Multiple Description Image and Video Coding for Noisy Channels

Manuela Pereira

Université de Nice Sophia Antipolis - UFR Sciences

École Doctorale STIC

Sciences et Technologies de l'Information et de la Communication

Thèse

présentée pour obtenir le titre de

Docteur en Sciences
de l'Universite de Nice Sophia-Antipolis

Spécialité : Automatique, traitement du signal et des images.

Présentée et soutenue par

Manuela Pereira

Multiple Description Image and Video Coding for Noisy Channels

Thèse dirigée par **M**arc ANTONINI et **M**ichel BARLAUD.

Soutenue publiquement à l'I3S le 18 de Juin de 2004 devant le jury composé par:

M:	Fernando	PEREIRA	IST	Président;
MM:	Christine	GUILLEMOT	IRISA	Rapporteur;
M:	Antonio	ORTEGA	USC	Rapporteur;
M:	Ferran	MARQUES	UPC	Rapporteur;
M:	Marc	ANTONINI	CNRS	Co-directeur;
M:	Michel	BARLAUD	I3S	Directeur.

To my parents

"O valor das coisas não está no tempo que elas duram, mas na intensidade com que acontecem. Por isso, existem momentos inesquecíveis, coisas inexplicáveis e pessoas incomparáveis."

Fernando Pessoa

Acknowledgments

First and foremost, I would like to dedicate this thesis to my parents.

Professors Marc ANTONINI and Michel BARLAUD, I want to thank their scientific and human qualities.

To all members of my jury: Christine GUILLEMOT, Ferran MARQUES, Antonio ORTEGA and Fernando PEREIRA, I want to thank their constructive feedback on my dissertation.

I want to thank all the friends that made these three years in France memorable. They made France my second home. To Mr. Sack, Fred Precioso, Adeline, Deni and Romain, Cristophe, Stéphanie, Manu and Tilio, Marc, Renata and Anais, Muriel, Lionel, Charles, Fred Payan, Valery, Joel, Annabelle, Marco, Guillaume, Carine, Frank, Pierre, Sara and Fabrice, Julien and Karen, Ludovic ...thanks.

For their help and availability I want to thank Micheline Hagnéré and Guy Teissier.

And finally I want to thank someone that believed on me more than myself: Simão; and to the reason of all this: Patrícia.

Contents

List of Figures

List of Tables

Notation and Symbols

$\#SB$	Subband' s number.
a_i	Quotient of the size of the subband divided by the size. of the whole image (e.g., $a_i = \frac{1}{2^{2i}}$ in the dyadic case).
C	Channel Capacity.
Δ_i	Optional weight for frequency selection.
D	Distortion.
D_0	Distortion associated with the central decoder.
$D_{n,n\in\mathbb{N}}$	Distortion associated with the n-esim channel.
$D(R)$	Distortion when a bit rate of R.
$D_{i,j}, (i,j \in \mathbb{N})$	Distortion associated with subband i, and descriptor j.
$H(x)$	Source entropy.
$H_y(x)$	Conditional entropy or equivocation.
$q_{i,j}, (i,j \in \mathbb{N})$	Quantization step associated with subband i, and descriptor j.
$\mathbb{N}$	Set of natural numbers.
r_N	Redundancy parameter.
$\mathbb{R}$	Set of real numbers.
R	Bit rate.
$R_{n,n\in\mathbb{N}}$	Bit rate associated with the n-th channel.
$R(D)$	Bit rate when a distortion of D.
$R_{i,j}, (i,j \in \mathbb{N})$	Bit rate associated with subband i, and descriptor j.
σ	standard deviation.
σ^2	variance.

Hamming weight (w_H) number of 1's in a word.

Minimum weight $w_{min} = \min\{w_H(c) | c \in \mathcal{C}, c \neq 0\}$

Hamming Distance The number of bits which differ between two binary strings. More formally, $d_H(a,b) = w_H(a \oplus b)$

Entropy $H(X) = \sum p(x) log_2 \frac{1}{p(x)}$.

Conditional entropy $H_y(X) = \sum_{x,y} p(x,y) log(p(x,y)/p(x)p(y))$.

Average Mutual Information $I(X;Y) = H(X) - H_y(x) = H(Y) - H_x(Y)$.

Channel Capacity $C = \max_{p(x)}(H(x) - H_y(x))$.

Single letter fidelity $d(x^n, \hat{x}^n) = \frac{1}{n}\sum_{i=1}^{n} d(x_i, \hat{x}_i)$.

Squared error distortion $d(x, \hat{x}) = (x - \hat{x})^2$.

SNR = $10log_{10}\frac{\sigma^2}{MSE}$.

PSNR = $10log_{10}\frac{2^{dyn}-1}{MSE}$, where dyn is the bits number each pixel is coded. By default we consider $dyn = 8$.

Abbreviations

3GGP	3^{rd} Generation Partnership Project
ARQ	Automatic Repeat reQuest
AWGN	Additive White Gaussian Noise
bpp	bits per pixel
bpss	bits per source symbol
BER	Bit Error Rate
BMDC	Balanced Multiple Description Coding
BSC	Binary Symmetric Channel
codec	Encoder and decoder
CCD	Charge Coupled Device
CDN	Content Delivery Network
CIF	Common Intermediate Format
DCT	Discrete Cosine Transform
DFT	Discrete Functional Theory
DWT	Discrete Wavelet Transform
DPCM	Differential Pulse Code Modulation
EBCOT	Embedded Block Coding with Optimized Truncation
EBWIC	Efficient Bit allocation Wavelet Image Coder
EC	Error Control
ETSI	European Telecommunication Standard Institute
EZW	Embedded Zerotree Wavelet
FEC	Forward Error Correction
FIR	Finite Impulse Response
GG	Generalized Gaussian
GOP	Group Of Pictures
GSM	Global System for Mobile communications
HPF	High Pass Filter
i.i.d.	Independent identically distributed
ITU	International Telecommunication Union
JPEG	Joint Photographic Experts Group
Kbps	Kilo bits per second
LEO	low-Earth-orbit
LOT	Lapped Orthogonal Transform

LPF	Low Pass Filter
MAI	Multiple Access Interference
Mbps	Mega bits per second
MD	Multiple Description
MD-FEC	Multiple Description through Forward Error Correction codes
MDBA	Multiple Description Bit Allocation
MDC	Multiple Description Coding
MDLVQ	Multiple Description Lattice Vector Quantizer
MDPT	Multiple Description via Polyphase Transform
MDSQ	Multiple Description Scalar Quantization
MDTC	Multiple Description Transform Coding
MPEG	Moving Picture Expert Group
MR	MultiResolution
MRC	Maximal Ratio Combining
MSE	Mean Square Error
nsec	nanoseconds
OFB	Oversampled Filter Bank
pdf	Probability Density Function
POCS	Projection Onto Convex Sets
PSNR	Peak Signal to Noise Ratio
QCIF	Quarter Common Intermediate Format (180×144)
QPSK	Quadrature Phase Shift Keying
ROI	Region Of Interest
SB	Subbands
SDC	Single Description Coding
SNR	Signal to Noise Ratio
SPIHT	Set Partitioning in Hierarchical Trees
UMDC	Unbalanced Multiple Description Coding
UMTS	Universal Mobile Telecommunications System
VLC	Variable-Length Coding
WCDMA	Wideband Code Division Multiple Access

Chapter 1

Introduction

1.1 Problem statement

Nowadays, the high performance of image/video coding techniques makes possible several new applications that were impossible before due to the high bit rates involved. Moreover, today's networks achieve high transmission speeds and support data rates sufficient for video applications even in mobile communications. These two facts together promise a whole new world of communications.

In fact, in the last decade the use of mobile communication and multimedia communication has seen an enormous increase, with the wireless channels considered as a transport medium for various types of multimedia information. Due to the high bit rates involved with multimedia, the scarcity of wireless bandwidth, the time-varying characteristics of the channel, and the power limitations of wireless devices, multimedia communications, specially the wireless ones, are a tremendous challenge.

The transmission delay is an important problem, specially for real time applications, since information that arrives too late at the decoder is considered as lost. High volumes of data greatly slows down transmission and, involves the use of lossy compression at low bit rates. For instance, video transmission is specially difficult due to the huge volume of data required to describe a video. Bandwidth limitations is another problem requiring compression for image and video transmission over wireless channels.

Despite the high performance of todays coding techniques we cannot ignore efficiency. Efficient codecs (encoder and decoder) are very important because the encoder/decoder operations cannot be overly complex, especially for real time applications.

The use of unreliable channels, such as wireless networks or the current Internet, implies that error-free delivery of data packets can only be achieved by allowing retransmission of lost or damaged packets, through error control mechanisms such as Automatic Repeat reQuest (ARQ) to ensure error-free delivery [76]. These techniques have been shown to be very effective and successfully when applied to wireless video transmission [96]. Retransmission of corrupted data frames, however, introduces additional delay, which might be unacceptable for real time transmissions that are delay sensitive. Such applications cannot easily make use of retransmission.

For applications involving transmission using unreliable channels the high

performance and efficiency of coder is not enough. Robust compression schemes are very useful especially for transmission at low bit rates. Therefore, it is important to devise encoding/decoding schemes that can make the compressed bitstream resilient to transmission errors. It is also necessary to design proper interfacing mechanisms between the codec and the network, so that the codec can adjust its operations based on the network conditions.

1.2 Error control: standard approaches

The goal of error control is to make the compressed bitstream resilient to transmission errors. The channel noise can occur in the form of random bit errors, burst bit errors or packet losses. Moreover, in the case of Internet or in the case of wireless communications, the network conditions are typically time varying. The codec must, in this case, be synchronized with the channel state. Video streams produced by standard codecs are specially sensitive to transmission errors. When using standard codecs, that use predictive coding and variable-length coding (VLC) in the source coder, a single erroneously recovered sample can lead to errors in the following samples in the same or following frames. Likewise, because of the use of VLCs, a single bit error can cause the decoder to lose synchronization, so that even successive correctly received bits become useless.

To make the compressed bitstream resilient to transmission errors one must add redundancy to the stream, so that it is possible to detect and correct errors. Typically, this is done at the channel by using Forward Error Correction codes (FEC) [166, 144, 24, 23, 159, 53, 160]. FEC involves the addition of redundant data to the compressed signal, which allows the decoder to correct errors up to a certain level. The classical Shannon information theory [156] states that one can separately design the source and channel coders, to achieve error-free delivery of a compressed bitstream, so long as the source is represented by a rate below the channel capacity. Therefore, the source coder should compress a source as much as possible (below the channel capacity) for a specified distortion, and then the channel coder can add redundancy through FEC to the compressed stream to enable the correction of transmission errors. Since, FEC implies the addition of redundant data it increases the total number of bits required reducing compression. The standard FEC presents high performances when developed for constant channels and channels that do not present burst errors. Nowadays different artifacts are added to FEC to be adapted to such kind of channels, as we present in the following.

FEC code must be designed with a worst case channel scenario in mind. For channels that have a highly variable quality, this worst case may imply the need for a very powerful coder, and hence a highly or even prohibitive amount of redundancy, which will severally reduce the compression performance. Furthermore, such a system will fail catastrophically whenever the FEC design limit is exceeded. To deal with this problem new systems use variable FEC codes instead of standard fix FEC codes. The first algorithm using variable FEC codes was proposed in [109]. This algorithm assigns unequal amount of

FEC depending on the importance of the data.

When dealing with channels with burst error characteristics the error correction capabilities of FEC are often exceeded or the block is error-free in which case additional redundancy is wasted. To overcome this limitation, FEC is often enhanced by a technique known as interleaving. For burst errors, this effectively reduces concentration of errors in single codewords, i.e., a burst of b consecutive symbol errors causes a maximum of b/M symbol errors in each codeword. Though interleaving can be implemented with low complexity it also suffers from increased delay, depending on the number of interleaved blocks M. Therefore interleaving is a frequently used technique for burst channels if additional delay is acceptable.

The typical error control techniques, ARQ, FEC or even both together [98, 76, 21], cannot be easily adapted to real time transmissions. Therefore, joint source and channel coding is often preferred. Such kind of scheme allocates a total amount of redundancy between source and channel coding. The error-resilient encoding methods working under this premise use this redundancy to recover from erroneous or missing bits. Thus, erroneous or missing bits will not have a disastrous effect in the reconstructed video quality.

For example, in error concealment, when an image or a block of samples are missing due to transmission errors, the decoder can estimate them based on surrounding received samples, by making use of inherent correlation among spatially and temporally adjacent samples [81, 86, 195]. Error concealment has, in contrast to error-resilient coding, the advantage of not employing (normally) any additional bit rate, but adds computational complexity at the decoder. This is a problem in real time communications.

Real time communications should combine optimization of source coding and channel coding, should present greater robustness and adaptability to adverse transmission condition and should make efficient use of limited network resources. The robustness and adaptability to adverse transmission is even more important in wireless communications since in this type of communications there are no guarantees that the packets arriving to the decoder are uncorrupted.

1.3 Multiple description approach

A particular joint source and channel coding method, known as Multiple Description Coding (MDC), has proven to be an effective way to provide error resilience with a relatively small reduction in compression ratio. The idea in MDC is to generate multiple independent descriptions of the source such that each description independently describes the source with a certain desired fidelity. When more than one description is available, they can be combined to enhance the quality. Note that while multiresolution (MR) approaches are sensitive to the position of losses in the bitstream, the multiple descriptions (MD) stream is insensitive to them and thus has the desired feature that delivered quality is only dependent on the fraction of descriptions delivered.

In the MDC problem (reduced to the simplest case of two descriptions), a source is described by two descriptions with side rates R_1 and R_2. These

two descriptions individually lead to reconstructions with side distortions D_1 and D_2, respectively; the two descriptions together yield a reconstruction with central distortion $D_0 \leqslant D_1$ and $D_0 \leqslant D_2$.

MDC is robust due to the redundancy of the MD of the same source and it may be scalable as each correctly received description improves the decoder performance. Also, MDC does not require prioritized transmission, as each description is independently decodable.

The MDC were created to solve a problem related with packet losses. Moreover, almost all multiple description codes to date assume the existence of multiple independent on-off channels between the transmitter and the receiver (e.g.: Internet). When a link is broken, all of the symbols or packets passing through that channel are lost and when it is functioning properly, the symbols are transmitted error free. Such standard MDC presents normally a large overhead implying that when channel loss rate is small, the reconstruction performance in the error free case dominates and a single description coding (SDC), without channel coding, perform best. Thus, a different MDC approach is needed when designing MDC schemes for new applications and todays communications. This explain the reduced amount of work dedicated to MDC for wireless channels that appears in the literature. So, as cited in [195] "A challenging task is how to design the MDC coder that can automatically adapt the amount of added redundancy according to underlying channel error characteristics". It is this challenge we will consider in the present work.

In this thesis we propose a joint source and channel coding method that presents robustness and adaptability to channel characteristics and state. Given its compression and synchronization capabilities it is suitable for real time transmission. In the MDC approach presented in this work the imposed redundancy between descriptions is highly dependent on channel model and state. The advantage of such MDC schemes is that they are well adapted for applications involving different channel models and time varying states, such as tele-conferencing, video telephony, virtual classrooms, video streaming, video downloading, etc. Nowadays MDC is not a method for Internet (or other on-off channels) communications but it becomes a method suitable for all kinds of communications involving unreliable channels. The present work will addresse some of the above applications.

1.4 Thesis overview

We start with a Multiple Description history in **chapter 2**. The state of the art of MDC for specific applications is presented in the chapters dedicated to these applications.

Our contributions begin in **chapter 3** where we present our MDC approach. Our goal is to find an optimal trade-off between efficient compression and robustness from losses due to communications using unreliable channels. The proposed method uses a MD scheme based on the Discrete Wavelet Transform (DWT) and an efficient bit allocation technique. The different MD are defined when

setting the bit allocation of each subband. We name it Multiple Description Bit Allocation (MDBA). The bit allocation for the successive sets of coefficients can be performed with respect to either rate or quality constraints. In both cases, the goal is to find a set of quantizers to apply in each subband, such that its performance lies on the convex hull of the global rate-distortion curve [161, 142, 54, 114]. To model the rate and distortion we use a non-asymptotic theoretical model for both rate and distortion [127]. The rate and distortion depends on the quantization step but also on the probability density function (PDF) of the wavelet coefficients. Assuming that the probability density model is accurate, this method provides optimal rate-distortion performances. Favoring the use of DWT is the fact that 2D DWT can be easily extended to 3D and thus applied to video coding. A 3D Scan-Based DWT video coder is presented in [120]. The use of a 3D Scan-Based DWT transform allows us to develop a stripe-based MDC and to use different redundancies to take into account changes in channel state while coding. This chapter ends with a comparison of the central PSNR versus side PSNR obtained with the proposed MDC and the most efficient MDC known to date. The work on this chapter resulted in several publications. In [127] we present the MDBA for image coding. In [128, 129] we present an extension for video coding. In [129] a low complexity scheme of the method in [128] is presented. We propose the automatic control of the amount of redundancy dispatched on the different descriptions by taking into account the channel model and state.

In **chapter 4** we explain how to use channel information to inject redundancy in the different descriptions. We take into account the Shannon theorem (Theorem 10) [156], and propose to define the redundancy using the equivocation $H_y(x)$. Indeed, in this theorem, Shannon states that the equivocation $H_y(x)$ is the amount of redundancy that the decoder needs to correct the received message. In this chapter we present some results for different kinds of channels for fixed or animated images. This work result in the following publications [131, 133, 134]. These papers present the adaptation of the proposed MDBA for wired [131] or wireless [133] communications. The papers [131] and [133] together yielded the journal publication [134].

In **chapter 5** we present the proposed MDC and a new one developed for coding quincunx images. We present an application of such MDC schemes for satellite transmission.

The quincunx arrangement is a way to improve image resolution by combining a pair of CCD linear arrays. Because each CCD array yields a classical image according to a square grid the systems using such acquisition model are tempted to treat each image isolated, disregarding the high redundancy between them.

We propose in this chapter a method of joint source-channel coding that takes into account the redundancy between the two images in source and/or channel coding. The proposed method uses the satellite channel characteristics when performing the source-channel coding. To compare with the proposed method we use the MDBA proposed in previous chapters. More specifically,

we process the quincunx sampled image with a well-suited transform to reduce the redundancies. After this step, we use the MDBA method proposed in chapters 3 and 4 and we adapt it for the case of satellite models. In the proposed method the redundancy between the two images is used to find a robust scheme. More precisely the different dyadic images are used to generate the two different descriptions in a MDC scheme and the difference between these two images is joined to both descriptions. This results in a highly robust scheme. As we expected, the standard method (using the MDBA) is better suited for lower levels of noise, while the proposed method perform best for higher levels of noise. We present this work in [135].

Finally, in **part III** we present general conclusions and perspectives.

We conclude that the proposed MDBA automatically adapted to the channel model and state, is efficient for transmission independently of the characteristics of the channel used. In this perspective, the proposed method is an alternative to methods that use error control schemes, such as FEC or ARQ, when limited delays are imposed or when time varying characteristics of the channels make difficult the use of such techniques.

MDC schemes designed to automatically adapt the amount of added redundancy according to underlying channel error characteristics, open new horizons for the MDC future. MDC can now be used for all kinds of applications demanding limit delays and error resilience. MDC becomes an option for new world communications.

Part I

State of the art

Chapter 2

Multiple description history

2.1 Problem statement

The MD problem was posed by Gersho, Witsenhausen, Wolf, Wyner, Ziv and Ozarow at the September 1979 IEEE Information Theory Workshop [69, 52] as a generalization of Shannon's problem of source coding with fidelity criterion [157]. More precisely, the posed problem was the follow. *"Suppose we wish to send a description of a stochastic process to a destination through a communication network. Assume that there is a chance that the description will be lost. Therefore we send two descriptions and hope that one of them will get through. Each description should be individually good. However, if both get through, then we wish the combined descriptive information to be as large as possible."*

This problem is summarized in figure 2.1. The difficulty in such a problem is that good individual descriptions must be close to the process, by virtue of their goodness, and necessarily must be highly dependent. Thus, after the reception of the first description, the second description will contribute little extra information. On the other hand, two independent descriptions must be far apart and thus cannot in general be individually good.

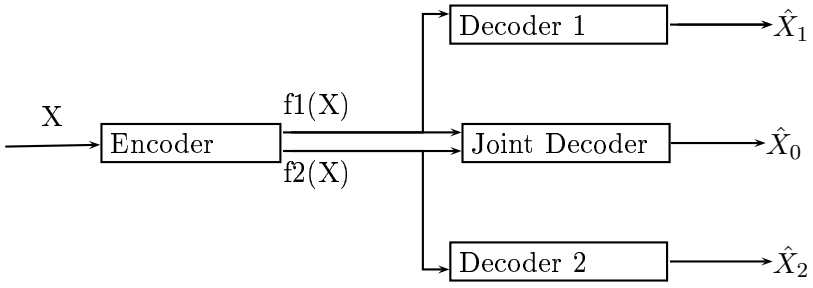

Figure 2.1: The channel splitting problem

2.2 Theoretical results

The first theoretical results appear in 1980 and try to characterize the set of achievable quintuples $(R_1, R_2, D_1, D_2, D_0)$. They were proposed by Witsen-

hausen [187], Ozarow [115], El Gamal and Cover [52] and Wolf, Wyner and Ziv [189]. We will briefly describe them.

In [187] Witsenhausen present a lower bound for side distortion when considered a memoryless binary symmetric source and used Hamming distance (i.e., probability of error) as distortion in Theorem 1.

Definition 1 *Suppose a block of $N = N_1 + N_2$ bits from a memoryless binary symmetric source is encoded into two signals U and V with respective alphabet sizes 2^{N_1} and 2^{N_2}. A receiver of the pair (U, V) is able to reconstruct the source block X_1^N without error. There are two other receivers, one receiving U only and producing a binary block $Y_1^N = F(U)$, and the other receiving V only and producing the binary block $Z_1^N = G(V)$.*

For each bit position $k(1 \leqslant k \leqslant N)$, the source bit X_k is compared with the decoded bits Y_k and Z_k. Define the error probabilities

$$p_u^k = Pr\{X_k \neq Y_k\}, p_v^k = Pr\{X_k \neq Z_k\}$$

Theorem 1 *Consider p_u^k and p_v^k as defined in Definition 1. For all k, the point (p_u^k, p_v^k) lies in the region of the (p_u, p_v) plane defined by $0 \leqslant p_u \leqslant 1$, $0 \leqslant p_v \leqslant 1$ and*

$$\left(p_u + \frac{1}{2}\right)\left(p_v + \frac{1}{2}\right) \geq \frac{1}{2}$$

In particular, if $p_u = p_v$, then $p_u \geq \frac{\sqrt{2}-1}{2}$

Also in [187] the following problem is posed: *" A memoryless source is encoded over n channels at rates $R_i (i = 1, ..., n)$. There are $2^n - 1$ decoders, one for each non void subset of channels. For a given distortion measure, the problem is to find the feasible combination of distortion rates."* The author concluded that for a certain value of $k, 0 < k < n$, if any k (or fewer) channel breaks down, $R = \frac{1}{n-k}$ is the rate required to obtain error-free operation.

In [189], Wolf, Wyner and Ziv, also considered the binary symmetric memoryless source and the Hamming distance as distortion. They proved that, if $(R_1, R_2, D_0, D_1, D_2)$ is achievable, then $R_1 + R_2 \geq 2 - h(D_0) - h(d_1 + 2d_2)$ and $R_1 + R_2 \geq 2 - h(D_0) - h(2d_1 + d_2)$, where

$$h(\lambda) = \begin{cases} 0, & \lambda = 0 \\ -\lambda log_2 \lambda - (1-\lambda) log_2 (1-\lambda), & 0 < \lambda \leqslant 1/2 \\ 1, & \lambda > 1/2. \end{cases} \tag{2.1}$$

From this theorem it follows that (with $R_1 = R_2 = 1/2, D_0 = 0), D_1 = D_2 \geq 1/6$.

From theorem 1, when $R_1 = R_2 = 1/2$, $D_0 = 0$ and $D_1 = D_2$, the rate distortion bound implies that $1 - h(D_1) \leqslant R_1 = 1/2$, or $D_1 \geq 0.11$. However, the theorem of Wolf, Wyner and Ziv, yields $h(3D_1) \geq 1$, or, $D_1 \geq 1/6$. The authors conclude that the bound under the Shannon assumptions is defined by the tangents to the hyperbola at the two points where it cuts the coordinate axis.

Work by El Gamal and Cover [52] has shown that, under the Shannon assumptions, all points above the hyperbola are achievable. The hyperbola is known as the achievable rate region of (R_1, R_2) pairs as a function of the distortion vector $D = (D_1, D_2, D_0)$ (see figure 2.2), for a memoryless source and a single-letter fidelity criterion as proved in Theorem 2.

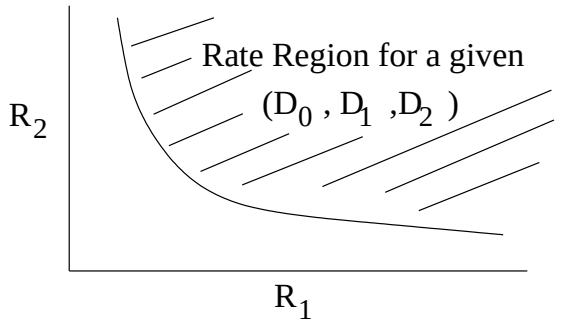

Figure 2.2: Achievable rate region of (R_1, R_2) pairs as a function of the distortion vector $D = (D_1, D_2, D_0)$.

The theorem 2 can be used to show that we can make $D_1 = D_2 = \frac{\sqrt{2}-1}{2} \approx 0.207$.

Theorem 2 *Let $X_1, X_2, ...$ be a sequence of i.i.d. finite alphabet random variables drawn according to a probability mass function $p(x)$. Let $d_i(.,.)$ be bounded. An achievable rate region for distortion $D = (D_1, D_2, D_0)$ is given by the convex hull of all (R_1, R_2) such that*

$$R_1 > I(X; \hat{X}_1),$$

$$R_2 > I(X; \hat{X}_2),$$

$$R_1 + R_2 > I(X; \hat{X}_1, \hat{X}_2, \hat{X}_0) + I(\hat{X}_1; \hat{X}_2),$$

for some probability mass function $p(\hat{x}, \hat{x}_0, \hat{x}_1, \hat{x}_2) = p(x)p(\hat{x}_0, \hat{x}_1, \hat{x}_2|x)$ such that

$$D_0 \geq Ed_0(X; \hat{X}_0),$$

$$D_1 \geq Ed_1(X; \hat{X}_1),$$

$$D_2 \geq Ed_2(X; \hat{X}_2).$$

Ozarow in [115] considered the case where the sources are Gaussian and the distortion is the squared-error criterion. The achievable rate region (the hyperbola) derived in [52], was proved to be, in fact, the rate distortion region for the source. This is done by obtaining the converse theorem. This theorem states that the achievable set of quintuples $(R_1, R_2, D_1, D_2, D_0)$ is given by the set of points satisfying $D_1 \geq e^{-2R_1}$, $D_2 \geq e^{-2R_2}$ and $D_0 \geq e^{-2(R_1+R_2)} \frac{1}{1-(\sqrt{(1-D_1)(1-D_2)}-\sqrt{D_1 D_2 - e^{-2(R_1+R_2)}})^2}$.

Witsenhausen and Wyner [188] have obtained an outer bound for the case of a binary symmetric source with the Hamming distortion and have compared it in one case to the achievable region of [52], but the bounds exceeds the achievable point.

Berger and Zhang in [14] defined $d = inf\{D : (\frac{1}{2}, \frac{1}{2}, D, D, 0)$ is achievable$\}$ for memoryless binary sources. They prove that $d = (\sqrt{2} - 1)/2$.

Ahlswede in [3] proved the tightness of the hyperbola bound in [52] on a case of no excess rate at D_0 ($R_1 + R_2 = R(D_0)$), for the binary symmetric memoryless source with an error frequency distortion criterion. Zhang and Berger in [199] disprove the conjecture that the achievable rate region given in [52] coincided with the rate distortion region (is tight) in case of binary symmetric source with Hamming distortion measure.

There have been no results to date for precisely determining the rate distortion region for non-Gaussian sources and for sources with memory. The latest results for non-Gaussian sources are from Zamir [196, 197]. Zamir develop an outer bound and an inner bound for the MD region for a general memoryless real source with squared-error distortion. These results are an extension of Shannon bounds on rate distortion function of a real source by the rate distortion function of the Gaussian source with the same variance / entropy.

Venkataramani, Kramer and Goyal have found bounds on the achievable performance region for MD coding with more than two descriptions [179].

An important special case of the MD problem was presented in [43] and is known as the problem of successive refinement of information or multiresolution (MR). The successive refinement problem is a special case of the MD problem in which there is no constraint on $Ed(X, \hat{X}_1)$ and in which $R_2 = R(D_2)$ and $R_1 + R_2 = R(D_0)$ is required. In this article, a necessary and sufficient condition is derived, such that, the rate distortion problem is successively refinable. The result follows from the tightness of the achievable region establish by El Gamal and Cover [52] for the no excess rate sum case [3].

At this moment there is a potential for applications of MD source codes in speech and video coding over packet-switched networks where packet losses can result in a degradation in signal quality. Another possible application is communication over fading multipath channels where diversity techniques are commonly used [139].

2.3 Practical coder designs for multiple descriptions

One of the first practical coder designs for multiple descriptions appears in the context of speech coding. In 1981 Jayant and Christensen, [80, 81], consider MD coding of DPCM speech for combating speech coding degradation due to packet losses. Information bits corresponding to even and odd samples are placed in separate packets. If only even (odd) sample packets are lost, data contained in the odd (even) packet is used to estimate the missing samples using the nearest neighbor interpolation.

Nowadays we can find different approaches of Multiple Description Coders. We will present each one of them in the order that they appear in the literature:

- Multiple Description Scalar Quantization.
- Multiple Description Transform Coding.
  - Square-Transform Based.
  - Frame Based.
- The new approaches of Multiple Description Coding.

There is also MD coding using forward error correcting (FEC) codes. Unlike others MDC approaches this one achieves MD property without modifying the source coding algorithm. Rather, correlation is reintroduced into the transmitted bitstream by applying different amounts of error protection to the sections of the bit stream produced by the source coder, and then combining these sections into equally important descriptions. Mohr at al. propose the use of error correcting codes of different strengths applied to different portions of a progressive bitstream such as that generated by SPIHT coder [109]. Also [140] considers the use of FEC in MDC. In these MDFEC systems, the reduction of distortion associated with any description actually depends on how many other descriptions are received. Thus we will not consider these methods in the MDC approaches we will present in the following.

2.4 Multiple description scalar quantization

The first approach, named Multiple Description Scalar Quantization (MDSQ), was pioneered by Vaishampayan in [173]. The MDSQ in [173] proposes that the rate of the descriptions can be traded of against the side distortions. This quantizer is obtained by a standard scalar quantizer followed by an index assignment that splits the signal into two descriptions. In this way, it sends information from each sample over both channels of the diversity system. Below we explain the MDSQ as presented in [173]. This design problem is posed as an optimization problem and necessary conditions for optimality are derived in [173]. Also, a design algorithm for quantizer design, is developed. Unlike a single channel scalar quantizer, the performance of a MD scalar quantizer is dependent on the index assignment. The author addresses the problem of index assignment and describes two families of index assignment matrices in which the maximal distortion between two indices sharing a description is minimized. Performance results and sample quantizer designs are presented for a memoryless Gaussian source.

MDSQ An (M_1, M_2)-level Multiple Description Scalar Quantizer maps the source sample x to the reconstruction levels $\hat{x}^0, \hat{x}^1, \hat{x}^2$ that take values in the codebooks, $\hat{\chi}^0 = \{\hat{x}^0_{i,j}, (i,j) \in \mathcal{C}\}, \hat{\chi}^1 = \{\hat{x}^1_i, i \in \mathcal{I}_1\}$ and $\hat{\chi}^2 = \{\hat{x}^2_j, i \in \mathcal{I}_2\}$, respectively, where $\mathcal{I}_1 = \{1, 2, ..., M_1\}, \mathcal{I}_2 = \{1, 2, ..., M_2\}$ and $\mathcal{C}$ is a subset of $\mathcal{I}_1 \times \mathcal{I}_2$. An MDSQ can be broken into two side encoders, $f_1 : \mathbb{R} \to \mathcal{I}_1$ and $f_2 : \mathbb{R} \to \mathcal{I}_2$ which select the indexes i and j, respectively, and three decoders, $g_0 : \mathcal{C} \to \mathbb{R}$ (central decoder), $g_1 : \mathcal{I}_1 \to \mathbb{R}$ and $g_2 : \mathcal{I}_2 \to \mathbb{R}$ (side decoders), whose outputs are the reconstruction levels

with indexes ij, i and j from the codebooks $\hat{\chi}^0, \hat{\chi}^1$, and $\hat{\chi}^2$, respectively. The rate of the encoder f_m is given by $R_m = log_2 M_m \, bpss, m = 1, 2$. The two encoders impose a partition $\mathcal{A} = \{A_{i,j}, (i,j) \in \mathcal{C}\}$ on $\mathbb{R}$, where $A_{i,j} = \{x : f_1(x) = i, f_2(x) = j\}$. The MDSQ is completely described by $\mathcal{A}, \hat{\chi}^0, \hat{\chi}^1$, and $\hat{\chi}^0$. The encoder is refereed as $\mho = (f_1, f_2)$, the decoder as $\eth = (g_0, g_1, f_2)$, $\mathcal{A}$ as the central partition, and the elements of $\mathcal{A}$ as the central cells. The determination of the central partition is crucial. Several methods by which this can be done are presented in [45].

If both indexes are received, the central decoder g_0 is used to reconstruct the source sample. On the other hand, if only $i(j)$ is received, then side decoder $g_1(g_2)$ is used to reconstruct the sample.

The central and side MSE's that can be achieved are determined by the index assignment. Figure 2.3 presents a simple example to illustrate the index assignment. Tables 2.1 and 2.2 present the reconstruction codebook and the central and side MSE for the central partition and index assignments of figure 2.3.

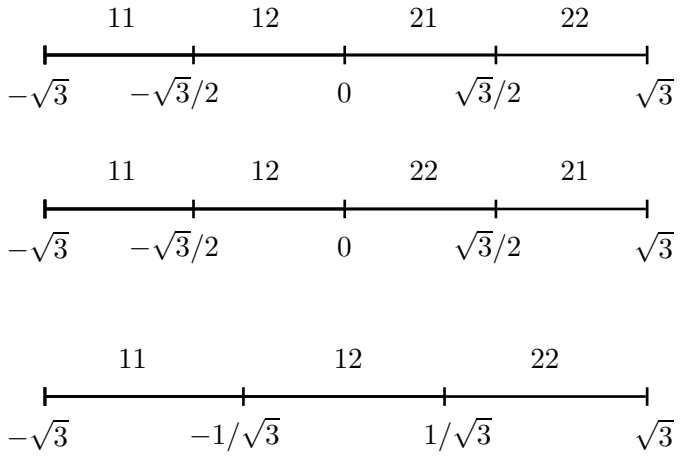

Figure 2.3: Three central partitions and index assignments for a uniformly distributed source

	$\hat{x}^0_{11}$	$\hat{x}^0_{12}$	$\hat{x}^0_{21}$	$\hat{x}^0_{22}$	$\hat{x}^1_1$	$\hat{x}^1_2$	$\hat{x}^2_1$	$\hat{x}^2_2$
(a)	$-3\sqrt{3}/4$	$-\sqrt{3}/4$	$\sqrt{3}/4$	$3\sqrt{3}/4$	$-\sqrt{3}/2$	$\sqrt{3}/2$	$-\sqrt{3}/4$	$\sqrt{3}/4$
(b)	$-3\sqrt{3}/4$	$-\sqrt{3}/4$	$\sqrt{3}/4$	$3\sqrt{3}/4$	$-\sqrt{3}/2$	$\sqrt{3}/2$	0.0	0.0
(c)	$-2/\sqrt{3}$	0.0	–	$2/\sqrt{3}$	$-1/\sqrt{3}$	$2/\sqrt{3}$	$-2/\sqrt{3}$	$1/\sqrt{3}$

Table 2.1: Reconstruction codebooks for figure 2.3

	$E(d_0)$	$E(d_1)$	$E(d_2)$
(a)	1/16	1/4	13/16
(b)	1/16	1/4	1
(c)	1/9	1/3	1/3

Table 2.2: Central and side MSE's for the central partition of figure 2.3 and the codebooks of table 2.1

Assuming that X is uniformly distributed over the interval $(-\sqrt{3}, \sqrt{3})$ and that $R_1 = R_2 = 1$ bps. Consider the MDSQ designs illustrated in figure 2.3 and Tables 2.1 and 2.2. In cases (a) and (b), each of the four codewords $(1,1),(1,2),(2,1),(2,2)$ is transmitted. Note from Table 2.2 that case (a) is clearly superior to case (b), but that in both cases the descriptions are unbalanced and one of them is poor, i.e., it has an MSE close to 1. In the third case (c), the codeword $(2,1)$ is not transmitted. Here, the descriptions are balanced. The distortion achieved by the joint description is larger than in case (a) and (b), however, both descriptions individually achieve a small MSE.

Vaishampayan and Domaszewicz in [175] extended the work in [173] to entropy constrained quantizers. They also used variable length codes (VLCs) instead of fixed length codes. With VLCs better performances are achieved, however, VLCs are very sensitive to errors (due to synchronization problems). In [70] the authors analyse the dependencies between the variables involved in the MDSQ coding chain and design an estimation strategy making use of part of the global model of dependencies each time. By analyzing the MDC system they evidence the most appropriate form of redundancy one should introduce in the context of VLC compressed streams in order to fight against de-synchronization when impaired by channel noise.

In [174] Vaishampayan and Batllo present an asymptotic analysis of MDSQ presented in [173]. Specifically, expressions are derived for the average side and central distortions and for entropy when the number of quantization levels is large. In this work they compare the distortion product D_0D_1 of the optimum level-constrained quantizer for a unit-variance Gaussian source with the one on the converse theorem. From the converse theorem, it can be shown that the multiple description rate distortion bound at large rates is given approximately by $D_0D_1 \approx \frac{1}{4}2^{-4R}$. The performance of the optimum level-constrained quantizer is given by $D_0D_1 \approx \frac{3\pi^2}{16}2{-}4R$ and of the optimum entropy-constrained quantizer by $D_0D_1 \approx \frac{\pi^2e^2}{144}2{-}4R$. These are important results because they show that for MDSQ both, the side and the central, distortion attain the optimal exponential rate of decay ($D_0 \approx 2^{-2R}, D_1 \approx 2^{-2R}$). The only sub-optimality of MDSQ at high rates is due to the use of a scalar quantizer which partitions the space into cubic regions instead of an ideal vector quantizer that would optimally partition

the space into spheres.

In [71] the authors consider the usage of Multiple Description Uniform Scalar Quantization (MDUSQ) for robust and progressive transmission of images over unreliable channels. They develop an index assignment which allows to improve the rate-distortion performance against previous proposed index assignments in the context of progressive and embedded bit streams. Thus, the MDUSQ proposed is well adapted for non stationary (varying bandwidth) communication environments.

Berger-Wolf and Reingold in [15] found an index assignment and a performance bound for MD scalar quantization for more than two descriptions. The index problem is formulated as a combinatorial optimization problem of arranging numbers in a matrix to minimize the maximum difference between the largest and the smallest number in any row or column. In the case of two descriptions transmitted at equal rates, the bounds (lower and upper bound) coincide, thus giving an optimal algorithm for the index assignment problem. In the case of three or more equal channels, the bounds are within a multiplicative constant.

Jafarkhani and Tarokh in [79] constructed MD trellis coded quantizers.

Vaishampayan in [172] describes an iterative algorithm similar to the generalized Lloyd algorithm that minimizes the Lagrangian of the rates and expected distortions R_1, R_2, D_1, D_2, D_0 and applied it to the optimization of multiple description vector quantizers. Non-balanced MD vector quantization was studied by Fleming and Effros [48] including more than two descriptions. This paper presents a new practical algorithm, based on a ternary tree structure, for the design of both fixed and variable rate multiple description vector quantizers for an arbitrary number of channels.

Some works proposing the design of MD lattice vector quantizers (MDLVQ) are [155, 178, 38] (we present a brief description of the MDLVQ below). The work in [38] has the particularity of considering asymmetric MD contrary to the former where the MD considered are always symmetric. In [61, 62] a method is introduced for a two channel MD coding that generalizes the MDLVQ developed in [155]. This last one uses a fine lattice Λ and a coarse sublattice Λ'. The former uses the index assignment of [155] and a coarse lattice Λ. With the slight increase in complexity, the convex hull of the operating points is improved. This extension of the algorithm in [155] provides a technique for more than two descriptions.

MDLVQ A Multiple Description Lattice Vector Quantizer is a triplet $\mathcal{Q} = (\Lambda, \Lambda', l)$ where:

- Λ is a lattice.
- Λ' is a sublattice that is geometrically similar to Λ (Two lattices Λ and Λ' are said to be *similar* iff there is a $c \in \mathbb{R}, c > 0$ and an orthogonal matrix $A \in \mathbb{R}^{n \times n}$, such that $\Lambda' = c\Lambda A$, i.e., if Λ and Λ' differ only by a rotation and a change of scale [35]).
- Each lattice point $\lambda \in \Lambda$ gets mapped by l to a pair of sublattice points $(\lambda'_{red}, \lambda'_{green})$ that uniquely identifies λ; i.e., l must be an

injection:

$$\Lambda \overset{1-1}{\longleftrightarrow} l(\Lambda) \subset \Lambda' \times \Lambda'.$$

l is referred to as the *index assignment*, and the pair of points in the image by l of a point λ are referred to as *red* and *green* descriptions.

The amount of redundancy in a lattice quantizer is controlled by $N = |\Lambda/\Lambda'|$, the index of Λ' in Λ. Given any sublattice point $\lambda' \in \Lambda'$, we require that l is such that the total number of distinct lattice points $\lambda \in \Lambda$ for which Λ' is used to describe λ is exactly N; i.e.,

$$|\{\lambda : \pi_{red}(l(\lambda)) = \lambda'\}| = |\{\lambda : \pi_{green}(l(\lambda)) = \lambda'\}| = N,$$

where, $\pi_{red}(\pi_{red}, \pi_{green}) = \pi_{red}$, and similarly for π_{green}. Lattice points are labeled with pairs of sublattice points (it is these sublattice points that actually get transmitted over each channel), and that each sublattice point is used exactly N times; the larger is N, the higher the uncertainty about the original lattice point when one of the channels fail.

A key property of good index assignments l is that the set of central cells that share a given label must be as localized in space as possible, in order to achieve low distortion in the case of channel failure. This is analogous to the idea that for a scalar quantizer the *spread* of a side cell must be minimized [173].

For real world sources such as speech and video, it is important to exploit the correlation in order to build efficient coders. MD quantizers can be used efficiently for sources with memory by using standard decorrelating transforms. Batllo and Vaishampayan name multiple description transform coding in [13] to an orthogonal transform followed by MDSQ to apply the quantizers to sources with memory. In [152, 153] Servetto, Ramchandran, Vaishampayan and Nahrstedt use the MDC in [13] to design a wavelet based image coder. Wavelet coding has been shown to achieve better compression than Discrete Cosine Transform (DCT) coding and moreover allows scalability. Some of the most successful wavelet coders [29, 83, 99, 150, 151, 154, 158, 190] derive their high coding performance from their ability to identify sets of coefficients with different statistics within image subbands, and then coding each of these sets with respect to an appropriate statistical model. Since these sets typically are image dependent, this information is not known a priori, and therefore must be somehow conveyed to the decoder. This can be done either explicitly, [83, 150, 154, 158, 190] or implicitly [29, 99, 151]. In the explicit case, side bits describing these sets are included in the bitstream; these are bits that do not convey information about the value of subband coefficients, but instead configure the decoder appropriately to decode such values. In the implicit case, the information regarding sets of coefficients is deduced only from data always available at the decoder, so that no explicit side bits are required.

As usually in the literature we use the name multiple description transform coding (MDTC) for the second and third approaches.

2.5 Multiple description transform coding

In the multiple description transform coding (MDTC) approach linear transforms are used to introduce a controlled amount of correlation among the transformed coefficients.

We can find two different MDTC methodologies. The first one, called square-transform based uses non-overlapping linear transforms while the second MDTC methodology, called frame based, uses overlapping transforms.

These two different methodologies are specified below.

MDTC Multiple Description Transform Coding of a source vector x is done in the following steps:

1. Use a decorrelate transform T_1 (e.g. KLT, DCT, ...);
2. Quantize the transformed coefficients;
3. Transform the quantized vector with an invertible, discrete transform $T_2 : \mathbb{C}^n \to \mathbb{C}^m$ ($m = n$ in the square-transform MDTC methodology and $m > n$ in the frame based MDTC methodology);
4. Entropy code the resultant components;
5. If the number of vector m is greater than the number of descriptions k, group them to be sent over the k channels.

When all components are received, the reconstruction process is to exactly invert the transform. The distortion is precisely the quantization error. If some components are lost, they are estimated from the received components using the statistical correlation introduced by the correlating transform. The estimation is then generated by inverting the transform as before.

It is very important to note that we can first quantize and then use a discrete transform. If a continuous transform is applied first and then quantized, the use of nonorthogonal transform lead to non cubic partition cells, which are inherently suboptimal among the class of partition cells obtainable with scalar quantization [54]. The configuration in [112] allows the use of discrete transforms derived from non-orthogonal linear transforms, and thus obtain better performance.

The MDTC systems focuses on the search for optimal redundancy rate-distortion points by designing the correlating transform T_2.

2.5.1 MDTC: square-transform based

The square-transform based MDTC was pioneered by Wang, Orchard, and Reibman in [183, 184]. MD quantizers are constructed by separately describing (i.e., quantizing and coding) the N coefficients of an $N \times N$ block linear transform, which has been designed to introduce a controlled amount of correlation between the transform coefficients. In this way, if one of the descriptions is lost, the other one can be statistically estimated using the introduced dependencies.

To circumvent the difficulty with designing a transform for $N \geq 2$ variable, the authors proposed to transform two variables at a time. They further considered how to form pairs given a set of N variables and how to allocate the redundancy among different pairs so that the reconstruction distortion is minimized for a given total redundancy rate. These articles also addresses several issues related to the optimality of the transforms used for encoding, and it is shown that nonorthogonal transforms perform better than orthogonal transforms in terms of redundancy rate-distortion gain.

In [112] Orchard et al. discuss MD coding of two dimensional Gaussian vectors using transform techniques. This work introduces a performance metric called *redundancy rate distortion function*, where the redundancy rate is defined as the number of extra bits required to match a given coding distortion, compared to a single description coding (SDC) system, and the distortion refers to the reconstruction distortion. More precisely, note that the performance of a MDC system can be measured with three parameters: the bit rate, the coding distortion, and the reconstruction distortion. The coding distortion refers to the error between the original signal and the decoded one from all descriptions, while the reconstruction distortion is defined as the error under a given channel loss profile. With conventional SDC, the goal is to maximize the coding efficiency which is equivalent to minimize the bit rate for a given coding distortion, or vice versa. With MDC, in order to reduce the reconstruction distortion, the coder must introduce a certain amount of correlation among separate descriptions, which will reduce the coding efficiency compared to that achievable by SDC.

In [63, 65, 64] the authors generalize the construction proposed in [112] by dealing with arbitrary N-dimensional vectors, and by expanding the set of transforms which are considered.

In [33] the authors developed a MDC encoder that generates multiple descriptions by splitting the coefficient blocks of a conventional LOT-based encoder. A maximally smooth image-recovery method is developed as part of the MDC decoder, which can recover the original signal from an incomplete set of coefficient blocks. The algorithm makes use of the constraints among the LOT coefficient blocks and the smoothness property of typical images and converts these constraints into an energy minimization problem, similar to the techniques previously developed for DCT based coders [185] and for the MDC coder using spatial subsampling [182].

In [82] the authors propose a two stage transform design technique for MDTC, i.e., structure design and magnitude design. The motivation is that protection properties of a MDTC system can be characterized by the output correlation matrix, i.e., which descriptions are correlated (structure) and to what extent they are correlated (magnitude). While the magnitude information can not, in general, be quantified for specific redundancy and distortion constraints, the structural information can be inferred from specific channel conditions or protection requirements. Consequently, the structure design will find admissible transforms (eigenmatrices of the output correlation matrix) using a Scaling-Rotation factorization and the magnitude design will search for the optimal transform from these admissible transforms. Such a design enables

the possibility of finding a structured transform solution using available channel information thus reducing both the design and implementation complexities.

2.5.2 MDTC: frame based

The frame based MDTC was pioneered by Goyal, Kovacevic, and Vetterli, MD quantizers are constructed by separately describing the N coefficients of an overcomplete $N \times K$ tight frame expansion [66]. Here, a linear transform from $\mathbb{R}^k$ to $\mathbb{R}^n$, followed by scalar quantization, is used to generate n descriptions of a k-dimensional source. The n descriptions are such that a good reconstruction can be computed from any k descriptions, but also descriptions beyond the kth are useful and reconstructions from less than k descriptions are easy to compute. In [65] preliminary image communication experiments are presented using the methods [63, 66]. In [27] a POCS-based (Projection Onto Convex Sets) algorithm for consistent reconstruction from MD of overcomplete expansions is developed. The POCS algorithm produces consistent reconstructions. Consistent reconstructions have smaller expected squared error distortion than inconsistent reconstructions [68]. The authors construct the frame from two complete transform bases. In this way, all projections can be expressed in terms of forward or inverse transforms. Since such transforms are usually efficient to compute, they can perform the reconstruction much faster than with previous methods.

Some authors dedicated their work to the construction and analysis of filter banks for MDTC or more generally for image coding and transmission over erasure channels. For example, in [12] a windowed Fourier method is used for a MDC based on overcomplete expansions, in [194] and [40] they designed biorthogonal filter banks for MD coding of Gaussian sources, with the difference that in [194] they use the correlating transform before quantization, and in [38] the quantization step is performed before the transform and approximated the continuous transform with a discrete one. In [137] a study of the performance of systems that use unitary filter banks [171] for the introduction of correlation is presented. In [101] oversampled block transform like the Discrete Fourier Transform (DFT) codes have been considered for MDC. It is shown in [143] that DFT codes are actually a special case of frames. Filter bank frame expansions have also been studied to achieve resilience to erasure [90, 39, 67, 102]. In [110] two channel oversampled filter banks (OFBs) and tree-structured oversampled filter banks which implement frame decomposition are considered . Tree-structured OFBs provide a natural framework for unequal loss protection. As seen above, frame expansions introduce redundancy in signal representation.

Several of the approaches mentioned above involve the design of specific transforms or quantizers that have to be matched to the desired level of protection. In these schemes, adapting to changing network conditions would entail having encoder and decoder both change the transform and/or quantizers they use. These approaches are thus limited in their ability to adapt to changing transmission conditions. The last approach of MDC tries to overcome this lim-

itation.

2.5.3 Multiple description coding using explicit redundancy

The new MDC approach, exploits the natural correlation between symbols for reconstruction. This approach is similar to the square-transform based MDTC approach above, except that the transform is not actively designed. An example is the Multiple Description via Polyphase Transform (MDPT) developed by Jiang and Ortega [82]. Their MDPT is an extension to SPIHT coder by separating Zerotrees into polyphase components. The SPIHT coder was first presented in [150]. This SPIHT algorithm uses the principles of partial ordering by magnitude, set partitioning by significance of magnitudes with respect to a sequence of octavely decreasing thresholds, ordered bit plane transmission and self-similarity across scale in an image wavelet transform. These are the principles of embedded zerotree wavelet (EZW) coding introduced by Shapiro in [158]. In [150] the realization of these principles in matched coding and decoding algorithms is presented. It is shown that the SPIHT algorithm is more effective than the previous implementation of the EZW coding. Rogers et al. in [147] propose to rearrange bits at the output of one configuration of the SPIHT coder, in such a way that the loss of one packet results in an error that does not propagate beyond the image region contained in that packet.

Miguel, Mohr and Riskin proposed a scheme using SPIHT in a generalized multiple description framework [106], called MD-SPIHT. This work is a generalization of the work in [82]. In [107] the authors extend the unequal loss protection framework of MD-SPIHT [106] by adding more redundancy to the ROI than to other parts of the image. In this way they present an efficient scheme for protecting a region of interest (ROI).

The approaches [82, 106] are related to earlier work on audio coding [73]. In these techniques, explicit redundancy is introduced, so that each sample in the input (for example each wavelet coefficient) is transmitted more than once and coded with different accuracy each time. This strategy has the drawback of leading to transmission of more samples than initially present in the source, and thus inefficiency in the case of error-free transmission.

In [148], Sagetong and Ortega demonstrate how these explicit redundancy techniques have the additional advantage of providing very simple mechanisms for adaptation to changing network conditions. The key observation is that the level of redundancy can be selected by determining the number of times a given sample (or wavelet coefficient) is transmitted, and how many bits should be used for each of the redundant representations. In this paper the authors show how a bit allocation problem can be defined, where the goal is to choose the best distribution of redundancy for a given packet loss rate. They provide techniques to solve this problem and show how indeed different loss rates require different levels of redundancy. Note that by using bit allocation to determine the level of redundancy, not only the encoder can adjust itself in a simple manner, but in addition the decoder can handle packets with different levels of redundancy without requiring any significant changes to its structure (e.g. the same transform, entropy coding, etc will be used). More specifically, the MDC technique

used in this paper generates the various descriptions through a polyphase transform. Consider for example the case of a scalar source. This polyphase-based MDC will divide this source into even and odd samples (or more sets if more than two descriptions are transmitted), and will compress each sample using two different quantization scales (coarse and fine). Then this approach will transmit groups of samples where a set of coarsely quantized odd samples is combined with a set of finely quantized even samples (and vice versa, i.e. fine odd with coarse even). The decoder operates by gathering the available information for each sample and then selecting for each polyphase component its highest quality copy to be used in the decoding; the remaining copies are discarded. In [149] the authors improve the system by using a *priority scaling factor* to introduce redundancy in each description.

The coders proposed above are designed for ideal MDC channel environments. In an ideal MDC channel environment, the channels are independent and data on each channel is either completely lost or received intact. In a packet network environment these ideal conditions may not hold true; packet losses can be correlated and only partial data (of either description) may be received at the decoder.

There has been limited work, on MD video coders for packet networks. Vaishampayan used MDC scalar quantizers [172] to develop robust image and video coders for packet loss environments. Recently, Reibman [145] has proposed a MD video coder for packet networks based on a rate allocation principle similar to the one presented in [34]. In [34] they proposed an unbalanced MDC (UMDC) system for transmission of video data over best effort packet networks. The system is unbalanced because the rate distribution among the various descriptions is not even; hence one description has high rate (high resolution/quality) and the other, low rate (low resolution). Most work in MDC has been on balanced systems, i.e., where each description is equally important, but for low packet loss rate conditions a UMDC system would be more useful. This is because the overhead in making descriptions balanced, which is particularly significant if the descriptions are to be coded in a standard syntax, would adversely affect the performance of balanced systems for low packet loss rates.

We finish this MD history remarking that the last MD coders are joint source channel coders that use redundancy adaptability to be adapted to changing network conditions. The existing methods using this approach are dedicated mostly to ideal MDC channel environment, and the last ones to packet lossy channels.

The MDC proposed in this manuscript belongs to the new MDC approach. As in the works of Jiang, Sagetong and Ortega, the proposed MDC adapts the explicit redundancy to changing networks transmission. Also, in the proposed MDC this is done by bit allocation allowing the automatic adjustment of the encoder and not needing any changes to the decoder.

The advantage of the proposed method over the former MDC of the new approach is that in the proposed MDC the explicit redundancy is dependent of the channel model and state. In this way, it is possible to develop a MDC

that performs well in different channels environments and that can be used even with time varying channels. Thus the proposed MDC is not only for ideal MDC channel environment neither only for packet lossy channels as the other methods into this approach. Our goal was to design a MDC method that can automatically adapt to any channel and state. Thus, the proposed MDC can be used for any application involving transmissions over unreliable channels. It is specially suited for real time application, where typical error control techniques, ARQ, FEC or even both together [98, 76, 21], cannot be easily adapted.

The proposed MDC is an option for the new world communications.

Part II

Multiple description coding using explicit redundancy

Chapter 3

Proposed multiple description bit allocation

This chapter presents the design of the proposed MDC coder. It bellows to the new approach of MDC. In this approach of MDC the redundancy is explicit. The novelty is that the proposed MDC can automatically adapt the amount of added redundancy dispatched on the different descriptions according to the error characteristics of the underlying channel . Thus, the proposed MDC is well suited for transmissions using unreliable channels.

We start this chapter presenting the proposed MDC in section 3.2. An important part of such MDC is the bit allocation procedure. It will be presented in section 3.3. In this section we verify that the proposed MDC is by nature unbalanced. However, we can make it balanced by introducing an additional constraint as will be presented in this section. The proposed algorithm for this bit allocation is presented in section 3.4. The bit allocation procedure is based on the non-asymptotic rate distortion models presented in section 3.5. We compare our application with the best Multiple Description Coding techniques reported to date in section 3.6. The proposed MDBA overcomes the referenced MDC.

3.1 Introduction

The proposed MDC is a joint source and channel coding method that presents error resilience using redundancy adaptability to be adapted to channel characteristics.

The proposed method uses a MD scheme based on the Discrete Wavelet Transform (DWT) and an efficient bit allocation technique. The different MD are defined when setting the bit allocation of each subband. We name it Multiple Description Bit Allocation (MDBA). The bit allocation for the successive sets of coefficients can be performed with respect to either rate or quality constraints. In both cases, the goal is to find a set of quantizers to apply in each subband whose performance lies on the convex hull of the global rate-distortion curve [161, 142, 54, 114]. To model the rate and distortion we use a non-asymptotic theoretical model for both rate and distortion [127]. The rate and distortion depends on the quantization step but also on the probability density function (pdf) of the wavelet coefficients. Assuming that the probability density model

is accurate, this method provides optimal rate-distortion performances.

3.2 Problem statement

The proposed MDC scheme focus on the special case in which there are two channels of equal capacity between a transmitter and a receiver. Thus, the proposed MDC scheme is a balanced MDC (BMDC). A BMDC framework generates descriptions of equal rate and importance. This property is well matched to communications systems with no priority mechanisms for data delivery, as for example Internet.

In such a scheme, a sequence of source symbols is given to an encoder to produce two independent bitstreams of equal rate and importance. These bitstreams are transmitted to three decoders over two noisy channels. One decoder (the central decoder) receives information sent over both channels while the remaining two decoders (the side decoders) receive information only from their respective channel (see figure 3.1).

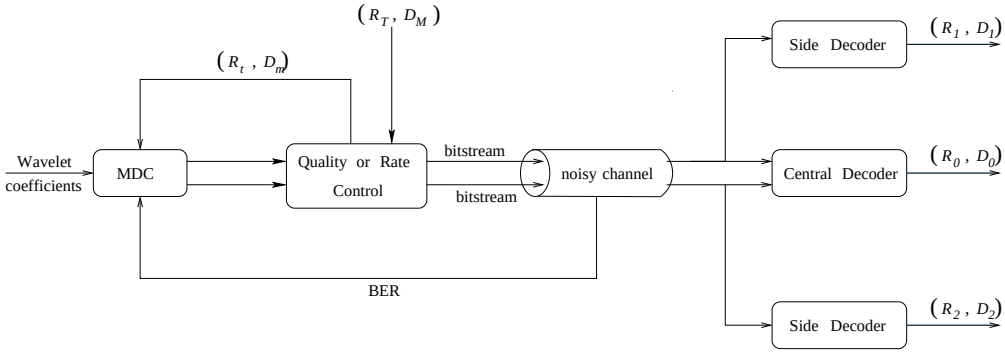

Figure 3.1: General Scheme

In the proposed method, as in the other MDC methods in the new approach, explicit redundancy is introduced, so that each sample in the input (each wavelet coefficient, in our case) is transmitted more than once and coded with different accuracy each time. In the proposed method the DWT is performed and then, the wavelet coefficients are repeated in both the descriptions. When a subband is finely coded in one description it will be coarsely coded in the other. In the following we will call primary subbands the finely coded subbands and redundant subbands the coarsely coded subbands.

We can see in figure 3.2 an example of division of the subbands in the descriptions.

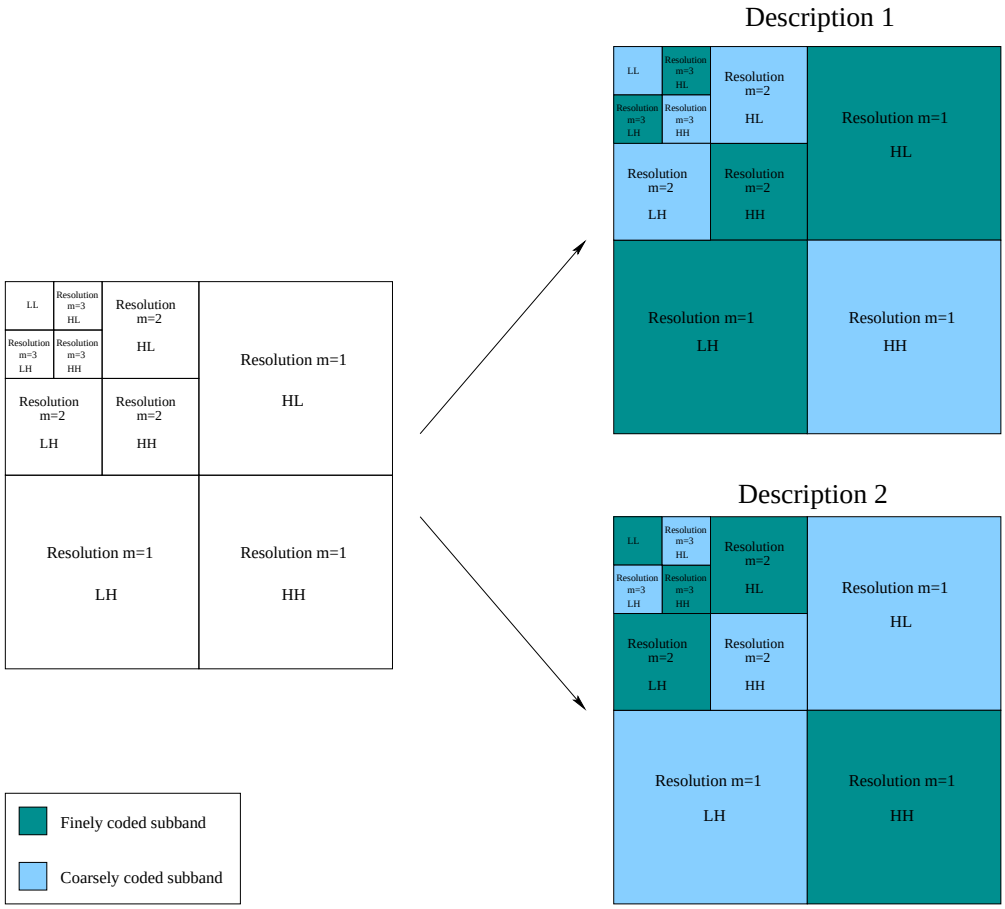

Figure 3.2: Example of division of the wavelet subbands between primary subbands (finely coded) and redundant subbands (coarsely coded) in the two description.

Since the division of the subbands into primary and redundant subbands will affect the methods performance it will be part of the problem. Other part of the problem is the decision of the amount of redundancy we pretend to dispatch on the different descriptions. This decision will be done by taking into account the channel model and state (or Bit Error Rate - BER). The general scheme we consider is presented in figure 3.1.

Given a total bit rate R_T and a maximal side distortion D_M, the generation of the two descriptions is constrained to three conditions that we detail in the following.

Condition 1 The central decoder has to reconstruct the original sequence from the two descriptions with minimal central distortion D_0;

Condition 2 A balanced MDC coder must generate two descriptions each with a side rate $R_1 = R_2 = R_T/2$;

Condition 3 When the channel is noiseless, the side decoders must reconstruct the original sequence from a single description with a side distortion $D_1 \leqslant$

D_M and $D_2 \leqslant D_M$.

The problem is then to minimize the central distortion D_0 (condition 1) when condition 2 and 3 are verified. That is, for N, the number of spatio-temporal subbands, we have to find the sets of bit rates $\{R_{i,1}\}, \{R_{i,2}\}$ that minimizes the central distortion D_0, where $R_{i,j}$ is the bit rate of subband $i \in \{1, .., N\}$ for description $j \in \{1, 2\}$. More precisely, we have to find the set of quantization steps $\{q_{i,1}\}, \{q_{i,2}\}$ that produces the sets of bit rates $\{R_{i,1}\}, \{R_{i,2}\}$. This problem is known as the *bit allocation problem* and we propose a solution in section 3.3.

3.3 Proposed bit allocation for MDC

3.3.1 Introduction of Lagrange operators

The problem is to find, for a given redundancy between the descriptions, the combination of scalar quantizers across the various wavelet coefficients subbands that will produce the minimum total central distortion while satisfying the side bit rates and side distortions constraints.

For a system considering N subbands of a wavelets decomposition, we intend to minimize the central distortion D_0 (condition 1) for a total bit rate R_T (condition 2). Thus, the purpose of our bit allocation for MD scheme is to determine the optimal sets of quantization steps $\{q_{i,1}\}, i = 1, ..., N$, $\{q_{i,2}\}, i = 1, ..., N$ for descriptions 1 and 2. This goal should be met while the side distortion is kept below a given distortion D_M (condition 3). The parameters R_T and D_M are given for the bit allocation (Fig. 3.1).

Some definitions and considerations are needed to formulate our bit allocation problem. They are presented in the following.

Definition 2 *Let define the functions* $f : \mathbb{R} \mapsto \mathbb{R}$ *and* $g : \mathbb{R} \mapsto \mathbb{R}$ *as*

$$f(x) = x - \frac{R_T}{2}. \tag{3.1}$$

$$g(x) = x - D_M. \tag{3.2}$$

The R_T parameter denotes the target output bit rate and D_M refers to the maximal side distortion imposed.

From Proposition 1 in [122] for a generalized Gaussian distribution with standard deviation σ and shape parameter α the bit rate R and quantization distortion D can be written depending only on α and the ratio $\frac{q}{\sigma}$ as presented in equations (3.3) and (3.4).

$$R = R\left(\alpha, \frac{q}{\sigma}\right) \tag{3.3}$$

$$D = \sigma^2 D\left(\alpha, \frac{q}{\sigma}\right). \tag{3.4}$$

In order to simplify the notation, we exclude from the formulas the shape parameter and we consider $\tilde{q}_i = \frac{q_i}{\sigma_i}$. Proposition 1 of [122] is described below, for proof or details about this proposition see [118].

Proposition 1 *(from [122]) When $p_{\alpha,\sigma}$ is a generalized Gaussian distribution with standard deviation σ and shape parameter α, there is a family of functions that verifies:*

$$\int_{q/2}^{q/2} x^n p_{\alpha,\sigma}(x)dx = \sigma^n f_{n,0}\left(\alpha, \frac{q}{\sigma}\right), \tag{3.5}$$

$$\int_{(m-1/2)q}^{(m+1/2)q} x^n p_{\alpha,\sigma}(x)dx = \sigma^n f_{n,m}\left(\alpha, \frac{q}{\sigma}\right) \ \forall m > 0 \tag{3.6}$$

with

$$f_{n,0}\left(\alpha, \frac{q}{\sigma}\right) = \int_{-(1/2)(q/\sigma)}^{+(1/2)(q/\sigma)} x^n p_{\alpha,1}(x)dx, \tag{3.7}$$

$$f_{n,m}\left(\alpha, \frac{q}{\sigma}\right) = \int_{m-(1/2)(q/\sigma)}^{+(m+1/2)(q/\sigma)} x^n p_{\alpha,1}(x)dx \tag{3.8}$$

Therefore, the bit rate R and the quantization distortion D depend only on the shape parameter α and the ratio $\frac{q}{\sigma}$.

$$R = R\left(\alpha, \frac{q}{\sigma}\right) \tag{3.9}$$

$$D = \sigma^2 D\left(\alpha, \frac{q}{\sigma}\right) \tag{3.10}$$

with,

$$R\left(\alpha, \frac{q}{\sigma}\right) = -f_{0,0}\left(\alpha, \frac{q}{\sigma}\right) log_2 f_{0,0}\left(\alpha, \frac{q}{\sigma}\right) - 2\sum_{m=1}^{+\infty} f_{0,m}\left(\alpha, \frac{q}{\sigma}\right) log_2 f_{0,m}\left(\alpha, \frac{q}{\sigma}\right), \tag{3.11}$$

$$D\left(\alpha, \frac{q}{\sigma}\right) = 1 - 2\sum_{m=1}^{+\infty} \frac{f_{1,m}\left(\alpha, \frac{q}{\sigma}\right)^2}{f_{0,m}\left(\alpha, \frac{q}{\sigma}\right)} \tag{3.12}$$

Using the definition 2 and proposition 1 our bit allocation problem can be resumed as:

$$(P)\begin{cases} \textbf{min } D_0\left(\{q_{i,1}, q_{i,2}\}\right) \\ \textbf{Constraints } f(R_1) \leqslant 0 \text{ and } f(R_2) \leqslant 0 \\ \textbf{Penalty } g(D_1) \leqslant 0 \text{ and } g(D_2) \leqslant 0 \end{cases} \tag{3.13}$$

where, $R_j, j = 1, 2$, is defined in equation (3.14) and $D_j, j = 1, 2$, is defined in equation (3.15).

$$R_j = \sum_{i=1}^{N} a_i R_{i,j}\left(\tilde{q}_{i,j}\right), \text{ for all } j \in \{1, 2\} \tag{3.14}$$

The parameter a_i in equation (3.14) is the size of the subband (i) divided by the size of the sequence, and $R_{i,j}(\tilde{q}_{i,j})$, is the output bit rate in bits per sample for the ith subband.

$$D_j = \sum_{i=1}^{N} \Delta_i w_i \sigma_{i,j}^2 D_{i,j}\left(\tilde{q}_{i,j}\right), \text{ for all } j \in \{1,2\} \tag{3.15}$$

In equation (3.15), Δ_i is an optional weight for frequency selection and w_i the weights used to take into account the nonorthogonality of filter bank [17].

The $\sigma_{i,j}^2 D_{i,j}(\tilde{q}_{i,j})$ for $j = 1, 2$ is the Mean Square Error (MSE) for the ith subband of description j, in the case of a Generalized Gaussian distribution.

The allocation problem (3.13) is a constrained problem which can be solved by introducing the Lagrange operators. The Lagrangian functional for the constrained optimization problem is given by equation (3.16).

$$J\left(\{q_{i,1}, q_{i,2}\}\right) = D_0 + \sum_{j=1}^{2} \lambda_j \, f(R_j) + \sum_{j=1}^{2} \mu_j \, g(D_j). \tag{3.16}$$

In equation (3.16) λ_j is the Lagrangian parameter and μ_j is the penalty parameter.

This Lagrangian functional 3.16, uses equations (3.1) and (3.2) defined in definition 2. For a source with generalized Gaussian distribution [6], the central distortion, D_0 in equation (3.16) can be written as presented in equation (3.17).

$$D_0 = \sum_{i=1}^{N} \Delta_i w_i \sigma_{i,0}^2 D_{i,0}\left(\tilde{q}_{i,1}, \tilde{q}_{i,2}\right). \tag{3.17}$$

The $\sigma_{i,0}^2 D_{i,0}(\tilde{q}_{i,1}, \tilde{q}_{i,2})$ is the central Mean Square Error (MSE) for the ith subband in the case of a Generalized Gaussian distribution. The expected central distortion is estimated based on the channels states and the a priori channels models as we will see in the next section.

3.3.2 Central distortion modeling

Recall that the central distortion is the distortion of the decoded image when using both descriptions at decoding. When the decoder receives both descriptions, each subband appears twice, with different bit rates (different associated quantization steps). In this case, if the subbands are noiseless, the central decoder for each couple of subbands chooses the one with the smaller quantization step and the other, the redundant subband, is only considered for side decoder. Hence, we can calculate the central distortion of the decoded image as

$$D_{i,0}(\tilde{q}_{i,1}, \tilde{q}_{i,2}) = \sum_{i=1}^{N} \min(\sigma_{i,1}^2 D_{i,1}(\tilde{q}_{i,1}), \sigma_{i,2}^2 D_{i,2}(\tilde{q}_{i,2})), \tag{3.18}$$

where $D_{i,1}, D_{i,2}$ are the distortions of subband i for descriptions 1 and 2, respectively. In the general case we have to take into account channel noise. So, we cannot ignore the redundant subbands as in the noiseless case. Actually, the

level of redundancy should increase when the BER increases, such as in the case of very noisy channels where the redundant subbands are as important as the others. In this last case the central distortion can simply be written as

$$\begin{aligned} D_{i,0}(\tilde{q}_{i,1},\tilde{q}_{i,2}) &= \sum_{i=1}^{N}\frac{1}{2}\left(\min(\sigma_{i,1}^2 D_{i,1}(\tilde{q}_{i,1}),\sigma_{i,2}^2 D_{i,2}(\tilde{q}_{i,1}))\right. \\ &\quad \left.+\max(\sigma_{i,1}^2 D_{i,1}(\tilde{q}_{i,1}),\sigma_{i,2}^2 D_{i,2}(\tilde{q}_{i,2}))\right) \\ &= \sum_{i=1}^{N}\frac{1}{2}\left(\sigma_{i,1}^2 D_{i,1}(\tilde{q}_{i,1})+\sigma_{i,2}^2 D_{i,2}(\tilde{q}_{i,2})\right). \end{aligned} \tag{3.19}$$

For the general case, we introduce a weighting parameter r_N to the redundant subbands, we call it *redundancy parameter*, and we propose to write the central distortion for a subband as presented in equation (3.20).

$$\begin{aligned} D_{i,0}\left(\tilde{q}_{i,1},\tilde{q}_{i,2}\right) = \frac{1}{\sigma_{i,0}^2}\frac{1}{r_N+1}\left[\min\left(\sigma_{i,1}^2 D_{i,1}\left(\tilde{q}_{i,1}\right),\sigma_{i,2}^2 D_{i,2}\left(\tilde{q}_{i,2}\right)\right)\right. \\ \left.+r_N\times\max\left(\sigma_{i,1}^2 D_{i,1}\left(\tilde{q}_{i,1}\right),\sigma_{i,2}^2 D_{i,2}\left(\tilde{q}_{i,2}\right)\right)\right] \end{aligned} \tag{3.20}$$

Equation (3.20) can be simplified as presented below.

$$\begin{cases} \frac{\sigma_{i,1}^2}{\sigma_{i,0}^2}\frac{1}{r_N+1}D_{i,1}(\tilde{q}_{i,1})+\frac{\sigma_{i,2}^2}{\sigma_{i,0}^2}\frac{r_N}{r_N+1}D_{i,2}(\tilde{q}_{i,2}), & \text{if } \sigma_{i,1}^2 D_{i,1}(\tilde{q}_{i,1})\leqslant\sigma_{i,2}^2 D_{i,2}(\tilde{q}_{i,2}) \\ \frac{\sigma_{i,2}^2}{\sigma_{i,0}^2}\frac{1}{r_N+1}D_{i,2}(\tilde{q}_{i,2})+\frac{\sigma_{i,1}^2}{\sigma_{i,0}^2}\frac{r_N}{r_N+1}D_{i,1}(\tilde{q}_{i,1}), & \text{otherwise.} \end{cases} \tag{3.21}$$

Taking in consideration what we said above, it is easy to conclude that the redundancy parameter domain is $[0,1]$. For $r_N=0$ equation (3.20) simplifies equation (3.18), and is used when the channel is noiseless. For $r_N=1$ equation (3.20) simplifies in equation (3.19), and is used when a very noisy channel is expected. The problem is how to choose intermediate redundancies, and implicitly intermediate values of r_N parameter. We want the amount of redundancy, i.e., the importance of the redundant subbands, to depend on the channel model and state (BER). In the next chapter we propose a method to compute the r_N parameter using channel characteristics.

3.3.3 Bit rate constraint

Finding the best bit allocation can be stated as a constrained optimization problem, where the R_i have to minimize the central distortion subject to a total bit rate constraint. The total bit rate constraint is posed in condition 2. This condition ($R_j \leqslant R_T/2$, $j=1,2$) has to be defined for each description. Using equation (3.14), we write condition 2 as a constraint F_j given in equation (3.22), for the different descriptions $j=1,2$.

$$F_j = \left(\sum_{i=1}^{N} a_i R_{i,j}\left(\tilde{q}_{i,j}\right) - R_T/2\right), \text{for all } j\in\{1,2\} \tag{3.22}$$

3.3.4 Side distortion penalty

We have stated earlier that the subbands are divided into primary and redundant subbands. If a subband is primary in one description it is redundant in the other and vice versa. The distinction between redundant and primary subbands is based on the redundancy parameter. The decision of which will be the redundant subbands in each description will be done by minimizing the central distortion. Thus, it is natural that in some cases the resultant descriptions are not of equal importance. For some applications this is not a drawback (as for example transmissions with low loss rates). When balanced descriptions are imposed by the application the MDC is attended, we have to join a penalty to our bit allocation problem. Is that the reason of condition 3.

Condition 3 ($D_j \leqslant D_M$, $j = 1, 2$) is forced using a penalty which has to be defined for each description. The side distortions D_1, D_2 are defined in equation (3.15) for a generalized Gaussian distribution.

The penalty method is simple and efficient. Consider a constraint $x > 0$. The penalty is written as

$$P(x) = \left(\frac{|x| - x}{2} \right)^2. \tag{3.23}$$

If the constraint is verified then $x \geqslant 0$ and $P(x) = 0$. Otherwise, $x < 0$ and $P(x) = x^2$. Considering the side distortions D_1, D_2 defined by (3.15), the constraint is

$$(D_j - D_M) \leqslant 0. \tag{3.24}$$

The penalty is then written as P_j in equation (3.25).

$$P_j = \left[\frac{|D_j - D_M| + (D_j - D_M)}{2} \right]^2, \text{ for all } j \in \{1, 2\} \tag{3.25}$$

The penalty function (3.25) allows us to find a solution with $D_j \leqslant D_M$, or more precisely, using equation (3.15),

$$\sum_{i=1}^{N} \Delta_i w_i \sigma_{i,j}^2 D_{i,j}\left(\tilde{q}_{i,j}\right) \leqslant D_M. \tag{3.26}$$

When equation (3.26) is verified we say that the penalty is verified.

In case of unbalanced MDC (UMDC) we can ignore this penalty. In such case the D_M should be fixed with a value higher or equal to σ^2 in such a way the penalty is always verified and the optimization problem is only performed in terms of central distortion and total bit rate. With UMDC we avoid the overhead in making descriptions balanced as presented in [34].

3.3.5 Solution of the problem

Considering the central distortion given by (3.20) the constraint (3.22) and the penalty (3.25), the Lagrangian functional (3.16) can be rewritten as in equation (3.27).

$$\begin{aligned} J\left(\{q_{i,1}, q_{i,2}\}\right) &= \sum_{i=1}^{N} \Delta_i w_i \sigma_{i,0}^2 D_{i,0}\left(\tilde{q}_{i,1}, \tilde{q}_{i,2}\right) \\ &+ \sum_{j=1}^{2} \lambda_j \left(\sum_{i=1}^{N} a_i R_{i,j}\left(\tilde{q}_{i,j}\right) - R_T/2 \right) \\ &+ \sum_{j=1}^{2} \mu_j \left[\frac{|D_j - D_M|}{2} + \frac{(D_j - D_M)}{2} \right]^2 \end{aligned} \tag{3.27}$$

The solution of (3.27) is obtained when

$$\begin{cases} \dfrac{\partial J\left(\{q_{i,1}, q_{i,2}\}\right)}{\partial q_{i,1}} = 0, & (a) \\ \dfrac{\partial J\left(\{q_{i,1}, q_{i,2}\}\right)}{\partial q_{i,2}} = 0, & (b) \\ \dfrac{\partial J\left(\{q_{i,1}, q_{i,2}\}\right)}{\partial \lambda_1} = 0, & (c) \\ \dfrac{\partial J\left(\{q_{i,1}, q_{i,2}\}\right)}{\partial \lambda_2} = 0. & (d) \end{cases} \tag{3.28}$$

The derivative of Lagrangian functional (3.27) with respect to λ_j, $j = 1, 2$ (equation (3.28 c) and equation (3.28 d)) is presented in equation (3.29).

$$\frac{\partial J\left(\{q_{i,1}, q_{i,2}\}\right)}{\partial \lambda_j} = 0 \Longleftrightarrow \sum_{i=1}^{N} a_i R_{i,j}\left(\tilde{q}_{i,j}\right) - R_T/2 = 0 \tag{3.29}$$

In the following the derivative of Lagrangian functional (3.27) with respect to $q_{i,1}$ (equation (3.28 a)) is given. The derivative of Lagrangian functional (3.27) with respect to $q_{i,2}$ (equation (3.28 (b)) is similar.

$$\begin{aligned} \frac{\partial J\left(\{q_{i,1}, q_{i,2}\}\right)}{\partial q_{i,1}} &= \Delta_i w_i \sigma_{i,0}^2 \frac{\partial}{\partial q_{i,1}} D_{i,0}\left(\tilde{q}_{i,1}, \tilde{q}_{i,2}\right) \\ &+ \lambda_1 a_i \frac{\partial}{\partial q_{i,1}} R_{i,1}\left(\tilde{q}_{i,1}\right) \\ &+ \mu_1 \frac{\partial}{\partial q_{i,1}} \left[\frac{|D_1 - D_M|}{2} + \frac{(D_1 - D_M)}{2} \right]^2 = 0 \end{aligned} \tag{3.30}$$

According to equation (3.21), $\frac{\partial}{\partial q_{i,1}} D_{i,0}\left(\tilde{q}_{i,1}, \tilde{q}_{i,2}\right)$ in equation (3.30) can be simplified as presented in equation (3.31).

$$\begin{cases} \dfrac{\sigma_{i,1}^2}{\sigma_{i,0}^2}\dfrac{1}{r_N+1}\dfrac{\partial}{\partial q_{i,1}}D_{i,1}(\tilde{q}_{i,1}), & \text{if } \sigma_{i,1}^2 D_{i,1}(\tilde{q}_{i,1}) \leqslant \sigma_{i,2}^2 D_{i,2})(\tilde{q}_{i,1}) \\ \dfrac{\sigma_{i,1}^2}{\sigma_{i,0}^2}\dfrac{r_N}{r_N+1}\dfrac{\partial}{\partial q_{i,1}}D_{i,1}(\tilde{q}_{i,1}), & \text{otherwise.} \end{cases} \tag{3.31}$$

In this way, the derivation of the Lagrangian functional (3.30) results in equation (3.32).

$$\begin{aligned} \frac{\partial J\left(\{q_{i,1}, q_{i,2}\}\right)}{\partial q_{i,1}} &= \frac{C_{i,1}}{1+r_N}\Delta_i w_i \sigma_{i,1}^2 \frac{\partial}{\partial q_{i,1}} D_{i,1}(\tilde{q}_{i,1}) \\ &+ \lambda_1 a_i \frac{\partial}{\partial q_{i,1}} R_{i,1}(\tilde{q}_{i,1}) \\ &+ \mu_1 \frac{\partial}{\partial q_{i,1}} \left[\frac{|D_1 - D_M|}{2} + \frac{(D_1 - D_M)}{2}\right]^2 = 0, \end{aligned} \tag{3.32}$$

The $C_{i,j}$ parameter in equation (3.32) comes from system (3.31) and is computed as presented in equation (3.33).

$$C_{i,j} = \begin{cases} 1, & \text{if } \min(\sigma_{i,1}^2 D_{i,1}(\tilde{q}_{i,1}), \sigma_{i,2}^2 D_{i,2}(\tilde{q}_{i,1})) = \sigma_{i,j}^2 D_{i,j}(\tilde{q}_{i,j}) \\ r_N, & \text{otherwise.} \end{cases} \tag{3.33}$$

Furthermore, using equation (3.15) $\frac{\partial}{\partial q_{i,1}} P_1$ can be written as in equation (3.34).

$$\begin{cases} 0, & \text{if } D_1 \leqslant D_M \\ 2 \times (D_1 - D_M)\, \Delta_i w_i \sigma_{i,1}^2 \dfrac{\partial}{\partial q_{i,1}} D_{i,1}(\tilde{q}_{i,1}), & \text{otherwise.} \end{cases} \tag{3.34}$$

Considering an E_1 parameter computed from system (3.34),

$$E_j = \begin{cases} 2 \times (D_j - D_M), & \text{if } D_j > D_M \\ 0 & \text{otherwise} \end{cases} \tag{3.35}$$

Then equation (3.32) becomes:

$$\begin{aligned} \frac{\partial J\left(\{q_{i,1}, q_{i,2}\}\right)}{\partial q_{i,1}} &= \frac{C_{i,1}}{1+r_N}\Delta_i w_i \sigma_{i,1}^2 \frac{\partial}{\partial q_{i,1}} D_{i,1}(\tilde{q}_{i,1}) \\ &+ \lambda_1 a_i \frac{\partial}{\partial q_{i,1}} R_{i,1}(\tilde{q}_{i,1}) \\ &+ \mu_1 E_1 \Delta_i w_i \sigma_{i,1}^2 \frac{\partial}{\partial q_{i,1}} D_{i,1}(\tilde{q}_{i,1}) = 0, \end{aligned} \tag{3.36}$$

Then, the derivative of equation (3.27) with respect to $q_{i,1}$ (equation (3.28 a)) yields:

$$\left(\frac{C_{i,1}}{1+r_N}+\mu_1 E_1\right)\Delta_i w_i \sigma_{i,1}^2 \frac{\partial}{\partial q_{i,1}} D_{i,1}(\tilde{q}_{i,1}) + \lambda_1 a_i \frac{\partial}{\partial q_{i,1}} R_{i,1}(\tilde{q}_{i,1}) = 0. \quad (3.37)$$

Equation (3.37) can be simplified to

$$\frac{\partial D_{i,1}}{\partial R_{i,1}}(\tilde{q}_{i,1}) = \frac{-\lambda_1 a_i}{\Delta_i w_i \sigma_{i,1}^2 \left(\frac{C_{i,1}}{1+r_N}+\mu_1 E_1\right)} \quad (3.38)$$

As we said above, the derivative of equation (3.27) with respect to $q_{i,2}$ (equation (3.28 b)) is similar. It yields:

$$\frac{\partial D_{i,2}}{\partial R_{i,2}}(\tilde{q}_{i,2}) = \frac{-\lambda_2 a_i}{\Delta_i w_i \sigma_{i,2}^2 \left(\frac{C_{i,2}}{1+r_N}+\mu_2 E_2\right)} \quad (3.39)$$

In resume, the solution of (3.27), is the system (3.40) that represents a two channel scheme.

Equation (3.40 a) comes from (3.38) that is the derivative of equation (3.27) with respect to $q_{i,1}$ (equation (3.28 a)).

Equation (3.40 b) comes from (3.39) that is the derivative of equation (3.27) with respect to $q_{i,2}$ (equation (3.28 b)).

Finally, equations (3.40 b) and (3.40 c) are derived from (3.29) that is the derivative of equation (3.27) with respect to λ_j for $j = 1$ (equation (3.28 c)) and $j = 2$ (3.28 d)).

$$\begin{cases} \dfrac{\partial D_{i,1}}{\partial R_{i,1}}(\tilde{q}_{i,1}) = \dfrac{-\lambda_1 a_i}{\Delta_i w_i \sigma_{i,1}^2 \left(\dfrac{C_{i,1}}{1+r_N}+\mu_1 E_1\right)} & (\mathbf{a}) \\ \dfrac{\partial D_{i,2}}{\partial R_{i,2}}(\tilde{q}_{i,2}) = \dfrac{-\lambda_2 a_i}{\Delta_i w_i \sigma_{i,2}^2 \left(\dfrac{C_{i,2}}{1+r_N}+\mu_2 E_2\right)} & (\mathbf{b}) \\ \sum_{i=1}^{N} a_i R_{i,1}(\tilde{q}_{i,1}) - R_T/2 = 0 & (\mathbf{c}) \\ \sum_{i=1}^{N} a_i R_{i,2}(\tilde{q}_{i,2}) - R_T/2 = 0 & (\mathbf{d}) \end{cases} \quad (3.40)$$

Resolution of the system (3.40) which has $2\times(N+1)$ equations and $2\times(N+1)$ unknowns gives us the optimal sets of quantization steps $\{q_{i,1}\}$, $\{q_{i,2}\}$, for a given r_N.

The proposed algorithm is based on modeling of R and D functions as we will show in section 3.4.

3.3.6 $C_{i,j}$ parameter

The $C_{i,j}$ parameter will define which subbands are primary or redundant subbands for a specific description. $C_{i,j} = 1$ defines that subband i is a primary subband in description j, thus it will be a redundant subband in the other description. $C_{i,j} = r_N$ defines that subband i is a redundant subband in description j, thus it will be a primary subband in the other description.

The division of the primary and redundant subbands between the descriptions can be done randomly. We present in the results a example where the division was the one presented in figure 3.2. However we intend to find for which division we have the minimum central distortion D_0. For that we have to compute the $C_{i,j}$ parameter as defined in equation (3.33).

In equation (3.33) the $C_{i,j}$ parameter depends on $\sigma_{i,1}^2 D_{i,1}$ and $\sigma_{i,2}^2 D_{i,2}$. One problem, is that the distortions involved are unknown before system (3.40) is solved due to their dependence on the quantization steps $\{q_{i,1}\}$, $\{q_{i,2}\}$. A solution is to perform the algorithm for all possible combinations of $C_{i,j}$ values, and then choose the combination that gives better results.

3.3.6.1 C parameter - optimal solution

- Compute the MDBA for all possible combinations;
- Choose the combination that results in the minimal central distortion.
  - $N = 10$: Nb $_{iterations} = 1024$;
  - $N = 40$: Nb $_{iterations} = 1099511627776$.

However this solution is not efficient (it is time consuming and depending of the number of subbands it can become prohibitive). Thus we propose an algorithm that, for the tested images gives similar results to the algorithm performing all possible combinations.

3.3.6.2 Proposed algorithm

The proposed algorithm initializes the $C_{i,j}$ parameters to 1, for $i = 1..N$ and $j = 1, 2$, and iteratively modifies their values according to the current $\sigma_{i,1}^2 D_{i,1}$ and $\sigma_{i,2}^2 D_{i,2}$.

More precisely, if we define $\mathcal{S}$ as the set of all possible subbands of description 1 and 2 we perform the following steps:

1. we initialize $C_{i,j} = 1$, for $i = 1..N$ and $j = 1, 2$;
2. we compute $q_{i,j}$, for $i = 1..N$ and $j = 1, 2$;
3. we compute $D_{i,j}(\tilde{q}_{i,j})$, for $i = 1..N$ and $j = 1, 2$;
4. we search the subband k in $\mathcal{S}$ with the highest distortion.
5. we set the correspondent $C_{k,j}$ value to r_N.

6. we redefine the set $\mathcal{S}$, excluding the subbands k of description $j = 1$ and $j = 2$.

7. if $\mathcal{S}$ is not empty we go to step 2.

It is easily verified that N iterations are always performed to compute all $C_{i,j}$ values (N is the total number of subbands of a wavelets decomposition). This algorithm is detailed in Fig. 3.3.

Begin

First iteration

Input

$C_{i,j}$	$i=1$	...	$i=k-1$	$i=k$	$i=k+1$	...	$i=N$
$j=1$	1	...	1	1	1	...	1
$j=2$	1	...	1	1	1	...	1

Search

k_1, l_1, such that $D_{k_1,l_1} \geq D_{i,j}, i \in \{1, ..., N\}, j = 1, 2$

Set

$C_{k_1,l_1} = r_N$

Second iteration

Input

$C_{i,j}$	$i=1$	...	$i=k_1-1$	$i=k_1$	$i=k_1+1$	...	$i=N$
$j=l_1$	1	...	1	r_N	1	...	1
$j=\{1,2\} \setminus l_1$	1	...	1	1	1	...	1

Search

k_2, l_2, such that $D_{k_2,l_2} \geq D_{i,j}, i \in \{1, ..., N\} \setminus \{k_1\}, j = 1, 2$

Set

$C_{k_2,l_2} = r_N$

Continue...

N iteration

Search

k_N, l_N, such that $D_{k_N,l_N} \geq D_{i,j}, i \in \{1, ..., N\} \setminus \{k_1, k_2, ..., k_{N-1}\}, j = 1, 2$

Set

$C_{k_N,l_N} = r_N$

End

Figure 3.3: Algorithm for the computation of $C_{i,j}$ parameter.

3.3.6.3 Comparison: Optimal solution vs. Proposed algorithm

- Optimal solution
  - $N = 10$: Nb $_{iterations} = 1024$;
    - $PSNR = 40.39$ dB, for Lena Image at 1 bpp;
  - $N = 40$: Nb $_{iterations} = 1099511627776$.
- Proposed algorithm
  - $N = 10$: Nb $_{iterations} = 10$;
    - $PSNR = 40.38$ dB, for Lena Image at 1 bpp;
  - $N = 40$: Nb $_{iterations} = 40$.

The proposed algorithm presents similar results to the optimal algorithm. Several other tests were performed with similar results. In the suite we use the proposed algorithm for the computation of $C_{i,j}$ parameter.

3.4 Proposed algorithm

The proposed algorithm is presented in figure 3.4. As can be seen in this figure we compute $R_{i,j}$ using the given parameters $C_{i,j}$, λ_j, μ_j, r_N, equation (3.40 a) and $\frac{\partial D}{\partial R}$ function. If the debit constraint (3.40 b) is not verified we recompute the $R_{i,j}$ using a new λ_j. If it is verified we compute: $q_{i,j}$ from $R_{i,j}$ and R function; $D_{i,j}(q_{i,j})$; D_j using $D_{i,j}(q_{i,j})$; and finally $C_{i,j}$ using $D_{i,j}(q_{i,j})$.

The above steps are performed N times. After N iterations all $C_{i,j}$ are computed and the algorithm can proceed.

The last step of bit allocation is the verification of the penalty (3.15). If this is not verified we restart the algorithm from the beginning with a new μ parameter computed using equation 3.41.

$$\mu_j^{new} = \mu_j^{old} + \eta_k d_j \tag{3.41}$$

where $d_j = D_j - D_M$. This is a sub-gradient method, for details see [108].

If the penalty is verified the algorithm stops and the output of bit allocation gives the optimal quantization steps $q_{i,j}$.

To compute $\frac{\partial D}{\partial R}$, R and D functions we use a model based implementation. This approach is very simple. We only need the statistic parameters for each subband. Contrary to other implementation, as for instance the signal based implementation used by JPEG2000 [167, 77], the model based implementation is adapted to the parallelization of the quantization and coding step of all subbands. The model based implementation is specified in section 3.5.

3.4.1 Bit allocation complexity

The complexity of our algorithm is the complexity of the EBWIC bit allocation presented in [127, 122]. The complexity of this coder is presented in [122] and detailed in [118]. In this work the authors conclude that the highest cost of

model based allocation method corresponds to the computation of generalized Gaussian distribution parameters for each subband. They need four operations (two additions and two multiplications) for wavelet coefficient to compute the σ and α parameters. Assuming that the complexity of the remaining part of the algorithm is lower than 1 operation for each image pixel. The authors conclude that the complexity of such allocation method is less than five simple arithmetic operations for each image pixel.

We can then conclude that the complexity of our bit allocation is less than five operations for each image pixel.

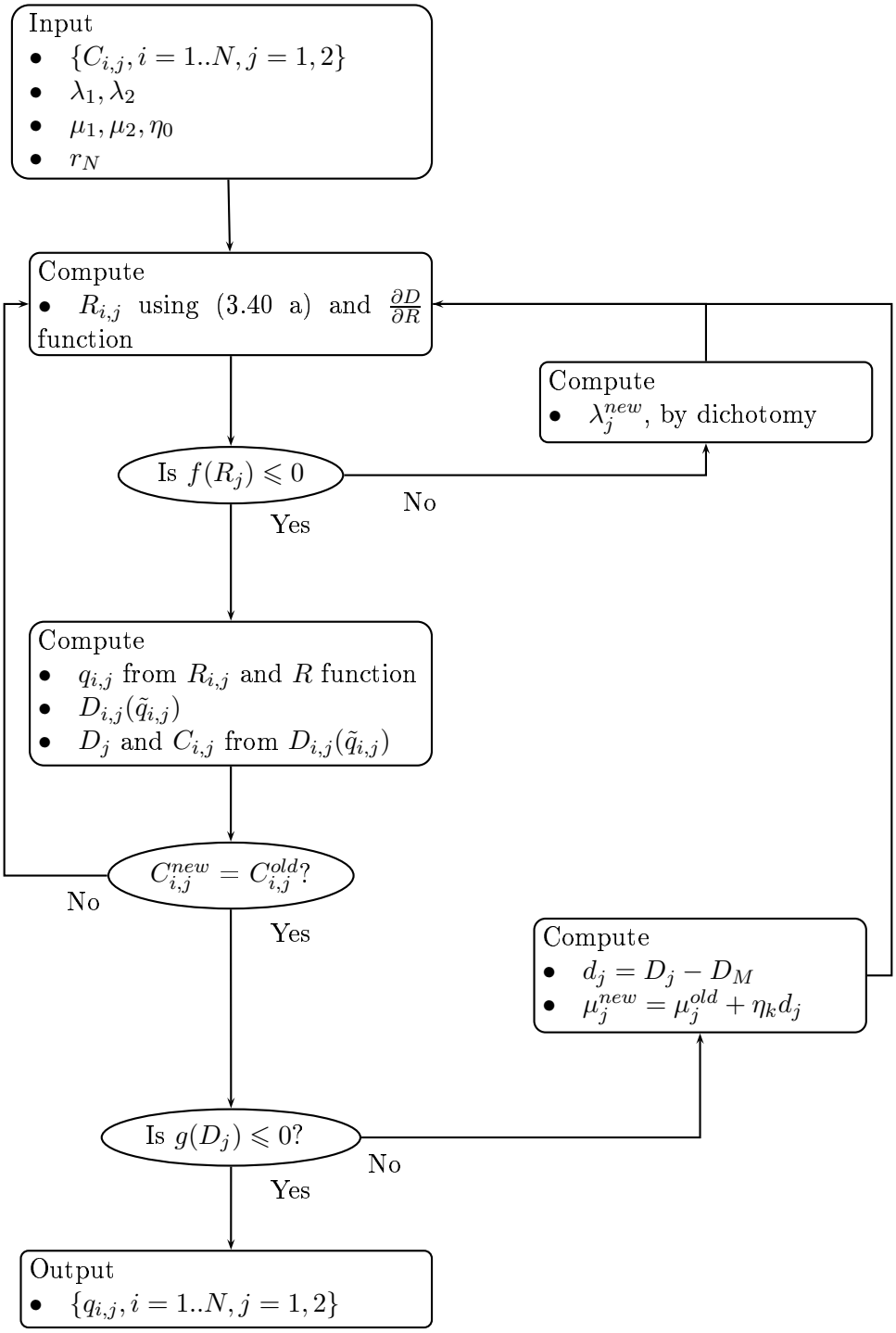

Figure 3.4: Global Bit Allocation Procedure

3.5 Model based implementation

To model the rate and distortion we use a non-asymptotic theoretical model [120]. The rate and distortion depends on the quantization step but also on the probability density function (pdf) of the wavelet coefficients. Assuming that the probability density model is accurate, this method provides optimal rate-distortion. In each subband the probability density function of the wavelet coefficients can be approximated with generalized Gaussian [6]. Therefore, for a shape α and a variance σ, the pdf is given by equation 3.42.

$$p_{\alpha,\sigma}(x) = ae^{-|bx|^{\alpha}} \tag{3.42}$$

with

$$b = \frac{1}{\sigma}\sqrt{\frac{\Gamma(3/\alpha)}{\Gamma(1/\alpha)}} \text{ and } a = \frac{b\alpha}{2\Gamma(1/\alpha)}.$$

For a given subband, the coder output bit rate R produced by the quantization step q, can be approximated by the entropy such that

$$R = -\sum_{m=-\infty}^{+\infty} Pr(m)log_2Pr(m), \tag{3.43}$$

where

$$Pr(m) = \int_{mq-\frac{q}{2}}^{mq+\frac{q}{2}} p_{\alpha,\sigma}(x)dx$$

is the probability of the quantization level m.

According to [164], the best decoding value, when using the Mean Square Error (MSE) as the distortion measure, for the quantization level m, is

$$\hat{x} = \frac{\int_{mq-\frac{q}{2}}^{mq+\frac{q}{2}} xp_{\alpha,\sigma}(x)dx}{Pr(m)}.$$

Then, the MSE can be expressed as:

$$D = \sum_{m=-\infty}^{+\infty} \int_{mq-\frac{q}{2}}^{mq+\frac{q}{2}} (x-\hat{x})^2 p_{\alpha,\sigma}(x)dx \tag{3.44}$$

Setting $\tilde{q} = \frac{q_{i,j}}{\sigma_{i,j}}$, it has been shown in [131, 122] that $\frac{\partial D_{i,j}}{\partial R_{i,j}}(\tilde{q})$ can be calculated as:

$$\frac{\partial D_{i,j}}{\partial R_{i,j}}(\tilde{q}) = \frac{\sum_{m=1}^{+\infty} \frac{2\frac{\partial f_{1,m}}{\partial \tilde{q}}(\alpha,\tilde{q})f_{1,m}(\alpha,\tilde{q})f_{0,m}(\alpha,\tilde{q}) - f_{1,m}(\alpha,\tilde{q})^2\frac{\partial f_{0,m}}{\partial \tilde{q}}(\alpha,\tilde{q})}{f_{0,m}(\alpha,\tilde{q})^2}}{\frac{p_{\alpha,1}(\tilde{q}/2)}{2}\left[\ln f_{0,0}(\alpha,\tilde{q})+1\right] + \sum_{m=1}^{+\infty}\frac{\partial f_{0,m}}{\partial \tilde{q}}(\alpha,\tilde{q})\left[\ln f_{0,m}(\alpha,\tilde{q})+1\right]} ln2 \tag{3.45}$$

with

$$\frac{\partial f_{n,m}}{\partial \tilde{q}}(\alpha, \tilde{q}) = \left[\left(m + \frac{1}{2} \right)^{n+1} p_{\alpha,1} \left(m\tilde{q} + \frac{\tilde{q}}{2} \right) - \left(m - \frac{1}{2} \right)^{n+1} p_{\alpha,1} \left(m\tilde{q} - \frac{\tilde{q}}{2} \right) \right] \tilde{q}^n. \tag{3.46}$$

The $\frac{\partial D}{\partial R}$, R and D functions are not easily invertible. However, for the generalized Gaussian model it has been shown in [119] that $\frac{\partial D}{\partial R}$, R and D functions can be tabulated in order to simplify the inversion. The tables for these functions can be found in [119, 122] for different generalized Gaussian shape parameters α sampled in $]0, 2]$. The tabulated $\frac{\partial D}{\partial R}$ and R functions are used in the proposed algorithm that is presented in the previous section 3.4.

3.6 Results

3.6.1 Central PSNR vs. side PSNR for still image

For 1 bpp central bit rate and 512×512 *Lena* image, central PSNR vs. side PSNR is plotted in Fig. 3.5 and Fig. 3.6 for various values of r_N between 0 and 1.

As said above, $r_N = 0$ represents no redundancy, that implies the higher central PSNR and the lowest side PSNR. Inversely, $r_N = 1$ represents the highest redundancy (there is no difference between primary and redundant subbands), that implies the lowest central PSNR and the highest side PSNR.

We compare our application with the best Multiple Description Coding techniques we know to date that are presented in [106, 82]. We present two different results. In the first one, without $C_{i,j}$ optimization, the bit allocation procedure was performed with $C_1 = \{1, r_N, 1, r_N, ...\}$ and $C_2 = \{r_N, 1, r_N, ...\}$, resulting in a very simple MDC scheme. In the second one, with $C_{i,j}$ optimization, we found the best set of C_j using the algorithm of section 3.4.

The proposed method with $C_{i,j}$ optimization provides the best results.

For high r_N values (near 1) the performances of the different methods are similar. We note that $r_N \approx 1$ means the highest redundancy. This meaning that the MDC performance for this values approximates the performance obtained when using an SDC and double the resultant bitstream. We can then conclude that the codec performance of the different methods used for comparisons are similar.

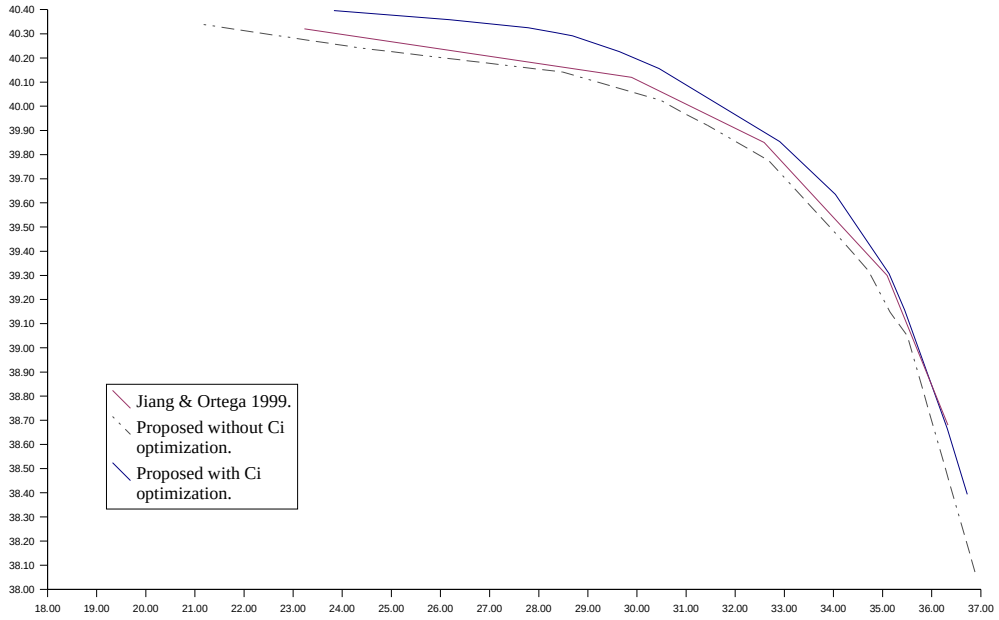

Figure 3.5: Side PSNR vs Central PSNR. Comparison of the proposed method with the method in[82]

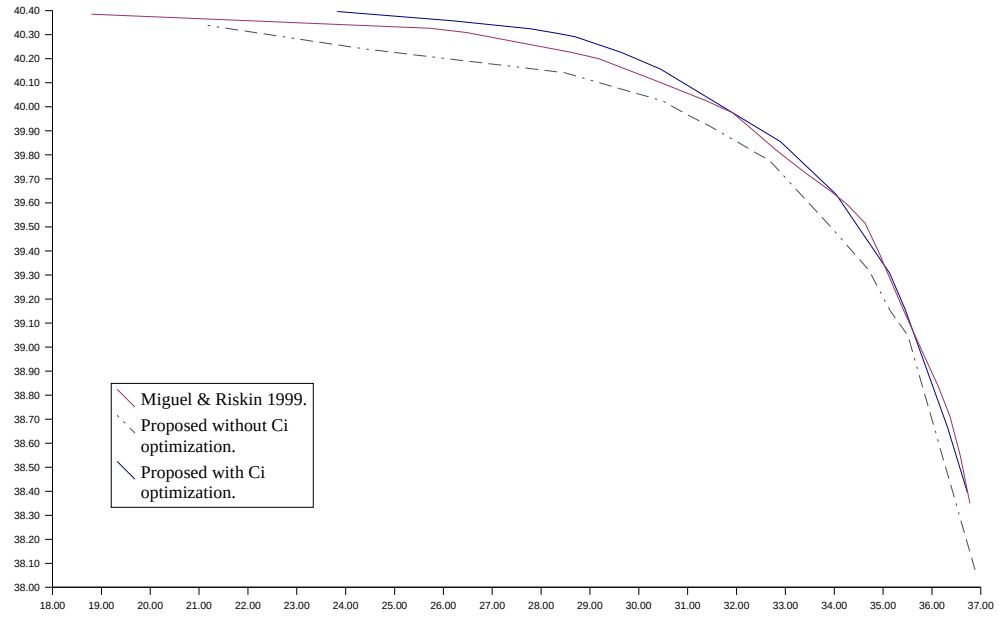

Figure 3.6: Side PSNR vs Central PSNR. Comparison of the proposed method with the method in[106]

3.7 Conclusions

In this chapter a joint source and channel MDC is presented. We will show next chapter that it can automatically adapt the amount of added redundancy according to underlying channel error characteristics.

The proposed method uses a MD scheme based on the Discrete Wavelet Transform (DWT) and an efficient bit allocation technique. The different MD are defined when setting the bit allocation of each subband. We name it Multiple Description Bit Allocation (MDBA).

The results presented show that the proposed MDC overcomes previous MDC schemes in the relation side PSNR vs. central PSNR.

The work on this chapter resulted in several publications. In [127] we presented the MDBA for image coding. In [128, 129] we presented an extension for video coding. In [129] we present a low complexity scheme of the method proposed in [128] that includes the $C_{i,j}$ computation presented in section 3.3.6.

Chapter 4

Channel adapted multiple description coding

In this chapter we made the MDBA presented in previous chapter automatically adapted to channel model and state. For this we use channel information to inject redundancy in the different descriptions. More precisely, we use channel information when computing the r_N parameter presented in previous chapter 3.

We take into account the Shannon theorem (Theorem 10) [156], and propose to define the redundancy parameter using the equivocation $H_y(x)$. Indeed, in this theorem, Shannon states that the equivocation $H_y(x)$ is the amount of redundancy that the decoder needs to correct the received message.

The chapter is organized as follow. Section 4.2 presents some of the existing MDC schemes that are adapted to channel models which are not on-off channels. In section 4.3 we expose how the redundancy parameter is computed. It is this parameter that allows the MDBA presented in previous chapter to be adapted to different channel models and states. This section is followed by section 4.4 that presents concisely the channel models used in this work and presents the specific redundancy computation for each of the presented channel models.

Section 4.5 presents the general MDC and explains how it can be used for image or for video coding. Simulation results for still image and for video, for different channel models are presented in section 4.6 and 4.7. We conclude in section 4.8.

4.1 Introduction

"A challenging task is how to design a MDC that can automatically adapt the amount of added redundancy according to underlying channel error characteristics" [195].

We consider that challenge in the present chapter. We present a method to control automatically the amount of redundancy dispatched on the different descriptions of the balanced MDBA presented in the previous chapter. The proposed method takes into account the channel model and state and can deal with time varying state channels (see figure 4.1).

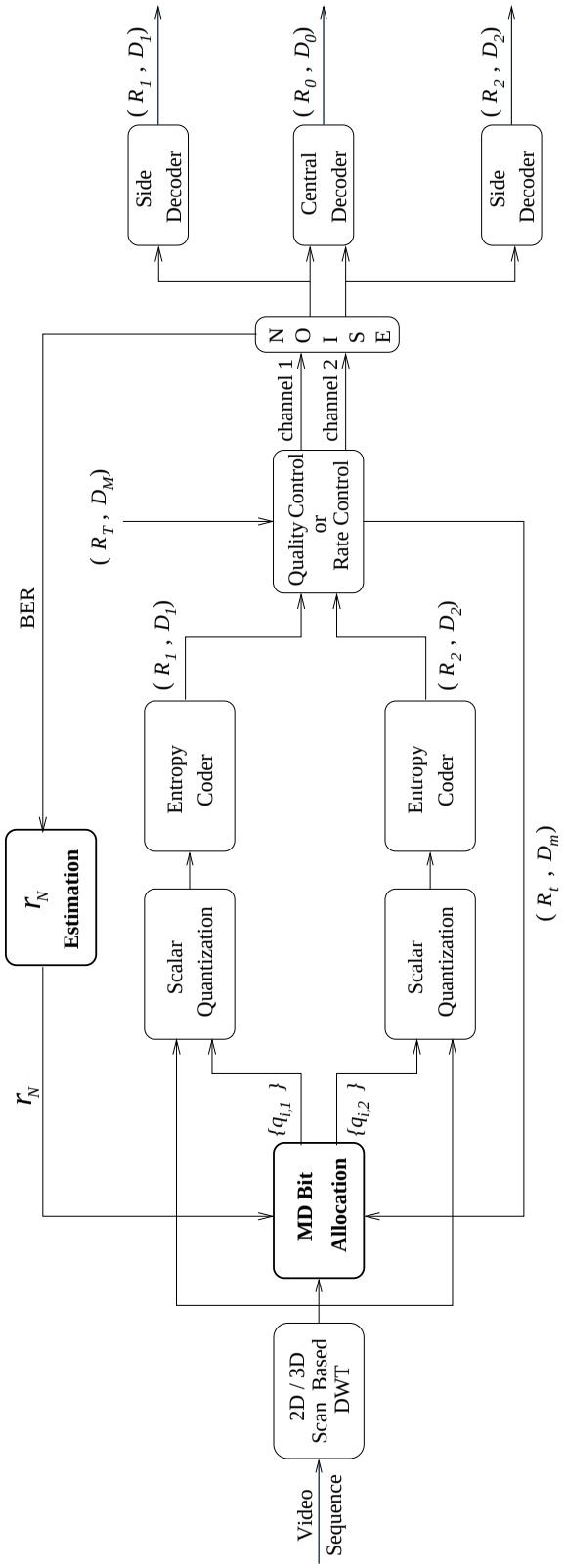

Figure 4.1: Complete Coding Scheme

To be automatically adapted to channels model and state the proposed method uses the channel capacity and the entropy of the input channel symbols to compute the redundancy parameter introduced in section 3.3.2 of chapter 3. These channel parameters (channel capacity and entropy of the input) are known for several channel models. In the present chapter we present some applications of image transmission over BSC, Gaussian channels. The method can be efficiently extended to other channels when the model that matches the channel behavior is known.

Favoring the use of DWT is the fact that 2D DWT can be easily extended to 3D and thus applied to video coding. A 3D Scan-Based DWT video coder is presented in [120]. In this chapter we present an extension of the proposed MDBA with automatic adaptation to channel model and state, for video. The 3D Scan-Based DWT transform allows us to develop a stripe-based MDC and to use different redundancies to take into account changes in channels state while coding. Some applications of video transmission over Gaussian, UMTS and Internet channels are presented.

4.2 Existing MDC for channel models different from on-off channels

Almost all multiple description codes to date assume the existence of multiple independent on-off channels between the transmitter and the receiver (ex: Internet). When a link is broken, all of the symbols or packets passing through that channel are lost; when it is functioning properly, the symbols are transmitted error free. Some exceptions are presented in the following.

In [85] the authors replace the on-off channel model with wireless channel, where they assume that the system will employ multiple transmit and multiple receive antennas. In this paper they observe that it is possible to improve average transmission error probability by a proper choice of the correlating transform. In [11] they introduce fading models, such as Rayleigh, Rician, or Nakagami channel, within the context of MDC. Their simulations show the efficiency of MDC for fading channels with multiple antennas.

Vaishampayan in [193] showed that for transmitting information from memoryless Gaussian source over a Rayleigh fading channel, the multiple description approach results in good performances at low interleaving delays as compared to standard channel coding approaches. This conclusion was extended to sources with memory in [89] where, on an equal interleaving delay basis, significant performance improvements are obtained over channel codes for speech transmission on Rayleigh fading channels.

The great amount of research dedicated to MDC for on-off channels vs. the very reduced amount of research dedicated to MDC for other kind of channels is explained by the relatively large overhead associated with existing MDC that implies that when channel loss rate is small, the reconstruction performance in the error free case dominates and a single description coding (SDC) without channel coding performs best.

We could see in the previous chapter that the proposed MDBA is comparable

with the SDC without channel coding when channel loss rate is small. Thus, it is suitable for transmissions even for different channels from the on-off channels usually used in MDC applications. We will show next, how we use channel information to dispatch redundancy between the descriptions. In this way we can adapt the coder to any channel since the model that matches the channel behavior is known.

4.3 Channel adaptation through redundancy parameter

4.3.1 MDBA recall

Remark that in figure 4.1 there are two main parts: the MD Bit Allocation bloc that was detailed in previous chapter 3; and the r_N Estimation bloc that will be detailed in this section.

We have shown in chapter 3 that the r_N parameter affects the central distortion. This central distortion was defined in previous chapter 3 as

$$D_{i,0}\left(\frac{q_{i,1}}{\sigma_{i,1}},\frac{q_{i,2}}{\sigma_{i,2}}\right)=\frac{1}{\sigma_{i,0}^2}\frac{1}{r_N+1}\left[min\left(\sigma_{i,1}^2D_{i,1}\left(\frac{q_{i,1}}{\sigma_{i,1}}\right),\sigma_{i,2}^2D_{i,2}\left(\frac{q_{i,2}}{\sigma_{i,2}}\right)\right)\right.$$
$$\left.+r_N\times max\left(\sigma_{i,1}^2D_{i,1}\left(\frac{q_{i,1}}{\sigma_{i,1}}\right),\sigma_{i,2}^2D_{i,2}\left(\frac{q_{i,2}}{\sigma_{i,2}}\right)\right)\right] \quad (4.1)$$

where r_N is the redundancy parameter and $\sigma_{i,j}^2D_{i,j}(\frac{q_{i,j}}{\sigma_{i,j}})$ is the Mean Square Error for the ith subband, in the case of a generalized Gaussian distribution.

Actually, the level of redundancy should increase when the BER increases, such as in the case of very noisy channels where the redundant subbands are as important as the others.

The amount of redundancy, i.e., the importance of the redundant subbands, depends on the channel model and BER (as can be seen in figure 4.1) and we propose in the following some strategies for the choice of r_N.

4.3.2 Redundancy computation

Taking in consideration what is said above, it is easy to conclude that the redundancy parameter domain is $[0,1]$. Parameter $r_N=0$ when the channel is noiseless and parameter $r_N=1$ when a very noisy channel is expected. The problem is how to choose intermediate redundancies, and implicitly intermediate values of r_N. Taking into account the Shannon theorem (Theorem 10) [156], we propose to solve this problem using the equivocation $H_y(x)$. Indeed, in this theorem, Shannon states that the equivocation $H_y(x)$ is the amount of redundancy that the decoder needs to correct the received message.

We propose to use equation (4.2) to compute the redundancy parameter.

$$r_N=\frac{H_y(x)}{\underset{p(x)}{max}(H(x))}, \quad (4.2)$$

In equation (4.2), $p(x)$ stands for the distribution of the input channel symbols and $H(x)$ for the entropy of the input.

$H_y(x)$ is unknown at encoding, however we can find some bounds for equivocation by stating the following proposition 2.

Proposition 2

$$\min_{p(x)}(H_y(x)) \leqslant \max_{p(x)}(H(x)) - C \leqslant \max_{p(x)}(H_y(x)) \tag{4.3}$$

where C is the channel capacity defined by equation (4.4).

$$C = \max_{p(x)}(H(x) - H_y(x)). \tag{4.4}$$

Proof :

Being

$$C = \max_{p(x)}(H(x) - H_y(x))$$

we can infer that

$$C \leqslant \max_{p(x)}(H(x)) - \min_{p(x)}(H_y(x)).$$

Thus,

$$\min_{p(x)}(H_y(x)) \leqslant \max_{p(x)}(H(x)) - C.$$

For the right bound, we start also from the channel capacity definition,

$$C = \max_{p(x)}(H(x) - H_y(x))$$

and we infer that

$$C \geq \max_{p(x)}(H(x) - \max_{p(x)}(H_y(x))) = \max_{p(x)}(H(x)) - \max_{p(x)}(H_y(x)).$$

Thus, we can conclude that

$$\max_{p(x)}(H_y(x)) \geq \max_{p(x)}(H(x)) - C.$$

Therefore, by proposition 2, instead of (4.2), we use (4.5) to compute the redundancy parameter.

$$r_N = \frac{\max_{p(x)}(H(x)) - C}{\max_{p(x)}(H(x))}. \tag{4.5}$$

We know that $0 \leqslant H_y(x) \leqslant H(x)$. Thus, $0 \leqslant C \leqslant \max_{p(x)}(H(x))$. Using (4.5) we can conclude that $0 \leqslant r_N \leqslant 1$ as pretended.

We will show in section 4.4 that the channel capacity is known for several interesting channel models. Thus, more details about the calculi of r_N parameters for different channel models will be done in that section.

4.4 Channel models and associated redundancies

Typical channel models used in the analysis of wireless transmission include memoryless channel, symmetric channel, additive white Gaussian noise (AWGN) channel and burst channel. Here we give a more detailed description of each of these models.

Memoryless channel is also known as a random bit error channel and is characterized by errors that are independent from one symbol to the next;

Symmetric channel for all values of i and j, the probability of receive a j when transmitted the symbol i is the same of receive a i when transmitted the symbol j. A commonly encountered example is the Binary Symmetric Channel (BSC) with a probability p of bit error;

Additive White Gaussian Noise (AWGN) channel is a memoryless channel in which the transmitted signal suffers from the addition of wide-band noise whose amplitude is a normally distributed random variable. AWGN is the most common form of a memoryless channel;

Burst channel is the channel where errors are characterized by periods of relatively high symbol error rate separated by periods of relatively low, or zero, error rate. A burst error means that the probability of error is dependent from one symbol to the next.

The two main elements which describe a channel are the transmission rate and the channel capacity. The transmission rate was defined by Shannon as presented in equation 4.6 for discrete or continuous channels.

$$R = H(x) - H_y(x) \tag{4.6}$$

Where $H(x)$ is the entropy of the input and $H_y(x)$ the conditional entropy or the equivocation. The channel capacity C is defined as the maximum of R when the input varies over all possible collection (equation (4.4)). This channel capacity was defined by Shannon for the BSC and the Gaussian cases. However, in most mobile radio systems, the channel exhibits Rayleigh fading, aggravated by typically log-normally distributed shadowing or slow fading, resulting in a time variant channel capacity. Lee [93] derived an estimate of the channel capacity in Rayleigh fading environments.

In this manuscript we focus on Binary Symmetric (BSC), Gaussian and Rayleigh channel models. We present below the computation of the r_N parameter for these channel models.

4.4.1 Binary symmetric channel

For the BSC case we have two possible symbols each with a probability p of coming through undisturbed, and $(1-p)$ of being changed into the other of the pair. The capacity (measured in bits/symbol) can be written as presented in equation (4.7) [156].

$$C = 1 + plog_2p + (1-p)log_2(1-p) \quad \text{bits/symbol} \tag{4.7}$$

For this channel model $max(H(x)) = 1$. Thus the redundancy parameter we present in section 4.3 can be estimated by equation (4.8).

$$r_N = \frac{\max\limits_{p(x)}(H(x)) - C}{\max\limits_{p(x)}(H(x))} = 1 - C = plog_2p + (1-p)log_2(1-p) \tag{4.8}$$

4.4.2 Additive white Gaussian noise channel

For a band limited Additive White Gaussian Noise (AWGN) the channel capacity, in bits/symbol, can be expressed as in equation (4.9) [156].

$$\frac{C}{B} = log_2(1+\gamma) \quad \text{bits/symbol} \tag{4.9}$$

In equation (4.9) B is the channel bandwidth in symbol/s and γ is the signal to noise ratio (SNR). The SNR γ is defined as $\gamma = \frac{S}{N}$, where S is the received signal power and N is the AWGN power within the channel bandwidth. For this channel model $max(H(x))$ depends on the modulation. For instance, if we consider a QPSK modulation $max(H(x)) = 2$ and in this case the redundancy parameter in section 4.3 is estimated by equation (4.10).

$$r_N = \frac{\max\limits_{p(x)}(H(x)) - C}{\max\limits_{p(x)}(H(x))} = \frac{2 - B\,log_2(1+\frac{S}{N})}{2} = 1 - \frac{B\,log_2(1+\frac{S}{N})}{2} \tag{4.10}$$

4.4.3 Rayleigh channel

In the case of Rayleigh models, an upper bound approximation for the normalized channel capacity was introduced by Lee's [93] as presented in equation (4.11).

$$\frac{C}{B} \approx log_2e.e^{\frac{-1}{\gamma}}(-e + ln\gamma + \frac{1}{\gamma}) \quad \text{bits/symbol} \tag{4.11}$$

For this channel model $max(H(x))$ depends also on modulation. For instance, if we consider a QPSK, $max(H(x)) = 2$ and the redundancy parameter can be written as equation (4.12).

$$r_N = \frac{\max\limits_{p(x)}(H(x)) - C}{\max\limits_{p(x)}(H(x))} = 1 - \frac{B\,log_2e.e^{-\frac{N}{S}}(-e + ln\frac{S}{N} + \frac{N}{S})}{2} \tag{4.12}$$

4.5 General coder

Standard transform based image coding algorithms consist of three stages: a linear decorrelating transform, followed by a quantization stage, and final entropy coding of the quantized data.

Figure 4.2 shows the coding scheme we use. We start with a DWT. Then the MDBA presented in chapters 3 and 4 is processed. This one is followed by a simple scalar quantization and the encoding of each subband uses a context based arithmetic coder [119, 167].

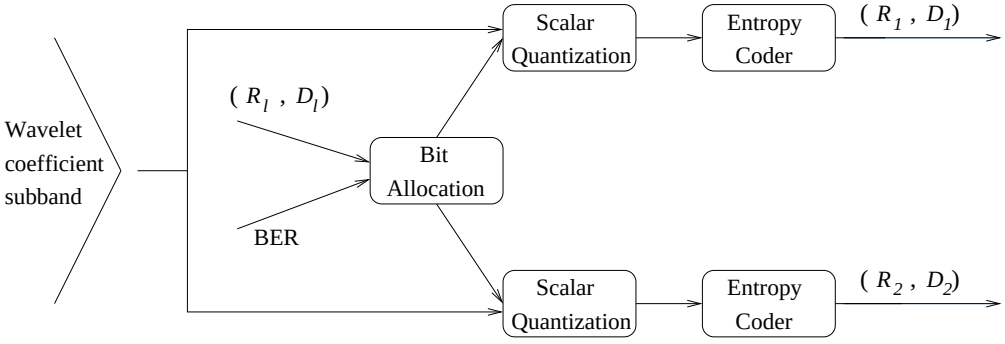

Figure 4.2: Transform based image coding scheme.

4.5.1 Entropy coding

Entropy coding is a reversible process, thus there is no approximation problem as in quantization. After quantization the variables take values drawn from a finite set $\{a_i\}$. The idea is to find a reversible mapping M to a new set $\{b_i\}$ such that the average number of bits/symbols is minimized. The parameters in searching for the mapping M are the probabilities of occurrence of the symbols $a_i, p(a_i)$.

Arithmetic coding converts a variable number of samples into a variable length codeword. In this way it can approach the entropy bound more closely than Huffman coding for common signals. In Huffman coding each sample (or group of samples) uses at least one bit. Therefore, for a very small alphabet, the bit rate cannot be lower than 1 bit/sample, in the case of scalar coding.

Our coder uses arithmetic coding. We explore this coding through what we call a *Smart Arithmetic Coding* to provide synchronization and minimize the error propagation in the case of channel errors. In the following we present the basis of arithmetic coding needed to understand the proposed *Smart Arithmetic Coding* presented in section 4.5.1.3.

4.5.1.1 Arithmetic coding

The idea in arithmetic coding is to represent a sequence of symbols by an interval in the line segment from 0 to 1, with length equal to the probability of the sequence. Because the probabilities of all possible sequences sum to 1, the

intervals corresponding to all possible sequences will fill up the entire line segment. The code bits for a sequence are essentially the binary representation of any point in the interval corresponding to the sequence.

We give in the following a simple example of arithmetic coding detailing the parameters needed for the algorithm. This example will contribute to the understanding of the proposed smart arithmetic coding presented in section 4.5.1.3.

4.5.1.2 Example

Consider the parameters and the table presented below. This table presents the probability and associated interval of each symbol.

Parameters :

I Coding interval;

Int Interval associated to each symbol

Icod Interval representing the coding of a specific sequence of symbols

Symbol	Probabilities	Int
A	0.2	$[0; 0.2]$
B	0.4	$[0.2; 0.6]$
C	0.2	$[0.6; 0.8]$
D	0.1	$[0.8; 0.9]$
/	0.1	$[0.9; 1.0]$

Table 4.1: Probability and associated interval of each symbol.

The coding of the sequence **CDCB**/ is done as presented bellow:

$$\begin{aligned}
\text{Starting point} \quad & I = [0; 1] \\
\text{Coding symbol C} \quad & I = [0.2; 0.6] \\
\text{Coding symbol D} \quad & I = [0.44; 0.52] \\
\text{Coding symbol C} \quad & I = [0.456; 0.488] \\
\text{Coding symbol B} \quad & I = [0.456; 0.4624] \\
\text{Coding symbol/} \quad & I = [0.46176; 0.4624]
\end{aligned}$$

Thus, in the end of the arithmetic coding process we have

$$Icod = [0.46176; 0.4624]$$

In practice our algorithm thus not use the interval $I = [I_{min}; I_{max}]$ but uses only I_{min} and the size of the interval that we call in the following the *A register*.

For decoding we start from the same starting point of the coding procedure and from $Icod$. We search $\{I|Icod \subset I\}$. We finish when $Icod = I$:

$I = [0; 1]$	Starting point
$Icod \subset I = [0.2; 0.6]$	Decoding symbol C
$Icod \subset I = [0.44; 0.52]$	Decoding symbol D
$Icod \subset I = [0.456; 0.488]$	Decoding symbol C
$Icod \subset I = [0.456; 0.4624]$	Decoding symbol B
$Icod = I = [0.46176; 0.4624]$	Decoding symbol/

Note that:

- the termination condition could be $Icod_{max} = I$, since the coding of the terminal symbol does not change I_{max};
- the decoding procedure only needs the $Icod_{max}$ to execute;
- and finally, the interval size in the end of the arithmetic coding and decoding process is the same.

In fact, the arithmetic coding works essentially with two registers, one representing the $Icod_{min}$ and the other, called *A register* representing the interval I size. The arithmetic decoding works also with two registers. One representing the $Icod_{max}$ and the other, is also the *A register*.

4.5.1.3 Proposed smart arithmetic coding

In order to provide synchronization and minimize the error propagation in the case of channel errors, each spatio-temporal subband is divided into blocks. Then, arithmetic coding is performed on each block independently. Because the block division requires side information, the block-size must be related to the channel BER such that high BER implies small block-sizes. For error detection, when the number of coded coefficients is known, it is possible to verify if the arithmetic coder stops correctly. In case of error and if the arithmetic coder is misplaced in the bitstream we synchronize the decoder to start at the beginning of next block.

It is well known that this usual method of error detection is largely insufficient. It is not rare to have the correct number of coefficients decoded and the arithmetic coder well placed, but wrong coefficients decoded. This problem is more difficult to solve. Some methods try to use the information of the neighbor blocks but their effectiveness is not optimal. Here, we propose to use some internal information included in the arithmetic code.

As explained above the arithmetic coder works with two main registers: the code string that represents the *base of the interval* (I_{min}) and the *interval* (*A register*). With each binary decision the current probability interval is subdivided into two sub-intervals, and the registers are modified in accordance. The states of registers depend only on the probability of the individual events. In

our scheme, we include the n last bits of the interval (*A register*) inside the final bitstream (in our experiments $n = 8$). We could see in the example above that this interval has the same value in coding and in decoding procedure. The decoder simply verifies if its final interval register corresponds to the one included by the coder inside the header of the bitstream. If this is not the case, it means an error occurred in the block and it is discarded.

4.5.2 Effect of noise in the headers

The proposed smart arithmetic coder avoids synchronization problems and minimize error propagation. With this method an error in normal data can be detected, in which case the block will be decoded using zeros or the average of the block (if decoding the LL subband).

Errors in the headers specifying crucial information can have small effect or be catastrophic. In the following we detail the headers information used and the error tolerance of such headers.

Image information (color space, image size, wavelet decomposition, block partition, etc). This header is introduced in both descriptions. If it is lost in both descriptions the image/video reconstruction is impossible, otherwise the image/video reconstruction is normal.

Size (in bits) of a block This header is used when some error occurs in a block. In such case the decoding process starts decoding the next block. If some error occurs in this header, in the data of the block and at same time the arithmetic decoding finishes in the wrong position, we have a synchronization problem.

Register A This header is used to verify that arithmetic decoding succeeded. In case of errors in this header the most probable is that the arithmetic decoding terminates with "not succeeded", even if it succeeds. Much improbable is that the arithmetic decoding terminates with "succeeded" when it not succeeded. In this last case we have a synchronization problem.

For each block, 24 bits for statistical information (in the case of a LL subband, 40 bits for statistical information). This header is introduced in both descriptions. If it is lost in both descriptions the block it corresponds is decoded erroneously in both descriptions.

For each block, 16 bits for the quantization step If this header is lost the block it corresponds is decoded erroneously.

We note that the most sensitive headers are repeated in both descriptions. The other headers are less sensitive since reconstruction is still possible and errors do not propagate to other blocks.

In the performed simulations the headers were protected (they are not afflicted by noise). In future work we intend to protect the most sensitive headers with FEC's.

4.6 Still image simulations

For spatial decomposition our coder uses a 9-7 biorthogonal filter [6] and a three level decomposition.

For simulations we use the 512×512 pixels *Lena* image. Lena image total bit rate was set to 1 bpp (i.e. $R_1 \approx R_2 \approx 0.5$ bpp) or 0.5 bpp (i.e. $R_1 \approx R_2 \approx 0.25$ bpp).

4.6.1 BSC and Gaussian channels

Figures 4.3, 4.4, 4.5, 4.6, 4.7 and 4.8 show some visual results of transmission of *Lena* image at 1 bpp and 0.5 bpp over BSC and Gaussian channels with $BER = 0.001$ and BER= 0.01. The mean PSNR we obtained for BSC channels and using *Lena* image at 0.5 bpp are tabulated in table 4.2.

As already mentioned, we can find a great amount of research dedicated to MDC for on-off channels but very reduced amount of research dedicated to MDC for channels as BSC or Gaussian. In [162] the authors constructed an MDC for networks with packet lost and/or bit errors. They provide a mechanism for bit allocation between the redundancy in terms of FEC and redundancy that is meant to correct for packet loss. In this work they present some results with *Lena* image compressed at 0.5 bpp. For BSC transmission with BER= 0.001 the PSNR is 27.5 dB and for a BER= 0.01 the PSNR obtained is 20.2 dB. This means that our MDC outperforms this one (see table 4.2). Moreover, the results of [162] where obtained with seven levels of decomposition and assuming the LL band uncorrupted.

	BSC channels	
BER	0.01	0.001
Proposed method	24.79 dB	28.62 dB
Method in [162]	20.2 dB	27.5 dB

Table 4.2: PSNR values for *Lena* image compressed to 0.5 bpp, when considering BSC channels.

Figure 4.3: *Lena* image coded at 0.5 bpp. BSC channel at 0.001 ber. PSNR=29.41;

Figure 4.4: *Lena* image coded at 1.0 bpp. BSC channel at 0.001 ber. PSNR=33.89 dB.

Figure 4.5: *Lena* image coded at 1.0 bpp. BSC channel at 0.01 ber. PSNR=24.99 dB.

Figure 4.6: *Lena* image coded at 1.0 bpp. Gaussian channel at 0.01 ber. PSNR=25.49 dB.

Figure 4.7: *Lena* image coded at 0.5 bpp. Gaussian channel at 0.001 ber. PSNR=31.27 dB.

Figure 4.8: *Lena* image coded at 1.0 bpp. Gaussian channel at 0.001 ber. PSNR=34.90 dB.

4.7 Extension of the multiple description coder to video

The MDC presented for images can be extended to video adding a 1D DWT in the time direction to the 2D DWT. This is named a 3D DWT. The 3D subband coding of video [87, 136, 47, 181, 192] provides encouraging results compared with MPEG [26, 169, 88, 111]. Furthermore there exist efficient 3D scan-based DWT and 3D scan-based motion compensated lifting DWT for video coding [120, 121, 122, 192, 168, 103, 5, 18] that are well suited for real time applications.

In the present work we use the 3D scan-based DWT presented in [120], without motion compensation. The 3D scan-based DWT transform allows us to develop a stripe-based MDC. More specifically, when using a 3D scan-based DWT, the bit allocation for the successive sets of temporally coherent coefficients can be performed with respect to either rate or quality constraints (see [122] for details). Thus, we can change the redundancy parameter each time a bit allocation starts for a new set of temporally coherent coefficients. In this way we can adapt the redundancy parameter to the channel state while coding. Thus, the proposed method automatically adapts the coding process to time varying states. The proposed MDBA is then suitable for transmission over time-varying channels.

4.7.1 Scan-based wavelet transform

Scan-based wavelet transform algorithms are meant to progressively compute the DWT to get low memory implementations of wavelet transforms. This problem was first addressed in [180] which only considered 1D transforms. In [30] the authors present a line based wavelet image compression. In this scheme the lines are read line by line and only the minimum required number of lines is kept on memory. This kind of algorithms only needs to store a small number of coefficients at the same time and allow a significant memory saving. Historically, first unpublished proposal for 2D images can be found in [22]. Other proposals are [124, 120, 122, 113]. An easy way to implement a scan based wavelet transform algorithm can be derived from the lifting scheme. Some algorithms using the lifting scheme are [31, 191, 192, 121].

4.7.2 The lifting scheme

The lifting scheme has been introduced by Sweldens [165, 37] as a light-weight computation method for performing any wavelet transform based on a biorthogonal filter bank. The usual low-pass/high-pass filter pair is replaced by a ladder structure as shown in figure 4.9.

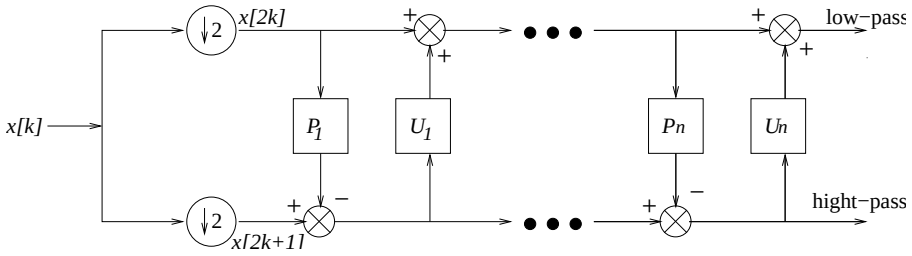

Figure 4.9: 1D DWT structure using the lifting scheme.

The lifting scheme starts by a "lazy" wavelet transform consisting of splitting the input vector $x[k]$ into two subsampled versions $x_{even}[k] = x[2k]$ and $x_{odd}[k] = x[2K+1]$. Several elementary transformation steps are applied to these two vectors:

lifting steps consist of subtracting from $x_{odd}[k]$ an estimated (or predicted) signal based on $x_{even}[k]$. Estimations are computed by applying the P_i filters to $x_{even}[k]$.

dual lifting steps consist of updating $x_{even}[k]$ by adding the result of U_i filters when applied to $x_{odd}[k]$.

After several lifting/dual lifting steps the detail and approximation coefficients of the 1D DWT of $x[k]$ is obtained. The low-pass coefficients are obtained after the last dual lifting step, while the high-pass coefficients are obtained after the last lifting step.

We use $(L_1, L_2, ..., L_N)$ to denote a lifting scheme with N steps, where L_{2n-1} is the coefficient number of the n prediction operator, and L_{2n} is the coefficient number of the n update operator. The biorthogonal 5-3 filter can then be implemented using a (2,2) lifting scheme, while the biorthogonal filter 9-7 needs a lifting scheme with four steps (2,2,2,2).

4.7.3 Previous MD coding dedicated to video

A MD coding dedicated to video was proposed by Vaishampayan in [176], where a predictive MD system was applied along with transform coding to construct an inter-frame balanced MD video coder based on the H.263 standard. In [9] Apostolopoulos and Wee show that MD coding and path diversity provide improved reliability in systems with multiple paths with equal or unequal bandwidths. In [146] Reibman et al. proposed MD video coders which use motion compensated predictions. In [120] we propose an extension of [127] for video that uses the 3D scan-based DWT.

4.7.4 Video simulations

For spatial decomposition our coder uses 9-7 biorthogonal filter [6] and three levels of decomposition. For temporal decomposition it uses the (2,2) filter and performs two levels of decomposition. The frames of the video sequence are acquired and processed on the fly to generate the 3D wavelet coefficients and

the data are stored in memory only until these coefficients have been encoded [131]. No motion compensation is performed.

4.7.4.1 Gaussian channel simulations

For Gaussian channel simulations we use 3 seconds of QCIF *silent* color video and 3s of QCIF *akyio* video. The *silent* video was compressed to 200 kbits/s (30 frames/s) and the *akyio* video was compressed to 300 kbits/s (30 frames/s). All channel simulations were performed ten times. Note that mean PSNR values are computed by averaging decoded MSE values and then converting the mean MSE to the corresponding PSNR values. Visual results of video present always the frames 1, 11, 21, 31, 41, 51, 61, 71, 81 and 91 of the video.

For comparison, we present results obtained with the proposed MDC with and without noise. We present also some results obtained with transmission of a singular description coder (SDC) using a similar codec without channel coding and an SDC with a Turbo Coder (SDC+TC) with a bit rate of 200kbits/s including the channel rate.

In table 4.3 the mean PSNR for BER= 0.001 for the proposed MDC is presented. We compare it with the case without noise. We can see that the performance of the proposed MDC in the presence of noise are similar to the performance without noise, especially for the U and V components. We can conclude from this table that the proposed MDC provides a gain of 3 dB over a standard method using SDC+TC.

	Y BER		U BER		V BER	
	0	10^{-3}	0	10^{-3}	0	10^{-3}
Proposed Method	33.17	31.50	40.72	40.00	41.51	41.15
SDC + TC	-	28.66	-	36.50	-	38.12

Table 4.3: Mean PSNR results for QCIF *silent* color video compressed to 200 kbits/s (30 frames/s), when channel transmission at 0.001 ber.

Figures 4.10, 4.11 and 4.12 present the mean PSNR of different frames for Y, U, and V, respectively, when transmission over Gaussian channel with a BER= 0.001. We can conclude that the proposed MDC is able to recover from channel losses. In these figures we can also see that the Y component is the most sensitive to noise.

Finally we show some visual results in figures 4.13, 4.14, for QCIF *silent* video and 4.15 and 4.16 for CIF *akyio* video. All these results consider Gaussian transmission for BER= 0.001. These images allow the comparison of the proposed MDC with the SDC+TC.

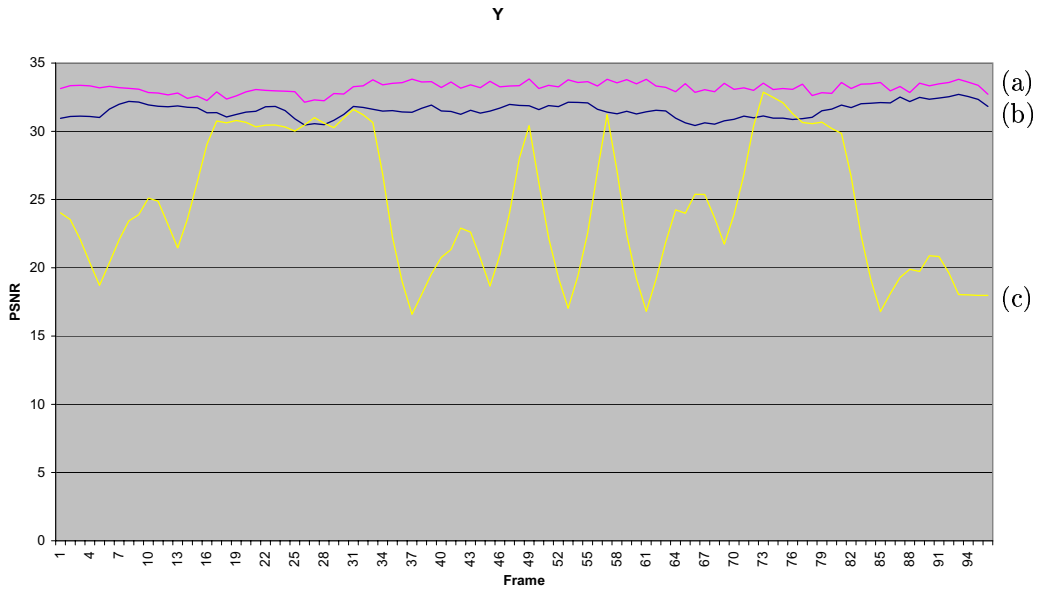

Figure 4.10: Mean PSNR's of each frame for Y component of QCIF *silent* color video compressed to 200 kbits/s (30 frames/s). Transmission over Gaussian channel at 0.001 ber; PINK (a): no noise; DARK BLUE (b): Proposed MDC; YELLOW (c): SDC.

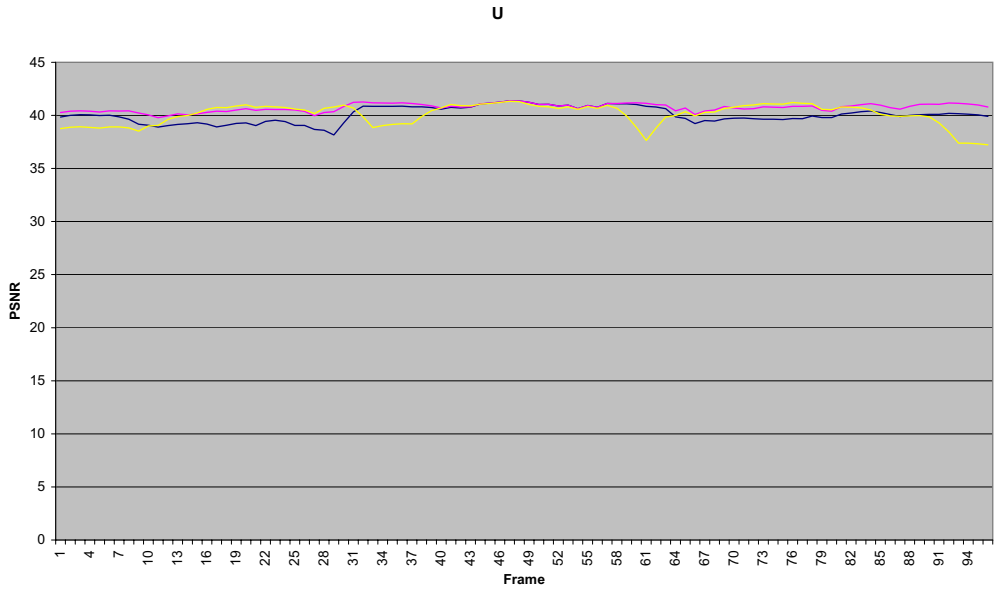

Figure 4.11: Mean PSNR's of each frame for U component of QCIF *silent* color video compressed to 200 kbits/s (30 frames/s). Transmission over Gaussian channel at 0.001 ber; PINK: no noise; DARK BLUE: Proposed MDC; YELLOW: SDC.

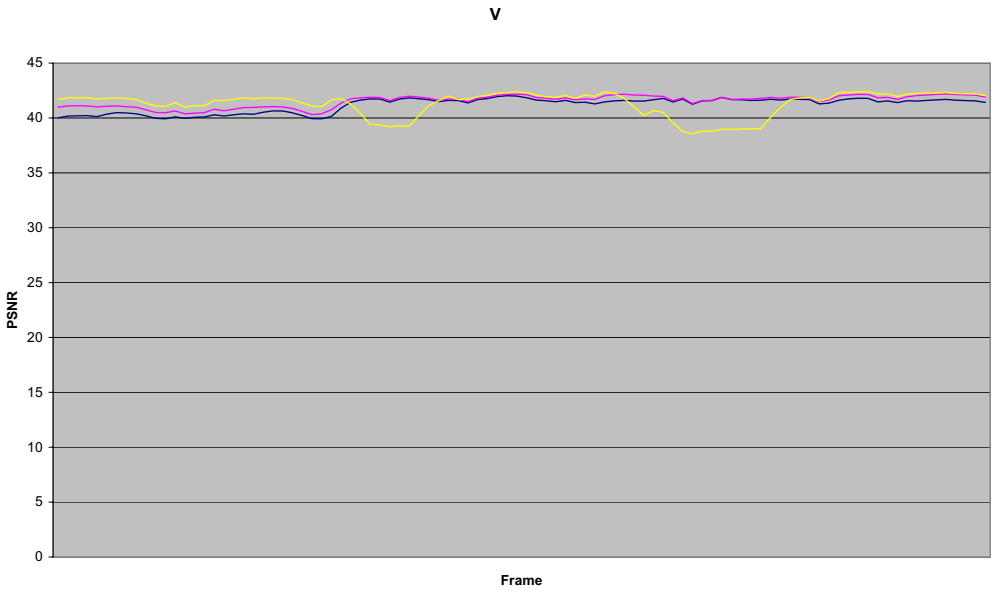

Figure 4.12: Mean PSNR's of each frame for U component of QCIF *silent* color video compressed to 200 kbits/s (30 frames/s). Transmission over Gaussian channel at 0.001 ber; PINK: no noise; DARK BLUE: Proposed MDC; YELLOW: SDC.

Figure 4.13: QCIF *silent* color video compressed to 200 kbits/s (30 frames/s). Transmission over Gaussian channel at 0.001 ber. Using the SDC+TC.

Figure 4.14: QCIF *silent* color video compressed to 200 kbits/s (30 frames/s). Transmission over Gaussian channel at 0.001 ber. Using the proposed MDC.

Figure 4.15: CIF *akyio* video compressed to 300 kbits/s (30 frames/s). Transmission over Gaussian channel at 0.001 ber. Using the SDC+TC.

Figure 4.16: CIF *akyio* video compressed to 300 kbits/s (30 frames/s). Transmission over Gaussian channel at 0.001 ber. Using the proposed MDC.

4.7.4.2 UMTS channel simulations

UMTS has born thanks to the ITU (International Telecommunication Union) organization. This organization made an effort to look for the recommendations and the standards allowing to future generations of mobile communication to answer to the new demands of system capacity and performance. This effort is summarize with the name " mobile communications of third generation (3G)".

UMTS channel[1] presents three different models: indoor, pedestrian, and vehicular. They are defined in [44], and we briefly expose them in Appendix A.

For simulations we use 3seconds of QCIF *silent* color video compressed to 200 kbits/s (30 frames/s). All channel simulations were performed ten times. Note that mean PSNR values are computed by averaging decoded MSE values and then converting the mean MSE to the corresponding PSNR values. Visual results of video present always the frames 1, 11, 21, 31, 41, 51, 61, 71, 81 and 91 of the video.

For comparison, we present results obtained with the proposed MDC with and without noise. The redundancy parameter was computed in this case using equation 4.12. Thus mean we consider the Rayleigh model to approximate the UMTS channel behavior.

We present also some results obtained for the transmission of a singular description coder (SDC) with a similar codec. We also use a SDC with a Turbo Coder (SDC+TC and SDC+TC+P) with a bit rate of 200kbits/s including the channel rate. The difference between the SDC+TC and the SDC+TC+P is that in the last one we adapt the redundancy of the TC based system by using a puncturing mechanism.

Table 4.4 and table 4.5 present the mean PSNR for BER= 0.01 and BER= 0.001, respectively, for the proposed MDC. We compare it with the case without noise and with the SDC+TC and the SDC+TC+P.

The proposed MDC presents better results. The differences in case of BER= 0.01 are smaller, however the BER in real UMTS transmissions is higher than 0.01.

Figures 4.17, 4.18 and 4.19 present the mean PSNR of different frames for Y, U, and V component respectively. In these simulations we consider transmission over pedestrian UMTS channel and BER=0.01.

Finally we show some visual results for SDC, SDC+TC and the proposed MDC in figures 4.20, 4.21, 4.22 and in figures 4.20, 4.21, 4.22 for a BER=0.01. In the first group of figures we consider transmission over an UMTS Indoor channel and in the second group of figures we consider transmission over an UMTS Pedestrian channel. As could be seen in table 4.4 and table 4.5 the Indoor channels are the one that implies the worst PSNR values and the Pedestrian channel the better PSNR values (in both cases MDC, SDC or SDC+TC simulations).

In figures 4.26, 4.27, 4.28 we consider a BER=0.001 and Indoor UMTS channel. As mentioned in chapter 1 the FEC code must be designed with a worst case channel scenario in mind. This reflects the bad performance of the SDC+TC in case of low BER.

[1]The authors wish to thank France Telecom for providing an UMTS simulator

		Y BER		U BER		V BER		Mean BER	
Channel	Method	0	10^{-2}	0	10^{-2}	0	10^{-2}	0	10^{-2}
Indoor UMTS	SDC+TC	28.66	27.96	36.50	36.48	38.12	38.11	34.43	**34.18**
	SDC+TC+P	30.40	21.8	38.10	37.46	39.44	38.72	35.98	**32.66**
	Proposed MDC	31.47	25.57	39.06	38.56	40.06	39.76	36.86	**34.63**
Pedestrian UMTS	SDC+TC	28.66	28.18	36.50	36.44	38.12	38.11	34.43	**34.24**
	SDC+TC+P	30.40	21.56	38.10	37.02	39.44	38.20	35.98	**32.26**
	Proposed MDC	31.47	28.02	39.06	38.67	40.06	39.89	39.76	**35.53**
Vehicular UMTS	SDC+TC	28.66	28.56	36.50	36.48	38.12	38.11	34.43	**34.38**
	SDC+TC+P	30.4	22.45	38.10	37.10	39.44	38.68	35.98	**32.74**
	Proposed MDC	31.47	26.26	39.06	38.58	40.06	39.68	39.76	**34.84**

Table 4.4: Mean PSNR (dB) results for QCIF *silent* color video compressed to 200 kbits/s (30 frames/s). UMTS channel transmission at 0.01 ber.

		Y BER		U BER		V BER		Mean BER	
Channel	Method	0	10^{-3}	0	10^{-3}	0	10^{-3}	0	10^{-3}
Indoor UMTS	SDC+TC	28.66	28.66	36.50	36.50	38.12	38.12	34.43	**34.43**
	SDC+TC+P	30.40	25.39	38.10	37.77	39.44	39.36	35.98	**34.17**
	Proposed MDC	31.54	30.08	39.11	39.07	40.09	40.09	36.91	**36.41**
Pedestrian UMTS	SDC+TC	28.66	28.66	36.50	36.50	38.12	38.12	34.43	**34.43**
	SDC+TC+P	30.40	27.22	38.10	37.80	39.44	39.32	35.98	**34.78**
	Proposed MDC	31.54	31.45	39.11	38.99	40.09	40.04	36.91	**36.83**
Vehicular UMTS	SDC+TC	28.66	28.66	36.50	36.50	38.12	38.12	34.43	**34.43**
	SDC+TC+P	30.4	27.86	38.10	37.90	39.44	39.40	35.98	**35.05**
	Proposed MDC	31.54	31.45	39.11	39.04	40.09	40.06	36.91	**36.85**

Table 4.5: Mean PSNR (dB) results for QCIF *silent* color video compressed to 200 kbits/s (30 frames/s). UMTS channel transmission at 0.001 ber.

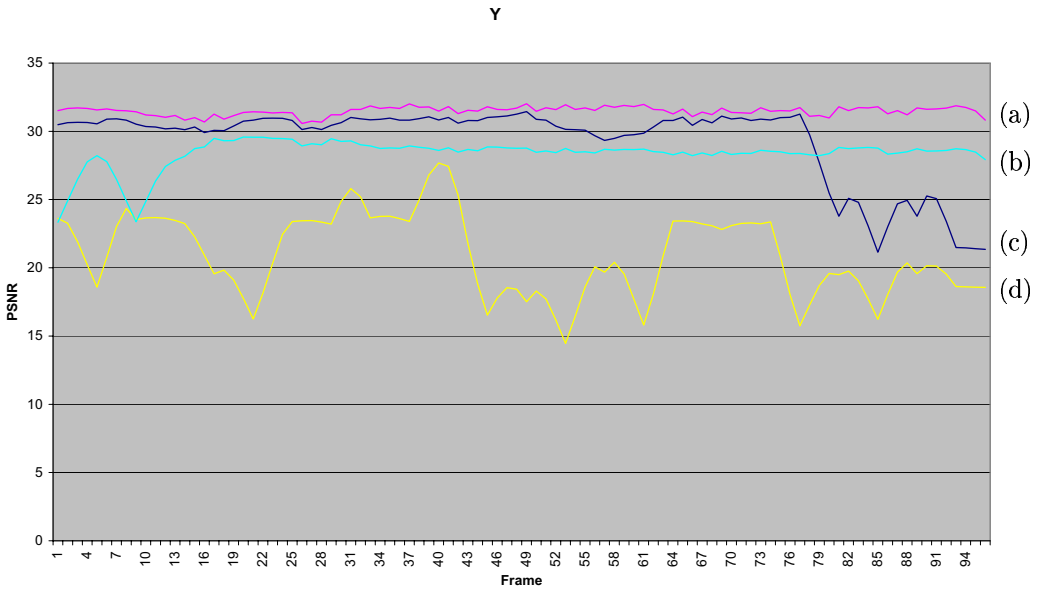

Figure 4.17: Mean PSNR's of each frame for Y component for QCIF *silent* color video compressed to 200 kbits/s (30 frames/s). Transmission over UMTS Pedestrian channel at 0.01 ber; PINK (a): no noise; DARK BLUE (c): Proposed MDC; YELLOW (d): SDC without channel coding; LIGHT BLUE (b): SDC+TC.

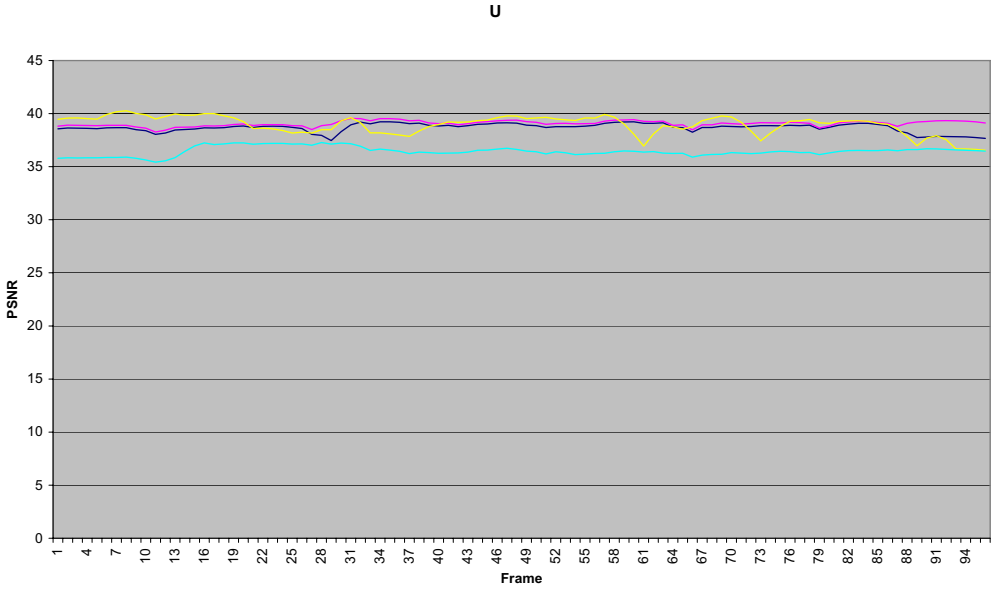

Figure 4.18: Mean PSNR's of each frame for U component for QCIF *silent* color video compressed to 200 kbits/s (30 frames/s). Transmission over UMTS Pedestrian channel at 0.01 ber; PINK: no noise; DARK BLUE: Proposed MDC; YELLOW: SDC without channel coding; LIGHT BLUE: SDC+TC.

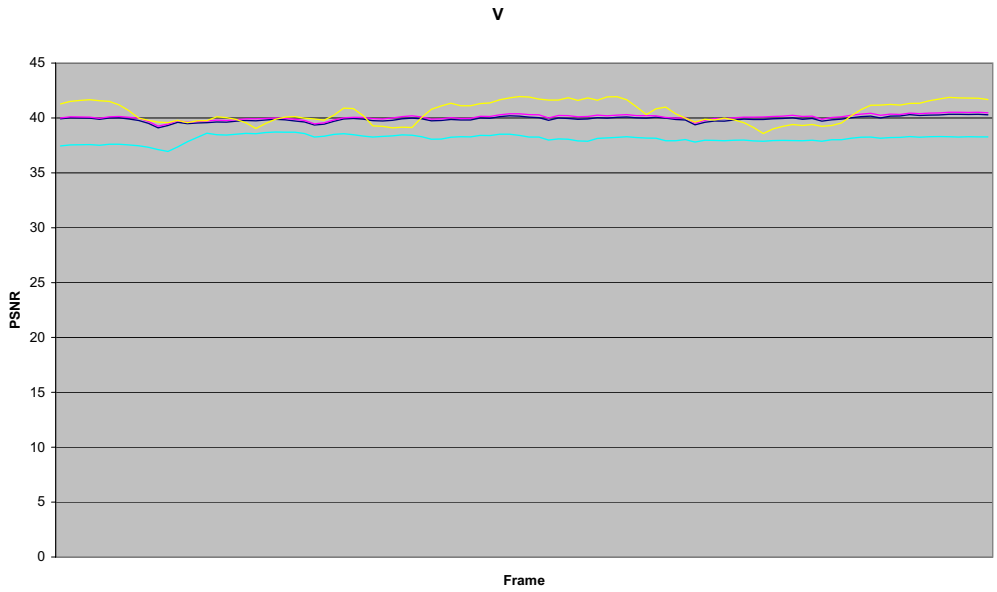

Figure 4.19: Mean PSNR's of each frame for V component for QCIF *silent* color video compressed to 200 kbits/s (30 frames/s). Transmission over UMTS Pedestrian channel at 0.01 ber; PINK: no noise; DARK BLUE: Proposed MDC; YELLOW: SDC without channel coding; LIGHT BLUE: SDC+TC.

Figure 4.20: QCIF *silent* color video compressed at 200 kbits/s (30 frames/s). Transmission over UMTS Indoor channel at 0.01 ber. Using the SDC without channel coding.

Figure 4.21: QCIF *silent* color video compressed at 200 kbits/s (30 frames/s). Transmission over UMTS Indoor channel at 0.01 ber. Using the SDC+TC.

Figure 4.22: QCIF *silent* color video compressed at 200 kbits/s (30 frames/s). Transmission over UMTS Indoor channel at 0.01 ber. Using the proposed MDC.

Figure 4.23: QCIF *silent* color video compressed at 200 kbits/s (30 frames/s). Transmission over UMTS Pedestrian channel at 0.01 ber. Using the SDC without channel coding.

Figure 4.24: QCIF *silent* color video compressed at 200 kbits/s (30 frames/s). Transmission over UMTS Pedestrian channel at 0.01 ber. Using the SDC+TC.

Figure 4.25: QCIF *silent* color video compressed at 200 kbits/s (30 frames/s). Transmission over UMTS Pedestrian channel at 0.01 ber. Using the proposed MDC.

Figure 4.26: QCIF *silent* color video compressed at 200 kbits/s (30 frames/s). Transmission over UMTS Pedestrian channel at 0.001 ber. Using SDC without channel coding.

Figure 4.27: QCIF *silent* color video compressed at 200 kbits/s (30 frames/s). Transmission over UMTS Pedestrian channel at 0.001 ber. Using SDC+TC.

Figure 4.28: QCIF *silent* color video compressed at 200 kbits/s (30 frames/s). Transmission over UMTS Pedestrian channel at 0.001 ber. Using the proposed MDC.

4.7.4.3 Internet channel simulations

Previous studies [198] show that a first-order Markov chain, such as the two state Markov model by Gilbert [55] and Elliot [42], can provide a good approximation in modeling the error process at the packet level.

Based in these studies we use the Internet simulator presented below. We also use this studies to compute the redundancy parameter of the MDBA for Internet channels. The redundancy parameter is in this case computed using the channel capacity presented in equation (4.14).

The simulations were performed for 10 seconds of QCIF *foreman* color video. This video was compressed to 200 kbits/s (30 frames/s).

To make comparisons, we present the results obtained with the proposed MDC whith and without being subject to noise. We present also some results obtained for the transmission of a singular description coder (SDC) with a similar codec. We also use an SDC with a Turbo Coder (SDC+TC) with a bit rate of 200kbits/s including the channel rate.

All channel simulations were performed 5 times. Visual results of video presents always the frames 10, 20, 30, 40, 50, ... of the video. The Y component is the most sensitive to noise. We only present results for this component.

Internet simulator

We use a $K = 2$ state Markov model to simulate the Internet channel, as suggested by [16]. The $K = 2$ state Markov model is represented in figure 4.29. The two states are represented by: "G" for "Good", where all packets are perfectly received, and "B" for "Bad", where all of them are lost. With this model, the global rate loss is given by

$$P(B) = \frac{pgb}{pgb - pbb + 1} \tag{4.13}$$

where $pgb = P(B|G)$ is the probability to move from "G" to "B" state, and pbb the probability to stay in the "B" state.

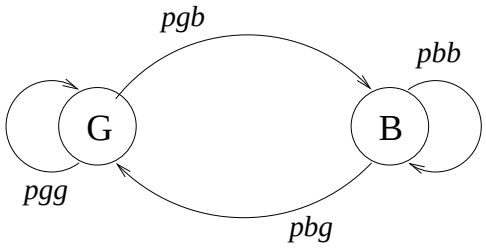

Figure 4.29: Model of the Internet channel. G stands for "Good" state and B for "Bad" state.

In order to produce realistic losses, we choose $pgb = 0.11$ and $pbb = 0.18$ for the first group of simulations and $pgb = 0.05$ and $pbb = 0.05$ for the second group of simulations [16].

Channel capacity

To compute the channel capacity we use [75]. In this article the authors compute $E[C(k)]$ (expected channel rate at time k) as presented in equation 4.14.

$$E[C(k)] = \bar{C} \times \pi_0(k). \tag{4.14}$$

Consider s_0 and s_1 the two states. s_0 represents the "Good" state and s_1 represents "Bad" state. $\pi_0(k)$ is the probability for the channel to stay in state s_0 at time k. $\pi_0(k)$ is computed using the equation presented in 4.15. $\bar{C}$ is the packet size in bits.

$$\pi(k) = [\pi_0(k), \pi_1(k)] = [\pi_0(t), \pi_1(t)].\mathbf{P}^{k-t}. \tag{4.15}$$

Assume that the observed channel state at time t is $S(t)$, then the initial state probability at time t can be written as

$$\pi_n(t) = \begin{cases} 1, & \text{when } S(t) = s_n; \\ 0, & \text{otherwise.} \end{cases}, \forall n \in \{0, 1\}. \tag{4.16}$$

$\mathbf{P}$ is the transition probability matrix. Using [16] for the first group of simulations the transition probability matrix **P1** presented in 4.17 is used. For the second group of simulations the transition probability matrix **P2** presented 4.18 is used.

$$\mathbf{P1} = \begin{bmatrix} 0.89 & 0.11 \\ 0.82 & 0.18 \end{bmatrix} \tag{4.17}$$

$$\mathbf{P2} = \begin{bmatrix} 0.771 & 0.05 \\ 0.771 & 0.05 \end{bmatrix} \tag{4.18}$$

Figures 4.30 and 4.31 present the mean PSNR of different frames for Y for the transmission over Internet. The packet loss is 10% unconditional loss probability for the probe packet and 18% of conditional probe loss probability in the first case and 5% for both in the second. We can see that when comparing with the SDC case we have a gain of 4 dB for the first case and more than 5 dB for the second one.

We perform some simulations with an SDC+TC. However the results were worst than the SDC without channel coding since the SDC+TC could not recover from packet losses.

We show some visual results in figures 4.32, 4.33. Figure 4.32 for the SDC case and figure 4.33 for the proposed MDC. We can conclude that with the proposed system we have a good quality since the average increase in gain is about 5 dB for the studied sequences.

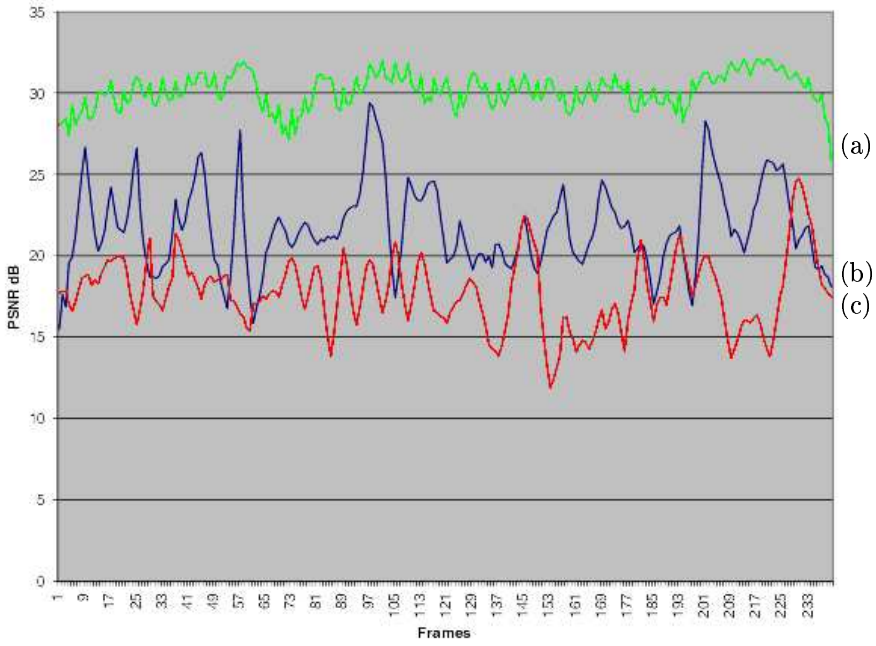

Figure 4.30: Mean PSNR's of each frame for Y component of QCIF *Foreman* color video compressed at 200 Kbps (30 frames/s). Transmission over Internet channel suffering from 10.89% packet loss. GREEN (a): Without retransmission (0% packet loss); Average Y-PSNR : 30.16 dB. BLUE (b): Proposed MDC; Average Y-PSNR: 21.78 dB. RED (c): Using SDC; Average Y-PSNR: 17.62 dB.

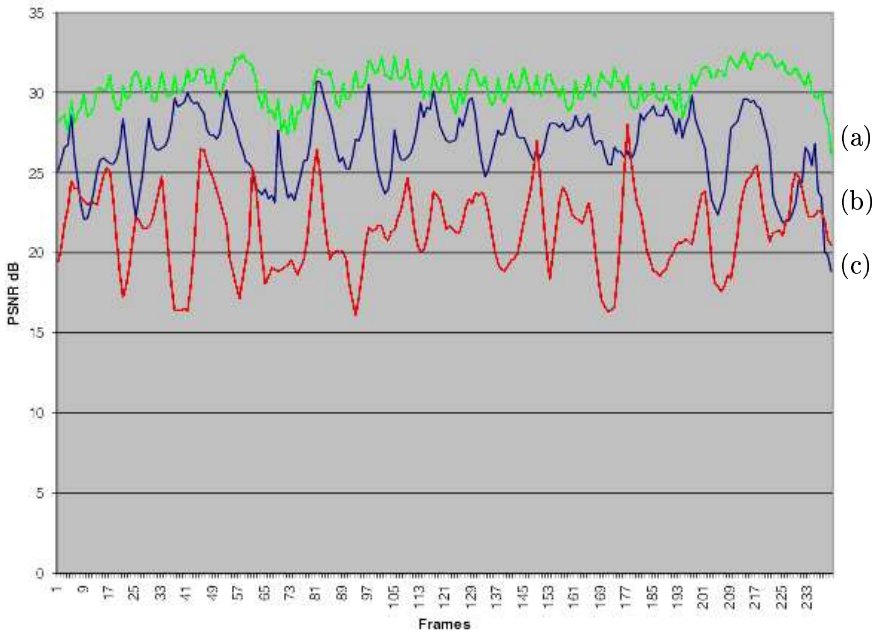

Figure 4.31: Mean PSNR's of each frame for Y component of QCIF *Foreman* color video compressed at 200 Kbps (30 frames/s). Transmission over Internet channel suffering from 5% packet loss. GREEN (a): Without retransmission (0% packet loss); Average Y-PSNR : 30.40 dB. BLUE (b): Proposed MDC; Average Y-PSNR: 26.74 dB. RED (c): Using SDC; Average Y-PSNR: 21.35 dB.

Figure 4.32: QCIF *Foreman* color video compressed at 200 Kbps (30 frames/s). Transmission over an Internet simulator suffering from 5 % packet loss. Using SDC. Y-PSNR: 20.50 dB; U-PSNR: 38.94 dB; V-PSNR: 40.62 dB.

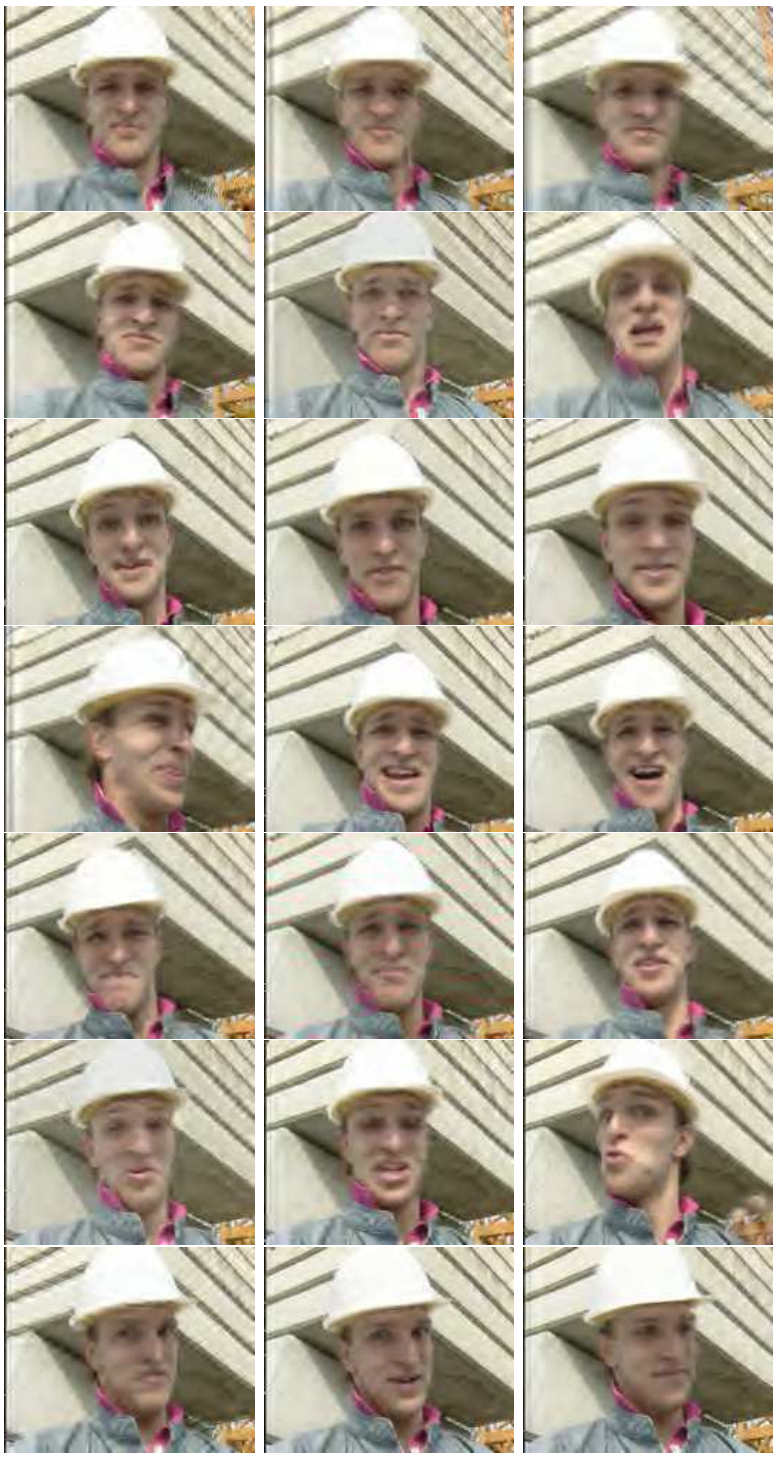

Figure 4.33: QCIF *Foreman* color video compressed at 200 Kbps (30 frames/s). Transmission over an Internet simulator suffering from 5 % packet loss. Using the proposed MDC. Y-PSNR: 26.76 dB; U-PSNR: 38.82 dB; V-PSNR: 40.52 dB.

4.8 Conclusion

In this chapter we proposed a method to estimate the amount of source redundancy dispatched between the different channels based on the channel state and the a priori channel model.

The proposed MDC uses the MDBA presented in chapter 3. The MDBA is based on DWT. One advantage of using the DWT is that we can extend the presented coder to video by adding to the 2D DWT a 1D DWT in the time direction. Furthermore there exist efficient 3D scan-based DWT and 3D scan-based motion compensated lifting DWT for video coding [120, 121, 122, 192, 168, 103, 5, 18, 4] that are well suited for real time applications and provide encouraging results compared with MPEG. Scan-based wavelet transform algorithms are meant to progressively compute the DWT to get low memory implementations of wavelet transforms.

The 3D scan-based DWT transform allows us to develop a stripe-based MDC. In this way we can adapt the redundancy parameter, presented in section 4.3, to the channel state while coding. Thus, the proposed method automatically adapts the coding process to time varying states. The proposed MDBA is then suitable for video transmission over time-varying channels as can be seen by presented simulations.

We conclude that the proposed MDC is a simple alternative for real time transmissions where methods that use error control schemes such as FEC or ARQ are not suitable for delay reasons.

With MDC, a long burst error or even the loss of an entire description does not have a catastrophic effect. Thus, one could even use fewer error control bits for each substream.

The work in this chapter resulted in the following publications [131, 133, 134]. These papers present the adaptation of the proposed MDBA for wired [131] or wireless [133] communications. These results together yield the journal publication [134].

Chapter 5

Multiple description coding for quincunx images. Application to satellite transmission

In this chapter we propose a method of joint source-channel coding for quincunx images. The quincunx arrangement is a way to improve image resolution by combining a pair of CCD linear arrays. The proposed MDC method uses the satellite channel characteristics when performing the source-channel coding. To compare with the proposed method we use the MDBA proposed in previous chapters and adapted to satellite models. We name it in the following as the standard MDC method.

The chapter is organized as follow. Section 5.2 presents the idea beyond quincunx images. Section 5.3 introduces the general MDC scheme. Method I and II are presented in sections 5.4 and 5.5, respectively. Results are presented in section 5.6 and we conclude in section 5.7.

5.1 Introduction

A way to improve image resolution is to combine a pair of CCD linear arrays in a quincunx arrangement. Because each CCD array yields a classical image according to a square grid the systems using such acquisition model are tempted to treat each image isolated, disregarding the high redundancy between them.

We propose a method of joint source-channel coding that take into account the redundancy between the two images in source channel coding. The proposed method uses the satellite channel characteristics when performing the source-channel coding.

The different dyadic images are used to generate the two different descriptions in a MDC scheme and the difference between these two images is joined to both descriptions. This results in a highly robust scheme.

We also present how to use the MDBA presented in previous chapters adapted for quincunx images. This method (we will call in the suite standard MDC method) processes the quincunx sampled image with a well-suited transform to reduce the redundancies. After this step, we use the MDBA method proposed in chapters 3 and 4 adapted here for the case of satellite models.

We present comparisons between the proposed method and the standard MDC method using the MDBA.

Using channel information and source redundancy when designing joint source channel coders results in a robust and efficient compression scheme as we will see in the experimental section.

5.2 Quincunx images

The increasing demand of satellite images (for regional planning, plane cartography and restitution of the relief, ecological monitoring, follow-up of the vegetation, etc...) justify the continuous efforts in order to improve the image quality provided. A way to improve image resolution is to combine a pair of CCD linear arrays in a quincunx arrangement. For instance, the earth observation satellite of CNES, SPOT5[91], provides a quincunx sampling image by using two different CCD linear arrays, shifted each other by 0.5 pixel in the direction of linear arrays, and $n + 0.5$ pixels ($n \in \mathbb{N}$) in the satellite motion direction (see figures 5.1 and 5.2). The emergence of such sampling techniques is due to the Modulation Transfer Function (MTF) of satellites equipped with CCD instruments. This MTF corresponds roughly to a low pass filter and has a frequency support close to the quincunx one [92]. The double linear arrays make a denser sampling grid with an optimal frequency support for this kind of acquisition scheme.

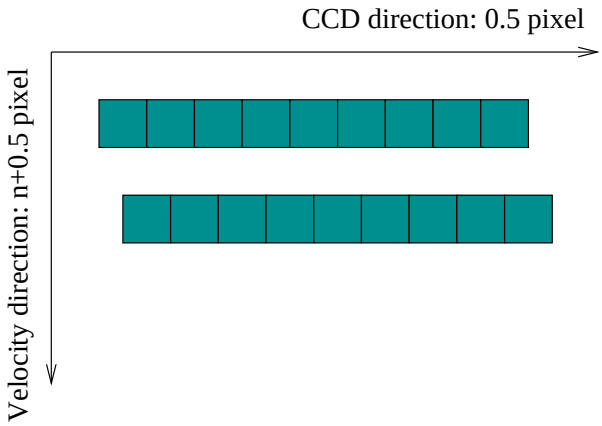

Figure 5.1: Representation of the two CCD linear arrays of a SPOT5 type acquisition system.

Each CCD linear array generates an image sampled on a square grid. This is the principal reason why models using such acquisition system process independently each of the two images. Traditionally, this kind of scheme performs Forward Error Correction (FEC) for each image independently to combat channel failures. For instance, SPOT5 uses Reed-Solomon channel codes more precisely RS(239,255) with interleaving. The drawback in this kind of model is that the dependencies between the pixels of two images are not taken into account neither in source nor in channel coding [92].

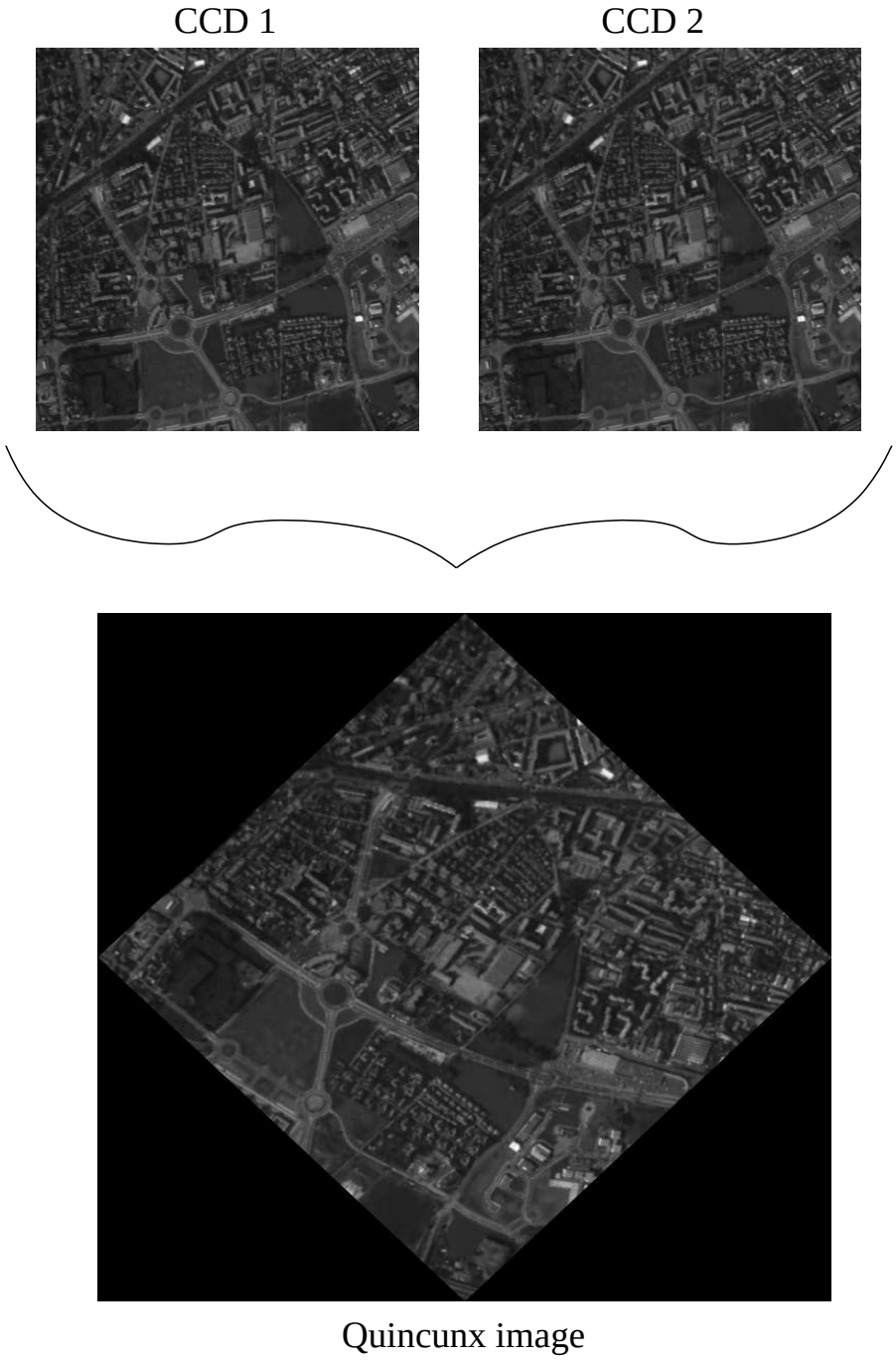

Figure 5.2: Combination of a pair of CCD linear arrays in a quincunx arrangement

We propose to perform joint source-channel coding to get the best image quality after transmission over satellite channel. Taking into account the redundancy between the two CCD arrays when performing joint source-channel coding results in a good trade off rate-quality-robustness.

5.3 General MDC scheme

The proposed method and the standard MDC uses MD schemes based on the Discrete Wavelet Transform (DWT) and an efficient bit allocation technique. Our goal is to find an optimal trade-off between efficient compression and robustness to errors due to communications using unreliable channels. For that, we propose to control automatically the amount of redundancy dispatched on the different descriptions by taking into account the satellite channel model and state (see figure 3.1 in chapter 3). The use of the Scan-Based DWT transform presented in [122, 123] allows the development of a stripe-based MDC and so, to use different redundancies to take into account changes in channels state while coding (see section 4.7 in chapter 4).

The two methods take into account the dependencies between the pixels of the two CCD arrays and use the noise characteristics, to be adapted to the satellite channel model, when performing the source-channel coding.

In Standard MDC approach, to the quincunx sampled image is applied a suited transform to reduce the redundancies. After this step we use the MDBA presented in chapters 3 and 4.

In the proposed model we use the redundancy of the two CCD arrays to find a robust scheme. More precisely, the different dyadic images are used to generate the two different descriptions in a MDC scheme. The difference of these two images is joined to both descriptions in order to find a highly robust scheme.

The objective of both methods is to find, for a given redundancy between the descriptions, which combination of scalar quantizers across the various wavelet coefficients subbands will produce the minimum total central distortion while satisfying the side bit rate constraint $R_1 = R_2 = R_T/2$. This allocation problem is a constrained problem which can be solved by introducing the Lagrange operators. For the standard method it was presented in previous chapter 3. The different initial point of the two methods (see figures 5.4 and 5.5) results in different expressions of the Lagrangian functional. Therefore, we expose each method in a dedicated section. For the standard method we only present the general scheme in section 5.4 since the MDBA was detailed in previous chapter 3. The new method will be detailed in section 5.5.

5.4 MDC for transmission of quincunx images - Standard MDC method

In this MDC approach, to the quincunx sampled image is applied a suited transform to reduce the redundancies. After this step we use the MDBA presented in chapters 3 and 4. We adapt in this case the r_N parameter to the satellite channel as will be specified in section 5.6.2. Here we only detail how the transform is performed in the quincunx sampled image.

The optimal way to reduce the redundancies is to process the quincunx sampled image with a well-suited transform. In the case of quincunx sampled images, we start from a semi-level of resolution and we must have a quincunx compressor, in order to avoid processing the two images separately. The used transform is a quincunx lifting scheme [57, 58, 59, 60] and the quincunx multiresolution is defined by adding an intermediate half resolution (see figure 5.3). The difference in resolution between two successive image approximations is equal to a factor $\sqrt{2}$ for the quincunx case [46] and a factor 2 for the common separable case.

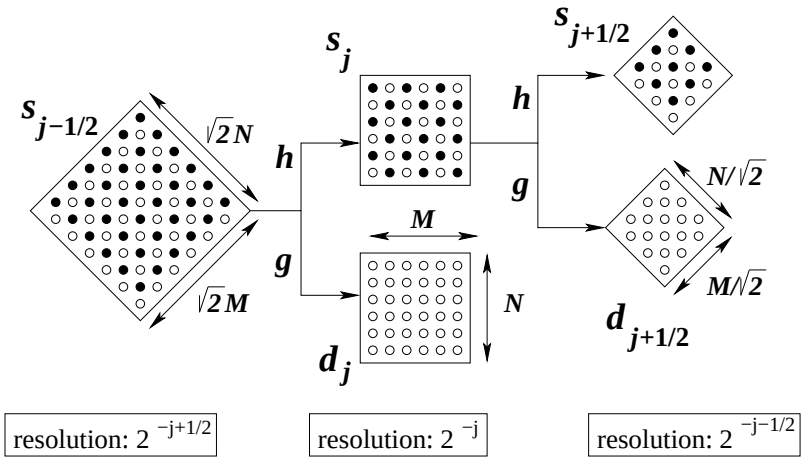

Figure 5.3: Quincunx multiresolution analysis.

Meyer showed that only one wavelet is necessary for multiresolution analysis and not three like in the separable case [104]. Thus, wavelet transform decomposes the signal into two subbands and not four. The factor of resolution differs from the bidimensional separable case. For a Mallat decomposition this factor is equal to 2, and for the quincunx case it is identified by the function L, which is a linear transform checking $L(x, y) = (x + y, x - y)$. One can observe that $L \circ L = 2Id$.

Taking the previous considerations, in this method we perform the quincunx sampled image with the transform presented in [58, 59, 60]. The resulting wavelet coefficients are used in the MDBA presented in chapter 3. Figure 5.4 presents the scheme of the standard MDC for quincunx images.

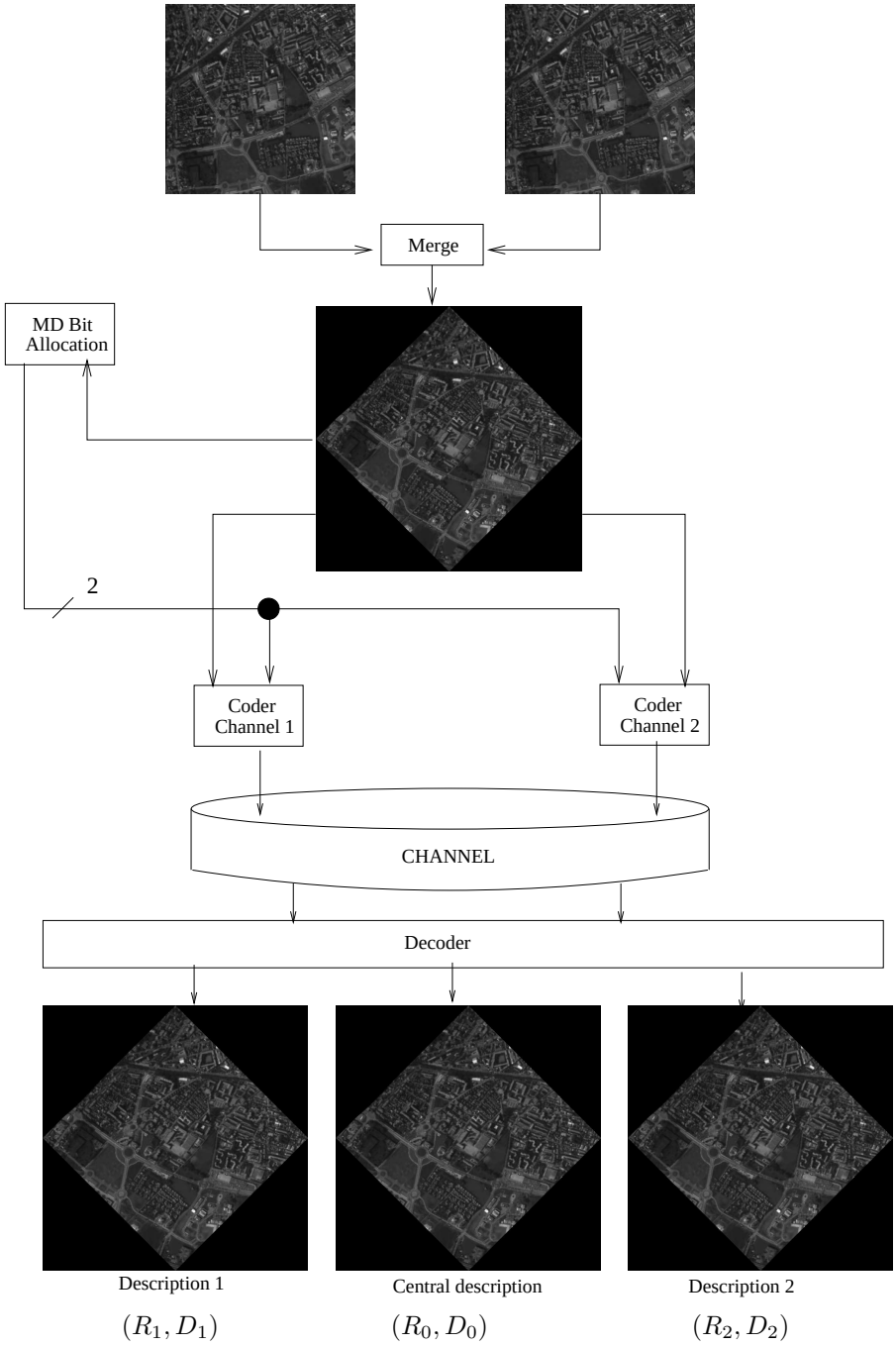

Figure 5.4: MDC for quincunx images. Standard MDC method.

5.5 MDC for transmission of quincunx images. Proposed method

In this model we use the redundancy of the two CCD arrays to reach robustness. More precisely, the different dyadic images are used to generate the two different descriptions in a MDC scheme and the difference of these two images is joined to both descriptions. Let us call $P1$ and $P2$ the two different CCD arrays, and ϵ their difference. As can be seen in figure 5.5, one description contains $P1$ and ϵ, while the other contains $P2$ and ϵ.

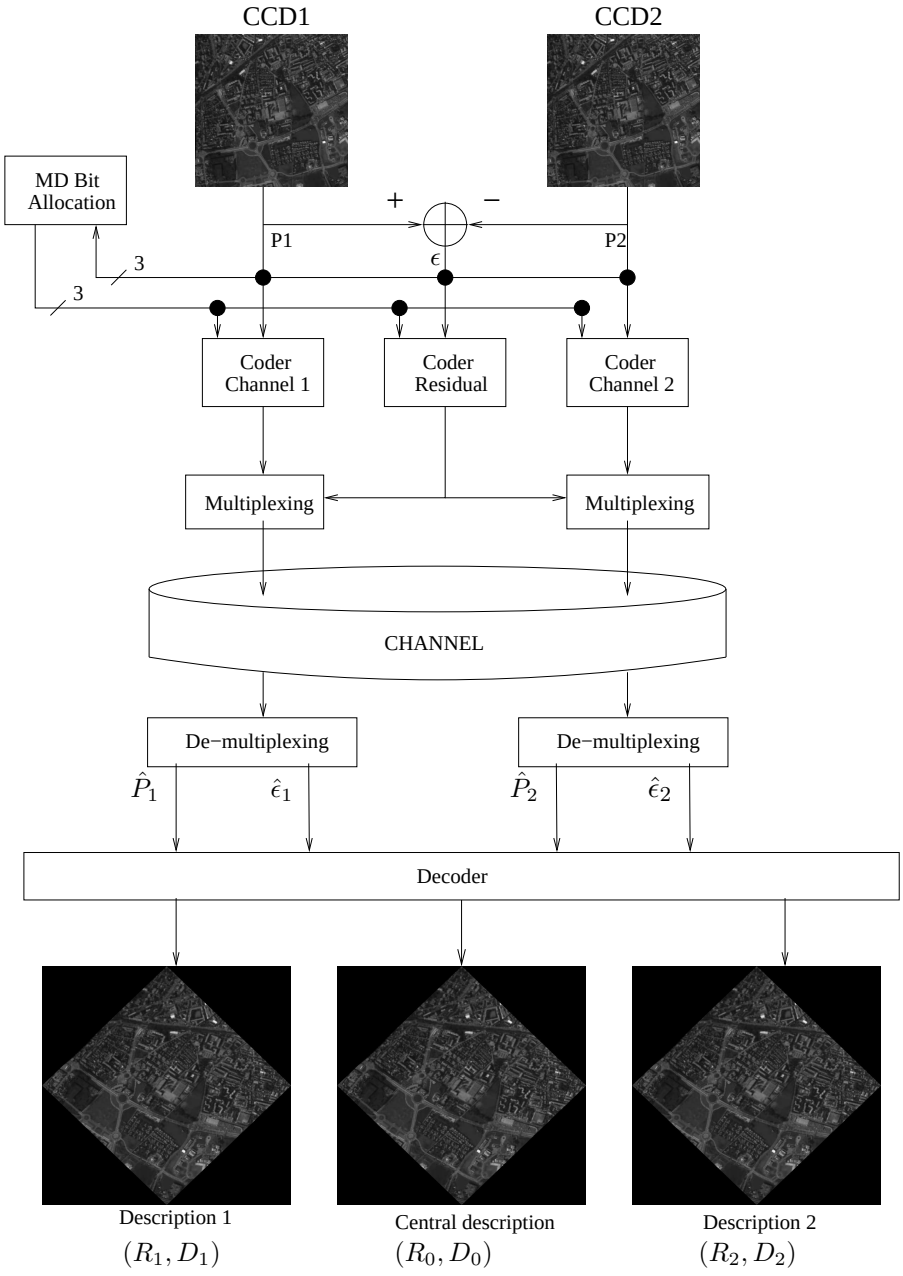

Figure 5.5: MDC for quincunx images. Proposed MDC method.

In this model we have always two different ways to recover an information: ϵ is coded twice; $P1$ can be recovered from description 1 or from $P2 + \epsilon$, and finally, $P2$ can be recovered from description 2 or from $P1 - \epsilon$. This results in a highly robust scheme.

In this model, the redundancy parameter is associated to the residual image ϵ. This parameter is computed in the same way as in the standard method, i.e., using r_N equation as proposed in chapter 4. In this approach we propose to apply the redundancy parameter to the residual image, because the use of the residual image by the decoder is dependent of the transmission losses.

5.5.1 Proposed bit allocation for MDC

5.5.1.1 Introduction of Lagrange operators

The problem is to find, for a given redundancy between the descriptions, which combination of scalar quantizers across the various wavelet coefficients subbands will produce the minimum total central distortion while satisfying the side bit rate constraint.

For a system considering N subbands of a wavelet decomposition, we intend to minimize the central distortion D_0 for a total bit rate R_T. Thus, the purpose of our bit allocation for MD scheme is to determine the optimal sets of quantization steps $\{q_{i,1}, i = 1, ..., N\}$, $\{q_{i,2}, i = 1, ..., N\}$ and $\{q_{i,\epsilon}, i = 1, ..., N\}$. The parameter R_T is given for the bit allocation.

Using definition 2 and proposition 1 presented in chapter 3 section 3.3, the bit allocation problem can be resumed as:

$$(P) \begin{cases} \textbf{min } D_0\left(\{\tilde{q}_{i,1}, \tilde{q}_{i,2}, \tilde{q}_{i,\epsilon}\}\right) \\ \\ \textbf{Constraints } f(R_1) \leqslant 0 \text{ and } f(R_2) \leqslant 0 \end{cases} \tag{5.1}$$

where, $R_j, j = 1, 2$ is defined in equation (5.2).

$$R_j = \sum_{i=1}^{N} a_i \left(R_{i,j}\left(\tilde{q}_{i,j}\right) + R_{i,\epsilon}\left(\tilde{q}_{i,\epsilon}\right)\right), \qquad j = 1, 2. \tag{5.2}$$

The parameter a_i in equation (5.2) is the size of the subband (i) divided by the size of the sequence and $R_{i,k}(q_{i,k}), k = 1, 2, \epsilon$, is the output bit rate in bits per sample for the ith subband.

The allocation problem (5.1) is a constrained problem which can be solved by introducing the Lagrange operators. The Lagrangian functional for the constrained optimization problem is given by equation (5.3).

$$J\left(\{q_{i,1}, q_{i,2}, q_{i,\epsilon}\}\right) = D_0 + \sum_{j=1}^{2} \lambda_j \left(f(R_j)\right) \tag{5.3}$$

For a source with GG distribution, D_0 in equation (5.3) can be written as

$$D_0 = \sum_{i=1}^{N} \Delta_i w_i \sigma_{i,0}^2 D_{i,0}\left(\tilde{q}_{i,1}, \tilde{q}_{i,2}, \tilde{q}_{i,\epsilon}\right), \tag{5.4}$$

The $\sigma_{i,0}^2 D_{i,0}(\tilde{q}_{i,1}, \tilde{q}_{i,2})$ is the central Mean Square Error (MSE) for the ith subband in the case of a Generalized Gaussian distribution. The parameter Δ_i is an optional weight for frequency selection and w_i the weights used to take into account the nonorthogonality of the filter bank [17]. The expected central distortion is estimated based on the channel states and the a priori channel models as we will see in the next section.

5.5.1.2 Central distortion modeling

Recall that the central distortion is the distortion of the decoded image when using both descriptions at decoding. One description contains $P1$ and ϵ, while the other contains $P2$ and ϵ.

When the decoder receives both descriptions, if the subbands are noiseless, the central decoder uses $P1$ and $P2$. Since the redundancy parameter is associated to ϵ it can be set to 0.

In the general case we have to take into account channel noise. In case of noise in a subband of $P1$ the decoder uses the same subband of $P2$ and of ϵ to estimate it. Inversely, in case of noise in a subband of $P2$ the decoder uses the same subband of $P1$ and of ϵ to estimate it. Thus, the importance of ϵ grows with the level of noise. In consequence, the redundancy parameter should increase when the BER increases.

Hence, we propose to calculate the central distortion of the decoded image as

$$\begin{aligned} D_{i,0}\left(\tilde{q}_{i,1}, \tilde{q}_{i,2}, \tilde{q}_{i,\epsilon}\right) &= \frac{1}{\sigma_{i,0}^2}\frac{1}{1+r_N}\left[\sigma_{i,1}^2 D_{i,1}\left(\tilde{q}_{i,1}\right)\right. \\ &\quad + r_N\left(\sigma_{i,2}^2 D_{i,2}\left(\tilde{q}_{i,2}\right) + \sigma_{i,\epsilon}^2 D_{i,\epsilon}\left(\tilde{q}_{i,\epsilon}\right)\right) \\ &\quad \left. + \sigma_{i,2}^2 D_{i,2}\left(\tilde{q}_{i,2}\right) + r_N\left(\sigma_{i,1}^2 D_{i,1}\left(\tilde{q}_{i,1}\right) + \sigma_{i,\epsilon}^2 D_{i,\epsilon}\left(\tilde{q}_{i,\epsilon}\right)\right)\right] \\ &= \frac{1}{\sigma_{i,0}^2}\left[\left(\sigma_{i,1}^2 D_{i,1}\left(\tilde{q}_{i,1}\right) + \sigma_{i,2}^2 D_{i,2}\left(\tilde{q}_{i,2}\right)\right)\right. \\ &\quad \left. + \sigma_{i,\epsilon}^2 \frac{2r_N}{1+r_N} D_{i,\epsilon}\left(\tilde{q}_{i,\epsilon}\right)\right] \end{aligned} \tag{5.5}$$

where $\sigma_{i,1}^2 D_{i,1}, \sigma_{i,2}^2 D_{i,2}$ and $\sigma_{i,\epsilon}^2 D_{i,\epsilon}$ are the MSE of subband i of $P1$, $P2$ and ϵ, respectively.

We have $D_0 \leqslant D_j$, $j = 1, 2$ for description $j = 1$ and $j = 2$, with D_j defined in equation (5.6).

$$D_j = \sum_{i=1}^{N}\left[\sigma_{i,j}^2 D_{i,j} + \left(\sigma_{i,j}^2 D_{i,j} + \sigma_{i,\epsilon}^2 D_{i,\epsilon}\right)\right], \qquad j = 1, 2. \tag{5.6}$$

where $\sigma_{i,j}^2 D_{i,j}$ is the MSE of subband i of P_j. The $\sigma_{i,\epsilon}^2 D_{i,\epsilon}$ is the MSE of subband i of ϵ, thus $(\sigma_{i,j}^2 D_{i,j} + \sigma_{i,\epsilon}^2 D_{i,\epsilon})$ is the MSE associated with the estimation of P_2 by description 1, when $j = 1$ or the estimation of P_1 by description 2, when $j = 2$.

As we said above, the redundancy parameter is computed in the same way as in the standard method, i.e., using r_N equation (4.5) as proposed in chapter

4. Since this parameter depends on the channel model and state (BER), in this chapter it will be adapted to satellite models that are presented in section 5.6.2.

5.5.1.3 Bit rate constraint

Finding the best bit allocation can be stated as a constrained optimization problem, where the R_i have to minimize the central distortion subject to a total bit rate constraint. The total bit rate constraint ($f(R_j) \leqslant 0, j = 1,2$ in equation (5.1)) has to be defined for each description. For the different descriptions $j = 1, 2$, we write condition 2 as a constraint Q_j given in equation (5.7).

$$Q_j = \left(\sum_{i=1}^{N} a_i \left(R_{i,j} \left(\tilde{q}_{i,j} \right) + R_{i,\epsilon} \left(\tilde{q}_{i,\epsilon} \right) \right) - R_T/2 \right), \qquad j = 1,2. \tag{5.7}$$

In equation (5.7) $R_{i,k}(\tilde{q}_{i,k}), k = 1, 2, \epsilon$, is the bit rate in bits per sample for the ith subband.

5.5.1.4 Solution of the problem

Considering (5.4), (5.5) and the constraint (5.7) the Lagrangian functional (5.3) can be rewritten as presented in equation 5.8.

$$J\left(\{q_{i,1}, q_{i,2}, q_{i,k}\}\right) = \sum_{i=1}^{N} \Delta_i w_i \sigma_{i,0}^2 D_{i,0} \left(\tilde{q}_{i,1}, \tilde{q}_{i,2}, \tilde{q}_{i,\epsilon} \right) + \sum_{j=1}^{2} \lambda_j Q_j \tag{5.8}$$

Solution of (5.8) is obtained when

$$\begin{cases} \dfrac{\partial J\left(\{q_{i,1}, q_{i,2}, q_{i,\epsilon}\}\right)}{\partial q_{i,1}} = 0, & (a) \\ \dfrac{\partial J\left(\{q_{i,1}, q_{i,2}, q_{i,\epsilon}\}\right)}{\partial q_{i,2}} = 0, & (b) \\ \dfrac{\partial J\left(\{q_{i,1}, q_{i,2}, q_{i,\epsilon}\}\right)}{\partial q_{i,\epsilon}} = 0, & (c) \\ \dfrac{\partial J\left(\{q_{i,1}, q_{i,2}, q_{i,\epsilon}\}\right)}{\partial \lambda_1} = 0, & (d) \\ \dfrac{\partial J\left(\{q_{i,1}, q_{i,2}, q_{i,\epsilon}\}\right)}{\partial \lambda_2} = 0. & (e) \end{cases} \tag{5.9}$$

The derivative of Lagrangian functional (5.8) with respect to λ_j, $j = 1, 2$ (equation (5.9 d) and equation (5.9 e)) is presented in equation (5.10).

$$\frac{\partial J\left(\{q_{i,1},q_{i,2},q_{i,\epsilon}\}\right)}{\partial \lambda_j}=0 \Longleftrightarrow \sum_{i=1}^{N} a_i\left(R_{i,j}\left(\tilde{q}_{i,j}\right)+R_{i,\epsilon}\left(\tilde{q}_{i,\epsilon}\right)\right)-R_T/2=0 \quad (5.10)$$

In the following we detail the derivative of Lagrangian functional (5.8) with respect to $q_{i,k}, k=1,2,\epsilon$ (equations (5.9 a), (5.9 b) and (5.9 c)).

$$\begin{aligned}\frac{\partial J\left(\{q_{i,1},q_{i,2},q_{i,\epsilon}\}\right)}{\partial q_{i,k}} &= \Delta_i w_i \sigma_{i,0}^2 \frac{\partial}{\partial q_{i,k}} D_{i,0}\left(\tilde{q}_{i,1},\tilde{q}_{i,2},\tilde{q}_{i,\epsilon}\right)\\ &+ \lambda_1 a_i \frac{\partial}{\partial q_{i,k}}\left(R_{i,1}\left(\tilde{q}_{i,1}\right)+R_{i,\epsilon}\left(\tilde{q}_{i,\epsilon}\right)\right)\\ &+ \lambda_2 a_i \frac{\partial}{\partial q_{i,k}}\left(R_{i,2}\left(\tilde{q}_{i,2}\right)+R_{i,\epsilon}\left(\tilde{q}_{i,\epsilon}\right)\right)=0 \quad (5.11)\end{aligned}$$

According to equation (5.5), $\Delta_i w_i \sigma_{i,0}^2 \frac{\partial}{\partial q_{i,k}} D_{i,0}\left(\tilde{q}_{i,1},\tilde{q}_{i,2},\tilde{q}_{i,\epsilon}\right)$ can be simplified as

$$\Delta_i w_i \sigma_{i,k}^2 C_k \frac{\partial}{\partial q_{i,k}} D_{i,k}(\tilde{q}_{i,k}) \quad (5.12)$$

with,

$$C_k=\begin{cases}1, & \text{if} \quad k=1,2\\ \frac{2r_N}{1+r_N}, & \text{if} \quad k=\epsilon\end{cases} \quad (5.13)$$

In this way, the derivation of the Lagrangian functional (5.11) results in equation (5.14).

$$\begin{aligned}\frac{\partial J\left(\{q_{i,1},q_{i,2},q_{i,\epsilon}\}\right)}{\partial q_{i,k}} &= \Delta_i w_i \sigma_{i,k}^2 C_k \frac{\partial}{\partial q_{i,k}} D_{i,k}(\tilde{q}_{i,k})\\ &+ \lambda_1 a_i \frac{\partial}{\partial q_{i,k}}\left(R_{i,1}\left(\tilde{q}_{i,1}\right)+R_{i,\epsilon}\left(\tilde{q}_{i,\epsilon}\right)\right)\\ &+ \lambda_2 a_i \frac{\partial}{\partial q_{i,k}}\left(R_{i,2}\left(\tilde{q}_{i,2}\right)+R_{i,\epsilon}\left(\tilde{q}_{i,\epsilon}\right)\right)=0 \quad (5.14)\end{aligned}$$

The derivative $\lambda_1 a_i \frac{\partial}{\partial q_{i,k}}\left(R_{i,1}\left(\tilde{q}_{i,1}\right)+R_{i,\epsilon}\left(\tilde{q}_{i,\epsilon}\right)\right)$ in (5.14) can be simplified as

$$\begin{cases}\lambda_1 a_i \frac{\partial}{\partial q_{i,k}}\left(R_{i,1}\left(\tilde{q}_{i,1}\right)\right) & \text{if} \quad k=1\\ 0 & \text{if} \quad k=2\\ \lambda_1 a_i \frac{\partial}{\partial q_{i,k}}\left(R_{i,\epsilon}\left(\tilde{q}_{i,\epsilon}\right)\right) & \text{if} \quad k=\epsilon\end{cases} \quad (5.15)$$

and similarly, the derivative $\lambda_2 a_i \frac{\partial}{\partial q_{i,k}} \left(R_{i,2} \left(\tilde{q}_{i,2} \right) + R_{i,\epsilon} \left(\tilde{q}_{i,\epsilon} \right) \right)$ in (5.14) can be simplified as

$$\begin{cases} 0 & \text{if} \quad k = 1 \\ \lambda_2 a_i \dfrac{\partial}{\partial q_{i,k}} \left(R_{i,2} \left(\tilde{q}_{i,2} \right) \right) & \text{if} \quad k = 2 \\ \lambda_2 a_i \dfrac{\partial}{\partial q_{i,k}} \left(R_{i,\epsilon} \left(\tilde{q}_{i,\epsilon} \right) \right) & \text{if} \quad k = \epsilon \end{cases} \tag{5.16}$$

Thus, the derivation of the Lagrangian functional (5.14) results in equation (5.17).

$$\begin{aligned} \frac{\partial J \left(\{ q_{i,1}, q_{i,2}, q_{i,\epsilon} \} \right)}{\partial q_{i,k}} &= \Delta_i w_i \sigma_{i,k}^2 C_k \frac{\partial}{\partial q_{i,k}} D_{i,k}(\tilde{q}_{i,k}) \\ &+ A_k a_i \frac{\partial}{\partial q_{i,k}} R_{i,k} \left(\tilde{q}_{i,k} \right) = 0 \end{aligned} \tag{5.17}$$

with

$$A_k = \begin{cases} \lambda_k, & \text{if} \quad k = 1, 2 \\ \lambda_1 + \lambda_2, & \text{if} \quad k = \epsilon \end{cases} \tag{5.18}$$

Finally, we can write the derivative (5.17) of the Lagrangian functional as

$$\Delta_i w_i \sigma_{i,k}^2 C_k \frac{\partial}{\partial q_{i,k}} D_{i,k} \left(\tilde{q}_{i,k} \right) + A_k a_i \frac{\partial}{\partial q_{i,k}} R_{i,k} \left(\tilde{q}_{i,k} \right) = 0 \tag{5.19}$$

that can be simplified in

$$\frac{\partial D_{i,k}}{\partial R_{i,k}} \left(\tilde{q}_{i,k} \right) = \frac{-A_k a_i}{\Delta_i w_i \sigma_{i,k}^2 C_k} \tag{5.20}$$

In resume, solution of (5.8), is given by the following system (5.21).

Equations (5.21 a), (5.21 b) and (5.21 c) come from equation (5.20). This equation (5.20) is the derivative of Lagrangian functional (5.8) with respect to $q_{i,k}$, $k = 1, 2, \epsilon$ (equation (5.9 a), equation (5.9 b) and equation (5.9 c)).

Equations (5.21 d) and (5.21 e) come from (5.10) that is the derivative of Lagrangian functional (5.8) with respect to λ_j, $j = 1, 2$ (equation (5.9 d) and equation (5.9 e)).

The system (5.21) is a two channel scheme.

$$\begin{cases} \dfrac{\partial D_{i,1}}{\partial R_{i,1}}\left(\tilde{q}_{i,1}\right) = \dfrac{-A_1 a_i}{\Delta_i w_i \sigma_{i,1}^2 C_1} & (a) \\ \dfrac{\partial D_{i,2}}{\partial R_{i,2}}\left(\tilde{q}_{i,2}\right) = \dfrac{-A_2 a_i}{\Delta_i w_i \sigma_{i,2}^2 C_2} & (b) \\ \dfrac{\partial D_{i,\epsilon}}{\partial R_{i,\epsilon}}\left(\tilde{q}_{i,\epsilon}\right) = \dfrac{-A_\epsilon a_i}{\Delta_i w_i \sigma_{i,\epsilon}^2 C_\epsilon} & (c) \\ \sum_{i=1}^{N} a_i \left(R_{i,1}\left(\tilde{q}_{i,1}\right) + R_{i,\epsilon}\left(\tilde{q}_{i,\epsilon}\right)\right) - R_T/2 = 0 & (d) \\ \sum_{i=1}^{N} a_i \left(R_{i,2}\left(\tilde{q}_{i,2}\right) + R_{i,\epsilon}\left(\tilde{q}_{i,\epsilon}\right)\right) - R_T/2 = 0 & (e) \end{cases} \tag{5.21}$$

Resolution of system (5.21) which has $3 \times N + 2$ equations and $3 \times N + 2$ unknowns gives us the optimal sets of quantization steps $\{q_{i,1}\}$, $\{q_{i,2}\}$ and $\{q_{i,\epsilon}\}$, for a given r_N.

The proposed algorithm is based on modeling of R and D functions as we will show in section 5.5.2.

5.5.2 Algorithm

As can be seen in Fig. 3.4 we compute $R_{i,k}, k = 1, 2, \epsilon$ using the given parameters λ_j, r_N, equation (5.21 a, b, c) and $\frac{\partial D}{\partial R}$ function. If the rate constraint (3.40 d, e) is not verified we recompute the $R_{i,k}$ using a new λ_j. If it is verified we compute: $q_{i,k}$ from $R_{i,k}$ and R function. The algorithm stops and the output of bit allocation gives the optimal quantization steps $q_{i,k}$.

To compute $\frac{\partial D}{\partial R}$, R and D functions we use, as in the MDBA presented in chapter 3, the model based implementation specified in chapter 3 section 3.5.

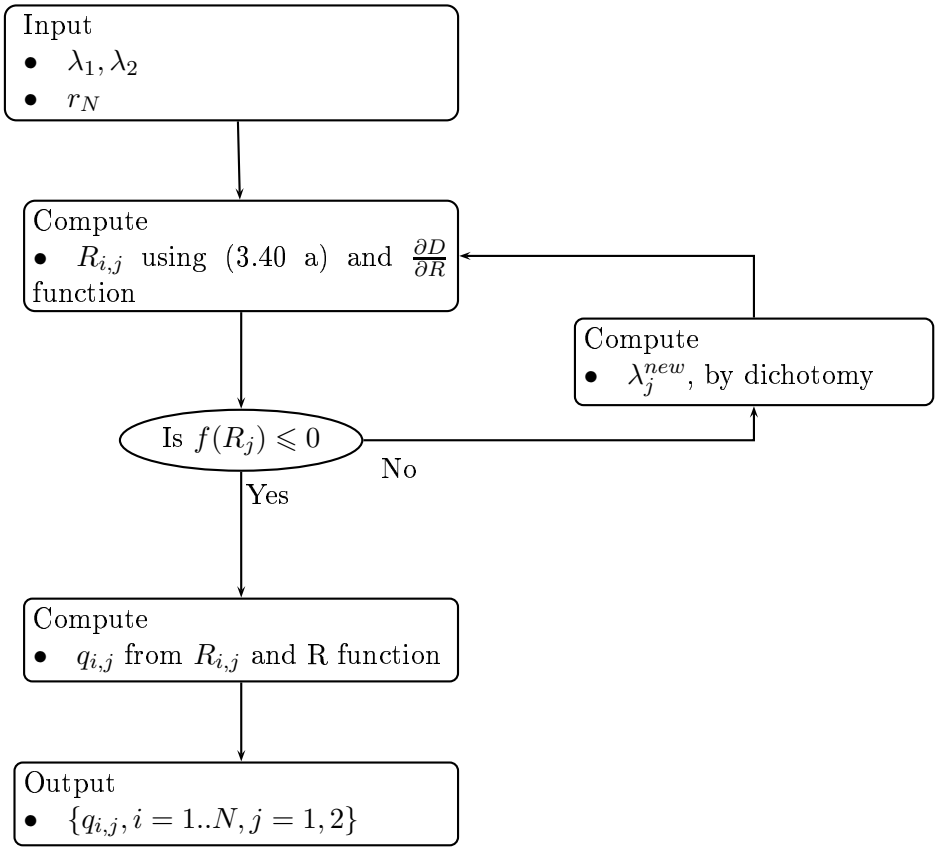

Figure 5.6: Global Bit Allocation Procedure

5.6 Results

5.6.1 Specifications

For spatial decomposition, in the first method, the coder uses (6,2) nonseparable lifting scheme and performs seven levels of decomposition [58, 60]. In the second method, the coder uses 9-7 biorthogonal filter [6] and performs three levels of decomposition. Quincunx transform adds an intermediate resolution level between two successive levels in the separable case, and then allows a twice as accurate multiresolution analysis as the separable one. Thus, the difference of resolution between two successive levels has the value 2 in the bidimensional separable case, and a value $\sqrt{2}$ in the nonseparable case. One can observe that the resolution of the quincunx sampling images is $\sqrt{2}$ times higher than the resolution of CCD1 and CCD2. To obtain the same resolution for all low frequency images, a n level separable decomposition has to correspond to an $2n + 1$ level nonseparable decompositions.

The images are acquired and processed on the fly to generate the wavelet coefficients and the data are stored in memory only until these coefficients have been encoded [122]. The bit allocation procedure is followed by a simple scalar quantization and the encoding of each subband uses context-based arithmetic bit-plane coder [119]. In order to provide synchronization and minimize the error propagation in the case of errors due to satellite communication, each spatio-temporal subband is divided into blocks. Then, arithmetic coding is synchronized on each block. For error detection, we use the Smart Arithmetic Coding method presented in section 4.5 of chapter 4. As satellite channel simulator we use the model proposed by Chee and Sweeney for LEO satellite channels [32], that we present in the following.

5.6.2 Satellite channel model

We compute the LEO satellite communication channel proposed in [32]. They show that for elevation angles of 23° and 52°, the burst statistics can be described by a two good state single error state Frichman model while a three-good state, single error state Frichman model can accurately describe the measured statistics for the rest of the elevation angles. Fritchman's partitioned Markov chain model is a generalization of Gilbert's model partitioned into k error free states and $N-k$ error states [51]. In this model, the interval length distribution between the errors is described by the sum of k exponentials, while the error burst distribution is described by the sum of $N-k$ exponentials. The derived transition probabilities for the three good state Fritchman model can be found in [32]. We show in figure 5.7 the Three-good state, single error state Fritchman model for 40° pass.

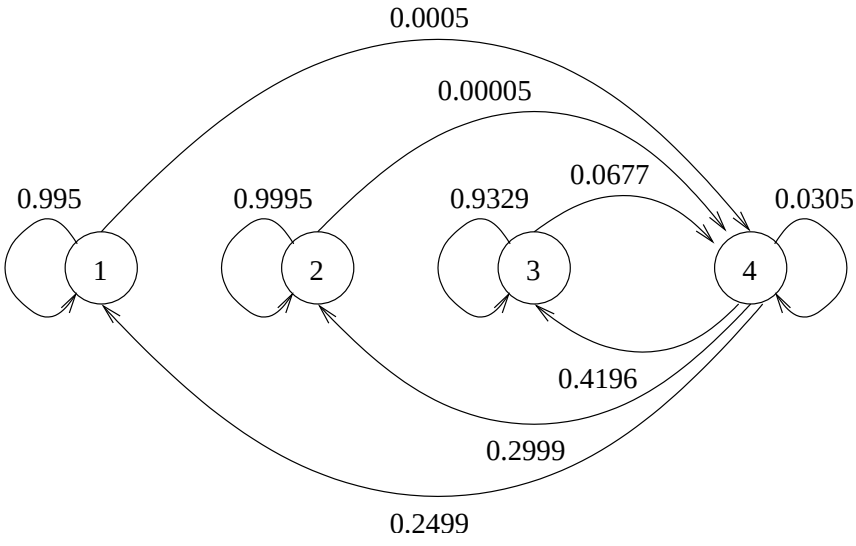

Figure 5.7: Three-good state, single error state Fritchman model for 40° pass.

5.6.2.1 Redundancy parameter for Satellite channel model

We use the studies presented in [76] to compute the redundancy parameter of equation (4.5) for Satellite channels. In these studies the authors present the channel rate estimation for a N-state Markov chain model with one "Good" state and all the others states representing "Bad" states. We present below an adaptation of this study to our case, of three "Good" states and one "Bad" state. The redundancy parameter is in this case computed using the channel capacity resulting from this adaptation.

Channel rate estimation

Assuming that at time t the channel state at time $t-b$, $S(t-b)$, is known.
In a 4-state Markov channel model. The transition probability matrix $\mathbf{P}$ considered will be the one presented in (5.22) when considering a transmission angle of 40° pass, and with transition probability matrix $\mathbf{P}$ presented in (5.23) when considering a transmission angle of 30° pass. We define the state probability $\pi_n(k|S(t-b))$ as the probability that the channel is in state s_n at time k given the channel state observation $S(t-b)$.

$$\mathbf{P} = \begin{bmatrix} 0.995 & 0 & 0 & 0.0005 \\ 0 & 0.9995 & 0 & 0.00005 \\ 0 & 0 & 0.9329 & 0.0677 \\ 0.2499 & 0.2999 & 0.4196 & 0.0305 \end{bmatrix} \tag{5.22}$$

$$\mathbf{P} = \begin{bmatrix} 0.9989 & 0 & 0 & 0.0011 \\ 0 & 0.9999 & 0 & 0.0001 \\ 0 & 0 & 0.9139 & 0.0861 \\ 0.5294 & 0.2299 & 0.2102 & 0.0304 \end{bmatrix} \tag{5.23}$$

A vector of state probabilities can be written as

$$\begin{aligned} \pi(k|S(t-b)) &= [\pi_1(k|S(t-b)), \pi_2(k|S(t-b)), \\ &\quad \pi_3(k|S(t-b)), \pi_4(k|S(t-b))]\,. \end{aligned}$$

The initial state probability $\pi(t-b|S(t-b))$ at time $t-b$ can be set up as

$$\pi_n(t-b|S(t-b)) = \begin{cases} 1, & \text{when } S(t-b) = s_n; \\ 0, & \text{otherwise.} \end{cases}, \forall n \in \{1,2,3,4\}. \tag{5.24}$$

In the Markov model, the state probabilities $\pi(k|S(t-b))$ at time k can be derived from the state probabilities $\pi(k-1|S(t-b))$ at the previous time slot and the transition probability matrix $\mathbf{P}$ as

$$\pi(k|S(t-b)) = \pi(k-1|S(t-b)).\mathbf{P}. \tag{5.25}$$

By recursively using equation (5.25), channel state probabilities at time k, where $k > t-b$, can then be calculated from $\pi(t-b|S(t-b))$ and $\mathbf{P}$ as

$$\pi(k|S(t-b)) = \pi(t-b|S(t-b)).\mathbf{P}^{k-t+b}. \tag{5.26}$$

In our channel model, packets are transmitted correctly ($\bar{C}$ bits are transmitted) when channel is in state s_1, s_2 or s_3, while error occurs (0 bits are transmitted) when channel is in state s_4. Therefore, $\pi_1(k) + \pi_2(k) + \pi_3(k)$ is the probability of correct transmission at time k. The expected channel rate $E[C(k)|S(t-b)]$ given the observation of channel state $S(t-b)$ can be calculated as

$$E[C(k)|S(t-b)] = \bar{C} \times (\pi_1(k) + \pi_2(k) + \pi_3(k)) . \tag{5.27}$$

5.6.3 Simulations

We use a 352×704 quincunx, that is generated using two dyadic 352×352 *Nimes* images, for the proposed method and two 352×352 dyadic *Nimes* images for the standard method. Each test was performed 10 times. Table 4.4 presents the average PSNR obtained by the side decoders and the central decoder, for method I and II at different bit rates and for different transmission angles (different ber). We can conclude from this table that the proposed method is better suited for lower ber, while the standard method is better suited for higher ber. Figures 6 and 7 present one of the simulations, at 2 bpp when transmission at an elevation angle of 40° (0.0005 ber).

	Side PSNR	Central PSNR
2 bpp		
Method I	32.71	40.26
Method II	31.29	38.74
3 bpp		
Method I	33.04	42.27
Method II	31.62	38.82

Table 5.1: PSNR values for *Nimes* image when considering transmission at an elevation angle of 40° (0.0005 ber).

	Side PSNR	Central PSNR
2 bpp		
Method I	33.92	37.76
Method II	31.88	39.48
3 bpp		
Method I	32.77	39.60
Method II	31.84	40.12

Table 5.2: PSNR values for *Nimes* image when considering transmission at an elevation angle of 30° (0.001 ber).

5.7 Conclusions

We propose a method of joint source-channel coding for quincunx images. The presented method is designed to get the best image quality after transmission over satellite channel.

Systems using a pair of CCD linear arrays in a quincunx arrangement to improve resolution treat each image independently, disregarding the highly redundancy between them. In the present work we take into account the redundancy between the two CCD arrays when performing joint source-channel coding. Furthermore, taking into account the satellite model characteristics presents a good trade off quality-robustness comparing with standard methods using FEC codes.

To make comparisons we also adapt the MDBA presented in previous chapters 3 and 4 for quincunx images. The proposed MDC for quincunx images only over-performs the MDBA when a very noisy channel is attended. This can be explained by the fact that the two CCD arrays that are used in the new MDC are highly redundant while the MDBA starts reducing this redundancy and only then join some controlled redundancy depending of the present channel state.

The method presented here and the MDBA adapted to quincunx images was presented in [135].

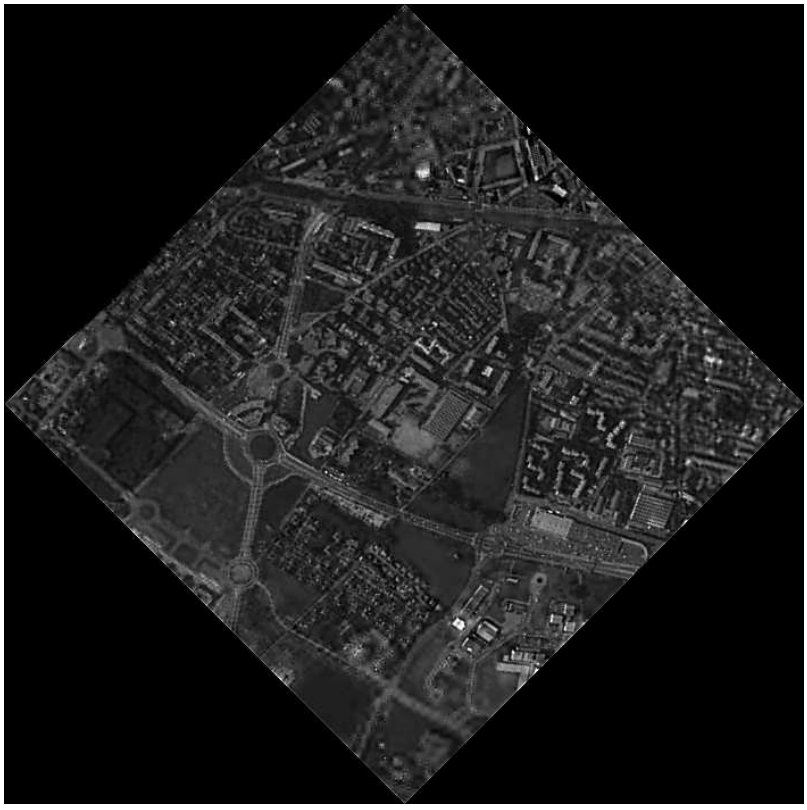

Figure 5.8: *Nimes* image compressed to 2 bpp when considering transmission at an elevation angle of 30° (0.001 ber). Standard MDC method. Side decoder 1.

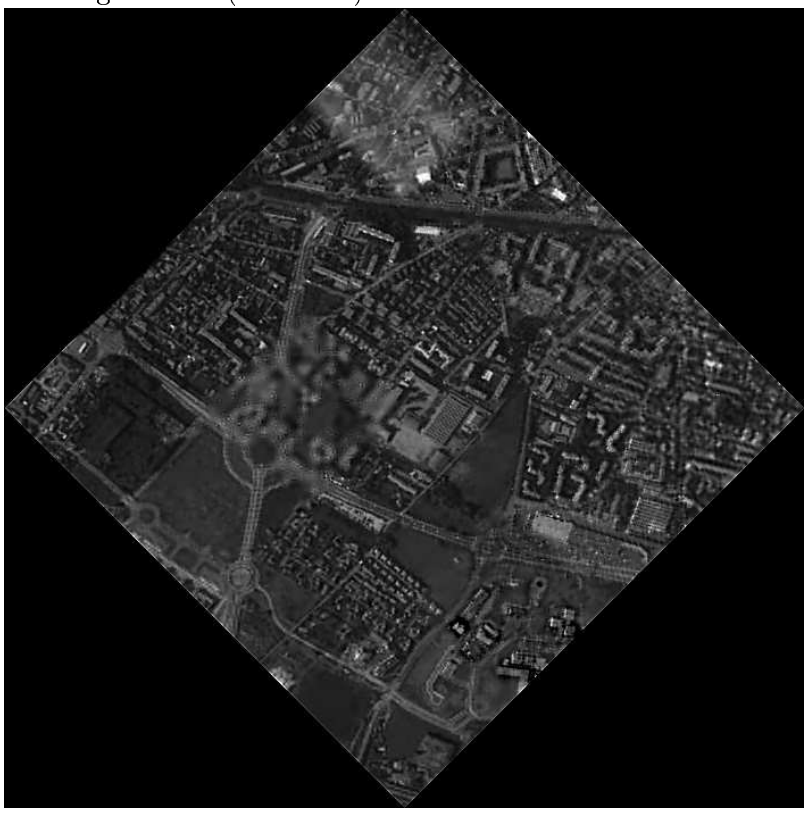

Figure 5.9: *Nimes* image compressed to 2 bpp when considering transmission at an elevation angle of 30° (0.001 ber). Standard MDC method. Side decoder 2.

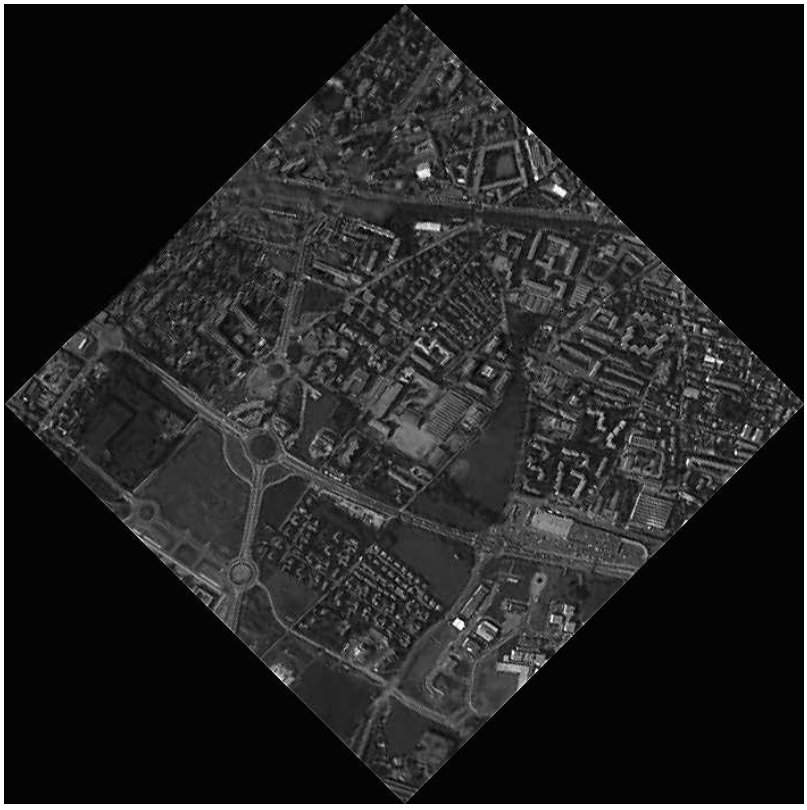

Figure 5.10: *Nimes* image compressed to 2 bpp when considering transmission at an elevation angle of 30° (0.001 ber). Proposed MDC method. Side decoder 1.

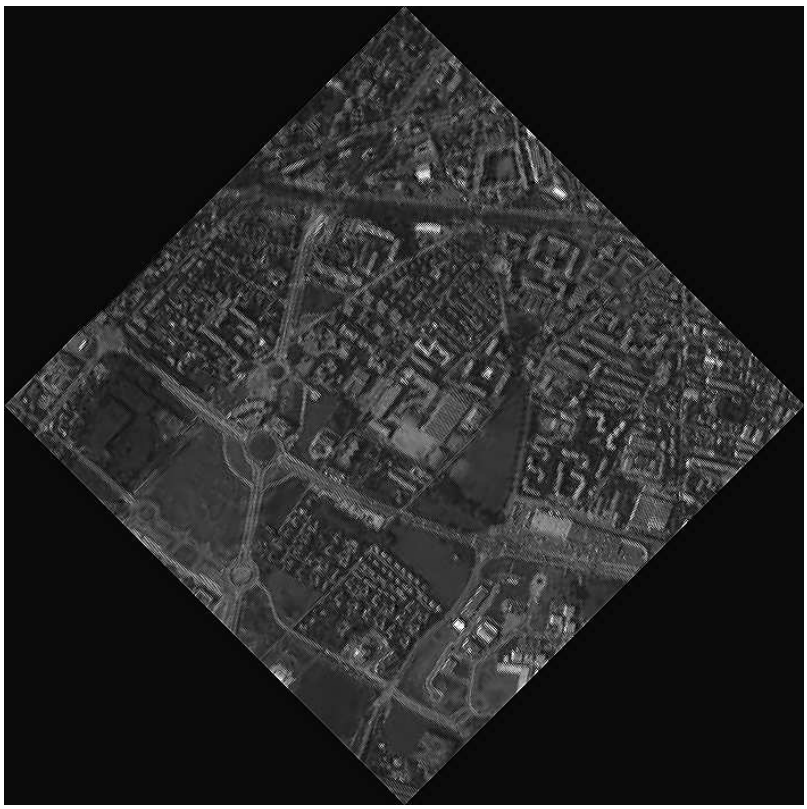

Figure 5.11: *Nimes* image compressed to 2 bpp when considering transmission at an elevation angle of 30° (0.001 ber). Proposed MDC method. Side decoder 2.

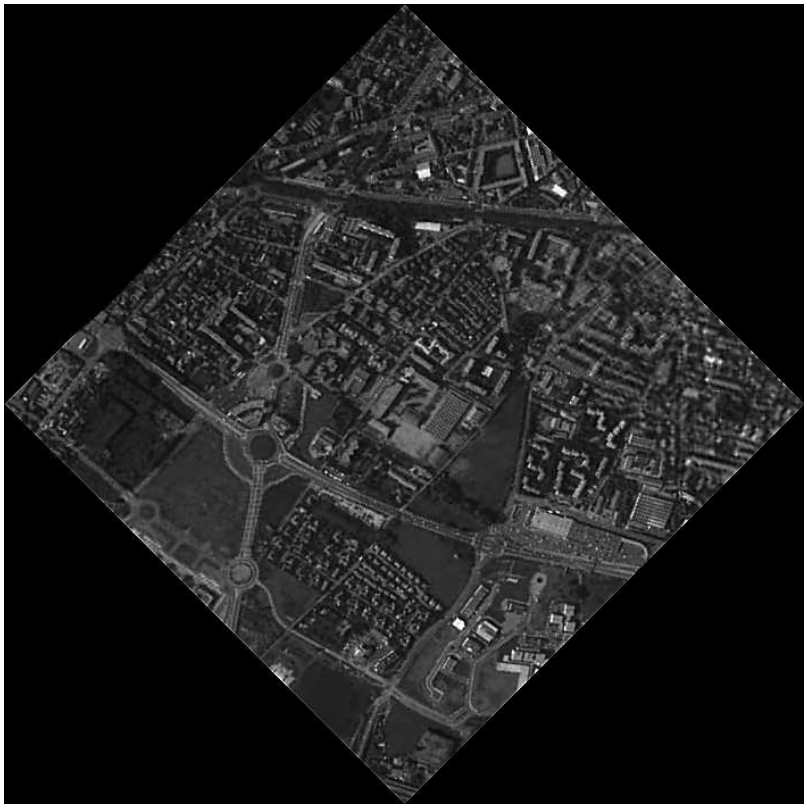

Figure 5.12: *Nimes* image compressed to 2 bpp when considering transmission at an elevation angle of 30° (0.001 ber). Standard MDC method. Central channel.

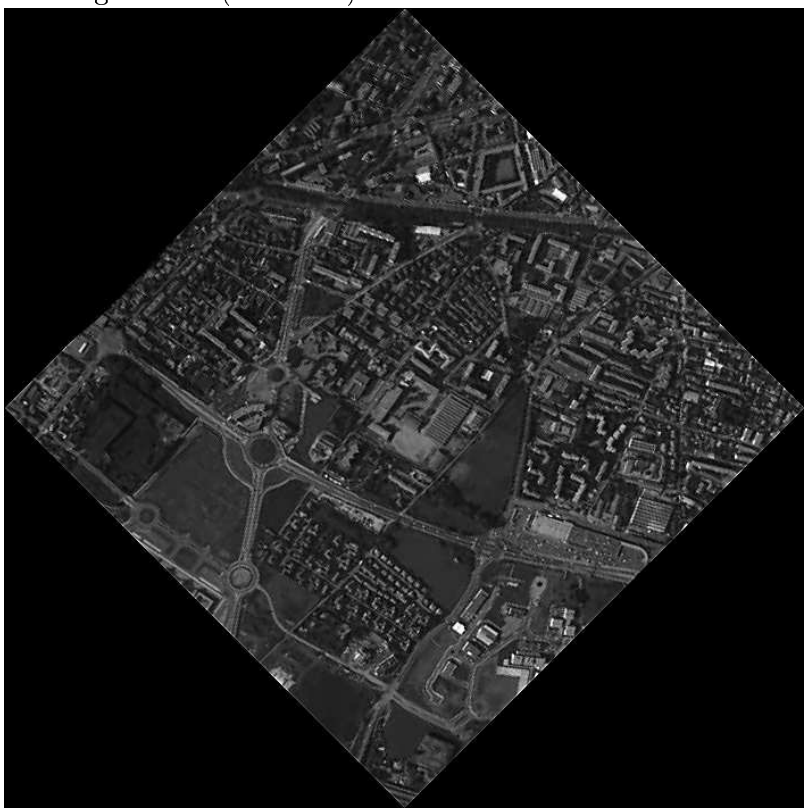

Figure 5.13: *Nimes* image compressed to 2 bpp when considering transmission at an elevation angle of 30° (0.001 ber). Proposed MDC method. Central channel.

Part III

Conclusions

Conclusions

The Multiple Description Coding (MDC) is a particular joint source and channel coding method whose main idea is to generate multiple independent descriptions of the source such that each description independently describes the source with a certain desired fidelity. When more than one description is available at the decoder side, they can be combined to enhance the quality. The MDC appears in the 80's to solve a problem related with packet losses. Moreover, almost all multiple description codes to date assume the existence of multiple independent on-off channels between the transmitter and the receiver (ex: Internet).

Several approaches of Multiple Description Coders were created since the appearance of the first practical coder in 1981 by Jayant and Christensen, [80, 81]. Standard MDC approaches (MDSQ, MDTC) involve the design of specific transforms or quantizers that have to be matched to the desired level of protection. In these schemes, adapting to changing network conditions would entail having encoder and decoder both change the transform and/or quantizers they use. These approaches are thus limited in their ability to adapt to changing transmission conditions. The last MDC approach (using explicit redundancy) tries to overcome this limitation. This MDC exploit the natural correlation between symbols for reconstruction. Explicit redundancy techniques have the additional advantage of providing very simple mechanisms for adaptation to changing network conditions.

In this dissertation a Multiple Description Coder is proposed (MDBA) that belongs to the new MDC approach. The proposed MDBA automatically adapts the explicit redundancy to changing networks transmission. This is done by bit allocation allowing the automatic adjustment of the encoder and not needing any changes to the decoder.

The advantage of the proposed MDBA over the former MDC, in the new approach, is the dependency of the explicit redundancy of the channel model and state. This propriety made possible the development of a MDC performing well in different channels environments and adapted even for time varying channels. The proposed MDC can be used for ideal MDC channel environment and for packet lossy channels, as the other methods into this approach, but can also be automatically adapted to any other channel model. The proposed MDC can then be used for any application involving transmissions over unreliable channels. It is specially suited for real time application, where typical error control techniques as ARQ, FEC or even both together cannot be easily adapted.

We start this dissertation presenting several applications of the new world of communications and the methods usually used to improve reliability of communications. We explain why for some of these applications it could be advantageous to use an MDC method instead of the standard methods usually used. This is presented in part I chapter (chapter 1) that pretend to explain our interest in the MDC methods. In part I chapter 2 we present the MDC history. The MDC evolutes since the 80' s. We present the history of MDC since is appearance to the new MDC approach. We expose how this evolution matches nowadays communications.

In the second part of the manuscript, chapter 3 and chapter 4 we present the proposed MDC approach. We call it MDBA since the redundancy is dispatched while performing the bit allocation procedure.

The chapter 3 present the general proposed MDC and compare it with the existent ones to prove the effectiveness of the proposed method. It is based on the Discrete Wavelet Transform (DWT) and on an efficient bit allocation technique. The different MD are defined when setting the bit allocation of each subband. We name it Multiple Description Bit Allocation (MDBA). The bit allocation for the successive sets of coefficients can be performed with respect to either rate or quality constraints. In both cases, the goal is to find a set of quantizers to apply in each subband whose performance lies on the convex hull of the global rate-distortion curve [161, 142, 54, 114]. To model the rate and distortion we use a non-asymptotic theoretical model for both rate and distortion [127]. The rate and distortion depends on the quantization step but also on the probability density function (pdf) of the wavelet coefficients. Assuming that the probability density model is accurate, this method provides optimal rate-distortion performances.

Chapter 4 shows how to automatically adapt the proposed MDBA to channel characteristics using channel information to inject redundancy in the different descriptions. For that, we take into account the Shannon theorem (Theorem 10) [156], and propose to dispatch the redundancy using the equivocation $H_y(x)$. Indeed, in this theorem, Shannon states that the equivocation $H_y(x)$ is the amount of redundancy that the decoder needs to correct the received message. We present in chapter 4 several application of the proposed MDBA for different channel transmissions. This serves to prove the channel adaptability of the proposed MDBA. The presented applications involve image or video transmission. Since the MDBA is based on DWT we can extend the presented coder to video by adding to the 2D DWT a 1D DWT in the time direction. This 3D subband coding of video provides encouraging results compared with MPEG. Furthermore there exist efficient 3D scan-based DWT and 3D scan-based motion compensated lifting DWT for video coding [120, 121, 122, 192, 168, 103, 5, 18, 4] that are well suited for real time applications. Scan-based wavelet transform algorithms are meant to progressively compute the DWT to get low memory implementations of wavelet transforms.

The 3D scan-based DWT transform allows us to develop a stripe-based MDC. In this way we can adapt the redundancy parameter, presented in section

4.3, to the channel state while coding. Thus, the proposed method automatically adapts the coding process to time varying states. The proposed MDBA is then suitable for video transmission over time-varying channels as can be seen by presented simulations.

The performed simulations are compared with an SDC with Turbo Codes. The Turbo Codes are very powerful codes used in channel coding. We could see in the results the high ability of the proposed method in adapting to different channel models.

Also in chapter 4 was proposed a method that explores arithmetic coding through what we call a *Smart Arithmetic Coding* to provide synchronization and minimize the error propagation in the case of channel errors.

In chapter 5 we present the proposed MDBA when used for quincunx images.

A way to improve image resolution is to combine a pair of CCD linear arrays in a quincunx arrangement. Because each CCD array yields a classical image according to a square grid the systems using such acquisition model are tempted to treat each image isolated, disregarding the high redundancy between them.

We proposed a joint source-channel MDC developed specifically for the quincunx images. We compare its performance with the above MDBA. It reach the best results for high levels of noise.

I finish this manuscript with the feeling that there is much more to be done. The proposed MDBA could and should evolute. Several applications can use advantageously the proposed MDBA. In the following we present some of the perspectives for the MDC future.

Perspectives

The generalization of the proposed method to n channels is an important point for the evolution of the MDBA. This generalization should consider transmission using similar channel models, but with different characteristics (different error probabilities). This problem really involves different points. The first one is the decision of the optimal number of descriptions to be used when a specific probability of error is expected. Other point is how to dispatch the redundancy between the different descriptions. For the first point we could use the work of Sagetong and Ortega in [148]. For the second we are making progresses.

Another important point for the MDBA evolution is related with the redundancy parameter. Alternatives to be used when the channel has memory and where the channel capacity reflects long term behavior are very important for future applications of the MDBA.

Other study to be done is the usage of motion compensation in video applications.

In the simulations presented we did not consider motion compensation. We could easily join motion compensation, using motion compensated temporal wavelet decomposition [5, 18], if we do not consider the transmission of vector motion in the MDC process (considering for instance that the vectors are transmitted using dedicated transmissions, that means without losses). Nevertheless, this process does not change the MDC itself. This is the reason why we were not interested to do it in the present work.

Consider the vector transmission with the rest of the data, implies differentiated protections. It is natural to think that motion information should have special redundancy when compared with other data. We intend to introduce vector transmission in future MDBA video applications.

These are some of points that will provide an evolution of the proposed MDBA. But as we said before some perspectives are simply the use and eventually adaptation of the proposed MDBA to new applications. Several new applications of the new mobile communication and multimedia communication appear in the last decade. The applications using data transmission over networks or diversity systems can take advantages in MDC usage.

For instance, in **chapter B** we present an application of the proposed MDBA for video streaming for wired or wireless networks.

The presented MDBA is effective for applications demanding a feedback-based encoder, as peer-to-peer video conferencing, wireless video, etc. However it does not fit in a real video streaming system where a server may be serving thousands of clients simultaneously. Nevertheless, the MDC in general and the MDBA in particular are well suited for streaming application due to their error resilience capabilities in packet loss transmissions and since the MDBA, for a given bit rate, dispatches the amount of redundancy on the different descriptions by tuning a redundancy parameter.

We explain the reasons that made the proposed MDBA suitable as base coder for video streaming. The proposed video streaming for wired or wireless networks present rate scalability by using multiple redundant representations and provide adaptability to time varying channel conditions. The main advantages of the proposed scheme when compared with the commercial SDC schemes are that the proposed video streaming scheme is error adaptable, it avoids retransmission and does not suffer from synchronization or error propagation problems. These last characteristics are due to the independence of the descriptions, and also due to the usage of the *Smart Arithmetic Coding*.

The presented result serves to prove the capability of the proposed scheme in adapting to varying networks.

We present the video streaming method proposed here in [130, 132] for wired and wireless communications, respectively.

Appendix

Appendix A

UMTS channels

The UMTS standard (Universal Mobile Telecommunications System) was created by the ITU organism (International Telecommunication Union). This organism pretends to find the standards and recommendations for future generation of mobile communications. They intend to respond to the new exigencies of systems capacity and performances. This effort results in the third generation mobile communication. The UMTS standardization in Europe starts in the 90' s in the European Telecommunication Standard Institute (ETSI). Nowadays, the UMTS technology is developed by the 3GGP (3^{rd} Generation Partnership Project).

3G is characterized by high-speed, high-bandwidth services that support a variety of applications, including high quality voice, high data rate, multimedia and video. The higher channel capacity (these systems allow bit rates that goes to 10Mbits/s in UMTS, compared with the 9.6 Kbits/s in GSM) is due to the introduction of spectrum spread technology (WCDMA). These systems are modulated accordingly the user needs or the new application needs as mobile video conference, the files transfer, the mobile Internet etc. These multimedia services need great flexibility systems: it is important to take care of services needing different bit rates and different reception quality. These services should be efficiently multiplexed to optimize the system resources.

France Telecom developed UMTS propagation models. These models reproduce the pertinent channel characteristics; time varying multipath channels are simulated. The emphases were given to the impact of different parameters such as the binary data rate, the mobile velocity, the imperfection of the channel estimation, etc. The numeric part of the baseband reception chain of a downlink communication (from the station to the mobile) was entirely simulated.

The test channels [44] were defined in the UMTS norm. There were developed two channel models:

- The static Gaussian channel (or AWGN) as reference channel;
- The multipath channel for different fading profiles.

Three terrestrial models with multipath propagation are defined by ETSI [44]. These test channels present six significant separable paths, Rayleigh fading amplitudes, uniform phases (between 0 and 2 π), and quasi-static paths.

Indoor channel A & B Corresponding to small cells (pico cells) with weak transmission power, particularly office interiors, base stations, pedestrians with reduced mobility, etc. The variable fading parameters depends on dispersion and attenuation dues to walls, floor, and metallic structures, etc. The indoor channel is almost similar to a fading situation with a single path, because the delay of propagation are confined to the chip duration (approximately).

Outdoor to Indoor and Pedestrian channel A & B (It is named simply by Pedestrian A & B in the follows). Corresponds to building exteriors and pedestrians with greater displacement velocity than in Indoor case. The fading variations level depends of displacement velocity and of reflection from moving vehicular. The Pedestrian B channel fade over approximately one symbol (flat fading situation).

Vehicular channel A & B Corresponding to large cells with most important transmission power. The fading nature depends on environment type: urban, rural, mountain, etc. and also on the vehicle velocity. Different of Indoor and Pedestrian channels, the Vehicular channel B extent's over four symbols (selective frequency fading situation).

The tables A.1, A.2 and A.3 resume, for the six significant path, the temporal delay relative to the first path and the mean power relative to the higher powered path. We recall that the temporal duration of one chip (T_c) is approximately 260 nsec.

Path	Indoor A channel		Indoor B channel	
	Relative delay [nsec]	Mean power [dB]	Relative delay [nsec]	Mean power [dB]
1	0	0	0	0
2	50	-3.0	100	-3.6
3	110	-10.0	200	-7.2
4	170	-18.0	300	-10.8
5	290	-26.0	500	-18.0
6	310	-32.0	700	-25.2

Table A.1: Indoor A & B multipath channel parameters.

Path	Pedestrian A channel		Pedestrian B channel	
	Relative delay [nsec]	Mean power [dB]	Relative delay [nsec]	Mean power [dB]
1	0	0	0	0
2	110	-9.7	200	-0.9
3	190	-19.2	800	-4.9
4	410	-22.8	1200	-8.0
5	-	-	2300	-7.8
6	-	-	3700	-23.9

Table A.2: Pedestrian A & B multipath channel parameter.

Path	Vehicular A channel		Vehicular B channel	
	Relative delay [nsec]	Mean power [dB]	Relative delay [nsec]	Mean power [dB]
1	0	0.0	0	-2.5
2	310	-1.0	300	0
3	710	-9.0	8900	-12.8
4	1090	-10.0	12900	-10.0
5	1730	-15.0	17100	-25.2
6	2510	-20.0	20000	-16.0

Table A.3: Vehicular A & B multipath channel parameters.

The profiles of Doppler spectra (the Doppler effect is due to transmitter or receptor movements) affecting each path are defined as in [44]:

- Classic Doppler for the six paths for the Pedestrian A & B and Vehicular A & B channels.

- Flat Doppler, defined by $S(f) = \frac{1}{2.f_D}$, $|f| < f_D$, for the six paths of Indoor A & B channels.

The A.4 table resume the maximum values for binary bit rates and displacement velocity for each one of the test environments.

Test environment (channel models)	Maximum bit rate	Maximal velocity
Vehicular A & B channel	144 Kbits/s	500 Km/h
Pedestrian A & B channel	384 Kbits/s	120 Km/h
Vehicular A channel		
Indoor A & B channel	2048 Kbits/s	10 Km/h
Pedestrian A channel		

Table A.4: Bit rate and velocity characteristics for each type of UMTS environment.

The transmitted signals are submitted to different fading forms.

- Multiple Access Interference (MAI): the interference of other users in the propagation channel (in the multiple access case by codes multiplexing).
- The thermal noise due the antenna or other components from the reception chain (amplifiers, filters, etc.).
- Sources of exterior radiation: industrial and urban parasitic emissions, atmospheric noise, etc.
- Interference sources extern to the system (as pirate emissions in tight band; this ones are not much harmful for spread spectra systems. They are note taken into account in the present simulator)

The different intern and extern noise contributions will be taken into account as an unique noise source located upstream the receptor. This noise is supposed to have a spectral density with uniform power (additive white noise). This can be model by a Gaussian random process, with zero mean, independent from the emitted signal, and will be called AWGN. This one will be quantified by his spectral density with mono-lateral power N_0 [Watt/Hz] centered to 2GHz frequency and 3.84MHz band width. Thus, the reception quality is expressed in function of the E_b/N_0 relation (relation between the mean of the received signal by the spectral mono-lateral noise density bit information).

As diversity technique, for better quality reception with UMTS, is used path diversity. The table A.5 resumes the number of maximal separable paths for each channel type, between the test models of 3GPP. The frequency diversity is intrinsic to signals transmitted by a wide band.

Model channel 3GPP	Number of paths in the model	Coherence band	Maximum number of separable paths for a bit rate of 4.0 Mchip/s
Indoor A channel	6	4.3 MHz	1
Indoor B channel	6	1.6 MHz	3
Pedestrian A channel	4	3.5 MHz	2
Pedestrian B channel	6	250 kHz	17
Vehicular A channel	6	430kHz	10
Vehicular B channel	6	53 kHz	78

Table A.5: Coherence band and maximal number of separable paths for the 3GPP test models UMTS environment.

When received the different de-correlated versions from the signal, they are combined into an only optimal signal. The optimization criterion of this combination is the received signal SNR, or simply the reception power. The used technique is the MRC (Maximal Ratio Combining). This technique combines, by ponderation, the different received versions. Thus, the combined signal is the sum of the different diversity branch, each one ponderated with a value depending of his SNR or his amplitude. This is the most complex technique. An phase and amplitude estimation is needed for each version of the signal. However, this technique present the better quality reception. This technique is adopted as adapted filter.

For more detailed information about the simulator or the UMTS models see [50, 1, 2, 36, 74, 138, 44].

Appendix B

Streaming video using multiple description bit allocation

In this chapter we present a TCP-Friendly streaming system based on the MDBA presented in chapters 3 and 4 of part 2.

Feedback-based encoders, as the MDBA, are effective in peer-to-peer video conferencing, wireless video, etc, but they don't fit in a real video streaming system. Considering the server may be serving thousands of clients simultaneously, it is unrealistic to expect the server to encode and output thousands of bitstreams that are adapted to each client.

This chapter presents the proposed MDC for video streaming that is rate scalable and quality adaptable. It uses multiple redundant representations to allow rate scalability and adaptability to varying networks. This is done without requiring computation at the media server. Furthermore, the proposed system avoids some problems related with layered approaches.

We start this chapter presenting video streaming related problems in section B.1. Also in this section, we present an overview of the proposed method and the main advantages of this method when compared with the existent state of the art.

The rest of the chapter is organized as follows. Section B.2 presents some of the most important existing streaming approaches. In section B.3 we propose a streaming scheme using the MDBA presented in chapter 3 of part 2. In this section we explain how we use the MDBA to design a streaming scheme and we present the advantages of using the MDBA. A general streaming scheme and some server specifications are also presented. Results are presented in section B.4 and conclusions are given in section B.5.

B.1 Introduction

The use of video streaming over Internet knew an enormous increase in the past few years being the design of Internet video streaming a challenging task due to the unpredictable and varying nature of network conditions. Furthermore, the increasingly access of Internet from wireless, often mobile terminal, comes up with additional problems. If in the former problem is considered that packets arriving to client are uncorrupted, in the case of wireless communications this is not guaranteed. Thus, to the problems related to Internet communications

(bandwidth, delay jitter and loss rates) we have an additional problem that demands effectiveness in presence of channel failures when performing wireless streaming.

Video streaming, when the network is afflicted by unpredictable and time-varying losses and available bandwidth, involves several problems as presented above. They can be resumed as *rate adaptation* and *error control and recovery* [100]. Conventional approaches for dealing with losses for static data, such as retransmission may not be possible in streaming context due to the real time nature of the content. Thus, additional mechanism are needed to provide streaming media delivery over Internet or wireless channels.

Commonly RTP protocols are used for streaming. RTP uses UDP as underlying transport protocol and has, unlike TCP, no built-in congestion control. This behavior has a potential congestion collapse problem. However, TCP transport protocols are not suitable for real time transmission due to the low latency objectives of such applications. One way to ensure fair competition with TCP for bandwidth is performing TCP friendly rate control at the server. The server ensures that the RTP packets are sent at a rate no higher than what is regarded as a TCP friendly share of the available bandwidth. TCP-friendly rate adaptation is provided in this case by the recently proposed TCP-friendly rate control (TFRC) protocol [49, 72].

In streaming video the client performs a demand to a server that transmits media packets over a network that serves fairly several clients. The server can implement intelligent transport mechanisms, by sending out the right packets at the right time, but the amount of computation that it can perform for each media stream is very limited due to the large number of streams to be served simultaneously. It is unrealistic to expect the server to encode and output thousands of bitstreams that are adapted to each client. Even the automatic conversion of the bit rate of a pre-encoded video bitstream, known as transcoding [125, 78, 56] is too complex.

Therefore, the task to compress video signal is left to the encoder. This task has to be done without the a priori knowledge of the channel conditions (bandwidth and loss rates). This is why representations that allow rate scalability must be adapted to varying network throughput without requiring computation at the media server. Multiple redundant representations are an easy way to achieve this task and will be used in the proposed scheme. Similar approaches for switching among different streams for conventional single description streaming are the actual industrial applications. RealSystem G2 was the first to support dynamic stream switching under the name of SureStream [97, 117]. The idea is that the server can switch among multiple streams to serve a client with one that best matches the client's available bandwidth.

At a panel on the "Future of Video Compression" at the Picture Coding Symposium held in April 1999, it was agreed that rate scalability was important for media streaming applications. It also appears that one may want to design a compression scheme that is tuned to the channel over which the video is transmitted [170]. MDC were considered as an option for this last task.

The MDBA, presented in previous chapters 3 and 4 of part 2, for a given bit rate dispatches the amount of redundancy on the different descriptions by

tuning a redundancy parameter. The proposed scheme is composed of multiple descriptions, generated by the MDBA, and is rate and quality adaptable. Switching among the multiple pre-encoded descriptions provide adaptability to changing bandwidth, such as in the industrial applications. But in the proposed video streaming scheme, switching among the multiple pre-encoded descriptions also provides adaptability to time varying channel conditions.

The multiple pre-encoded descriptions generated by the MDBA are independent and error resilient. Moreover, retransmission is avoided decreasing network congestion, and suppressing additional delays (the server will not have to wait for packet delivery information). These are two problems of standard single description streaming schemes that use layered representations. The standard single description streaming schemes, based on DCT and using predictive coding in temporal domain have additional problems when transmitting over wireless channels. These problems are synchronization and error propagation. When an error occurs it propagates to other frames due to predictive coding. For example, the emerging MPEG-4 video standard (see [41, 126] for details) supports mechanisms such as resynchronization markers, data partition, and reversible VLC for intraframe error control. When these features are implemented in a compressed bitstream, the decoder can regain synchronization in case of bit errors, and continue to decode a video frame even if a portion of the frame is damaged or lost. However, any intraframe error control technique would become useless if a whole video frame is lost completely what is often the case in low bit rate video streaming. The standard H263+ also has several error resilience modes at source coding level, including Slice-Structured Mode, Independent Segment Decoding Mode and Reference Picture Selection Mode. The first two techniques are designed for intraframe error control, thus, they are not very effective in controlling errors that result from a packet loss in video streaming. The last technique involves a feedback-based encoder that, as mentioned before, is inadequate for video streaming.

Our MDBA coder uses wavelets in temporal domain and uses the *smart arithmetic coding* presented in previous chapter 4 to avoid problems related with synchronization and error propagation.

Experimental results shows that our system adapts well to time varying bandwidth and to time varying channels.

B.2 Existing streaming approaches

Almost all current approaches use Forward Error Correction (FEC) and/or Automatic Repeat reQuest (ARQ). For instance, in [163] a two-layer scalable video coder combined with unequal error protection is used. In [141] a generalized MD ($N > 2$ descriptions) coding through the usage of FEC codes for streaming video was proposed. In [19] they use a hybrid FEC/ARQ approach known as incremental redundancy. In this paper they compute the rate distortion optimized transmission policy using the Iterative Sensitivity Adjustment (ISA) algorithm introduced in [28]. The ISA algorithm involves estimation of the probability that a single packet will be communicated with an error function of the expected

redundancy, or cost, used to communicate the packet. The [20] work is an extension of the work in [28]. It incorporates models for packet path diversity. Also in [95] the authors exploits path diversity. In [105] a media layer representation for transmission over current heterogeneous networks is used. The authors propose a framework for scalable streaming media delivery, that involves a scheduling algorithm called expected runtime Distortion Based Scheduling (EDBS) which decides the order in which packets should be transmitted in order to improve client playback quality in the presence of channel losses. In [84] they use adaptive media playout to reduce the delay introduced by the client buffer. They consider retransmission of lost media packets. To minimize retransmission in [94] the authors use pre-stored representations of certain frames at the server such that the chosen representation only uses previous frames, as reference, received with very high probability.

The approaches above use standard video coders (MPEG-2/4, H.263). In such schemes retransmission of lost media packets is essential for a video streaming application over error-prone channels. Continuous video playout at the receiver can only be guaranteed if all packets are available due to the interdependency of successive video packets introduced by motion compensated prediction. In [25] the three-dimensional (3-D) SPIHT with a new method of partitioning the wavelet coefficients into spatio-temporal tree blocks is used to achieve error resilience. However, they had to sacrifice the progressiveness of the first block of the bitstream to get error resilience. The method also has the problem that if a portion at the beginning of the bitstream is lost they can not reconstruct anything from the bitstream. This is also a problem of the above layered approaches. Such approaches essentially prioritize data and thereby support intelligent discarding of the data (the enhancement data can be lost or discarded while maintaining usable video). However the base layer bitstream is critically important - if it is lost, then the other bitstream(s) are useless. Layered approaches do not seem to be adequate for Internet where all packets are equally likely to be lost, so video can be completely lost if there is an error in the base layer.

An alternative to layered methods are joint source and channel coding methods as multiple description coding. In MDC each description or MD stream is independent of each other and is typically of roughly equal importance such that, if either description is received it can be used to decode baseline quality video, and both descriptions can be used to decode improved quality video. This is in contrast to conventional video coders (e.g. MPEG-1/2/4, H.261/3/4, which produce a single stream that does not have these MD properties; we refer to these methods as single description coding (SDC).

A number of MD video coding algorithms have recently been developed, which provide different tradeoffs in terms of compression performance and error resilience [186, 177, 146, 7].

Some MDC schemes dedicated to streaming video coding and using path diversity have recently been developed. In [10] MD coding is used to code a media stream into multiple complementary descriptions, which are distributed across the edge servers in a Content Delivery Network (CDN). They exploit path diversity provided by the different network paths that exists between a

client and its nearby edge servers. Also in [8] a MDC is proposed, when path diversity is considered. This coding scheme assumes that there are several parallel channels between the source and destination, and that each channel may be temporarily down or suffering from long burst errors. Furthermore, the error events of different channels are independent so that the probability that all channels simultaneously experience losses is small. These channels could be physically distinct paths between the source and destination in, for example, a wireless multipath network or a packet switched network. Even when only a single physical path exists between the source and destination, the path can be divided into several virtual channels by using time interleaving frequency division, etc.

The proposed method uses MDC schemes, and it can easily be adapted to take advantage of path diversity. Here, we introduce the method for the case when only one path is considered.

B.3 Proposed MDC for video streaming

B.3.1 General proposed scheme

As we said above the proposed video streaming scheme uses multiple redundant representations of the original signal. These multiple representations allow rate scalability and adaptability to varying network throughput without requiring computation at the media server.

Similar approaches for switching among different streams for conventional single description streaming are the actual industrial applications. The problem of such approaches are that they use standard coders, thus they present some problems related with such coders.

We use the MDBA proposed in previous chapters 3 and 4 to generate the multiple pre-encoded files. These different descriptions are independent and error resilient. Moreover, retransmission is avoided decreasing network congestion, and suppressing additional delays (the server will not have to wait for packet delivery information). These are two problems of standard single description streaming schemes that use layered representations. The standard single description streaming schemes, based on DCT and using predictive coding in temporal domain have additional problems when transmitting over wireless channels. These problems are synchronization and error propagation. When an error occurs it propagates to other frames due to predictive coding. Our MDBA coder uses wavelets in temporal domain and uses the *smart arithmetic coding* presented in previous chapter 4 to avoid these problems.

In the following we will detail the proposed video streaming scheme based on MDBA.

B.3.2 Generation of the multiple redundant representations

We divide our signal into sub-signals or Group Of Pictures (GOP) representing the video at different time (for instance see figure B.1).

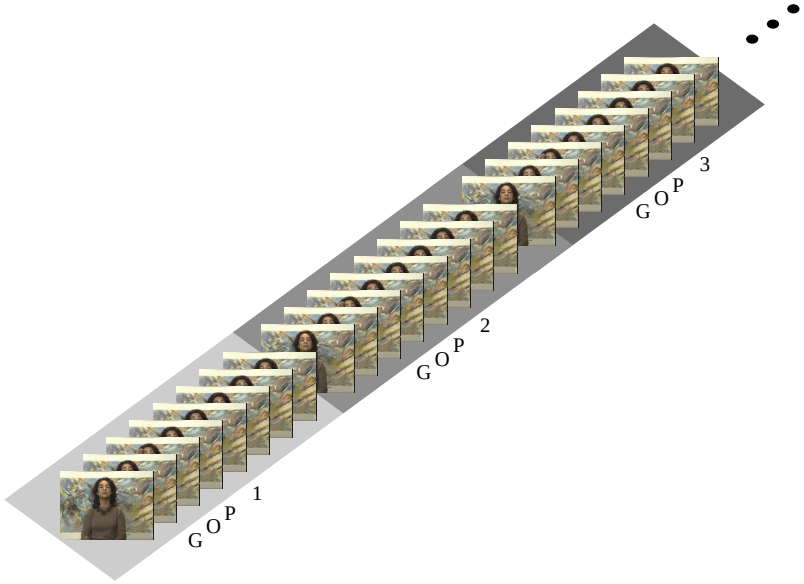

Figure B.1: Signal division into sub-signals or Group Of Pictures (GOP)

For each GOP, the MDBA coder presented in previous part 2, chapters 3 and chapter 4 is performed for different bit rate and redundancy parameters. Each MDBA execution results in two descriptions (we call them in the following base descriptions) adapted for a specific bit rate and channel state. The different descriptions generated are downloaded into the server. The server possesses several descriptions with the format presented in figure B.2 and organized as presented in table B.1.

Each column in table B.1 presents descriptions adapted to a specific bit rate. This means that each base description into the same column have been coded with the bit rate parameter specified in the top of the column.

Each line in table B.1 presents base descriptions adapted to a specific channel state. Thus, base descriptions into the same line has been coded with the redundancy parameter specified in the left of the line.

Since each GOP is coded independently, it is possible to switch of description between GOP's. Switching among these multiple pre-encoded descriptions results in a rate scalable and time varying channel adapted scheme.

In the following we present the advantages of using the MDBA when generating the multiple redundant representations of the signal.

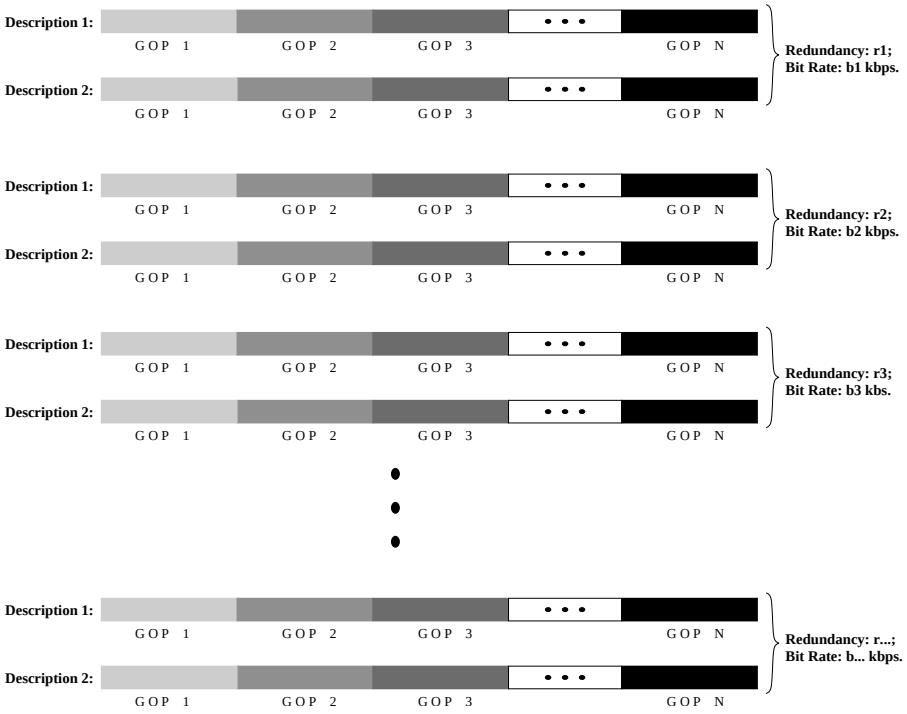

Figure B.2: Descriptions resultant from MDBA

	Bit Rate			
Redundancy	**100Kbps**	**200Kbps**	**...**	**X Mbps**
1	$\text{Description}_{1,1,1}$ $\text{Description}_{2,1,1}$	$\text{Description}_{1,1,2}$ $\text{Description}_{2,1,2}$	... ...	$\text{Description}_{1,1,n}$ $\text{Description}_{2,1,n}$
0.5	$\text{Description}_{1,3,1}$ $\text{Description}_{2,3,1}$	$\text{Description}_{1,3,2}$ $\text{Description}_{2,3,2}$	... ...	$\text{Description}_{1,3,n}$ $\text{Description}_{2,3,n}$
0.1	$\text{Description}_{1,5,1}$ $\text{Description}_{2,5,1}$	$\text{Description}_{1,5,2}$ $\text{Description}_{2,5,2}$	... ...	$\text{Description}_{1,5,n}$ $\text{Description}_{2,5,n}$
...	... ...	... ...	... ...	... ...
0	$\text{Description}_{1,m,1}$ $\text{Description}_{1,m,1}$	$\text{Description}_{1,m,2}$ $\text{Description}_{1,m,2}$	... ...	$\text{Description}_{1,m,n}$ $\text{Description}_{1,m,n}$

Table B.1: Different pre-stored descriptions for an interval time of video (for a GOP).

B.3.3 Advantages of using the MDBA

Video streaming systems have to deal with different problems:

- delivery constraints;
- time variable available bandwidth;
- time variable channel conditions.

Due to delivery constraints, retransmission should be avoided. Thus MDC schemes are preferred to layered or multi resolution schemes since in the former the descriptions are independent what is not true in other methods. The layered or multi resolutions schemes need retransmission to work properly.

In the following we present some of the advantages of MDC in general and of the MDBA in particular:

- In MDC methods, streams are independent. Thus, as soon as a stream reach the decoder it can be decoded. The decoder does not need to verify if other streams with higher priority had been already decoded as in layered approaches. Also due to the stream independence, retransmission is avoided. The server does not need streams deliver information, since each packet are independent from the others.

- It was shown in chapter 3 that high compression efficiency is achieved with the MDBA coder. MDBA give similar results than SDC (without channel coding) in the noiseless case. It was shown in chapter 4 that with the MDBA coder, a long burst error or even the loss of an entire description does not have a catastrophic effect. Since streams are error resilient, the decoder can always use the received bitstream, unless it is completely lost. In this last case the other streams will be entirely used.

- Time variable available bandwidth implies adaptability of compression parameters to vary the rate of the bitstream to available network bandwidth. As presented in previous section, each MDBA execution is performed for specific bit rate and redundancy parameters. Moreover, switching among the multiple pre-encoded descriptions provide bit rate scalability, thus is adapted to time varying available bandwidth.

- Time variable channel conditions implies adaptability of compression parameters to vary the redundancy in the bitstream to network conditions. Since each MDBA execution is performed for specific redundancy parameters switching among the multiple pre-encoded descriptions provide adaptability of redundancy parameter to time varying channel conditions.

- Furthermore, our MDBA coder uses wavelets in temporal domain to avoid error propagation problems. Bitstream synchronization is also very important, since in wireless transmission we have no guarantee that the stream reach the decoder without errors. We have presented in part 2, chapter 4 a *smart arithmetic coding* to provide synchronization and minimize error

propagation in case of errors in the bitstreams. We use this *smart arithmetic coding* in the present streaming system. Standard SDC streaming schemes are based on DCT and use predictive coding in temporal domain. When an error occurs it propagates to other frames due to predictive coding.

In resume, the main advantages of using the MDBA as base coder in a streaming system are streams independence that results in a retransmission free system, streams resilience, automatic scalability to time variable available bandwidth and automatic adaptability to time variable channel conditions.

We can conclude that the MDBA coder is well suited for applications attending high error rates or burst errors as is the Internet or the wireless streaming.

We present next some considerations of the base description generation.

B.3.4 Base descriptions

The base MDBA coder for a given bit rate dispatches the amount of redundancy on the different descriptions by tuning a redundancy parameter.

We show in chapters 3 and chapter 4 that the MDBA coder with automatic redundancy control is error resilient. This resilience is due to an automatic adaptation of the amount of added redundancy (through the redundancy parameter) to the underlying channel error characteristics.

To automatically compute the redundancy parameter we proposed to use equation (4.5) in chapter 4. Remark that using equation (4.5) redundancy 1 gives maximal robustness and redundancy 0 gives maximal quality. Intermediate values of redundancies give a trade-off between quality and robustness. This equation uses channel model and state information to compute the best redundancy to apply. Specific equations for BSC, Gaussian and 3G channels are presented in chapter 4.

In this chapter since we use UMTS channels and/or Internet channels, the computation of the channel capacity is performed using the Rayleigh model (equation 4.11 in chapter 4 part 2) or using the two state Markov model by Gilbert [55] and Elliot [42]. Moreover, as the UMTS channel simulator we use considers QPSK modulation, we use equation (4.12) and equation (4.11) to compute the redundancy parameter, when considering UMTS transmission. For the internet simulator it was shown in chapter 4, section 4.7.4.3 that the Markov model provides a good approximation in modeling the error process at the packet level. We show that the redundancy parameter is in this case computed using the channel capacity presented in equation (4.14).

Both equations (4.12) and (4.14) exploit channel state information. Using channel information when computing the redundancy to dispatch between descriptions results in an MDC automatically adapted to channel models and states.

We propose to use this MDBA coder to code the video signal at different bit rates to get rate scalability.

B.3.5 Principle of the proposed scheme

We propose to download the different base descriptions generated for different bit rates (adapted to variable bandwidth) and different redundancies (adapted to loss rate) in the server. The descriptions will be organized as presented in table B.1.

Each column in table B.1 presents base descriptions adapted to a specific bit rate. This mean that all base descriptions into the same column has been coded with the same bit rate. This bit rate is the one specified in the top of the column. Moreover, each base descriptions into the same column have been coded with different redundancy parameters.

We can see in figure B.3 the relation between distortion and bit rate. In this example, redundancy parameter was calculated for internet channel with 10 % packet loss (transmission was not considered). This figure represents the distortions associated with one line of the table B.1. As expected, for the same redundancy parameter, higher bit rates imply smaller distortions.

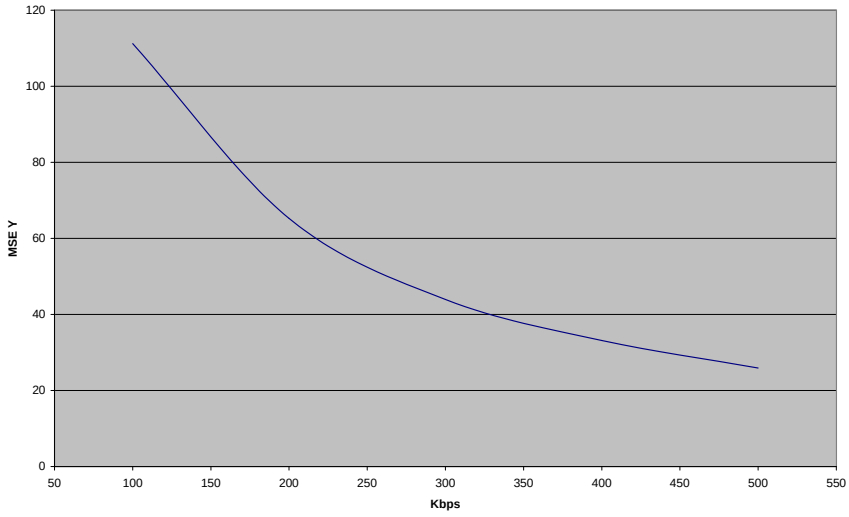

Figure B.3: Y component of QCIF Foreman video for bit rates from 100 to 500 Kbps. 10 % probability of packet loss.

Each line in table B.1 presents base descriptions adapted to a specific channel state. Thus, base descriptions into the same line has been coded with the same redundancy parameter. This redundancy parameter is the one specified in the left of the line. Moreover, each base descriptions into the same line has been coded with different bit rate parameters.

Figure B.4 shows the relation between distortion and packet loss. In this example, bit rate was fixed at 200Kbps (transmission is not considered). This

represents the distortions associated with one column of the table B.1. In this figure, for the same bit rate, higher redundancy implies higher side distortion.

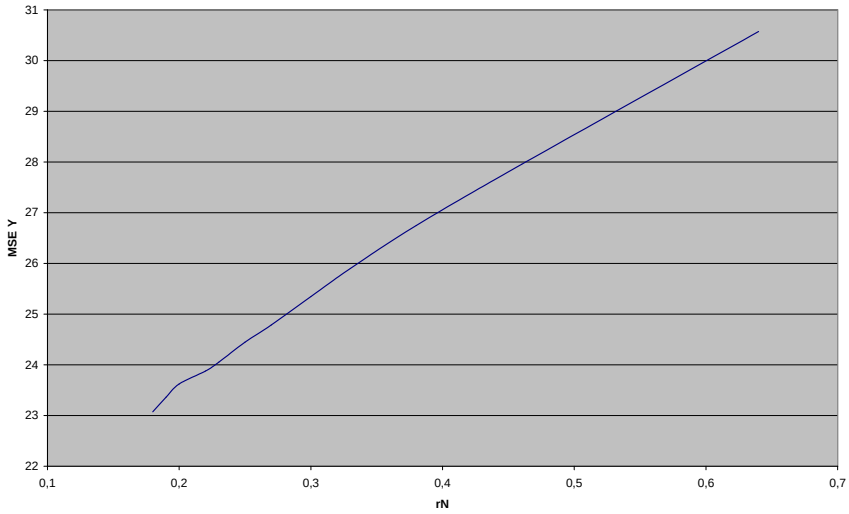

Figure B.4: Y component of QCIF Foreman video compressed at 500 Kbps for different probabilities of packet loss (i.e. different r_N values). $r_N = 1$ means high packet losses.

The server has the only task to switch between lines and columns.

Switch between lines to adapt to the network conditions. To choose the optimal line, the server computes the redundancy parameter based in equations (4.12) and (4.14) as explained in section B.3.4.

Switch between columns to adapt to the bandwidth (when a TCP-friendly application is pretended the server estimates the current TCP's throughput and sends the data bounded by this value). More details are presented in next section, that give some server specifications.

Switching among the multiple pre-encoded descriptions provide adaptability to time varying available bandwidth and to time varying channel conditions.

Since descriptions are independent all descriptions that reach the decoder can be used to enhance the video quality. For its compression and synchronization capabilities, the proposed streaming system based in the MDBA is suitable for real time transmission.

B.3.6 Server specification

These multiple descriptions generated by the coder are downloaded to the server. The server only has to choose sending out the right description at the right time depending of channel conditions (bandwidth and loss rate).

The server can choose the description with the bit rate that is more suitable for each client and which redundancy is most adapted to the present loss rate. Furthermore, the server can continue to send descriptions corresponding to the same GOP of video but with less redundancy (better quality) when there is time for it. With this scheme we avoid the added delay of methods using retransmission that have to wait for the packet delivery state. Then, the server will repeat the above process for the next GOP of the video. The server will always start a GOP with the most suitable redundancy for the present channel condition and continues sending descriptions that will enhance the video quality.

The computation of the number of bits the server can send for each GOP can be done based on the last packet delays as proposed in proposition 3.

Proposition 3 *Let us define δ as the interval of emission between two successive packets, rtt the average of the last packets delays and pbr the packet size (in bits). Considering t the time we have to send the GOP, the number n of bits to send in order to meet real time decoding is computed by:*

$$n \leq \frac{(t+\delta)(pbr)}{rtt+\delta}. \tag{B.1}$$

Proof We call n to the number of bits we have time to send (based on previous time delays). Considering we are sending packets with pbr bits, $\frac{n}{pbr}$ will be the number of packets we have to use to send the n bits. Considering rtt the average of the last packets delays and δ the interval of emission between two successive packets, equation (B.2) defines the time we need to send these n bits.

$$\frac{n}{pbr} \times rtt + \frac{n}{pbr} \times \delta - \delta \tag{B.2}$$

If we consider the needed decoding time as initial buffering, we have t seconds to send the n bits. Moreover, from equation (B.2) we can write the inequality (B.3).

$$\frac{n}{pbr} \times rtt + \frac{n}{pbr} \times \delta - \delta < t. \tag{B.3}$$

This inequality is easily simplified in the proposed equation to compute the number of bits to send for a GOP, equation (B.1).

The server can use n in equation (B.1) to decide if it is possible to send more descriptions with less redundancy or if it starts sending descriptions for the next GOP in order to meet real-time decoding.

B.3.6.1 TCP throughput

To achieve fairness between TCP and non-TCP flows, and improve overall network utilization, we use TCP throughput equation. The server estimates the current TCP's throughput and sends the data bounded by this value. The TCP

throughput equation is given in B.4 as a function of packet loss rate (p) and round trip time (rtt) as presented in [116].

$$TCP_{throughput} \leq \frac{1.22pbr}{rtt\sqrt{p}} \quad \text{bit/second} \tag{B.4}$$

More precisely, if the client bandwidth is smaller than the TCP throughput, the server uses this bandwidth to chose the right column in table B.1. Otherwise it uses the TCP's throughput.

The estimation of rtt and p parameters used above is done using feedback information.

Thus, the rate control is done by the server that adapts the video rate to the available bandwidth. This bandwidth is variable, but since we download descriptions with different bit rates for each GOP, and these descriptions are independent, server can adapt the video rate to variable bandwidth.

B.4 Results

For spatial decomposition our coder uses 9-7 biorthogonal filter [6] and a three level decomposition. For temporal decomposition it uses the (2,2) filter and performs a two level decomposition. For more coder specifications see chapter 4, section 4.5.

We present results obtained with the proposed MDC with and without redundancy adaptability, in Figure B.5. We considered transmission over a Pedestrian UMTS channel.

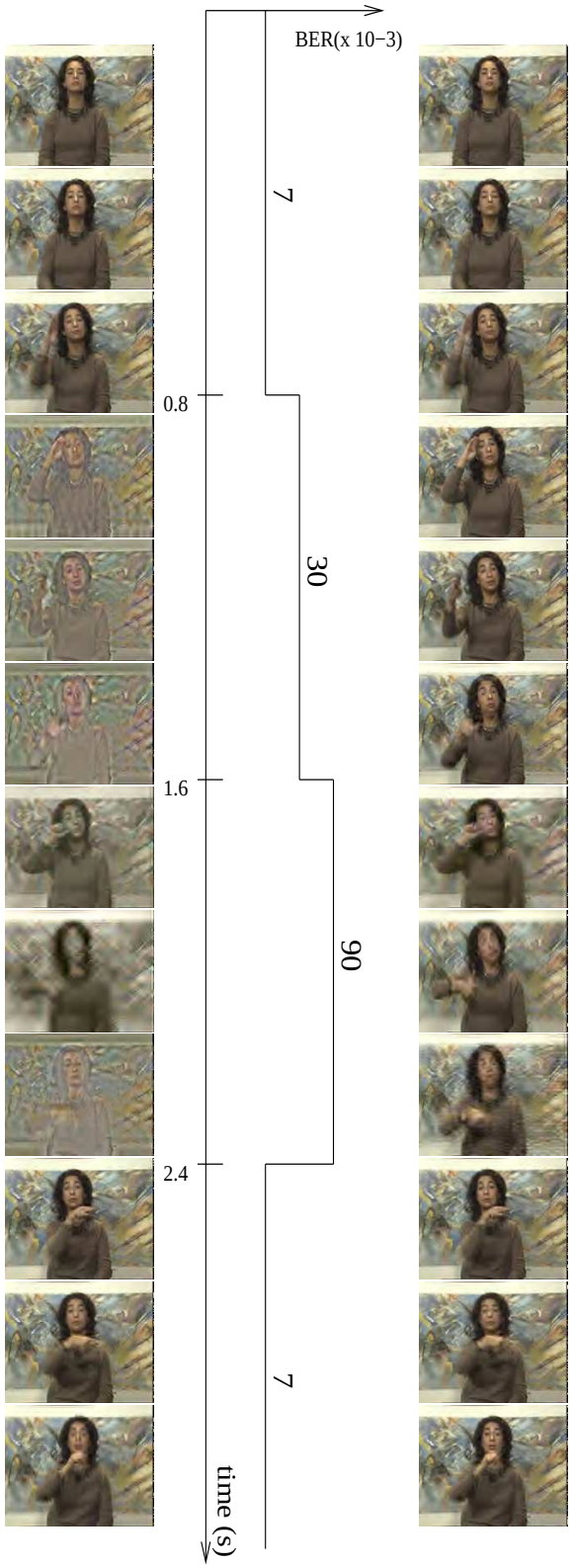

Figure B.5: Silent video compressed at 200 Kbps and transmitted over a UMTS simulator. Left column: without redundancy adaptability; Right column: with redundancy adaptability.

B.5 Conclusions

We proposed a streaming system that downloads multiple descriptions in the server. The server switches among the multiple pre-encoded descriptions to provide adaptability to time varying available bandwidth and to time varying channel conditions. The proposed MDC video streaming system avoids problems related with standard approaches.

Experimental results shows that our system presents rate scalability and channel noise adaptability.

This is a work in progress. It already yielded the publications [130, 132] for wired and wireless communications, respectively.

Appendix C

Original images and videos

C.1 Images

Figure C.1: 512×512 pixels *Lena* image.

Figure C.2: 352×360 pixels CDD1 *Nimes* image.

Figure C.3: 352×360 pixels CDD2 *Nimes* image.

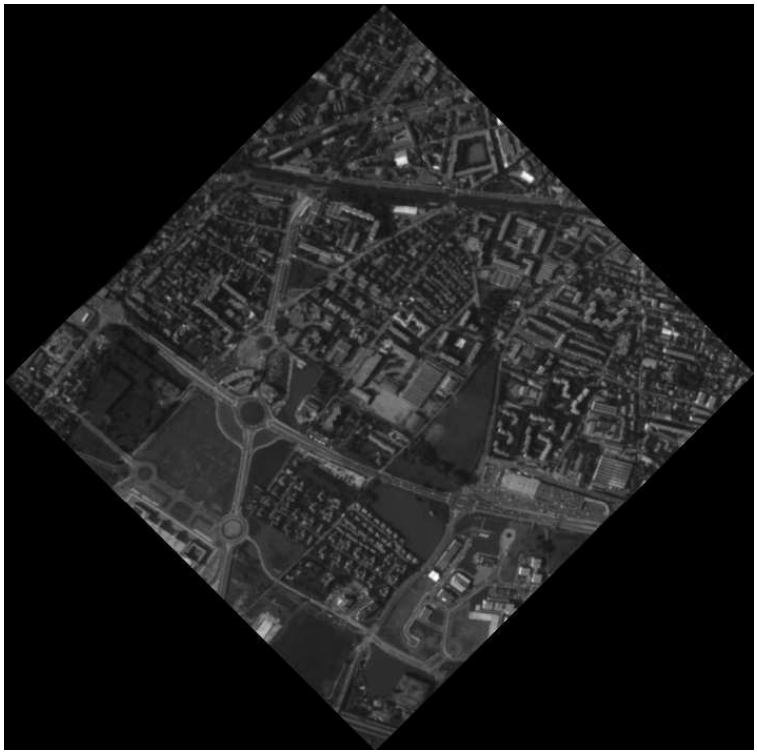

Figure C.4: Quincunx *Nimes* image.

C.2 Videos

Figure C.5: QCIF *silent* color video. Frames: 1, 11, 21, 31,...

Figure C.6: CIF *akiyo* video. Frames: 1, 11, 21, 31, ...

Figure C.7: QCIF *foreman* color video. Frames: 10, 20, 30,...

Bibliography

[1] *3GPP, Technical Specification Group: Physical channels and mapping of transport channels onto physical channels (TDD)*, TS25.221 v3.2.0 edition, Mars 2000.

[2] *3GPP, Technical Specification Group: Spreading and modulation (TDD)*, TS25.223 v3.2.0 edition, Mars 2000.

[3] R. Ahlswede. The rate-distortion region for multiple descriptions without excess rate. *IEEE Transaction on Information Theory*, IT-31(6):721–726, November 1985.

[4] T. André, M. Antonini, and M. Barlaud. Codage vidéo par transformée en ondelettes 3d au fil de l'eau et compensée en mouvement. In *Premier Congrès IEEE International de Signaux Circuits et Systèmes*, Monastir, Tunisie, Mars 2004.

[5] T. André, M. Cagnazzo, M. Antonini, M. Barlaud, N. Bozinovic, , and J. Konrad. (n,0) motion-compensated lifting-based wavelet transform. In *IEEE International Conference in Acoustics, Speech and Signal Processing*, Montreal, Quebec, Canada, May 2004.

[6] M. Antonini, M. Barlaud, P. Mathieu, and I. Daubechies. Image coding using wavelet transform. *IEEE Transaction on Image Processing*, 4(8):1053–1060, August 1992.

[7] J. Apostolopoulos. Error-resilient video compression via multiple state streams. In *International Workshop on Very Low Bitrate Video Coding (VLVB)*, pages 168–171, October 1999.

[8] J. Apostolopoulos. Reliable video communication over lossy packet networks using multiple state encoding and path diversity. In *SPIE Visual Communication Image Processing Conference*, January 2001.

[9] J. Apostolopoulos and S. Wee. Unbalanced multiple description video communication using path diversity. In *IEEE International Conference Image Processing*, Thassaloniki, Greece, October 2001.

[10] J. Apostolopoulos, T. Wong, Wai tian Tan, and S. Wee. On multiple description streaming with content delivery networks. In *IEEE Infocom*, June 2002.

[11] N. At and Y. Altunbasak. Multiple description coding for wireless channels with multiple antennas. In *GLOBECOM*, pages 2040–2044, San Antonio, TX, November 2001.

[12] R. Balan, I. Daubechies, and V. Vaishampayan. The analysis and design of windowed Fourier frame based multiple description source coding schemes. *IEEE Transaction on Information Theory*, 46(7):2491–2537, 2000.

[13] J. Batllo and V. Vaishampayan. Asymptotic performance of multiple description transform codes. *IEEE Transaction on Information Theory*, 43(2):703–707, March 1997.

[14] T. Berger and Z. Zhang. Minimum breakdown degradation in binary source encoding. *IEEE Transaction on Information Theory*, IT-29(6):807–814, November 1983.

[15] T. Berger-Wolf and E. Reingold. Index assignment for multichannel communication under failure. *IEEE Transaction on Information Theory*, 48(10):2656–2668, October 2002.

[16] J. Bolot. Characterizing end-to-end packet delay and loss in the Internet. *Journal of High-Speed Networks*, 2(3):305–323, December 1993.

[17] B.Usevitch. Optimal bit allocation for biorthogonal wavelet coding. In *IEEE Data Compression Conference*, pages 387–395, Snowbird, UT, April 1996.

[18] M. Cagnazzo, T. André, M. Antonini, and M. Barlaud. (n,0) motion-compensated lifting-based wavelet transform. In *IEEE International Conference Image Processing*, Submitted to Singapore, October 2004.

[19] J. Chakareski, P. Chou, and B. Aazhang. Computing rate-distortion optimized policies for streaming media to wireless clients. In *IEEE Data Compression Conference*, Snowbird, UT, April 2002.

[20] J. Chakareski and B. Girod. Rate-distortion optimized packet scheduling and routing for media streaming with path diversity. In *IEEE Data Compression Conference*, Snowbird, UT, April 2003.

[21] V. Chande, N. Farvardin, and H. Jafarkhani. Image communication over noisy channels with feedback. In *IEEE Data Compression Conference*, Snowbird, UT, April 1999.

[22] P. Charbonnier, M. Antonini, and M. Barlaud. Implantation d'une transformée en ondelettes 2D dyadique au fil de l' eau. Technical report, CNES (French Space Agency), October 1995.

[23] Q. Chen and T. Fisher. Image coding using robust quantization for noisy digital transmission. *IEEE Transaction on Image Processing*, 1996.

[24] Q. Chen and T. Fisher. Robust quantization for image coding and noisy digital transmission. In *IEEE Data Compression Conference*, pages 3–12, Snowbird, UT, March 1996.

[25] S. Cho and W. Pearlman. Error resilience and recovery in streaming of embedded video. *Signal Processing, Special Issue on Image and Video Coding Beyond Standards*, 82:1545–1558, November 2002.

[26] S. Choi and J. Woods. Motion-compensated 3D subband coding of video. *IEEE Transaction on Image Processing*, 8(2):155–167, 1999.

[27] P. Chou, S. Mehrotra, and A. Wang. Multiple description decoding of overcomplete expansions using projections onto convex sets. In *IEEE Data Compression Conference*, Snowbird, UT, 1999.

[28] P. Chou and Z. Miao. Rate-distortion optimized streaming of packetized madia. *IEEE Transaction on Multimedia*, 2001.

[29] C. Chrysafis and A. Ortega. Efficient context-based entropy coding lossy wavelet image compression. In *IEEE Data Compression Conference*, Snowbird, UT, March 1997.

[30] C. Chrysafis and A. Ortega. Line based reduced memory wavelet image compression. *IEEE Transaction on Image Processing*, 9(3):378–389, March 2000.

[31] C. Chrysafis and A. Ortega. Minimum memory implementations of the lifting scheme. In *SPIE*, San Diego, US, August 2000.

[32] V. Chu and P. Sweeney. Channel modelling and error control strategies for the LEO satellite channel. In *International Symposium on Communications Theory and Applications (ISCTA)*, Ambleside, UK, July 1999.

[33] D.-M Chung and Y. Wang. Multiple description image coding using signal decomposition and reconstruction based on lapped orthogonal transforms. *IEEE Transaction on Circuits and Systems for Video Technology*, 9(6):895–908, September 1999.

[34] D. Comas, R Singh, A. Ortega, and F. Marqué. Unbalanced multiple description video coding with rate-distortion optimization. *EURASIP Special Issue Multimedia Signal Processing*, 2003(1):81–90, January 2003.

[35] J. Conway and N. Sloane. *Sphere Packing, Lattices and Groups*. Springer Verlag, 3rd edition edition, 1998.

[36] E. Dahlman, B. Gudmundson, M. Nilsson, and J. SKlöd. UMTS/IMT-2000 based on wideband CDMA. *IEEE Communication Magazine*, pages 70–80, 1998.

[37] I. Daubechies and W. Sweldens. Factoring wavelet transforms into lifting steps. *Journal of Fourier Analysis and Applications*, 4(3):247–269, 1998.

[38] S. Diggavi, N Sloane, and V. Vaishampayan. Asymmetric multiple description lattice vector quantizers. *IEEE Transaction on Information Theory*, 48(1):174–191, January 2002.

[39] P. Dragotti, J. Kovacevic, and V. Goyal. Quantized oversampled filter banks with erasure. In *IEEE Data Compression Conference*, Snowbird, UT, march 2001.

[40] P. Dragotti, S. Servetto, and M. Vetterli. Optimal filter banks for multiple description coding: Analysis and synthesis. *IEEE Transaction on Information Theory*, 48(7):2036–2052, July 2002.

[41] L. Ducla-Soares and F. Pereira. Error resilience and concealment performances for MPEG-4 frame-based video coding. *Signal Processing: Image Communication*, 14(6-8):447–472, May 1999.

[42] E. Elliot. Estimates of error for codes on burst-noise channel. *Bell System Technical Journal*, 42:1977–1997, September 1963.

[43] W. Equitz and T. Cover. Sucessive refinement of information. *IEEE Transaction on Information Theory*, 37(2):269–275, March 1991.

[44] ETSI - European Telecommunications Standards Institute. *Selection procedures for the choice of radio transmission technologies of the UMTS*, TR 101 112 v3.2.0 edition, April 1998.

[45] N. Farvardin and V. Vaishampayan. Optimal quantizer design for noisy channels: An approach to combined source-channel coding. *IEEE Transaction on Information Theory*, 33(6):827–838, November 1987.

[46] J. C. Feauveau. Analyse multirésolution pour les images avec un facteur de résolution $\sqrt{2}$. *Traitement du signal*, 7(2), 1990.

[47] B. Felts and B. Pesquet-Popescu. Efficient context modeling in scalable 3D wavelet-based video compression. In *IEEE International Conference Image Processing*, Vancouver, Canada, September 2000.

[48] M. Fleming and M. Effros. Generalized multiple-description vector quantization. In *IEEE Data Compression Conference*, Snowbird, Utah,, March 1999.

[49] S. Floyd, M. Handley, J. Padhye, and J. Widmer. Equation-based congestion control for unicast application. In *SIGCOM*, August 2000.

[50] J. Francis. *Etude des systèmes de traitement numérique de canal pour la reception radio-mobile UMTS/TDD*. PhD thesis, Institut National Polytechnique de Grenoble, France, September 2002.

[51] B. Frichman. A binary channel characterisation using partitioned Markov chains. *IEEE Transaction on Information Theory*, IT-13:221–227, April 1967.

[52] A. Gamal and T. Cover. Achievable rates for multiple description. *IEEE Transaction on Information Theory*, IT-28(6):851–857, November 1982.

[53] J. Garcia-Frias and J. Villasenor. An analytical treatment of channel-induced distortion in entropy coded image subbands. In *IEEE Data Compression Conference*, Snowbird, UT, March 1997.

[54] A. Gersho and R. Gray. *Vector Quantization and Signal Compression.* Kluwer Academic Publishers, Boston, MA, 1992.

[55] E. Gilbert. Capacity of burst-noise channel. *Bell System Technical Journal*, 39:1253–1265, September 1960.

[56] G.Keesmman, R.Hellinghuizen, F.Hoeksema, and G.Heidman. Transcoding of MPEG bitstreams. *Signal Processing: Image Communication*, 8:481–500, September 1996.

[57] A. Gouze, M. Antonini, and M. Barlaud. Quincunx filtering lifting scheme for image coding. In *SPIE Visual Communication Image Processing Conference*, volume 1, pages 1–11, Janvier 1999.

[58] A. Gouze, M. Antonini, and M. Barlaud. Quincunx lifting scheme for image compression. In *IEEE International Conference Image Processing*, volume 1, pages 665–668, Vancouver, septembre 2000.

[59] A. Gouze, M. Antonini, M. Barlaud, and B. Macq. Optimized lifting scheme for two-dimensional quincunx sampling images. In *IEEE International Conference Image Processing*, volume 2, pages 253–258, Thessaloniki, Greece, October 2001.

[60] A. Gouze, M. Antonini, M. Barlaud, and B. Macq. Signal-adapted multidimensional lifting scheme. *Acepted for publication: IEEE Transaction on Image Processing*, May 2003.

[61] V. Goyal, J. Kelner, and J. Kovacevic. Multiple description lattice vector quantization: Variations and extensions. In *IEEE Data Compression Conference*, Snowbird, UT, March 2000.

[62] V. Goyal, J. Kelner, and J. Kovacevic. Multiple description vector quantization with a coarse lattice. *IEEE Transaction on Information Theory*, 48(3):781–788, March 2002.

[63] V. Goyal and J. Kovacevic. Optimal multiple description transform coding of gaussian vectors. In *IEEE International Conference Image Processing*, Chicago, IL, October 1998.

[64] V. Goyal and J. Kovacevic. Generalized multiple description coding with correlating transforms. *IEEE Transaction on Information Theory*, 47(6):2199–2224, September 2001.

[65] V. Goyal, J. Kovacevic, R. Arean, and M. Vetterli. Multiple description transform coding of image. In *IEEE International Conference Image Processing*, Chicago, USA, October 1998.

[66] V. Goyal, J. Kovacevic, and M. Vetterli. Multiple description transform coding: robustness to erasures using tight frame expansions. In *International Symposium Information Theory*, page 408, Cambridge, MA, August 1998.

[67] V. Goyal, J. Kovacevic, and M. Vetterli. Quantized frame expansions as source-channel codes for erasure channels. In *IEEE Data Compression Conference*, Snowbird, UT, march 1999.

[68] V. Goyal, M. Vetterli, and N. Thao. Quantized overcomplete expansions in $\mathbb{R}^n$: Analysis, synthesis, and algorithms. *IEEE Transaction on Information Theory*, 44(1):16–31, January 1998.

[69] R. Gray and D.L. Neuhoff. Quantization. *IEEE Transaction on Information Theory*, 44(6):2325–2384, October 1998.

[70] T. Guionnet, C. Guillemot, and E. Fabre. Soft decoding of multiple descriptions. In *IEEE International Conference in Multimedia & Expo*, pages 26–29, Lausanne, Suisse, August 2002.

[71] T. Guionnet, C. Guillemot, and S. Pateux. Embedded multiple description coding for progressive image transmission over unreliable channels. In *IEEE International Conference Image Processing*, pages 7–10, Thessalonique, Grece, October 2001.

[72] M. Handley, J. Padhye, S. Floyd, and J. Widmer. TCP friendly rate control (TFRC): Protocol specification. In *RFC 3448, Proposed Standard*, January 2003.

[73] V. Hardman, A. Sasse, M. Handley, and A. Wason. Reliable audio for use over the Internet. In *INNET*, 1995.

[74] H. Holma and A. Toskala. *UMTS, Les réseaux de troisième génération.* Osman Eyrolles Multimedia, Paris, 2000.

[75] C.-Y. Hsu and A. Ortega. A lagrangian optimization approach to rate control for delay-constrained video transmission over burst-error channels. In *IEEE International Conference in Acoustics, Speech and Signal Processing*, New York, NY, April 1998.

[76] C.-Y. Hsu, A. Ortega, and M. Khansari. Rate control for robust video transmission over burst-error wireless channels. *IEEE Journal Selected Areas on Communication*, 17(5):1–17, May 1999.

[77] ISO. *JPEG 2000 image coding system - part 1: Core coding system.* ISO/IEC 15444-1 Information Technology, 2000.

[78] S. Jacobs and A. Eleftheriadis. Streaming video on the WEB using dynamic rate shaping. *Journal of Visual Communication and Image Representation*, 9(3):211–222, September 1998.

[79] H. Jafarkhani and V. Tarokh. Design of successively refinable trellis coded quantizers. *IEEE Transaction on Information Theory*, 45(5):1490–1497, July 1999.

[80] N. Jayant. Subsampling of a DPCM speech channel to provide two "self-contained" half-rate channels. *Bell System Technical Journal*, 60(4):501–509, April 1981.

[81] N. Jayant and S. Christensen. Effects of packet losses in waveform coded speech and improvements due to an odd-even sample-interpolation procedure. *IEEE Transaction on Communication*, COM-29(2):101–109, February 1981.

[82] W. Jiang and A. Ortega. Multiple description coding via polyphase transform and selective quantization. In *SPIE Visual Communication Image Processing Conference*, 1999.

[83] R. Joshi, H. Jafarkani, J. Kasner, T. Fiscer, N. Farvardin, M. Marcellin, and R. Barbenger. Comparasion of different methods of classification in subband coding images. *IEEE Transaction on Image Processing*, 6(11):1473–1486, November 1997.

[84] M. Kalman, E. Steinbach, and B. Girod. Adaptive media playout for low delay video streaming over error-prone channels. *IEEE Transaction on Circuits and Systems for Video Technology*, 2001.

[85] N. Kamaci, Y. Altunbasak, and R. Mersereau. Multiple description coding with multiple transmit and receive antennas for wireless channels: The case of digital modulation. In *GLOBECOM*, pages 3272–3276, San Antonio, TX, November 2001.

[86] G. Karlsson and M. Vetterli. Subband coding of video for packet networks. *Optical Engineering*, 27:574–586, 1988.

[87] G. Karlsson and M. Vetterli. Three-dimensional subband coding of video. In *IEEE International Conference in Acoustics, Speech and Signal Processing*, pages 1100–1103, New York, NY, April 1988.

[88] B.-J Kim and W. Pearlman. An embedded wavelet video coder using three-dimensional set partitioning in hierarchical trees (SPIHT). In *IEEE Data Compression Conference*, pages 251–260, Snowbird, Utha, March 1997.

[89] K. Kintzley. An application of multiple description scalar quantizers to speech coding on correlated fading channels. Master's thesis, Electrical Engineering Departement, Texas A&M University, College Station, TX, August 1995.

[90] J. Kovacevic, P. Dragotti, and V. Goyal. Filter bank frame expansions with erasures. *IEEE Transaction on Information Theory*, 48:1439–1450, June 2002.

[91] C. Lambert-Nebout, C. Latry, G. Moury, C. Parisot, M. Antonini, and M. Barlaud. Image compression for future high resolution optical remote sensing missions. In *SPIE*, San Diego, 2000.

[92] C. Latry and B. Rougé. SPOT5 THR MODE. In *SPIE Visual Communication Image Processing Conference*, USA, 1998.

[93] W. Lee. Estimate of channel capacity in Rayleigh fading environnment. *IEEE Transaction on Vechicular Technology*, 39:187–189, August 1990.

[94] Y. Liang, , and B. Girod. Low-latency streaming of pre-encoded video using channel-adaptive bitstream assembly. In *IEEE International Conference Image Processing*, Rochester, NY, US, September 2002.

[95] Y. Liang, E. Setton, and B. Girod. Channel-adaptive video streaming using packet path diversity and rate-distortion optimized reference picture selection. In *IEEE International Workshop on Multimedia Signal Processing*, St. Thomas, Virgin Island, December 2002.

[96] S. Lin, D. Costello, and M. Miller. Automatic repeat request error control schemes. *IEEE Communication Magazine*, pages 5–17, 1984.

[97] A. Lippman. Video coding for multiple target audiences. In *SPIE Visual Communication Image Processing Conference*, volume 3653, pages 780–782, January 1999.

[98] H. Liu and M. Zarki. Performance of H263 video transmission over wireless channels using hybrid ARQ. *IEEE Journal Selected Areas on Communication*, 15(9):1775–1785, December 1997.

[99] S. LoPresto, K. Ramchandran, and M. Orchard. Image coding based on mixture modeling of wavelet coefficients and a fast estimation-quantization framework. In *IEEE Data Compression Conference*, Snowbird, UT, March 1997.

[100] J. Lu. Signal processing for Internet video streaming: A review. *SPIE Image and Video Communication and Processing*, January 2000.

[101] T. Marshall. Coding of real-number sequences for error correction: A digital signal processing problem. *IEEE JSAC*, 2(2):381–392, March 1984.

[102] S. Mehrotra and P. Chou. On optimal frame expansions for multiple description quantization. In *International Symposium Information Theory*, Sorrento, Italy, June 2000.

[103] N. Mehrseresht and D. Taubman. Adaptively weighted update steps in motion compensated lifting based on scalable video compression. In *IEEE International Conference Image Processing*, Barcelone, Spain, September 2003.

[104] Y. Meyer. *Wavelets and Operators*. Cambridge University Press, 1992.

[105] Z. Miao and A. Ortega. Optimal scheduling for streaming of scalable media. In *Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, CA, October-November 2000.

[106] A. Miguel, A. Mohr, and E. Riskin. SPIHT for generalized multiple description coding. In *IEEE International Conference Image Processing*, volume 3, pages 842–846, 1999.

[107] A. Miguel and E. Riskin. Protection of regions of interest against data loss in a generalized multiple description framework. In *IEEE Data Compression Conference*, Snowbird, UT, March 2000.

[108] M. Minoux. *Mathematical Programming: Theory and Algorithms*. Wiley: New York, 1986.

[109] A. Mohr, E. Riskin, and R. Ladner. Graceful degradation over packet erasures channels through forward error correcting. In *IEEE Data Compression Conference*, Snowbird, UT, March 1999.

[110] R. Motwani and C. Guillemot. Tree-structured oversampled filter banks as joint source-channel codes: Applications to image transmission over erasure channels. *Under Review with IEEE Transactions on Signal Processing*, March 2003.

[111] J.-R Ohm. Three dimensional subband coding with motion compensation. *IEEE Transaction on Image Processing*, 3(5):559–571, 1994.

[112] M. Orchard, Y. Wang, V. Vaishampayan, and R. Reibman. Redundancy rate-distortion analysis of multiple description coding using paiwise correlating transforms. In *IEEE International Conference Image Processing*, 1997.

[113] E. Ordentlich, D. Taubman, M. Weinberger, G. Seroussi, and M. Marcelin. Memory efficient scalable line-based image coding. In *IEEE Data Compression Conference*, Snowbird, Utha, March 1999.

[114] A. Ortega. *Compressed Video over Networks*, chapter Variable Bit-rate Video Coding, pages 343–382. M. Sun and A. Reibman Eds, Marcel Dekker, New York, NY, 2000.

[115] L. Ozarow. On a source-coding problem with two channels and three receivers. *Bell System Technical Journal*, 59(10):1909–1921, December 1980.

[116] J. Padhye, V. Firoiu, D. Towsley, and J. Kurose. Modeling TCP throughput: a simple model and its empirical validation. Technical Report Tr98-008, UMASS CMPSCI, Feb. 1998.

[117] RealNetworks White Papers. Surestream^{tm}- delivering superior quality and reliability. http://www.realnetworks.com/devzone/library/whitepapers/surestrm.html.

[118] C. Parisot. *Allocations Bases Modéles et Transforme en Ondelettes au Fil de l'Eau pour le Codage d'Images et de Vidéos.* PhD thesis, University of Nice - Sophia Antipolis, France, January 2003. http://www.i3s.unice.fr/ parisot/.

[119] C. Parisot, M. Antonini, and M. Barlaud. Ebwic: A low complexity and efficient rate constrained wavelet image coder. In *IEEE International Conference Image Processing*, Vancouver, Canada, September 2000.

[120] C. Parisot, M. Antonini, and M. Barlaud. 3D scan-based wavelet transform for video coding. In *IEEE International Workshop on Multimedia Signal Processing*. IEEE, October 2001.

[121] C. Parisot, M. Antonini, and M. Barlaud. Motion-compensated scan based wavelet transform video coding. In *Tyrrhenian International Workshop on Digital Communication*, Capri, Italy, September 2002.

[122] C. Parisot, M. Antonini, and M. Barlaud. 3D scan-based wavelet transform and quality control for video coding. *EURASIP Special Issue Multimedia Signal Processing*, 2003(1):56–65, January 2003.

[123] C. Parisot, M. Antonini, M. Barlaud, C. Lambert-Nebout, C. Latry, and G. Moury. On-board stripe-based wavelet image coding for future space missions. In *IEEE IGARSS : Special Session on Remote Sensing Compression*, Honolulu, Hawaii, Juillet 2000.

[124] C. Parisot, M. Antonini, M. Barlaud, C. Lambert-Nebout, C. Latry, and G. Moury. On board stripe-based wavelet image coding for future space remote sensing missions. In *IEEE International Geoscience and Remote Sensing Symposium*, volume 6, pages 2651–2653, Honolulu, US, Juillet 2000.

[125] P.Assunção and M. Ghanbari. A frequency-domain video transcoder for dynamique bit-rate reduction of MPEG-2 bit streams. *IEEE Transaction on Circuits and Systems for Video Technology*, 8(8):953–967, December 1998.

[126] F. Pereira. Visual data representation: recent achievements and future developments. In *EURASIP European Signal Processing Conference*, Tampere, Finland, September 2000.

[127] M. Pereira, M. Antonini, and M. Barlaud. Channel adapted multiple description coding scheme using wavelet transform. In *IEEE International Conference Image Processing*, Rochester, NY, US, September 2002.

[128] M. Pereira, M. Antonini, and M. Barlaud. Channel adapted scan-based multiple description video coding. In *IEEE International Conference in Multimedia & Expo*, Lausanne, Switzerland, August 2002. IEEE.

[129] M. Pereira, M. Antonini, and M. Barlaud. Low complexity multiple description coding scheme using wavelet transform. In *EURASIP European Signal Processing Conference*, Toulouse, France, September 2002.

[130] M. Pereira, M. Antonini, and M. Barlaud. Multiple description coding for Internet video streaming. In *IEEE International Conference Image Processing*, Barcelona, Spain, September 2003.

[131] M. Pereira, M. Antonini, and M. Barlaud. Multiple description coding for noisy-varying channels. In *IEEE Data Compression Conference*, Snowbird, US, March 2003.

[132] M. Pereira, M. Antonini, and M. Barlaud. Multiple description coding for video streaming over wireless networks. In *IEEE HSNMC*, Estoril, Portugal, July 2003.

[133] M. Pereira, M. Antonini, and M. Barlaud. Multiple description video coding for UMTS. In *Picture Coding Symposium*, Saint-Malo, France, April 2003.

[134] M. Pereira, M. Antonini, and M. Barlaud. Multiple description image and video coding for wireless channels. *EURASIP Signal Processing: Image Communication, Special issue on Recent Advances in Wireless Video*, Accepted for publication, 2003.

[135] M. Pereira, A. Gouze, M. Antonini, and M. Barlaud. Multiple description coding for quincunx images. application to satellite transmission. In *IEEE HSNMC*, Estoril, Portugal, July 2003.

[136] C. Podilchuk, N. Jayant, and N. Farvardin. Three-dimensional subband coding of video. *IEEE Transaction on Image Processing*, 4(2):125–139, 1995.

[137] S. Pradhan and K. Ramchandran. On the optimality of block orthogonal transforms for multiple description coding of gaussian vector sources. *IEEE Signal Processing Letters*, 7(4):76–78, April 2000.

[138] R. Prasad. *Universal Wireless Personal Communications*. Artech House, Boston-London, 1998.

[139] J. Proakis. *Digital Communications*. Mc Graw-Hill, New York, 1983.

[140] R. Puri, K. Ramchandran, and I. Kozintsev. Multiple description source coding using forward error correction (FEC) codes. In *Asilomar Conference on Signals, Systems and Computers*, Pacific Groove, CA, October 1999.

[141] R. Puri, K. Ramchandran, K. Lee, and V. Bharghavan. Application of FEC based multiple description coding to Internet video streaming and multicast. In *Packet Video Workshop*, Sardinia, Italy, May 2000.

[142] K. Ramchandran and M. Vetterli. Best wavelet packet bases in a rate-distortion sense. *IEEE Transaction on Image Processing*, 1(2):160–176, 1993.

[143] G. Rath and C. Guillemot. Performance analysis and recursive syndrome decoding of dft codes for bursty erasure recovery. *Accepted for publicastion in IEEE Transactions on Signal Processing*, 2002.

[144] D. Redmill and N. Kingsbury. Still image coding for noisy channels. In *IEEE International Conference Image Processing*, pages 95–99, 1994.

[145] A. Reibman. Optimizing multiple description video coders in a packet loss environment. In *Packet Video Workshop*, Pittsburgh, USA, March 2002.

[146] A. Reibman, H. Jafarkhani, Y. Wang, M. Orchard, and R. Puri. Multiple description video coding using motion-compensated prediction. In *IEEE International Conference Image Processing*, Kobe, Japan, October 1999.

[147] J. Rogers and P. Cosman. Robust wavelet zerotree image compression with fixed length packetization. *IEEE Signal Processing Letters*, 5(5):105–107, May 1998.

[148] P. Sagetong and A. Ortega. Optimal bit allocation for channel-adaptive multiple description coding. In *Image and Video Communication and Processing*, San Jose, CA, January 2000.

[149] P. Sagetong and A. Ortega. Analytical model-based bit allocation for wavelet coding with applications to multiple description coding and region of interest. In *IEEE International Conference in Multimedia & Expo*, Tokyo, Japan, August 2001.

[150] A. Said and W. Pearlman. A new, fast, and efficient image coded based on set partitioning in hierarchical trees. *IEEE Transaction on Circuits and Systems for Video Technology*, 6(3):243–250, 1996.

[151] S. Servetto, K. Ramchandran, and M. Orchard. Image coding based on morphological representation of wavelet data. *IEEE Transaction on Image Processing*, 1996.

[152] S. Servetto, K. Ramchandran, V. Vaishampayan, and K. Nahrstedt. Multiple description wavelet based image coding. In *IEEE International Conference Image Processing*, Chicago, USA, October 1998.

[153] S. Servetto, K. Ramchandran, V. Vaishampayan, and K. Nahrstedt. Multiple description wavelet based image coding. *IEEE Transaction on Image Processing*, 9(5):813–826, 2000.

[154] S. Servetto, J. Rosenblatt, and K. Ramchandran. A binary Markov model for the quantized wavelet coefficients of images and its rate/distortion optimization. In *IEEE International Conference Image Processing*, Santa Barbara, CA, 1997.

[155] S. Servetto, V. Vaishampayan, and N. Sloane. Multiple description lattice vector quantization. In *IEEE Data Compression Conference*, Snowbird, UT, 1999.

[156] C. Shannon. A mathematical theory of communication. *Bell System Technical Journal*, 27:379–423, 623–656, July, October 1948.

[157] C. Shannon. Coding theorems fos a discrete source with fidelity criterion. *IRE Nat. Conv. Rec.*, 7(4):142–163, March 1959.

[158] J. Shapiro. Embedded image coding using zerotrees of wavelets coefficients. *IEEE Transaction on Signal Processing*, 41(12):3445–3462, December 1993.

[159] P. Sherwood and K. Zeger. Progressive image coding for noisy channels. *IEEE Signal Processing Letters*, 4:189–191, July 1997.

[160] P. Sherwood and K. Zeger. Error protection for progressive image transmission over memoryless and fading channels. *IEEE Transaction on Communication*, 46(12), December 1998.

[161] Y. Shoham and A. Gersho. Efficient bit allocation for an arbitrary set of quantizers. *IEEE Transaction on Acoustics, Speech, and Signal Processing*, 36(9):1445–1453, 1988.

[162] S. Somasundaram and K. Subbalakshmi. Exploiting path diversity and forward error correction for robust transmission of images. In *Picture Coding Symposium*, pages 259–262, Saint-Malo, France, April 2003.

[163] K. Stuhlmuller, M. Link, and B. Girod. Scalable Internet video steaming with unequal error protection. In *Packet Video Workshop*, NY, US, April 1999.

[164] G. Sullivan. Efficient scalar quantization of exponential and laplacian random variable. *IEEE Transaction on Information Theory*, 42:1365–1374, September 1996.

[165] W. Sweldens. The lifting scheme: A new philosophy in biorthogonal wavelet constructions. In A. Laine and M. Unser, editors, *Wavelet Applications in Signal and Image Processing III*, pages 68–79. SPIE 2569, 1995.

[166] N. Tanabe and N. Farvardin. Subband image coding using entropy coded quantization over noisy channels. *IEEE Journal Selected Areas on Communication*, 10:926–943, June 1992.

[167] D. Taubman and M. Marcellin. *Jpeg2000 : Image Compression Fundamentals, Standards, and Practice*, volume 642 of *Kluwer International Series in Engineering and Computer Science*. Kluwer Academic Publishers, Boston, November 2001.

[168] D. Taubman and A. Secker. Highly scalable video compression with scalable motion coding. In *IEEE International Conference Image Processing*, Barcelone, Spain, September 2003.

[169] D. Taubman and A. Zakhor. Multirate 3-D sbband coding of video. *IEEE Transaction on Image Processing*, 3(5):572–588, 1994.

[170] L. Torres and E. Delp. New trends in image and video compression. In *EURASIP European Signal Processing Conference*, Tampere, Finland, September 2000.

[171] P. Vaidyanathan. *Multirate systems and Filter Banks*. Englewood Clifs, NJ, Prentice-Hall, 1993.

[172] V. Vaishampayan. Vector quantizer design for diversity systems. In *CISS*, 1991.

[173] V. Vaishampayan. Design of multiple description scalar quantizers. *IEEE Transaction on Information Theory*, 39(3):821–834, 1993.

[174] V. Vaishampayan and J. Batllo. Asymptotic analysis of multiple description quantizers. *IEEE Transaction on Information Theory*, 44(1):278–283, January 1998.

[175] V. Vaishampayan and J. Domaszewicz. Design of entropy-constrained multiple description scalar quantizers. *IEEE Transaction on Information Theory*, 40(1):245–250, 1994.

[176] V. Vaishampayan and S. John. Interframe balanced-multiple-description video compression. In *Packet Video Workshop*, Columbia University, NY, 1998.

[177] V. Vaishampayan and S. John. Interframe balanced-multiple-description video compression. In *IEEE International Conference Image Processing*, October 1999.

[178] V. Vaishampayan, N. Sloane, and S. Servetto. Multiple description vector quantization with lattice codebooks: Design and analysis. *IEEE Transaction on Information Theory*, 47(5):1718–1734, 2001.

[179] R. Venkataramani, G. Kramer, and V. Goyal. Bounds on the achievable region for certain multiple description coding problems. In *International Symposium Information Theory*, Washington, DC, 2001.

[180] M. Vishwanath. The recursive pyramid algorithm for the discrete wavelet transform. *IEEE Transaction on Signal Processing*, 42:673–676, March 1994.

[181] A. Wang, Z. Xiong, P. Chou, and S. Mehrotra. Three-dimensional wavelet coding of video with global motion compensation. In *IEEE Data Compression Conference*, pages 404–414, Snowbird, Utha, March 1999.

[182] Y. Wang and D.-M Chung. Robust image coding and transport in wireless networks using nonhierarchical decomposition. In *3 rd International Workshop Mobile Multimedia Communications*, New Brunswick, NJ, September 1996.

[183] Y. Wang, M. Orchard, and A. Reibman. Multiple description image coding for noisy channels by paring transform coefficients. In *IEEE International Workshop on Multimedia Signal Processing*, 1997.

[184] Y. Wang, M. Orchard, and A. Reibman. Optimal pairwise correlating transform for multiple description coding. In *IEEE International Conference Image Processing*, Chicago, IL, October 1998.

[185] Y. Wang, Q.-F Zhu, and L. Shaw. Maximally smooth image recovery in transform coding. *IEEE Transaction on Communication*, 41:1544–1551, October 1993.

[186] S. Wenger, G. Knorr, J. Ott, and F. Konssentini. Error resilience support in H.263+. *IEEE Transaction on Circuits and Systems for Video Technology*, pages 867–877, November 1998.

[187] H. Witsenhausen. On source networks with minimal breakdown degradation. *Bell System Technical Journal*, 59(6):1083–1087, July-August 1980.

[188] W. Witsenhausen and A. Wyner. Source coding for multiple descriptions II: A binary source. *Bell System Technical Journal*, 60(10):2281–2292, December 1981.

[189] J. Wolf, A. Wyner, and J. Ziv. Source coding for multiple descriptions. *Bell System Technical Journal*, 59(8):1417–1427, October 1980.

[190] Z. Xiong, K. Ramchandran, and M. Orchard. Space-frequency quantization for wavelet image coding. *IEEE Transaction on Image Processing*, 6(5):677–693, May 1997.

[191] J. Xu, S. Li, Z. Xiong, and Y.-Q Zhang. On boundary effects in 3-D wavelet video coding. In *Symposium on Optical Science and Technology*, San Diego, US, Juillet 2000.

[192] J. Xu, S. Li, Y.-Q Zhang, and Z. Xiong. A wavelet video coder using three dimensional embedded subband coding with optimized truncation (3D ESCOT). In *IEEE Pacific-Rim Conference on Multimedia*, Sydney, Australia, December 2000.

[193] S. Yang and V. Vaishampayan. Low delay communication for Rayleigh fading channels: An application of the multiple description quantizer. *IEEE Transaction on Communication*, 43:2771–2783, November 1995.

[194] X. Yang and K. Ramchandran. Optimal subband filter banks for multiple description coding. *IEEE Transaction on Information Theory*, 46(7):2477–2490, November 2000.

[195] Y.Wang and Q. Zhu. Error control and concealment for video communication: A review. *Proc. of the IEEE*, 86(5):974–997, May 1998.

[196] R. Zamir. Gaussian codes and Shannon bounds for multiple descriptions. *IEEE Transaction on Information Theory*, 43(5):2629–2635, November 1999.

[197] R. Zamir. Shannon-tupe bounds for multiple descriptions of a stationary source. *Journal of Combinatorics, Information & System Sciences*, pages 1–15, 2000.

[198] M. Zorzi, R. Rao, and L. Milstein. On the accurancy of a first-order markov model for data transmission on fading channels. In *ICUPC*, November 1995.

[199] Z.Zhang and T. Berger. New results in binary multiple description. *IEEE Transaction on Information Theory*, IT-33(4):502–521, July 1987.

Codage à Description Multiple d'Images et de Vidéos pour des canaux bruités

Résumé

Les travaux développés dans cette thèse apportent un nouveau regard sur les techniques de codage par descriptions multiples (MDC).

Nous proposons une méthode (MDBA) de codage MDC source/canal conjoint robuste adaptée à la transmission d'images et de vidéos sur des canaux non stationnaires. Le principal avantage de cette méthode est qu'elle est bien adaptée pour des applications de transmission sur des canaux peu fiables et variables dans le temps. De plus, grâce à ses capacités de compression et de synchronisation, elle permet de faire de la transmission en temps réel.

Nous montrons que la méthode proposée présente les meilleurs résultats en terme de rapport signal à bruit et de qualité visuelle lorsqu'on la compare avec d'autres méthodes de descriptions multiples issues de l'état de l'art. De plus, elle s'avère bien adaptée pour des applications où les méthodes standard de contrôle d'erreur ne sont pas capables de s'adapter facilement aux caractéristiques du canal.

La méthode est validée sur différents modèles de canal (BSC, AWGN, Internet, UMTS, satellite) dans le cadre de la transmission d'images et de vidéos.

La méthode MDBA proposée est bien adaptée pour des applications qui ont besoin d'un codeur qui utilise l'information venant d'un canal de "feedback" comme par exemple "peer-to-peer video conferencing", vidéo sans fil, etc, mais s'avère insuffisante quand on veut faire du "streaming" vidéo. C'est pour cette raison que nous avons développé un système dédié au "streaming" de vidéo. Ainsi ce manuscrit termine avec une extension de la méthode pour faire du "streaming" vidéo robuste aux erreurs du canal. La méthode proposée permet à la fois de s'adapter à bande passante du canal (débit variable) et de s'adapter aux erreurs de transmission liés au niveau de bruit présent sur le canal.

Mots-clefs: MDC, adaptabilité, scalabilité, codage robuste, transmissions variables, ondelettes, analyse multirésolution, allocation de débits.

Multiple Description Image and Video Coding for Noisy Channels

Abstract

This PhD thesis provides a new ways of looking multiple description methods (MDC).

We propose a robust joint source channel MDC method that is adapted to the transmission of images and videos on non stationary channels. The main advantage of this method is that it is well adapted for applications implying transmissions using unreliable and variable channels. Moreover, for its compression and synchronization capabilities it is appropriate for real time transmissions.

We show that the proposed method present the best results, in terms of signal to noise ratio and of visual quality, when compared with other multiple description methods existing in the state of the art. Moreover, it is well adapted for applications where the standard methods of error control cannot easily adapt to the characteristics of the channel.

The method is validated on various channel models (BSC, AWGN, Internet, UMTS, satellite) when considering image and video transmissions.

The proposed MDBA is well adapted for application demanding a feedback-based encoder, as peer-to-peer video conferencing, wireless video, etc, but it is insufficient in a real video streaming system where the server may be serving thousands of clients simultaneously. It is that the reason why we develop a system based in the MDBA but dedicated to video streaming. The last chapter of this dissertation presents an extension of the MDBA method to make streaming video robust to channel errors. The proposed method is rate adaptable, as the actual commercial methods, and it is also quality adaptable that is presented in the literature as an important characteristic of video streaming systems.

Keywords: MDC, adaptability scalability, robust coding, non stationary channels, wavelets, multiresolution analysis, bit allocation.