# PARENTAL FINITE STATE VECTOR QUANTIZER AND VECTOR WAVELET TRANSFORM-LINEAR PREDICTIVE CODING

By

Lam Chi Wah

A THESIS

SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF MASTER OF PHILOSOPHY

DIVISION OF INFORMATION ENGINEERING

THE CHINESE UNIVERSITY OF HONG KONG

December 1997

# Acknowledgement

I would like to take this opportunity to express my sincere gratitude to my supervisor, Prof. K.W. Victor Wei, for his guidance. I have learnt a lot from him. Prof. Wei is a demanding supervisor. He set high standard for me but because of this I have grown both in research ability and in self-confidence. His innovative ideas and extensive technical knowledge give much inspiration to me. It has been rewarding, fruitful and above all happy to work with him.

I also want to thank all the colleagues in the Information Integrity Laboratory. They provide a healthy and friendly working environment for me. Especially thanks go to S. L. Cheng, C. Lai and L. Wong for their wonderful source of both friendship and useful technical interactions. In addition, my gratitude also goes to Derek Chan, K. Y. Chan, Joe Siu, Andy Yuen, Albert Sung, Freeman Hui and Samantha Chan for their friendship. They give many valuable advice to me and share many of their experience to me.

# 簡介

圖像和視像壓縮在現今變得愈來愈重要。現今有許多壓縮技術，而其中小波(Wavelet)壓縮可算是一個不可或缺的工具。在這論文中，所提及到的技術，都是建築在小波壓縮的範圍上。

一向以來， Vector Quantization 比 Scalar Quantization 表現得更好。但是傳統上 Image Transformation 的設計和 Quantization 的設計是分別考慮的。在這篇論文中，將會指出假若將這兩個設計一齊考慮會有更好的表現。我們將介紹一個這樣的例子，叫做 Vector Wavelet Transform (VWT)。另外，也會介紹它的改良版，叫作 Vector Wavelet Transform-Linear Predictive Coding (VWT-LPC)。這兩個技術，都是利用小波壓縮和 Vector Quantization ，而且還有很好的表現。

此外，這論文也會介紹一個主要用在小波壓縮上的 Bit Allocation Algorithm 。它稱爲 Inter-Band Bit Allocation (IBBA)。最外，這論文會介紹一個 Finite State Vector Quantizer 。它是用 Parent Vector 的 Code Book Index 作爲決定 State 的因素。它稱爲 Parental Finite State Vector Quantizer (PFSVQ)。之後也會有一些示範,結果和總結。

# Abstract

Image and video coding are becoming more and more important nowadays. There are many sophisticated image coding algorithms. Among them, wavelet coding is one of the most promising tools. It is expected that it will hold a more important role in the next generation's image coding algorithms. In this thesis, all the algorithms are based on wavelet coding.

By Shannon's rate distortion theory, vector quantization (VQ) has better coding performance than scalar quantization. In fact, vector quantization is a common technique that used by many image coding researchers. However, conventionally, we treat the design of image transformation and the design of image quantization as two separate problems. In this thesis, we will show that by considering these two problems together, both the efficiency of the VQ and the overall coding performance can be improved. We will describe an image coding algorithm, in which, the design of image transformation is related to the design of vector quantization. This algorithm is called Vector Wavelet Transform (VWT). However, there is a drawback in VWT. To further boost the coding performance of VWT, another coding algorithm called Vector Wavelet Transform-Linear Predictive Coding (VWT-LPC) is developed. It is similar to VWT except that LPC is using together with VWT in the new algorithm. It can avoid the drawback of VWT and at the same time, make the variance of the vectors smaller so that VQ can be done more efficiently. Both VWT and VWT-LPC use wavelet transform and vector quantization as the basic components in the algorithms. Experiments shows that VWT-LPC has better performance than many existing algorithms.

We will also describe a bit allocation algorithm for wavelet subbands. This algorithm is called Inter-band Bit Allocation (IBBA) algorithm. It aims at distribute a total number of bit among different subbands of wavelet coefficients. The last part of the thesis is about finite state vector quantizer (FSVQ). We developed a finite state vector quantizer, called Parental Finite State Vector Quantizer (PFSVQ) especially suitable for wavelet image coding. In PFSVQ, the current state of the current vector is determined by the codebook index of its parent vector. It is because there is correlation between a vector and its parent vector in wavelet subbands, therefore, in each state, the vectors have some similarities. As a consequence, VQ can be done in a more efficient way.

Chapter 1 of this thesis is an introduction to data compression and image coding. Chapter 2 is an introduction to subband coding and wavelet transform. Vector Quantization will be discussed in chapter 3. The VWT and VWT-LPC will be described in chapter 4. Chapter 5 is a description of the IBBA algorithm. Chapter 6 will introduce the PFSVQ. Finally, the simulation results and conclusion will be found in chapter 7 and chapter 8 respectively.

# Table of Contents

# List of Tables

# List of Figures

# Chapter 1

# Introduction to Data Compression and Image Coding

## 1.1 Introduction

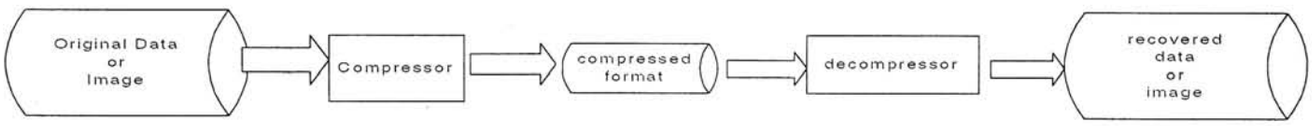



Fig. 1.1 Data compression and decompression model

Today, we are living in an information era. The advance in information technology changes our lives so much and makes our daily lives very convenient. Watching films at home comfortably or talking to our friends through the Internet are very common today. Information becomes easy to handle and process when it is in digital format. However, as the technology growths, the demand for faster communication speed, bigger data storage capacity, better data security also increasing tremendously [1]. The task of data compression is to reduce to the storage size or bit rate of an original data as much as possible. This can save storage space and make communication faster. Therefore data compression is an important area in information technology. The block diagram of data compression and decompression is shown in fig. 1.1.

The compressor compresses the original data into a compressed format which has small size compared to the original data. To recover the original data, a decompressor decompresses the compressed format. There are two categories of data compression, namely lossless compression and lossy compression. In lossless compression, the recovered or reconstructed data and the original data must be the same, i.e. without any error. On the contrary, lossy compression allows controlled amount of error in the recovered data. By carefully controlling lossy of fidelity, lossy compression can have higher compression ratio, ratio of the size of the original file to the size of the compressed format (1.1).

$$compression\ \ ratio = \frac{size\ \ of\ \ original\ \ data}{size\ \ of\ \ compressed\ \ format} \tag{1.1}$$

## 1.2 Fundamental Principle of Data Compression

If we have a source of data and is modelled as a random process $X$, the alphabet set of $X$ is defined as $\chi$. The entropy of $X$ is defined as

$$H(X) = -\sum_{x \in \chi} p(X = x) \cdot \log p(X = x) \tag{1.2}$$

The physical meaning of the entropy of $X$ is the minimum number of bits per symbol needed to represent $X$. It is also the theoretical lower bound of the average codeword length of $X$. The average codeword length of X is defined as:

$$L(X) = \sum_{x \in \chi} p(X = x) \cdot l(x) \tag{1.3}$$

where $l(x)$ is the codeword length or the number of bits of the alphabet $x$. It can be easily seen that equation (1.3) is actually a weighted sum of the codeword length of each alphabet by its probability of occurrence.

Lossless data compression, sometimes called **entropy coding**, is able to reduce the data rate and at the same time without introducing any error in the decompressed data. This is possible by using better organization of the data. That is to reorganize the original data into a more compact form. If we use few bits for more frequent alphabets or symbols, and more bits for less frequent symbols, by equation (1.3), it is expected that the overall bit rate (which is a weighted sum of the codeword length of each alphabet by its probability of occurrence) will be reduced. Therefore we come up with the fundamental principle of data compression:

**Encode more frequent symbols by shorter codewords**
**Encode less frequent symbols by longer codewords**

On the other hand, the principle of lossy data compression is different from that of lossless data compression. It is done by removing the redundancy of the data that is insignificant to people. As a consequence, the size of the compressed file can be reduced and at the same time the error is not easily noticeable by people.

## 1.3 Some Data Compression Algorithms

Huffman Coding and Arithmetic Coding are the lossless data compression algorithms that we used frequently in this thesis. Both are lossless compression algorithms and are commonly used today, for example in Joint Photographic Experts Group, JPEG and Motion Picture Experts

Group, MPEG. Huffman and Arithmetic Coding are also classified as variable rate entropy coding which means different symbols are coded with different codeword lengths.

## 1.4 Image Coding Overview

Digital image is a 2 dimensional array of real or complex numbers represented by a finite number of bits. Digital image takes up huge amount of storage space. For example, a 512 x 512 8-bit grayscale image needs 262144 Bytes to store. Color images and high definition images like medical images take up even more space. In digital video coding, most of the bits is used to code the intra-frame, which is a single frame of digital image. Better digital image coding algorithm can improve the coding performance of digital video. Therefore image coding is very important.

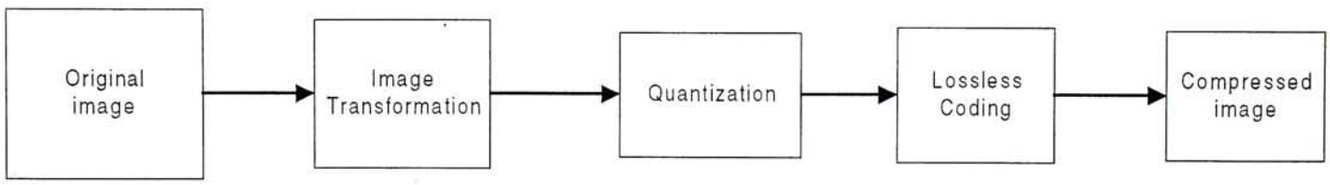



Fig. 1.2 Classical TQC model of image coding

Most of the data compression techniques can be applied to image coding. In addition, our human visual system is sensitive to noise in some regions of an image while insensitive to noise in other regions. For example, our eyes are more sensitive to noise in constant luminance region than that in the edge region. Therefore by carefully allowing some errors in the decompressed/reconstructed image, we can achieve higher compression ratio. This is lossy image coding. However, if there are too many errors in the reconstructed image, the visual quality will be adversely affected. Actually, there is a trade-off between high compression ratio and the amount of artifacts in the reconstructed image.

Fig. 1.2 shows the classical model of image coding, namely the Transform-Quantization-Lossless Coding (TQC) model. Obviously in some image coding algorithms not all the three components are present. If the second stage, quantization, is omitted, this is lossless image coding. It is because, in general, quantization is the only lossy part of the TQC model. (In fact, transformation can be a lossy part. For example in wavelet transform, if the filters are not perfect reconstruction filters, then the transform is lossy.) Lossless image coding is important in some areas, like medical imaging, where the reconstructed image needs to be identical to the original image.

## 1.5 Image Transformation

The first component of the TQC model is image transformation. It has been studied extensively during the last two decades and has become a very popular compression method for still-image coding [5]. The purpose of image transform is to transform an image from spatial domain that is difficult to compress to another domain that is easier to compress. The idea behind is to decorrelate the image pixels so that spatial redundancy can be remove more efficiently in the transform domain. Many transforms have the ability to pack a large fraction of the average energy of the image into a relatively few components of the transform coefficients, so that many of the transform coefficients will contain very little energy. This is known as the energy compactness of the transform. Some common transform algorithms are Discrete Cosine Transform (DCT) and wavelet transform. Fig. 1.3a shows the Lena image, and fig. 1.3b shows the wavelet coefficients of Lena image after 2-level wavelet transform using 9/7 biorthogonal filter. It can be easily seen that in wavelet domain (fig. 1.3b) the energies are grouped in a small number of coefficients on the left upper corner. This shows the energy compactness of the wavelet transform.

Fig. 1.3a 512x512 8-bit Lena image

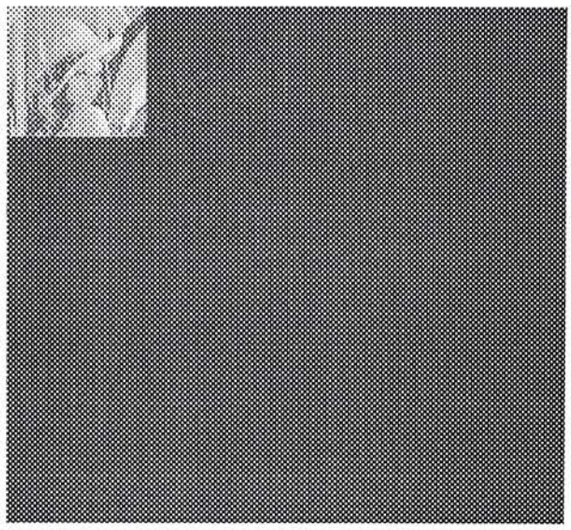


Fig. 1.3b Wavelet coefficient of Lena image

The first step in image transformation normally is to subtract a "d.c." value or constant value from the pixels. We call this step as Amplitude Shift. The d.c. value equals to the $2^{B-1}$, where $B$= number of bits per pixel of the original image. This is equivalence to shift the pixels from $[0, 2^B-1]$ to signed integers with range $[-2^{B-1}, 2^{B-1}-1]$. The amplitude shift affects the overall coding performance. We will show this effect through some simulations and comparisons in chapter 7. The next step is to split the image into $NxN$ non-overlapping block. The transformation can be viewed as a linear, separable, unitary matrix operation on the block $b$ using a transform $NxN$ unitary matrix $A$ and obtaining a $NxN$ coefficients c by:

$$c = AbA^T \qquad (1.4)$$

The reconstruction of block $b$ is done by reverse the transformation process. Mathematically, this is described as (1.5):

$$b = A^{*T} cA^{*} \qquad (1.5)$$

Note that since $A$ is unitary, therefore $AA^{*T}=I$ and $A^{T}A^{*}=I$. For example in $NxN$ discrete cosine transform (DCT), the transform matrix $A=\{a(k,n)\}$ is defined as:

$$a(k,n)=\begin{cases}\dfrac{1}{\sqrt{N}}, & k=0;\ \ 0\le n\le N-1\\ \sqrt{\dfrac{2}{N}}\cos\dfrac{\pi(2n+1)k}{2N}, & 1\le k\le N-1,\ \ 0\le n\le N-1\end{cases} \qquad (1.6)$$

The operation of wavelet transform is similar to DCT. In chapter 2 contains more discussions on wavelet tranform.

## 1.6 Quantization

With appropriate choice of which transform coefficients to be coded and which to be removed, high compression ratio with relatively high peak signal to noise ratio can be achieved. This is the idea of the second stage, Quantization. Peak Signal to Noise Ratio (PSNR) is the method to assess the likeliness of the original and the reconstructed images. PSNR is given by:

$$PSNR=10\log\left[\frac{1}{MN}\cdot\sum_{x=0}^{N}\sum_{y=0}^{M}\frac{A^{2}}{\left[f(x,y)-\hat{f}(x,y)\right]^{2}}\right] \qquad (1.7)$$

where $NxM$ = image dimension, $A$= image magnitude, $f(x,y)$= original image pixel, f^(x,y)= reconstructed image pixel. Another quality assessment method is the Mean Square Error (MSE). It is given as

$$MSE=\frac{1}{MN}\sum_{x=0}^{N}\sum_{y=0}^{M}(f(x,y)-\hat{f}(x,y))^{2} \qquad (1.8)$$

Quantization is an many to one mapping and is the lossy part of any image coding systems. It can remove the subjective redundancy of the pixels or coefficients. The output of quantization is an index. In order to achieve compression, the number of possible values of the input must be bigger than the number of output indices. Input can be one coefficient or image pixel. In this case, this is know as scalar quantization. If the coefficients or pixels are grouped together into a vector and outputs an index for each vector, this is known as vector quantization. Chapter 3 contains more descriptions on vector quantization

## 1.7 Lossless Coding

To further compress the image, lossless coding is used to code the indices of the quantization output. Statistical redundancy of the indices is removed in this stage. Commonly used algorithms are Huffman Coding and Arithmetic Coding.

# Chapter 2

# Subband Coding and Wavelet Transform

## 2.1 Subband Coding Principle

Signal expansion or decomposition was first appeared in 19th century since Fourier invented harmonic trigonometric series [26][30] and since then many kinds of signal expansion appeared. Subband coding (SBC) and wavelet coding is a kind of signal expansion. They were found to be very useful in signal processing and compression [7][23][30]. We will discuss SBC in this section and wavelet transform in section 2.4.

Subband coding was first introduced for speech coding [18], and was then applied in image coding. The aim is to decompose an image into several frequency bands by filter bank and allocate different number of bits to each band according to the human visual perceptual criteria. For simplicity, we will first consider the case of one-dimensional signals. SBC for images is same as for one dimensional signals, except the filters are becoming 2 dimensional filters or simply apply the one dimensional filters to each row and column of the image separately if the filters are separable.

Fig. 2.1 shows the block diagram of a subband coding system. In the encoder, the original signal $x[n]$, with length $N$, passes to a bank of $K$ analysis filters $H_1$, $H_2$, ..., $H_{K-1}$, $H_K$ . Usually, $H_1$ is a low pass filter and $H_K$ is a high pass filter, filter $H_2$, $H_3$, ... $H_{K-1}$ are band-pass filters. Then each output of the filters is decimated by a factor of $K$. As a result, a set of $K$ signals each of length $N/K$ is obtained. Finally, $Q_1$, $Q_2$, .., $Q_K$ different quantizers are used to quantize the decimation output. In the decoder, reconstruction is done firstly by $R_1$, $R_2$, ..., $R_K$ on each of the $K$ encoded signals and then interpolation by a factor of $K$ and filtering by synthesis filters $F_1$, $F_2$, ..., $F_K$ is done in sequence and finally the $K$ output reconstructed signals are added together to form the reconstruction signal $\hat{x}[n]$.

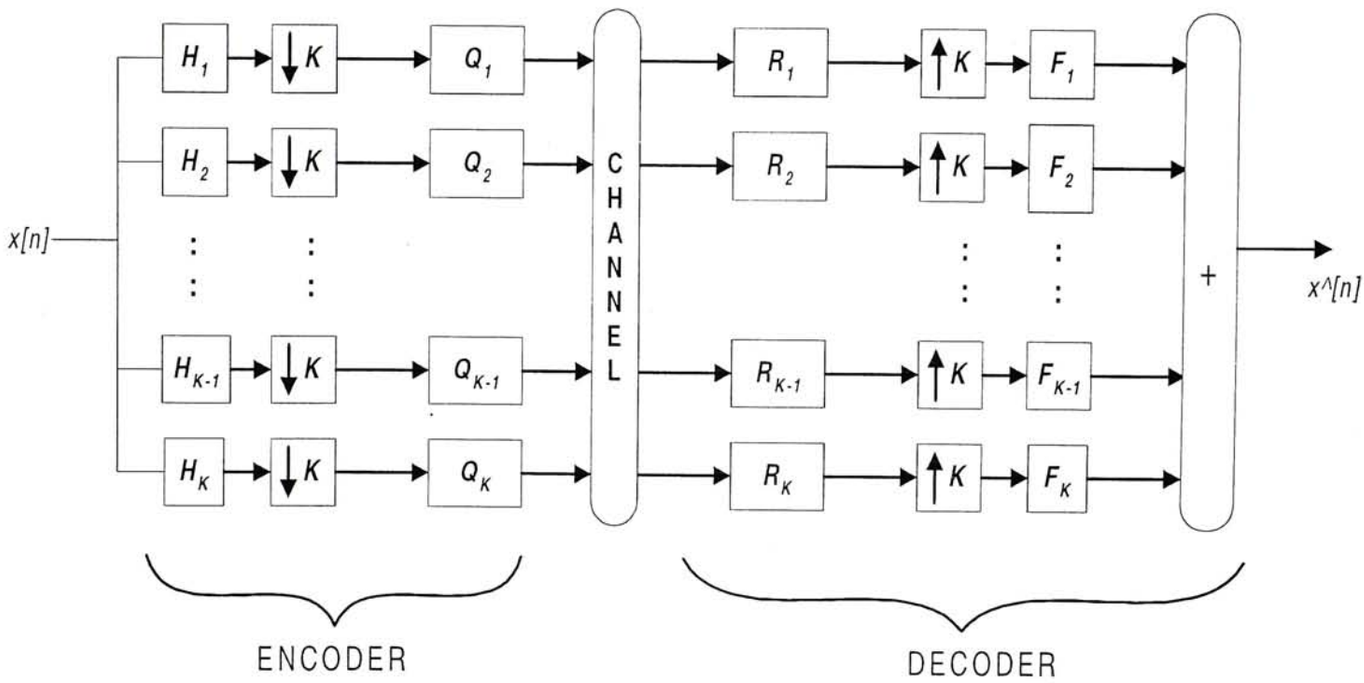



Fig. 2.1 Block diagram of subband coding system

## 2.2 Perfect Reconstruction

In addition to the quantization error, the analysis and synthesis filter banks can also cause $\hat{x}[n]$ differ from $x[n]$. The inability of perfect reconstruction caused by the filter banks can lead to frequencies distortion, phase distortion and aliasing in the reconstructed signal. Therefore perfect reconstruction filter banks are important in SBC.

To study perfect construction, first we consider the simple case of two bands SBC system without quantizers. Therefore fig. 2.1 is simplifed as fig. 2.2.

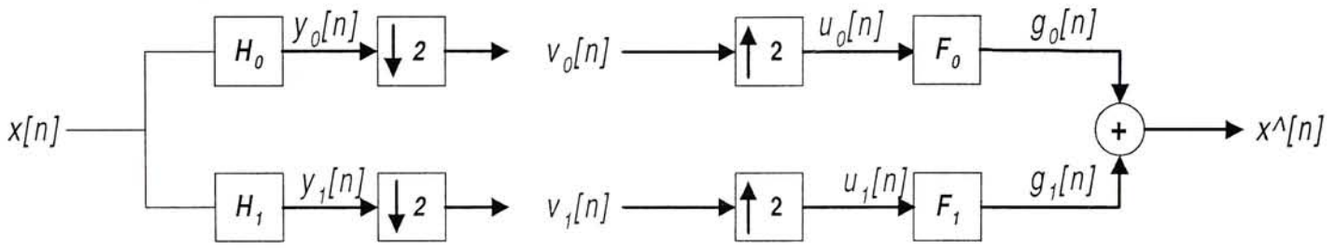



Fig. 2.2 Two bands subband coding system

$Y_0(Z)$ and $Y_1(Z)$ are the z-transform of the $y_0[n]$ and $y_1[n]$ respectively.

$$Y_0(Z) = H_0(Z)X(Z)$$
$$Y_1(Z) = H_1(Z)X(Z) \qquad (2.1)$$

$V_0(Z)$ and $V_1(Z)$ is the z-transform of the downsampling output of $y_0[n]$ and $y_1[n]$ respectively and are actually equal:

$$V_0(Z) = \frac{1}{2}\left( Y_0(Z^{\frac{1}{2}}) + Y_0(-Z^{\frac{1}{2}}) \right)$$
$$V_1(Z) = \frac{1}{2}\left( Y_1(Z^{\frac{1}{2}}) + Y_1(-Z^{\frac{1}{2}}) \right) \qquad (2.2)$$

However, the interpolation result, $u_0[n]$ and $u_1[n]$ equals:

$$U_0(Z) = V_0(Z^2) = \frac{1}{2}( \ Y_0(Z) + Y_0(-Z) \ )$$
$$U_1(Z) = V_1(Z^2) = \frac{1}{2}( \ Y_1(Z) + Y_1(-Z) \ ) \qquad (2.3)$$

Now, consider the reconstruction signal, $\hat{x}[n]$. It equals the summation of $g_0[n]$ and $g_1[n]$ and the z-transform is equal to:

$$\hat{X}[Z] = G_0(Z) + G_1(Z)$$
$$= U_0(Z)F_0(Z) + U_1(Z)F_1(Z) \qquad (2.4)$$

Now, substitute equation (2.3) into equation (2.4):

$$\hat{X}(Z) = \frac{1}{2}[Y_0(Z) + Y_0(-Z)]F_0(Z) + \frac{1}{2}[Y_1(Z) + Y_1(-Z)]F_1(Z)$$
$$= \frac{1}{2}[H_0(Z)F_0(Z) + H_1(Z)F_1(Z)]X(Z) +$$
$$\frac{1}{2}[H_0(-Z)F_0(Z) + H_1(-Z)F_1(Z)]X(-Z) \qquad (2.5)$$

It can be easily seen that the first term represents perfect reconstruction and the second term is the aliasing term. Therefore if the second term is eliminated, perfect reconstruction can be achieved.

A famous type of filter banks called Quadradure Mirror Filter (QMF) satisfies the above conditions. The requirement of QMF is as follows:

$$h_0[n] = h[n] \qquad h_1[n] = (-1)^n h[n]$$
$$f_0[n] = 2h[n] \qquad f_1[n] = -2(-1)^n h[n] \qquad (2.6)$$

where *h[n]* is a FIR, symmetric, even-length low pass filter of length *N*. More details about QMF can be found in [9][30].

## 2.3 Multi-Channel System

The above QMF is classified as two-channel system. Multi-channel systems can be built by using one low pass filter, one high pass filter and several band pass filters. Alternatively, it can be built by using cascade of several sets of two-channel systems. Fig. 2.3 shows an example of 4-channel system built by cascading of two-channel systems. The output bands are *LL, LH, HL* and *HH*. This kind of filter band is also called uniform band decomposition system. On the other hand, in some cases, further decomposition is done only in the lower band and leave the higher bands unchanged (fig. 2.4). This is called octave band decomposition.

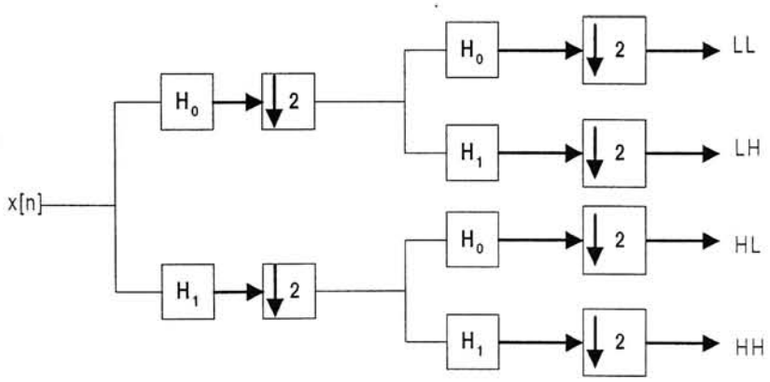



Fig. 2.3 Cascade of 2-channel system

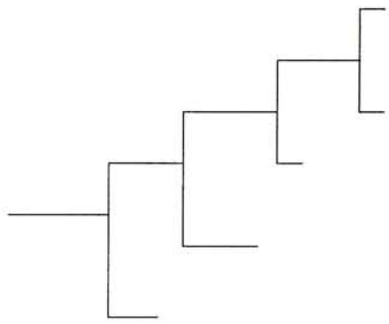


Fig. 2.4 Octave-band decomposition

## 2.4 Discrete Wavelet Transform

Discrete wavelet transform (DWT) or wavelet transform (WT) is a new generation of image coding technique. It is identical to subband coding, except it uses special choice of filters. Therefore sometimes wavelet coding is also called wavelet subband coding. The filters should be linear phase, orthogonal or biorthogonal, short, regularity, frequency selective [30]. Linear phase filters can be used in cascade to form pyramidal structures without the need for phase compensation [7]. In wavelet transform, octave band decomposition is usually used and an image is decomposed in row and then in column in sequence if the wavelet filter is separable. For example, in fig. 2.5, an image is decomposed into four subbands in the first level of decomposition. The four subbands are $LL_1$, $LH_1$, $HL_1$, $HH_1$. For example, $LL_1$ means the row and column are low frequency and the subscript 1 means in the first level of decomposition. Similarly $LH_1$ means column in low frequency and row in high frequency in the 1st level. Further decomposition only happens in the lowest (or coarsest) subband (in case of octave band decomposition), i.e. the $LL_1$, and obtains another 4 subbands called $LL_2$, $LH_2$, $HL_2$, $HH_2$.

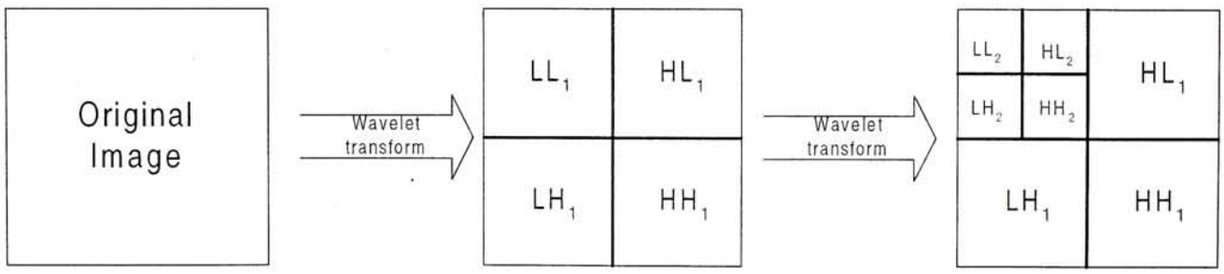



Fig. 2.5 Two levels octave band decomposition of image using wavelet transform

In most of the traditional signal decomposition techniques, like discrete cosine transform (DCT), the decomposed signals are localized in frequency. Wavelet transform has the additional properties of localization in space [16]. This is very important because typical images have most of energy localize in low frequencies and the rest energy of the high frequencies mostly come from edges. Wavelet transform provides the ability of compacting most energy into few low-

frequency coefficients and representing high-frequency energy in a few, spatially-clustered high frequency coefficients [17][23]. Figure 2.6 is an example of two-level wavelet transform of Lena image. It can be seen easily that most energy is concentrated in those coefficient in the $LL_2$ frequency band and at the same time the each band is also spatially localized.

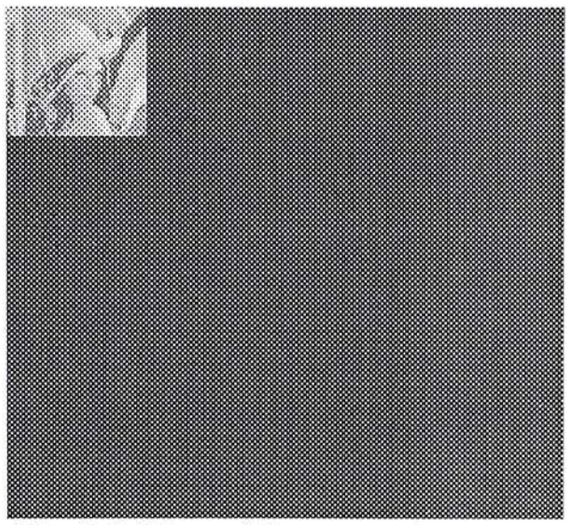


Fig. 2.6 Most of the energy is compacted in coefficients in $LL_2$ subband

In addition to have good compression performance, wavelet transform is also suitable for progressive transmissions and multi-resolution applications. This thesis uses wavelet transform for image decomposition and most of the techniques used in this paper exploit the inter and intra band correlation of the wavelet subbands. More details will be introduced later. The filters that frequently used in this thesis are Daubechies 4 (*DB4*) orthonormal filter and the 9/7 biorthogonal filter.

# Chapter 3

# Vector Quantization (VQ)

## 3.1 Introduction

A new video coding standard MPEG-4 is targeted to release in November 1998. In the new standard, the bit rate is targeted in two regions. They are the Very Low Bit Rate Core (VLBV Core) at 5-64kbit/s and the High Bit Rate Core (HBV Core) at 64-2Mbits/s [5]. Very low bit rate image coding is an important area in today's video coding technology. In order to achieve such a low bit rate, vector quantization (VQ) is a promising tool for use.

According to the result of Shannon's rate-distortion theory, the performance in coding a vector of information is better than that of coding a scalar of information, in term of rate-distortion criteria, even though the information source is memoryless. The performance of coding a vector of information improves as the vector dimension increases. This is because the correlation within the vector is exploited [27][6]. Image vector quantization bases on this theory to group image pixels or transform coefficients into vectors and gives a code or index for each vector. In next section, the basic procedure of VQ will be introduced. Section 3.3 gives a description of codebook searching and the LBG algorithm. Section 3.4 describes some problems in VQ and introduces two variations of VQ. The last section is about VQ on wavelet coefficients.

## 3.2 Basic Vector Quantization Procedure

To vector quantize an image, the image is first partitioned into non-overlapping blocks of dimension *N*x*N*. Each block represents a vector in the $R^{N^2}$ Euclidean space. The vector quantizer compares each input vector with a set of standard code vectors, called codebook, to find the best match. An index from the codebook identifying the best match (called codeword) is the final output of the quantizer. Mathematically, codeword $i$ is the output if:

$$d(v,r_i) \le d(v,r_j) \qquad for \quad i \neq j, \quad i,j \in [1,M] \qquad (3.1)$$

where $v$= input vector, $r_i$, $r_j$ are code vectors in the codebook, $M$= size of the codebook and $d$ is any distortion measurement such as Mean Square Error (MSE). To reconstruct the image at the receiver end, each codeword $i$ in the input stream is replace by the code vector $r_i$ in the codebook and the image is reconstructed block by block. Therefore, the bit rate of the coded image is equal to $[log_2(M)]/(N\text{x}N)$ bits/pixel. The encoding and decoding procedure is summarized in fig 3.1.

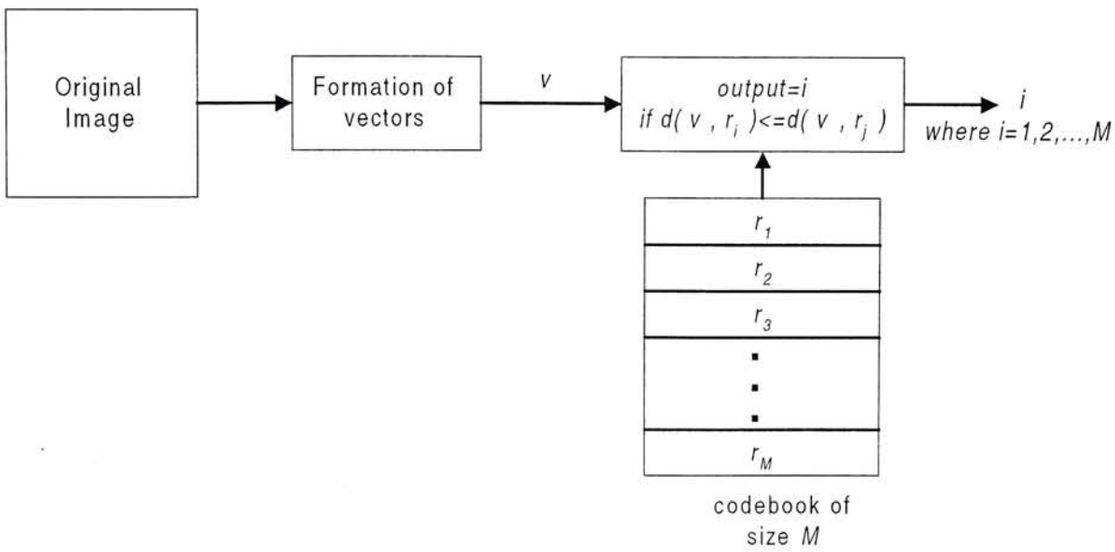



Fig. 3.1a Vector quantization encoding procedure

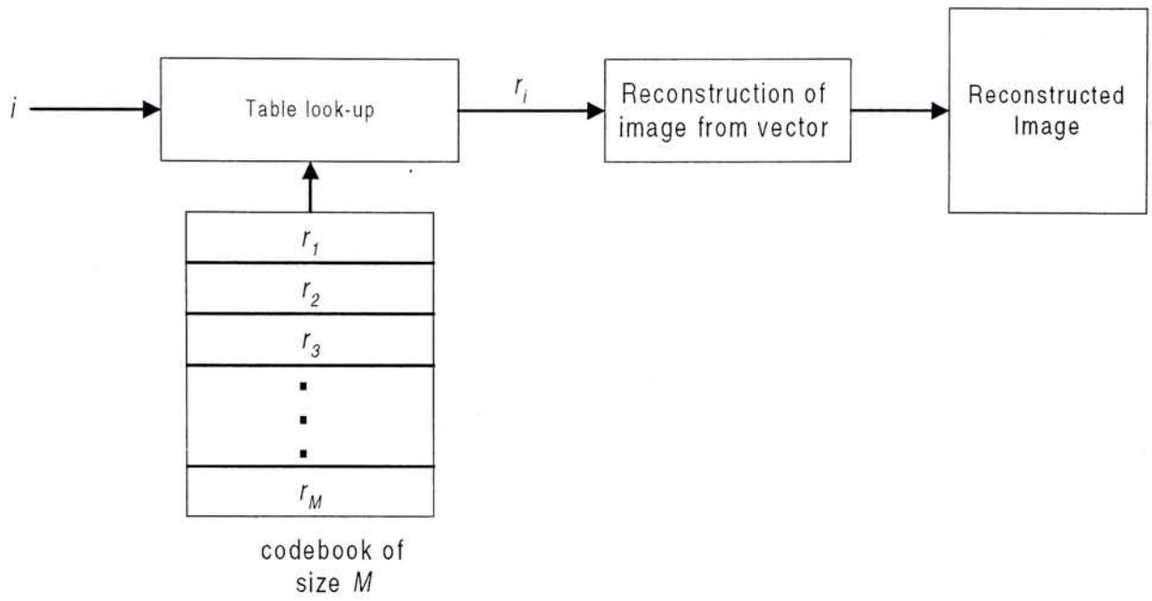



Fig. 3.1b Vector quantization decoding procedure

## 3.3 Codebook Searching and the LBG Algorithm

### 3.3.1 Codebook

The VQ routine involves the comparisons of each input vector to all code vectors in the codebook. Therefore, the ratio-distortion performance of VQ is highly affected by the codebook. In fact, codebook searching is the most important and computationally intensive part of the whole VQ procedure. The algorithms for finding the codebook are called clustering algorithms because this involves grouping vectors together (from a set of vectors, called training set) to form Voronoi cells and find the best representation for each cell as the code vectors.

Obviously, the best training set for coding an image is all the vectors of the same image. In this case, the codebook obtained is called local codebook. However, the use of local codebook is not efficient because a new codebook is need to be found for each image to be coded and this process requires much memory and computational power. In addition, the transmitter needs to

transmit the codebook to the receiver. This adversely affects the efficiency of the transmission. For example if the block size is 4x4 (i.e. vector dimension is 16), codebook size is 1024, image dimension is 512x512 and each pixel is 8 bit, then the number of bits per pixel needed for transmit the codebook equals:

$$\frac{N^2 \cdot M \cdot 8}{p^2} = 0.5 \quad bits / pixel$$

where $N$= 4, $M$= 1024, $p$= 512

However, the number of bits per pixel for coding the image equals:

$$\frac{\log_2 M}{N^2} = 0.625 \quad bits / pixel$$

The example reveals that the number of bits required for the transmission of side information, that is the codebook, is comparable to that used for the encoded image. Therefore local codebook reduces the efficiency of the VQ greatly.

On the other hand, instead of including the all vectors of the image to be coded in the training set, we can use a large enough set of vectors as the training set to obtain a statistically adequate description of any arbitrary images. The resultant codebook is called global codebook and is stored in a ROM in the encoder and decoder, therefore does not need to be transmitted.

### 3.3.2 LBG Algorithm

There are two main approach for codebook searching. The first approach uses some subset of a lattice to form a highly structured codebook. The VQ using this kind of codebook is referred as

lattice VQ [25][23][19][20][21]. The second approach is the vector clustering approach that we mentioned in section 3.3.1. In this thesis, all VQ used belongs to the second approach.

The most common and well-known vector clustering algorithm is the Linde-Buzo-Gray (LBG) algorithm (sometimes called generalized Lloyd algorithm (GLA)) [6][11][25]. The LBG algorithm for a codebook of size $M$ is summarized as follows:

Step 1. Given a training set $S$. Initialize $M$ code vectors ($r_1^{(1)}$, $r_2^{(1)}$, ..., $r_M^{(1)}$) for the codebook. Choose a distortion threshold $\varepsilon$ ($\varepsilon>0$) and a distortion measurement method $d$. Set the initial average distortion value $D^{(0)}$ to a very large number. Set counter $L=1$.

Step 2. For all vectors $v$ in $S$, find $d(v, r_i^{(L)})$. If $d(v, r_i^{(L)}) \leq d(v, r_j^{(L)})$ where $i \neq j$, $i,j \in [1, M]$, then group the vectors $v$ into Voronoi cell $C_i^{(L)}$ We therefore group all vectors into $M$ Voronoi cells and the total number of vectors in cell $i$ is denoted as $N(i)$.Then it is followed by updating the code vectors in the codebook: the new code vector $r_i^{(L+1)}$ is the centroid of Voronoi cell $C_i^{(L)}$, where $i \in [1,M]$. Mathematically, it is equal to:

$$r_i^{(L+1)} = \frac{1}{N(i)} \sum_{\forall v_k \in C_i^{(L)}} v_k$$

Find the average distortion value $D^{(L)}$:

$$D^{(L)} = \frac{1}{W} \sum_{i=1}^{M} \left( \sum_{\forall v_k \in C_i^{(L)}} d(v_k, v_i^{(L)}) \right)$$

where $W$ = total number of vectors in $S$.

If

$$\frac{D^{(L-1)} - D^{(L)}}{D^{(L-1)}} \leq \varepsilon$$

finish and the final code vectors are $r_i^{(L+1)}$, $i=1,2,...,M$.

Else go to step 3.

Step 3. Set $L=L+1$ and go to step 2.

In order to make the training set having good descriptions to any images, normally, the number of vectors required in the training set $S$ is at least $40M$ ($M$=codebook size) [24].

## 3.4 Problem of VQ and Variations of VQ

The formation of codebook requires a lot of vector comparisons and hence requires a lot of memory and computational power. This is the main drawback of VQ. In addition, the LBG algorithm is actually a minimization process, however, it can only find the local minimum [28]. Therefore the first step, initialization of code vector is important. There are several methods to initialize the codebook [28][29][11]. The VQ used in this thesis uses randomly chosen vectors from the training set as the initial codebook.

Here we are giving an example of VQ. The image under consideration is 512x512, 8-bits Lena image. Fig. 3.2a is the reconstructed image. The compression is at 0.5 bits/pixel with PSNR= 31.29 dB. Fig 3.2b is a close-up of fig. 3.2a. It can be easily seen that the perceptual quality of the reconstructed image is good. However, there is blocking effect and staircase effect, around the edges of the reconstructed image.

Fig. 3.2a Lena image after VQ at 0.5bpp

Fig. 3.2b Close-up of Lena image after VQ at 0.5bpp

To reduce the complexity, smaller vector dimensions can be used. However, the aim of VQ is to exploit the correlation among pixels or coefficients. Reducing the vector dimension means sacrifices the intra-vector dependency. Another approach to reduce complexity uses constrained VQ which impose some constraints to the codebook. There are many constrained VQ [25][6]. In addition to reduce the computation complexity, some also have other improvements, such as reduce saw-tool effect around edges. Here, we will discuss two commonly used constrained VQ.

## 3.4.1 Classified VQ (CVQ)

In classified VQ (CVQ), the vectors in the training set are first classified into $Q$ different classes according to the vectors' properties. For example, the classification can base on texture properties, edge directions, features of diagnostic or scientific importance or irrelevance, perceptual masking, or a variety of other criteria [23]. Then LBG is done on each class, forming $Q$ different codebooks. This method can reduce the computational complexity since for each

incoming vector, it only needs to compare with the vectors of the same class. For instance, by classifying the vectors into shades and edges, the saw-tool effect around the edges can be reduced. In the encoder, the incoming vector is first classified into a class, then it is compared with the code vectors in the codebook of the corresponding class. Therefore, two parameters are needed to transmit for each vector; one is the index specifying the codebook being used and the other is the index of the code vector being used. A block diagram of the CVQ is shown in fig. 3.3.

### 3.4.2 Finite State VQ (FSVQ)

In classified VQ mentioned in the last session, side information specifying which codebook being used is need to be transmitted for each input vector. To avoid the transmission of side information, finite-state VQ (FSVQ) is a choice for use. FSVQ is similar to CVQ. They both use multiple codebooks. However, no side information is needed for FSVQ. The side information is replaced by using next state rule. The next state rule can be based on, for example, the energy of previously decoded vector, or the texture properties of previously decoded vector. The important point is both the encoder and the decoder know the next state function which is used for deciding the choice of codebook. Chapter 6 of this thesis will introduce an innovative finite state VQ exploiting the inter-vector correlation between vectors in difference wavelet subbands.

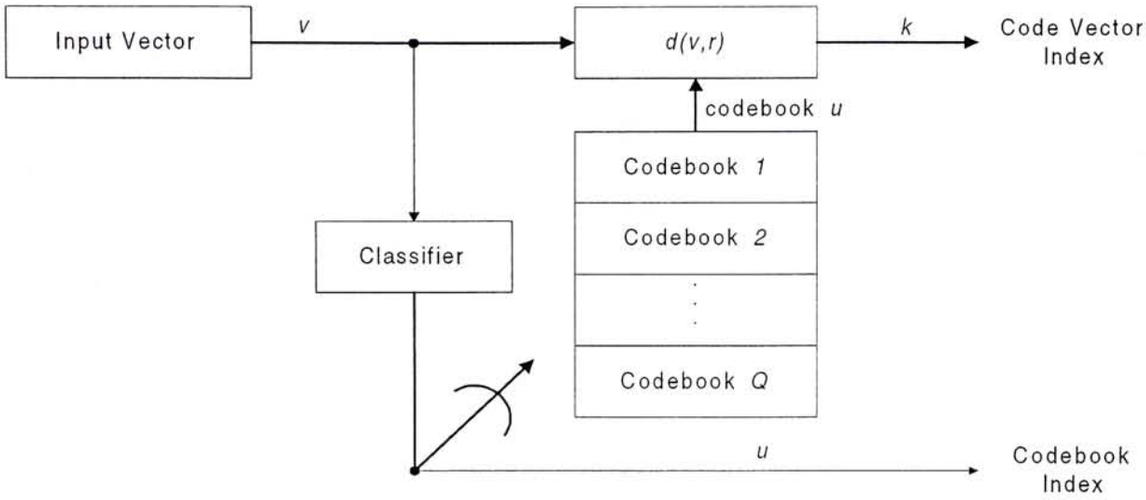



Fig. 3.3a Classified VQ encoding process

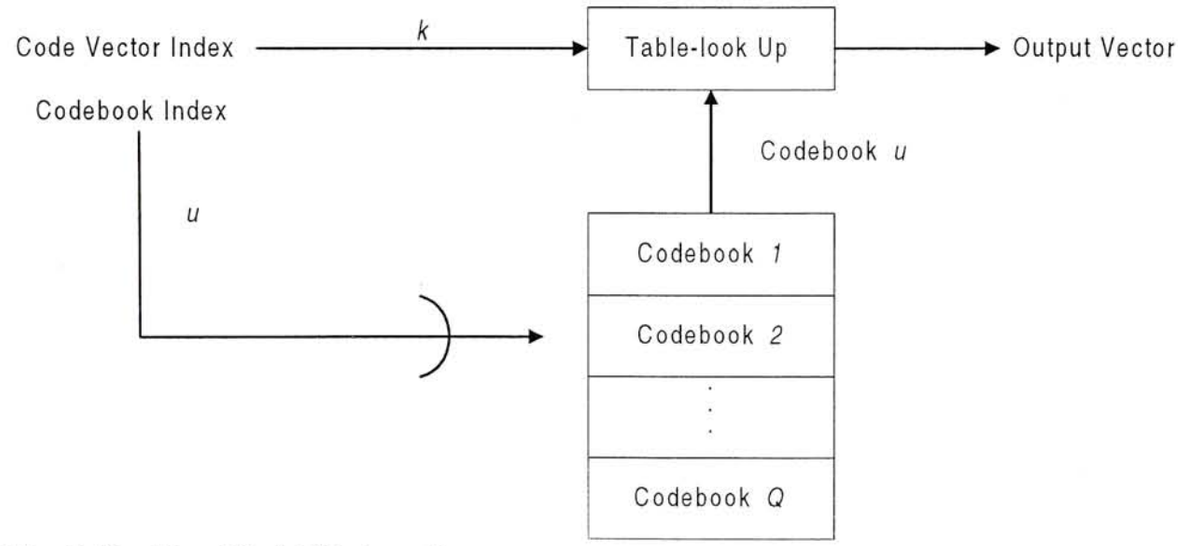



Fig. 3.3b Classified VQ decoding process

## 3.5 Vector Quantization on Wavelet Coefficients

Vector Quantization on wavelet coefficients or Wavelet VQ is classified as transform VQ. Instead of vector quantizing the pixels of an image, it uses the transform coefficients as the input of the quantizer. VQ is especially good for encoding wavelet coefficients. Actually, it has the

better performance in coding wavelet coefficients than to code image pixels [22]. In addition, we can also exploit the correlation of the wavelet coefficients between different wavelet subbands and within the same wavelet subband to boost the performance. We will discuss improvements based on these correlations in Chapter 4. By modifying the image transformation into a way that 'fits' the use of VQ, the traditional TQC model (see chapter 1) of image coding can be improved with the use of VQ. Next chapter will have a more detail description on these modifications.

# Chapter 4

# Vector Wavelet Transform-Linear Predictor Coding

## 4.1 Image Coding Using Wavelet Transform with Vector Quantization

In chapter one, we introduced the basic concept of image coding. The classical Transform-Quantization-Lossless Coding (TQC) model is a powerful method to encode an image. In chapter 2 and 3, wavelet and subband coding and vector quantization were introduced respectively. The last section of chapter 3 concluded that vector quantization is especially suitable for coding wavelet coefficients. It is expected that image coding using TQC model with wavelet transform and vector quantization together will give good coding performance over rate-distortion criteria. In fact, there are many literatures and new algorithms based on using wavelet transform and vector quantization together and they produce excellent performance.

### 4.1.1 Future Standard

The new video coding standard, ISO Motion Picture Expert Group Phase 4 standard (MPEG-4), aims at establishing a universal, efficient coding of different forms of audio-visual data, called

audio-visual objects which can be natural or synthetic. This aim will be achieved by defining two basic elements. The first is a set of coding tools for audio-visual objects capable of providing support to different functionalities such as object based interactivity and scalability, and error robustness, in addition to efficient compression. The second is a syntactic description of coded audio-visual objects, providing a formal method for describing the coded representation of these audio-visual objects and the method used to code them [20]. In real applications, most of the codec systems are asymmetric. That means the encoder does much computations than the decoder. For example the encoding complexity for digital TV is very high and is done by TV stations. However, users can enjoy the digital TV only by using simple decoding circuits in their TV sets. The VQ encoding and decoding system is also an asymmetric system because the decoding complexity is much lower than encoding complexity. There is a very high probability that the wavelet transform coding and vector quantization will be included in the coding tools.

### 4.1.2 Drawback of DCT

Discrete cosine transform (DCT) is being applied in many existing image and video coding standards, like JPEG for digital images, MPEG-1 for full-motion video with VHS quality, MPEG-2 for full-motion video at various quality levels including broadcasting, studio, and HDTV quality. DCT became popular since 1980s because of its low complexity and effective bit allocation can be done. Also, the evolution of Fast Fourier Transform (FFT) and Fast Cosine Transform (FCT) enables real-time application of DCT. DCT performs well at high bit rate. However, at moderate to low bit rate, it suffers severe blocking effects because the basis functions of DCT for reconstruction are short [9]. Fig. 4.1a shows a JPEG image of the Lena at 1.3 bits per pixel. The blocking effect of DCT can be seen easily in fig. 4.1b. It is the close-up of fig. 4.1a. The blocking effect makes the image looks blur.

### 4.1.3 Wavelet Coding and VQ, the Future Trend

Subband coding using wavelets, as mentioned in chapter 2, provides an efficient way to decompose an image because both the spatial and frequency aspects of the images are taken into considerations. Furthermore, wavelet transform can avoid blocking effect at medium bit rate because its basis functions have variable lengths [9]. Long basis functions represent flat background, i.e. low frequency. Short basis functions represent regions with texture. Although wavelet transform suffers ringing effect especially around edges at low bit rate, there is no doubt that wavelet transform will hold an important role in image and video coding.

Thanks to the development of electronic technology, the price of memory and computational power is greatly reduced when compared to the past. This helps the development of VQ. Today, we have many advance algorithms for VQ and many of them have very good performance. Therefore, wavelet subband coding and vector quantization will be one of the most powerful tools for further image and video coding technology. The algorithms introduced in this thesis all based on wavelet transform together with the use of vector quantization.

Fig. 4.1a JPEG image of Lena at 1.3 bpp

Fig. 4.1b Blocking effect of DCT coded image

## 4.2 Mismatch between Scalar Transformation and VQ

Image transformation is an indispensable part in image coding. Normally, it is done by separating the image by blocks and then transform each block one by one. This kind of image transform is scalar-based transformation. It works very well together with scalar quantizer and works better with vector quantizer. However, we can further boost the power of VQ by using image transformation that is vector-based. When people design scalar quantizers, they want to make it suitable for scalar-based transform. In other works, the development of scalar-based image transformation and scalar quantizers has been an interactive process [15]. On the contrary, the use of scalar-based image transform together with vector quantization does not consider the requirement of VQ. This means the design of scalar-based image transform and VQ is not an interactive process. Therefore, there is much room to boost the power of VQ by making the image transform 'fit' for VQ. Here we define this kind of image transformation as vector-based image transformation. Thus the development of vector-based image transform and vector quantizer is an interactive process. These relationships are shown in fig. 4.2. For simplicity, the term vector-based image transformation has the same meaning as vector transformation or vector transform (VT). Similarly, scalar-based image transformation and scalar transformation or scalar transform (ST) refers to the same meaning.

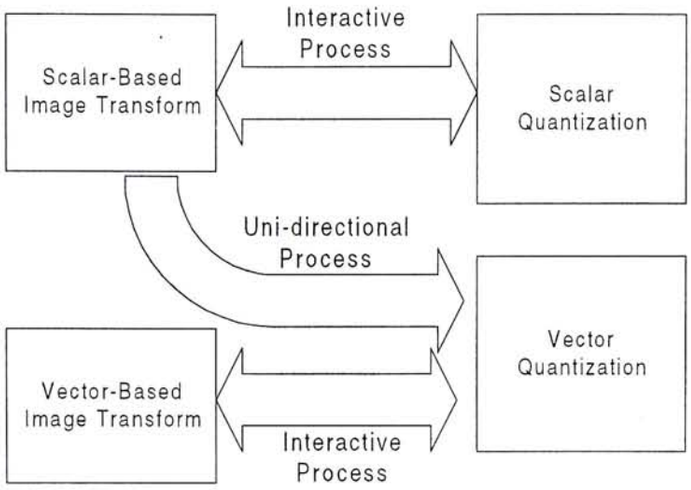



Fig. 4.2 Relationship between the design process of scale(vector)-based transform and scale(vector) quantizer

## 4.3 Vector Wavelet Transform (VWT)

Vector wavelet transform is a kind of vector transformation developed by Li and Zhang [15]. It has the properties and merits of vector transform. The implementation of vector wavelet transform is very simple. To make it easy to understand, it can be considered as a modification of scalar wavelet transform. The following block diagrams are the procedures of VWT. Equations showing how VWT works can be found in [15]. Here is the procedures of VWT:

Step 1: Subsample the original image at a rate of M vertically and horizontal to form MxM subimages.

Step 2: Each subimage is transform by wavelet filters.

Step 3: Regroup the wavelet coefficients (from the subimages) of the same frequency and spatial location into vectors of dimension MxM. And the vectors are input to the vector quantizer.

Fig. 4.3 shows these procedures.

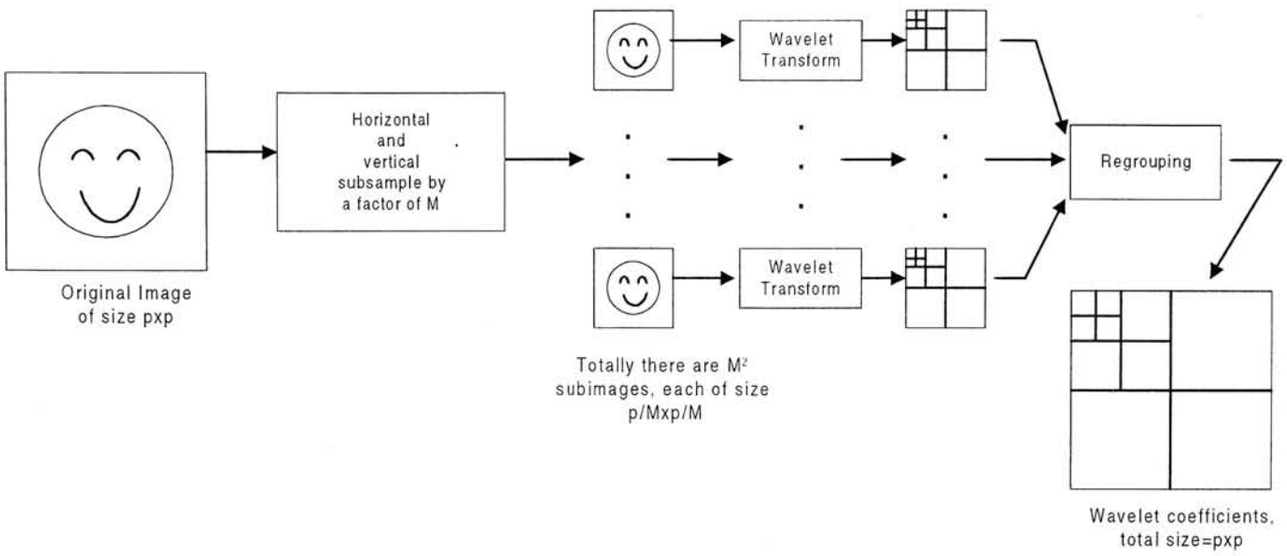



Fig. 4.3a Forward Vector Wavelet Transform

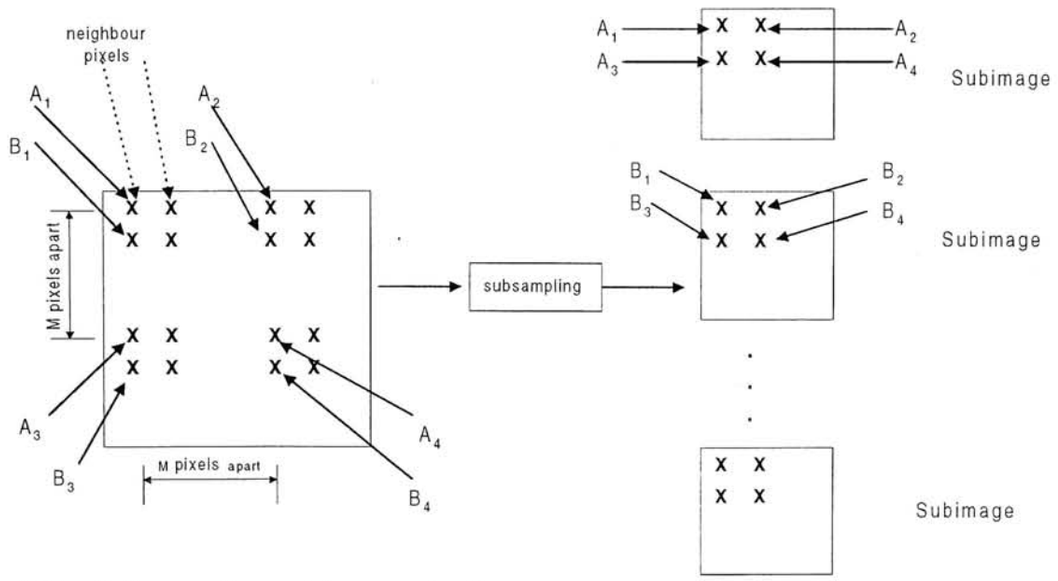



Fig. 4.3b Subsampling of image pixels

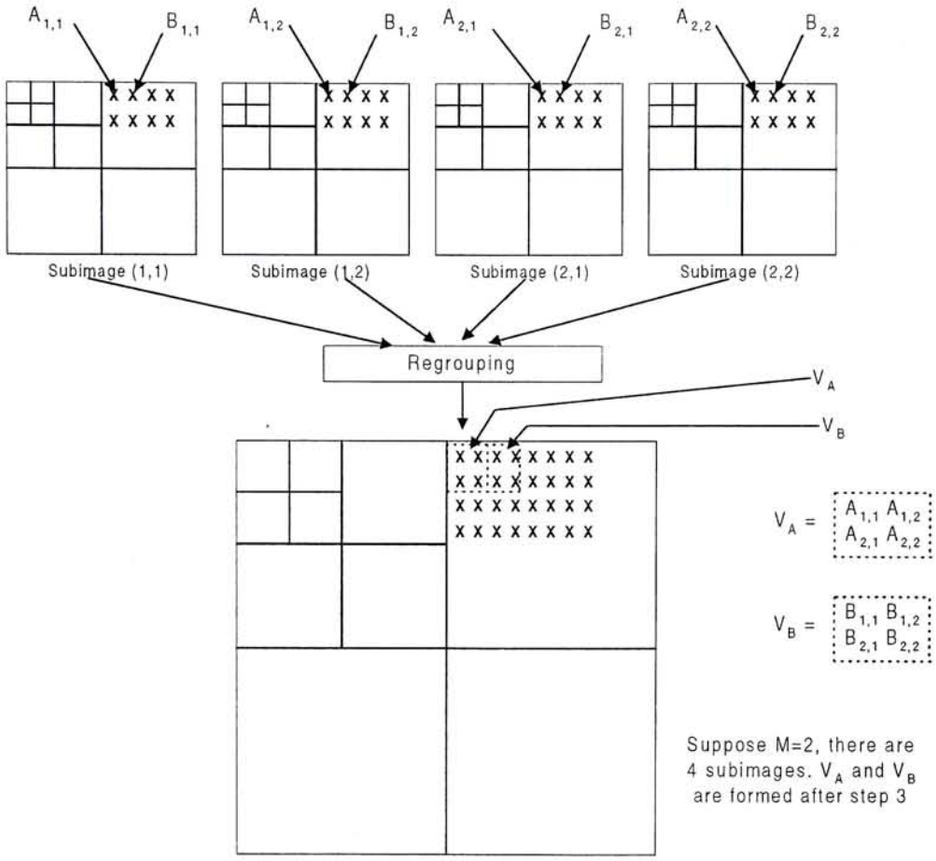



Fig. 4.3c Regrouping of wavelet coefficients from different subimages

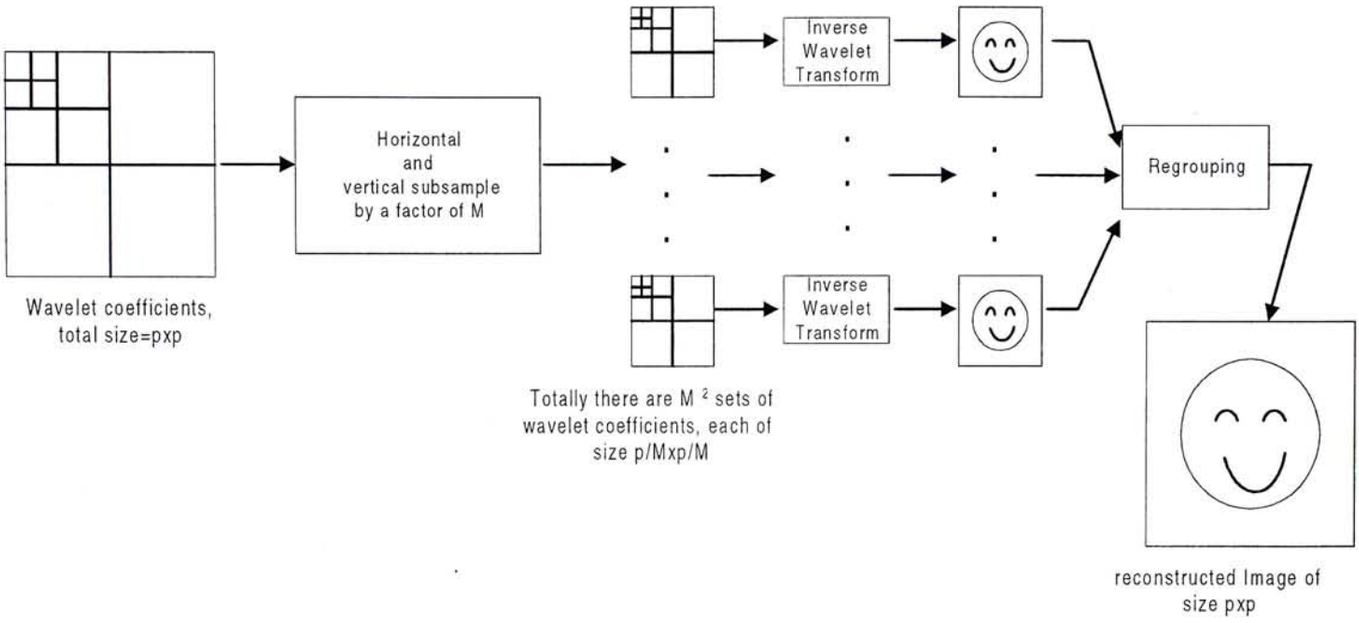



Fig. 4.3d Inverse Vector Wavelet Transform

The first step, image subsampling forms $M^2$ subimages. They look alike. However, within each subimage, the pixels are less correlated than that in the original image. Therefore after the wavelet transform in the second step, the correlation between coefficients within the same subimage is very low but the correlation between coefficients from different subimages at the same spatial and frequency location is high. ([12] provides evidence for above arguments.) For example, in fig. 4.3c, the correlation among $A_{1,1}$ , $A_{1,2}$ , $A_{2,1}$ , $A_{2,2}$ in $V_A$ and the correlation among $B_{1,1}$ , $B_{1,2}$ , $B_{2,1}$ , $B_{2,2}$ in $V_B$ are high. But correlation between elements in $V_A$ and elements in $V_B$ is low.

This is because the neighbour pixels within the same subimage are originally M pixels apart in the original image, therefore the wavelet coefficients within the subimage is lower than that in case of wavelet transform of the original image. On the contrary, the pixels at the same spatial location of neighbour subimages are actually neighourhoods in the original image. Therefore

correlation between wavelet coefficients from different subimages at the same frequency and spatial location is high.

The last step aims at producing vectors that has high intra-vector correlation and low inter-vector correlation. This is done by regrouping the wavelet coefficients from different subimages at the same frequency and spatial location to form vectors. Since in step 1, the subsampling rate is M (vertically and horizontal), therefore, now each vector has dimension MxM. This step can achieve our aim because the elements of each vector come from pixels of the original image of M pixels apart (horizontally or/and vertically).

The complete image coding procedure is finished by putting the vectors into vector quantizer, converting them to indices and finally losslessly code these indices using algorithms like Huffman coding or Arithmetic coding.

## 4.4 Example of Vector Wavelet Transform

To illustrate the use of vector wavelet transform, an example is given in this section using 512x512 8-bit Lena image. In our example, the subsampling rate M= 4. An subimage is formed by taking pixel alternately 4 pixels apart both vertically and horizontally. Therefore after step 1, there are 16 subimages (fig. 4.4a). Each subimage has dimension (512/4)x(512/4)= 128x128. These 16 subimages are then transformed by 9/7 biorthogonal wavelet filter in 2 levels. Each subimage is converted to 7 bands of wavelet subbands (fig. 4.4b). Finally, the wavelet coefficients are regrouped together (fig. 4.4c). Each 4x4 coefficients form a vector, which has high intra-vector correlation. Each vector is then VQ and losslessly coded.

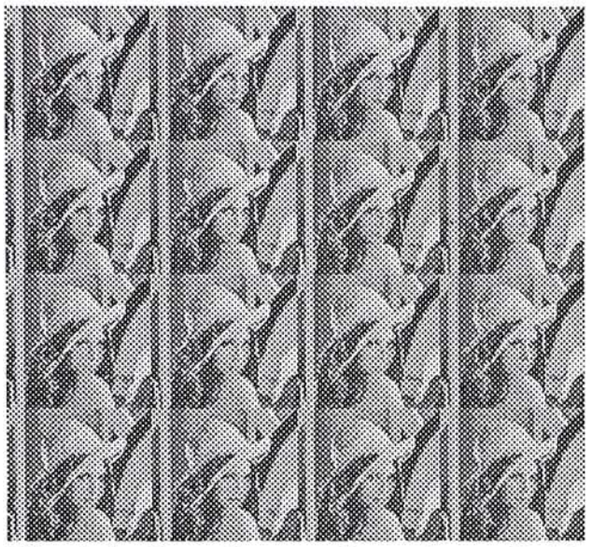


Fig. 4.4a Lena image after subsampled in rate of 4 horizontally and vertically

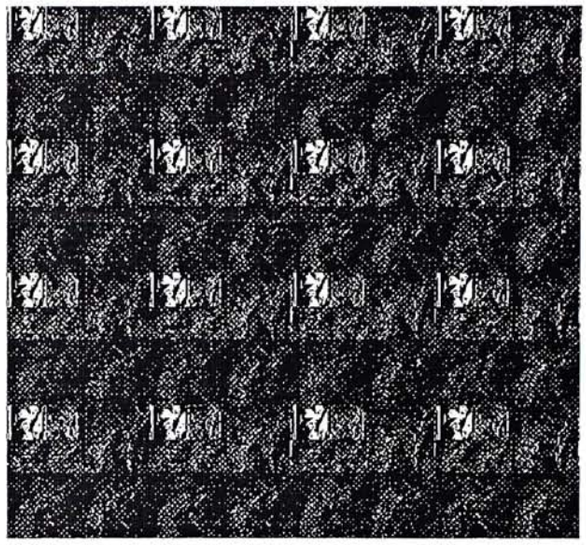


Fig. 4.4b Result of subsampled Lena image after wavelet transform using 9/7 biorthogonal filter in 2 levels

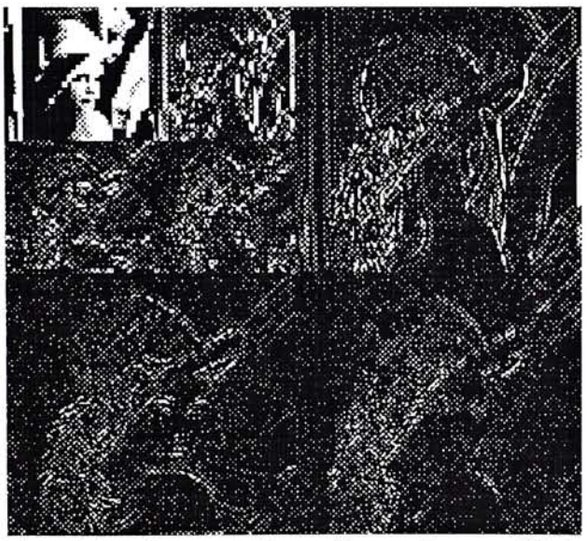


Fig. 4.4c Regrouping of wavelet coefficients from the subimages of fig. 4.4b

## 4.5 Vector Wavelet Transform - Linear Predictive Coding (VWT-LPC)

Vector wavelet transform, described in last section, successfully decorrelates the inter-vector correlation and preserves the intra-vector correlation. However, in the procedure of VWT, although the first step, subsampling of pixels helps the decorrelation of the inter-vector correlation, but at the same time, it produce much energy in the higher wavelet subbands after the wavelet transform in the second step. This is because the similarity of neighbour pixels in the subimages is smaller than that in the original image. This reduction in similarity causes the energy in the lower subbands becomes smaller and that in higher subbands becomes higher. In other words, the energy of the wavelet coefficients is less compact. This is shown in fig. 4.5.

These two figures compare the energy compactness of wavelet coefficients obtained by wavelet transform the Lena image directly and wavelet coefficients of VWT (i.e. from fig. 4.4c). The vertical axis is the amplitude of the wavelet coefficients and the horizontal axis is the scan line. The lower the wavelet subband the coefficients belong to, the more the coefficients near to the left side of the scan line. Fig. 4.5a is the scan line of wavelet coefficient of Lena image. Fig. 4.5b is the scan line of wavelet coefficient from VWT (from fig 4.4c). It can be easily seen that the coefficients in fig. 4.5b have higher amplitudes around the right side of the scan line (i.e. higher subbands). This means that the energy of the wavelet coefficient of VWT is less compacted.

Here we introduce a modification of VWT, called Vector Wavelet Transform-Linear Predictive Coding (VWT-LPC). The evolution of VWT-LPC is to alleviate the above problem by adding linear predictive coding (LPC) after vector wavelet transform. The LPC takes the vectors produced by VWT as its input and "filter" each vector by a linear predictor. The "filtering" can be viewed as taking difference between the vector and the predictor of the vector. The result is

called residual, which is also a vector (in our case) and is the output of the LPC. The residual is then vector quantized. The block diagram of VWT-LPC is shown in fig. 4.6.

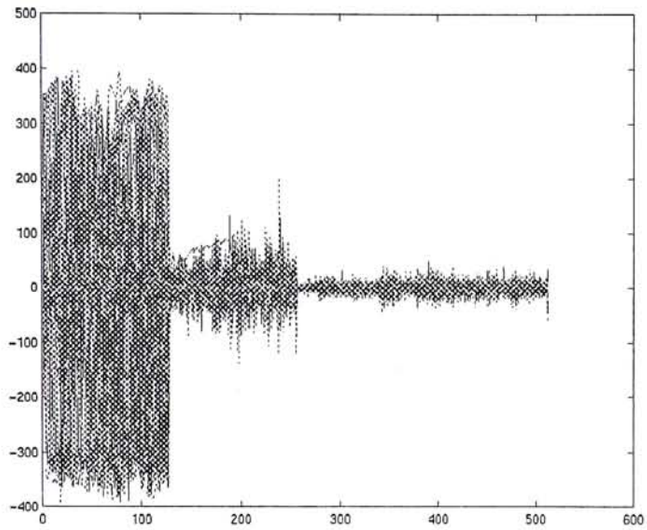



Fig. 4.5a Scan line of wavelet coefficient of Lena image

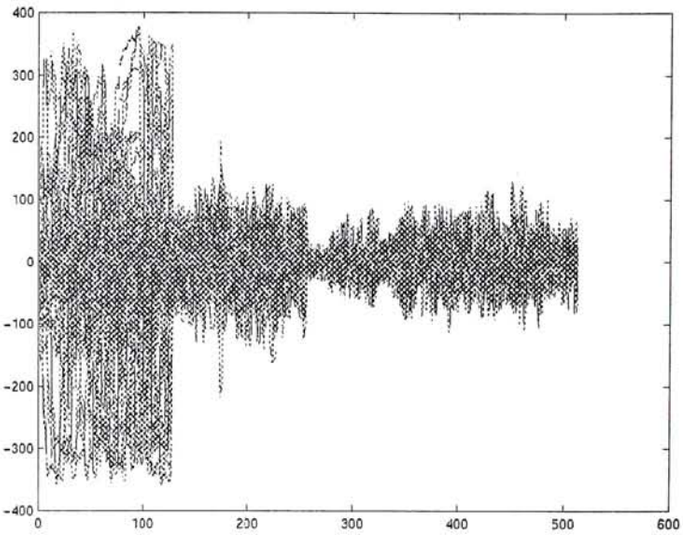



Fig. 4.5b Scan line of wavelet coefficient from vector wavelet transform

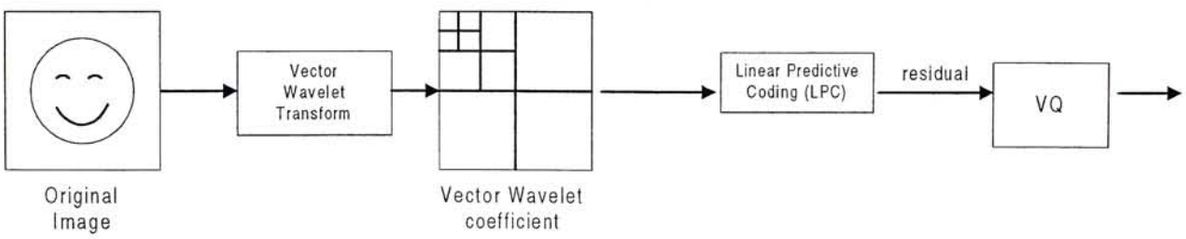



Fig. 4.6 Block diagram of Vector Wavelet Transform-Linear Predictive Coding

In this algorithm, the predictor of the LPC is carefully chosen so that the residual has smaller variance than the vector wavelet coefficients. In that case, all the residual vectors are more clustered in the $R^{M^2}$ Euclidean space (if vector dimension is MxM). As a result with the same number of code vectors, the codebook obtained with the use of LPC gives a better representation

to the vectors than that without the use of LPC. Therefore the rate-distortion performance will be improved by using LPC together with VWT. One may argue that the use of linear predictor on the vectors (from VWT) may reduce the intra-vector correlation. In previous section, we argued that higher intra-vector correlation will enhance the peformance of VQ. Actually, there is a trade-off between using LPC to reduce the variance and not using LPC to preserve the intra-vector correlation. Our simulation in chapter 7, shows that VWT-LPC has better performance than VWT at various bit rates.

## 4.6 An Example of VWT-LPC

VWT-LPC can be implemented using commonly used predictor and it still performs better than VWT. In this example, we choose the intra-vector mean as the predictor. Mathematically, the predictor $p_v$, (for vector $v$) can be written as:

$$p_v = \frac{1}{K} \sum_{x_i \in v} x_i \qquad (4.7)$$

where $K$= total number of coefficients in each vector $v$

The difference between each wavelet coefficient (of vector $v$) and the predictor (of the vector $v$) is taken and the output is a vector and goes to the vector quantizer. The predictor is coded separately (Fig. 4.7).

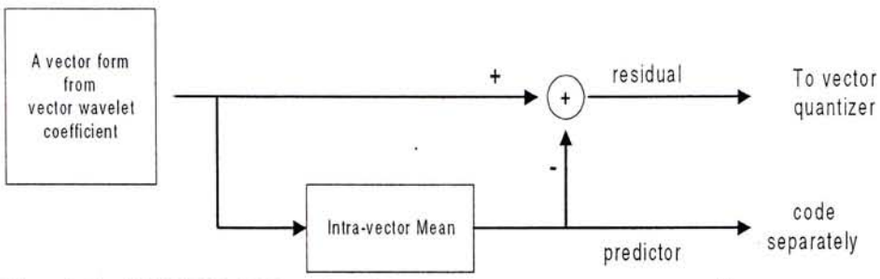



Fig. 4.7 VWT-LPC using intra-vector mean as predictor

Again we use 8-bit 512x512 Lena image in this example. VWT is first done on the Lena image. The wavelet filter used is 9/7 biorthogonal filter and wavelet transform is done in 2 levels, and obtains 7 subbands. The subsampling rate M=4. Fig. 4.8 is the residual after LPC. Comparing to fig. 4.4c, the coefficients in fig. 4.8 have smaller variance. Details of the coding performance of VWT-LPC can be found in chapter 7.

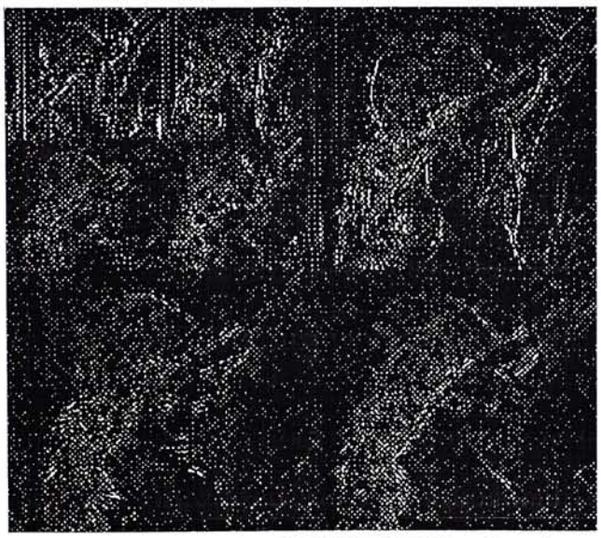


Fig. 4.8 Residual of VWT-LPC using intra-vector mean predictor

# Chapter 5

# Vector Quantizaton with Inter-band Bit Allocation (IBBA)

## 5.1 Bit Allocation Problem

In transform coding, the original image is decomposed into different frequencies or subbands. Each has different degree of significance to our human visual system. For the best reconstruction result, it is obvious to encode all the subbands losslessly. In this case, the encoding system is lossless. However, in many cases, the total number of bits available for the encoding is limited and /or we want to achieve higher compression ratio, therefore we need to have quantization.

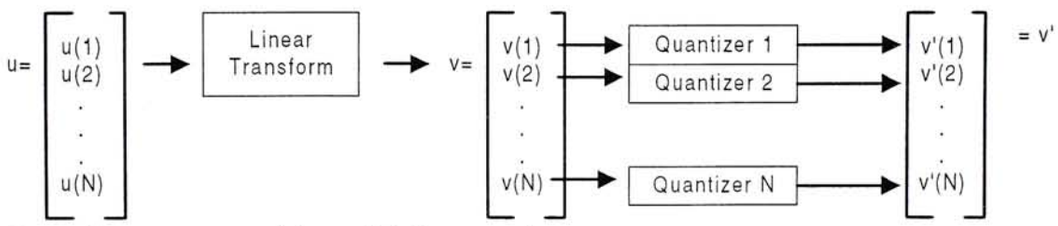



Fig. 5.1 Encoding system with multiple quantizers

Quantization is a lossy process that removes signal redundancy or removes any signal that is insignificance to the receiver. To effectively encode the coefficients according to their degree of

significance, an image coding system often contains several different quantizers each of which quantizes the transform coefficients in different accuracies (different number of bits) of reproduction. In particular, in subband coding, the variances of the transform coefficients of different subbands are generally unequal, therefore each subband requires a different number of quantization bits. The task of distributing a given quota of bits to various quantizers to optimize the overall coding performance is known as bit allocation. Suppose now the transform coefficient is denoted as $\{v_k: 0 \le k \le N-1\}$ and the number of bits assigned for coefficient $v_k$ *is* $b_k$ and $b_k \ge 0$ .Let $W_k(b_k)$ be the mean square error incurred in optimally quantize $v_k$ with $b_k$ bits. Then the problem of bit allocation is to minimize $D$ and under the constraint of (5.1), where $B$ is total bits available for the encoder.

$$D = \sum_{k=0}^{N-1} W_k(b_k)$$
$$B \le \sum_{k=0}^{N-1} b_k \qquad (5.1)$$

To solve the problem of bit allocation, one can solve the above equations directly. However, this requires exhaustive computations. In addition, the optimal distortion functions $W_k(b_k)$ are rarely known exactly and this will adversely affect the degree of the optimality [25]. Some researchers allocate the bits in arbitrary ways and some researchers develop other algorithms for bit allocation, for example the equal-slope method [23]. In this chapter, we will present a bit allocation algorithm for encoding vector-based wavelet transform coefficients [12]. We call it Inter-band Bit Allocation (IBBA).

In next section, we will describe the Inter-band Bit Allocation (IBBA) algorithm. In chapter 6, we will describe an algorithm that can improve the coding result of IBBA.

## 5.2 Bit Allocation for Wavelet Subband Vector Quantizer

### 5.2.1 Multiple Codebooks

In wavelet subband coding, the using a single codebook for all the subbands is not as good as using multiple codebooks, one codebook for each subband. This is because using multiple codebooks, each codebook can be trained to adapt the statistical properties of each individual subband. In addition, the number of vectors in each subband is less than the total number of vectors in all the subbands and therefore the complexity of codebook training is highly reduced. The vectors in each subband are more clustered in the $R^{N^2}$ Euclidean space (NxN is the vector dimension). Therefore the multiple codebooks give better representations to the vector in their corresponding subbands.

Using multiple codebooks for wavelet subband VQ coding is therefore better than using single codebook. These many of codebooks are trained once only, based on a set of training images. The encoding and decoding of each subbband do not require a search of more than one codebook because each band has its corresponding codebook. Therefore, multiple codebooks wavelet subband VQ does not create big efficiency problem.

### 5.2.2 Inter-band Bit Allocation (IBBA)

In this section, the word "bit allocation" is equivalent to "codebook size allocation" in multiple codebooks wavelet subband VQ. Each codebook has different size (i.e. different number of code vectors). This is necessary because each codebook represents different degree of importance to our human visual system. Larger codebook size is assigned to those more "important" codebooks and less "important" codebooks are assigned with smaller size. Hence, the total number of bits is allocated among different subbands of the same image. Here, we call this kind of bit allocation

method as Inter-band Bit Allocation (IBBA). It is first developed by Li and Zhang [12] for encoding vector-based wavelet transform coefficients. We are now introducing this bit allocation algorithm and apply it to scalar-based wavelet transform coefficients. The algorithm is as follows:

Let:

$d$ = index indicates wavelet decomposition level.

$(u,v)$ = index indicates the frequency band.

$R_d(u,v)$ = number of bits per pixel in subband $C_d(u,v)$

$D$ = total number of subband decomposition levels.

R = total number of bits available

(fig. 5.2)

If the subbands are in the same level of the wavelet pyramid, then they are allocated the same number of bits, that is:

$$R_d(0,1) = R_d(1,0) = R_d(1,1)$$

The bit allocated to the top level $R_0(0,0)$ is a factor $\alpha$ times of that for the next level $R_0(u,v)$.

The bit allocated to level d, $R_d(u,v)$ is a factor $\beta$ times of that for the next level $R_{d+1}(u,v)$, that is:

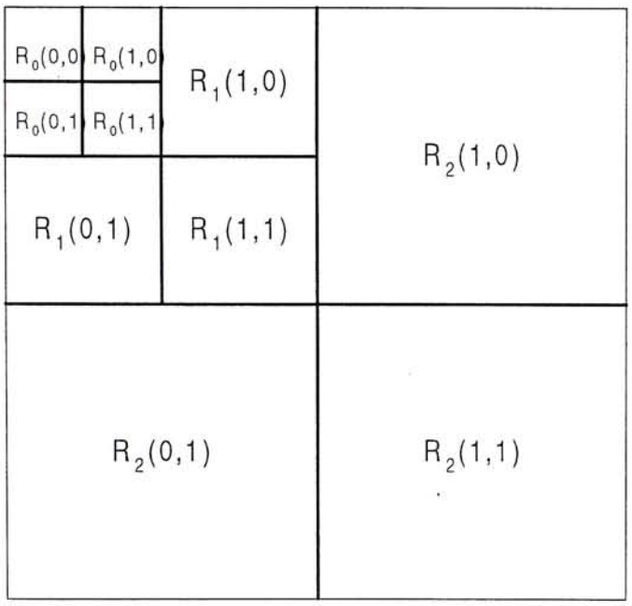
$$R_0(0,0) = \alpha \cdot R_0(u,v)$$
$$R_d(u,v) = \beta \cdot R_{d+1}(u,v)$$

The formula for bit allocation is [12]:

$$R = (\frac{1}{4})^D \cdot \beta^{D-1} R_{D-1}(u,v) \cdot \left[ \alpha + 3 \sum_{d=0}^{D-1} \left( \frac{4}{\beta} \right)^d \right]$$

For example, if we have 7 subbands, that is, we have two wavelet decomposition levels ($D$=2) or 3 levels of pyramids. The total bits for allocation, $R$= 0.25 bpp (bits per pixel). Let $\alpha$= 2 and $\beta$= 4, the bit allocation is therefore (fig. 5.3):

$$R_1(0,1) = R_1(1,0) = R_1(1,1) = 0.125 \quad bpp$$
$$R_0(0,1) = R_0(1,0) = R_0(1,1) = 4 \times 0.125 = 0.5 \quad bpp$$
$$R_0(0,0) = 2 \times 0.5 = 1 \quad bpp$$




D= 2

Fig. 5.2 Nomenclature of the bit allocation when D=2

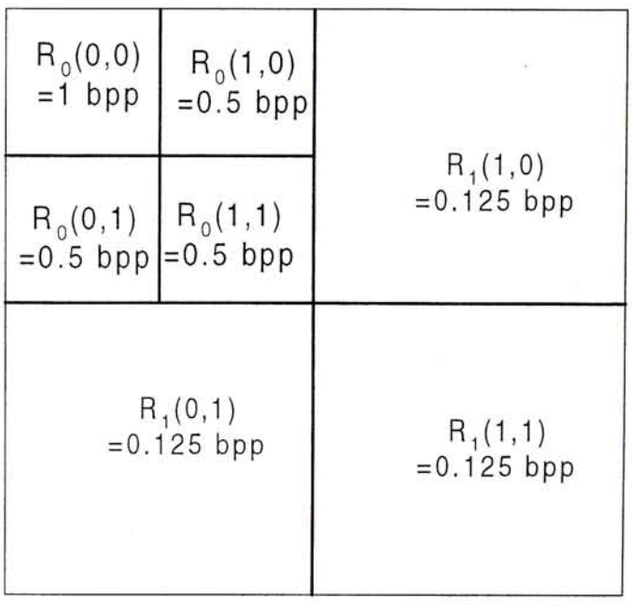



Fig. 5.3 An example of bit allocation

# Chapter 6

# Parental Finite State Vector Quantizers (PFSVQ)

## 6.1 Introduction

This chapter aims at improving the performance of the fixed-rate wavelet coding using vector quantization mentioned in last chapter. In last chapter, we used unconstrained VQ on each wavelet subband separately, that is, we used multiple codebooks for each subband. The size of each codebook is assigned by using Inter-band Bit Allocation (IBBA) algorithm. Our target in this chapter is to boost the fixed-rate wavelet coding performance by using constrained VQ instead of unconstrained VQ.

Many researchers use classified VQ (CVQ) in their coding algorithms. This is one of a good choice of constrained VQ because different vectors of the image can be classified into different classes according to the features of the vectors and each class has its own codebook. However, as stated chapter 3, one of the drawbacks of CVQ is that side information (denoting which codebook is being used) is need to be transmitted. This reduces the efficient use of bandwidth.

Finite state VQ (FSVQ) uses the similar principle as CVQ. However, instead of classifying the vectors into classes, the vectors are classified into states according to a finite state

function which is known to both the encoder and the decoder. In other words, no side-information is required. The example in this chapter will show how this can be achieved.

The finite state VQ used in this chapter is called Parental Finite State Vector Quantization (PFSVQ). It is designed to use together with wavelet subband coding because the finite state function is based on the information of the vectors in the parent subband.

In the last chapter we use VQ on separate bands and use inter-band bit allocation to assign the codebook size. The vectors of one subband do not affect the codebook training of another subband. In this chapter, our PFSVQ algorithm, the codebook training of the current subband is affected by the vectors of its parent. The VQ using this kind of subband structure is called VQ using interband information for intraband vectors.

In the next sections, we will describe the parent-child relationship between subbands. Section 6.3 will introduce different subband vector structures for VQ. Section 6.4 will introduce the Parental Finite State Vector Quantization algorithm.

## 6.2 Parent-Child Relationship Between Subbands

In a wavelet octave subband system, the image is firstly decomposed into four subbands $LL_1$, $LH_1$, $HL_1$, and $HH_1$ by using separable horizontal and vertical filters (assume the wavelet filters are separable). These are the finest scales of the wavelet subbands. To obtain next coarser scale, the lowest frequency subband, i.e. $LL_1$, is taken for decomposition. $LL_1$ is decomposed into 4 coarser subbands, $LL_2$, $LH_2$, $HL_2$, $LL_2$. In the same way, each wavelet decomposition at level n results in four subbands $LL_n$ , $LH_n$ , $HL_n$ , $HH_n$ . To further decompose to next coarser level the subbands, $LL_n$ , is taken for decomposition. The n level wavelet decomposition is also called n-scale wavelet decomposition and results in (n+1) levels of wavelet pyramids. Fig. 6.1 gives an example of n=2.

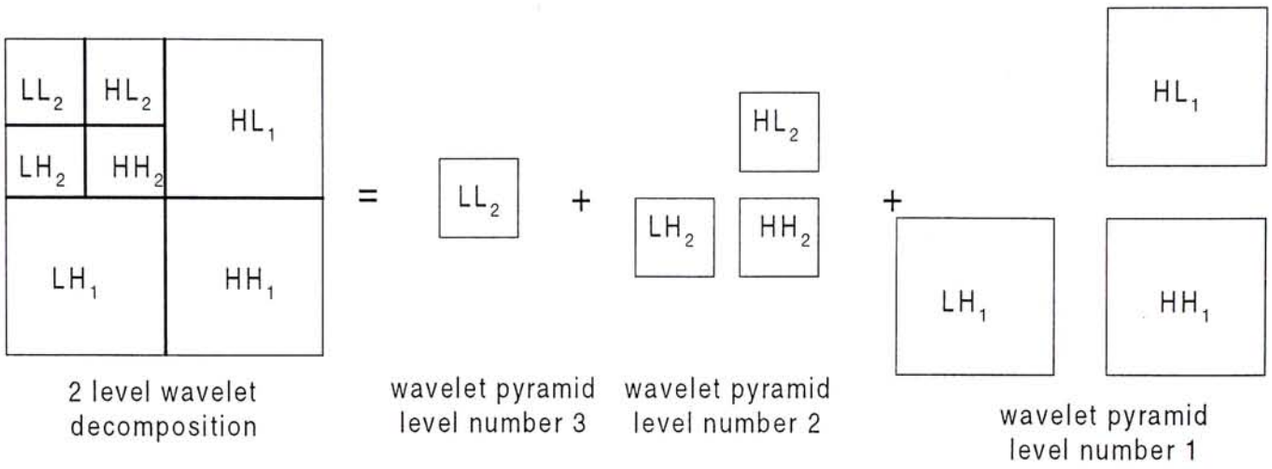



Fig. 6.1 Relationship between wavelet decomposition level and wavelet pyramid level, given wavelet decomposition level, n, equals 2.

Now every coefficient at a given scale has a parent-child relationship with a set of coefficients at the next finer scale of similar orientation. The word "parent" refers to the coefficient at the coarser scale. The word "child" refers to all coefficients corresponding to the same spatial location at the next finer scale of similar orientation. For a given parent, the word "descendants" refers to the set of all coefficients at all finer scales of similar orientation. Finally, for a given child, the word "ancestors" refers to the set of all coefficients at all coarser scales of similar orientation. Fig. 6.2 shows this relationship.

From the figure, all parents have four children except that in the lowest frequency subband in which all the parents have only three children, each one located on the higher frequency subband within same wavelet level.

The above gives an overview of the parent-child relationship of a wavelet octave subband system. For more detail, please refer to [31]. The PFSVQ algorithms introduced in this chapter exploits this relationship to define the next state function.

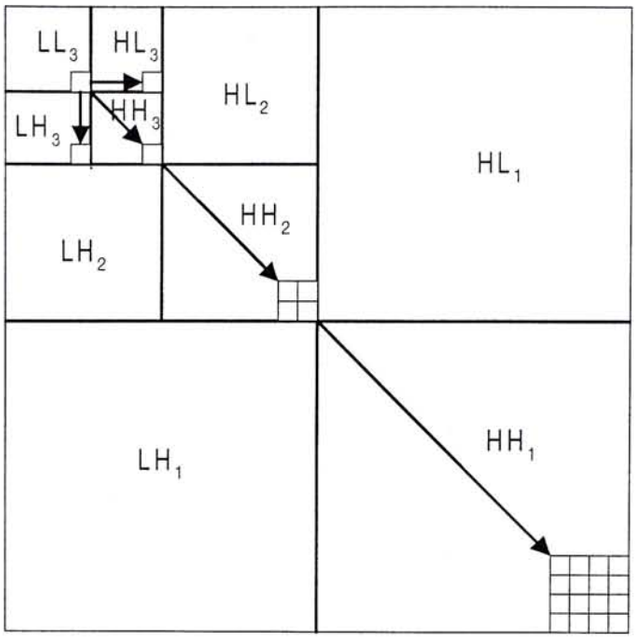



Fig. 6.2 Parent-Child relationship of an octave wavelet subband system

## 6.3 Wavelet Subband Vector Structures for VQ

Wavelet subband structures for VQ refers to how we organize the vectors from the same subband or different subbands for the purpose of vector quantization. It is, therefore, obvious that wavelet subband vector structure affects the way of how we do the VQ very much and hence affects the coding performance very much. Coding performance will be improved if we can organize the vector structure in a way that is suitable for the vector quantization algorithm being used. In summary, there are three kinds of structures. They are "VQ on Separate Bands", "Interband Information for Intraband Vectors", "Cross Vector Method" [23]. We will introduce them in the following sections.

### 6.3.1 VQ on Separate Bands

In general, there is high correlation between neighbour coefficients within a subband. Vector quantization is done on each subband separately. The vectors in one band do not affect the VQ on the other band. Vectors are formed by grouping wavelet coefficients from the same subband. Usually, the size of the vectors is 4x4. This can exploit both horizontal and vertical correlation of the coefficients. If the size of the vector is too small, the correlation between coefficient is not effectively used. On the contrary, if the size is too large, the computational complexity is very large. The total number of bits available is allocated among each subband.

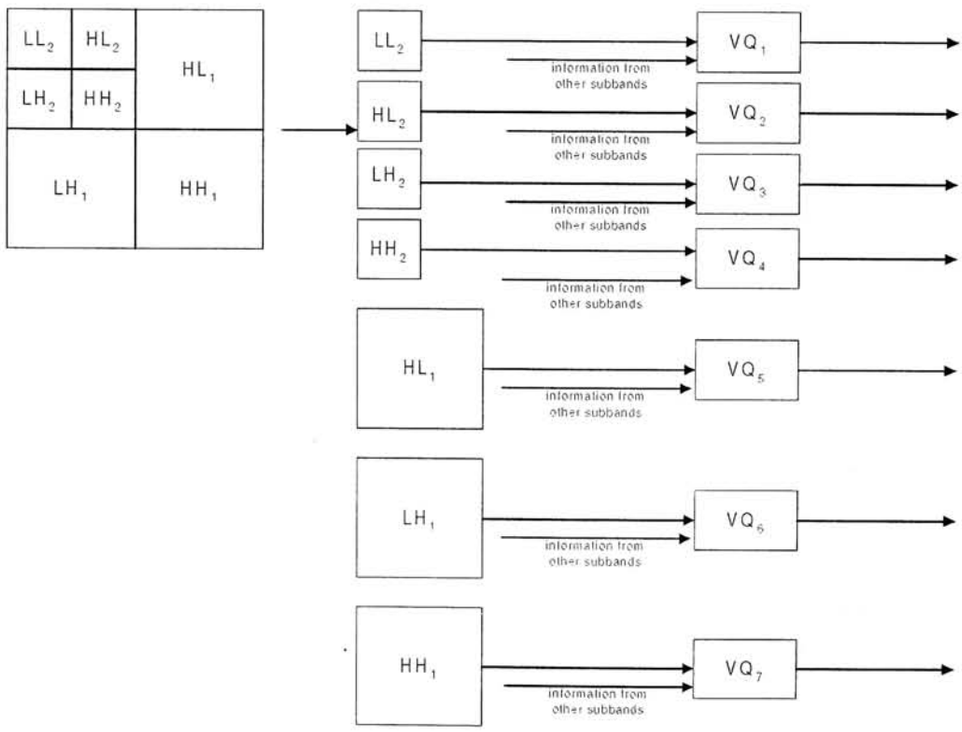



Fig 6.3 Structure of wavelet subband coding with the use of interband information for intraband vectors, for wavelet decomposition level= 2

## 6.3.2 InterBand Information for Intraband Vectors

The correlation between coefficients in different subbands is taken into consideration. Therefore the VQ of the current subband is affected by the vectors or coefficients of the other subbands. Each vectors are still formed from coefficients of the same subband but the VQ use

the information from other subbands, that is interband information, for codebook training, encoding and decoding. For example finite state vector quantizer with VQ done on each subband separately but the state of a vector is obtained from the information of other subbands. The Parental Finite State VQ is one of this kind. Fig. 6.3 shows the structure of wavelet subband coding using interband information for intraband vectors for wavelet decomposition level equals 2.

### 6.3.3 Cross band Vector Methods

Vectors are formed by grouping coefficients from different subbands. We call this kind of vectors as cross band vectors. The correlation between coefficients of different subband is exploited in each vector. Usually, the lowest frequency subband is encoded separately because energy of the coefficients of this band is much bigger than that in the other subbands. In addition, the lowest subband has very different characteristics, both perceptually and statistically, from the other subbands. If the coefficients from the lowest subband are included in the cross band vector, the intra vector correlation of the cross band vector is not as good as that without including the lowest subband.

There are several methods of grouping coefficients from different bands. For example in the case of uniform decomposition (section 2.3), the cross band vector is formed by take one coefficients from the same orientation of each subband (fig. 6.4a). In case of octave band decomposition, cross band vectors can be formed by grouping the childern and the parents together (fig. 6.4b) or by grouping the coefficients from the subbands in the same wavelet pyramid level and in the same orientation (fig. 6.4c). The last step, regrouping, of the vector wavelet transform (VWT) described in chapter 4 is one kind of cross band vector grouping.

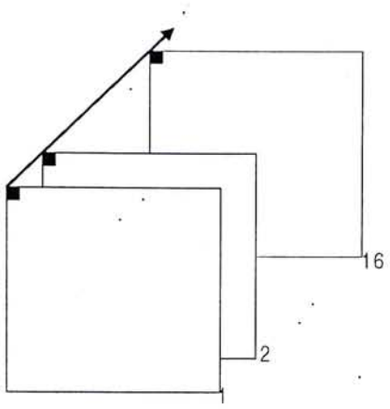
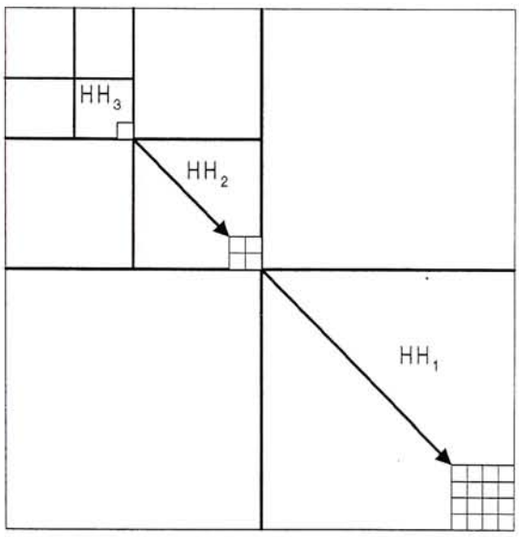
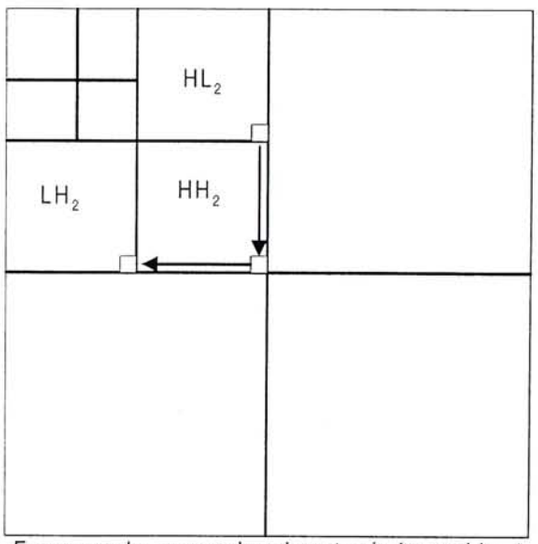



Fig. 6.4a Cross band vector form from16 uniform subbands.




For example, a crossband vector is formed by 1 coefficient in $HH_3$, 4 coefficient in $HH_2$ and 16 coefficients from $HH_1$

Fig. 6.4b Cross band vector form by grouping parents and children together




For example, a crossband vector is formed by 1 coefficient in $HL_2$, 1 coefficient in $HH_2$ and 1 coefficients from $LH_2$
Note that they come from subbands of the same wavelet pyramid level and at the same orientation

Fig. 6.4c Cross band vector form from coefficients of the same wavelet pyramid level at the same orientation.

## 6.4 Parental Finite State Vector Quantization Algorithms

Before describing the Parental Finite State Vector Quantization algorithm, we first define some terms for future use. We extend the relationship described in section 6.2 from scalar scope to vector scope. For a given subband, the word "parent subband" refer to the subband one scale coarser and "child subband" refer to the subband one scale finer. For a given vector, the word "parent vector" refers to the vector formed from the all parent coefficients of the current vector components and the word "child vector" refers to the vectors formed from all child coefficients of the current vector components.

Therefore except in the lowest frequency subband, all NxN dimensional vectors have one parent vector of dimension N/2xN/2 and have one child vector of dimensional 2Nx2N. Each NxN dimensional vector in the lowest frequency subband has three child vectors each of dimension NxN.

### 6.4.1 Scheme I: Parental Finite State VQ with Parent Index Equals Child Class Number

Suppose the codebook size of the VQ of the parent vectors is M, then after encoding, all the parent vectors are mapped into one of the M symbols. In other words, the parents vectors are all clustered into M different groups and each group is represented by one centroid vector. Therefore, we expect that within each group, all vectors have some similarities. This means we can group/classify vectors into classes/states according to their similarities.

These inter vector similarities within each subband exists in both the parent subband and the child subband. In addition, there is correlation between parent vector and child vector. As a result, we can group child vectors into groups according to the grouping result of the parent vectors in the parent subband. That is, vectors can be classified into states according the VQ

encoding result of the parent vectors. If the codebook size of the VQ in the parent subband is M, then the child vectors can be classified into M states. Here is the rule of classifying:

For the current vector $v$ having parent vector $v_p$, if the VQ encoding output of $v_p$ is $x$, then $v$ belongs to state $x$.

In words of finite state VQ, the VQ encoder output of parent vectors give the state of the child vectors. If codebook size of the codebook of parent subband is M, there are M states in the finite vector quantizer for the child subband. Suppose the VQ encoder output, i.e. that the index, of the parent subband is already transmitted to the decoder, then the decoder can easily determine the state of the current vector without the need of any transmission of side information.

Now each vector is classified into M states, then a codebook is searched for each state by LBG algorithm. In the encoder, the VQ encoding result or index of the parent vector is first found. The index determines the state of the current vector. If the index is i, then codebook i is used for encoding. In the decoder, from the current input index, we know which vector is being decoded, therefore, we know where its parent located. As a result, the parent VQ index is also known. The parent VQ index indicates which codebook is going to be used for decoding. The codebook training method, encoding and decoding is summarized in fig. 6.5a, fig. 6.5b and fig. 6.5c respectively.

For example (we do not consider the case when the parent subband is the lowest frequency subband), if the size of parent subband is pxp and the size of parent vectors is NxN, then there are pxp/(NxN) parent vectors. The size of the corresponding child subband is 2px2p and the size of child vectors is 2Nx2N and there are totally 2px2p/(2Nx2N)= pxp/(NxN) child vectors. If the codebook size of the parent subband is M, then there are M states for the

PFSVQ of the child subband. Vectors (of each state) are coded separately by VQ using codebooks which are trained from vectors of the corresponding states.

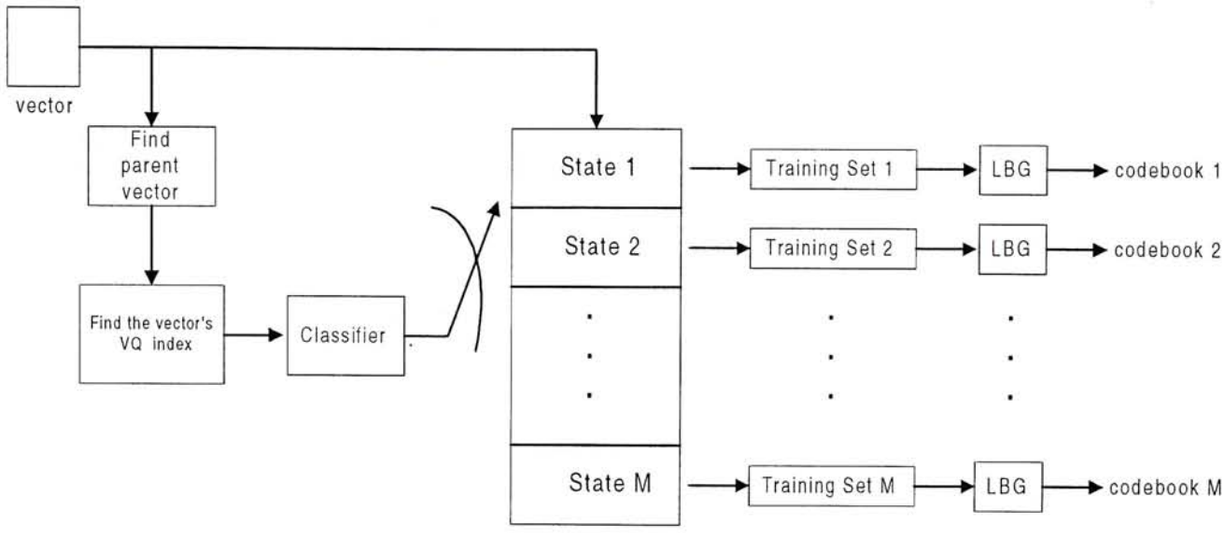



Fig. 6.5a Codebook training of scheme I of PFSVQ

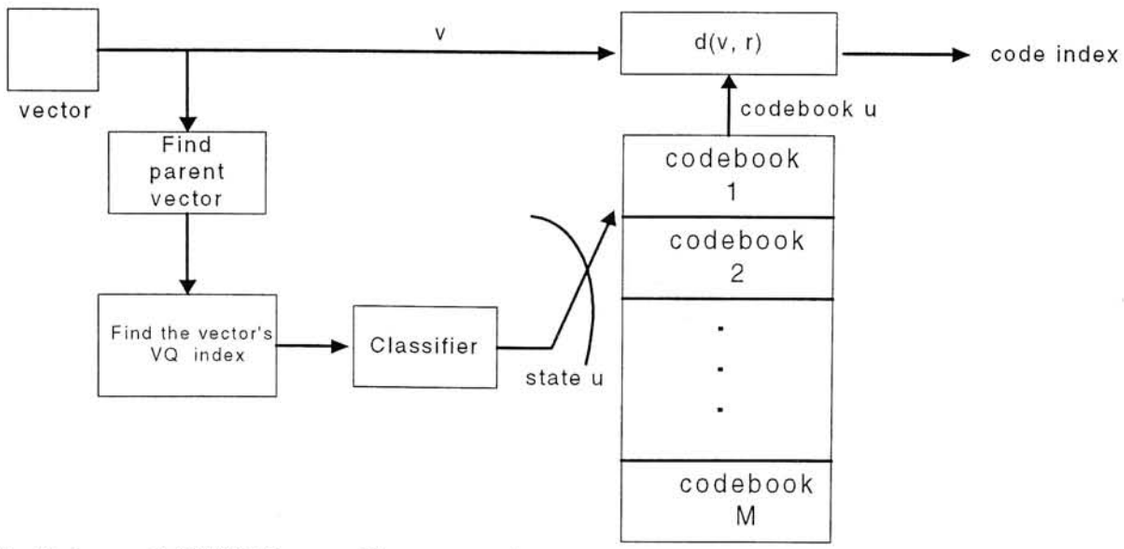



Fig. 6.5b Scheme I PFSVQ encoding procedure

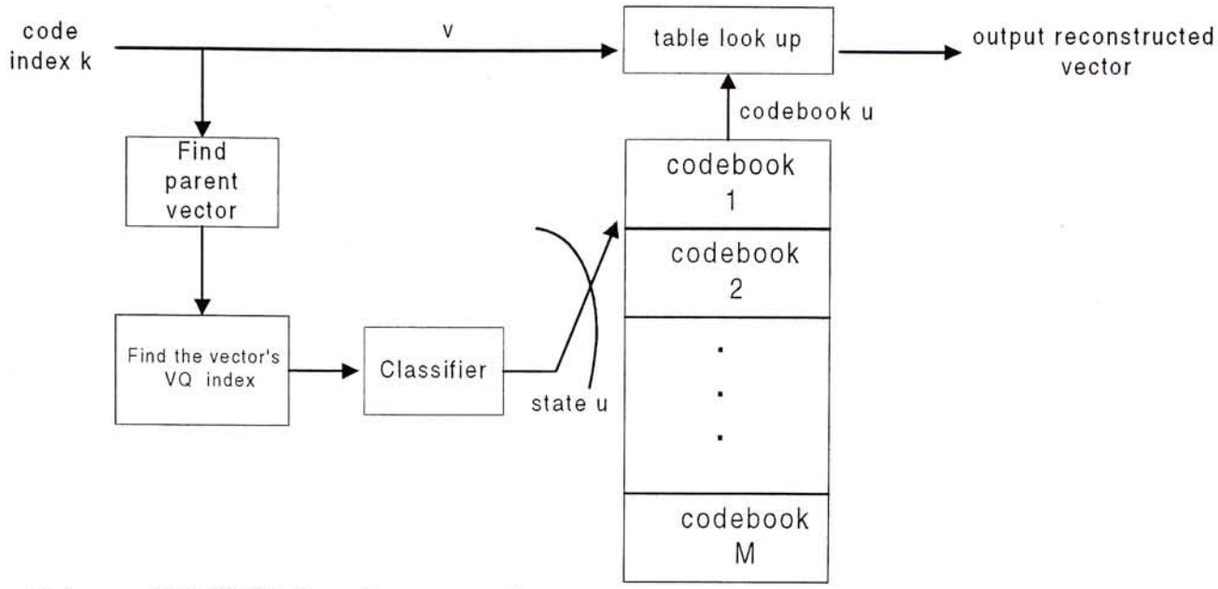



Fig. 6.5c Scheme I PFSVQ decoding procedure

## 6.4.2 Scheme II: Parental Finite State VQ with Parent Index Larger than Child Class Number

In scheme I, we can classified the current vectors into different states by the VQ encoding results of the corresponding parent vectors. For example, if the codebook size of the parent subband, M, equals 256 and the size of the parent subband is 128x128, the size of the parent vector is 4x4. Then there will have 256 states for the child vectors. This means there are totally 256 codebooks for the child subband. If M= 512, the total number of child subband codebooks doubles. Sometimes, the encoder and decoder want to keep the number of codebooks smaller. In this case, scheme I seems to be undesirable. In the above example, the child subband has size 512x512 and the size of child vector is 8x8, then the total number of child vector equals 4096. Since the total number of states equals 256, therefore the total number of child vectors per state equals 16. That means for each image in the training set, there are on average 16 training vectors for each state. If we want to train a codebook of size 256, then we need at least 256*40 (section 3.3.2) training vectors. The total number of training image required is 256*40/16= 640 !!! This is unrealistic in most cases. However, the

training set for a codebook affects the training result very much. In order to solve this problem, we reduce the number of states.

In scheme II, the number of states is less than the number of code vectors in the codebook of parent subband, or in other words, the number of states is less than the number of code index of the VQ of parent subband. In order to achieve this aim, the parents vectors are firstly VQ encoded and decoded. A large set of training vectors formed by the reconstructed vectors in the parent subband and it is used for another LBG codebook training. In this codebook training, the number of code vectors is set to the desired number of states (for encoding the child vectors using PFSVQ). We call this new codebook as State-Reduced codebook (SR codebook). The parent reconstructed vectors are then encoded by the SR codebook. As a result, the parent vectors have a new and smaller set of code index. Now in PFSVQ, the state of the current vector is based on the new code indices. Similarly, in the encoding and decoding process, the states of current vectors are determined by the new VQ encoding result of the parent reconstructed vectors (which is based on the SR codebook). Scheme II is summarized in fig. 6.6.

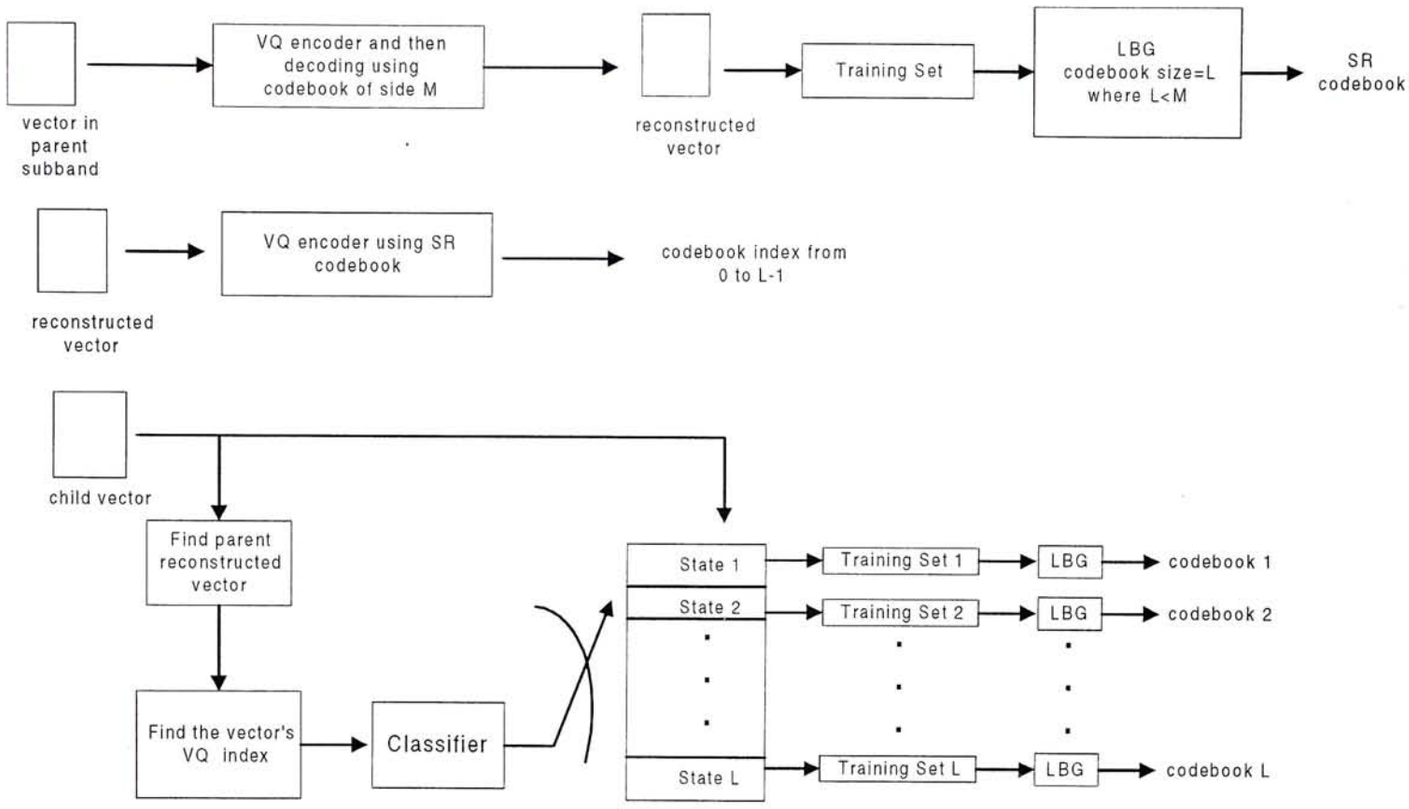



Fig. 6.6 Scheme II of PFSVQ

# Chapter 7

# Simulation Result

## 7.1 Introduction

This section gives the simulation results on the algorithms described in early chapters. Section 7.2 is the simulation results of Vector Wavelet Transform (VWT). Section 7.3 is the simulation results of Vector Wavelet Transform-Linear Predictive Coding (VWT-LPC). Section 7.4 is the results of using the Inter-band Bit Allocation (IBBA) on wavelet subband image. Section 7.5 is the simulation results of Parental Finite State VQ (PFSVQ).

Generally, the simulation parameters are as follows unless specially stated. 512x512 Lena image and Aeroplane image, 8 bits per pixel are used for the simulations. The training set is a set of 18 images (exclude Lena image and Aeroplane image) and global codebook is used in the simulation. .The subsampling rate used in step one of the VWT or VWT-LPC procedure is 4. The thresholds used in VQ codebook training are 0.1 and 0.01 and the vectors are formed by non-overlapping 4x4 blocks. We try two wavelet filters, the Daubechies-4 (DB-4) orthonormal filter and the 9/7 biorthogonal filter. All wavelet transforms are done in two levels. The output of the VWT and VWT-LPC, are either transmitted directly, i.e. no lossless coding, or lossless coded. We simulate the coding performance of five lossless coding algorithms: (i) Lempel-Ziv coding; (ii) Arithmetic Coding using Markov order equals 3; (iii)

Arithmetic Coding using Markov order equals 2; (iv) Arithmetic Coding using Markov order equals 1; (v) Arithmetic Coding using Markov order equals 0.

## 7.2 Simulation Result of Vector Wavelet Transform (VWT)

Table 7.1 is the result of VWT using Lena image. The subsampling rate is 4. The wavelet filter is Daubechies-4. No amplitude shifted is done before wavelet transform. The VQ used is unconstrained VQ using threshold 0.1 and codebook sizes are 512, 256, 128 and 64.

Table 7.2 is the result of VWT using Aeroplane image. The subsampling rate is 4. The wavelet filter is Daubechies-4. No amplitude shifted is done before wavelet transform. The VQ used is unconstrained VQ using threshold 0.1 and codebook sizes are 512, 256, 128 and 64.

Table 7.3 is the result of VWT using Lena image. The subsampling rate is 4. The wavelet filter is 9/7 biorthogonal filter. No amplitude shifted is done before wavelet transform. The VQ used is unconstrained VQ using threshold 0.1 and codebook sizes are 512, 256, 128 and 64.

Table 7.4 is the result of VWT using Aeroplane image. The subsampling rate is 4. The wavelet filter is 9/7 biorthogonal filter. No amplitude shifted is done before wavelet transform. The VQ used is unconstrained VQ using threshold 0.1 and codebook sizes are 512, 256, 128 and 64.

Table 7.5 is the result of VWT using Lena image. The subsampling rate is 4. The wavelet filter is 9/7 biorthogonal filter. Amplitude shifted is done before wavelet transform. The VQ used is unconstrained VQ using threshold 0.1 and codebook sizes are 512, 256, 128, 64 and 32.

Table 7.6 is the result of VWT using Lena image. The subsampling rate is 4. The wavelet filter is 9/7 biorthogonal filter. Amplitude shifted is done before wavelet transform. The VQ used is unconstrained VQ using threshold 0.01 and codebook sizes are 512, 256, 128, 64 and 32, 16.

The performance of using Lena image, using different parameters are plotted in fig. 7.1. and the performance of using Aeroplane image, using different parameters are plotted in fig. 7.2. In both curve, the performance is measured by PSNR versus bit rate (bits per pixel). In each curve, for the same PSNR, the smallest bit rate (corresponds to the best lossless coding scheme) is taken to be the data point in the performance curve. From example, we take the data 0.45 bpp at PSNR=28.43 dB (in table 7.1) to put on the performance curve (fig. 7.1) instead of other bit rates at the same PSNR.

It can been seen clearly that the performance is better when we used smaller threshold for the VQ, especially at lower bit rate. 9/7 biorthogonal filter performs better than Daubechies-4 filter at lower bit rate, however, their difference becomes smaller at moderate bit rate. The amplitude shift before wavelet transform helps to improve the performance at both low bit rate and moderate bitrate. The performance of VWT is comparable to other algorithms. At bit rate of 0.37 bpp, VWT has PSNR around 28.17 dB while the PSNR stated in [32] is 27.27dB. At bit rate of 0.5bpp, VWT has PSNR around 30dB.

The reconstructed images of the Aeroplane image using VWT, 9/7 biorthogonal filter, VQ threshold = 0.1 and no amplitude shift is shown in fig 7.5. The reconstructed images of Lena image using VWT, 9/7 biorthogonal filter, VQ threshold= 0.01 and with amplitude shift is shown is fig. 7.6. The reconstructed images have good visual reconstruction performance.

## 7.3 Simulation Result of Vector Wavelet Transform-Linear Predictive Coding (VWT-LPC)

### 7.3.1 First Test

In the first test, amplitude shift is done before the wavelet transform and we use Intra-vector Mean Predictor for the prediction in LPC. The residual is unconstrained VQ with vector dimension 4x4. The threshold of the VQ is 0.01. The predictor is separately coded by uniform quantization with step size 10.0, 15.0 and 20.0. The output of the quantizer is then losslessly coded. Table 7.7a is the coding result of the predictor of VWT-LPC. Table 7.7b is the overall coding result of the VWT-LPC.

### 7.3.2 Second Test

In the second test, no amplitude shift is done before the wavelet transform. We also use Intra-Vector Mean Predictor for the predictor in LPC. The coding of residual is same as that in the first test except the threshold of the VQ is now 0.1. The predictor is coded by uniform quantizer with step size 16.0. Table 7.8a is the coding result of the predictor. Table 7.8b is the coding result of the residual. Table 7.8c is the overall coding results. In this table, the result is obtained by using the best predictor coding scheme ($2^{nd}$ last row of table 7.8a) together with the best residual coding scheme in each codebook size of table 7.8b.

### 7.3.3 Third Test

In this test, amplitude shift is done before the wavelet transform and we still use Intra-vector Mean Predictor for the prediction in LPC. The residual is unconstrained VQ with vector dimension 4x4. The threshold of the VQ is 0.1. The predictor is separately coded by uniform quantization with step size 10.0, 15.0 and 20.0. The output of the quantizer is then losslessly coded. Table 7.9a is the coding result of the predictor of VWT-LPC. Table 7.9b is the overall coding result of the VWT-LPC.

We plot the performance curve of VWT-LPC in fig. 7.3. From the diagram, the gain in coding performance can be achieved by using amplitude shift before the image transform and obtained by using small VQ threshold. This performance gain is nearly constant over a range of bit rate. The comparison of VWT and VWT-LPC is plotted in fig. 7.4. It can be seen easily that the VWT-LPC has much better performance than VWT. For example at bit rate of 0.3 bpp, the VWT-LPC has PSNR= 30 dB while for VWT, it achieve PSNR= 30 dB at bit rate of 0.5 bpp and the PSNR of VWT at 0.3 bpp is only 27.5dB. The most important thing is the improvement is obtained by only adding a very little complexity to the VWT.

Comparing to other coding algorithms, at the bit rate of 0.25 bpp, using Lena image, the coding performance of other algorithms are [12]:

VQ coding = 26.2 dB
DCT-based transform coding = 27.3 dB
Subband coding with VQ = 28.3dB
Subband coding together with DCT and VQ = 27.2
Subband coding with other vector image transform = 29.2

Our algorithm using VWT-LPC has PSNR= 29.42 dB at the same bit rate. The most important thing is VWT-LPC is very easy to implement. The reconstructed images of Lena using VWT-LPC are shown in fig. 7.7. These images have good visual performance.

## 7.4 Simulation Result of Vector Quantization Using Inter-band Bit Allocation (IBBA)

In this simulation, we use the IBBA algorithm to allocate the total bit rate of 0.25 bpp around the seven bands of a 2-level wavelet decomposed Lena image. Lena image is firstly transformed by 9/7 biorthogonal filter and is decomposed into seven subbands. The IBBA algorithm finds the number of bits for each band, that is the codebook size for the VQ of each subband. Then separate VQ are done on each subband. The VQ threshold is 0.01.

First, we train the codebook by a training set of 19 images, where the Lena image is also included. The result is put in table 7.10. Then we do the same test, however, this time, Lena image is excluded from the training. The result is put in table 7.11. The results show that the performance is much better if Lena image is included in the training set. The improve in performance is over 5 dB. Result also shows that in case of Lena image not in the training set, the highest level of the pyramid, that is subband $LL_2$ has the worse performance. We expect that the improve in performance in $LL_2$ will greatly improve the overall performance.

It is because $LL_2$ has very large variance and therefore it is difficult to encode $LL_2$ using VQ. $LL_2$ is the coarsest scale of the subband. To avoid any blocking effect in the image, we code it separately using other algorithm. In our experiment, we use the SPIHT algorithm [33], to encode the $LL_2$ at the bit rate of 1 bpp. The result is shown in table 7.12. The PSNR of using SPIHT on $LL_2$ at 1 bpp is 21.83 dB. Compare to that obtained by VQ of the $LL_2$ with Lena image excluded from the training set, the new encoding method has 3 dB improvement. The overall PSNR increases from 26.41dB to 29.34dB. Fig. 7.8 shows the reconstructed Lena image using IBBA algorithm.

## 7.5 Simulation Result of Parental Finite State Vector Quantizers (PFSVQ)

In this part, we use Lena image for testing. It is first wavelet transformed by 9/7 biorthogonal filter and each band, except the $LL_1$, is separately VQ coded. The codebook size for the VQ in each subband is obtained by IBBA algorithm. The $LL_1$ subband is separately coded by SPIHT algorithm. The result is shown in table 7.12.

The simulation in this part is aimed at improve the result shown in table 7.12. From the result in table 7.12, we find that among the subbands in the second wavelet pyramid ($HL_2$, $LH_2$, $HH_2$), $HL_2$ has the worse coding performance. Among the first wavelet pyramid ($HL_1$, $LH_1$, $HH_1$), $HL_1$ has the worse coding performance. Therefore, in order to improve the overall performance, the performance of the subband $HL_1$ and $HL_2$ must be improved.

In addition, we observed that the $HL_2$ is the parent subband of $HL_1$. Therefore, in this simulation, we try to improve the coding performance of $HL_1$ by using Parental Finite State Vector Quantizers (PFSVQ), both Scheme I and Scheme II will be used for testing. The coding performance of $HL_1$ using PFSVQ Scheme I and PFSVQ Scheme II is shown is table 7.13. It shows that the PFSVQ Scheme I successfully improve the coding performance of $HL_1$ from 31.74 dB to 32.29dB and PFSVQ Scheme II also improve the performance to 32.22dB. However, scheme II achieves this result by using bit rate doubles of that of scheme I. This simulation shows that scheme I is good enough in providing the improvement. But scheme II does have its merit of require less storage space for codebooks.

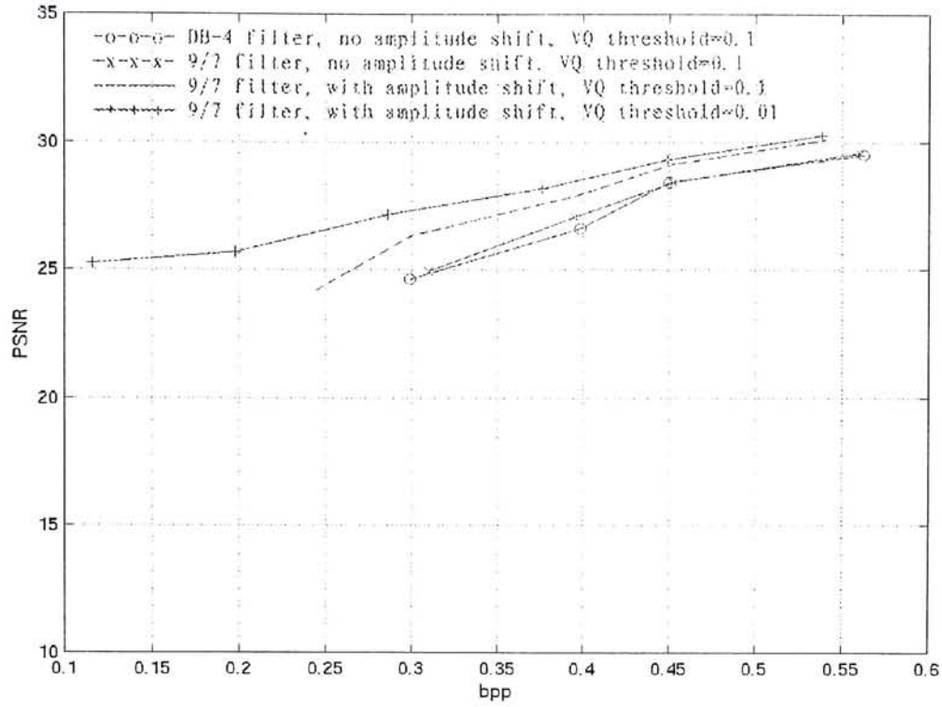



Fig. 7.1 PSNR vs bit rate of VWT using Lena image and using different parameters.

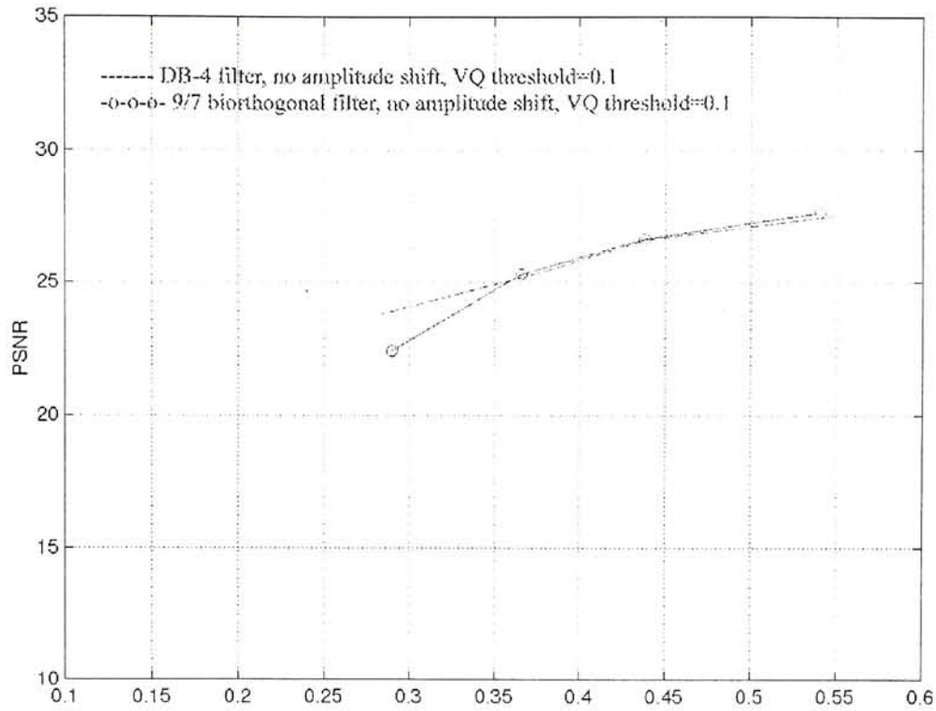



Fig. 7.2 PSNR vs bit rate of VWT using Aeroplane image and using different parameters.

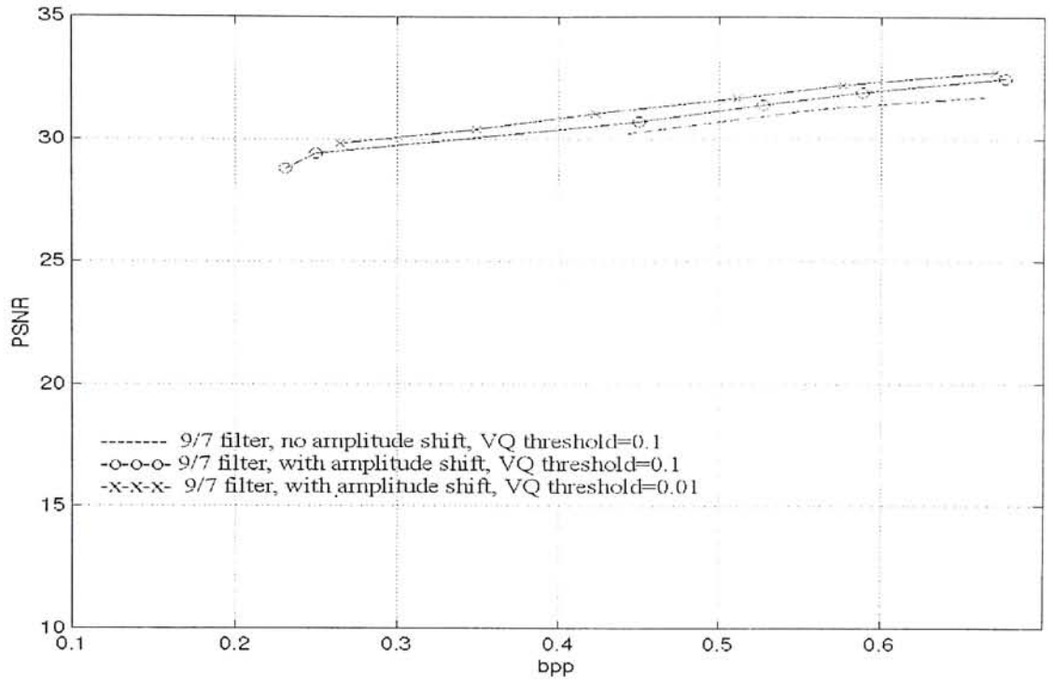



Fig. 7.3 PSNR vs bit rate of VWT-LPC using Lena image and using different parameters.

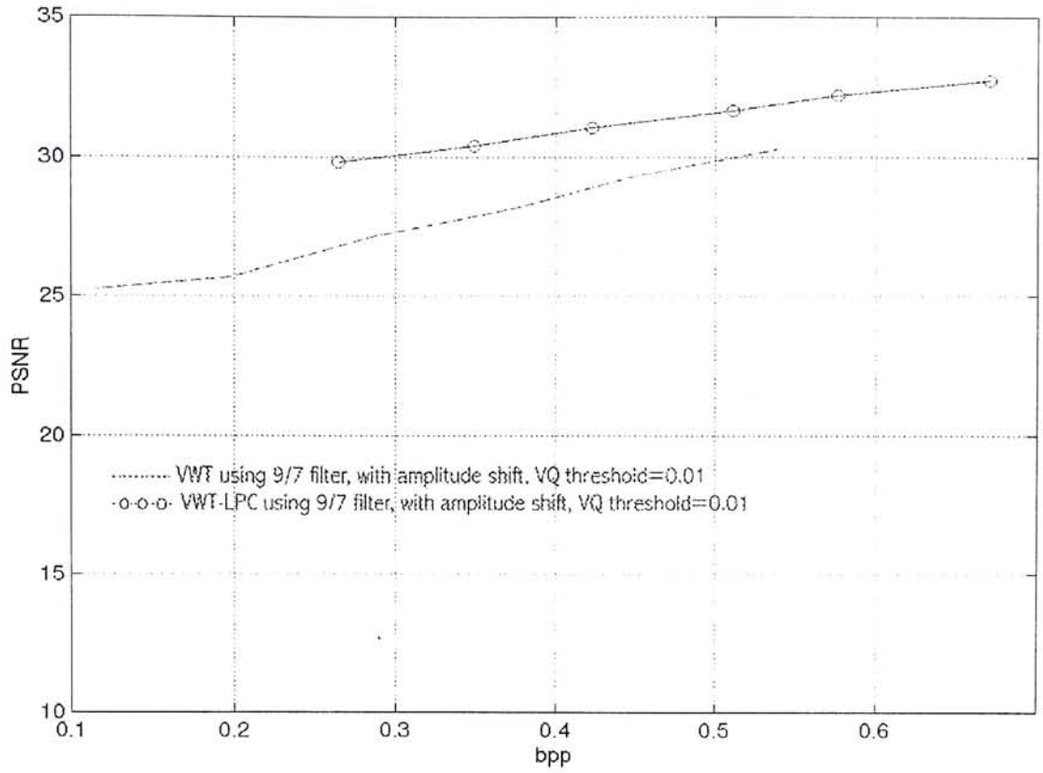



Fig. 7.4 Comparison of performance of VWT and VWT-LPC using Lena image.

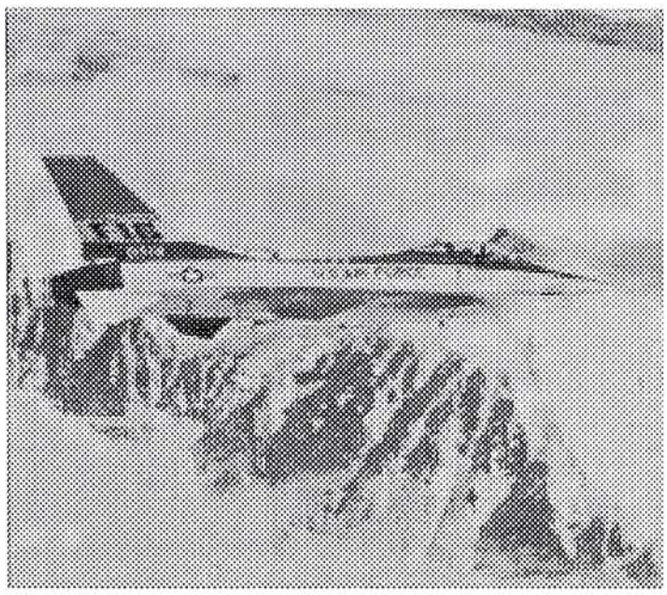


Fig. 7.5a VWT result of Aeroplane image, 0.547bpp at 27.51dB

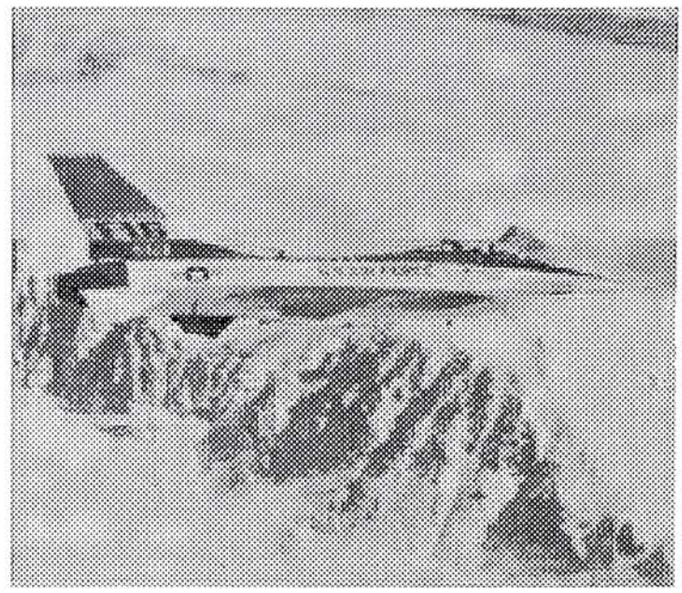


Fig. 7.5b VWT result of Aeroplane image 0.438bpp at 26.61dB

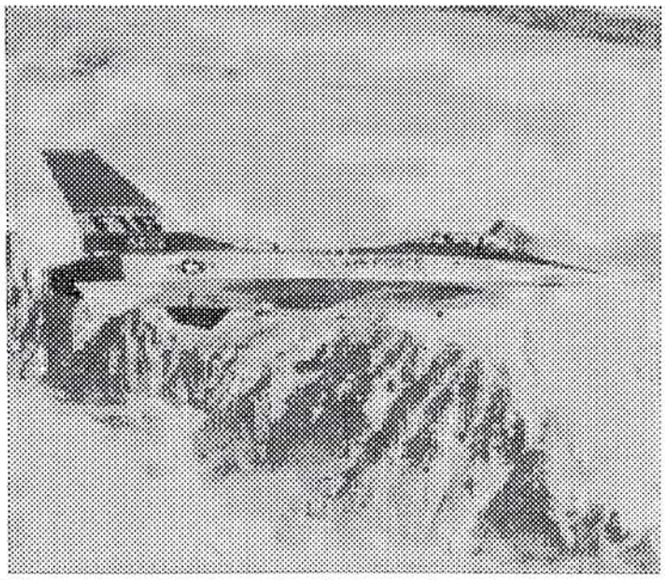


Fig 7.5c VWT result of Aeroplane image, 0.377bpp, 25.37dB

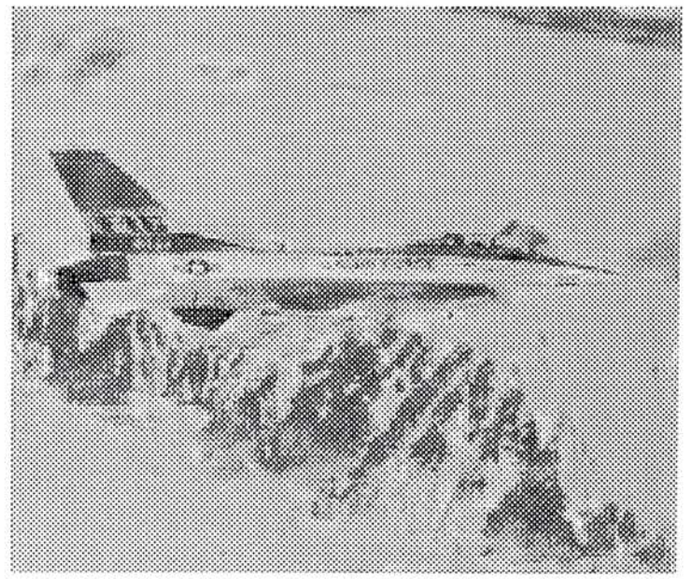


Fig 7.5d VWT result of Aeroplane image, 0.284bpp, 23.82dB

Fig 7.6a VWT result of Lena image,
0.538bpp, 30.27dB

Fig 7.6b VWT result of Lena image,
0.449bpp, 29.29dB

Fig 7.6c VWT result of Lena image,
0.376bpp, 28.17dB

Fig 7.6d VWT result of Lena image,
0.286bpp, 27.17dB

Fig 7.6e VWT result of Lena image, 0.198bpp, 25.70dB

Fig. 7.6f VWT result of Lena image 0.116bpp, 25.23dB

Fig. 7.7a VWT-LPC result of Lena image 0.67bpp, 32.76dB

Fig. 7.7b VWT-LPC result of Lena image, 0.576bpp, 32.2dB

Fig. 7.7c VWT-LPC result of Lena image 0.511bpp, 31.65dB

Fig. 7.7d VWT-LPC result of Lena image, 0.423bpp, 31.04dB

Fig 7.7e VWT-LPC result of Lena image 0.349bpp, 30.39dB

Fig 7.7f VWT-LPC result of Lena image 0.265bpp, 29.83dB

Fig. 7.8a Coding result of Lena image using IBBA and local codebook, 0.25bpp, 31.62dB

Fig. 7.8b Coding result of Lena image, using IBBA and global codebook, 0.25bpp, 26.41dB

Fig. 7.8c Coding result of Lena image using IBBA and local codebook, the lowest frequency subband is coded by SPIHT, 0.25bpp, 29.34dB

Table 7.1 Coding result of VWT of Lena image using subsampling rate of 4, Daubechies-4 filter in two levels decomposition, no amplitude shift before image transform. Threshold of VQ= 0.1

| Codebook size | PSNR | bpp (no lossless coding) | bpp (with lossless coding) |
|---|---|---|---|
| 512 | 29.5 | 0.5625 | |
| | | | |
| 256 | 28.43 | 0.5 | 0.45 (a) |
| | | | 0.47 (b) |
| | | | 0.47 (c) |
| | | | 0.47 (d) |
| | | | 0.486 (e) |
| | | | |
| 128 | 26.60 | 0.4375 | 0.416 (a) |
| | | | 0.418 (b) |
| | | | 0.412 (c) |
| | | | 0.398 (d) |
| | | | 0.409 (e) |
| | | | |
| 64 | 24.63 | 0.375 | 0.3397 (a) |
| | | | 0.33 (b) |
| | | | 0.32 (c) |
| | | | 0.299 (d) |
| | | | 0.32 (e) |
| | | | |

(a) = Lempel-Ziv coding (LZ77)
(b) = Arithmetic coding using Markov order of 3
(c) = Arithmetic coding using Markov order of 2
(d) = Arithmetic coding using Markov order of 1
(e) = Arithmetic coding using Markov order of 0

Table 7.2 Coding result of VWT of Aeroplane image using subsampling rate of 4, Daubechies-4 filter in two levels decomposition, no amplitude shift before image transform. Threshold of VQ= 0.1

| Codebook size | PSNR | bpp (no lossless coding) | bpp (with lossless coding) |
|---|---|---|---|
| 512 | 27.51 | 0.5625 | 0.547 (a) |
| | | | 0.57 (b) |
| | | | 0.57 (c) |
| | | | 0.569 (d) |
| | | | 0.58 (e) |
| | | | |
| 256 | 26.61 | 0.5 | 0.438 (a) |
| | | | 0.45 (b) |
| | | | 0.45 (c) |
| | | | 0.44 (d) |
| | | | 0.47 (e) |
| | | | |
| 128 | 25.37 | 0.4375 | 0.406 (a) |
| | | | 0.3899 (b) |
| | | | 0.385 (c) |
| | | | 0.377 (d) |
| | | | 0.399 (e) |
| | | | |
| 64 | 23.82 | 0.375 | 0.326 (a) |
| | | | 0.314 (b) |
| | | | 0.304 (c) |
| | | | 0.284 (d) |
| | | | 0.313 (e) |
| | | | |

(a) = Lempel-Ziv coding (LZ77)
(b) = Arithmetic coding using Markov order of 3
(c) = Arithmetic coding using Markov order of 2
(d) = Arithmetic coding using Markov order of 1
(e) = Arithmetic coding using Markov order of 0

Table 7.3 Coding result of VWT of Lena image using subsampling rate of 4, 9/7 biorthogonal filter in two levels decomposition, no amplitude shift before image transform. Threshold of VQ= 0.1

| | | | use bit file |
|---|---|---|---|
| Codebook size | PSNR | bpp (no lossless coding) | bpp (with lossless coding) |
| 512 | 29.59 | 0.5625 | 0.56 (a) |
| | | | 0.608 (b) |
| | | | 0.6073 (c) |
| | | | 0.607 (d) |
| | | | 0.608 (e) |
| | | | |
| 256 | 28.43 | 0.5 | 0.452 (a) |
| | | | 0.475 (b) |
| | | | 0.473 (c) |
| | | | 0.471 (d) |
| | | | 0.487 (e) |
| | | | |
| 128 | 27.05 | 0.4375 | 0.418 (a) |
| | | | 0.417 (b) |
| | | | 0.411 (c) |
| | | | 0.3951 (d) |
| | | | 0.410 (e) |
| | | | |
| 64 | 24.95 | 0.375 | 0.356 (a) |
| | | | 0.348 (b) |
| | | | 0.33 (c) |
| | | | 0.31 (d) |
| | | | 0.337 (e) |

(a) = Lempel-Ziv coding (LZ77)
(b) = Arithmetic coding using Markov order of 3
(c) = Arithmetic coding using Markov order of 2
(d) = Arithmetic coding using Markov order of 1
(e) = Arithmetic coding using Markov order of 0

Table 7.4 Coding result of VWT of Aeroplane image using subsampling rate of 4, 9/7 biorthogonal filter in two levels decomposition, no amplitude shift before image transform. Threshold of VQ= 0.1

| Codebook size | PSNR | bpp (no lossless coding) | bpp (with lossless coding) |
|---|---|---|---|
| 512 | 27.64 | 0.5625 | 0.54 (a) |
| | | | 0.57 (b) |
| | | | 0.569 (c) |
| | | | 0.562 (d) |
| | | | 0.579 (e) |
| | | | |
| 256 | 26.63 | 0.5 | 0.438 (a) |
| | | | 0.448 (b) |
| | | | 0.447 (c) |
| | | | 0.443 (d) |
| | | | 0.468 (e) |
| | | | |
| 128 | 25.31 | 0.4375 | 0.38 (a) |
| | | | 0.382 (b) |
| | | | 0.376 (c) |
| | | | 0.366 (d) |
| | | | 0.389 (e) |
| | | | |
| 64 | 22.40 | 0.375 | 0.318 (a) |
| | | | 0.327 (b) |
| | | | 0.315 (c) |
| | | | 0.29 (d) |
| | | | 0.321 (e) |
| | | | |

(a) = Lempel-Ziv coding (LZ77)
(b) = Arithmetic coding using Markov order of 3
(c) = Arithmetic coding using Markov order of 2
(d) = Arithmetic coding using Markov order of 1
(e) = Arithmetic coding using Markov order of 0

Table 7.5 Coding result of VWT of Lena image using subsampling rate of 4, 9/7 biorthogonal filter in two levels decomposition. Threshold of VQ= 0.1

| Codebook Size | PSNR | bpp (no Lossless Coding) | bpp (With Lossless Coding) |
|---|---|---|---|
| 512 | 30.10 | 0.5625 | 0.54 (a) |
| | | | 0.69 (b) |
| | | | 0.69 (c) |
| | | | 0.69 (d) |
| | | | |
| 256 | 29.11 | 0.5 | 0.45 (a) |
| | | | 0.486 (b) |
| | | | 0.486 (c) |
| | | | 0.479 (d) |
| | | | 0.489 (e) |
| | | | |
| 128 | 27.86 | 0.4375 | 0.4 (a) |
| | | | 0.413 (b) |
| | | | 0.412 (c) |
| | | | 0.395 (d) |
| | | | 0.395 (e) |
| | | | |
| 64 | 26.33 | 0.375 | 0.33 (a) |
| | | | 0.35 (b) |
| | | | 0.34 (c) |
| | | | 0.3 (d) |
| | | | 0.32 (e) |
| | | | |
| 32 | 24.18 | 0.3125 | 0.29 (a) |
| | | | 0.27 (b) |
| | | | 0.257 (c) |
| | | | 0.244 (d) |
| | | | 0.26 (e) |

(a) = Lempel-Ziv coding (LZ77)
(b) = Arithmetic coding using Markov order of 3
(c) = Arithmetic coding using Markov order of 2
(d) = Arithmetic coding using Markov order of 1
(e) = Arithmetic coding using Markov order of 0

Table 7.6 Coding result of VWT of Lena image using subsampling rate of 4, 9/7 biorthogonal filter in two levels decomposition. Threshold of VQ= 0.01

| Codebook size | PSNR | bpp (no lossless coding) | bpp (with lossless coding) |
|---|---|---|---|
| 512 | 30.27 | 0.5625 | 0.538 (a) |
| | | | 0.69 (b) |
| | | | 0.69 (c) |
| | | | 0.69 (d) |
| | | | 0.69 (e) |
| | | | |
| 256 | 29.29 | 0.5 | 0.449 (a) |
| | | | 0.47 (b) |
| | | | 0.47 (c) |
| | | | 0.458 (d) |
| | | | 0.465 (e) |
| | | | |
| 128 | 28.17 | 0.4375 | 0.38 (a) |
| | | | 0.399 (b) |
| | | | 0.403 (c) |
| | | | 0.376 (d) |
| | | | 0.378 (e) |
| | | | |
| 64 | 27.17 | 0.375 | 0.311 (a) |
| | | | 0.32 (b) |
| | | | 0.308 (c) |
| | | | 0.286 (d) |
| | | | 0.302 (e) |
| | | | |
| 32 | 25.70 | 0.3125 | 0.229 (a) |
| | | | 0.222 (b) |
| | | | 0.2107 (c) |
| | | | 0.198 (d) |
| | | | 0.2185 (e) |
| | | | |
| 16 | 25.23 | 0.25 | 0.147 (a) |
| | | | 0.127 (b) |
| | | | 0.120 (c) |
| | | | 0.116 (d) |
| | | | 0.135 (e) |

(a) = Lempel-Ziv coding (LZ77)
(b) = Arithmetic coding using Markov order of 3
(c) = Arithmetic coding using Markov order of 2
(d) = Arithmetic coding using Markov order of 1
(e) = Arithmetic coding using Markov order of 0

Table 7.7a Coding result of the predictor of VWT-LPC of Lena image using subsampling rate of 4, 9/7 biorthogonal filter in two levels decomposition. Prediction is intra-vector mean prediction.

| Uniform Scalar Quantizer Step Size | PSNR | bpp (with lossless coding) |
|---|---|---|
| 10.0 | 40.537 | 0.16 (a) |
| | | 0.14 (b) |
| | | 0.14 (c) |
| | | 0.13 (d) |
| | | |
| 15.0 | 37.6 | 0.13 (a) |
| | | 0.106 (b) |
| | | 0.1 (c) |
| | | 0.12 (d) |
| | | |
| 20.0 | 35.61 | 0.11 (a) |
| | | 0.09 (b) |
| | | 0.09 (c) |
| | | 0.09 (d) |

(a) = Lempel-Ziv coding (LZ77)
(b) = Arithmetic coding using Markov order of 3
(c) = Arithmetic coding using Markov order of 2
(d) = Arithmetic coding using Markov order of 1

Table 7.7b Overall Coding result of VWT-LPC of Lena image using subsampling rate of 4, 9/7 biorthogonal filter in two levels decomposition. Prediction is intra-vector mean prediction. Threshold of VQ= 0.01. Predictor is encoder using uniform quantizer, step size of 10.0 and arithmetic coding of Markov order = 1

| Codebook size | Overall PSNR | bpp (residual coded with lossless coding) |
|---|---|---|
| 512 | 32.76 | 0.67 (a) |
| | | |
| 256 | 32.2 | 0.576 (a) |
| | | 0.607 (b) |
| | | 0.607 (c) |
| | | 0.604 (d) |
| | | 0.604 (e) |
| | | |
| 128 | 31.65 | 0.519 (a) |
| | | 0.5466 (b) |
| | | 0.5391 (c) |
| | | 0.519 (d) |
| | | 0.511 (e) |
| | | |
| 64 | 31.04 | 0.449 (a) |
| | | 0.469 (b) |
| | | 0.448 (c) |
| | | 0.423 (d) |
| | | 0.43 (e) |
| | | |
| 32 | 30.39 | 0.386 (a) |
| | | 0.379 (b) |
| | | 0.367 (c) |
| | | 0.349 (d) |
| | | 0.360 (e) |
| | | |
| 16 | 29.83 | 0.3 (a) |
| | | 0.28 (b) |
| | | 0.27 (c) |
| | | 0.265 (d) |
| | | 0.278 (e) |

(a) = Lempel-Ziv coding (LZ77)
(b) = Arithmetic coding using Markov order of 3
(c) = Arithmetic coding using Markov order of 2
(d) = Arithmetic coding using Markov order of 1
(e) = Arithmetic coding using Markov order of 0

Table 7.8a Coding result of the predictor of VWT-LPC of Lena image using subsampling rate of 4, 9/7 biorthogonal filter in two levels decomposition, no amplitude shift before image transform. Prediction is intra-vector mean prediction.

| Uniform Scalar Quantizer Step Size | PSNR | Bit rate | Lossless Compression |
|---|---|---|---|
| 16.0 | 37.18071 | 0.13 | Lempel Ziv Coding (LZ77) |
| 16.0 | 37.18071 | 0.108 | Arithmetic Coding, Markov Order= 3 |
| 16.0 | 37.18071 | 0.104 | Arithmetic Coding, Markov Order= 2 |
| 16.0 | 37.18071 | 0.103 | Arithmetic Coding, Markov Order= 1 |
| 16.0 | 37.18071 | 0.119 | Arithmetic Coding, Markov Order= 0 |

Table 7.8b Coding result of the residual of VWT-LPC of Lena image using subsampling rate of 4, 9/7 biorthogonal filter in two levels decomposition, no amplitude shift before image transform. Prediction is intra-vector mean prediction. Threshold of VQ= 0.1

| Codebook size | PSNR (overall) | PSNR (residual) | bpp of residual (no lossless coding) | bpp of residual (no lossless coding) |
|---|---|---|---|---|
| 512 | 31.72 | 33.256 | 0.5625 | 0.56 (a) |
| | | | | 0.7 (b) |
| | | | | 0.7 (c) |
| | | | | 0.7 (d) |
| | | | | |
| 256 | 31.22 | 32.54 | 0.5 | 0.458 (a) |
| | | | | 0.5 (b) |
| | | | | 0.5 (c) |
| | | | | 0.499 (d) |
| | | | | |
| 128 | 30.75 | 31.9209 | 0.4375 | 0.402 (a) |
| | | | | 0.422 (b) |
| | | | | 0.424 (c) |
| | | | | 0.407 (d) |
| | | | | |
| 64 | 30.19 | 31.17 | 0.375 | 0.337 (a) |
| | | | | 0.42 (b) |
| | | | | 0.424 (c) |
| | | | | 0.407 (d) |

(a) = Lempel-Ziv coding (LZ77)
(b) = Arithmetic coding using Markov order of 3
(c) = Arithmetic coding using Markov order of 2
(d) = Arithmetic coding using Markov order of 1

Table 7.8c Overall coding result of VWT-LPC of Lena image using subsampling rate of 4, 9/7 biorthogonal filter in two levels decomposition, no amplitude shift before image transform. Prediction is intra-vector mean prediction.

| Codebook Size | PSNR | bpp (with lossless coding) |
|---|---|---|
| 512 | 31.72 | 0.663 |
| 256 | 31.22 | 0.561 |
| 128 | 30.75 | 0.505 |
| 64 | 30.19 | 0.44 |

Note: The bit rate is obtained by using the best lossless coding schemes that used in the predictor and residual.

Table 7.9a Coding result of the predictor of VWT-LPC of Lena image using subsampling rate of 4, 9/7 biorthogonal filter in two levels decomposition. Prediction is intra-vector mean prediction.

| Uniform Scalar Quantizer Step Size | PSNR | bpp (with lossless coding) |
|---|---|---|
| 10.0 | 40.537 | 0.16 (a) |
| | | 0.14 (b) |
| | | 0.14 (c) |
| | | 0.13* (d) |
| | | |
| 15.0 | 37.6 | 0.13 (a) |
| | | 0.106 (b) |
| | | 0.1* (c) |
| | | 0.12 (d) |
| | | |
| 20.0 | 35.61 | 0.11 (a) |
| | | 0.09 (b) |
| | | 0.09 (c) |
| | | 0.09 (d) |

(a) = Lempel-Ziv coding (LZ77)
(b) = Arithmetic coding using Markov order of 3
(c) = Arithmetic coding using Markov order of 2
(d) = Arithmetic coding using Markov order of 1
The meaning of * and **, please refer to table 7.7b

Table 7.9b Overall coding result of VWT-LPC of Lena image using subsampling rate of 4, 9/7 biorthogonal filter in two levels decomposition. Prediction is intra-vector mean prediction. Threshold of VQ= 0.1

| Codebook size | PSNR (overall) | bpp (residual code with lossless coding) |
|---|---|---|
| 512 | 32.48* | 0.676 # |
| | | |
| 256 | 31.89* | 0.588 (a) |
| | | 0.63 (b) |
| | | 0.63 (c) |
| | | 0.629 (d) |
| | | 0.629 (e) |
| | | |
| 128 | 31.4* | 0.53 (a) |
| | | 0.551 (b) |
| | | 0.554 (c) |
| | | 0.537 (d) |
| | | 0.527 (e) |
| | | |
| 64 | 30.71* | 0.467 (a) |
| | | 0.49 (b) |
| | | 0.476 (c) |
| | | 0.49 (d) |
| | | 0.45 (e) |
| | | |
| 16 | 29.42** | 0.3 (a) |
| | | 0.264 (b) |
| | | 0.254 (c) |
| | | 0.25 (d) |
| | | 0.26 (e) |
| | | |
| 8 | 28.8* | 0.275 (a) |
| | | 0.236 (b) |
| | | 0.231 (c) |
| | | 0.232 (d) |
| | | 0.238 (e) |

(a) = Lempel-Ziv coding (LZ77)
(b) = Arithmetic coding using Markov order of 3
(c) = Arithmetic coding using Markov order of 2
(d) = Arithmetic coding using Markov order of 1
(e) = Arithmetic coding using Markov order of 0
* = coding of predictor use is uniform quantized of step size 10.0 and using arithmetic coding of Markov order 1.
** = coding of predictor use is uniform quantized of step size 15.0 and using arithmetic coding of Markov order 2.
# = no lossless coding has been done

Table 7.10 Vector Quantization on separate wavelet subband with IBBA, using Lena image, 9/7 biorthgonal filter in two levels and no lossless coding. Lena image is in the training set. The overall bit rate is 0.25 bpp. Threshold of VQ = 0.01

| Band | vector size | codebook size | bpp | PSNR |
|---|---|---|---|---|
| LL2 | 4x4 | 65536 | 1.0 | 41.145167 |
| HL2 | 4x4 | 256 | 0.5 | 24.716 |
| LH2 | 4x4 | 256 | 0.5 | 28.672 |
| HH2 | 4x4 | 256 | 0.5 | 30.04697 |
| HL1 | 4x4 | 4 | 0.125 | 31.559 |
| LH1 | 4x4 | 4 | 0.125 | 35.651 |
| HH1 | 4x4 | 4 | 0.125 | 39.0339 |

Final overall PSNR= 31.62dB

Table 7.11 Vector Quantization on separate wavelet subband with IBBA, using Lena image, 9/7 biorthgonal filter in two levels and no lossless coding. Lena is not in the training set. The overall bit rate is 0.25 bpp. Threshold of VQ = 0.01

| Band | vector size | codebook size | bitrate (bpp) | PSNR |
|---|---|---|---|---|
| LL2 | 4x4 | 65536 | 1.0 | 16.407339 |
| HL2 | 4x4 | 256 | 0.5 | 24.305837 |
| LH2 | 4x4 | 256 | 0.5 | 28.251178 |
| HH2 | 4x4 | 256 | 0.5 | 29.708291 |
| HL1 | 4x4 | 4 | 0.125 | 31.735694 |
| LH1 | 4x4 | 4 | 0.125 | 35.533961 |
| HH1 | 4x4 | 4 | 0.125 | 38.958959 |

Final overall PSNR= 26.405371dB

Table 7.12 Vector Quantization on separate wavelet subband with IBBA, using Lena image, 9/7 biorthgonal filter in two levels and no lossless coding. Lena is not in the training set. The overall bit rate is 0.25 bpp. Threshold of VQ = 0.01. The LL2 subband uses SPIHT

| Band | vector size | codebook size | bitrate (bpp) | PSNR |
|---|---|---|---|---|
| LL2 | use SPIHT | use SPIHT | 1.0 | 21.826818 |
| HL2 | 4x4 | 256 | 0.5 | 24.305837 |
| LH2 | 4x4 | 256 | 0.5 | 28.251178 |
| HH2 | 4x4 | 256 | 0.5 | 29.708291 |
| HL1 | 4x4 | 4 | 0.125 | 31.735694 |
| LH1 | 4x4 | 4 | 0.125 | 35.533961 |
| HH1 | 4x4 | 4 | 0.125 | 38.958959 |

Final overall PSNR= 29.336752dB

Table 7.13 Parental Finite State VQ on HL1 subband of the Lena image.

| Scheme | Band | No. of states | vector size | codebook size | bitrate (bpp) | PSNR of HL1 |
|---|---|---|---|---|---|---|
| I | HL1 | 256 | 8x8 | 256 | 0.125 | 32.291062 |
| II | HL1 | 4 | 4x4 | 16 | 0.25 | 32.224610 |

# Chapter 8

# Conclusion

In this thesis, we described two image coding algorithms, Vector Wavelet Transform (VWT) and Vector Wavelet Transform-Linear Predictive Coding (VWT-LPC). In addition, we introduced a bit allocation algorithm, the Inter-band Bit Allocation (IBBA), for wavelet subband coefficients. Finally, a constrained vector quantization algorithm, the Parental Finite State Vector Quantizer (PFSVQ), is introduced.

The VWT and VWT-LPC exploit both the intra vector correlation and inter vector correlation to improve the performance of image coding using VQ. Both algorithms use wavelet transform and vector quantization to encode an image. In general, the coding performance of VQ can be improved if the vectors have small inter-vector correlation and large intra-vector correlation. The merit of VWT and VWT-LPC is that they have the ability to increase the intra vector correlation and reduce the inter-vector correlation. VWT has reasonably good coding performance, for example, at the bit rate of 0.37 bpp, the PSNR is 28.17 dB. Another important thing is that this good performance is obtained with very little complexity.

To further improve the performance of VWT, a linear predictive coding is used together with VWT. This is the VWT-LPC. Although VWT successfully exploits the intra- and inter-vector correlation, however, at the same time, the energy compactness of the wavelet coefficients are reduced. This affects the coding performance. In VWT-LPC, a linear predictive coding is used to reduce the variance of the coefficients. As a consequence the codebook can give a

better representation of all the vectors. Therefore the coding performance can be improved. At the bit of 0.3 bpp, the VWT-LPC has PSNR= 30dB while that of VWT is 27.5 dB. There is much improvement of using VWT-LPC. VWT-LPC also has good performance when compared to other algorithms. At the bit rate of 0.25 bpp, the PSNR of VWT-LPC is 29.42 dB while that of subband coding with VQ is only 28.3 dB. Also another important thing of VWT-LPC is that it is very easy to implement.

The Inter-band Bit Allocation (IBBA) algorithm is designed for wavelet subband codings. It successfully allocates a total number of bits among different subbands. If each subband is encoded by separate VQ, then the number allocated to each subband refers to the codebook size for each subband. At the bit rate of 0.25 bpp, two level wavelet decomposition, the IBBA algorithm distributes 1 bpp to $LL_2$, 0.5 bpp to $HL_2$, $LH_2$, $HH_2$ and 0.125 bpp to $HL_1$, $LH_1$, $HH_1$. The encoding result using this algorithm has a PSNR of 31.62 dB at 0.25 bpp.

The last algorithm introduced in this thesis is the Parental Finite State Vector Quantization (PFSVQ) and we also described two schemes for it. The PFSVQ algorithm is an innovative finite state vector quantization algorithm and is used together with wavelet coding. In this algorithm, the state of the finite state vector quantizer for the current vector (under encoding) is the codebook index of the vector's parent vector. It is because if the parent vectors are coded with same codebook index, then they must have some similarities and there is correlation between a vector and its parent vector. Therefore, the PFSVQ can classified vectors into states according to their parents' codebook index, so that within each state, the vectors will have some similarities and can be coded more efficiently by VQ. Scheme I and scheme II are very similar, except in scheme I, the total number of states used in encoding a subband is equal to the total number of codebook index of the parent subband, however, in scheme II, the total number of states does not need to be equal to the total number of codebook index of the parent subband.

PFSVQ has reasonably good coding performance. In an experiment using two level wavelet transform and IBBA algorithm, applying PFSVQ on the subband $HL_1$, the PSNR of the subband $HL_1$ is 32.29 dB at the bit rate (of $HL_1$) of 0.125bpp. PFSVQ is a fixed-rate coding algorithm, therefore it has good error resilience performance. In addition, the decoder of PFSVQ is also very simple. This makes PFSVQ suitable for broadcasting or distribution applications like digital TV and digital video CD. These applications allow high complexity encoders but the decoders must have low complexity.

Finally, for the VWT-LPC, more research can be done on designing the LPC. It is because in our current research, the design of LPC is independent of the design of VWT. If the design of LPC and VWT become an inter-related issue, it is expected that the new LPC can improve the overall coding performance. Also, further research can be done on designing more sophisticated Parental Finite State VQ so that more information from the parent vector can be used for the determination of the state of an vector. It is expected that PFSVQ can give very good coding performance especially for wavelet transform coding because the parent-child relationship can be exploited in PFSVQ.

# REFERENCE

[1] Yih-Fang Huang, Che-Ho Wei, "Circuits and Systems in the Information Age," Short Courses at the 1997 International Symposium on Circuits and Systems.

[2] R. E. Ziemer and W. H. Tranter, "Principles of Communication: systems, modulation, and noise," Houghton Mifflin, c1976.

[3] Anil K. Jain, "Fundamentals of Digital Image Processing," Prentice-Hall International Editions 1989.

[4] Kuei-Ann Wen, Chung-Yen Lu, Ming Chas Tsai, "The transform image codec based on fuzzy control and human visual system," IEEE trans. on fuzzy systems, vol. 3, no. 3, August 1995.

[5] Thomas Sikora, "MPEG Digital Video-Coding Standards," IEEE signal processing magazine, Sep 97, vol. 14, No. 5, p.82-100.

[6] Nasser M. Nasrabadi, and Robert A.King, "Image Coding Using Vector Quantization: A Review," IEEE Trans. Commun., vol. COM-36, pp.957-971, Aug. 1988.

[7] Marc Antonini, Michel Barlaud, Pierre Mathieu, and Ingrid Daubechies, "Image Coding Using Wavelet Transform," IEEE Trans. on Image Processing, vol. 1, No. 2, Apr. 1992.

[8] Kwok-Tung Lo, Shing-Mo Cheng, "New Addressing Scheme for Vector Quantization of Images," Proc. IEEE Int. Conf. on Image Proc., vol. 3, p435-438.

[9] Gilbert Strang, Truong Nguyen, "Wavelets and Filter Banks," Wellesley-Cambridge Press, 1996.

[10] Rajan L. Joshi, Hamid Jafarkhani, James H. Kasner, Thomas R. Fischer, Nariman Farvardin, Michael W. Marcellin, and Reberto H. Bamberger, "Comparison of Different Methods of Classification in Subband Coding of Images," IEEE Trans. on Image Processing, vol. 6, no. 11, Nov. 1997.

[11] Robert M. Gray, "Vector Quantization," IEEE ASSP magazine, April, 1984.

[12] Weiping Li, Ya-Qin Zhang, "A Study of Vector Transform Coding of Subband-Decomposed Images," IEEE Trans. on Circuits and Systems for Video Tech., vol. 4, No. 4, Aug 1994.

[13] J. Wus, W. Li, Y.-Q. Zhang, "Image Coding Using Vector Filter Bank and Vector Quantization," Proc. IEEE ISCAS 1994 , pp. 133-136.

[14] M. A, Cay, W. Li, and Y.-Q. Zhang, "On the Optimal Transform for Vector Quantizaton of Images," Proc. IEEE Int. Symp. on Circuits and Syst., 1993, pp. 687-690.

[15] Weiping Li, Ya-Qin Zhang, "Vector-Based Signal Processing and Quantization for Image and Video Compression," Proc. of the IEEE, vol. 83, no. 2, Feb 1995.

[16] William H. Press, "Numerical Recipes in C: the art of scientific computing," Cambridge University Press, 1988.

[17] Z. Xiong, K. Ramchandran, and M. T. Orchard, "Space-frequency Quantization for Wavelet Image Coding," to appear in IEEE Trans. Image Processing, 1997.

[18] R.E. Crochiere, S.A. Webber, and J. L. Flanagan, "Digital coding of speech in subbands, " BELL. Syst. tech. J., vol. 55, pp. 1065-1085, Oct. 1976.

[19] J.H.Conway, N.J.A. Sloannem, "Sphere Packings, Lattices, and Groups," Springer-Verlag, second edition, 1992

[20] M. Antonini, M. Barlaud, P. Mathieu, "Image coding using lattice vector quantization of wavelet coefficients," Proc. IEEE ICASSP, 1991, pp.2273-2277.

[21] D.G. Jenog, J.D. Gibson, "Lattice Vector Quantization for Image Coding, " Proc. IEEE ICASSP, 1989, pp.1743-1746.

[22] ISO/IEC JTC1/SC29/WG11 N1387, "Short MPEG-4 Description," July 1996.

[23] Pamela C. Cosman, Robert M. Gray, Martin Vetterli, "Vector Quantization of Image Subbands: A Survey," IEEE Trans on Image Processing, vol 5, No 2, Feb 1996.

[24] Allen Gersho, Bhaskar Ramamurthi, "Image Coding Using Vector Quantization," Proc. IEEE Conf. Acoust., Speech, Signal Processing, vol. 1, pp. 428-431, May 1982.

[25] Allen Gersho, Robert M. Gray, "Vector Quantization and Signal Compression," Kluwer Academic Publishers, 1991.

[26] Alan V.Oppenheim, Alan S. Willsky with Ian T. Young, "Signals and Systems," Prentice-Hall.

[27] Thomas M. Cover, Joy A. Thomas, "Elements of Information Theory," John Wiley & Sons, Inc. 1991.

[28] William H. Equitz, "A New Vector Quantization Clustering Algorithm," IEEE Trans. Acoust., Speech, Signal Processing, vol. 37, pp1568-1575, Oct. 1989.

[29] Y. Linde, A. Buzo, R.M. Gray, "An Algorithm for Vector Quantizer Design," IEEE Trans. Commun., vol. COM-28, pp84-95, Jan. 1989.

[30] Martin Vetterli, Jelena Kovacevic, "Wavelet and Subband Coding", Prentice Hall PTR 1995.

[31] Jerome M. Shapiro, "Embedded Image Coding Using Zerotrees of Wavelet Coefficients," IEEE Trans. on Signal Proc., vol. 41, no. 12, Dec. 1993.

[32] W. Li and Y-Q. Zhang, "New Insights and Results on Transform Domain VQ of Images," Proc. IEEE Int. Conf. on Acoust,, Speech, and Signal Processing, Apr. 1993, pp. V609-V602.

[33] Amir Said, William A. Pearlman, "A New, Fast and Efficient Image Codec Based on Set Partitioning in Hierarchical Trees," IEEE Trans. on Circuits and Systems for Video Tech., vol. 6, no. 3, June 1996.

[34] J. Foster, R. M. Gray, and M. O. Dunham, "Finite-state vector quantization of waveform coding," IEEE Trans. Inform. Theory, vol. IT-31, p348-359. May 1985.

[35] M.J.T. Smith and S. L. Eddins, "Subband coding of images with octave band tree structure," Proc. ICASSP, p1382-1386, April 1987.