original wavelet transform, the inverse-transformation will yield a better quality reconstruction of the original image. SPIHT generates excellent image quality and performance due to several properties of the coding

algorithm. The properties are in partial ordering by coefficient value, taking advantage of redundancies between different wavelet scales and transmitting data in bit-plane order following a wavelet transformation. The SPIHT algorithm does not need any data training, supports the multi-coding rate, and has higher signal-to-noise ratio and better quality of the recovered image [57].

Vector Quantization [119, 120] is a clustering technique by which an input space is divided into a number of distinct regions and for each region a reconstruction vector is defined. Self Organizing Feature Maps have been extensively applied to vector quantization [52, 121] to solve the main problem associated to the classical VQ techniques, which are rather sensitive to codeword errors. Due to the capability of Self Organizing Maps to form ordered topological feature maps, the SOFM's weight vectors are spatially ordered in an array such that the neighboring vectors in the map are more alike than the more distant ones resulting in optimal codebook and partition-design. The fundamental concept of VQ (referred to as memoryless VQ) exploits the statistical redundancy between pixels within the same vector to reduce the bit-rate [122].

A SOFM consists of a two layer network as shown in Figure 2.10. The weights of neurons are initialised to some random values. Then the network is fed with large number of training vectors. For a training sample, the Euclidean distance measure with all the weight vectors is computed. The neuron with weight vector closest to the input vector is called the winning vector. Then, all the weight vectors are updated using the weight updation rule given by (2.3). This process is repeated for each training vector till convergence is achieved. After the training is over, a set of all the final code vectors is called SOFM code-book. The main interesting properties of SOFM are [123]:

- Self-organizing algorithm: It does not need to classify the training data (unsupervised learning);
- Ability to form ordered topological feature maps [131];

- Quantization is performed in the gray-level (spatial domain).

3.2 Hybrid Coding Scheme for Effectual Image Compression, SPIHT-VQ [P1,P3]

The proposed hybrid scheme integrates two classes of image compression schemes: transform based image coding and vector quantization. The two schemes integrated in the proposed hybrid scheme are:

- Wavelet transform-based SPHIT coding scheme;
- SOFM based Vector Quantization.

The positive aspects of both the schemes are exploited in the design of the proposed hybrid scheme. The block diagram of the encoder and decoder involved in the proposed hybrid coding scheme are depicted in Figure 3.1 and Figure 3.2 respectively.

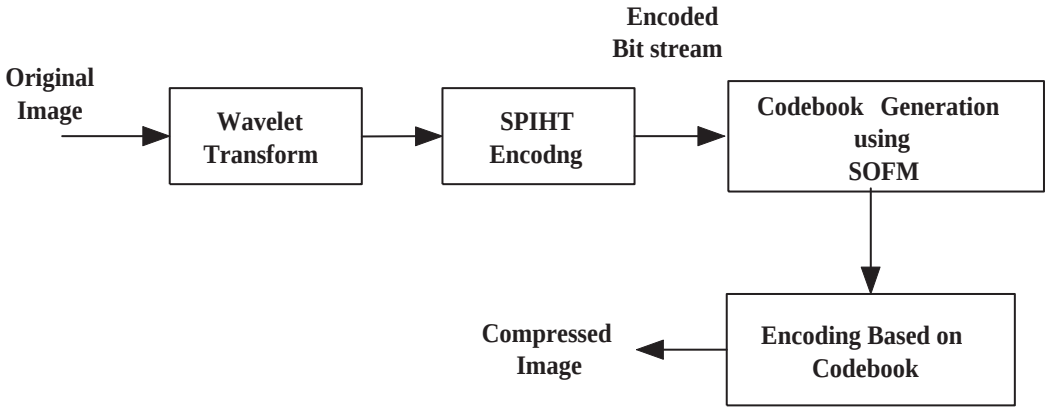


Figure 3.1: Block Diagram of Proposed Encoder

The input to the proposed coding scheme is a dyadic square ($2^n \times 2^n$) image. First, the input image is subjected to transform-based SPIHT encoding. Primarily, SPIHT is a wavelet-based image compression coder; wavelet transform has been

Figure 3.2: Block Diagram of Proposed Decoder

made use of for image decomposition. The SPIHT coder encodes the decomposed image to a bit stream. Subsequently, the bit stream obtained from the SPIHT coder is fed to the SOFM based VQ coding for further compression. In the first phase, the SOFM generates a codebook based on the bit stream resulted from the SPIHT coder. In the second phase, SOFM produces an encoded data based on the codebook generated. The vector quantizer used here is as described in [52]. This vector quantizer uses the generic code book as described in Chapter 2. The code book is used for encoding the SPIHT bit stream output using the following steps.

1. The input bit stream to vector quantizer is divided into blocks.
2. The blocks are provided to the SOFM.
3. Inner neuron is selected as the neuron having the minimum Euclidean distance.
4. The index of the winner neuron for each input block is stored.
5. The set of indices of all the winner neurons for the blocks and the corresponding codebook forms the compressed form of the data.

The image can be reconstructed from the encoded data by applying the linear combination of the processes involved in the encoder.

3.3 Selection of Wavelet for SPIHT and SOFM based Vector Quantization

Wavelet transform [132–134, 137], due to its localization property, has become an indispensable signal and image processing tool for variety of applications, including compression and denoising [138–140, 143]. A wavelet is a mathematical function used to decompose a given function or continuous time signal into different frequency components and study each component with a resolution that matches its scale. The wavelets are scaled and translated copies (known as daughter wavelets) of a finite-length or fast decaying oscillating waveform known as mother wavelet or basis function.

Wavelet transforms are classified into continuous wavelet transform (CWT) and discrete wavelet transform (DWT). The continuous wavelet transform [3, 11–13, 144, 145] has received significant attention for its ability to perform a time scale analysis of analog signals. On the other hand, DWT is an implementation of wavelet transform using a set of discrete wavelet scales and translations obeying some definite rules. In DWT the image is decomposed using discrete set of wavelet coefficients using an orthogonal set of basis functions. Integer wavelet transform (IWT)and DWT are popular in the field of compression. IWT leads to much larger degradation than DWT, due its large quantization error [147].

The construction of the majority of useful wavelet bases is a consequence of the design of some 2π periodic functions called wavelet filters by the signal processing community. Any wavelet filter with finite or infinite length is a finite impulse response (FIR) or infinite impulse response (IIR) filter, respectively. The properties of wavelets that suit image compression are compact support, symmetry, orthogonality, regularity and degree of smoothness. Compact support leads to efficient implementation. Symmetry property helps in avoiding dephasing in image processing. Orthogonality allows fast implementation of algorithm. Regularity and degree of smoothness properties are related to filter order and length [126].

Higher order filters have a good frequency localization. But, blurring of image takes place due to high degree of smoothness. There are more vanishing moments with higher order filters and more energy compaction. Lower order filters have good time localization, preserve edge information, have less energy compaction and demonstrate more blockiness. Smooth functions are preferred for wavelets used in image compression. Complexity of DWT increases with increase in filter length [27]. Hence, in image compression applications, balance of filter length, degree of smoothness and computational complexity is required.

Many wavelet families are proposed in literature. The first set of orthogonal basis functions is developed by Harr in 1910 [148]. Then several orthogonal and orthonormal basis functions have been reported in literature. Stromberg developed the first orthonormal basis which has exponential decay in time and frequency in 1982 [151]. In 1985, Meyer constructed the bases that are compactly supported in frequency domain [152]. Lemarie has constructed a wavelet basis with bounded regularity and exponential decay [153]. The DWT can be implemented by passing the signal through a combination of low-pass and high-pass filters and down sampling by a factor of two to obtain a single level of decomposition. Multiple levels of the wavelet transform are performed by repeating the filtering and down sampling operation on low-pass branch outputs. The coefficients of orthogonal filters are real numbers. The filters are of the same length and are not symmetric. The low pass filter, G_0 and the high pass filter, H_0 are related to each other by $H_0(z) = z^{-N}G_0(z^{-1})$. The two filters are alternated flip of each other. The alternating flip automatically gives double-shift orthogonality between the low pass and high pass filters, i.e., the scalar product of the filters, for a shift by two is zero.

$$\sum G(k)H(k-2l) = 0 \tag{3.1}$$

where $(k, l\ \varepsilon\mathbb{Z})$.

Filters that satisfy (3.1) are known as Conjugate Mirror Filters (CMF). Perfect

reconstruction is possible with alternating flip. Also, for perfect reconstruction, the synthesis filters are identical to the analysis filters except for a time reversal. Orthogonal filters offer a high number of vanishing moments. This property is useful in many signal and image processing applications. They have regular structure which leads to easy implementation and scalable architecture.
In the case of the bi-orthogonal wavelet filters, the low-pass and the high-pass filters do not have the same length. The low-pass filter is always symmetric, while the high pass filter could either be symmetric or anti-symmetric. The coefficients of the filters are either real numbers or integers. For perfect reconstruction, bi-orthogonal filter bank has all odd length or all even length filters. The two analysis filters can be symmetric with odd length or one symmetric and the other anti symmetric with even length. Also, the two sets of analysis and synthesis filters must be dual. The linear phase bi-orthogonal filters are the most popular filters for data compression applications. Each wavelet family is characterised by by an integer N that denotes the filter order. Daubechies wavelets are denoted by dbN and are tested for filter order of 1 to 44. Coiflet wavelet is denoted by CoifN and is tested for filter order of 1 to 5. Biorthogonal wavelets use filters of similar or dissimilar order for decomposition and reconstruction. N_d denotes decomposition filter, N_r denotes reconstruction filter and the biorthogonal wavelets are denoted by $bior N_r N_d$.

Among the many available wavelets, choice of wavelet is crucial for coding performance in image compression. The wavelet function chosen should be adjusted to spectral activity in the images. The compression performance for images with high spectral activity is fairly insensitive to choice of compression scheme whereas coding performance for the images having moderate spectral activity is more sensitive to the choice of compression scheme [27]. Therefore, to select an optimal wavelet basis function, it is required to select the test images with moderate spectral activity. For the purpose of analysis, the commonly used Harr wavelet (also called db1), Daubechies wavelet, bi-orthogonal wavelets and Coi-flets are

considered. These wavelets are tested on the test image Lena that exhibits moderate spectral activity. This is based on the objective picture quality measures: PSNR(dB), VIFP, MSSIM and UQI. Simulations are carried out on the test im-

(a) (b)

Figure 3.3: Visual quality comparison of decompressed Lena image using different wavelets at a bit rate of 0.0625 bpp (a)bior4.4, PSNR = 22.28dB, VIFP =0.4945 (b) db1, PSNR = 29.36dB, VIFP = 0.3957

age Lena with different wavelets keeping the target bit rate of 0.25 bpp for SPIHT coder. The bit stream output of SPIHT is further compressed by the vector quantizer and the overall compression ratio is 0.0625 bpp. The results are tabulated in Table 3.1. All the wavelets use a target bit rate of 0.25 bpp so that overall compression ratio at hte output of the vector quantizer is 0.0625 bpp.

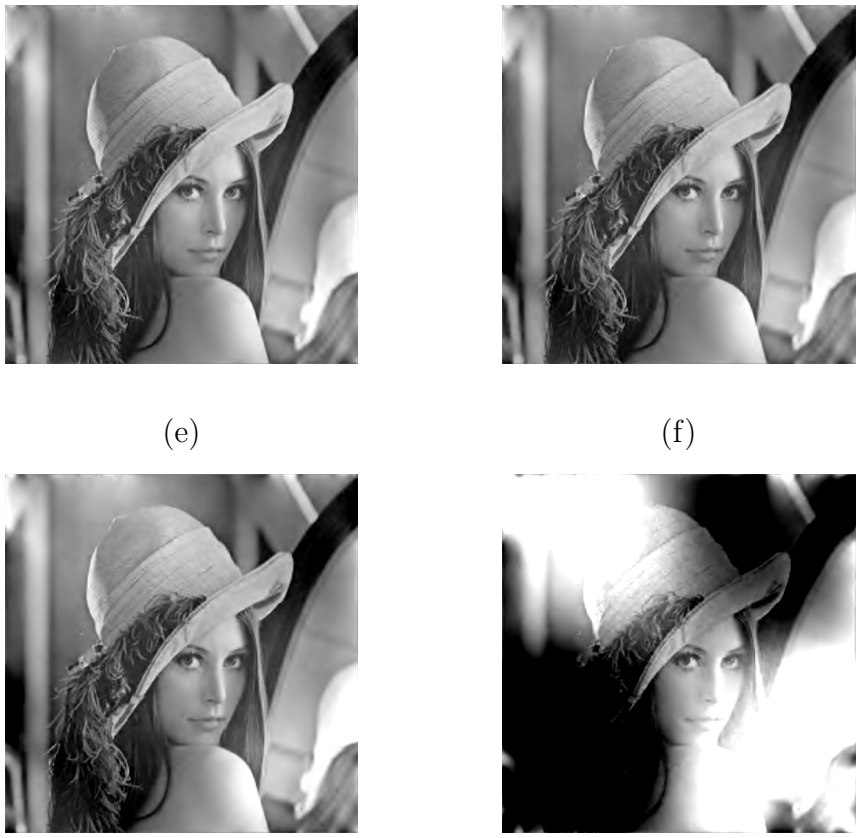
It is observed that the wavelet functions: bior1.5, bior1.3, bior1.1 and db1 perform better in terms of PSNR. The wavelet functions db1 and bior1.1 are the best in terms of PSNR performance with a value of 29.36 dB. But the performance of these wavelets in terms of other quality metrics, VIFP, MSSIM and UQI is poor as compared to bior4.4 wavelet. It is observed that bior4.4 wavelet performs better than all the other wavelets in terms of VIFP, MSSIM, and UQI although its PSNR performance is only 22.28 dB. Figure 3.3(a) and (b) give the visual quality of image Lena for wavelets bior4.4 and db1. It is observed that although db1 has good PSNR performance, there is lot of blocking artifact visibility in the image

of Figure 3.3(b) as compared to Figure 3.3(a). This degradation is more severe at still lower bit rates. Hence bior4.4 wavelet is used to evaluate the proposed scheme SPIHT-VQ. The visual quality for all the wavelets is shown in Figure 3.8. The wavelet function bior3.1 performs poorly in in terms of all the metrics.

(a) (b)

(c) (d)

Figure 3.4: Visual quality comparison of decompressed Lena image using different wavelets at a bit rate of 0.0625 bpp (a)bior6.8 (b) bior5.5 (c) bior4.4 (d) bior3.9



(e) (f)

(g) (h)

Figure 3.5: Visual quality comparison of decompressed Lena image using different wavelets at a bit rate of 0.0625 bpp (e) bior3.7 (f)bior3.5 (g) bior3.3 (h) bior3.1

(i) (j)

(k) (l)

Figure 3.6: Visual quality comparison of decompressed Lena image using different wavelets at a bit rate of 0.0625 bpp (i) bior2.8 (j) bior2.6 (k) bior2.4 (l) bior2.2

(m)

(n)

(o)

(p)

Figure 3.7: Visual quality comparison of decompressed Lena image using different wavelets at a bit rate of 0.0625 bpp (m) bior1.5 (n) bior1.3 (o) bior1.1 (p) db1

(q)

(r)

(s)

Figure 3.8: Visual quality comparison of decompressed Lena image using different wavelets at a bit rate of 0.0625 bpp (q) db2 (r) coiflet1 (s) symlet2

Table 3.1: Rate Distortion Performance of Different Wavelets tested on test image Lena at bit rate of 0.0625bpp

Wavelet	PSNR(dB)	VIFP	MSSIM	UQI
bior6.8	22.08	0.4909	0.8478	0.5703
bior5.5	20.67	0.4519	0.8304	0.5446
bior4.4	22.28	0.4945	0.8473	0.5681
bior3.9	21.31	0.4353	0.8358	0.5661
bior3.7	20.85	0.4350	0.8307	0.5622
bior3.5	19.83	0.4300	0.8187	0.5541
bior3.3	17.27	0.4071	0.7993	0.5225
bior3.1	13.42	0.2135	0.2449	0.1012
bior2.8	22.07	0.4695	0.8406	0.5568
bior2.6	22.19	0.4759	0.8425	0.5583
bior2.4	22.37	0.4778	0.8434	0.5602
bior2.2	22.79	0.4840	0.8445	0.5619
bior1.5	27.59	0.3539	0.7853	0.4843
bior1.3	28.53	0.3694	0.7954	0.4670
bior1.1	29.36	0.3957	0.7992	0.4670
dB1 (Haar)	29.36	0.3957	0.7992	0.4670
dB2	20.49	0.4274	0.8119	0.5077
coiflet1	22.92	0.4547	0.8292	0.5348
symlet2	20.49	0.4274	0.8119	0.5077

3.4 Coding Performance of SPIHT-VQ

The coding performance of SPIHT-VQ is tested for three different vector quantizers, VQ1, VQ2 and VQ3 mentioned in Chapter 2. The bit-rate for the hybrid scheme is varied from 0.0156 bpp to 0.5 bpp and the results are summarised in the tables: Table 3.2 through Table 3.5. The bold figures in the table indicate the better performance of SPIHT-VQ2.

The SPIHT-VQ scheme cascades two lossy compression schemes. The objective is to improve the compression ratio while maintaining the visual quality of the image. It is observed from Table 3.2 that variation in PSNR values is around

Table 3.2: Coding Performance of the proposed scheme SPIHT-VQ in terms of PSNR(dB)

Lena						
bpp	0.0156	0.0313	0.0625	0.125	0.25	0.5
SPIHT-VQ1	19.80	20.44	21.22	21.84	22.27	22.53
SPIHT-VQ2	21.18	21.80	22.23	22.49	**22.84**	23.15
SPIHT-VQ3	20.43	21.22	21.85	22.28	22.53	22.68
Barbara						
SPIHT-VQ1	17.78	18.40	18.97	19.40	20.15	20.81
SPIHT-VQ2	18.97	19.40	20.15	20.81	21.32	21.96
SPIHT-VQ3	18.40	18.98	19.40	20.16	20.82	21.25
Peppers						
SPIHT-VQ1	18.04	18.82	19.42	19.91	20.17	20.3
SPIHT-VQ2	19.42	19.90	20.16	20.30	20.37	20.86
SPIHT-VQ3	18.82	19.42	19.91	20.17	20.30	20.37
Baboon						
SPIHT-VQ1	18.64	18.92	19.42	20.05	21.00	22.38
SPIHT-VQ2	19.41	20.05	21.00	22.38	24.16	25.35
SPIHT-VQ3	18.92	19.42	20.05	21.00	22.38	24.17

Table 3.3: Coding Performance of the proposed scheme SPIHT-VQ in terms of Visual Information Fidelity(VIFP)

Lena						
bpp	0.0156	0.0313	0.0625	0.125	0.25	0.5
SPIHT-VQ1	0.1531	0.2037	0.2648	0.3469	0.4361	0.5374
SPIHT-VQ2	0.2641	0.3471	0.4361	0.5374	**0.6532**	0.6618
SPIHT-VQ3	0.2469	0.3222	0.4065	0.4945	0.5875	0.6693
Barbara						
SPIHT-VQ1	0.1211	0.1647	0.2028	0.2544	0.3250	0.4389
SPIHT-VQ2	0.2027	0.2544	0.3250	0.4389	0.5943	0.6165
SPIHT-VQ3	0.2262	0.2571	0.3137	0.3728	0.4789	0.6055
Peppers						
SPIHT-VQ1	0.1476	0.1937	0.2528	0.3216	0.4025	0.4851
SPIHT-VQ2	0.2525	0.3266	0.4025	0.4851	0.5717	0.5960
SPIHT-VQ3	0.2446	0.3042	0.3770	0.4601	0.5394	0.5890
Baboon						
SPIHT-VQ1	0.1783	0.2038	0.2377	0.2956	0.3970	0.4905
SPIHT-VQ2	0.0879	0.1300	0.1765	0.2671	0.3988	0.4998
SPIHT-VQ3	0.1432	0.1438	0.1923	0.2358	0.3214	0.4140

Table 3.4: Coding Performance of the proposed scheme SPIHT-VQ in terms of Mean Structural Similarity Index Measure (MSSIM)

Lena						
bpp	0.0156	0.0313	0.0625	0.125	0.25	0.5
SPIHT-VQ1	0.4310	0.4757	0.5359	0.5918	0.6406	0.6791
SPIHT-VQ2	0.7374	0.7966	0.8473	0.8907	**0.9202**	0.9300
SPIHT-VQ3	0.6856	0.7380	0.7966	0.8473	0.8907	0.9262
Barbara						
SPIHT-VQ1	0.3212	0.3686	0.4127	0.4577	0.5372	0.6242
SPIHT-VQ2	0.5732	0.6355	0.7350	0.8373	0.9114	0.9423
SPIHT-VQ3	0.5198	0.5732	0.6335	0.7350	0.8373	0.9114
Peppers						
SPIHT-VQ1	0.4093	0.4544	0.4997	0.5411	0.5762	0.6019
SPIHT-VQ2	0.7066	0.7508	0.7910	0.8255	0.8905	0.9125
SPIHT-VQ3	0.6596	0.7066	0.7508	0.7910	0.8225	0.8705
Baboon						
SPIHT-VQ1	0.1783	0.2038	0.2377	0.2956	0.3970	0.4905
SPIHT-VQ2	0.3580	0.4284	0.5521	0.6857	0.7864	0.8875
SPIHT-VQ3	0.3091	0.3580	0.4284	0.5521	0.6857	0.7864

3.35 dB for image Lena, 3 dB for image Barbara, 1.5 dB for image Peppers, and 5 dB for image Baboon. These values of PSNR are small as compared to those mentioned in the tables: Table 2.4 through Table 2.7 in Chapter 2.

Figure 3.9 gives the graphical representation of PSNR with bit rates. It is observed that SPIHT outperforms the proposed scheme for all the test images. But from graphs of Figure 3.10, Figure 3.11 and Figure 3.12 it is clear that our proposed scheme has shown better performance than the existing scheme SPIHT at the same bit rates. In terms of visual quality measures, VIFP, MSSIM and UQI performances are better for our proposed schemes. This indicates that the visual quality improvement is achieved by SPIHT-VQ scheme at lower bit rates although the PSNR performance is poor. PSNR is not a good measure for similarity of images. If there is a shift in image pixels the resulting image has a very poor PSNR. This perceptual quality is better measured by UQI, MSSIM and VIFP which are HVS based metrics. Out of the three SPIHT-VQ schemes proposed, the performance of SPIHT-VQ2 scheme is better than the others.

A bit-rate of 0.25 bpp may be good enough to preserve the required quality for many applications as we can see from the values of MSSIM of 0.9202 and PSNR of 22.84 dB with SPIHT. If we can tolerate a little more distortion, then we may compromise the quality for a bit-rate of 0.125 bpp by which we obtain double the compression.

Table 3.5: Coding Performance of the proposed scheme SPIHT-VQ in terms of Universal Quality Index (UQI)

Lena						
bpp	0.0156	0.0313	0.0625	0.125	0.25	0.5
SPIHT-VQ1	0.1567	0.2156	0.2907	0.3685	0.4380	0.5040
SPIHT-VQ2	0.3928	0.4860	0.5681	0.6494	**0.7412**	0.7563
SPIHT-VQ3	0.3026	0.3985	0.4857	0.5961	0.649	0.7312
Barbara						
SPIHT-VQ1	0.1403	0.1978	0.2547	0.3048	0.3962	0.4960
SPIHT-VQ2	0.3623	0.4376	0.5567	0.6804	0.7864	0.8125
SPIHT-VQ3	0.2865	0.3623	0.4376	0.5567	0.6804	0.7865
Peppers						
SPIHT-VQ1	0.1332	0.1870	0.2487	0.2984	0.3499	0.3938
SPIHT-VQ2	0.3842	0.4518	0.5176	0.5815	0.7009	0.7245
SPIHT-VQ3	0.3024	0.3862	0.4518	0.5179	0.5815	0.7009
Baboon						
SPIHT-VQ1	0.0678	0.1004	0.1386	0.2075	0.2976	0.4145
SPIHT-VQ2	0.2123	0.3084	0.4498	0.6125	0.7864	0.7989
SPIHT-VQ3	0.1487	0.2123	0.3084	0.4498	0.6125	0.7864

We observe that a bit-rate of 0.125 bpp yields slightly less promising quality in terms of PSNR, VIFP, MSSIM and UQI. So, we may draw the following conclusions.

1. SPIHT-VQ2 is the best among all the three proposed schemes.
2. A bit-rate of 0.25 bpp (CR = 32) is good enough to preserve the image details with quite less distortion.
3. A bit-rate of 0.125 bpp (CR = 64) may be accepted if we tolerate little more distortion.

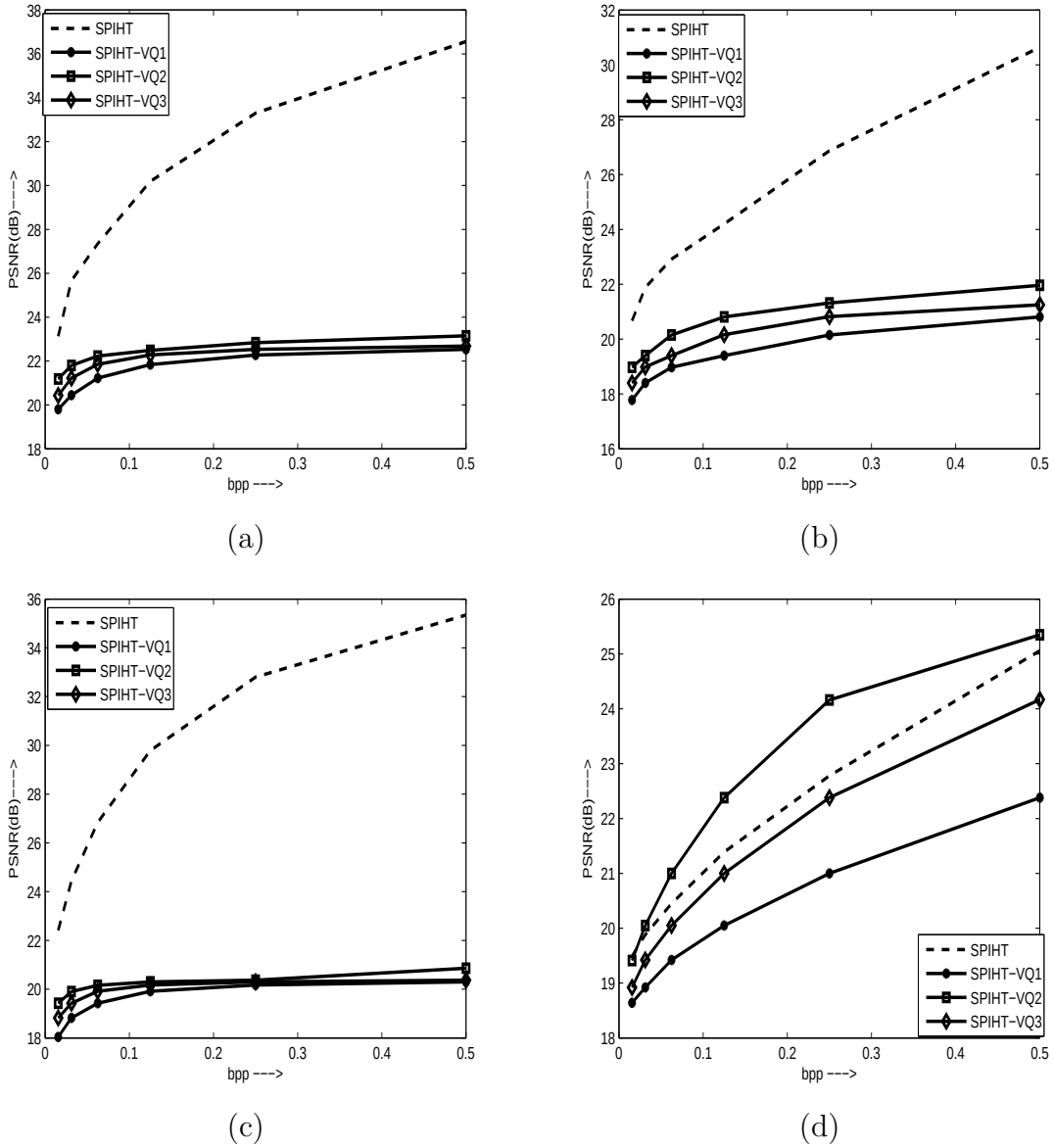


Figure 3.9: Rate distortion performance of the proposed scheme in terms of PSNR(dB)(a)Lena,(b) Barbara,(c) Peppers,(d) Baboon

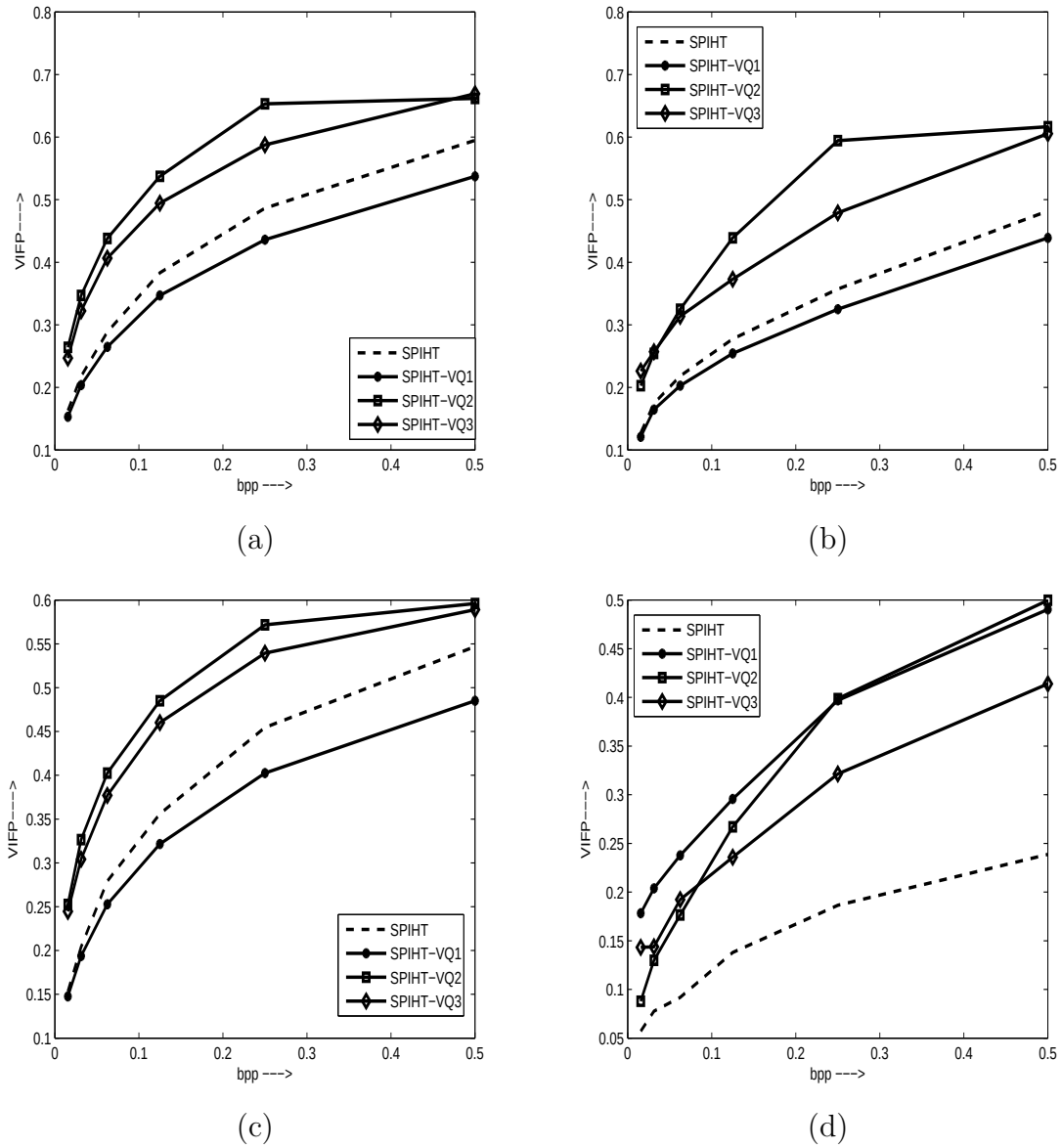


Figure 3.10: Rate distortion performance of the proposed scheme in terms of VIFP(a)Lena,(b) Barbara,(c) Peppers,(d) Baboon

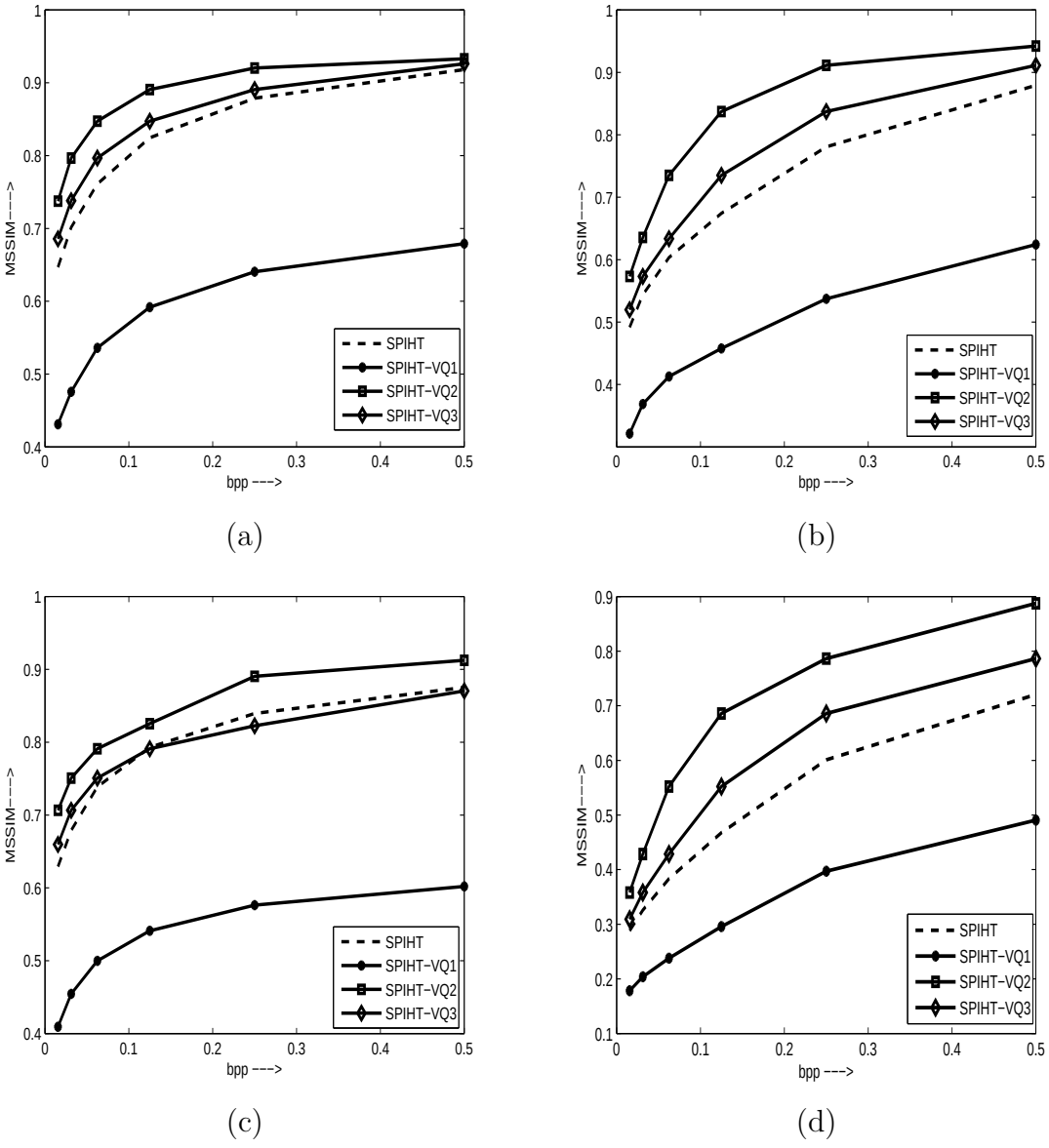


Figure 3.11: Rate distortion performance of the proposed scheme in terms of MSSIM (a)Lena,(b) Barbara,(c) Peppers,(d) Baboon

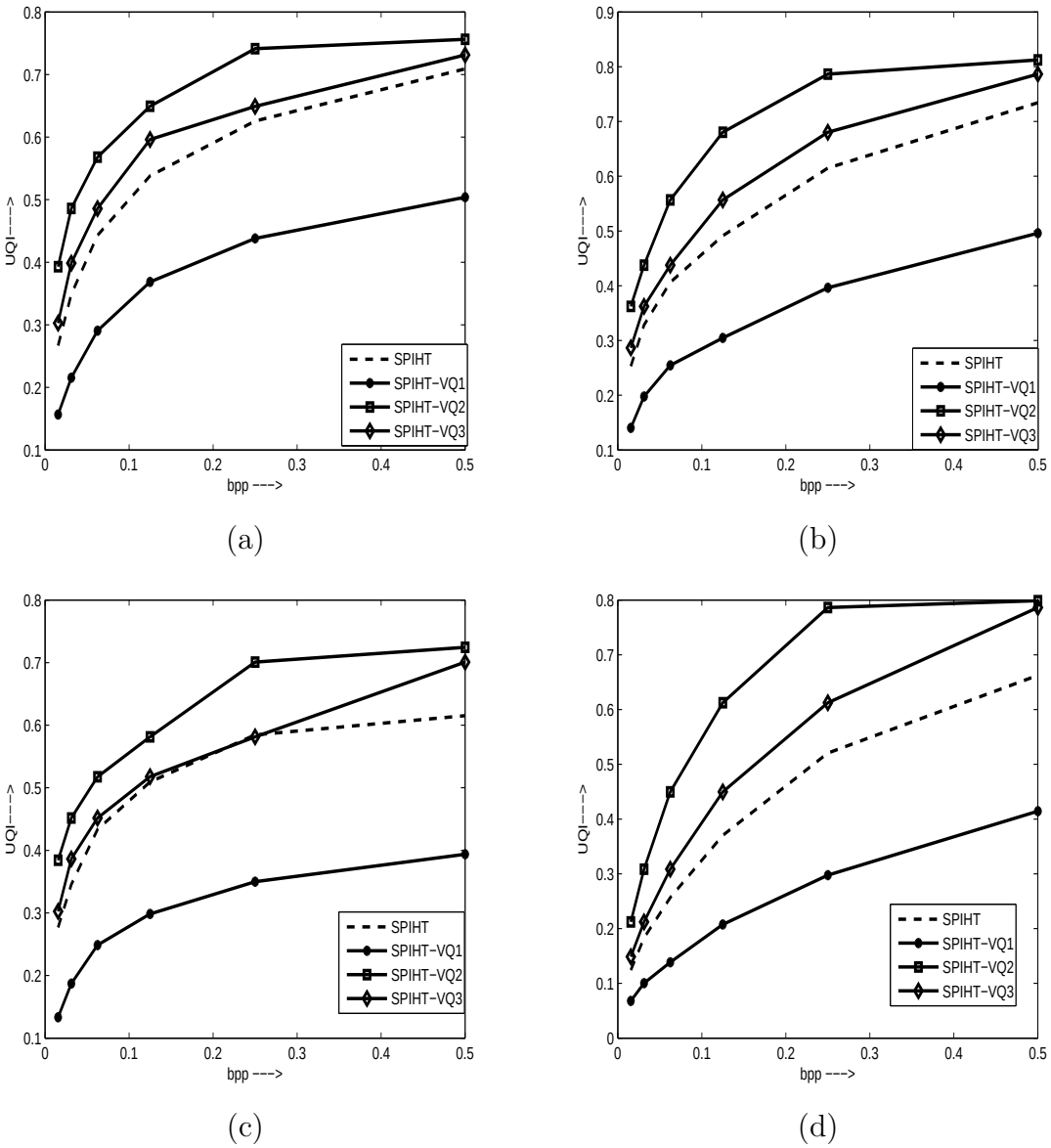


Figure 3.12: Rate distortion performance of the proposed scheme in terms of PSNR(dB)(a)Lena,(b) Barbara,(c) Peppers,(d) Baboon

3.5 Improved Hybrid Coding Scheme for Effectual Image Compression using Human Visual Characteristics [P2]

The application of a human vision system (HVS) model to image compression is a recent research area [128, 129]. From the observations of a person's visual phenomena and the research of physiological and psychological vision, the effect of vision masking has been discovered. The full use of the effect of vision masking in the image coding process permits a bigger actual distortion under the same condition in the subjective sensation distortion. Lower bit-rates may be obtained while maintaining the subjective quality of the image unchanged if HVS is employed in association with an image compression algorithm. The various observations on HVS are:

1. Human vision is very sensitive to information distortion of the image fringe area;

2. Human vision is quite sensitive to information distortion of the image smooth area;

3. Human vision is insensitive to information distortion of the image texture area.

The sensitivity of a person's eye to the area which are fringe area, smooth area and texture area in the image is different. It means that there are differences in the importance of image information among the three kinds of different regions from image coding aspect.
Attempting to compress an image as a whole would not yield optimal results. Hence, the proposed scheme first divides the image into a plurality of blocks of size $n \times n$. This allows the algorithm to exploit the fact that similar regions tend to appear together in small parts of an image. Different sensitivities suggest

that different perceptual weights should be assigned to different blocks of the image [128]. For this purpose, image blocks are classified as edge block, smooth block and textured block on basis of entropy and variance. The entropy value of smooth block is smaller than the edge and textured block. Variance of flat areas should be very low.

Classification of blocks

The image is divided into square blocks. The values of entropy and variance for each block of image is calculated using equations.

Using the calculated values of entropy, calculate minimum and maximum values of the entropy as $e_{min}$and e_{max}, and minimum and maximum values of the entropy as $var_{min}$and var_{max}

Now calculate

$median_{entropy} = (e_{max} - e_{min})/2$

if

$median_{entropy} < individualentropyvalue$

the block is SMOOTH BLOCK.

Otherwise

$calculatemedian_{variance} = (var_{max} - var_{min})/2$

if

$individualvariance < median_{variance}$

the block is EDGE BLOCK.

elseif

$individualvariance > median_{variance}$

the block is TEXTURE BLOCK.

The block diagram of the proposed HVS based SPIHT-VQ (HSPIHT-VQ) scheme is shown in Figure 3.13. For incorporating HVS, the changes are there only in the encoder side. The decoder remains same as shown in Figure 3.2.

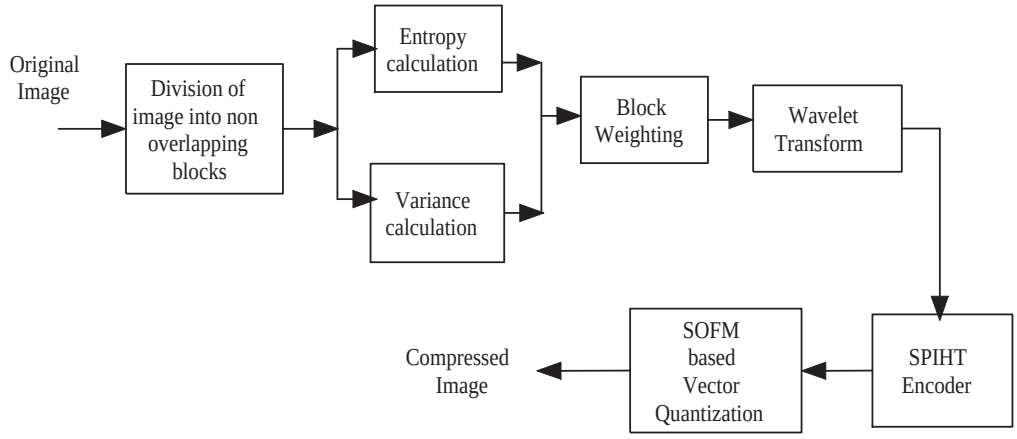


Figure 3.13: HVS based SPIHT-VQ image coder

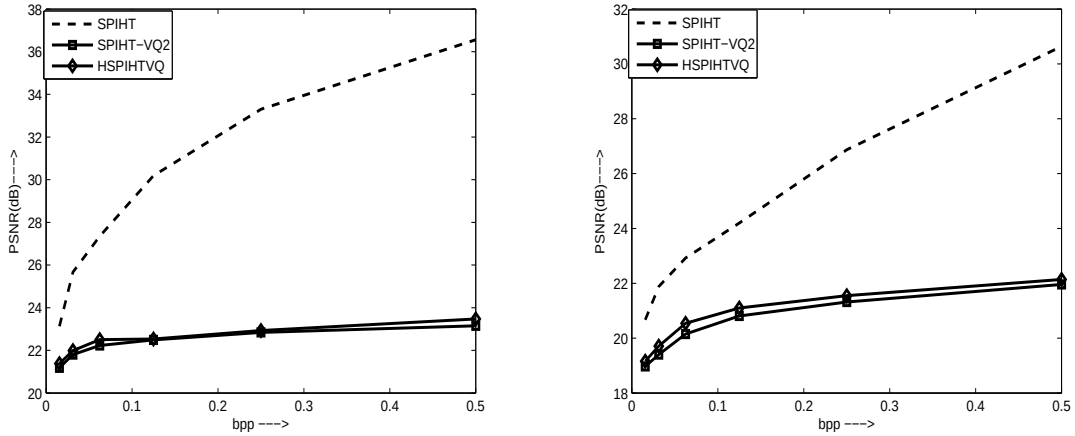
Table 3.6: Performance of HVS based SPIHT-VQ (HSPIHT-VQ) using SPIHT-VQ2 for Lena image

bpp	PSNR	VIFP	MSSIM	UQI
0.0156	21.38	0.2666	0.7388	0.3058
0.0313	21.99	0.4415	0.7989	0.4001
0.0625	22.50	0.4477	0.8497	0.5047
0.125	22.53	0.5952	0.8927	0.6018
0.25	**22.93**	**0.6555**	**0.9249**	**0.6535**
0.5	23.47	0.6641	0.9361	0.7368

Table 3.7: Performance of HVS based SPIHT-VQ (HSPIHT-VQ) using SPIHT-VQ2 for Barbara image

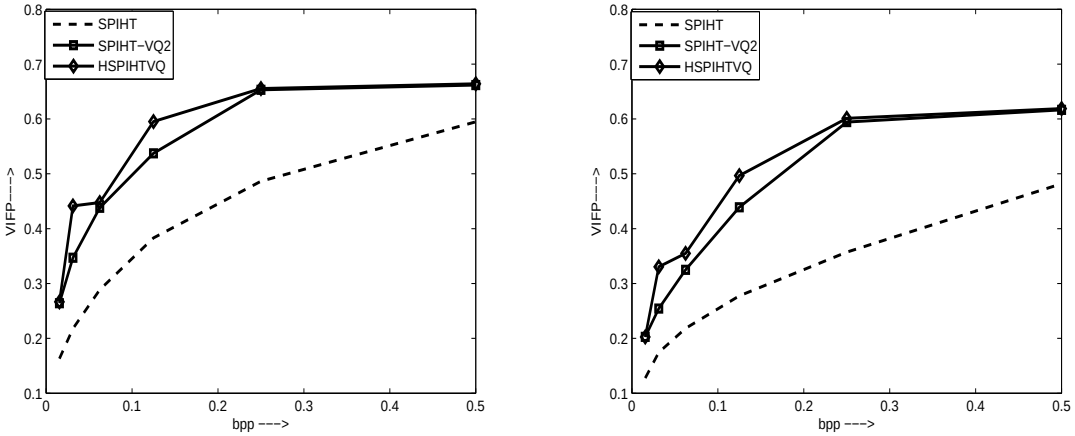
bpp	PSNR	VIFP	MSSIM	UQI
0.0156	19.16	0.2052	0.5764	0.2913
0.0313	19.71	0.3304	0.6371	0.3464
0.0625	20.54	0.3550	0.7540	0.4400
0.125	21.10	0.4967	0.8430	0.5587
0.25	**21.55**	**0.6011**	**0.9159**	**0.6851**
0.5	22.14	0.6188	0.9479	0.7926

To evaluate the performance of the proposed HVS based hybrid image coding algorithm, experiments are conducted on Lena and Barbara images. Lena image is a smooth image while Barbara image a is texture-image. It is observed from previ-



(a) (b)

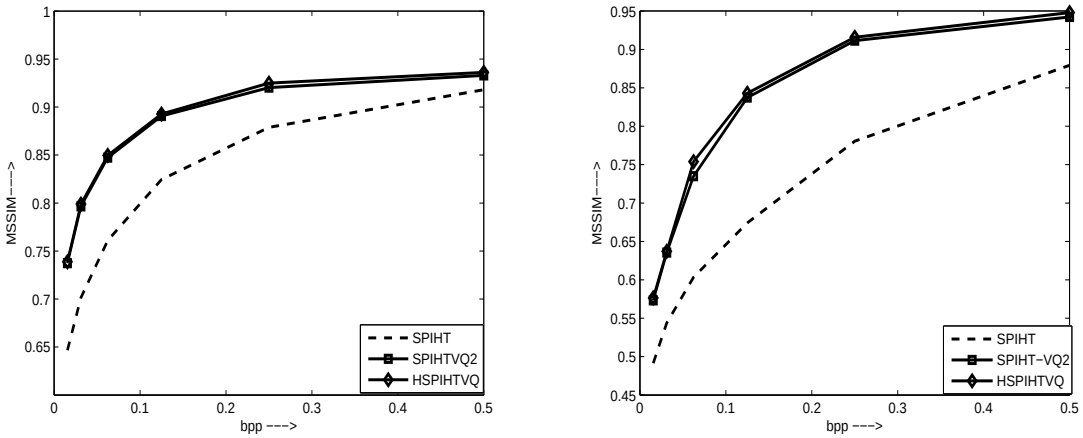
Figure 3.14: Rate distortion performance of the scheme HSPIHT-VQ in terms of PSNR (dB)(a)Lena (b) Barbara



(a) (b)

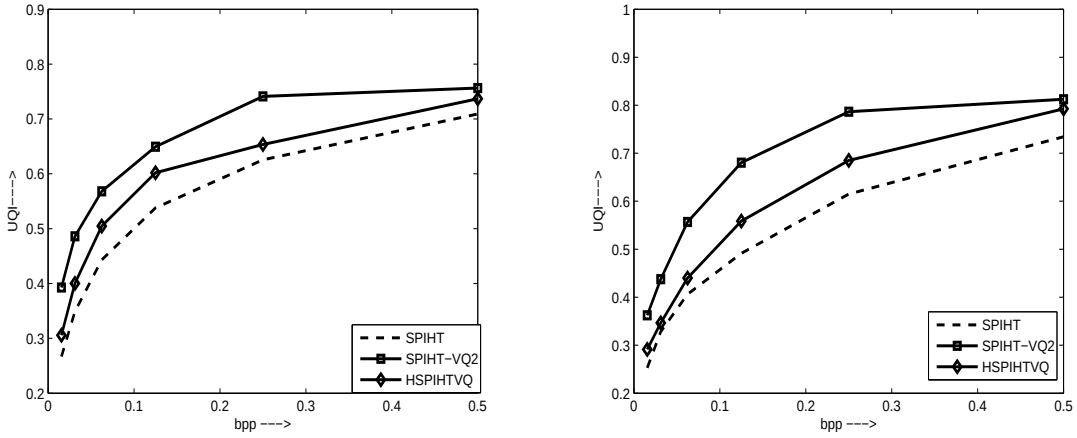
Figure 3.15: Rate distortion performance of the scheme HSPIHT-VQ in terms of VIFP (a)Lena (b) Barbara

ous section that SPIHT-VQ2 performs better than SPIHT-VQ1 and SPIHT-VQ3. Hence SPIHT-VQ2 is used for evaluating the performance. For fair comparison, the bit-rates of SPIHT-VQ2 and HSPIHT-VQ are kept same. Table 3.6 and Table 3.7 summarize the performance of the proposed scheme HSPIHT-VQ. The



(a) (b)

Figure 3.16: Rate distortion performance of the scheme HSPIHT-VQ in terms of MSSIM (a)Lena (b) Barbara



(a) (b)

Figure 3.17: Rate distortion performance of the scheme HSPIHT-VQ in terms of UQI (a)Lena (b) Barbara

graphical comparison is done in Figure 3.14 through Figure 3.17. A marginal improvement is observed in performance of HSPIHT-VQ over SPIHT-VQ in terms of the quality metrics: PSNR,VIFP, and MSSIM. But it shows a significant improvement in terms of UQI.

Thus, it may be concluded that there is considerable improvement in image quality if HVS is combined with SPIHT-VQ.

3.6 Conclusion

In this chapter, a hybrid image compression scheme based on SPIHT and SOFM based vector quantization is developed and analyzed. The SPIHT coder is cascaded with vector quantizer. The output of SPIHT coder is further compressed by the vector quantizer. The vector quantizer uses a generic code book. The effectiveness of the scheme at low bit-rate is demonstrated using image quality parameters set specifically for yielding low bit-rates. The visual quality of the images is improved by incorporating human visual characteristics in the proposed compression scheme.

Chapter 4

Development of Hybrid Image Compression Scheme using DCT and Fractals

Preview

Image Transform methods using orthogonal kernel functions are commonly used in image compression. One of the most widely known image transform methods is DCT, which is employed in JPEG compression standard. The computing devices such as personal digital assistants, digital cameras and mobile phones require a lot of image transmission and processing. Though various efficient compression techniques have been reported, the wide range of multimedia applications demands for further improvement in compression quality. In this chapter, a simple and efficient DCT block based algorithm is proposed. The proposed algorithm reduces the bits in the compressed bitstream while maintaining the rate-distortion performance of the image.

4.1 Introduction

Two dimensional image transforms render the image data in the form that is easily compressible. DCT based JPEG and DWT based JPEG2000 are well known existing standards. Research still continues to get more compression without much degradation in the performance. Although wavelets are capable of more flexible space-frequency resolution trade off's than DCT, DCT is still widely used in many practical applications because of its compression performance and computational advantages. In this chapter, a modified block DCT based compression scheme incorporating the idea of self similarity borrowed from fractals is proposed.

4.2 The proposed DCT-Fractal (DCT-F) Scheme [P4]

4.2.1 The Compression Process

A novel image compression scheme is developed here to compress the images combining DCT and the idea of fractal image compression. The proposed scheme divides the input image $f(x, y)$ into sub-images $g(x, y)$. Each sub-image $g(x, y)$ is further tiled into blocks of size 8×8. The 8×8 block of data is transformed using two dimensional DCT and the transformed values are quantized using standard JPEG quantization table.

The entire quantized coefficients of the 8×8 block are rearranged in a zig-zag manner as shown in Figure 4.1.

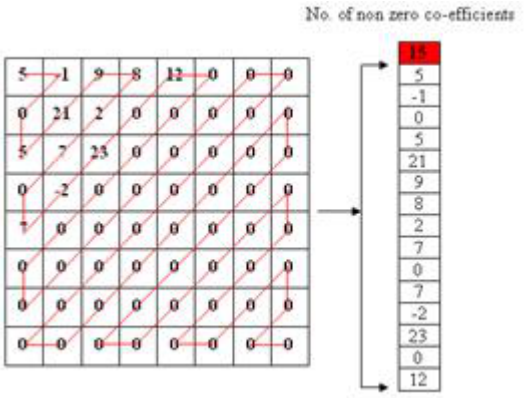


Figure 4.1: Zig-Zag Arrangement of Coefficients

Most of the high frequency coefficients (lower right corner) become zeros after quantization. A zig-zag scan of the matrix yielding long strings of zeros is used to exploit the number of zeros. The current coder acts as filter and passes only the string of non-zero coefficients. A list of non-zero coefficients of the blocks in the order of their count will be obtained at the end of this process. It is repeated for all the blocks of sub-image, $g(x, y)$.

The concept of self similarity is used to prevent performing repetitive compression on the sub-image $g(x, y)$. Euclidean distance is measured between the central (range) block and the other blocks in sub image $g(x, y)$ to determine the similarity.

If $g_1(x_1, y1)$ and $g_2(x_2, y_2)$ are blocks of sub-image $g(x, y)$ then Euclidean distance is given by

$$d(g_1, g_2) = \sqrt{(y_1 - x_1)^2 + (y_2 - x_2)^2} \tag{4.1}$$

The proposed method calculates the similarity of the $b_{m,n}$ block for the fractal image compression by comparing the distance measure of the $b_{r,s}$ block with that of a (range block) block as depicted in Figure 4.2

where

$r = m - 1 : m + 1,$

$s = n - 1 : n + 1,$

and $(r, s) \neq (m, n)$

A flag is set according to the threshold, D_{tsh}, after the distance measure is cal-

$b_{m-1,n-1}$	$b_{m,n-1}$	$b_{m+1,n-1}$
$b_{m-1,n}$	$b_{m,n}$	$b_{m+1,n}$
$b_{m-1,n+1}$	$b_{m,n+1}$	$b_{m+1,n+1}$

Figure 4.2: Range Blocks $b_{m,n}$ and neighboring blocks in sub-image $g(x, y)$

culated. The calculated distance D is compared with the threshold, D_{tsh}, by

$$I_F^b = \begin{cases} 1, & D < D_{tsh} \\ 0, & otherwise \end{cases} \tag{4.2}$$

where b and F represent the image block and the flag value of each block of the image, respectively. Both $b_{r,s}$ and $b_{m,n}$ blocks are said to be similar if the block yields a flag value of 1, when it is compared with the $b_{m,n}$ block. Otherwise, the blocks are said to be dissimilar. This is illustrated in Figure 4.3

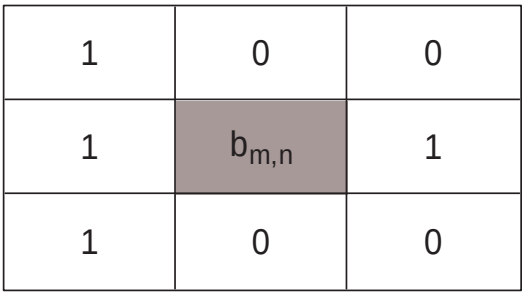


Figure 4.3: Flags assigned to neighboring blocks(Flag value of 1 corresponds to domain block)

The indices of the blocks similar to the $b_{m,n}$ block are stored. In fractal image compression, $b_{m,n}$ block is the range block and the analogous similar blocks are domain blocks. Instead of all the similar domain blocks, only the range block is used once the indices of range block and its corresponding domain blocks are accumulated. The coefficients of the range(central) block are compressed and sent to encoder. The indices of the blocks that satisfy the similarity measure are sent along with the central block. Other non similar blocks are compressed separately and sent to the encoder. The time and memory requirements for the overall compression process are decreased due to this.

All the steps mentioned above are repeated for all the sub-images obtained in the first step. The first array consists of information of sub-image number and the non-zero pixels of the central block. The second array consists of the information of sub-image number and the index of blocks that are similar to range block in same sub-image. The third array consists of indices of non similar blocks with their non zero pixel values in DCT domain. All the arrays are concatenated to obtain a single array. Huffman encoding is applied to the composite array to obtain further compression. Thus, an array of compressed image data $\phi(x, y)$ is obtained.

4.2.2 Decompression Process

The compressed data stream $\phi(x, y)$ is first passed through a Huffman decoder to remove lossless encoding. From the data in decompressed array, the data corresponding to the first sub-image is extracted. Using the indices and values of range blocks the data is arranged in zigzag form for 8×8 blocks. The 8×8 blocks are dequantized using the same quantization table. Inverse 8×8 DCT is taken on each block. This procedure is repeated for all sub-images and the decompressed image is obtained.

4.3 Coding Performance of DCT-F Scheme

The coding performance of the DCT-F scheme is evaluated on standard test images: Lena and Barbara. The bit-rate for BDCT based compression scheme is varied by varying the scale factor. The rate-distortion performance of BDCT scheme is evaluated by varying the scale factor from 1 to 7 for all the test images and is summarised in Table 4.1.

Table 4.1: Rate distortion performance of BDCT for various scale factors for test image, Lena

Scale Factor	bpp	PSNR(dB)	VIFP	MSSIM	UQI
1	0.9426	42.93	0.9810	0.9182	0.9992
2	0.6139	36.25	0.8928	0.8841	0.9998
3	0.4870	35.70	0.8070	0.8549	0.9996
4	0.4024	30.76	0.7261	0.8274	0.9976
5	0.3463	27.28	0.6524	0.8037	0.9880
6	0.3071	26.51	0.5997	0.7844	0.9856
7	0.2790	26.13	0.5560	0.7635	0.9830

For the proposed scheme, DCT-F, the bit-rate can be varied by changing the scale factor and the distance threshold D_{tsh}. For the test image, Lena, tables:

Table 4.2: Rate distortion performance of BDCT for various scale factors for test image, Barbara

Scale Factor	bpp	PSNR(dB)	VIFP	MSSIM	UQI
1	1.3075	34.94	0.9634	0.9287	0.9998
2	0.9128	25.29	0.8431	0.8608	0.9646
3	0.7215	25.08	0.7714	0.8113	0.9492
4	0.5973	25.18	0.7138	0.7753	0.9512
5	0.5158	25.04	0.6541	0.7324	0.9357
6	0.4468	25.02	0.6121	0.7011	0.9394
7	0.3951	24.81	0.5571	0.6543	0.8974

Table 4.3: Rate distortion performance of DCT-F scheme for various scale factors using $D_{tsh} = 3$ and sub-image size of 3×3 for test image, Lena

Scale Factor	bpp	PSNR	VIFP	MSSIM	UQI
1	0.9076	42.88	0.9733	0.9154	0.9992
2	0.5807	35.80	0.8630	0.8737	0.9997
3	0.4475	35.01	0.8356	0.7621	0.9993
4	0.3651	30.43	0.6719	0.8050	0.9971
5	0.3117	28.48	0.6018	0.7784	0.9869
6	0.2724	27.18	0.5436	0.7529	0.9893
7	0.2418	26.66	0.4935	0.7277	0.9926

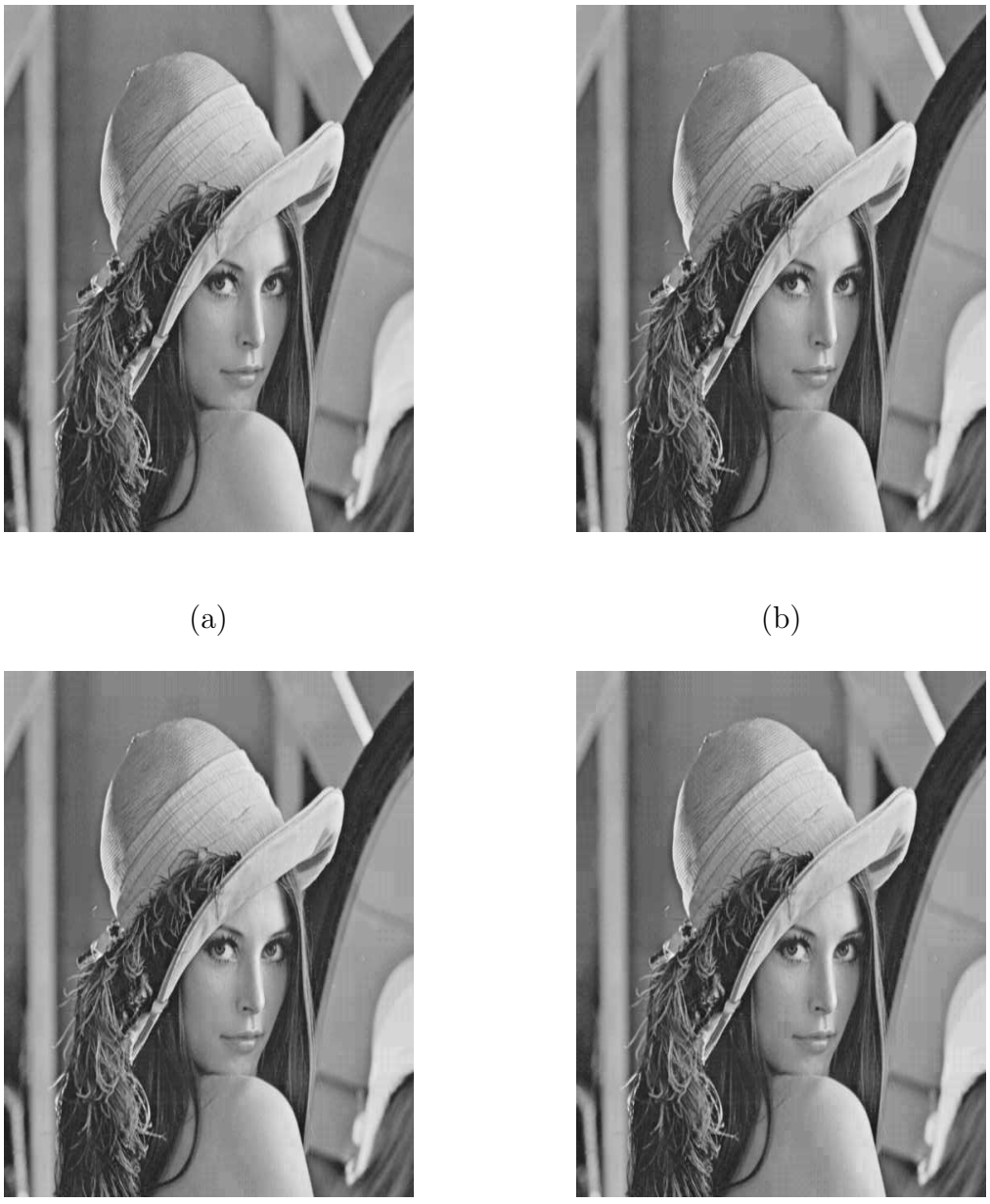
Table 4.3 through Table 4.5 illustrate the performance of DCT-F scheme for various scale factors and $D_{tsh} = 3$ when sub-image size is 3×3. It is observed by comparing Table 4.1 with Table 4.3 that an additional compression of 4% and 10% is achieved at scale factors of 1 and 7 respectively. The cost of achieving additional compression of 4% is very marginal as compared to that of achieving 10% additional compression.

(a) (b)

Figure 4.4: Visual quality comparison of decompressed Lena image using scale factors at $D_{tsh} = 3$ (a) scale factor=1, bpp=0.9076, (b) scale factor=7, bpp=0.2418

Table 4.4: Rate distortion performance of DCT-F scheme for various scale factors using $D_{tsh} = 5$ and sub-image size of 3×3 for test image, Lena

Scale Factor	bpp	PSNR(dB)	VIFP	MSSIM	UQI
1	0.8857	42.76	0.9640	0.9118	0.9992
2	0.5807	35.80	0.8630	0.8737	0.9997
3	0.4475	35.01	0.8356	0.8356	0.9993
4	0.3394	30.07	0.6228	0.7813	0.9964
5	0.2847	28.13	0.5455	0.7513	0.9849
6	0.2473	27.15	0.4948	0.7270	0.9878
7	0.2187	26.44	0.4378	0.6959	0.9877



(a) (b)

(c) (d)

Figure 4.5: Visual quality comparison of decompressed Lena image using BDCT and DCT-F at scale factor = 1 with $D_{tsh} = 3$ (a) bpp = 0.9426,(b) bpp = 0.9076, (c) bpp = 0.8857, (d) bpp = 0.8674

Table 4.5: Rate distortion performance of DCT-F scheme for various scale factors using $D_{tsh} = 7$ and sub-image size of 3×3 for test image, Lena

Scale Factor	bpp	PSNR(dB)	VIFP	MSSIM	UQI
1	0.8674	42.57	0.9528	0.9080	0.9992
2	0.5352	34.56	0.8032	0.8516	0.9992
3	0.4061	33.70	0.6917	0.8024	0.9985
4	0.3232	29.88	0.5907	0.7644	0.9955
5	0.2686	28.08	0.5126	0.7334	0.9835
6	0.2303	27.09	0.4578	0.7068	0.9859
7	0.2037	26.40	0.4016	0.6788	0.9880

Table 4.6: Rate distortion performance of DCT-F scheme for various scale factors using $D_{tsh} = 3$ and sub-image size of 3×3 for test image, Barbara

Scale Factor	bpp	PSNR(dB)	VIFP	MSSIM	UQI
1	1.2928	34.92	0.9634	0.9275	0.9998
2	0.8851	25.30	0.8268	0.8546	0.9646
3	0.6887	25.09	0.7443	0.8014	0.9491
4	0.5619	25.20	0.6824	0.7623	0.9510
5	0.4685	25.07	0.6116	0.7103	0.9359
6	0.3915	25.16	0.5591	0.6702	0.9379
7	0.3322	24.86	0.6702	0.9379	0.8967

Table 4.7: Rate distortion performance of DCT-F scheme for various scale factors using $D_{tsh} = 5$ and sub-image size of 3×3 for test image, Barbara

Scale Factor	bpp	PSNR(dB)	VIFP	MSSIM	UQI
1	1.2828	34.90	0.9600	0.9262	0.9998
2	0.8646	25.31	0.8108	0.8486	0.9646
3	0.6642	25.12	0.7178	0.7897	0.9494
4	0.5297	25.24	0.6462	0.7407	0.9509
5	0.4329	25.11	0.5709	0.6840	0.9359
6	0.3528	25.22	0.5062	0.6387	0.9377
7	0.2950	24.90	0.4436	0.5852	0.8967

Table 4.8: Rate distortion performance of DCT-F scheme for various scale factors using $D_{tsh} = 5$ and sub-image size of 3×3 for test image, Barbara

Scale Factor	bpp	PSNR(dB)	VIFP	MSSIM	UQI
1	1.2684	34.84	0.9539	0.9233	0.9997
2	0.8474	25.33	0.7948	0.8411	0.9646
3	0.6401	25.138	0.6960	0.7789	0.9493
4	0.5002	25.27	0.6117	0.7205	0.9503
5	0.3939	25.55	0.5341	0.6651	0.9584
6	0.3120	25.29	0.4558	0.6053	0.9371
7	0.3120	25.29	0.4558	0.6053	0.9371

The blocking artifacts are very clearly visible at higher compression ratios (i.e. scale factor of 7) as seen from Figure 4.5. For the value of $D_{tsh} = 5$ and $D_{tsh} = 7$ reduction in bit-rate of 6% and 7.9% is observed at scale factor of 1 from Table 4.4 and Table 4.5. Figure 4.5 gives the visual quality comparison of decompressed images using BDCT and DCT-F.

Table 4.2 and Table 4.6 through Table 4.8 depict the rate-distortion performances of BDCT and DCT-F scheme respectively for test image Barbara. It is observed that there is very less reduction in bit-rates as compared to Lena image. Also, as the bit-rate drops below 1 bpp, the PSNR performance is almost constant for all the values of threshold and scale factors. In that case, the reduction in decompressed image quality is to be judged from the other performance parameters: VIFP and MSSIM which show a numerical reduction with bit-rate reduction.

So, it may be concluded that BDCT and our proposed scheme, DCT-F perform well at low bit rates for low-complexity images like Lena. But, their performance is not so promising for high complexity and texture images like Barbara.

4.4 Conclusion

A hybrid image compression scheme based on DCT and fractals, DCT-F, has been developed and analysed in this chapter. The scheme aims at providing extra compression to the image with minimal loss in quality in terms of distortion metrics. For this purpose, the idea of self similarity used in fractal image compression has been incorporated with BDCT. The self similarities between the analogous blocks is found by using the Euclidean distance measure. This eliminates the continual compression of analogous blocks and provides extra compression. This scheme gives promising results with a slight extra compression.

Chapter 5

Hybrid Image Compression based on CALIC and Spatial Prediction Structures

Preview

Spatial prediction in images is a key component for efficient lossless image compression [1, 2]. In general the lossless image compression schemes attempt to predict image data using the the pixels in the spatial neighbourhood. The efficiency of image compression reduces by this process. Lossless schemes based on predictive coding are widely used from early work on television signal coding to modern lossless image compression schemes. It uses the idea of using previous pixel to predict current pixel value. The basic concept is to use the previous data to predict current pixel and then only prediction error should be encoded. Various forms of predictive coding like differential pulse code modulation (DPCM), difference mapping, etc. are available in the literature [3]. Lossless compression limits the compression ratio between one and three, while lossy compression with some degradation in quality yields compression ratios greater than twenty. Therefore, there is a need to have a compression technique, which not only compresses well, but also retains the quality acceptable for certain applications.

5.1 Introduction

Structure components such as edges, contours, and textures are found in natural images. These components repeat themselves at various locations and scales. Therefore, an image prediction scheme that exploits this type of image correlation in spatial domain can be developed. This type of image correlation has been exploited in spatial prediction structures. The spatial structure prediction algorithm [146] breaks the neighborhood constraint, attempting to find an optimal prediction of structure components from the previously encoded image regions. It borrows the idea of motion prediction from video coding, which predicts a block in the current frame using its previous encoded frames. To improve the compression, the image is classified into two types of regions: namely the *structure*

regions and the *non-structure regions*. Non-structure regions are smooth image areas that can be efficiently represented [149] with the spatial transforms, such as KLT (Karhunen Loeve transform), DCT and DWT [1, 2, 7, 10]. The structure regions, on the other hand, consist of high-frequency components and curvilinear features in images, such as edges, contours, and texture regions, which cannot be efficiently represented by the linear spatial transforms. The structured regions are hard to compress and consume a majority of the total encoding bits. The structure regions are encoded with spatial prediction structures while non structure regions can be efficiently encoded with conventional image compression method, CALIC [7, 9, 10]. There is no codebook requirement in the compression scheme, since the best matches of structure components are simply searched within encoded image regions. In this work, an efficient image compression scheme based on spatial prediction of structural units is discussed and analyzed. Extensive experimental results demonstrate that the scheme is very competitive and even outperforms the state-of-the-art image compression methods.

5.2 Spatial Prediction Structures [P5]

The idea of spatial prediction structure comes from motion prediction used in video coding [6]. In motion prediction, (see Figure 5.1) an area in the reference frame is searched to find the best match of the current block based on some distortion metric. The chosen reference block becomes the predictor of the current block. The prediction residual and the motion vector are then encoded and sent to the decoder.

In spatial prediction structures, regions within the previously encoded image regions are searched to find the prediction of an image block as shown in Figure 5.2. The reference block that results in the minimum block difference in terms of SAD is selected as the optimal prediction. Figure 5.3 shows the block diagram of the hybrid image compression scheme based on spatial prediction of structural

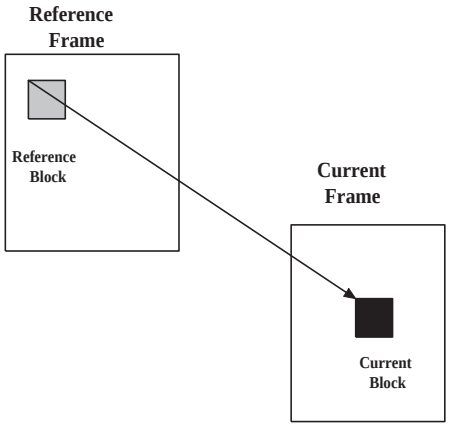


Figure 5.1: Motion Prediction used in Video Coding

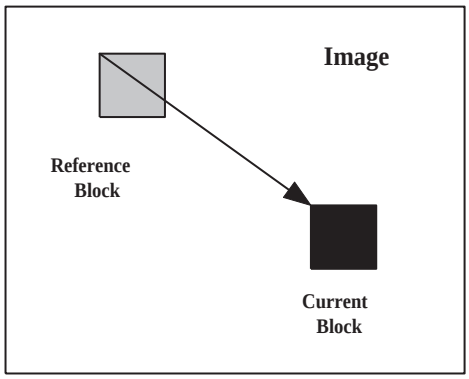


Figure 5.2: Spatial Prediction Structure

components. First, we classify the input data into two categories: structural regions and non-structural regions, where structure regions consist of high frequency regions, and non structure regions consists of smooth areas. The upper output of CALIC to the combiner represents the bitstream corresponding to non structure regions. The process of classification, employed here, is described in detail in the following section.

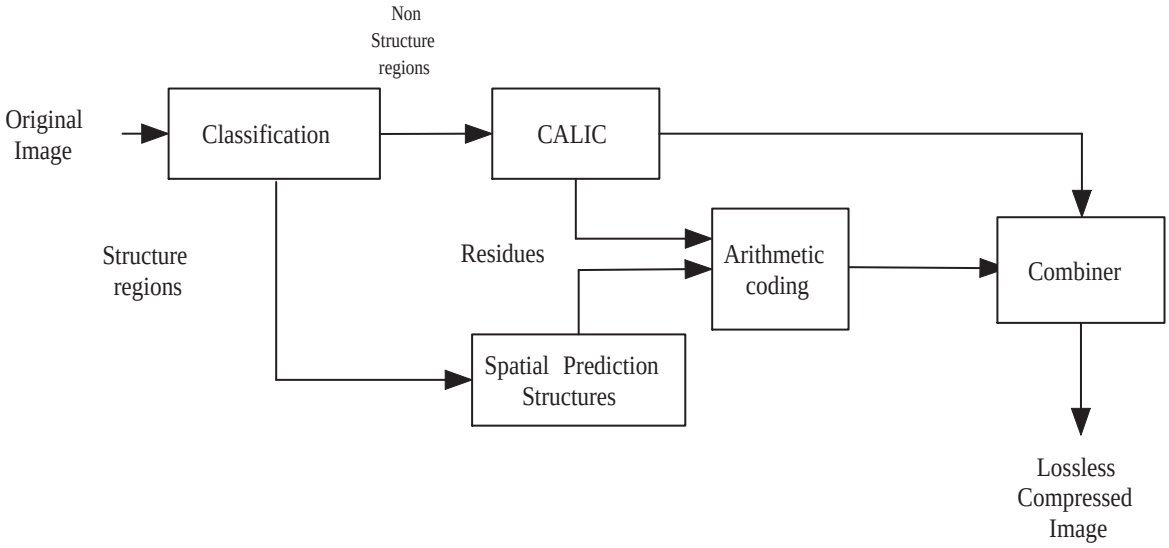


Figure 5.3: The hybrid image compression scheme CALIC-SPS

5.2.1 Classification of Structural Components

The size of the prediction unit is an important parameter in the structure prediction scheme. When the unit size is small, the amount of prediction and coding overhead will become very large. However, if a larger prediction unit is used, the overall prediction efficiency decreases. In this work, an attempt to find a good tradeoff between these two is made and spatial image prediction is performed on block basis. Here a block of size 4×4 is chosen.

Estimation of Threshold

A threshold is required while comparing the current block with the previous encoded block region. This threshold value should be so decided that it will give best compression performance. For classification purpose every 4×4 block of original image is compared with corresponding 4×4 block in theCALIC compressed image using SAD. The SAD values are arranged in descending order. The threshold is chosen based on the percentage of blocks with top SAD values that are to be encoded with spatial prediction structures. For example, if 50% of blocks are to be encoded as structure regions, the top 50% blocks with highest SAD values are

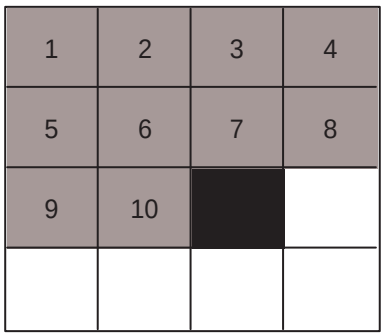


Figure 5.4: Direct prediction method for optimal prediction

chosen.

5.2.2 OptimalPrediction of Structural Components

In the hybrid scheme- CALIC and Spatial Prediction Structures(CALIC-SPS), after the image is classified into structure and non-structure regions, the structure region is encoded using spatial prediction structure. The prediction is carried out in two different methods, namely, *direct prediction* and *intra-prediction* [33, 95].

Direct prediction

In direct prediction method, the current image block is predicted from previously encoded image region and the block that results in minimum block difference in terms of SAD is selected as the optimal prediction. Figure 5.4 shows the direct prediction method used, in which, black color block represents the current block and gray blocks are the previously encoded blocks. The current block will be compared with all other blocks numbered from 1 to 10 only and the block that results in minimum value of SAD is chosen.

Intra-prediction

In addition to direct prediction, additional prediction modes are also used. In this prediction method, using 4×4 blocks, nine modes of prediction are present [76].

A 4×4 block of pixels labeled a through p are predicted from a row of eight

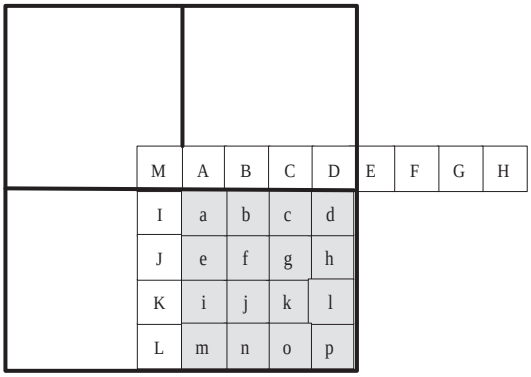


Figure 5.5: Relationship of 16 pixels labelled a through p of current 4×4 block

pixels labeled A through H above the current block and a column of four pixels labeled I through L to the left of the current block as well as a corner pixel labeled M, as shown in Figure 5.5. The nine modes of 4×4 intra-prediction are

1. Mode 0 (vertical prediction)
2. Mode 1 (horizontal prediction)
3. Mode 2 (DC prediction)
4. Mode 3 (diagonal down/left prediction)
5. Mode 4 (diagonal down/right prediction)
6. Mode 5 (vertical-right prediction)
7. Mode 6 (horizontal-down prediction)
8. Mode 7 (vertical-left prediction)
9. Mode 8 (horizontal-up prediction)

These nine modes are shown in Figure 5.6 to its neighboring row of pixel on the top and column of pixels to the left of the current block

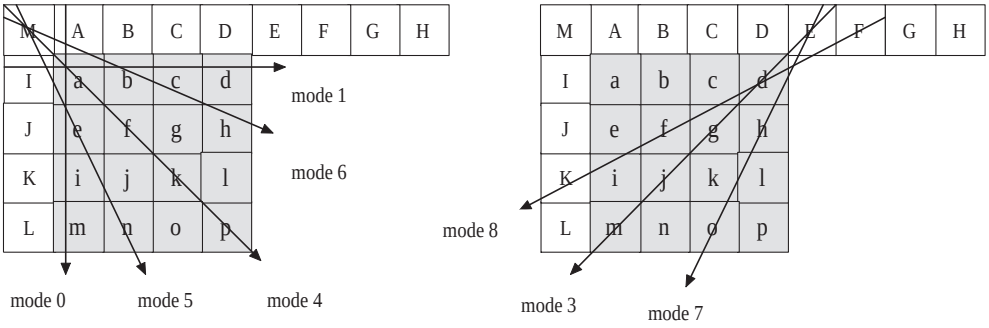


Figure 5.6: Relationship of 16 pixels labelled a through p of current 4×4 block

5.2.3 Residue Encoding

It is required to send the residues along with the compressed image for the scheme to be lossless. This will increase the payload size and thereby the compression ratio decreases. The residues are encountered in two places: The output of CALIC and the output of spatial prediction structures. Arithmetic coding scheme [10] is used to transmit the residues to further reduce the size of the overhead data per block.

5.3 Coding Performance of CALIC-SPS

The coding scheme is evaluated for compression performance for both the gray-scale and color images. Bit-rate is used as metric to specify the compression performance. The test images used for simulation are of size 512×512, with tonal resolution of 8 bits per pixel. The bit-rate has been calculated for all test images when only CALIC is applied and when the CALIC is applied in combination with spatial prediction structures, using both direct prediction mode and intra-prediction modes.

The compression performance of spatial prediction structures prediction with CALIC is calculated for various test images using both direct prediction mode and intra-prediction modes. The bit-rate for each image is tabulated in Table 5.1

Table 5.1: Compression Performance comparison of CALIC-SPS using Direct Prediction Mode with CALIC in bpp

Image	bpp		bit-rate saving, Δbpp
	CALIC	CALIC-SPS	
Lena	3.1509	3.1231	0.0278
Barbara	3.4804	2.9729	0.5075
Baboon	4.2813	3.4811	0.8002
Boat	3.5804	3.3528	0.2276
Goldhill	3.4663	3.3419	0.1244
Peppers	3.3524	3.3186	0.0338

and Table 5.2. As seen from Table 5.1, the bit-rate savings is more for Baboon image when compared to other images. This is because the Baboon image has lots of structural regions when compared to others. Thus, images that contain large number of structure region are compressed more and hence yield reduced bit-rate. The bit-rate of 3.4811 bpp is observed with CALIC-SPS for the test image, Baboon.

Table 5.2: Compression Performance comparison of CALIC-SPS using intra-prediction Mode with CALIC in bpp

Image	bpp		bit-rate saving Δbpp
	CALIC	CALIC-SPS	
Lena	3.1509	3.0711	0.0798
Barbara	3.4804	2.3218	1.1586
Baboon	4.2813	2.9738	1.3705
Boat	3.5804	3.0941	0.4863
Goldhill	3.4663	2.4866	0.9797
Peppers	3.3524	2.9746	0.3778

As observed in Table 5.2, where the intra-prediction modes are applied, the bit-rate savings are more for the test image Baboon image compared to others. Thus, optimal prediction of image blocks is better than the standard lossless compression algorithm, CALIC. Next, the bit-rate saving is more for Barbara. Table 5.3 shows that by using different optimal prediction methods, CALIC-SPS

with intra-prediction modes gives better saving over bits per pixel when compared to CALIC-SPS using direct prediction mode where the image blocks are directly predicted from previous encoded image regions.

Table 5.3: Comparison of bit-rate using different modes of prediction in bpp

Image	bpp		bit-rate saving, Δbpp
	CALIC-SPS Direct prediction	CALIC-SPS Intra-prediction	
Lena	3.1231	3.0711	0.052
Barbara	2.9729	2.3218	0.6511
Baboon	3.4811	3.2860	0.2425
Boat	3.3528	3.0941	0.2587
Goldhill	3.3419	2.4866	0.8533
Peppers	3.3186	2.9746	0.3440

The bit-rate is varied with respect to percentage of structural regions and the results are shown graphically for each test image. From Figure 5.7(a), for Lena image the variation of bit rate graph is more between 50% to 65% and 80% to 90% of structure region and it is more steep in these regions. With the increase in structure regions, the compression is more and the bit-rate keeps on decreasing. The best case result value from this graph is tabulated in Table 5.2 which is comparable to CALIC.

Similarly, the plot in Figure 5.7 (b) depicts the decrease in bit rate for Barbara image with the variation of structure region. The bit-rate has decreased almost to 3 bpp between 50% to 65% of structure regions.

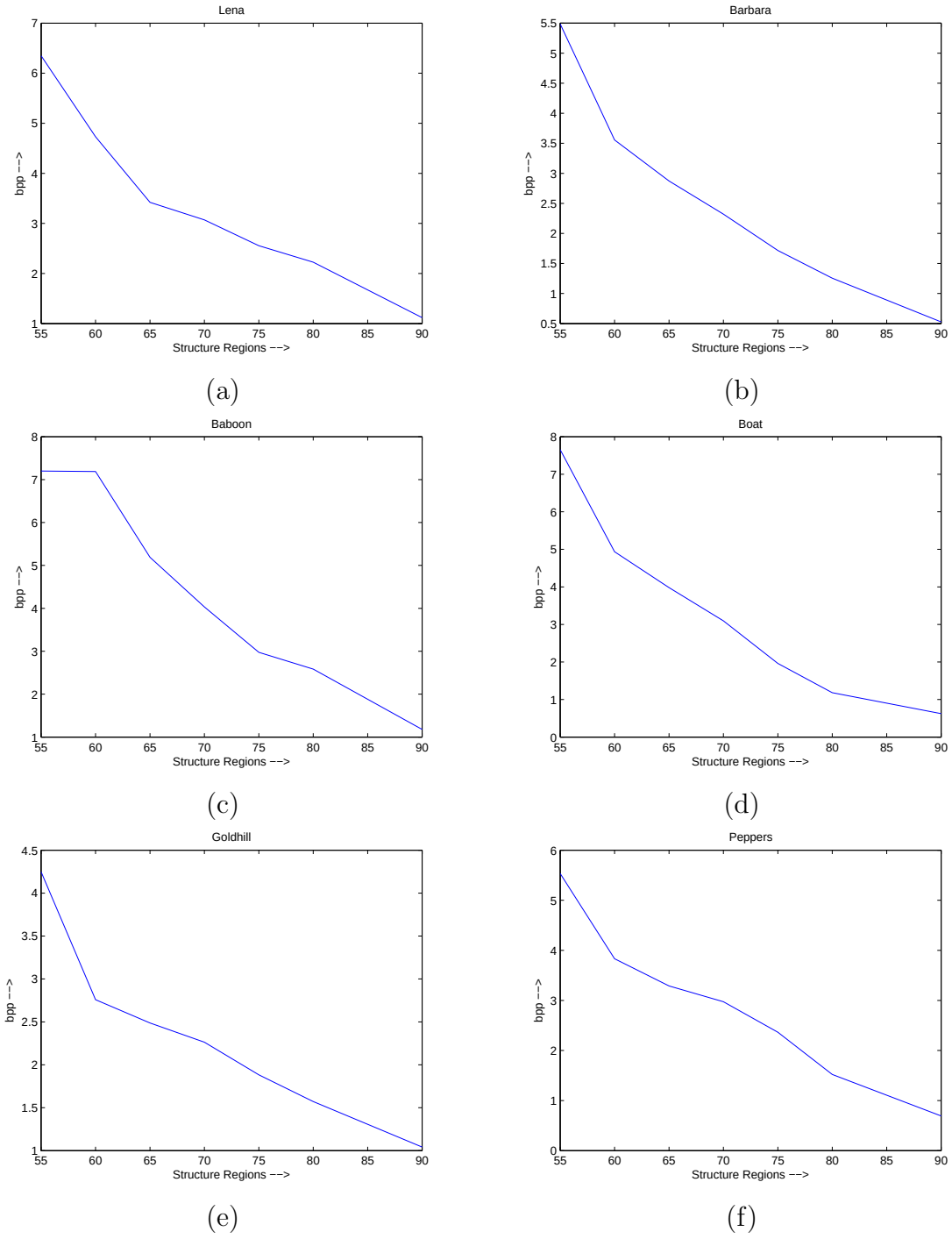


Figure 5.7: Variation of bit rate with percentage of structure regions (a) Lena, (b) Barbara, (c) Baboon, (d) Boat, (e) Goldhill, (g) Peppers

Table 5.4: Compression performance comparison for each of the R, G and B components for CALIC-SPS using direct prediction mode with CALIC in bpp

Image	bpp						bit-rate saving, Δbpp		
	CALIC			CALIC-SPS					
	R	G	B	R	G	B	R	G	B
Lena	3.9924	4.5581	4.8587	5.4001	2.2309	2.3736	-1.4077	2.3272	2.4851
Aircraft	3.862	3.9259	3.9306	2.4806	2.2131	3.5127	1.3814	1.7128	0.4179
Baboon	6.0389	6.2115	6.1609	4.2115	5.44783	3.3597	1.8274	0.7337	2.8012
Lake	5.1924	5.4294	5.1297	3.0973	3.5068	4.7533	2.1131	1.9226	0.3764
Peppers	5.0632	4.6796	4.6846	5.2476	3.3365	1.4489	-0.1844	1.3431	3.2357

Table 5.5: Compression performance comparison of CALIC-SPS using direct prediction mode with CALIC for color images in bpp

Image	bpp		bit-rate saving, Δbpp
	CALIC	CALIC-SPS Direct-prediction	
Lena	13.4092	10.0046	3.9817
Aircraft	11.7185	8.2064	3.5121
Baboon	18.4113	13.049	5.3623
Boat	15.7514	11.3394	4.412
Peppers	14.4274	10.033	4.3944

The CALIC-SPS algorithm is extended for color image compression. The color images are decomposed into individual R, G and B components and the compression scheme is run separately on each color plane. The bit-rate saving for each of the R, G and B components and the combined bit-rate saving for the color image using CALIC and CALIC-SPS that employ direct prediction is tabulated in Table 5.4 and Table 5.5. Our proposed scheme, CALIC-SPS using intra-prediction mode is compared with CALIC for each of components: R, G and B and bit-rate savings are tabulated in Table 5.6 and combined bit-rate savings are tabulated in Table 5.7. Comparison of bit-rate using different methods of prediction is tabulated in Table 5.8.

Table 5.6: Compression performance comparison for each of the R, G and B components for CALIC-SPS using intra-prediction mode with CALIC in bpp

Image	bpp						bit-rate saving, Δbpp		
	CALIC			CALIC-SPS					
	R	G	B	R	G	B	R	G	B
Lena	3.9924	4.5581	4.8587	5.5984	2.062	1.767	-1.606	2.4961	3.0917
Aircraft	3.862	3.9259	3.9306	2.0859	2.1256	3.4805	1.7761	1.6003	0.4501
Baboon	6.0389	6.2115	6.1609	3.7886	5.0375	2.8752	2.2503	1.174	3.2857
Lake	5.1924	5.4294	5.1297	2.9302	4.4655	3.9063	2.2622	0.9639	1.2234
Peppers	5.0632	4.6796	4.6846	5.2581	4.2229	0.98	-0.1949	0.4567	3.7046

Table 5.7: Compression performance comparison of CALIC-SPS using intra-prediction mode with CALIC bit-rate bpp

Image	bpp		bit-rate saving, Δbpp
	CALIC	CALIC-SPS Intra-prediction	
Lena	13.4092	9.4275	3.9817
Aircraft	11.7185	7.692	4.0265
Baboon	18.4113	11.70	6.7100
Boat	15.7514	11.302	4.4494
Peppers	14.4274	10.461	3.9644

Table 5.8: Comparison of bit-rate using CALIC-SPS for different modes of prediction in color images

Image	bpp		bit-rate saving, Δbpp
	CALIC-SPS Direct prediction	CALIC-SPS Intra-prediction	
Lena	10.0046	9.4275	0.5771
Aircraft	8.0246	7.692	0.5144
Baboon	13.049	11.70	1.3477
Boat	11.3394	11.302	0.0374
Peppers	10.033	10.461	0.428

5.4 Making the threshold adaptive

To obtain an adaptive threshold we have done the following study. We consider two parameters: local region(4×4) variance, σ_i^2 and the global maximum of local variances, $\max(\sigma_i^2)$ and thus compute normalised local variance, $\sigma_{iN}^2 = \sigma_i^2 / max(\sigma_i^2)$. Then, with simulation, we have found some relation between bpp and σ_{iN}^2 for various test images tabulated in Table 5.9 and Figure 5.8.

From Table 5.9 and Figure 5.8, it is observed that the bpp is almost steady in the range of 0.23 to 0.29, say nearly 0.25 for $\sigma_i^2 > 0.5$, whereas it increases appreciably beyond that value.

Table 5.9: Variation of bit rate, bpp with variance threshold

Normalised Variance Threshold	Lena	Peppers	Baboon	Barbara
0.1	0.5493	0.5219	1.7573	0.8910
0.2	0.3893	0.3824	0.8133	0.4835
0.3	0.3401	0.3	0.5281	0.3551
0.4	0.3134	0.2699	0.3636	0.2944
0.5	0.2842	0.2564	0.2889	0.2706
0.6	0.2721	0.2424	0.2666	0.2604
0.7	0.2586	0.2402	0.2493	0.2422
0.8	0.2421	0.2402	0.2373	0.2397
0.9	0.2373	0.2351	0.2348	0.2349

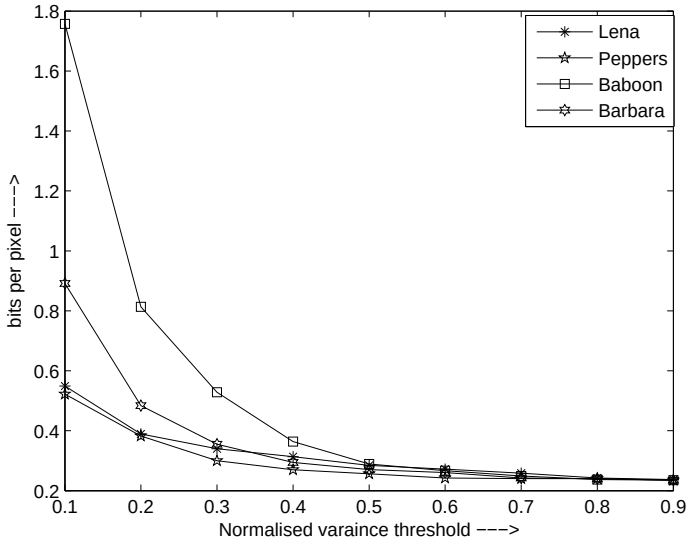


Figure 5.8: Bit rate variation with variance threshold

From this study, we infer the following :

For a structural region :

local variance $\geq 0.5\times$ global maximum of local variances.

Therefore, we develop the following scheme to obtain an adaptive threshold that significantly reduces the bpp(almost by a factor of 10) compared to its non- adaptive counterpart.

Algorithm

FOR i = 1 : number_of_blocks
IF
local variance of i^{th} block $\geq 0.5\times$ global maximum of local variance
THEN
i^{th}*_block* is structural region
ELSE
i^{th}*_block* is non-structural region
END
END

5.5 Conclusion

A hybrid lossless image compression scheme based on spatial prediction structures and CALIC providing extra bit-rate reduction when compared to standard lossless compression CALIC is analyzed. The scheme is motivated by motion estimation in video coding. It works well for images with significant amount of structure regions. Taking CALIC as the base code, the image has been classified into two regions and then the CALIC-SPS is applied to structure regions while the simple CALIC is applied to non-structure regions.

The extensive experimental results demonstrate that the hybrid scheme is very efficient in lossless image compression, especially for images with significant structure components. One major drawback of the hybrid CALIC-SPS scheme is its computational complexity due to expensive structure prediction. Sub-optimum

algorithms can be designed to provide a good trade-off between complexity and coding efficiency.

Finally an adaptive scheme is developed to automatically vary the threshold considering a local region as structural region or not. This has been employed in our proposed CALIC-SPS algorithm. With adaptive threshold, this scheme yields highly promising results with bpp almost ten times less than that its non- adaptive counterpart provides.

Chapter 6

Conclusion

6.1 Introduction

In this chapter, the overall conclusions are presented and the contributions are summarised. The algorithms, proposed in this thesis, have been developed for providing high compression rates (lower bit rates) while maintaining the visual quality of the decompressed image at high levels. The results yielded by the compression schemes are analysed in the next section.

6.2 Result Analysis

A thorough experimental analysis of SPIHT and VQ has been presented in Chapter 2 in terms of traditional quality metric, PSNR and the recently developed image quality metrics: VIFP, MSSIM and UQI. These two standard compression schemes are hybridised and a new image compression scheme SPIHT-VQ has been proposed in Chapter 3. The compression performance of SPIHT-VQ is compared with the state-of-art scheme SPIHT. The comparison shows that there is a satisfactory improvement in the image quality in terms of human visual system based metrics: VIFP, MSSIM and UQI at low bit rates, although the PSNR performance is not so promising. Further improvement in the visual quality is achieved by incorporating HVS in SPIHT and then cascading the VQ. The scheme called HSPIHT-VQ shows the improvement in image quality over SPIHT-VQ. This has been presented in Section 3.3.

Though the wavelet based coders provide good image quality at high compression ratios, they have received very less attention from digital camera manufacturers and software developers. BDCT is the Work horse on which even the latest MPEG video coding standards rely. JPEG still dominates consumer market and nearly all the pictures on internet are compressed using JPEG. An efficient image compression scheme called DCT-F is developed that hybridises BDCT with fractals. This has been presented in Section 4.2. It has been verified from simu-

lations that additional 5% bit rate reduction can be achieved over BDCT based compression by this scheme at a very marginal cost in terms of the image quality metrics used.
A modified point-wise shape adaptive DCT algorithm that yields smoothing and deblocking is developed and presented in Section 2.6. This post-processing scheme can be used with decoders at the receiver end to improve the visual performance of highly compressed BDCT and JPEG images.

A lossless hybrid image compression scheme, CALIC-SPS, which breaks the neighborhood constraint for the image prediction, is discussed in Chapter 5. The scheme is motivated from motion estimation in video coding. It provides promising results as compared to CALIC scheme.

To have a bird's eye view on the performance of all the proposed schemes, their results in terms of PSNR, VIFP, MSSIM and UQI are presented in Table 6.1 for test image Lena.

From Table 6.1, it is observed that the SPIHT scheme at a bit-rate of 0.0625 bpp yields a PSNR of 27.35 dB, but its visual quality is very poor as indicated by VIFP value. On cascading with SOFM based VQ (SPIHT-VQ scheme), though the PSNR value drops to 21.85 dB, the visual quality is better, as indicated by the metrics: VIFP, MSSIM and UQI. For the BDCT scheme, the settings are adjusted to give compression ratio of 0.5 bpp. The scheme DCT-F is simulated with same settings. The results in Table 6.1 indicate that the proposed scheme: DCT-F performs better in terms of bit-rate with very less degradation in performance in terms of the distortion metrics. The hybrid CALIC-SPS method provides more bit rate savings than the existing method. CALIC and CALIC-SPS being lossless schemes, the results in terms of distortion metrics are not mentioned. In the proposed modified point-wise SA-DCT based algorithm, the objective metrics depict improvement in the deblocking performance of the visual quality of the image.

Table 6.1: Compression Performance of the existing and proposed schemes for the test image: Lena

	Scheme	PSNR(dB)	VIFP	MSSIM	UQI	bpp
Lossy	SPIHT	27.35	0.2286	0.7610	0.4431	0.0625
	VQ	35.72	0.8203	0.8444	0.9997	0.5
	SPIHT-VQ	21.85	0.4065	0.7966	0.4857	0.0625
	HSPIHT-VQ	22.50	0.4415	0.8121	0.5047	0.0625
	BDCT	36.97	0.8382	0.8743	0.9988	0.5
	DCT-F	36.36	0.7976	0.8660	0.9987	0.415
	Point-wise SA-DCT based deblocking	32.12	0.8368	1.0000	0.5968	—
	Modified Point-wise SA-DCT based deblocking	31.99	0.8513	1.0000	0.5791	—
Lossless	CALIC	–	–	–	–	3.1509
	CALIC-SPS	–	–	–	–	3.0711

6.3 Conclusion

The analysis, presented in the previous section, leads us to draw the following conclusion.

- The proposed scheme, SPIHT-VQ2 performs better than the other schemes at 0.25 bpp . A bit-rate of 0.25 bpp (CR = 32) is good enough to preserve the image details with quite less distortion. A bit-rate of 0.125 bpp (CR = 64) may also be accepted in some applications where we can tolerate little more distortion.

- The proposed scheme DCT-F gives promising results, maintaining the quality in terms of objective metrics with slightly extra compression.
- The modified point-wise SA-DCT approach for deblocking provides the improvement in visual appearance of the image.
- Our proposed scheme, CALIC-SPS provides better compression than the existing standard method CALIC for lossless compression by yielding lower bit-rate.

Finally, it may be concluded that we may select CALIC-SPS algorithm for lossless compression whereas SPIHT-VQ2 will be a better candidate for lossy compression systems.

6.4 Scope for Future Work

The research work in this thesis can further be extended in the following directions:

- Adaptive HVS and modified SPIHT can be used with VQ to improve the performance of SPIHT-VQ by a large margin. Post processing techniques may be used to further enhance the subjective quality of images.
- DCT-F algorithm can be extended for color image/video coding by exploiting correlations among different color planes.
- Adaptive threshold may be employed to improve the coding efficiency of CALIC-SPS.

References

[1] Barni Mauro, *Document and image compression,* CRC press, 2006.

[2] David, Salomon,*Data compression: The Complete Reference,* Springer, USA, ISBN 10:1846286026, 2007.

[3] R.C.Gonzalez and R.E.Woods,*Digital Image Processing,* Pearson Prentice Hall, third edition, 2008.

[4] Bennamoun Mohammed and George J.Mamic, *Object recognition: fundamentals and case studies,* Springer, 2002.

[5] Sheikh H.R. and A.C.Bovik. *Information theoretic approaches to image quality assessment,* Handbook of Image and Video Processing, Elsevier Academic Press, pp 975-92,2005.

[6] Ohm, J. R., Sullivan, G. J., Schwarz, H., Tan, T. K., Wiegand, T. *Comparison of the coding efficiency of video coding standardsincluding high efficiency video coding (HEVC).* IEEE Transactions on Circuits and Systems for Video Technology, Vol 22,No.12, pp 1669-1684, 2012

[7] Gonzalez Rafael C, Richard E. Woods and Steven L. Eddins, *Digital image processing using MATLAB,* Upper Saddle River, N.J: Pearson Prentice Hall 2004.

[8] Gersho Allen and Robert M. Gray"Vector quantization and signal compression," Springer, 1992.

[9] Blanchet Gerard and Maurice Charbit, *Digital Signal and Image processing using MATLAB,* Vol.666, John Wiley & Sons, 2010.

[10] Sayood Khalid,*Data compression,* Morgan Kaufmann Publishers, San Francisco, CA 2000.

[11] Mallat Stephane, *A wavelet tour of signal processing,* Academic press, 1999.

[12] Strang Gilbert and Truong Nguyen, *Wavelets and filter banks,* SIAM, 1996.

[13] Daubechies Ingrid, *Ten lectures on wavelets,* Philadelphia: Society for industrial and applied mathematics, Vol.61, 1992.

[14] Scharstein Daniel, and Richard Szeliski,"A taxonomy and evaluation of dense two-frame stereo correspondence algorithms," International journal of computer vision, Vol.47, no.1-3, pp 7-42, 2002.

[15] K.Veeraswamy, S.Srinivaskumar and B.N.Chatterji, "Designing quantization table for Hadamard transform based on human visual system for image compression," ICGST-GVIP, Vol.7, pp. 31-38, Nov 2007

[16] Al Sayeed, Choudhury A. and Abul Bashar M.Ishteak HossainImage compression using frequency-sensitive competitive neural network,Photonics Asia 2004. International Society for Optics and Photonics, 2005.

[17] Bairagi Vinayak K.,and Ashok M Sapkal, "ROI-based DICOM image compression for telemedicin," Sadhana 38, no.1, pp 123-131, 2013.

[18] Monteagudo-Pereira, Jose Lino, Joan Bartrina-Rapesta, Francesc Auli-Llinas, Joan Serra-Sagrista, Alaitz Zabala, and Xavier Pons, "Interactive transmission of spectrally wavelet-transformed hyperspectral images," In Optical Engineering Applications, pp. 708405-708405, International Society for Optics and Photonics, 2008.

[19] Bouzerdoum Abdesselam,"Image compression using a stochastic competitive learning algorithm (SCOLA)," In Sixth International Symposium on Signal Processing and its Applications, Vol.2, pp 541-544, 2001.

[20] Avramovic Aleksej and Goran Banjac,"On predictive-based lossless compression of images with higher bit depths," Telfor Journal, Vol.4, no.2, pp 122-127, 2012.

[21] Wu Xiaolin and Nasir Memon"CALIC-a context based adaptive lossless image codec," IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP-96., Vol.4, 1996.

[22] Kornblum Jesse D,"Using JPEG quantization tables to identify imagery processed by software," digital investigation 5, pp S21-S25 ,2008.

[23] Jacquin Arnaud E,"Image coding based on a fractal theory of iterated contractive image transformations," IEEE Transactions on Image Processing, Vol.1, no.1, pp 18-30, 1992.

[24] Ueno Hirokazu and Yoshitaka Morikawa,"A New Distribution Modeling for Lossless Image Coding Using MMAE Predictors," The 6^th International Conference on Information Technology and Applications, 2009.

[25] Manjunatha D.V. and G.Sainarayanan, "Low power 8×8 Sum of Absolute Difference engine for Motion Estimation in video compression," In IEEE International Conference on Emerging Trends in Computing, Communication and Nanotechnology, pp. 68-73, 2013.

[26] He Xiao Cheng, Sheng Yu, and Jing Li Zhou, "Context and HVS based multiwavelet image coding using SPIHT framework," Circuits, Systems and Signal Processing, Vol.24, no.2 pp 117-134, 2005.

[27] Kharate Gajanan K, Varsha H.Patil, and Niranjan L.Bhale,"Selection of mother wavelet for image compression on basis of nature of image,"Journal of Multimedia, Vol.2, no.6, pp 2007.

[28] Hafner, Ullrich, Frank S Unger M, and Albert J, "Hybrid weighted finite automata for image and video compression," Technical Report 160, Department of Computer Science, University Wurzburg, 1997.

[29] Delp Edward J, Salama Paul, Asbun Eduardo, Saenz N and Shen Ke,"Rate scalable image and video compression techniques," In IEEE 42^{nd} Midwest Symposium on Circuits and Systems, Vol. 2, pp 635-638, 1999.

[30] Ujjaval Y Desai, Ichiro Masaki Marcelo M. Mizuki and Berthold K.P.Horn,"Edge and mean based image compression," In Massachussetts Institute of Technology Artificial Intelligence Labortry, no.1584, 1996.

[31] Uli Graseman and Risto Miikkulainen,"Effective image compression by using evolved wavelets," In Proceedings of 2005 Conference on Genetic and Evolutionary Computation, pp 1961-1968, Washington DC, 2005.

[32] Laurent Demaret and Arim Iske,"Advances in digital image compression by adaptive thinning," In Annals of MCFA, pp 105-109, 2004.

[33] Grois, Dan, Detlev Marpe, Amit Mulayoff, Benaya Itzhaky, and Ofer Hadar. *Performance comparison of H. 265/MPEG-HEVC, VP9, and H. 264/MPEG-AVC encoders.* In IEEE Picture Coding Symposium (PCS), pp. 394-397 , 2013.

[34] Davis Geoffrey M. and Aria Nosratinia,"Wavelet-based image coding: an overview," Applied and computational control, signals, and circuits. Birkhuser Boston, pp 369-434, 1999.

[35] Mohamad Dzulkifli,"Image Compression Using Hybrid Technique," In Information and Knowledge Management, Vol.2, no.7, pp 10-16, 2012.

[36] Giuseppe Placidi,"A novel adaptive lossless compression algorithm for efficient medical image archiving and transmission" ENMI, 1(1), pp 1204-1209, 2005.

[37] Li Wern Chew, Li-Minn Ang, and Kah Phooi Seng,"Lossless image compression using tuned degree-k zerotree wavelet coding," In Proceedings of the International Multi Conference of Engineers and Computer Scientists,Vol 1,pp 779-782, 2009.

[38] Chakrapani Y. and K.Soundara Rajan,"Genetic algorithm applied to fractal image compression," ARPN Journal of Engineering and applied Sciences, Vol.4, no.1, pp 53-58, 2009.

[39] Gray Robert M. and David L. Neuhoff. "Quantization,"IEEE Transactions on Information Theory, Vol.44, no.6 pp 2325-2383, 1998.

[40] Marta Mrak, Sonja Grgic and Mislav Grgic,"Picture quality measures in image compression systems," In EUROCON 2003, Computer as a Tool. The IEEE Region 8, 2003.

[41] Deepak Mishra, N.Subhash Chandra Bose, Arvind Tolambiya, Ashutosh Dwivedi, Prabhanjan Kandula, Ashiwani Kumar, Prem K. Kalra, "Color Image Compression with Modified Forward-Only Counterpropagation Neural Network: Improvement of the Quality using Different Distance Measures," International Conference on Information Technology, pp 139-140, 2006.

[42] Borisagar Viral H and Mukesh A.Zaveri,"Disparity Map Generation from Illumination Variant Stereo Images Using Efficient Hierarchical Dynamic Programming," The Scientific World Journal 2014 .

[43] Zhang Liangbin and Lifeng Xi,"Hybrid image compression using fractal-wavelet prediction," WSEAS Transactions on Systems,Vol.6, no.3, pp 556, 2007.

[44] Zhou Wang and Alan C.Bovik,"A universal image quality index,"IEEE Signal Processing Letters, Vol.9, no.3 , pp 81-84, 2002.

[45] Zhou Wang, Alan C.Bovik, Hamid R.Sheikh and Eero P.Simoncelli, "Image quality assessment: from error visibility to structural similarity," IEEE Transactions on Image Processing, Vol.13, no.4, pp. 600-612, 2004.

[46] Linde Yoseph, Andres Buzo and Robert M.Gray,"An algorithm for vector quantizer design," IEEE Transactions on Communications, Vol.28,no.1, pp 84-95, 1980.

[47] Shen Day-Fann, and Kuo-Shu Chang,"Fast PNN algorithm for design of VQ initial codebook," International Society for Optics and Photonics , Photonics West'98 Electronic Imaging, pp 842-850, 1998.

[48] Flanagan J.K, D R. Morrell, R.L.Frost, Christopher J.Read, and Brent E.Nelson, "Vector quantization codebook generation using simulated annealing," International Conference on Acoustics, Speech, and Signal Processing, ICASSP-89, pp 1759-1762, 1989.

[49] C.K. Ma and C.K.Chan,"Maximum descent method for image vector quantization," Electronics Letters, Vol.27, no.19, pp 1772-1773, 1991.

[50] C.K.Ma and C.K.Chan,"A fast method of designing better codebooks for image vector quantization,"IEEE Transactions on Communications, Vol.40, no.2, pp 237-242, 1994.

[51] Equitz William H,"A new vector quantization clustering algorithm," IEEE Transactions on Speech and Signal Processing, Vol.37, no.10, pp 1568-1575, 1989.

[52] A.Laha, N.R.Pal and B.Chanda,"Design of vector quantizer for image compression using self-organizing feature map and surface fitting," IEEE Transactions on Image Processing, Vol.13, no.10, pp 1291-1303, 2004.

[53] Yan Wanga, Amine Bermaka, Abdesselam Bouzerdoumb and Brian Ngc, "FPGA implementation of a predictive vector quantization image compression algorithm for image sensor applications," In 4th IEEE International Symposium on Electronic Design, Test & Applications, pp 431-434, 2008.

[54] Butler D,"Distortion equalized fuzzy competitive learning for image data vector quantization,"In IEEE International Conference on Acoustics, Speech, and Signal Processing ICASSP-96, Vol 6, pp 3390-3396, 1996.

[55] Lai Chung Fu and Tong Lee, "Fuzzy competitive learning, Neural Networks," Vol.7, no.3, pp 539-551, 1994.

[56] Banerjee Arindam, and Joydeep Ghosh,"Frequency-sensitive competitive learning for scalable balanced clustering on high-dimensional hyperspheres," IEEE Transactions on Neural Networks, Vol.15, no.3, pp 702-719, 2004.

[57] Amir.Said and William A.Pearlman,"A new, fast, and efficient image codec based on set partitioning in hierarchical trees," IEEE Transactions on Circuits and Systems for Video Technology, Vol.6, no.3, pp 243-250, 1996.

[58] Jiang J, "Image compression with neural networksa survey," Signal Processing: Image Communication, Vol.14, no.9, pp 737-760, 1999.

[59] Zhe-Ming L.U.and P.E.I.Hui, "Hybrid image compression scheme based on PVQ and DCTVQ," IEICE transactions on information and systems, Vol.88, no.10, pp 2422-2426, 2005.

[60] Lenni Yulianti and Tati R. Mengko, "Application of hybrid fractal image compression method for aerial photographs," MVA2OOO IAPR Workshop on Machine Vision Applications University of Tokyo, Japan, 2000.

[61] Nilesh Singh V.Thakur and Dr.O.G.Kakde,"Color image compression with modified fractal coding on spiral architecture," Journal Of Multimedia, Vol.2, no.4, pp 55-66, August 2007.

[62] Se-Kee Kil, Jong-Shill Lee, Dong-Fan Shen, Je-Goon Ryu, Eung-Hyuk Lee,Hong-Ki Min, and Seung-Hong Hong,"Lossless medical image compression using redundancy analysis," IJCSNS International Journal of Computer Science and Network Security, Vol.6, no.1A, pp 50-56, January 2006.

[63] C.Hemasundara Rao and M.Madhavi Latha."A novel VLSI architecture of hybrid image compression model based on reversible blockade transform." International Journal of Electronics, Circuits and Systems, Vol.3, no.1, pp 25-31, 2009.

[64] Shapiro Jerome M."Embedded image coding using zerotrees of wavelet coefficients," IEEE Transactions on Signal Processing,Vol.41, no.12, pp 3445-3462, 1993.

[65] M.J.Weinberger, G.Seroussi and G.Shapiro," Loco-I: A low complexity,context-based, lossless image compression algorithm," In Proceedings of 1996 Data Compression Conference, pp 140-149, (Snowbird, Utah, USA), March 1996.

[66] Xiaolin Wu, "Lossless compression of continuous-tone images via context selection quantization and modeling", IEEE Transactions On Image Processing, Vol.6, no.5, pp 656-664, May 1997.

[67] I.Tabus and J.Astola,"Adaptive boolean predictive modelling with application to lossless image coding," In SPIE - Statistical and Stochastic Methods for Image Processing II, pp 234-245, San Diego California, July 1997.

[68] K.Balashov, D.Akopian and J.Astola,"Lossless compression of natural images using regressive analysis," In Proceedings of Eusipco 2000, Vol.2, pp 1181-1183, Tampere Finland, September2000.

[69] S.E. Ghrare M.A. Mohd. Ali K. Jumari and M. Ismail," An efficient low complexity lossless coding algorithm for medical images," American Journal of Applied Sciences, Vol.6, no.8, pp 1502-1508, 2009.

[70] Wang Zhou, and Alan C.Bovik, "Modern image quality assessment," Synthesis Lectures on Image, Video, and Multimedia Processing Vol.2, no.1, pp 1-156, 2006.

[71] V.Lukin, M Zriakov, S Krivenko, N Ponomarenko, and Z Miao,"Lossy compression of images without visible distortions and its applications," In Proceedings of ICSP 2010, pp 694-697, October 2010.

[72] Zhou Wang and A.C.Bovik," Mean squared error: Love it or leave it ?," IEEE Signal Processing Magazine, pp 98-117, January 2009.

[73] A.C.Bovik and H.R.Sheikh, "Image and video quality assessment research at LIVE."[Online].Available: http://live.ece.utexas.edu/research/quality.

[74] H.R.Sheikh and A.C.Bovik"Image information and visual quality," IEEE Transactions on Image Processing,Vol.15, no.2, pp 430-444, February 2006.

[75] Zhou Wang and A.C.Bovik,"Image quality assessment: From error visibility to structural similarity," IEEE Transactions on Image Processing, Vol.13, no.4, pp 600-612, April 2004.

[76] K.S.Thyagrajan, *Still Image and Video Compression with MATLAB.* John Wiley and Sons Inc, 2011.

[77] Starosolski Roman,"Simple fast and adaptive lossless image compression algorithm," Software: Practice and Experience, Vol.37, no.1, pp 65-91, 2007.

[78] John F. Tilki and A.A.(Louis) Beex, "Image data compression using multiple bases representation,"In 26th IEEE Southeastern Symposium on System Theory(SSST'94) Athens OH, pp 457-461, March 20-22 1994.

[79] Sunil Kumar and R.C.Jain,"Low complexity fractal-based image compression techniques," IEEE Transactions on Consumer Electronics, Vol.43, no.4, pp 987-993, 1997.

[80] Vijaya Prakash A. and Gurumurthy K.“ A novel VLSI architecture for digital image compression using discrete cosine transform and quantization,” International Journal of Computer Science and Network Security, Vol.10, no.9, pp 175-182, September 2010.

[81] Singh.S.and Kumar.S.Maejo,“Mathematical transforms and image compression: A review, International Journal of Science and Technology,” Vol.4, no.2, pp 235-249, 2010.

[82] Sarantos Psycharis, “The didactic use of digital image lossy compression methods for the vocational training sector,” In IV International Conference on Multimedia and ICTs in Education, University of Agean, pp 2066-2069, Seville (Spain), 2006.

[83] Alexander Wong and William Bishop,“A Flexible Content Based Approach to Adaptive Image Compression,” In IEEE International Conference on Multimedia and Expo, pp 713-716, Toronto, 2006.

[84] Ashutosh Dwivedi, N.Subhash Chandra Bose, Ashiwani Kumar, Prabhanjan Kandula, Deepak Mishra and Prem K.Kalra,“A novel hybrid image compression technique: Wavelet-MPFOCON,” In The 9th Asian Symposium on Information Display ASID-06 New Delhi India, 2006.

[85] Mascher-Kampfer A, Herbert Stgner, and Andreas Uhl, “Comparison of compression algorithms impact on fingerprint and face recognition accuracy,” Electronic Imaging 2007, International Society for Optics and Photonics, 2007.

[86] Muhammad Azhar Iqbal, Muhammad Younus Javed and Usman Qayyum,“Curvelet-based image compression with SPIHT,” In International Conference on Convergence Information Technology, pp 961-965, IEEE, 2007.

[87] Osman.G.Sezer, Oztan Harmanci and Onur.G.Guleryuz,“Sparse orthonormal transforms for image compression,” In 15th IEEE International Conference on Image Processing, pp 149-152, 2008.

[88] Takahiro Nakayama, Masahiro Konda, Koji Takeuchi, Koji Kotani and Tadahiro Ohmi,“Still image compression with adaptive resolution vector quantization technique,” Intelligent Automation and Soft Computing, Vol.10, no.2, pp 155-166, 2004.

[89] Ian Berry, Julie Wilson, Chris Mayo, Jon Diprose, and Robert Esnouf, “The effect of image compression on classification and storage requirements in

a high-throughput crystallization system," In International conference on intelligent data engineering and automated learning, IDEAL 2004, Springer Berlin Heidelberg, pp 117-124, 2004.

[90] A.Kumar Kombaiya and V.Palanisamy,"Wavelet based image compression using ROI SPIHT coding," International Journal of Computational Intelligence Research, Vol.5, no.1, pp 67-74, 2009.

[91] K.Veeraswamy and S.Srinivas Kumar,"Adaptive ac-coeffcient prediction for image compression and blind watermarkiing," Journal of Multimedia, Vol.3, no.1, pp 16-22, May 2008.

[92] Liangbin Zhang and Lifeng Xi,"Hybrid image compression using fractal-wavelet prediction," In Proceedings of the 5th WSEAS International Conference on Information Security and Privacy, pp 112-117, Venice Italy, 2006.

[93] Zhe-Ming Lu and Hui Pei,"Hybrid image compression scheme based on PVQ and DCTVQ," IEICE Transactions on Information and Systems,Vol.E88D, no.10, pp 2422-2426, 2005.

[94] Martin K.and Lukac R. and Plataniotis K.N,"SPIHT based coding of the shape and texture of arbitrarily shaped visual objects," IEEE Transactions on Circuits and Systems for Video Technology, Vol.16, no.10, pp 1196-1208, 2006.

[95] Wiegand Thomas, Gary J.Sullivan, Gisle Bjontegaard and Ajay Luthra, "Overview of the H. 264/AVC video coding standard," IEEE Transactions on Circuits and Systems for Video Technology, Vol.13, no.7, pp 560-576, 2003.

[96] Diri Banu and Songul Albayrak,"Color image compression using self organizing feature map," Proceedings of the 24th IASTED International Conference on Artificial Intelligence and Applications. ACTA Press, 2006.

[97] Dinesh.K.and Gaur Loveleen and Okunbor D,"Image compression and feature extraction using Kohonen's self-organizing map neural network," Journal of Strategic E-Commerce, Vol.5, no.1, pp 1-4, 2007.

[98] Zhang Xing-hui, Jing-lei Guo, Kuan-sheng Zou and Zhi-dong Deng,"Improved SPIHT algorithm based on associative memory neural network and human visual system," In IEEE International Conference on Intelligent Computation Technology and Automation ICICTA- 2008, Vol.1, pp. 200-203, 2008 .

[99] Sung Tze Yun and Hsin H.C,“ A hybrid image coder based on spiht algorithm with embedded block coding,” IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences, Vol.90, no.12, pp 2979-2984, 2007.

[100] U.S.Mohammed,“Highly scalable hybrid image coding scheme,” Digital Signal Processing, Vol.18, no.3, pp 364-374, 2008.

[101] Ramya R, and K. Mala, A hybrid Compression Algorithm for Compound Images, IEEE International Conference on Computational Intelligence and Multimedia Applications, Vol.3, pp 68-72, 2007.

[102] Chen Yen-Yu,“ Medical images compression for remote diagnosis using modified SPIHT data organization and fidelity enhancement filter,” International Journal of Imaging Systems and Technology, Vol. 17, no.2, pp. 49-61, 2007.

[103] Kazuyuki Tanaka, Norihiro Hoshi and Tsuyoshi Horiguchi,“Color image compression algorithm using self-organizing feature map,” Interdisciplinary Information Sciences,Vol.9, no.2, 2003.

[104] Pei Soo-Chang and You-Shen Lo,“Color image compression and limited display using self-organization kohonen map,” IEEE Transactions on Circuits Systems and Video Technology, Vol.8, no.2, pp 191-205, 1998.

[105] Annadurai.S and M.Sundaresan“Wavelet based color image compression using vector quantization and morphology,” In Proceedings of the International Conference on Advances in Computing, Communication and Control, pp. 391-396. ACM, 2009.

[106] S.A.Durai and E.A.Saro,“An improved image compression approach with SOFM Network using Cumulative Distribution Function,” In International Conference on Advanced Computing and Communications, ADCOM, pp 304-307, 2006.

[107] Kuo Liang Chung, Liu Y and Yan W,“ A hybrid gray image representation using spatial and DCT-based approach with application to moment computation,” Journal of Visual Communication and Image Representation, Vol.17, no.6, pp 1209-1226, December 2006.

[108] Pandian.S.and J.Anitha“A neural network approach for color image compression in transform domain,” International Journal of Recent Trends in Engineering, Vol.2, no.2, pp 152-154, November 2009.

[109] Krikor Lala, Baba S, Arif T and Shaaban Z,"Image encryption using DCT and stream cipher," European Journal of Scientific Research, Vol.32, no.1, pp 47-57, 2009.

[110] Khalil.M,"Image compression using new entropy coder," International Journal of Computer Theory and Engineering, Vol.2, no.1, pp 1793-1801, February 2010.

[111] Meng Meng and Zong Meijuan,"A new zerotree structure for color image compression based on DWT and VQ," The 2nd IEEE International Conference on Information Management and Engineering, ICIME 2010, pp. 339-342, 2010.

[112] Wang Yan, Amine Bermak, Abdesselam Bouzerdoum and Brian Ng,"FPGA implementation of a predictive vector quantization image compression algorithm for image sensor applications," In 4th IEEE International Symposium on Electronic Design, Test and Applications, DELTA 2008, pp. 431-434, 2008.

[113] Foi Alessandro, Vladimir Katkovnik and Karen Egiazarian, "Pointwise shape-adaptive DCT for high-quality denoising and deblocking of grayscale and color images," IEEE Transactions on Image Processing, Vol.16, no.5, pp 1395-1411, 2007

[114] Foi Alessandro, Vladimir Katkovnik, Karen Egiazarian and Jaakko Astola, "A novel anisotropic local polynomial estimator based on directional multiscale optimizations," In Proceedings of 6th IMA International Conference On Mathematica and Signal processing, pp. 79-82, 2004.

[115] Katkovnik V , Alessandro Foi, Karen Egiazarian and Jaakko Astola, " Directional varying scale approximations for anisotropic signal processing," In Proceedings of XII European Signal Processing Conference EUSIPCO 2004, pp 101-104, September 2004.

[116] Katkovnik Vladimir,"A new method for varying adaptive bandwidth selection," IEEE Transactions on Signal Processing, Vol.47, no.9, pp 2567-2571, 1999.

[117] Goldenshluger A and A.Nemirovski,"On spatially adaptive estimation of nonparametric regression," Mathematical methods of Statistics, Vol.6, no.2, pp 135-170, 1999.

[118] Averbuch Amir Z, Alon Schclar and David L. Donoho, "Deblocking of block-transform compressed images using weighted sums of symmetrically aligned

pixels," IEEE Transactions on Image Processing, Vol.14, no.2, pp 200-212, 2005.

[119] N.M Nasrabadi and R.A.King, "Image coding using vector quantization: A Review," IEEE Transactions on Communication, Vol.36, no.8, pp 957-971, 1988.

[120] R. M. Gray, "Vector Quantization," IEEE ASSP Magazine, Vol.1, pp 4-29, April 1984.

[121] Nasrabadi Nasser M and Yushu Feng,"Vector quantization of images based upon the Kohonen self-organizing feature maps," In IEEE International Conference on Neural Networks, pp 101-108, 1988.

[122] Ibrahim Oz Cemil and Oz Nejat Yumuayak,"Image Compression Using Multiple Level 2-D Discrete Wavelet Transform," Technical report, Sakarya University, Faculty of Engineering, Department of Electrical & Electronics Engineering, Sakarya, Turkey, 2000.

[123] Veerakumar.T, S Esakkirajan, R.Sudhakar and V.Senthil Murugan, "Fingerprint compression using contourlet transform and self organizing feature map," Iranian Journal of Electrical and Computer Engineering, Vol.6, no.2, pp 133, 2000.

[124] Equitz William H,"A new vector quantization clustering algorithm," IEEE Transactions on Acoustics, Speech and Signal Processing, Vol.37, no.10, pp 1568-1575, 1989.

[125] Liew Alan W.C, and Hong Yan,"Blocking artifacts suppression in block-coded images using overcomplete wavelet representation," IEEE Transactions on Circuits and Systems for Video Technology, Vol.14, no.4, pp 450-461, 2004.

[126] Sonja Gergic, Kresimir Kers, Mislav Grgic,"Image compression using wavelets," Proceedings of the IEEE International Symposium on Industrial Electronics, ISIE99, Vol.48, no.3, pp. 682-695, June 2001.

[127] Kohonen Teuvo,"The self-organizing map," Proceedings of the IEEE, Vol.78, no.9, pp 1464-1480, 1990.

[128] M.G.Ramos, S.S.Hemami, and M.A.Tamburro,"Psychovisually-based multiresolution image segmentation," In Proceedings of the International Conference on Image Processing (ICIP 97), Vol.3, Santa Barbara, CA, pp 6669, October 26-29 1997.

[129] Nadenau Marcus J, Julien Reichel and Murat Kunt,"Wavelet-based color image compression: exploiting the contrast sensitivity function," IEEE Transactions on Image Processing, Vol.12, no.1,pp 58-70, 2003.

[130] Al-Gindy A., H.Al-Ahmad, R.Qahwaji, and A.Tawfik, "A novel blind image watermarking technique for colour RGB images in the DCT domain using green channel," In IEEE International Conference on Communications, Computers and Applications, MIC-CCA 2008, Mosharaka, pp 26-31, 2008.

[131] Aitsab.O, R.Pyndiah and B.Solaiman, "Joint optimization of multidimensional SOFM codebooks with qam modulations for vector quantized image transmission," In Proceedings IWISPO, Vol.96, pp. 3-6, 1996.

[132] Mallat, Stephane G, "A theory for multiresolution signal decomposition: the wavelet representation," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.11, no.7, pp 674-693, 1989.

[133] Vetterli Martin and Cormac Herley,"Wavelets and filter banks: Theory and design," IEEE Transactions on Signal Processing, Vol.40, no.9, pp 2207-2232, 1992.

[134] Stanhill David, and Yehoshua Y.Zeevi, "Two-dimensional orthogonal and symmetrical wavelets and filter-banks," Conference Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing ICASSP-96, Vol.3, 1996.

[135] McGinnity De-Shuang Huang Martin and Laurent Heutte Xiao-Ping Zhang,*Advanced Intelligent Computing Theories and Applications,* Springer 2010.

[136] Howard Paul G, and Jeffrey S.Vitter, "Fast progressive lossless image compression," In International Symposium on Electronic Imaging: Science and Technology, SPIE 1994, pp 98-109, International Society for Optics and Photonics, 1994.

[137] Cooklev Todor and Akinori Nishihara, "Biorthogonal coiflets," IEEE Transactions on Signal Processing, Vol.47, no.9, pp 2582-2588, 1999.

[138] Donoho David L, "De-noising by soft-thresholding," IEEE Transactions on Information Theory, Vol.41, no.3, pp 613-627, 1995.

[139] Donoho David L and Iain M.Johnstone, "Adapting to unknown smoothness via wavelet shrinkage," Journal of the American statistical association, Vol.90, no.432, pp 1200-1224,1995.

[140] Donoho David L, "Ideal spatial adaptation via wavelet shrinkage," Biometrica, Vol.81, pp 424-455, 1994.

[141] Luisier Florian, Thierry Blu and Michael Unser, " A new SURE approach to image denoising: Interscale orthonormal wavelet thresholding," IEEE Transactions on Image Processing, Vol.16, no.3, pp 593-606, 2007.

[142] Chang S.Grace, Bin Yu, and Martin Vetterli, "Adaptive wavelet thresholding for image denoising and compression," IEEE Transactions on Image Processing, Vol.9, no.9, pp 1532-1546, 2000.

[143] Bhoi Nilamani, "Development of Some Novel Spatial-Domain and Transform-Domain Digital Image Filters," PhD dissertation, National Institiute of Technology, Rourkela, India, 2009.

[144] Goswami Jaideva C and Andrew K Chan,"Fundamentals of wavelets: Theory, Algorithms and Applications," Vol.233, John Wiley & Sons, 2011.

[145] Meyer Yves," Wavelets-algorithms and applications," Society for Industrial and Applied Mathematics Translation, Vol.142, 1993.

[146] Zhao Xiwen Owen and Zhihai Henry He, "Lossless image compression using super-spatial structure prediction," IEEE Signal Processing Letters, Vol.17, no.4, pp 383-386, 2010.

[147] Reichel Julien, Gloria Menegaz, Marcus J.Nadenau, and Murat Kunt," Integer wavelet transform for embedded lossy to lossless image compression," IEEE Transactions on Image Processing, Vol.10, no.3, pp 383-392, 2001.

[148] Servetto Sergio D, Kannan Ramchandran and Michael T.Orchard, "Image coding based on a morphological representation of wavelet data," IEEE Transactions on Image Processing, Vol.8, no.9, pp 1161-1174, 1999.

[149] Zhao Xiwen and Zhihai He," Local structure learning and prediction for efficient lossless image compression," In IEEE International Conference on Acoustics Speech and Signal Processing, ICASSP, pp. 1286-1289, 2010.

[150] Shen Day-Fann and Chang Kuo Shu,"Fast PNN algorithm for design of VQ initial codebook," In International Society for Optics and Photonics, Photonics West'98 Electronic Imaging, pp. 842-850, 1998.

[151] Stromberg Jan-Olov, "A modified Franklin system and higher order spline systems on R^n as unconditional bases for Hardy spaces," Fundamental Paper in Wavelet Theory, pp 197-215, 2006.

[152] Meyer Yves, "Principe d'incertitude, bases hilbertiennes et algebres d'operateurs," Seminaire Bourbaki, Vol.28, pp 209-223,1985 .

[153] Lemarie, Pierre-Gilles,Ondelettesa localisation exponentielle, J. Math. Pures Appl, Vol.67, no.3, pp 227-236, 1988.

Dissemination of Research Outcome

[**P1**]**Chandan Singh Rawat**, Sukadev Meher, A Hybrid Coding Scheme Combining SPIHT and SOFM Based Vector Quantization for Effectual Image Compression, European Journal of Scientific Research, ISSN 1450-216X,Vol.38, No.3, pp 425-440, 2009.

[**P2**]**Chandan Singh Rawat**, Sukadev Meher A Hybrid image compression scheme using HVS characteristics: Combining SPIHT and SOFM based Vector Quantization, International Journal of Signal and Imaging Systems Engineering, Vol.5, No.3, pp 175-186, 2012

[**P3**]**Chandan Singh Rawat**, Sukadev Meher, Selection of wavelet for image compression in hybrid coding scheme combining SPIHT and SOFM based Vector Quantization International Journal of Signal and Imaging Systems Engineering, Vol.7, No.1, 2014.

[**P4**]**Chandan Singh Rawat**, Sukadev Meher, A Hybrid Image Compression Scheme using DCT and Fractal Image Compression International Arab Journal of Information Technology, ISSN: 1683 -3198, Vol.10, No.6, Nov 2013.

[**P5**]**Chandan Singh Rawat**, Seema G. Bhateja, Sukadev Meher, A Similar Structure Block Prediction for Lossless Image Compression, International Journal of Computer Science and Communication Networks, (ISSN-2249-5789), Vol.1, No.3, Dec 2011.

[**P6**]**Chandan Singh Rawat**, Sukadev Meher, A Novel Algorithm of Spatial in RGB color Space for Image Compression, International Journal of Scientific Engineering and Research ,ISSN 2259-5518,Vol. 3, Issue 2, Feb 2012.

[**P7**]**Chandan Singh Rawat**, Sukadev Meher, Comparative Study of One Dimensional DCT versus Block DCT for Multiresolution Technique, International Conference on Computers, Communication and Control (ICCCI) on 2-3 Jan 2009 at VESIT, Mumbai-71

[**P8**]**Chandan Singh Rawat**, Rohan Shambharkar, Sukadev Meher, Modified Point wise Shape-Adaptive DCT for High-Quality Deblocking of Compressed Images, International Journal of Advanced Research in Computer Science and Electronics Engineering (IJARCSEE), Vol.1, no.2, pp-134, 2012

[**P9**]**Chandan Singh Rawat**, Sukadev Meher, Deblocking of Highly Compressed Images using Image Denoising Technique, National Conference on Recent Trends in Computer Engineering (RTCE-09), 29-30 Dec 2009 at Sinhagad College of Engineering, Pune-41.

[**P10**] **Chandan Singh Rawat**, Sukadev Meher, Reducing Blocking Artifacts Using Denoising Algorithm with Shape Adaptive DCT in Color Images, National conference on ICT: Theory, Applications and Practices, 5-.6 March 2010 at Sir Padampat Singhania University, Udaipur 313601

[**P11**]**Chandan Singh Rawat**, Sukadev Meher, Human Vision system Models in Digital Image Compression, National Conference on New Advances in Core Computing and their Challenges (NACCTC-2010), 20-21 March 2010 at M.B.M College of Engineering, JNV University, Jodhpur 342011.

Author's Biography

Name:	**Chandan Singh D Rawat**
Sex:	Male
Date of Birth:	20 November 1969
Father's Name:	Late Shri Dau Singhji Rawat
Nationality:	Indian
Present Address:	PhD Scholar Dept. of Electronics & Communication Engineering, National Institute of Technology Rourkela Rourkela - 769008 (INDIA)
Permanent Address:	302, A-wing, Building No.1, Adarsh Park, Kalyan Road, Dombivli(East), Thane - 421201 (INDIA)
E-mail:	csrawat3@gmail.com

Educational Qualifications:

Sl No.	Institute/University	Degree	Year
1	Nagpur University, Nagpur	B.E (Electronics Engineering)	1991
2	Mumbai University	M.E. (Electronics & Telecommunication)	2004
3	National Institute of Technology, Rourkela	Ph. D. (Image Processing)	2007-Continuing

Publications:

- Published 07 papers in International Journals.
- Published 04 papers in National and International Conferences