DISTRIBUTED VIDEO STREAMING USING MULTIPLE DESCRIPTION CODING

A DISSERTATION

Submitted in partial fulfillment of the requirements for the award of the degree

of

MASTER OF TECHNOLOGY

in

ELECTRONICS AND COMMUNICATION ENGINEERING

(With Specialization in Communication Systems)

By

AJNESH SINGH

DEPARTMENT OF ELECTRONICS AND COMPUTER ENGINEERING
INDIAN INSTITUTE OF TECHNOLOGY ROORKEE
ROORKEE - 247 667 (INDIA)

JUNE, 2011

CANDIDATE'S DECLARATION

I hereby declare that the work, which is presented in this dissertation, entitled **"Distributed Video Streaming Using Multiple Description Coding"** being submitted in partial fulfillment of the award of the degree of **Master of Technology** with specialization in **Communication Systems**, in the Department of Electronics and Computer Engineering, Indian Institute of Technology, Roorkee is an authentic record of my own work carried out from July 2010 to June 2011, under the guidance of **Dr. Debashis Ghosh,Associate Professor, Department of Electronics and Computer Engineering, Indian Institute of Technology, Roorkee.**

The results embodied in this dissertation have not submitted for the award of any other Degree or Diploma.

Date: 30 **June 2011**

Place: Roorkee

Ajnesh

(AJNESH SINGH)

CERTIFICATE

This is to certify that the statement made by the candidate is correct to the best of my knowledge and belief.

Date: 30 **June 2011**

Place: Roorkee

D Ghosh

Dr.DEBASHIS GHOSH

Associate Professor, E&CE Department

Indian Institute of Technology, Roorkee

Roorkee – 247667, (India)

ACKNOWLEDGEMENTS

I express my foremost and deepest gratitude to **Dr. Debashis Ghosh**, Department of Electronics & Computer Engineering, Indian Institute of Technology Roorkee, Roorkee for his valuable guidance, support and motivation throughout this work. The valuable hours of discussion and suggestions that I had with him have undoubtedly helped in supplementing my thoughts in the right direction for attaining the desired objective. I consider myself extremely fortunate for having got the opportunity to learn and work under his able supervision over the entire period of my association with him.

My sincere thanks to all faculty members of Communication Systems for their constant encouragement, caring words, constructive criticism and suggestions towards the successful completion of this work. My sincere thanks to laboratory staff to access the computers and other resources at will for the completion of this work.

Last but not the least, I'm highly indebted to my parents, whose sincere prayers best wishes, moral support and encouragement have a constant source of assurance, guidance, strength, and inspiration to me.

I am greatly indebted to all my friends, who have graciously applied themselves to the task of helping me ample morale support and valuable suggestions. Finally, I would like to extend my gratitude to all those persons who have directly or indirectly helped me in the process and contributed towards this work.

AJNESH SINGH

ABSTRACT

As developments in wireless networks continue, there is an increasing demand to support multimedia communications in such networks. The recent advances in multiple description (MD) video coding have made it highly suitable for multimedia applications in these kind of networks. Multiple description coding (MDC) has emerged as an attractive approach for video applications where retransmission is unacceptable. Video transmission over unreliable networks such as Internet or wireless networks suffers from various kinds of difficult conditions such as bandwidth fluctuation, burst error contamination, packet loss, and excessive packet delay due to network congestion. By taking advantage of multiple logical channels provided by digital networks, a MD video coding scheme is implemented to combat with varying network environments and to meet quality of service requirements.

In this thesis algorithms namely, SPIHT Video Coding as Single Description Coding (SDC) and SPIHT Video coding as Multiple Description Coding(MDC) is discussed in terms of rate distortion performance. The embedding bit-stream is obtained by exploiting the 3-D SPIHT algorithm. The coding efficiency of the video coder is optimized by removing spatio-temporal redundancy from video signals and constructing a more compact hierarchical zero-tree for 3-D SPIHT algorithm.

CONTENTS

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1 VIDEO CODING

In few decades with development of mobile and wireless communications, the user density increases rapidly but we have limited communication resources like bandwidth, power. The main communication resource is bandwidth; because bandwidth is limited we have to use it carefully. Video coding is used to reduce the bit rate at the output of encoder so that the bandwidth requirement is reduced to transmit the signal through a band limited channel. Coding in mainly used. In communication systems, Source Coding is the process of encoding the output of an information source into a format that can be transmitted digitally to a receiver as a series of code words such that the average length of the code words is as small as possible. This reduction in redundancy is also known as data compression. The smallest theoretically achievable average code word length for lossless compression is equal to the entropy of the source. Further compression is possible only if it is lossy. Coding is mainly classified in two types.

1.1.1 LOSSY CODING

In this type of coding digital audio, image, video where some errors or loss can be tolerated, it uses both data redundancy and human perception properties. Some forms of data compression are lossless. This means that when the data is decompressed, the result is a bit-for-bit perfect match with the original.

1.1.2 LOSSLESS CODING

In this types of coding exploit only data redundancy. This type of coding used mainly in computer programs. Lossless compression (coding) of video is possible, it is not used much, because lossy compression results in far higher compression ratios at an acceptable level of quality.

The compression is possible because of inherent redundancy present in the video signal. There are two types of redundancy present in the video, spatial redundancy and temporal redundancy. Spatial redundancy is found within the frame and temporal redundancy is found between different frames. The main task of video coding is to remove this inherent redundancy so that compression is achieved.

1.2 MULTIPLE DESCRIPTION VIDEO CODING

Multiple Description Coding (MDC) has recently emerged as attractive video or image coding technique for robust transmission over error prone channels. In multiple description coding we generate multiple sub-bit streams which is individually called a description. These descriptions are transmitted over independent channels to the receiver. The basic idea to generate multiple correlated description of the source such that each description independently describe the source with acceptable quality of video signal, When all descriptions are received, a high-quality reconstruction is possible and when we lost some descriptions due to channel error, we still get reasonable quality or acceptable quality of image or video. Video communication over wireless links therefore requires high compression and robustness against errors. Achieving both high compression and robustness against errors is a challenge because there is a big trade off between these two . The study carried out in [1] combats errors encountered in the wireless channel by coding the highly compressed video data into multiple independently decodable streams.

Multiple Description problem arises at Bell Systems Technical General in 1980, by N.S.Jayant as a channel splitting problem as "It is required to split a B bit/s speech code sequence into two Self Contained B/2 bits/s components, either of which can be used to reproduce acceptable speech,also if both components are available at a receiver,it must be possible to reproduce speech with full B bits/s quality". He experimented 3.2 KHz speech signal sampled at 12 KHz and coded using Pulse Code Modulation Technique. The output sequence of codewards is split into odd and even word sequences. A central receiver which uses both channels simply reconstitutes the out as prior to decoding, while a side decoder receiver with only the odd(or even) subchannel estimates the even(or odd) components by nearest neighbor interpolation.

The efficiently compressed visual data has to be transmitted over communication channels, such as wireless channels or networks. This raises the problem of error protection, since most of these channels are error-prone. Compressed images and especially video sequences are very susceptible to transmission errors. If the error occurs in a video frame, it may propagate further into subsequent frames because of motion-compensated prediction. Multi-view video compression methods utilize comprehensive temporal and inter-view prediction structures and therefore channel errors occurring in one view can propagate not only to the subsequent frames of the same view but also to the other views.

A common approach to error protection is to consider it as a pure channel problem, separate from the source compression problem. This approach is based on Shannon's work, which states that in principle the source and channel coding tasks can be carried out independently with no loss of efficiency. Following this approach, raw source sequences are processed in a way which reduces the data rate as much as possible. Reliable transmission of the bitstream to the receiver is provided then by a channel coder. The transport mechanism has to be perfect since a single error in the compressed bitstream might severely damage the reconstructed signal.

One method to combat the effect of errors present in the channel is Automatic Repeat Request(ARQ). There error-free transmission is achieved by retransmitting packets that have been lost or corrupted. A problem with such a mechanism is that it causes delays and thus requires larger memory buffers. The delay is at least a packet round-trip time. But each retransmission can also be lost, and the delay can be arbitrarily large. Because of delay we cannot use this method in real time video streaming (video conferencing).

Another approach for reliable transmission over lossy channels is so-called Forward Error Correction (FEC). the transmission of redundant packets that allow recovery of lost packets at the receiver side. The compressed bitstream data is distributed between packets and protected by block channel codes. The data from the lost packets can be reconstructed from the received packets. The choice of the block code length is important. In terms of efficiency, long blocks are preferred since short blocks generate bitstreams with a relatively large number of additional symbols. Just as with the previous approach based on retransmissions, a problem of delays and large memory buffers exists, this time caused by long blocks.

Two above-mentioned approaches tolerate no errors. They assume that all the data transmitted is correctly received. Consequently, they spend a considerable amount of resources to guarantee this. These expenses grow with the amount of data to be transmitted. One may consider multi-view video as a case of such a growing amount of data compared to single-view video. Pure channel coding approaches might not be quite feasible in such cases. As an alternative, one can tolerate channel losses. Assuming that not all data sent has reached the decoder, one can concentrate on ensuring efficient decoding of the correctly received data. In this case, one needs to change the source coding accordingly and, more broadly, to consider the error protection problem as a joint source-channel problem.

Multiple description coding (MDC) is a coding approach for communicating a source over unreliable channel. The source is encoded into several descriptions, which are sent to the decoder independently over different channels. A certain amount of controlled redundancy is added to the compressed descriptions to make them useful even when received alone. The decoder can reconstruct the source from one description with low, yet acceptable quality. The more descriptions received, the higher is the reconstruction quality. Usually, the descriptions are balanced; that is, the descriptions are of equal rate and importance. In that case, the reconstruction quality depends on the number of received descriptions only and not on which particular descriptions are received. The first multiple description problem arises in 1982 by Gamal and T. M. Cover [2],who given the achievable rates for multiple description coding.

The factor that makes MDC perform well even when only one description is successfully received is redundancy. MDC works by introducing redundant information among the descriptions. In general, the more the redundancy, the higher the quality of received video when decoded with only one of the descriptions. Of course, redundant data also consumes more channel capacity or transmission bandwidth. Thus, the research effort in MDC focuses on achieving high quality with as little redundancy as possible. Wireless links consist of bandwidth limited and error prone channels. Each channel may have different bandwidths.

In a communication system where the bandwidth and transmit power are set by regulation authorities, the channel capacity C is limited by the channel noise. This means that whenever the information source has an entropy $H > C$, we need to introduce a lossy source coder that reduces the entropy to a level below the capacity. Otherwise, we have no control over the distortion introduced in the communication process. The first step for a source coder is to remove all deterministic components, called redundancy. After removing natural redundancy which is present in image or video signal, we add some redundancy to combat the effect of errors which is present in the channel. This is called source coding(e.g.convolutional code).

General model for multiple description coding is as follows: A source sequence $\{ x_k \}$ is input to a coder, which outputs m streams at rates $R_1, R_2, \ldots R_m$. These streams are sent on m separate channels. There are many receivers, and each receives a subset of the channels and uses a decoding algorithm based on which channels it receives. There are $2^m - 1$ receivers, one for each distinct subset of streams except for the empty set, and each experiences some distortion. This is equivalent to communicating with a single receiver when each channel may be working or broken, and the status of the channel is known to the decoder but not to the encoder.

MDC Methods are given below

1.Multiple Description Scalar Quantizer (MDSQ) principle[3,4]

2.Multiple description coding using pairwise correlating transform[5]

3.MDC Using Pre –Post Processing [6]

4.Multiple Description using Zero Padding[7]

5.MD Coding using SPIHT

1.3 Multiple Description Coding As Joint Source Channel (JSC) Coding

When operating under a delay-constraint on a time-varying channel, it is generally no longer optimal to regard the two coders(source and channel coder) separately and we have to jointly optimize the source coder and the channel coder and the result is some sort of joint source-channel coding (JSC). Multiple description transform coding is a technique which can be considered a Joint Source Channel (JSC) coding for error prone wireless channels .The basic idea is to introduce correlation between transmitted coefficients in a known, controlled manner so that erased or lost coefficients can be statistically estimated from received coefficients. This correlation is used at the decoder at the coefficient level, as opposed to the bit level, so it is fundamentally different from schemes that use information about the transmitted data to produce likelihood information for the channel decoder. The multiple description coding is a common element of Joint Source Channel coding systems.

1.5 STATEMENT OF PROBLEM

Increasing multimedia applications through wireless communication networks, because wireless networks are error prone channels. So we have to use some technique to combat the errors, which are present in such channels. Forward Error Correction and ARQ is the techniques which used to combat the errors which present in the channels.

Multiple Description Coding (MDC) is immerged as new technique to combat the errors in channels. The objective of this dissertation is to study about Multiple Description Coding. The success of SPIHT algorithm on still image, motivates to apply it on video signals. In this thesis we discussed about use of SPIHT algorithms in MDC. We compared the Rate Distortion performance of MDC with Single Description Coding.

1.6 ORGANIZATION OF THE THESIS

The dissertation report has been presented in 6 chapters. Chapter 1 gives introductory discussion on the Multiple Description Coding and discussion about its application as Joint Source Channel (JSC) Coding. Then the motivation for the selection of the topic, about the problem. Chapter 2 discussed about the Multiple Description Coding(MDC) and performance measure Rate Distortion(R-D) criterion. Chapter 3 discussed about the brief introduction of different types of Multiple Description Coding. Chapter 4 discussed about Wavelet and multiresolution analysis, Embedded Zero Wavelet (EZW) Coding and Set Partitioning In Hierarchical Tree(SPIHT) its extension to 3-D SPIHT as Multiple Description Coding. Chapter 5 discussed about the results. Chapter 6 discussed about the conclusion and future scope of Multiple Description Coding.

Chapter 2

MULTIPLE DESCRIPTION CODING (MDC)

2.1 INTRODUCTION

Multiple descriptions coding (MDC) has emerged as an effective method for video transmission over unreliable networks or error prone networks where the channel is varying with time (i.e. sometimes we have good channel and sometime we have worst channel . Multiple Description Coding is a image and video coding, which is used in error prone networks or channels. In MDC we create n outputs bit streams from a video source which we want to code. The different bit streams, which is called a description. After the coding process, we send the descriptions or transmit the description through a single channel or multiple channel to achieve path diversity. Path Diversity is used in modern communication systems for reliable transmission of image and video through error prone channel where the channel having non-zero probability of error. At the receiving end or decoder, we can reconstruct the transmitted image or video by using single description or any subset of descriptions. The quality of transmitted video depends on the descriptions available at the output. If we get all descriptions correctly at the decoder then we get high quality of video as we transmitted at the transmitter. when we lost a description, due to some packet loss in the channel, we can still get the reasonable quality of video at the decoder using other descriptions.

In few recent years because of increasing multimedia applications, the reliable transmission of the video stream over lossy environments has become an important challenge. Multiple Description Coding (MDC) has become a vital video coding approach to enhance the error resilience of a video multimedia system. MDC corresponds to the introduce the redundancy at the source in order to generate multiple correlated streams with equal rate and equal importance that can be transmitted over different channels and decoded independently. So that each description alone provides low but acceptable quality and both descriptions together lead to higher quality. The first reason for the increasing popularity of MDC is that it can provide reasonable quality of video without packet retransmissions. This implies we can use MDC mainly for real-time applications such as conferences or TV transmission, because in these situations we don't want to retransmit the packets because of delay problems. The transmission of digital information over communication channels has become very challenging with the increasing demand of services such as multimedia applications. These applications naturally require significant amount of bandwidth and storage. This has been a strong motivation for development of efficient source coding techniques in order to reduce the required memory and bandwidth. Source coding methods can be classified as either lossless or lossy. The goal of both types of source coding is to encode the source into a compressed digital representation that can be used for transmission or storage.

The primary aim in the lossless source coding is to generate a compressed representation of the source that can be decoded to reconstruct the original signal without any single error. Because of high bit rate, In lossless coding, we require high bandwith and storage space for transmission of image and video. We use lossy video coding techniques to overcome these problems. Though lossy source coding technique does not generate the exact reconstruction of the original source, an acceptable approximation can be generated that achieves higher compression ratio than the lossless techniques. The performance of lossy source coding mainly depends on the data itself. In general, if we have a source in which we have more redundant information so we can achieve high compression ratio, because of redundancy, with reasonable quality of reconstructed image or video at the decoder.

If consider a video source coding problem in which we want to transmit source information over a communication channel which having some probability of error. We can say, the encoder sends many packets of data through the communication channel, and at the receiving end or decoder, we either receive the packet correctly or the packet is lost due to non zero probability of error of channel. To provide more reliability for such a data communication scenario, we may provide a diversity system with more than one channel for packet delivery. In this case, the probability of all channels breaking down is much less than that in the one-channel case. We are hoping that in the case of multiple channels at least some packets are received by the decoders and, as a result, we can better avoid the situation where the receiver does not receive anything at all. Consequently, we need to generate more than one description of the source packet, and we refer to this as the multiple description coding problem.

If we consider a Multiple Description Coding system with two channels. Now if we send the same information over both of the channel and suppose we get both the descriptions at the decoder, then fifty percent of the received information has no meaning. This implies the importance of sending different information over each channel. Every description in this case must be generated such that if one description lost due to some channel probability of error and we receive only one description correctly at the receiver decoder side, the received description at decoder is sufficient to satisfy a minimum fidelity criterion. However, if both descriptions are received at the receiver, the information received from one channel can be used to refine the information from the other channel in order to achieve a higher fidelity.

This Multiple Discription Coding problem and its relation to information theorey were first discussed by papers [13] in 1980 IEEE Information Theory Workshop. Though , the practical motivations of multiple description coding goes back to 1970s . Like other communication technology, the Multiple Description Coding was also invented at Bell Laboratories for speech communication over the telephone network. The telephone system needs to be reliable. But failure of transmission links are common. Thus, a mechanism is required to manage this failure in transmission and achieve high reliability in telephone network. In this context of speech coding, Jayant and Christensen developed a technique for combating speech quality degradation due to packet losses. Two different packets are created by even and odd sampling of a pulse amplitude modulation (PCM) bit stream. For instance, if only an even sample packet is lost, data contained in the odd packet is used to estimate the missing samples like predictive coding.

Digital video or multimedia applications have become very popular nowadays, the demand of image and video data transmitted over networks is rapidly increasing due to internet and other communication systems. International standards have been developed , such as MPEG and H.263, to meet this increasing demand OF video applications . These standards are made to achieve a high compression ratio and in addition we get the encoded bit stream which is more error resilient to transmission errors and packet loss. Several techniques have been developed so far to protect the video signal against transmission errors. As an example, error concealment methods which is proposed by Chen, Weng in 1997, which attempts to conceal the erroneous or lost blocks by making use of received information in the adjacent blocks. However, in the context of packet networks, the lost packets could result in a loss of synchronization, error propagation in the video signal, and severe data damage. In such cases, all the blocks are corrupted until the next synchronization codeword. This makes the error concealment technique inefficient, since no information is available from the adjacent blocks. Scalable video coding is another approach that encodes the video signal using two layers . The base-layer bit stream consists of the essential information such as the motion vectors, and is transmitted with a high level of priority. The second-layer bit stream contains the less important information, which may be ignored if it is lost. The problem with this approach is that the base-layer can also be lost in a highly congested network and cause severe degradation in the decoded video quality.

Multiple Description (MD) video coding is an alternative approach to enhance the robustness of the transmitted video signal, and combat the effect of packet loss. Similar to a general MDC scheme, the video signal is encoded into two correlated descriptions, and then transmitted over separate channels to the decoder. If both descriptions are received, the decoder provides a high-quality reconstruction of the original video. On the other hand, if one of the descriptions is lost during the transmission, the decoder estimates it from the received description, and then provides a lower but acceptable quality reconstruction.

MDC has emerged as a promising approach to enhance the error resilience of communication systems, especially, multimedia delivery systems. A primary reason for the increasing popularity of MDC is that it can offer sufficient reconstruction quality while retransmission of the lost information is not required. This makes MDC particularly appealing for real-time interactive multimedia applications such as video conferencing, since retransmission often incurs unacceptable delay. In recent years, MDC has been extensively explored in the area of multimedia communication, e.g., multiple description (MD) image coding , MD video coding , MD audio coding and MD speech coding .

The MDC problem can be depicted with a simple two-description model in Figure. A source X is encoded with two descriptions and sent to the receiver through two noisy channels. It is assumed that loss can occur in either channel. The central decoder processes information sent over both channels while the side decoders receive information only from one channel. Let $R1$ and $R2$ be the transmission rates for channel 1 and channel 2, respectively. MDC poses conflicting requirements. If a good description at rate $R1$ is sent over channel 1 and another good description at rate $R2$ is sent over channel 2, there is no reason for the two descriptions together to be as good as a single description at rate $R1 + R2$. Similarly, a good compressed representation at rate $R1 + R2$ cannot be easily split into two useful descriptions. The MDC model leads to several challenges in information theory and practical compression.

A generic multiple description (MD) video codec for two descriptions is shown in Fig.2.1 If one description is received, the side decoder reconstructs the source with a higher but reasonable or acceptable distortion . If side decoder 1 receive description then distortion is D1 ,and D2 is the distortion corresponding to side decoder 2. D0 is the distortion at the decoder when both the description received at the decoder. Rate distortion theory [1] says that D0 is always less then D1 and D0 is also less than D2.

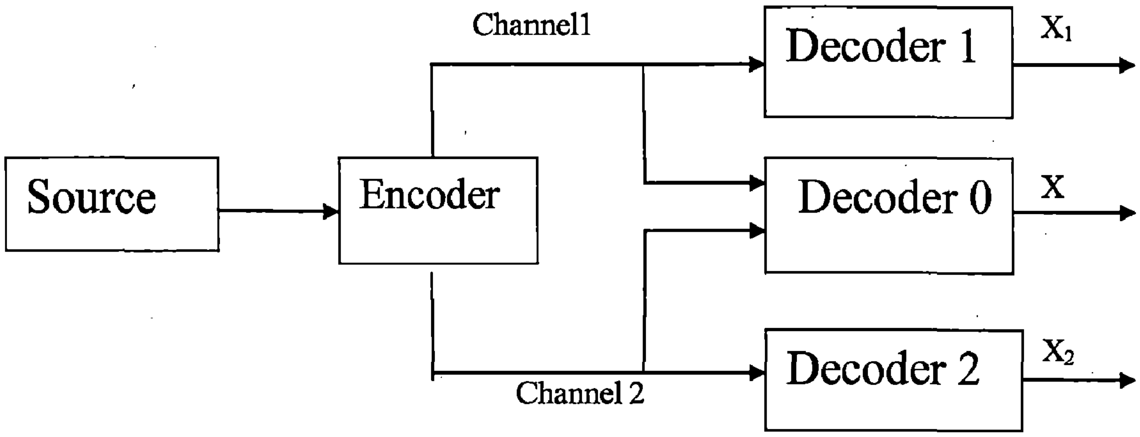


Fig. 2.1 Scenario for MD coding with two channels and three receivers[19]

2.2 MULTIPLE DESCRIPTION RATE DISTORTION REGION

Multiple description coding has its roots in information theory. A number of studies have addressed the information-theoretic aspects of MDC and studied bounds on achievable rates and distortions thus providing tools for optimizing practical MDC schemes. In this section I will discuss about the information theoretic works on MDC to form a basis for better understanding the problems discussed in the subsequent sections.

In source coding problem we basically concentrate on the compression. In lossy coding technique we use quantizer which lossy part (because there is no longer one to one mapping between input and output of quantizer) of coder. So we introduce distortion by compression. Rate Distortion(R-D) Theory theory calculates the minimum transmission bit-rate R for a required picture quality. Qualitatively we can say the distortion is proportional to compression,if we compress more ,more will bwe the distortion.

In single description coding, a rate-distortion pair (R,D)[2] is called achievable if there exist a source code with rate R and distortion D. The rate distortion (RD) region is then defined as a closure of the set of all achievable rate-distortion pairs (R,D) . Theoretical works on MDC consider generalizations of the RD region to the case of multiple descriptions. Most of these studies consider a classical multiple description (MD) case with two channels and three decoders. Some papers describe achievable regions in the case of many descriptions. Unfortunately, even for the two-channel case, there is no general result for the multiple description RD regions yet. Nevertheless, the MD RD region has been found for several special cases of interest.

A classical MDC scenario is shown in Fig 2.1. Consider a sequence of i.i.d. random variables $X_1, X_2, \ldots X_N$. A variable $\mathbf{X} = (X_1, X_2, \ldots X_N)$ is coded into two descriptions: Description 1 and Description 2 with rates R1 and R2 respectively. The descriptions are independently sent over Channel 1 and Channel 2. The receiving side has three decoders. Decoder 1 gets the information from Channel 1 only, while Decoder 2 gets the information from Channel 2. Decoder 0 gets the information from both channels. Decoder 0 is also called a central decoder, while Decoder 1 and Decoder 2 are called side decoders. The Decoder i estimates the variable **X** as **^Xi**, where i = 0, 1, 2. The distortion measures d1, d2, and d0 are given and the distortions corresponding to each description are

$$d_i(x^N, x_i^{N}) = E[\frac{1}{N}\sum_{k=1}^{N} d_i(x_k, x_{ik})] \leq D_i \qquad \text{i=0,1,2} \tag{2.2}$$

Where $X^N = (X_1, X_2, \ldots, X_N)$, and $\hat{X}^{N_i} = (\hat{}xi1, \hat{}xi2, \ldots, \hat{}xiN)$. Unlike the single description case, the rate distortion relations for multiple descriptions have not been defined in general in terms of information-theoretic quantities such as entropy, mutual information etc. The MD region can be defined as a set of all achievable quintuples $(R1, R2, D0, D1, D2)$.

2.3 ACHIEVABLE REGION IN MDC

Consider the classical MDC scenario with two channels and three decoders (Fig 2.1). A pair (R1,R2) is called an achievable rate for a given distortion **D** = (D1,D2,D0) if there exists a sequence of N pairs of descriptions f1($\mathbf{x}_N$), f2($\mathbf{x}_N$) with rates R1 and R2 and reconstruction functions $\hat{}\mathbf{x}_{N0}$ (f1, f2), $\hat{}\mathbf{x}_{N1}$ (f1), $\hat{}\mathbf{x}_{N2}$ (f2) such that, for sufficiently large N, the following inequality is met [26].

$$E[\frac{1}{N}\sum_{k=1}^{N} d_i(x_k, \hat{x}_{ik})] \leq D_i \qquad \text{i=0,1,2......} \tag{2.3.1}$$

The rate distortion region R(**D**) for distortion **D** = (D1,D2,D0) is defined as a closure of the set of achievable rate pairs (R1,R2) satisfying (2). An achievable rate region is any subset of the rate distortion region.

The theoretical studies were going on in 1980s to find rate-distortion bounds governing the tradeoff between the dual channel distortion ($D0$) and single channel distortion ($D1;D2$) for various sources.An achievable rate region has been defined by El Gamal and Cover [2] as follows. For a sequence of i.i.d. finite alphabet random variables $X_1,X_2, \ldots$ with the probability function $p(x)$, and the distortion measure di an achievable rate region for distortion **D** = (D_1,D_2,D_3) is given by the convex hull of all (R_1,R_2) satisfying
El Gamal and Cover [2] found an achievable rate distortion region represented by rate-distortion sets (R_1,R_2,D_0,D_1,D_2) for an i.i.d. source:

$$\begin{aligned} &R_1 \geq I(X;\hat{X}) \\ &R_2 \geq I(X;\hat{X}_2) \\ &R_1 + R_2 \geq I(X;X_1,\hat{X}_2,\hat{X}_0) \end{aligned} \qquad (2.3.2)$$

$E[d(X,\hat{X}_i) \leq D_i$, i=1,2,3.........and $I(.)$ is Shannon mutual information. The achievable rate region as defined by (2.3.2) gives sufficient conditions for the quintuple (R_1,R_2,D_0,D_1,D_2) to be in the MD rate distortion region. To know the rate distortion region completely, one would need to find necessary and sufficient conditions.

This has been done for the interesting case of "no excess rate sum", i.e. R1+R2 = R(D0). Ahlswede has shown [14] that for the no excess rate case the El Gamal–Cover's conditions (2.3.2) are necessary as well as sufficient. Thus, for this particular case, these conditions define the MD rate distortion region.

It has been guess that El Gamal–Cover's region (2.3.2) is tight in general and provides a complete RD region. By a counterexample, Zhang and Berger have shown that sometimes these conditions are not tight when R1+R2 > R(D0). Thus, generally, (2.3.2) does not provide the complete RD region for multiple descriptions and should be considered just as an inner bound.

The MD region has been found for the special case of a memoryless Gaussian source and mean squared error distortion. For this source and distortion measure, Ozarow has shown that the El Gamal–Cover's achievable rate region is also a complete RD region [2].

The rate distortion function of any memoryless source can be bounded by the rate distortion function of a Gaussian source with the same variance. Therefore, the knowledge of the RD region of a Gaussian source is rather important and can be utilized in practical MDC designs. Consider the source in Fig. 1 as a sequence of i.i.d. random variables {Xk} having Gaussian distribution with variance σ2. The distortion measure for all tree decoders is squared error di(x, ˆxi) = (x − ˆxi)2, i = 0, 1, 2. The distortion-rate function D(R) of a Gaussian source is [8]

$$D(R) = \sigma^2 2^{-2R} \tag{2.3.4}$$

The obvious outer bound is

$$D_i \geq D(R_i) = \sigma^2 2^{-2R_i} \tag{2.3.5}$$

$$D_0 \geq \sigma^2 2^{-2(R_1+R_2)} \tag{2.3.6}$$

The Shannon's source coding theorem implies that the RD function can be approached arbitrarily close when the codeblock length N approaches infinity [8]. If this held for the MD case, then (2.3.4,5,6) would define the achievable region. However, Ozarow has shown [15] that the actual achievable set of quintuples (R_1,R_2,D_0,D_1,D_2) for Fig.2.3 is formed by points satisfying the below given conditions.the achievable MD region is completely known only for special sources. Ozarow [15] shows that the achievable region for memoryless Gaussian sources with zero mean and unit variance and squared-error distortion criterion is given by the set of points satisfying.

$$D_1 \geq 2^{-2R_1} \tag{2.3.7}$$

$$D_2 \geq 2^{-2R_2} \tag{2.3.8}$$

$$D_0 \geq 2^{-2(R_1+R_2)} \frac{1}{1-(\sqrt{\Pi}-\sqrt{\Delta})^2} \tag{2.3.9}$$

Where $\Pi = (1-D_1)(1-D_2)$ and $\Delta = D_1 D_2 - 2^{-2(R_1+R_2)}$

The last term in the expression for D_0 shows that there is a penalty on D_0 for small values of distortions D_1 and D_2. The lowest central distortion is achieved when Π=Δ and therefore $D_1+D_2=\sigma^2(1+2^{-2(R1+R2)})$. Consequently, if either D_1 or D_2 is small, the other side distortion must be near σ2. That means that the second description is almost useless by itself [15]. Conversely, for $R_1 = R_2$ and $D_1 = D_2$, the central distortion is not better than half of the side distortion. For a small value of D this is far worse than the value $D0 \geq 1\ \sigma^2 D2$ given by the bound (2.3.7, 8, 9).

In other words, two descriptions together are twice as good as one if they do not need to be individually very good. On the contrary, if the side distortion constraints are severe, then two descriptions will not work better than one, because they are in fact the same description [15].

In case of squared error distortion the inner and outer bounds for MD region have been established by Zamir . A sketch of the inner and outer bounds is given in Fig.2.3. The inner bound is shown to be the Ozarow's MD region for the Gaussian source[15].

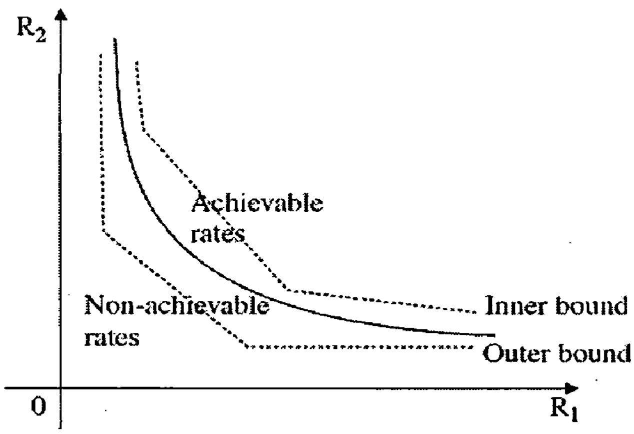


Fig. 2.3 Inner and outer bounds for rate region[15]

2.4 SUCCESSIVE REFINEMENT PROBLEM

Successive refinement of information is a special case of the multiple description problem. In the successive refinement problem, Decoder 2 is removed from the MDC scheme as given in Fig.3 Decoder 1 gets only the information from Channel 1 at rate R_1. Decoder 0 is said to refine the information from Channel 1. It gets the information from both Channel 1 and Channel 2 with rates R_1 and R2, respectively. The corresponding distortions are D1 and D_0. It is said that we are successively refining a sequence of random variables $\{X_k\}$ from distortion D_1 to distortion D_0 if the description of the source is optimal in every stage, i.e. $R_1 = R(D_1)$ and $R = R_1$

+ $R_2 = R(D_0)$. The problem is said to be successively refinable in general if the successive refinement from distortion D_1 to distortion D_0 is achievable for every $D_1 \geq D_0$.

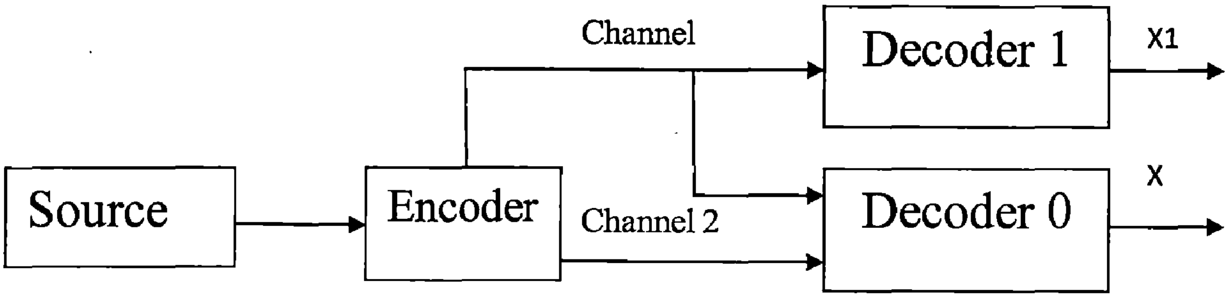


Fig.2.4. Successive refinement problem[16]

Chapter 3

A REVIEW ON MULTIPLE DESCRIPTION CODING

3.1 MULTIPLE DESCRIPTION SCALAR QUANTIZER (MDSQ)

A simple way to add error-protecting redundancy to the compressed image bitstream is to do this at the quantization stage, i.e. the stage where loss of insignificant information occurs. This idea has been extensively developed in the works of Vaishampayan, who has suggested a theory of multiple description scalar quantizers [3] .

The first practical design of multiple description quantizer was the multiple description scalar quantizer (MDSQ) proposed by Vaishampayan in 1993 . His method allows flexible adjustment of the weighting between central and side distortions. The key mechanism of Vaishampayan's technique is an index assignment (IA) scheme. In the case of two descriptions, the IA scheme labels each codeword of central quantizer by a pair of indices, one for each side quantizer. The process of MDSQ first quantizes a signal sample to a central quantizer codeword, then via index assignment maps this codeword to a pair of side quantizer indices. The two indices are encoded by fixed-length code in [3], so the corresponding scheme is called fixed rate MDSQ. The indices can also be encoded by a variable length code , the corresponding scheme is called entropy-constrained MDSQ.

Multiple description scalar quantization (MDSQ) works as follows. Two side (coarse) quantizers with overlapping cells operate in parallel at the quantization stage. The quantized source can be reconstructed from the output of either quantizer with lower quality. When the outputs of the two quantizers are combined, they produce higher quality reconstruction due to the resulting smaller quantization cells. In a practical scheme, the encoder first applies a regular scalar quantizer, mapping the input variable x to a quantization index I. Then, in a second step, an index assignment is applied, mapping each index I to a codeword index pair (i, j) in a codebook.

Figure3.1(a) [3] presents the index assignment matrix for the case of staggered index assignment. The cells of the quantizer corresponding to the index I are numbered in the matrix from 0 to 14. The row and column indices of the index assignment matrix form the index pair (i, j). Index i is included in Description 1 whereas index j is included in Description 2. The central decoder reconstructs the exact value of index I and the corresponding value ˆX0. The side decoders estimate X as an expected value when one of the indices is fixed. Thus, the quality of the side reconstruction is determined by a number of diagonals in the index assignment matrix. In Fig. 3.1(a), only 15 out of 64 cells in the index assignment matrix are occupied. Unoccupied cells constitute coding redundancy. Figure 3.1 (b) shows an index assignment with three diagonals filled and lower redundancy. The highest redundancy is achieved when only the main diagonal of the index assignment matrix is filled. This corresponds to duplication of all data in both descriptions. The side distortions are equal to the central distortion E[d0] = E[d1] = E[d2] and consequently the bitrate is doubled. If the index assignment matrix is full, there is no redundancy, resulting in high side distortions.

	0	1	2	3	4	5	6	7
0	0							
1	1	2						
2		3	4					
3			5	6				
4				7	8			
5					9	10		
6						11	12	
7							13	14

Fig.3.1 Index assignment: (a) Staggered quantization cells[3]

	0	1	2	3	4	5	6	7
0	0	2						
1	1	3	4					
2		5	6	8				
3			7	9	10			
4				11	12	14		
5					13	15	16	
6						17	18	20
7							19	21

(b)

Fig.3.2 Index assignment: **(b)** Higher spread quantization cells [3]

The high-rate analysis of MDSQ has been presented by V. Vaishampayan and the performance of MDSQ has been compared to rate-distortion bound [2] for squared error distortion and memoryless Gaussian source . Comparing the optimal entropy-constrained quantizer with the theoretical bound, a 3.07 dB gap between the product of the average central and side distortions d0d1 has been identified . It has been conjectured that this gap is caused by the non-spherical form of quantization cells. Therefore, the gap could be closed by constructing quantizers with more spherical like cells that is, cells with a smaller normalized second moment than a hypercube. Several solutions have been proposed, including trellis-coded quantization and multiple description lattice vector quantizer (MDLVQ). For the single description problem, one of the most important results of the quantization theory is that vector quantization is more efficient than scalar quantization, even if the source samples are independent random variables. It has been shown that MD vector quantizers are capable of closing the 3.07 dB gap when vector dimensions tend to infinity ($N \rightarrow \infty$). An improvement of 0.3 dB is achieved in the two-dimensional case using an MD hexagonal lattice quantizer .

Originally, MDLVQ was limited to the balanced case (equal rates $R_1 = R_2$ and equal distortions $D_1 = D_2$). Diggavi et al. generalized this method to asymmetric multiple description vector quantizers that cover the entire spectrum of the distortion profile, ranging from balanced to successively refinable descriptions. The improvements for MDLVQ provide more operating points and more flexible rate-distortion trade-off by the slight increase in complexity. A generalized MDVQ for more than two descriptions has also been designed [4].

3.2 MULTIPLE DESCRIPTION CODING USING PAIRWISE CORRELATING TRANSFORMS

MD approach considers adding redundancy immediately after the stage of transform coding by means of so-called pairwise correlating transform (PCT) [5]. The general framework is the following. First, the input signal is decorrelated using proper transform (e.g. DCT). The resulting coefficients are ordered according to their variances and coupled into pairs. These pairs undergo correlating transform, i.e. two uncorrelated coefficients at the PCT input give rise to two correlated coefficients at the PCT output. One transform coefficient is sent to Description 1 and the other is sent to Description 2. By the explicitly added redundancy within the pairs of coefficients, a lost coefficient from the pair can be estimated from the received one. When both descriptions are received, the exact values of variables can be determined by taking the inverse transform.

The basic idea here is to create correlation between the transmitted information. This correlation can be exploited in the case of a packet loss on one of the channels since the received packet due to the correlation will contain information also about the lost packet. In order to do this a piecewise correlating transform T [16] is used such that

$$\begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix} = T \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} \tag{3.2.1}$$

$$T = \begin{bmatrix} r_1 \cos\theta_1 & r_2 \sin\theta_2 \\ -r_1 \sin\theta_1 & r_2 \cos\theta_2 \end{bmatrix} \tag{3.2.2}$$

Here r1 and r2 will control the length of the basis vectors and $\theta 1$ and $\theta 2$ will control the direction. The transform is invertible so that

$$\begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = T^{-1} \begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix} \tag{3.2.3}$$

Based on the choice of r1, r2,$\theta 1$, $\theta 2$ a controlled amount of correlation, i.e. redundancy, will be introduced in Y1 and Y2 which are transmitted over the channels. The more redundancy introduced the lower side distortion will be obtained at the cost of an increased central distortion.To get some intuition about the behavior of the original method some of the theoretical results of [3] are reviewed: For the case when R1 = R2 = R and using high rate approximations it can be showed that when no redundancy is introduced between the packets, i.e T equals the identity matrix, the performance will behave approximately as

$$D^*{}_0 = \frac{\pi e}{6} \sigma_1 \sigma_2 \, 2^{-2R} \tag{3.2.4}$$

$$D^*{}_s = \frac{1}{4}(\sigma_1{}^2 + \sigma_2{}^2) + \frac{\pi e}{12} \sigma 1 \sigma_2 \, 2^{-2R} \tag{3.2.5}$$

where D*s is the average side distortion between the two channels. Using the transformation matrix.

We get

$$D_0 = \Gamma D^*_0 \qquad (3.2.6)$$

$$T = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \qquad (3.2.7)$$

$$D_s = \frac{1}{\Gamma^2} \frac{1}{4} (\sigma_1{}^2 + \sigma_2{}^2) + \Gamma \frac{\pi e}{12} \sigma_1 \sigma_2 \, 2^{-2R} \qquad (3.2.\ 8)$$

where

$$\Gamma = \frac{(\sigma_1{}^2 + \sigma_2{}^2)/2}{\sigma_1 \sigma_2}$$

Hence, the central distortion is increased (assuming $\sigma 1 > \sigma 2$) and the constant term in the average side distortion is decreased at the same time as the exponential term is increased. Two conclusions can be drawn; firstly MDCPC is not of interest when the rate is very high, since Ds is bounded by a constant term. Secondly, the method also requires unequal variances for the sources X1 and X2 since otherwise $\tau = 1$. The method is however efficient in increasing robustness with a small amount of redundancy.

An important tool for evaluating performance of MD codes is the redundancy rate-distortion (RRD) function [2]. The redundancy is determined as $\rho = R - R^*$. Here, R is the resulting rate of the MD coder with a central channel distortion D_0, and R* is the rate of the best single description coder for a given distortion D0. Thus, ρ is an additional bitrate needed for one-channel reconstruction. The RRD function is then defined as ρ(D1;D0), where D1 is the averaged one-channel distortion. Since ρ depends weakly on D_0, the RRD function could be written as ρ(D1) .

3.3 MULTIPLE DESCRIPTION USING FORWARD ERROR CORRECTION (MD-FEC)

A general approach is to equip multiple descriptions with forward error correction (MD-FEC) [9]. Its basic idea is to assign unequal numbers of FEC symbols to different parts of the compressed bit stream, depending on the information importance of these parts and their contribution to the overall reconstruction quality. This idea is best applied to so-called embedded bit streams, where the bytes of the compressed source are ordered according to their importance. The wavelet-based SPIHT encoder is an example of compression algorithm generating an embedded bit stream . Its first bytes are the more important ones and any subsequent byte refines the decoded image. Thus, the bit stream can be truncated based on the given bit budget leaving the reconstruction still possible. In connection with FEC, the first bytes should be better protected than the later bytes.

We illustrate how MD-FEC works by means of an example. Seventeen data symbols are coded using eight FEC symbols, thus, a total of 25 symbols is to be transmitted, with these broken into five codes, as shown in Table 3.1. FEC is implemented by means of Reed-Solomon (RS) codes. Stronger RS codes are applied to the data located at the beginning of the bit stream, i.e. to the more important data. Namely, (5, 2) codes are applied to symbols 1 and 2, (5, 3) codes are applied to symbols 3 to 8, (5,4) codes are applied to symbols 9 to 12, and symbols 13 to 17 are left unprotected. Then, the symbols, including FEC ones, are grouped vertically into multiple descriptions (packets). Each packet is protected with a parity code enabling error detection and sent to the receiver.

	D1	D2	D3	D4	D5
Code 1	**1**	**2**	**FEC**	**FEC**	**FEC**
Code 2	**3**	**4**	**5**	**FEC**	**FEC**
Code 3	**6**	**7**	**8**	**FEC**	**FEC**
Code 4	**9**	**10**	**11**	**12**	**FEC**
Code 5	**13**	**14**	**15**	**16**	**17**

Table 3.3 Example of MD-FEC

At the receiver side, the decoder detects erroneous descriptions and uses RS codes to reconstruct the lost data. As (5, 2) RS code can sustain a loss of three symbols, receiving any two descriptions makes it possible to decode symbols 1 and 2. Similarly, receiving any three descriptions makes it possible to decode symbols 1 to 8. When no descriptions are lost, all 17 symbols are reconstructed. For reconstruction with RS codes, it does not matter which descriptions are lost, only the number of lost descriptions matters. Thus, MD-FEC generates inherently balanced descriptions (having the same size and resulting in the same distortion when lost).

3.4 MDC USING PRE –POST PROCESSING

Another simple way of generating multiple descriptions is based on the theory of zero padding. Padding zeros in time domain can result in interpolation in frequency domain. On the other hand, padding zeros in frequency domain results in interpolation in time domain.

The coding scheme proposed in [6] pads certain number of zeros in both horizontal and vertical directions in DCT domain. Coefficients of new size are inverse transformed using IDCT and results in an over-sampled frame with bigger size. Pixels are more correlated now and then video source is divided into two descriptions by spatial sub-sampling. The two descriptions are independently coded at the encoder. If one description is lost, it can be estimated by the other one. This is proved to be much better than directly spatial sub-sampling the original frame. It performs very well especially at low bit rate, and is simple to implement.

The redundancy can also be added at the preprocessing stage before the polyphase transform . Such a pre-processing procedure is shown in Fig. The input image is transformed to DCT domain to an array of size $D \times D$, which is then padded with zeros to the size of $(D \times M)$ $(D \times M)$. The obtained $(D \times M)$ $(D \times M)$ representation is transformed back to spatial domain, where it is split into multiple descriptions by polyphase sub sampling. Clearly, the redundancy can be adjusted by the padding parameter M.

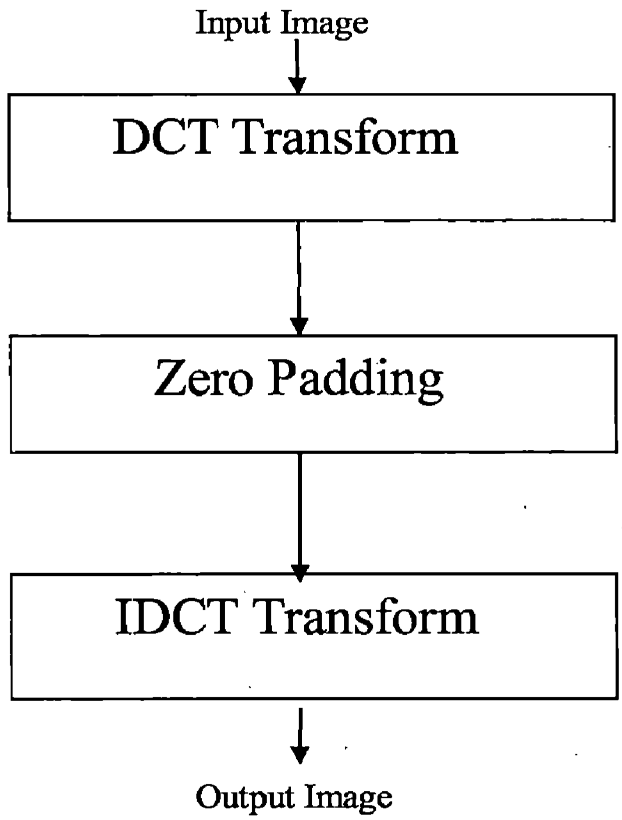


Fig.3.5 Pre and Post Processing[6]

3.5 MDC USING ZERO PADDING

A new method base on one-dimensional DCT has been proposed in [7]. Each frame is transformed using one-dimensional DCT on each column, and then padded with zeros vertically. After one-dimensional IDCT on new-sized columns, the new frame is sub-sampled into two sub frames and independently coded. As have been mentioned, after zero padding, pixels are more correlated and can lead to higher side decoding quality when one description is lost. In central decoder the two reconstructed sub frames are merged together. Then DCT, removing padded zeros and IDCT is performed to get the frame needed. When only one description is received, the other description is estimated using correctly received description and then merged. Results show that this new method performs better than the two-dimensional padding technique, at a much lower computational complexity.

3.6 MD VIDEO CODING USING MOTION COMPENSATED TEMPORAL PREDICTION

The coding scheme which is called MD video coding using motion compensated temporal prediction (MDMCTP) [10] utilizes MD transform coding to generate two descriptions and has three prediction loops. To reduce mismatch between encoder and decoder, this coders uses one central prediction loop and two side prediction loops at the encoder, to get the three possible scenarios at the decoder: (1) both descriptions are received. (2) only description 1 is received, and (3) only description 2 is received. Side prediction is based on a single description and the mismatch signals are quantized. At the decoder, when one description is lost, reconstructed mismatch signals are added to the available side predicted frame. At the central prediction loop, pair wise correlating transform (PCT) is used for paired DCT coefficients to generate two descriptions, The PCT operation has the effect of re-introducing a controllable amount of correlation between pairs of coefficients. Therefore, a wide range of tradeoff between coding efficiency and error resilience can be easily achieved. Simulation results show that it is important to have side prediction loops and transmit some redundancy information about the mismatch. This coder performs significantly better than traditional SD video coders over packet lossy channels. However, the achievable redundancy range is quite high for acceptable qualities. The coder proposed in [9] also has similar three-prediction-loop structure, yet it employs a second-order predictor for motion compensation, which predicts a current frame from two previously coded frames. This coder generates two description, containing the coded even and odd frames, respectively, which is similar to VRC. Compared to MDMCTP coder, it can operate at a much lower redundancy range for the same range of side distortion.

3.7 MD CODING USING SPIHT

Success of SPIHT algorithm at still image opens new scope for MD video coding. The details of algorithm and application as Multiple Description are given in next chapter.

CHAPTER 4

MDC USING 3-D SPIHT

4.1 WAVELET AND MULTIRESOLUTION ANALYSIS

It is common in signal processing to investigate signal properties in the frequency domain. Qualitatively we can say the variation of the signal is related to some frequency components. If at some place the signal changes fast in time domain, it can be thought as having high frequency components at that point, on the other hand if signal varies slowly at some place, it can be said as having low frequency components at that point.

In an image, the boundary of an object usually has an abrupt change from its background. So high frequency components are found at such boundaries. The background image varies slowly, implying low frequency components. A transform captures these low frequency background information and high frequency boundary information by concentrating the energy of the image into coefficients at the respective frequency sub bands.

The traditional method of obtaining the spectral domain is by the fourier transform, however, In fourier transform domain, although the energy information at a particular frequency is provided, no time information is provided about time. For example, if an impulse function is examined in the frequency domain, nothing can be say at what time the impulse occurred because its spectrum spans the whole frequency axis. The short time fourier transform(STFT) attempts to solve this problem by cutting the signal into several parts and analyzing each each part. The results are frequency time domain representation with a uniform time and frequency division. The Heisenberg uncertainty principle says that either the time resolution can be very high or the frequency resolution cab be very high, but not both. If one wants to examine a signal carefully at some particular locations while coarsely at other location, STFT is not a good choice.

The wavelet transforms [11] offers a solution remedy, the high frequency part is analyzed with finer time resolution and the lower frequency part is analyzed with coarser time resolution. This is desired since usually we care less about the exact locations of low frequency components, but we do care about the exact locations of high frequency components, where the boundary information is preserved.

The discrete wavelet transform used in this coding is similar to hierarchical sub band decomposition. The decomposition is done by dividing the image into four sub bands. The low frequencies bandwidth range is $0<|\omega|<\pi/2$, and the high frequencies bandwidth range representation is $\pi/2<|\omega|<\pi$. The four sub bands are obtained when horizontal and vertical filters are applied to the image. The sub bands LH1, HH1 and HL1 represent the very finest details of the image. For the second level of decomposition LL1 band is further divided into sub bands LL2, HL2, LH2 and HH2 as shown in figure.

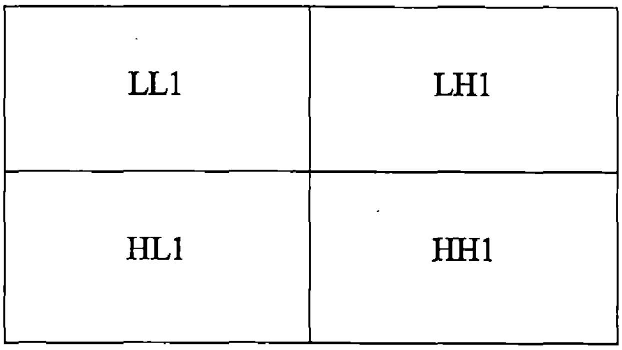


Fig4.1 First level hierarchical subband decomposition[11]

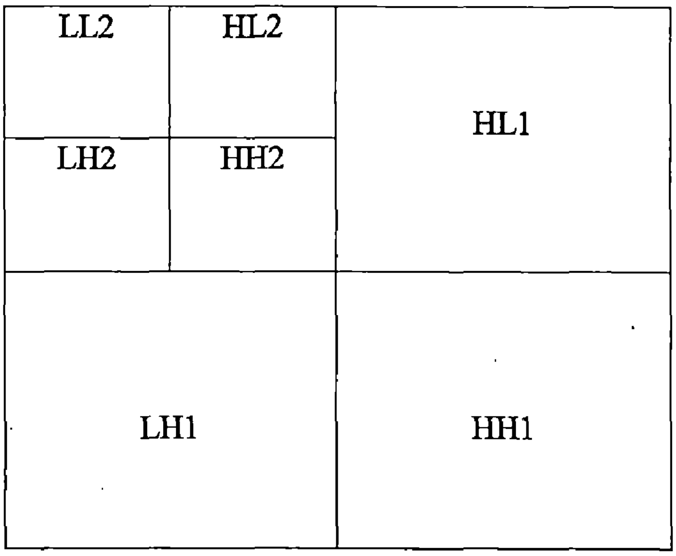


Fig.4.2Second level hierarchical subband decomposition[11]

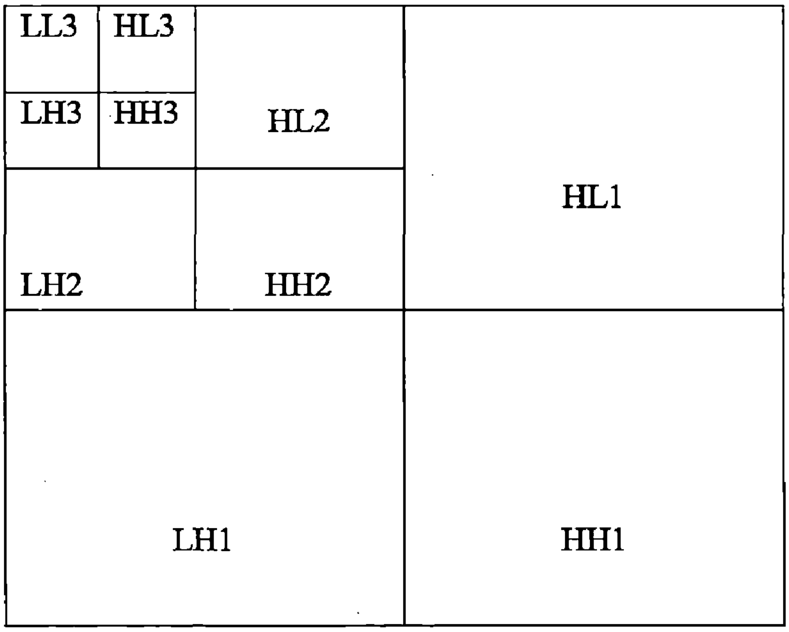


Fig.4.3 third level hierarchical subband decomposition[11]

To obtain more levels of decomposition, the LL subband is further sent through the filters and further divided into subbands. Each coefficient after decomposition is said to occupy 2x2 area of the total original image.

4.2 EMBEDDED ZERO-TREE WAVELET (EZW) CODING

J.M Shapiro [12] proposed this simple yet remarkable image compression algorithm which is embedded in nature the bits in the bit stream are generated according to the order of importance. The embedded coding compression algorithm lets the encoder to cease the encoding bit stream at any time when the target bit rate is met. Also, the decoder can also cease the decoding at any point and can still get back the same image as the image we get from decoding the entire bit stream.

The embedded zero tree wavelet algorithm (EZW) [12] is an effective image compression algorithm. This new technique produces a fully embedded bit stream for image coding. Also, the compression performance of this algorithm is competitive with virtually all known techniques. Furthermore, this technique requires definitely no training, no pre-stored codebooks or tables and requires no preceding knowledge of the image source. The EZW algorithm is based on four principal concepts: a discrete wavelet transform or hierarchical sub-band decomposition, prediction of the absence of significant information across scales, entropy coded successive-approximation quantization and universal lossless data compression which is accomplished via adaptive arithmetic coding. The EZW algorithm includes: a discrete wavelet transform (DWT) [11]. Adaptive arithmetic coding is used to achieve a fast and efficient method for entropy coding. We want to call attention to the following sentence. 'This algorithm runs consecutively and stops whenever a target bit rate or a target distortion is met. Certainly, one of the most significant parts of this algorithm is the encoding process. As mentioned above, in our case, discrete wavelet transform is used. The original image is passed through discrete wavelet transformation to produce transform coefficients. The aim of this transformation is to produce de-correlated coefficients and remove dependencies between samples as much as possible. Besides, the more significant bits of precision of most coefficients are allowed by this transformation to be efficiently encoded as part of exponentially growing zero trees. This transformation is considered to be lossless. The produced transform coefficients are then quantized to produce a stream of symbols which will be used in compression. Almost all of the information loss occurs in the quantization stage. For embedded coding, the successive approximation quantization is used which allows the coding of multiple significance maps using zero trees, and allows the encoding and decoding to stop at any point. Therefore, in quantization stage the threshold value is arranged to determine significance. The coding process allows the entropy coder to incorporate learning into the bit stream itself. As a final part of encoding, the data compression stage takes the stream of symbols and attempts to represent the data stream with no loss as efficiently as possible. In decoding operation converse of these concepts will be applied. In the decoding process after receiving the compressed image data, entropy decoding, de-quantization and inverse DWT will be applied consecutively. In the decoding operation, each decoded symbol refines and reduces the width of the uncertainty interval in which the true value of the coefficient may occur. A brief literature search has only found two published results where authors generate an actual bit stream that claims higher Peak Signal to Noise Ratio (PSNR) performance at rates between 0.25 and 1 bit/pixel, the latter of which is a variation of the EZW. The performance of the EZW coder was also compared widely available version of JPEG. It is seen that JPEG does not allow the user to select a target bit rate.

The important features of EZW compression algorithm are

1. The discrete wavelet transform that gives the multiresolution representation of the image.

2. The unique zero tree coding of the multiresolution representation which is wavelet transformed provides the significant maps by which we can understand the positions of the significant bits. The zero tree coding predicts the significant coefficients to represent the exponentially growing trees.

3. The successive approximation in it allows the multi precision representation of the significant coefficients thus making the algorithm embedded in nature.

4. The wavelet coefficients are prioritized, and ordering is done by considering the magnitude and the spatial location of the coefficients. The larger magnitude coefficients are given more importance than the smaller.

5. The multilevel adaptive arithmetic coding provides the entropy coding of the strings and the sequential coding of the algorithm. It stops whenever the target bit rate is achieved.

4.3 SET PARTITIONING IN HIERARCHICAL TREE (SPIHT)

4.3.1 INTRODUCTION

The Embedded zero tree wavelet image compression was proposed by Shapiro, which is a best image compression method in terms of compression efficiency and computational complexity.

Before EZW the image compression technique are based such that the compression efficiency is proportional to the complexity (e.g. for high compression we need more complex circuitry) but Shapiro interrupted this process by proposing EZW.

Fig.4.4 Block Diagram of SPIHT Coder[20]

4.3.2 Coefficient Ordering in Progressive Image Transmission

The hierarchical subband transformation [20] may be expressed in a general form as

$$C = \Omega(P) \tag{4.1}$$

where, P is the original image array, c is the transformed coefficient array and Ω is the unitary hierarchical subband transformation. Both the original image array and the coefficient array have the same dimensions. The encoder transmits the coefficients and the decoder updates itself according to the received bit-stream. From the approximated coefficient array $\hat{C}$, it is possible to reconstruct an approximated form of image through the inverse transformation, given by

$$P = \Omega^{-1}(C) \tag{4.2}$$

The mean-squared reconstruction error of the image at the decoder is given by

$$D_{mse}(p - \hat{p}) = \frac{1}{N}\sum_{n1}\sum_{n2}(P_{n1,n2} - \hat{P}_{n1,n2})^2 \tag{4.3}$$

where, N is the total number of pixels and $P_{n1,n2}$ is the pixel intensity at the location (n1,n2).

The subband transformation being lossless, the mean square error will be invariant to the transformation, This implies that in an embedded bit-stream, the large-valued coefficients should be transmitted first as they contribute more to the reduction of mean-square error and better reconstruction quality. This justifies why in an embedded bit-stream generation, the transform coefficients should be ordered according to the magnitude. We had already seen that the EZW algorithm follows this by ordering the significant coefficients in the subordinate pass. The idea of coefficient ordering can be extended to bit-planes if the coefficients are ranked according to their binary representations and the most significant bits are transmitted first.

4.3.3 SPATIAL ORIENTATION TREE

The spatial orientation tree, illustrated in Fig.4.5 defines the spatial relationship between the subbands in the form of a pyramid composed of a recursive four-band split. Each node of the tree corresponds to a pixel and is identified by its pixel coordinate. Other than the leaves, each node of the tree has four offspring corresponding to the pixel at the same position in the next finer level of the pyramid of same orientation, as shown by arrows in the diagram. The only exceptional case is the LL subband existing at the highest level of pyramid. Pixels in this subband form the root and groups of adjacent 2x2 pixels are composed. Other than one of the pixels (marked as '*') out of these four, all remaining three pixels have their four offspring in the HL, LH and HH subbands of the same scale, as shown. One out of the four is obviously left out since only three subbands exist for determining the descendants. The spatial orientation tree discussed as above has close resemblance with the hierarchical tree structure that was used to examine the zerotrees and the zerotree roots in EZW algorithm. However, there is a major difference too. In the hierarchical tree of EZW, every LL subband pixel at the highest level has three offspring at the HL, LH and HH subbands, whereas the offspring relationship of LL subband pixels in the spatial orientation tree is as what has been already discussed in the last paragraph. Spatial orientation trees are groups of wavelet transform coefficients organized into trees rooted in the lowest frequency or coarsest scale subband with offspring in several generations along the same spatial orientation in the higher frequency (resolution) subbands. Fig.4.5 depicts the key for parent-offspring relationship of coefficients to tree nodes for a 2-D wavelet transform with two levels of decomposition. In the spatial orientation trees, each node consists of adjacent pixels, and each pixel in the node has four offspring, except at the highest level of the pyramid, where one pixel in a node indicated by "*" in this figure does not have any offspring. Spatial orientation trees were introduced to exploit self-similarity and magnitude localization properties in a 2-D wavelet-transformed image. In particular, if a coefficient magnitude in a certain node of a spatial orientation tree does not exceed a given threshold, it is very likely that none of its descendants will exceed that threshold.

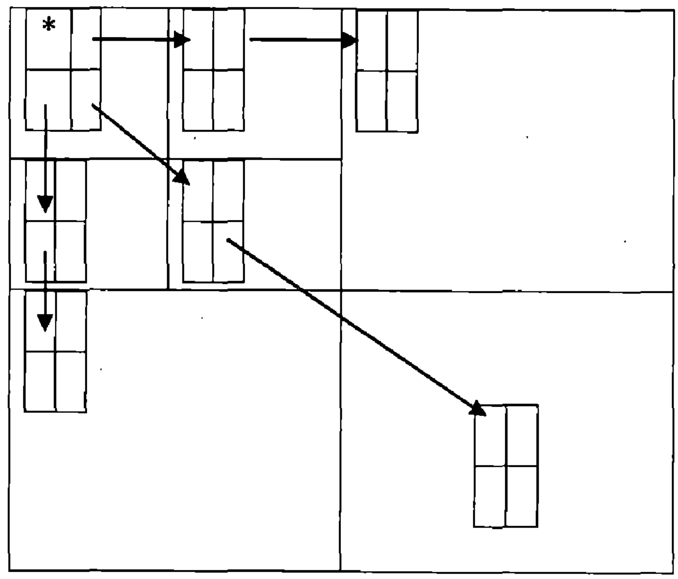


Fig. 4.5 The dydic wavelet decomposition(spatial orientation tree)[20]

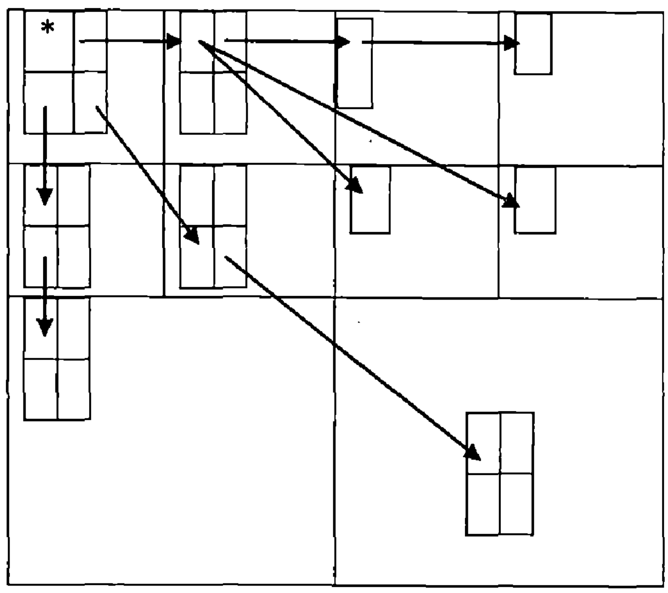


Fig.4.6 A wavelet packet decomposition case[20]

In the case of a wavelet packet transform, frequency bands other than the lowest may be recursively split. A node in the tree at a given level then becomes associated with one pixel at the same corresponding spatial location in each of the four subbands generated from the split. An example of the parent–offspring relationships in such a tree is shown in Fig. 4.6, where one of the high frequency bands is further split.

4.3.4 SET PARTITIONING IN HIERARCHICAL TREE(SPIHT)

The SPIHT algorithm utilizes three basic concepts:

(1) Searching for sets in spatial-orientation trees in a wavelet transform.

(2) partitioning the wavelet transform coefficients in these trees into sets defined by the level of the highest significant bit in a bit-plane representation of their magnitudes.

(3) coding and transmitting bits associated with the highest remaining bit planes first.

In set partitioning approach, the ordering information is not be explicitly transmitted. Instead, the encoder and the decoder follow the same execution path and if the decoder receives the results of magnitude comparisons from the encoder, it can recover the ordering information from the execution path. In set partitioning, no explicit sorting of coefficients is done. Instead, for a given value of *n*, the coefficients are examined if they fall within $2^n \le \mathrm{mod}(C_{n1,n2}) < 2^{n+1}$. If $2^n < \mathrm{mod}(C_{n1,n2})$, it is significant. Otherwise, it is insignificant.

The set partitioning rule is designed to work in the subband hierarchy. The objective of the set partitioning algorithm should be such that the subsets expected to be insignificant contain larger number of elements and the subsets expected to be significant should contain only one element.

SPIHT consists of two main stages, sorting and refinement. In the sorting stage, SPIHT sets a magnitude threshold , where is called the level of significance, and seeks to identify three entities in the spatial-orientation trees: isolated coefficients significant at level (magnitude no less than); isolated coefficients insignificant at level (magnitude less than); and sets of coefficients insignificant at level (all their magnitudes less than). For a given , the algorithm searches each tree, partitioning the tree into sets of the three entities above and moves their coordinates respectively to one of three lists: 1) the list of isolated significant pixels (LSP); 2) the list of isolated insignificant pixels (LIP); and 3) the list of insignificant sets (LIS). The last set can be identified by a single coordinate, due to the partitioning rule in the search, where the set of descendants having a significant member is split into its (four) offspring and a subset of all descendants of offspring. When a coefficient is tested and found insignificant, a "0" bit is emitted to the output bit stream and its coordinate is moved to the LIP for subsequent testing at lower . When a coefficient is found significant, a "1" bit and a sign bit are emitted and its coordinate is moved to the LSP. When an LIS set is tested for significance at level , a "0" bit is emitted if insignificant. But when found significant, a "1" bit is emitted and the set is partitioned into offspring and descendants of offspring. The offspring are moved to the end of the LIP and subsequently tested for significance at the same and also to the LIS as roots of their descendant sets that are subsequently tested for significance at the same .

The bit significance number is successively lowered in unit increments from the maximum of the largest magnitude coefficient. At a given , the th of every member of the LSP found significant at a higher is emitted to the codestream, adding to the "1's" in the th bit of the coefficients just found significant for the same . This is called the refinement stage of the algorithm. When is decremented, the LIP is tested for significance, and the test result is emitted as a "0" or "1" bit for insignificant or significant, respectively. If significant, its coordinates are moved to the LSP and a sign bit is emitted. Then the LIS is visited and its tree sets are partitioned according to the results of the significance tests. The process terminates when the desired bit rate or quality level is reached.

The decoder of the code bitstream receives the outputs of the significance tests and can therefore build the same lists, the LIP, LIS, and LSP, as in the encoder. Therefore, as input bits are read from the codestream, it reconstructs the magnitude and sign bits of LSP members seen by the encoder. The coefficients of the final LIP and LIS sets are set to zero. In the wavelet transform of an image, large sets of zero values exist which are identified efficiently by SPIHT with a single bit. Moreover, significant coefficients are never represented by more bits than needed in their natural binary representation, since the highest "1" bit is always known.

4.3.5 SPIHT CODING ALGORITHM

1) Initialization: output $n = \log_2 (\max(_{i,j})$ { mod $C(_{i,j})$})];

set the LSP as an empty list, and add the coordinates

(i, j) E *X* to the LIP, and only those with descendants

also to the LIS, as type *A* entries.

2) **Sorting Pass:**

2.1) for each entry (2, *J)* in the LIP do:

2.1.1) output *Sn(2, J) ;*

2.1.2) if S,(z, j) = 1 then move (2, j) to the LSP

and output the sign of *et,* ;

2.2) for each entry *(i, j)* in the LIS do:

2.2.1) if the entry is of type A then

Output *Sn(D(2, J)) ;*

if *Sn(D(z, 1))* = 1 then

for each *(k , I)* E O(z, *J)* do:

output *Sn(k,* 1);

if *Sn(k,I)* = 1 then add *(k , E)* to the

if *S,(k, 1)* = 0 then add *(I C , I)* to the

if *C(i, j)* # 0 then move ***(i,** j)* to the

end of the LIS, as an entry of type *B,*

and go to Step *2.2.2);* otherwise, remove

entry *(i, j)* from the LIS;

LSP and output the sign of *c k ;*

end of the LIP;

2.2.2) if the entry is of type *B* then

output *Sn(L(2, J)) ;*

if *Sn(L(z, j))* = 1 then

* add each *(k , I)* E O(z, *j)* to the end of

the LIS as an entry of type A;

* remove ***(i,** j)* from the LIS.

3) **Refinement Pass:** for each entry (2, *j)* in the LSP,

except those included in the last sorting pass (i.e with same *n),*

output the nth most significant bit

4) **Quantization-Step Update:** decrement *n* by 1 and go

to Step *2.*

4.3.6 3-D SPIHT WITH MDC

Now in this section we will study about the the extension of the concept of SPIHT still image coding to 3-D video coding. The algorithm[17] is very simple as case of 2-D SPIHT, while still providing high performance, full embeddedness, and precise rate control as 2-D spiht case.

3-D SPIHT scheme extended from the 2-D SPIHT, having the following three similar characteristics:

(1) partial ordering by magnitude of the *3-D wavelet* transformed video with a *3-D set partitioning* algorithm.

(2) ordered bit plane transmission of refinement bits; and

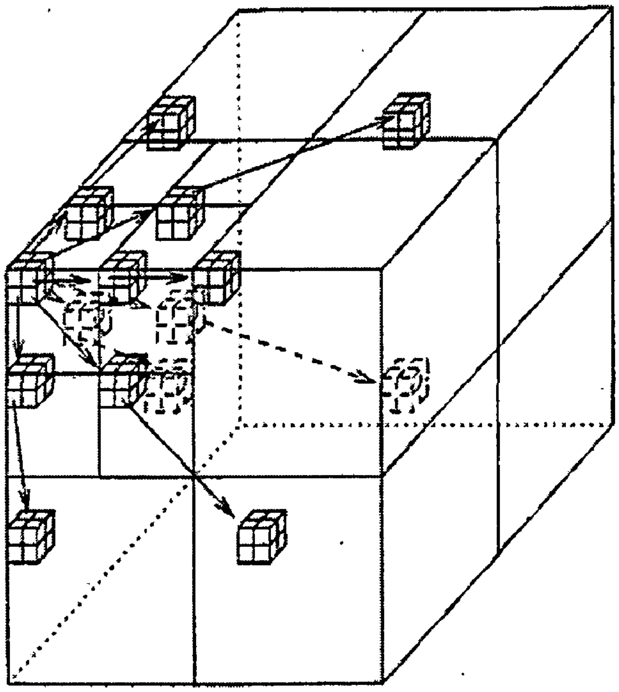
(3) exploitation of self-similarity across spatio-temporal orientation trees.

The resulting compressed bit stream is completely embedded, so that a single file for a video sequence can provide progressive video quality, i.e., the algorithm can be stopped at any compressed file size or let run until nearly lossless reconstruction is obtained, which is desirable in many applications including HDTV.

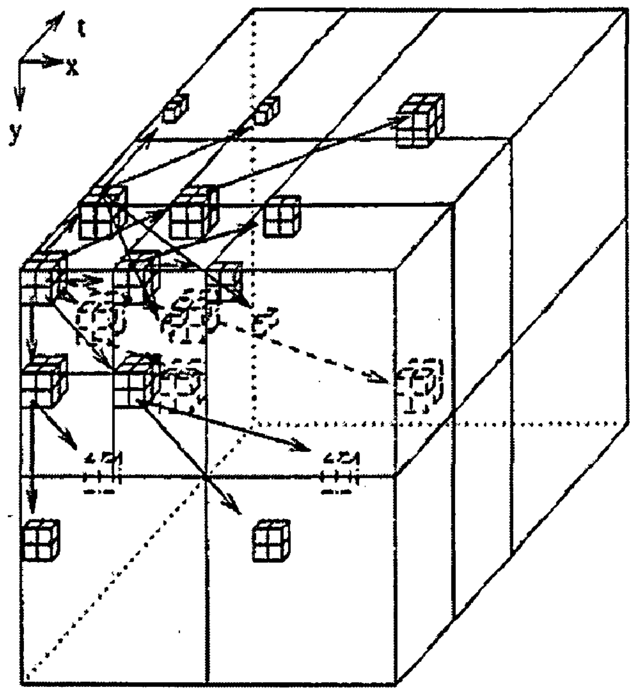
In the previous section, I discussed about the basic concepts of 2-D SPIHT. We have seen that there is no constraint to dimensionality in the algorithm itself, as pixels are sorted regardless of dimensionality. If all pixels are lined up in magnitude decreasing order, then what matters is how to transmit significance information with respect to a given threshold. In 3-D SPIHT, sorting of pixels proceeds just as it would with 2-D SPIHT, the only difference being 3-D rather than 2-D tree sets. Once the sorting is done, the refinement stage of 3-D SPIHT will be exactly the same.

A natural question arises as to how to sort the pixels of a three dimensional video sequence. Recall that for an efficient sorting algorithm, 2-D SPIHT utilizes a 2-D subband/wavelet transform to compact most of the energy to a small number of pixels, and generates a large number of pixels with small or even zero value. Extending this idea, one can easily consider a 3-D wavelet transform operating on a 3-D video sequence, which will naturally lead to a 3-D video coding scheme.

In 3-D SPIHT, we define a new 3-D spatio temporal orientation tree and its parent-offspring relationships. When pass the sequence through spatial filter and then through temporal filter,again through spatial filter then temporal filter,we get the decomposition which is is purely dyadic, a straightforward extension from the 2-D case is to form a node in 3-D SPIHT as a block



(a)



(b)

Fig.4.7. Spatio Tempral Orientation tree

with eight adjacent pixels, two extending to each of the three dimensions, we get a node of $2\otimes2\otimes2$pixels. The node which is at lowest subband is called root nodes .Root node have one pixel with no descendants and the other seven pointing to eight offspring in a $2\otimes2\otimes2$cube at corresponding locations at the same level. A pixel has eight offspring in a cube $2\otimes2\otimes2$one level below in the pyramid for nonroot and nonleaf node. We can see in Fig. 4.7 the parent-offspring relationships in the case of a two-level dyadic 3-D decomposition with 15 subbands which is generated by a once repeated spatial-horizontal, spatial-vertical, and temporal splitting, in that order. A wavelet packet transform, on the other hand, may produce a split of a given subband at any level into a number of smaller subbands, so that the $2\otimes2\otimes2$offspring nodes are split into pixels in these smaller subbands at the corresponding orientations in the nodes at the original level. Fig. 4.7 illustrates the parent-offspring relationships in 21 subbands produced by two levels of spatial-horizontal and vertical-dyadic splitting followed by a two-level temporal dyadic splitting. We have chosen packet transform in this 3-D SPIHT, this gives a different number of decompositions between the spatial and temporal dimensions , because of it Packet transform achieves better compression results compare to the purely dyadic decompositions. Therefore, we can decompose into more spatial levels than temporal ones. It is advantageous to limit the number of frames to be buffered to form a coding unit in order to save memory and coding delay. We do not want to limit the spatial decompositions to the same small number, since more spatial decompositions usually produce noticeable coding gains.

we can see in given Fig.4.8 that all trees are originates at the lowest spatio-temporal subband. When a tree originates from lowest subband new branches are adding in tree from different subband of wavelet dicompostion. The important point about this kind of tree structure is that we can selectively prune the branches of a tree and still get a reduced quality picture. The main idea in this kind of tree structure is that,it can be used for multiple description coding in which a description is generated by using branch pruning technique .The one way to make a description is ,we can use the branches which is originating only at even-indexed coefficients in a frame row can be retained in a description, and the other description can be made by using tree which is originating odd-indexed coefficients. The amount of redundancy can be controlled by retaining certain important branches which originating fron lower subband. The tree structure which used for multiple description coding is given in the Fig.4.8,in which 2 spatial and 3 temporal decomposition is used.

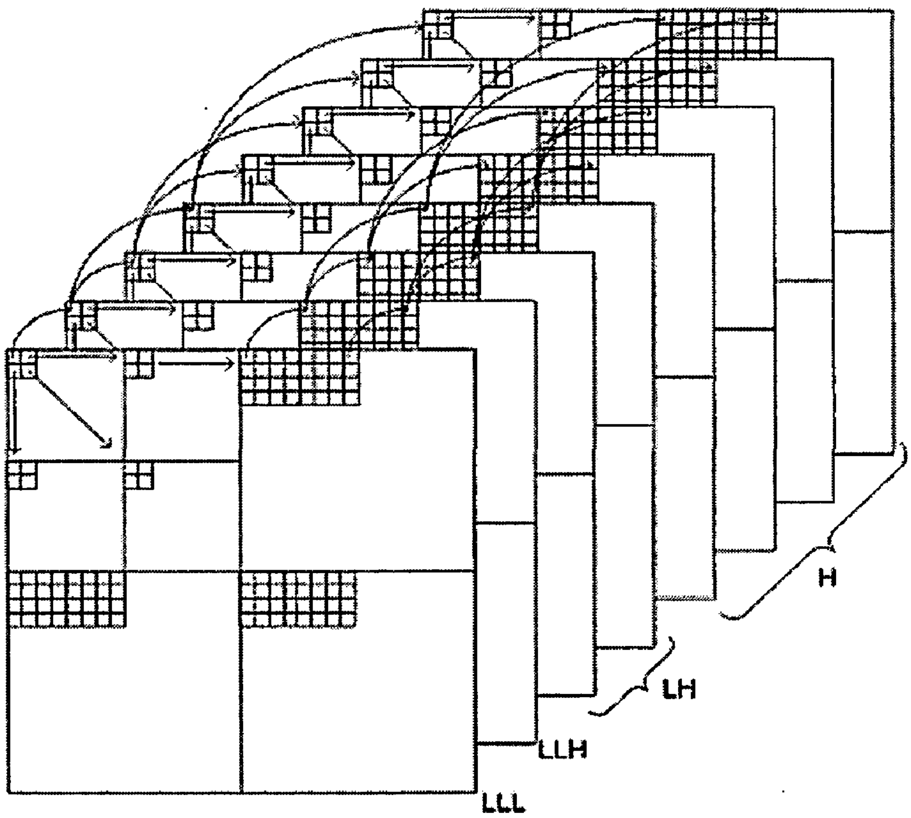


Fig.4.8 Two spatial and three temporal levels [19]

At lower subband ,we have tree structure or parent offspring relationship which is same as 3D(AT)-SPIHT, when we move toward upper subband,we modify the tree structure in both temporal and spatial directions. Both spatial and temporal high frequency coefficients are important, so we take care of both temporal and spatial high frequency coefficients. There are more branches in coefficients which belongs to lower subband. We don't ignore the coefficients which lie in lower subbands ,which makes redundancy in descriptions, so clusterd loss problem is not arise in this case.

CHAPTER 6

RESULTS

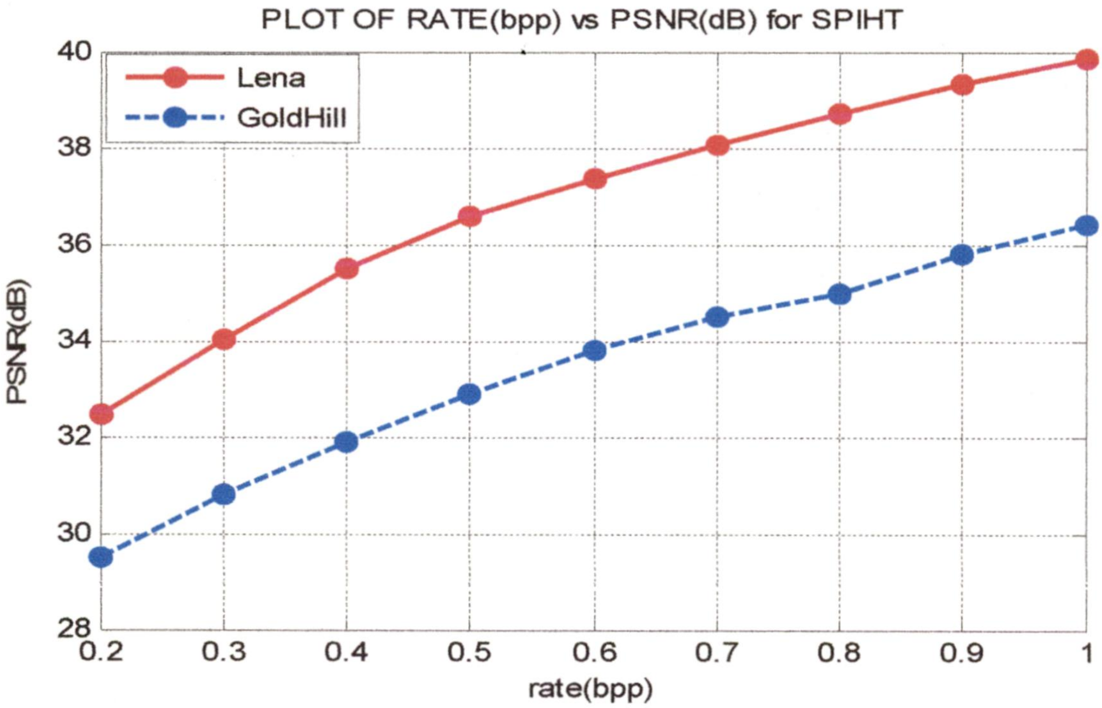


Fig.6.1 Comparision result of PSNR value between lena and goldhill image

Input Image	Wavelet Name	Decomposition level	Bitrate(kbps)	PSNR(dB)
Lena(512x512)	9/7	5	0.3	34
Goldhill(512X512)	9/7	5	0.3	30.4

Table.6.1 Parameters used in SPIHT

The above Fig.6.1 shows the implementation results of SPIHT algorithm for image LENA and GOLDHILL.. Which uses 9/7 biorthogonal filter as wavelet decomposition. Filter 9/7 have the best energy compaction properties. The resulting wavelet coefficients applied to SPIHT coder and resulting bit strem is the entropy coded by arithmetic coding. If we increase the Decomposition level the efficiency of algorithm increases.

The formula which is used to calculate the PSNR is

$$PSNR = 10\log_{10}(\frac{255^2}{MSE})\,dB \tag{6.1}$$

The algorithm performs better for lena image compared to goldhill image. The filters and the formule for calculating PSNR is same for the below results.

The below Fig 6.2 and Fig 6.3 shows the simulation result for FOREMAN and CARPHONE video. I have taken 8 frames in a Group Of Frames (GOF). Then performed 2 level of spatial decomposition and 3 level of temporal decomposition. Each experiment is performed for 296 frames.

Algorithm	Test Video	Filter	Spatial Decomposition	Temporal Decomposition	Avg. PSNR(dB)	Avg. Bitrate(kbps)
SPIHT SDC	Foreman	9/7 bior	2	3	27.7	180
SPIHT MDC	Foreman	9/7 bior	2	3	28	180
SPIHT SDC	Carphone	9/7 bior	2	3	29.7	180
SPIHT MDC	Carphone	9/7 bior	2	3	30.3	180

Table 6.2 Comparisons of SDC SPIHT and MDC SPIHT for different Test signals

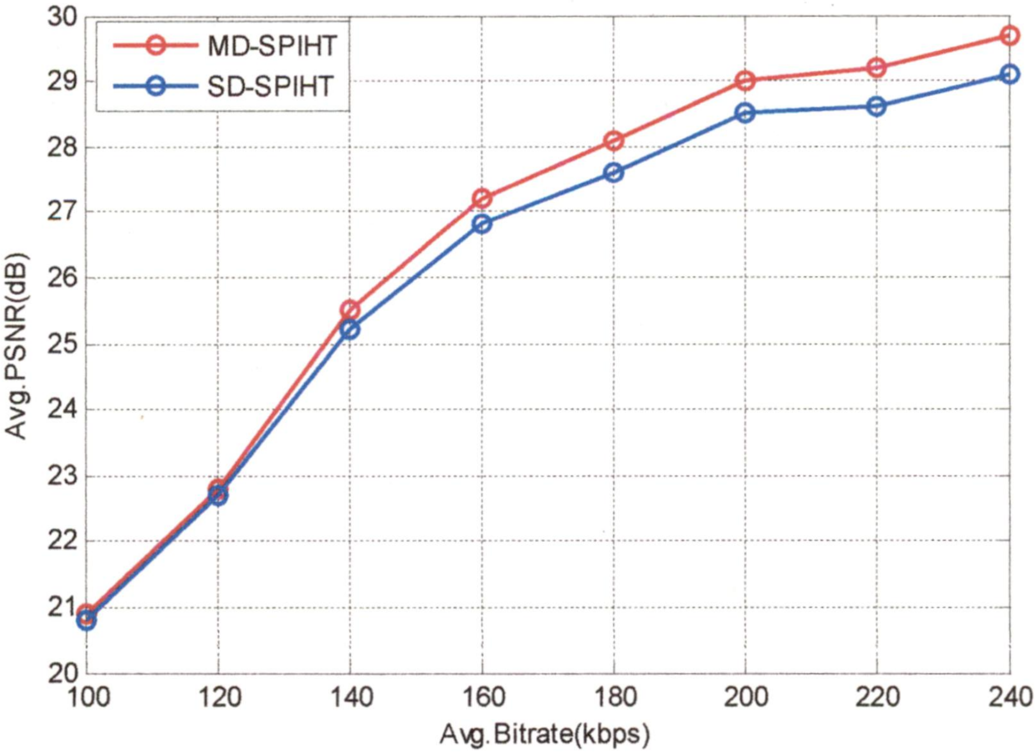


Fig.6.2 Comparision between SDC and MDC for Foreman

In Fig 6.2 I have compared the PSNR performance Multiple Description Coding(MDC) SPIHT with Single Description Coding(SDC) SPIHT.I have taken FOREMAN video as test signal. We can see the MDC SPIHT performs better then SDC SPIHT. We get 0.1-0.3 dB gain at low bitrate and we get gain of 0.3-0.6 dB at high bitrate.

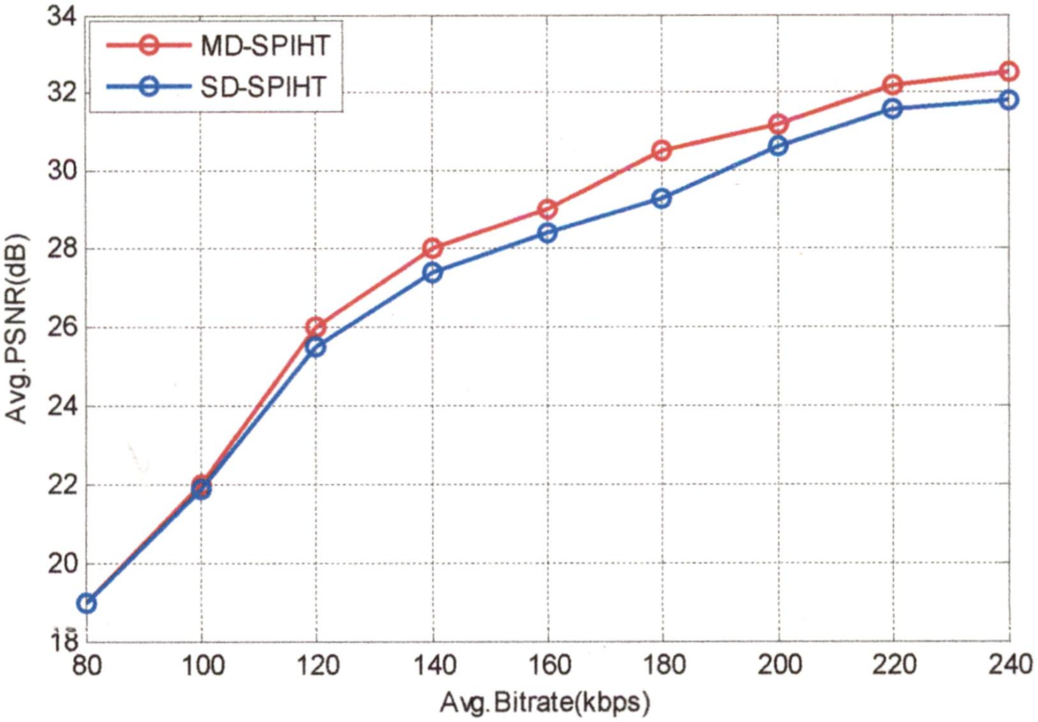


Fig.6.3 Comparision between SDC and MDC for Carphone

In Fig 6.3 I have compared the PSNR performance Multiple Description Coding(MDC) SPIHT with Single Description Coding(SDC) SPIHT. We have taken CARPHONE video as test signal. We can see the MDC SPIHT performs better then SDC SPIHT. We get 0.1-0.3 dB gain at low bitrate and we get gain of 0.3-0.6 dB at high bitrate. This is shown in Fig 6.3 that this case this method performs better in medium bitrates as compared to previous case.

CHAPTER 6

CONCLUSION AND FUTURE SCOPE

Channel diversity is used for reliable communication in wireless networks. MDC can be considered as a Joint Source channel coding, so this is best suited in wireless networks. Transmission over unreliable networks such networks suffers from various kinds of adverse conditions such as bandwidth fluctuation, burst -error contamination, packet loss, and excessive packet delay due to network congestion. By taking advantage of multiple logical channels provided by digital networks, a multiple description video coding scheme is used to meet quality of service requirements.

In this thesis we have studied about multiple description coding (MDC), which can be used in presence channel diversity wireless communication networks. The coding efficiency of the MDC is optimized by removing spatio-temporal redundancy from video signals and constructing a more compact hierarchical zero-tree for 3-D SPIHT algorithm. We have seen that Multiple Description (MD)-SPIHT performs better then Single Description (SD)-SPIHT

FUTURE SCOPE

The future scope may be in 4G wireless communication, where we want high data rate over error porne channel. Scalable coding is used for various data rates according to the channel conditions. Scalability can be used with MDC to offer both rate scalability and error resilience. This can be achieved either by creating Multiple Description coders in which each description is scalable or by designing layered coders in which each layer is MD with a different amount of redundancy.

REFERENCES

[1] J. G. Apostolopoulos, "Error Resilient Video Compression Through The Use of Multiple States," *IEEE Int'l Conf. on Image Processing*, vol. 3, Vancouver, BC, Canada, , pp. 352–355. September 2000

[2] A.A. El Gamal and T. M. Cover, "Achievable Rates For Multiple Descriptions," *IEEE Trans. Inform. Theory.*, vol. IT-28, pp. 851-857, Nov. 1982.

[3] V. A. Vaishampayan, "Design Of Multiple Description Scalar Quantizers," *IEEE Trans. Inf. Theory*, vol. 39, no. 3, pp. 821–834, May 1993.

[4]V. A. Vaishampayan, N. J. A. Sloane, and S. D. Servetto, "Multiple-Description Vector Quantization With Lattice Codebok: Design And Analysis," *IEEE Trans. Inf. Theory*, vol. 47, no. 7, pp. 1718–1734, Jul. 2001.

[5] Y. Wang, M. T. Orchard, V. A. Vaishampayan, and A. R. Reibman, "Multiple Description Coding Using Pairwise Correlating Transforms," *IEEE Trans. Image Process.*, vol. 10, no. 3, pp. 351–366, Mar. 2001.

[6] S. Shirani, M. Gallant, and F. Kossentini, "Multiple Description Video Coding Using Pre-And Post-Processing," *Int. Conf. IT: Coding & Comput.*, Las Vegas, NV, , pp. 35–39, Apr. 2001

[7] D.Wang, N. Canagarajah, D. Redmill, and D. Bull, "Multiple Description Video Coding Based On Zero Padding," in *Proc. ISCAS*, vol. 2, pp. 205–208, May 2004

[8] V.K. Goyal,; J.Kovacevic,.; , "Generalized Multiple Description Coding With Correlating Transforms ," *IEEE Transactions on Information Theory,* , vol.47, no.6, pp.2199-2224, Sep 2001

[9] A.E. Mohr; E.A.Riskin, R.E.Ladner, "Generalized Multiple Description Coding Through Unequal Loss Protection," *International Conference on Image Processing, 1999.* vol.1, no., pp.411-415 vol.1, 1999

[10] A.R. Reibman, H. Jafarkhani, Yao Wang,M.T.Orchard, R. Puri, , "Multiple-Description Video Coding Using Motion-Compensated Temporal Prediction," *IEEE Transactions on Circuits and Systems for Video Technology,* , vol.12, no.3, pp.193-204, Mar 2002

[11] S.Mallat, "A Theory For Multiresolution Signal Decomposition: The Wavelet Representation," *IEEE Trans. Pattern Anal. Much. Intell.*, vol. 11, pp. 674-693, July 1989.

[12] J. M. Shapiro, "Embedded Image Coding Using Zerotrees Of Wavelets Coefficients," *IEEE Trans. Signal Processing,* vol. 41, pp. 3445-3462, Dec. 1993.

[13]J. Wolf, A.Wyner, and J. Ziv, "Source Coding For Multiple Descriptions," Bell Syst. Tech. J., vol. 59, no. 8, pp. 1417-1426, Oct. 1980.

[14] R. Ahlswede, The Rate-Distortion Region For Multiple Descriptions Without Excess Rate," *IEEE Trans. Inform. Theory*, vol. IT-31, pp. 721-726, Nov. 1985.

[15]. L. Ozarow, On A Source-Channel Coding Problem With Two Channels And Three Receivers, Bell Syst. Tech. J. **59**, no. 10, 1909–1921, 1980.

[16] Goyal, V.K.; , "Multiple Description Coding: Compression Meets The Network," *IEEE Signal Processing Magazine,* , vol.18, no.5, pp.74-93, Sep 2001

[17] Kim Beong-Jo; W.A.Pearlman, "An Embedded Wavelet Video Coder Using Three-Dimensional Set Partitioning In Hierarchical Trees (SPIHT)," *Data Compression Conference, 1997. DCC '97. Proceedings* , vol., no., pp.251-260, 25-27 Mar 1997

[18] S. Cho and W. A. Pearlman, "Error Resilient Video Coding With Improved 3-D SPIHT And Error Concealment," in *Proc. SPIE,* 2003, vol. 5022, pp. 125–136.

[19] M. Biswas, M.R. Frater, J.F Arnold, "Multiple Description Wavelet Video Coding Employing a New Tree Structure," *Circuits and Systems for Video Technology, IEEE Transactions on* , vol.18, no.10, pp.1361-1368, Oct. 2008

[20] A. Said and W. A. Pearlman, "A New Fast And Efficient Image Codec Based On Set Partitioning In Hierarchical Trees," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 6, no. 3, pp. 243–250, Jun. 1996.