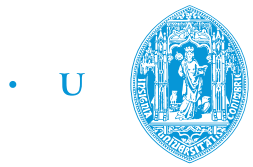

Universidade de Coimbra
Faculdade de Ciências e Tecnologia
Departamento de Engenharia Electrotécnica e de Computadores

Multiple Description Coding for Path Diversity Video Streaming

Pedro Daniel Frazão Correia

2012

Multiple Description Coding for Path Diversity Video Streaming

Pedro Daniel Frazão Correia

Submitted in partial fulfillment of
the requirements for the degree of
Doctor of Philosophy

Universidade de Coimbra
Faculdade de Ciências e Tecnologia
Departamento de Engenharia Electrotécnica e de Computadores

under supervision of

Prof. Dr. Vitor Manuel Mendes da Silva
Prof. Dr. Pedro A. Amado Assunção

In memory to my mother, Maria Celeste

To Sara, Samuel, Simão and Paula

Acknowledgments

This page evokes all of those who have collaborated, directly or indirectly in this work. First of all, I would like to thank to professors Pedro Assunção and Vitor Silva, my scientific supervisors, by their constant support, pragmatism and also by their fruitful discussions allowing me to pursue the research work with success.

My gratitude to Instituto de Telecomunicações, in special to Leiria branch, by providing me the physical, material and financial conditions necessary to develop the research activities.

I want to thank also to Instituto Politécnico de Tomar, my work place, for its support in order to combine the academic activities with the research work, and in this way to make possible this thesis.

My acknowledgment also to Fundação para a Ciência e a Tecnologia (FCT), that has supported this work with grants SFRH/BD/30087/2006 and SFRH/BD/50035/2009, last one under the program "Programa de Apoio à Formação Avançada de Docentes do Ensino Superior Politécnico" (PROTEC).

In a personal plan, I want to thank to my friends Ana Cristina, Ana Vieira and Gabriel Pires by their encouragement and motivation during this period. Also, I want to thank to Lino Ferreira, to his friendship and for his collaboration in part of presented work. My gratitude to my brothers of Missionary Community "Servidores do Evangelho", by their constant presence.

I sincerely want to thank to my close family, specially to my father Daniel, for all support and care that have given in all this time. Lastly, my deep gratefulness to my daughter Sara, my sons Samuel and Simão, and my wife Paula, for their constant comprehension, support and love.

I would like to dedicate this work in memory to my mother, Maria Celeste.

Agradecimentos

Esta página lembra todos aqueles que directa e indirectamente colaboraram para que este trabalho pudesse chegar a este ponto. Em primeiro lugar gostaria de agradecer aos professores Pedro Assunção e Vitor Silva, meus orientadores científicos, pelo incentivo constante, pelo seu pragmatismo e pelas frutuosas discussões que permitiram que o trabalho desenvolvido chegasse a bom porto.

O meu agradecimento ao Instituto de Telecomunicações, em especial à Delegação de Leiria, pelas condições físicas, materiais e financeiras proporcionadas para o desenvolvimento do trabalho de investigação.

Agradeço também ao Instituto Politécnico de Tomar, instituição que a pertenço como Assistente, pelas condições proporcionadas para conciliar a actividade docente com a actividade de investigação, conducente ao trabalho apresentado nesta tese.

Quero agradecer o apoio financeiro da Fundação para a Ciência e a Tecnologia (FCT) através das bolsas SFRH/BD/30087/2006 e SFRH/BD/50035/2009, esta última no âmbito do Programa de Apoio à Formação Avançada de Docentes do Ensino Superior Politécnico (PROTEC).

Num plano mais pessoal quero agradecer aos meus amigos, Ana Cristina, Ana Vieira e Gabriel Pires, pelo encoragamento e motivação ao longo de todo o período que decorreu este trabalho. Ao Lino Ferreira pela amizade e pela colaboração em parte do trabalho realizado. A minha gratidão a todos os irmãos da Comunidade Missionária Servidores do Evangelho, pela sua presença constante.

Quero também agradecer a toda a minha família, em especial ao meu pai Daniel, por todo o apoio e incentivo que dedicaram ao longo deste tempo. E por fim o meu agradecimento profundo, aos meus filhos, Sara, Samuel e Simão e à minha esposa Paula, por toda a sua compreensão, o seu apoio constante, pelo seu amor.

Dedico este trabalho em memória da minha mãe, Maria Celeste.

Abstract

In the current heterogeneous communications environments, the great variety of multimedia communications combined with fast evolution of networking architectures and topologies, give rise to new research problems related to the various elements of the communication chain. This includes, the ever present problem in video communications, which results from the need for coping with transmission errors and losses. In this context, video streaming with path diversity appeared as a novel communication framework, involving different technological fields and posing several research challenges. The research work carried out in this thesis is a contribution to robust video coding and adaptation techniques in the field of Multiple Description Coding (MDC) for multipath video streaming.

The thesis starts with a thorough study of MDC and its theoretical basis followed by a description of the most important practical implementation aspects presented in available literature. Additionally, a review of Multiple Description (MD) video coding is presented, where the relevant methods recently developed regarding this issue are explained, covering different video coding approaches. In MDC, a video signal is typically encoded into several independent descriptions, i.e., compressed streams, where each one can be delivered over different channels making use of path diversity communication environments. A high quality video representation is achieved when all descriptions are available at decoder, whereas a lower but still acceptable quality is obtained when only one description is received. These interesting features of MDC are investigated in this thesis and compared to classic Single Description Coding (SDC).

A research evaluation study of MDC for Advanced Video Coding (AVC) in regard to coding efficiency, distortion and error resilience is presented. Starting with open-loop MD video coding architectures, this study evaluates the effects of distortion propagation that happens when individual decoding of each description is performed. A novel multi-loop architecture for AVC is then proposed, which prevents drift distortion accumulation, generating a controlled amount of additional information. In particular, the MD video coding architecture is based on Multiple Description Scalar Quantization (MDSQ), including a new method for generating descriptions with different rates, *i.e.* unbalanced descriptions. This research extends the current state-of-the-art methods using balanced MDSQ, developing new MDC capabilities in different application scenarios without losing coding efficiency neither robustness.

This thesis also extends the current concept of multiple description coding (MDC) to the compressed domain, by proposing efficient splitting of standard single description coded (SDC) video into a multi-stream representation. A novel multiple description video splitting (MDVS) scheme was developed to operate at network edges, for increased robustness in path diversity video streaming across heterogeneous communications chains. The proposed scheme is able to effectively control drift distortion in both intra and inter predictive coding, even when only one description reaches the decoder. This is achieved by generating a controlled amount of relevant side information to compensate for drift accumulation, whenever any description is lost in its path, achieving significant quality improvement at reduced redundancy cost.

Additionally, and taking into account the new achievements, a novel high-accurate rate control method is proposed for MDSQ video coding. Known models based on linear relation between the output rate and the percentage of zeros of the quantized transform coefficients are extended to MDSQ video coding. In particular, this research demonstrates that the linear relation is maintained when MDSQ is used, making possible its use in linear rate control methods for MDC. Taking into account the balancing rate between descriptions, i.e. the percentage of the overall rate that is given to each description, this new method has the ability to choose the appropriate coding parameters to accurately achieve the target MD bitrate.

Overall, the new methods investigated in this thesis along with the good performance obtained from the experimental results, demonstrate the relevancy of its contribution to the field of MDC and practical usefulness in new robust multimedia services and applications using path diversity channels.

Resumo

Nos ambientes heterogéneos das redes de comunicação acuais, a grande variedade de tecnologias de rede, assim como a rápida evolução das suas arquiteturas e topologias, tem dado origem ao aparecimento de novos métodos que permitam solucionar os problemas de erros de transmissão e perdas. A difusão de vídeo com diversidade de canais é vista como um esquema interessante de comunicação para fazer face a este cenário, envolvendo vários aspetos tecnológicos e colocando também vários desafios para investigação. Enquadrando-se neste contexto, esta tese apresenta novas técnicas de codificação e adaptação de vídeo com robustez a erros usando Codificação de Vídeo com Múltiplas Descrições (MDC)

Esta tese faz um estudo das técnicas MDC existentes, abordando o assunto nos seus aspetos teóricos, como também descrevendo os aspectos relacionados com a sua implementação. Adicionalmente, a tese apresenta uma revisão da literatura das aplicações MDC no contexto de codificação de vídeo. Esta técnica de codificação consiste na representação de uma fonte de vídeo por vários fluxos independentes *i.e.*, descrições, podendo ser descodificados com uma qualidade aceitável. No entanto, estes fluxos podem ser conjugados entre si, obtendo-se uma qualidade de reconstrução mais elevada. Esta característica interessante do tópico MDC é investigada nesta tese, comparando o seu desempenho com a codificação tradicional de descrição única.

Neste trabalho, é realizada uma avaliação de desempenho da aplicação das técnicas MDC em Codificação Avançada de Vídeo(AVC), tendo em conta a eficiência de codificação, a distorção e a resiliência a erros. Partindo de arquiteturas MDC para vídeo em malha-aberta e do estudo da influência da propagação de distorção nos sinais descodificados, este trabalho propõe uma nova arquitetura multi-malha para AVC. Esta permite prevenir a propagação da distorção quando apenas uma descrição é descodificada de forma independente, gerando informação adicional redundante que permite compensar a propagação de erros, aumentando significativamente a qualidade de reconstrução das sequencias de vídeo. Em particular, a nova arquitetura permite gerar descrições múltiplas baseadas em Quantificação Escalar (MDSQ), cujos descrições possam ser codificadas com débitos distintos. Este novo método apresenta-se como uma evolução do estado da arte dos métodos MDSQ em codificação de vídeo, cujas de descrições são codificadas com o mesmo débito.

Esta tese apresenta também o conceito de MDC para o domínio comprimido. É proposto um novo método que permite que diferentes fluxos de vídeo possam ser gerados a partir

de um fluxo de vídeo comprimido normalizado. Este novo esquema deve operar em nós de rede que funcionem como pontos de adaptação, de modo a aumentar a robustez a erros em redes e serviços com diversidade de canais em ambientes heterogéneos. A arquitetura proposta, possui a capacidade de controlar a propagação de erros na descodificação de uma descrição, para os vários modos de predição existentes na codificação. É gerada informação adicional a partir do fluxo de vídeo existente, permitindo assim, o controlo eficiente da propagação de distorção na descodificação de uma descrição de forma independente.

Tendo como cenário de aplicação os novos métodos propostos, esta tese apresenta também um novo método de controlo de débito para codificação de vídeo usando MDSQ. O método proposto adapta os modelos conhecidos de controlo de débito baseados na relação linear entre o débito de saída e a percentagem de coeficientes quantificados de valor nulo resultantes da utilização de transformadas na codificação. Este trabalho mostra que a relação se mantém quando a operação MDSQ é colocada na malha de codificação, permitindo a implementação de métodos precisos de controlo de débito no âmbito de codificação robusta de vídeo com MDC. O método proposto tem em consideração o débito que cada descrição deve apresentar escolhendo os parâmetros de codificação apropriados de modo a atingir de forma precisa o débito global de saída desejado.

Tendo em conta o desempenho dos métodos propostos, é possível afirmar que este trabalho apresenta contributos relevantes na área de Codificação de Vídeo com Descrições Múltiplas, mostrando a sua utilidade em novos serviços e aplicações com robustez a erros onde a diversidade de caminhos possa ser utilizada.

Contents

1 Introduction **1**

1.1 Motivation . . . 3

1.2 Main Contributions . . . 4

1.3 Outline of the thesis . . . 6

2 Multiple Description Coding Techniques **9**

2.1 Theoretical Basis . . . 10

2.1.1 Single Description Rate-Distortion Region . . . 11

2.1.2 Multiple Description Rate-Distortion Region . . . 13

2.1.3 Multiple Description Bounds for a Gaussian Source . . . 16

2.1.4 Representation of MD Rate-Distortion Region . . . 18

2.1.5 Redundancy . . . 20

2.2 Practical Implementation . . . 22

2.2.1 Multiple Description with Correlating Transforms . . . 22

2.2.2 Multiple Description Sub-band Coding . . . 27

2.2.3 Polyphase Decomposition and Selective Quantization . . . 28

2.2.4 Multiple Description Scalar Quantisation . . . 29

2.2.5 MDC Based on Error Correction Codes . . . 33

3 Multiple Description Video Coding and Networking **35**

3.1 Basic Video Coding Techniques . . . 36

3.2 Multiple Description Video Coding-an overview . . . 40

3.2.1 MDC with pre-processing . . . 43

3.2.2 Embedded MDC . . . 46

3.2.3 High Dimensional MDC . . . 50

3.2.4 Multiple Description Scalable Coding . . . 51

3.2.5 Stereo/Multiview MDC . . . 52

3.2.6 Unbalanced MDC . . . 54

3.2.7 Multiple Description Distributed Video Coding . . . 55

3.3 Multiple Description Video Streaming . . . 56

3.3.1 MDC Video Streaming over Content Delivery Networks . . . 57

3.3.2 MDC Streaming over Peer-to-peer networks . . . 59

3.3.3 Scheduling in path diversity networks 62

4 MDSQ for Advanced Video Coding 65

4.1 Balanced MDSQ 66

4.2 Unbalanced MDSQ 67

4.3 MD Video Architecture 69

4.3.1 Open-loop MD Encoder 69

4.3.2 Open-loop MD Decoder 70

4.3.3 Drift Analysis 71

4.3.4 Multi-loop MD Encoder 74

4.3.5 Multi-loop MD Decoder 76

4.4 R-D Performance - Two Descriptions 77

4.4.1 Balanced MDSQ 77

4.4.2 Unbalanced MDSQ 84

4.5 R-D Performance - Single Description 91

4.5.1 Balanced MDSQ 91

4.5.2 Unbalanced MDSQ 101

4.6 Conclusions 106

5 Multiple Description Video Splitting of Coded Streams 107

5.1 MDVS Scenario 109

5.2 Classic MDVS 111

5.2.1 Drift Analysis 111

5.3 MDVS with drift compensation 113

5.3.1 MDVS architecture 113

5.3.2 Simplified MDVS 115

5.4 MDVS drift performance 118

5.4.1 Intra predicted frames 118

5.4.2 MC predicted frames 118

5.4.3 Generic regular GOP 119

5.4.4 The overall effect of side information 123

5.5 MDVS streaming with path diversity 124

5.6 MDVS vs MDC: comparative discussion 128

5.7 Conclusion . . . 129

6 MDC Rate Control and Priority Video Streaming 131

6.1 MDC Rate Control . . . 132
6.1.1 The ρ model for MDC . . . 133
6.1.2 Rate Control . . . 134
6.1.3 R-D Performance . . . 134
6.1.4 Rate Control Accuracy . . . 136
6.2 Priority MDC video streaming . . . 137
6.2.1 MD priority streaming scenario . . . 138
6.2.2 Packetization and Priority Streaming . . . 139
6.2.3 Optimal binary picture classification . . . 141
6.2.4 Simulation Results . . . 144
6.3 Conclusions . . . 147

7 Conclusions and future work 149

7.1 Conclusions . . . 149
7.2 Future Work . . . 152

A Published Papers 155

A.1 Journal Papers . . . 155
A.2 Conference Papers . . . 155

References 157

List of Tables

4.1 Balanced MDSQ,k=2 67
4.2 Unbalanced MDSQ, k=1 and $Z = 0$ 68

5.1 PSNR vs Side Information Redundancy (one description). 123
5.2 Average PSNR in transmission with different packet loss rates for 1Mbps@30Hz 128

6.1 Rate control accuracy 137

List of Figures

2.1 SDC rate-distortion encoder and decoder. 11
2.2 MD source coding model for $N = 2$ descriptions. 14
2.3 MD source coding model for $N = 2$ descriptions. 20
2.4 Lower Bounds on side distortions $D_1 = D2$ when the base rate is r and fixed distortion $D_0 = 2^{-2r}$ [Goyal 2001b]. 21
2.5 MDCT Generic Scheme. 24
2.6 Multiple Description Subband Coding. 27
2.7 MDC with polyphase transform and selective quantisation. 28
2.8 Generic MDSQ scheme. 29
2.9 MDSQ example 1 - (a) Matrix representation ; (b) Linear representation . 31
2.10 MDSQ example 2 -(a) Matrix representation; (b) Linear representation . . 31
2.11 MDC FEC. 34
2.12 MDC FEC. 34

3.1 Hybrid video encoder. 37
3.2 3-D sub-band video encoding. 38
3.3 3-D sub-band decomposition. 39
3.4 Distributed video encoding. 40
3.5 Generic MD encoder 41
3.6 Pre-processing MDC. 43
3.7 Open-loop multiple description video encoder. 46
3.8 Data Partitioning among channels 47
3.9 Tree-loop multiple description video encoder. 49
3.10 Path diversity in content delivery networks. 59
3.11 Path diversity by multiple-tree topology. 61
3.12 Path diversity using mesh topology. 62

4.1 Open-loop MDSQ encoder. 70
4.2 MDSQ open-loop video decoder. 71
4.3 Distortion accumulation within an intra frame - *Coastguard* 73
4.4 Distortion accumulation in MC predicted frames-*Coastguard* 74
4.5 MDSQ encoder with drift compensation. 75
4.6 MDSQ video decoder. 76
4.7 Central Distortion *Foreman* sequence, I Frame 78
4.8 Central Distortion *bus* sequence, I Frame 78

4.9 Redundancy of *Foreman* sequence, I Frame 80
4.10 Redundancy of *bus* sequence, I Frame 80
4.11 Central Distortion *Foreman* sequence, IPBB Frame 82
4.12 Central Distortion *bus* sequence, IPBB Frame 82
4.13 Redundancy of *Foreman* sequence, IPBB Frame 83
4.14 Redundancy of *bus* sequence, IPBB Frame 83
4.15 *Bus* sequence: QP=10;k=1 84
4.16 *Bus* sequence: QP=20;k=1 85
4.17 Central Distortion *foreman* sequence, Intra Frame 87
4.18 Central Distortion *bus* sequence, Intra Frame 87
4.19 Central Distortion *foreman* sequence, Intra Frame 88
4.20 Central Distortion *bus* sequence, Intra Frame 88
4.21 Central Distortion *foreman* sequence, IPBB Frame 89
4.22 Central Distortion *bus* sequence, IPBB Frame 89
4.23 Redundancy of *foreman* sequence, IPBB Frame 90
4.24 Redundancy of *bus* sequence, IPBB Frame 90
4.25 Side-distortion *Foreman* sequence, Intra Frame CIF, MDSQ with 2 diagonals 92
4.26 Side-distortion *bus* sequence, Intra Frame CIF, MDSQ with 2 diagonals . . 92
4.27 Side-distortion *Foreman* sequence, Intra Frame CIF, MDSQ with 3 diagonals 94
4.28 Side-distortion *Bus* sequence, Intra Frame CIF, MDSQ with 3 diagonals . . 94
4.29 Side Distortion *Foreman* sequence,Intra Frame CIF, MDSQ with 5 diagonals 95
4.30 Side Distortion *Bus* sequence,Intra Frame CIF, MDSQ with 5 diagonals . . 95
4.31 Side Distortion *Foreman* sequence, CIF 10HZ, MDSQ with 2 diagonals . . 97
4.32 Side Distortion *bus* sequence, CIF 10HZ, MDSQ with 2 diagonals 97
4.33 Side Distortion *Foreman* sequence, CIF 10HZ, MDSQ with 3 diagonals . . 99
4.34 Side Distortion *Bus* sequence, CIF 10HZ, MDSQ with 3 diagonals 99
4.35 Side Distortion *Foreman* sequence, CIF 15HZ, MDSQ with 5 diagonals . . 100
4.36 Side Distortion *Bus* sequence, CIF 15HZ, MDSQ with 5 diagonals 100
4.37 Side Distortion *foreman* sequence, IPBB Frame 102
4.38 Side Distortion *foreman* sequence, IPBB Frame 102
4.39 Side Distortion *foreman* sequence, IPBB Frame 103
4.40 Side Distortion *foreman* sequence, IPBB Frame 103
4.41 Side Distortion *Bus* sequence, IPBB Frame 104
4.42 Side Distortion *Bus* sequence, IPBB Frame 104
4.43 Side Distortion *Bus* sequence, IPBB Frame 105
4.44 Side Distortion *Bus* sequence, IPBB Frame 105

5.1 MDVS application scenario. 110
5.2 Classic MDVS scheme. 111
5.3 Two-loop MDVS architecture. 113
5.4 Single-loop MDVS architecture. 115
5.5 Equivalent single-loop MDVS architecture. 116
5.6 PSNR of Intra Frame macroblocks. 120
5.7 PNSR for MC predicted frames (IPPP...) for *coastguard* and *foreman*. . . . 120
5.8 PNSR for generic regular GOP (IPBBP...) for *coastguard*. 121
5.9 PNSR for generic regular GOP (IPBBP...) for *foreman*. 122
5.10 PNSR for generic regular GOP (IPBBP...) for *bus*. 122
5.11 Average PNSR for *bus* at 1.25 Mbit/s (Burst length BEL=4). 126
5.12 Average PNSR for *bus* at 1.25 Mbit/s (Burst length BEL=12). 126
5.13 Average PNSR for *bus* at 1.8 Mbit/s (Burst length BEL=4). 126
5.14 Average PNSR for *bus* at 1.8 Mbit/s (Burst length BEL=12). 126
5.15 Average PNSR for *bus* at 2.16 Mbit/s (Burst length BEL=4). 127
5.16 Average PNSR for *bus* at 2.16 Mbit/s (Burst length BEL=12). 127
5.17 Frame by frame PSNR for *bus* sequence 127

6.1 I frames 135
6.2 P frames 135
6.3 Unbalanced R-D performance 136
6.4 MD priority streaming scenario. 139
6.5 NAL packetization. 140
6.6 Priority MDC streaming scheme. 141
6.7 Simulation results for *Mother & daughter* sequence. 145
6.8 Simulation results for *Foreman* sequence. 145
6.9 Simulation results for *Bus* sequence. 146

List of Abbreviations

AVC	Advanced Video Coding
CBR	Constant Bit Rate
CDN	Content Delivery Networks
CIF	Common Intermediate Formate
DCT	Discrete Cosine Transform
DPCM	Differential Pulse Coding Modulation
DVC	Distributed Video Coding
DWT	Discrete Wavelet Transform
EZW	Embedded Zerotree Wavelet
FEC	Forward Error Correction
FGS	Fine Granularity Scalability
FMO	Flexible Macroblock Ordering
GOP	Group of Pictures
HP	High Priority
HVS	Human Visual System
IP	Internet Protocol
IPTV	Internet Protocol Television
JPEG	Joint Photographic Experts Group
KLT	Karhunen-Loeve Transform
LAN	Local Area Networks
LC	Layered Coding
LP	Low Priority

MC	Motion Compensation
MCTF	Motion Compensated Temporal Filtering
MD	Multiple Description
MD-FEC	Multiple Description with Forward Error Correction
MDC	Multiple Description Coding
MDCT	Multiple Description with Correlating Transforms
MDDVC	Multiple Description with Distributed Video Coding
MDSQ	Multiple Description Scalar Quantization
MDVS	Multiple Description Video Splitting
ME	Motion Compensation
MP	Matching Pursuits
MPEG	Moving Picture Expert Group
MV	Motion Vector
NALU	Network Abstraction Layer Unit
OBMC	Overlapped Block Motion Compensation
P2P	Peer-to-Peer
PCT	Pairwise Correlating Transform
PLR	Packet Loss Rate
PSNR	Peak to Peak Signal-to-Noise Ratio
QCIF	Quarter Common Intermediate Formate
QoS	Quality of Service
QP	Quatization Parameter
R-D	Rate-Distortion

R-S	Reed-Solomon
SD	Single Description
SDC	Single Description Coding
SVC	Scalable Video Coding
UEP	Unequal Error Protection
VBR	Variable Bit Rate
VoD	Video on Demand
WLAN	Wireless Local Area Networks
WZ	Wyner-Ziv

Chapter 1

Introduction

Contents

1.1 Motivation 3

1.2 Main Contributions 4

1.3 Outline of the thesis 6

In the current heterogeneous communications environments, the great variety of multimedia communications combined with fast evolution of networking architectures and topologies, give rise to new research problems related to the various elements of the communication chain. At network level, such as in Internet, in wireless LAN or in last generation wireless access networks, different quality of service constraints are used, namely bandwidth, delay and jitter which pose new problems to ensure an appropriate level of error robustness. On the other hand, different terminal capabilities and the convergence between traditional TV sets, Internet TV and video mobile streaming, puts new problems in content adaptation to specific features and constraints, such as processing capacity, memory, display size and power autonomy.

Additionally, new emerging video streaming services and applications using both centralized Content Delivery Networks(CDN) and distributed infrastructures, such as Peer-to-Peer (P2P) Networks have been developed. In this scenario, path diversity is posed as an efficient transmission scheme to give high adaptability to network reconfiguration, and also to provide error robustness and load balancing. In CDN schemes, a high number of access users leads to adopt streaming schemes with server diversity. On the other hand, P2P networks have to deal with problems related to network management and configuration. For instance, an user is a network receiver, and at the same time, can be a network

node to other peer. Also, each node can leave the network at any instant, forcing to a new network reconfigurations. Furthermore, taking into account the network dynamics, the massive dissemination of high quality multimedia contents and also the access ubiquity, new challenges to video coding methods in terms of efficiency, error resilience and quality of service constraints need to be solved.

Multiple Description Coding (MDC) has recently been given particular attention by the research community as a promising approach to improve the quality of multimedia streaming over error-prone networks with path diversity. In MDC, a video signal is typically encoded into several independent descriptions, i.e., compressed streams, where each one can be delivered over a separate channel making use of available path diversity. If joint decoding of all descriptions is done at the receiver, then the quality of the reconstructed signal is higher than that obtained from individual decoding of any single description. These interesting features of MDC are accomplished at the cost of higher coding rate, i.e., redundancy, when compared to classic single description coding (SDC) [Goyal 2001a]. This is essentially driven by networks with multiple available paths from the sender to the receiver (e.g., mesh and overlay networks) and multiple source coding representations (i.e., MDC) that go beyond the classical paradigm of SDC, where one source is encoded into one single representation [Frossard 2008].

The use of MDC in path diversity video networking is a competitive alternative with other network-adaptation schemes, such as scalable video coding (SVC) for robust streaming over time-varying networking conditions. In general, the main difference between SVC and MDC lies in the inter-dependency of SVC layers in contrast with independent MDC descriptions. MDC is also an interesting approach whenever retransmission schemes are not a viable solution, because, for instance, when transmission delay is not acceptable. Also comparing with the channel coding approaches usually refereed to as forward error correction codes (FEC), the MDC techniques are more efficient for high packet loss rates where long code sizes are needed, which creates the difficulties associated with delay overhead.

Further research of error robustness video coding and adaptation techniques are needed, in order to adapt the quality of decoded signals to the constraints and needs posed by network transmission infrastructure. Particularly, in the MDC field, novel coding and adaptation schemes are needed to improve the error robustness on live and pre-encoded video signals, aiming to rise the quality of decoded signals at each user terminal.

1.1 Motivation

Traditionally, the MDC schemes use description of equal importance due to the principle that each description can be independent decodable and an acceptable quality must be provided to the decoder. On the other hand, the heterogeneity of the networks and applications requires the adaptability of the source coding schemes in terms of the rate-distortion performances. Although several works have been proposed in order to optimally obtain the best coding parameters in order to obtain the minimum distortion based on channel conditions, these works are mainly focus on balanced descriptions, *i.e* with same rate. Therefore, the investigation on unbalanced MDC schemes and also rate control methods for MDC solutions have not been extensively exploited in the literature, which is an interesting topic to evaluate in order to cope with content delivery chains with different resources.

Whenever a coded video stream is processed, the predictive nature of video coding algorithms must be taken into account, because drift leads to distortion accumulation at the end-user decoding terminal. In MDC, the effect of drift can be explained as follows. In the absence of errors or data loss, all descriptions are decoded and the reconstructed blocks/frames are then used as reference for others by providing accurate predictions. However, if any description is lost, then the predictions reconstructed in the decoder do not match those originally used at encoder or used to encode the original SDC stream in case of compressed stream adaptation. This mismatch is the origin of drift by adding distortion to decoded video, which is further accumulated in the reconstruction loop and propagated throughout all subsequent predicted blocks. New problems need to be solved when the most recent advanced predictive techniques are expanded into MDC methods, which means that further research work needs to be done in order to propose new efficient MDC methods adapted to advanced video coding schemes.

On the other hand, the combination of MDC with path diversity has always been used in communication chains typically comprised of an uncompressed source signal feeding an MDC encoder, followed by multiple transmission paths to the receiver [Mao 2005], [Akyol 2007] or by streaming multiple complementary descriptions distributed across the edge servers of content delivery networks [Apostolopoulos 2002], [Ahuja 2008]. A shortcoming of such communication model is that it does not take into account the typical scenario of current heterogeneous networking, where single path routes co-exist with multiple paths in the same delivery chain.

The concept of multiple description video splitting of coded streams (MDVS) fills the existing gap in heterogeneous video communications where an SDC stream is transmitted over a single path network and then needs to be split into several MDC streams. This might be particularly useful at edge nodes to benefit from path diversity over different networks where multiple paths are available. Since MDVS operates on coded streams, any networking node with such processing capability can split an incoming SDC stream into the different outgoing paths that can be used from that particular network node to the end user terminal. A recent work highlighting the advantages of using MDVS for robust video streaming and to deal with handoff over wireless local area networks (WLAN) is presented in [Chen 2007b].

Since MDVS suffers from the same intrinsic problem of drift as MDC, i.e., accumulation of decoding distortion when any description is lost in the network, a novel research topic that needs to be addressed is the capability for limiting such type of distortion. New types of low processing complexity MDVS architectures must be investigated in order to efficiently generate MDC streams from the incoming SDC video streams.

Although the improving of error resilience in video transmission chains due to path diversity, intermediate nodes must be able to use scheduling algorithms in order to increase the quality of experience at user due to network congestions. Based on scheduling decisions, one network node needs to decide which packet must be dropped or should be re-transmitted. Usually these methods use rate-distortion optimization in order to maximize the overall decoding quality and define a prioritization set for the transmitted packets. Looking for to the video coding schemes, SVC is usually used for scheduling schemes because of its nature where different layers are inherently set with an hierarchial importance order. The research of scheduling and prioritization schemes for applications based on MDC is still an open issue and needs further investigation in order to find methods to better adapt the MDC error resilience capabilities to the dynamic nature of the path diversity video delivery networks.

1.2 Main Contributions

Drift-free multiple description for advanced video coding A multiple description coding scheme based on a multi-loop structure is proposed, which prevents drift distortion accumulation in both intra and inter predicted slices. The drift is compensated by

generating a controlled amount of side information used by the decoder whenever any description is lost. The experimental results show that predicted slices do not suffer from drift and their quality is significantly improved at reduced redundancy cost in comparison with the open-loop implementation.

Unbalanced Multiple Description Scalar Quantisation A new method is proposed to generate unbalanced descriptions using Multiple Description Scalar Quantisation (MDSQ). For this purpose, a new index assignment method is proposed based on new set of tables that dynamically generate two descriptions with different unbalancing rates including the traditional balanced case. The performance evaluation show that proposed scheme gives better rate-distortion performance comparing with balanced scheme with also a good performace in packet loss simulation scenarios.

Multiple Description Video Splitting An extension of the current concept of MDC to the compressed domain is proposed, by developing an efficient splitting of standard SDC video into a multi-stream representation. A novel Multiple Description Video Splitting (MDVS) scheme is proposed to operate at network edges, for increased robustness in path diversity video streaming across heterogeneous communications chains. The proposed scheme is able to effectively control drift distortion in both intra and inter predictive coding, even when only one description reaches the decoder. The simulation results show that any individual description can be decoded on its own without producing drift, achieving significant quality improvement at reduced redundancy cost.

Rate control for multiple description video coding The problem of rate control in unbalanced MDSQ is addressed and an extension of the ρ model is developed based on the evidence that the linear relationship between the rate of each description and the corresponding percentage of zeros of transform coefficients is maintained when moving from the SDC into the MDC domain. The simulation results show that the rate-distortion performance of the proposed unbalanced MDC rates is better than the equivalent balanced MDC. The rate control algorithm exhibits high accuracy for a given target bitrate and unbalanced ratio between two descriptions. The proposed method is useful in MDC video streaming over asymmetric transmission paths, where each description rate can be dynamically adapted to different channel conditions.

Optimal priority MDC video streaming for networks with path diversity A robust video streaming scheme for priority networks with path diversity is proposed, based on a combined approach of MDC with optimal picture classification into two priorities. A binary picture classification algorithm is proposed to define high (HP) and low (LP) priority network abstraction layer units (NALU), which in turn define the packet priorities. An optimization algorithm finds HP pictures based on minimization of the packet loss concealment distortion. The proposed algorithm is able to effectively improve the decoded video quality without increasing the MDC stream redundancy.

1.3 Outline of the thesis

This thesis is formed by seven chapters and wants to deal with the main aspects in MD video coding research area. The thesis is divided in two parts, in which the reader follows a comprehensive path of the main MDC topics and also the author contributions for this research area. The first part of the thesis is formed by Chapter 2 and Chapter 3 that describes the theory and the state of the art of MD coding respectively. The second part is formed by Chapters 4,5,6 and 7 in which are described the main contributions of this thesis.

Chapter 2 is focus on MDC techniques. First, the theoretical definitions and optimal rate-distortion bounds are explained. Following, the MDC techniques used in the literature are described, and includes, Multiple Description with Correlating Transforms (MDCT), Multiple Description Subband Coding, based on Polyphase Decomposition and Selective Quantisation, Multiple Description Scalar Quantisation (MDSQ), and MDC based on the Error Correction Codes.

Chapter 3 describes the main MDC video architectures are described for the different video coding schemes including scalable coding, multi-view and stereo video coding, and also for the MDC video applications with $N > 2$ descriptions. In the second part of the chapter, the different path diversity topologies and the MDC networking applications are described.

Chapter 4 presents the main author contributions for MDC in advanced video coding. First the overall MDSQ video architecture is presented. A rate-distortion evaluation of balanced MDSQ is explored, comparing the performance of the proposed architecture

with open-loop proposal. A new set of index assignment tables is proposed in order to generate unbalanced descriptions. Comparison between balanced and unbalanced schemes are done, taking into account, the overral R-D performance and also the side-information coding efficiency. Finally the study of error-resiliency features are presented for distinct packet-loss rates, and available channel bitrates.

Chapter 5 addresses MDVS of coded video. This chapter proposes a novel MDVS scheme based on MDSQ, using side information to control drift in both spatial and temporal prediction. The side information is generated from the original stream and its rate is controlled with an independent quantisation parameter which also controls redundancy. Then, a simplified architecture is devised to reduce the overall complexity in regard to the number of processing functions and memory requirements.

Chapter 6 is divided in two different parts. The first part deals with the problem of rate control in unbalanced MD video coding. A new method is proposed to ensure an asymmetric target bit rate in each description. An extension of the ρ model is developed based on the evidence that the linear relationship between the rate of each description and the corresponding percentage of zeros of transform coefficients is maintained when moving from the single description domain (SDC) into the MDC domain. In the second part proposes a robust video streaming scheme for priority networks with path diversity, based on a combined approach MDC with optimal picture classification into two priorities. An optimization algorithm finds high priority frames based on minimization of the packet loss concealment distortion. The proposed algorithm is able to effectively improve the decoded video quality without increasing the MDC stream redundancy.

Finally, Chapter 7 concludes the thesis and gives a set of topics for future work in the Multiple Description Video Coding.

Chapter 2

Multiple Description Coding Techniques

Contents

2.1 Theoretical Basis **10**

2.1.1 Single Description Rate-Distortion Region 11

2.1.2 Multiple Description Rate-Distortion Region 13

2.1.3 Multiple Description Bounds for a Gaussian Source 16

2.1.4 Representation of MD Rate-Distortion Region 18

2.1.5 Redundancy 20

2.2 Practical Implementation **22**

2.2.1 Multiple Description with Correlating Transforms 22

2.2.2 Multiple Description Sub-band Coding 27

2.2.3 Polyphase Decomposition and Selective Quantization 28

2.2.4 Multiple Description Scalar Quantisation 29

2.2.5 MDC Based on Error Correction Codes 33

This chapter presents a review of Multiple Description Coding (MDC) principles, where the relevant theoretical definitions and optimal rate-distortion bounds are explained. The main MDC techniques documented in the literature are described, including Multiple Description with Correlating Transforms (MDCT), Multiple Description Subband Coding, based on Polyphase Decompositon and Selective Quantisation, Multiple Description Scalar Quantisation (MDSQ), and MDC based on Error Correction Codes.

2.1 Theoretical Basis

The initial motivation to the forthcoming of MDC has been the need of high reliability in the telephone networks in order improve the mechanisms to handle outages. The use of path diversity is an efficient approach to provide robustness in transmission networks. In the case of telephone networks, path diversity was achieved using redundant standby transmission links. One way to improve the robustness of the network without using standby links was to split the speech signal among two different links or paths. Hence, in normal operation, the communication would be done with good quality, and when some link or path failed, the other channel link would give a communication without breaks with a lower, yet acceptable quality.

The theoretical formulation of this practical constraint was done by Ozarow in [Ozarow 1980], by defining MDC as *the source coding problem with two channels and three receivers*. The starting point is the Shannon rate-distortion theory, which is named as *Single Description (SD) rate-distortion region* in this context. The focus on this theory is the derivation of performance bounds, that is, determining the region of achievable points in the rate-distortion trade-off for certain limited statistical sources.

While traditional rate-distortion theory intends to find the minimum achievable rate for a given distortion, MDC rate-distortion approach increases the dimension of the problem and needs to find the optimal jointly achievable distortions for three different decoded representations that are encoded in two different descriptions of the source with the respective rates. This is different from the traditional rate-distortion formulation because the distortion achieved by joining two streams does not correspond to the Shannon limits for a single source encoded at the same rate as that of both descriptions added together.

Thus, the MDC theoretical characterization has not been formulated in a closed form, but only some bounds have been given to approximated it. This was done in [Gamal 1982], with the definition of the *Multiple Description (MD)* rate-distortion region, evaluated for Guassian sources and quadratic distortion. In MDC context, if each description is individually encoded, each one with a different rate, the performance rate-distortion region corresponds to the combination of achievable distortions of each description individually decoded, and at the same time, to get a minimum as possible distortion when all descriptions are merged. This formulation is presented in section 2.1.2.

These optimal bounds of MD rate-distortion region cannot be established for situations

of practical relevance. Moreover, in general these bounds provide useful information to benchmark practical implementations of MDC techniques.

2.1.1 Single Description Rate-Distortion Region

In order to better understand the fundamental MDC theoretical statements, some basic definitions of rate-distortion theory are explained following the approach used in [Cover 2006]. The main concern of rate-distortion theory is how to represent a source with the minimum possible number of bits for a given reproduction quality. This formulation requires a measure of the source fidelity which leads to the definition of a distortion function based on a distance metric between the source and its coarse representation.

Fig.2.1 shows the general rate-distortion model for a SD coding scheme. The source signal is represented by a finite alphabet χ in the form of a vector of n independent identically distributed *(i.i.d.)* random variables $X_1, X_2, ..., X_n$ with a probability mass function $p(x), x \in \chi$. The encoder describes the source vector X^n by an index $f_n(X^n) \in \{2^0, 2^1, ..., 2^{nR}\}$, where nR is the *rate* in bits per alphabet symbol. The decoder represents X^n by an estimate $\hat{X}^n \in \hat{\chi}$. Usually, the reproduction alphabet $\hat{\chi}$ is the same as χ.

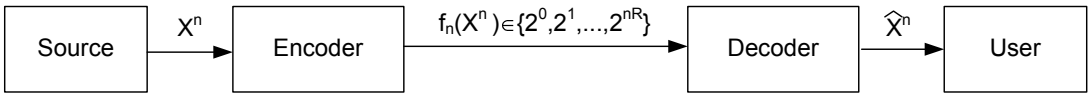

Figure 2.1: SDC rate-distortion encoder and decoder.

A distortion measure is a mapping function,

$$d : \chi \times \hat{\chi} \rightarrow \Re^+ \tag{2.1}$$

from the set of source pairs reproduction into the set of non-negative real numbers. The distortion $d(x, \hat{x})$ measures the cost of representing symbol x by symbol $\hat{x}$.

In general, several different distortion metrics can be used, but usually the squared-error, $d(x, \hat{x}) = (x - \hat{x})^2$ is chosen for rate-distortion analysis. This distortion function can be extended for vectors of symbols as follows:

$$d(x^n, \hat{x}^n) = \frac{1}{n} \sum_{i=1}^{n} d(x_i, \hat{x}_i) \tag{2.2}$$

The distortion measure for a sequence is the average of the per symbol distortion taking all the elements of the sequence into account.

A rate-distortion code defined as $(2^{nR}, n)$ consists in an encoding function,

$$f_n = \chi^n \rightarrow \{2^0, 2^1, ..., 2^{nR}\}, \tag{2.3}$$

and a decoding (reproduction) function,

$$g_n = \{1, 2, ..., 2^{nR}\} \rightarrow \hat{\chi}^n. \tag{2.4}$$

The distortion associated with $(2^{nR}, n)$ code is

$$D = E[d(X^n, g_n(f_n(X^n)))] = \sum_{x^n} p(x^n) d(x^n, g(f(x^n))), \tag{2.5}$$

where the set of n-tuples $g(1)$,$g(2)$,...,$g(2^{nR})$, denoted by $\hat{X}(1)$, $\hat{X}(2)$, $\hat{X}(2^{nR})$ constitutes the *codebook* and $f^{-1}(1)$,$f^{-1}(2)$,...,$f^{-1}(2^{nR})$ are the associated *assignment regions*.

A rate-distortion pair (R, D) is said to be *achievable* if there exists a sequence of $(2^{nR}, n)$ rate-distortion codes (f, g) such that

$$\lim_{n \rightarrow \infty} E[d(X^n, g_n(f_n(X^n)))] \leq D. \tag{2.6}$$

After defining the *achivability* concept for a given (R, D) pair, the definition of *SD rate-distortion region* for a source is the closure of the set of the achievable rate-distortion pairs (R, D).

Finally, let the *rate-distortion function* $R(D)$ be the *infimum* of rates R such that (R, D) lies in the rate-distortion region of the source for a given distortion D.

Another way of looking the *rate-distortion function* is with the definition of the *distortion-rate function* $D(R)$ which is the *infimum* of all distortions D such that (R, D) is in the *rate-distortion region* of the source for a given R.

Functions $R(D)$ and $D(R)$ both provide equivalent definitions for dealing with the boundary of the *rate-distortion region*.

The *rate-distortion function* for an i.i.d. source X with distribution $p(x)$ and distortion function $d(x^n, \hat{x}^n)$ is equal to

$$R(D) = \inf I(X; \hat{X}) \tag{2.7}$$

where

$$I(X; \hat{X}) = \sum_{x,\hat{x}} p(x, \hat{x}) \log(p(x, \hat{x})/p(x)p(\hat{x})), \tag{2.8}$$

is the *mutual information* between X and $\hat{X}$. The *infimum* is taken over all joint probability mass functions $p(\hat{x}, x)$ such that

$$\sum_{x,\hat{x}} p(x, \hat{x}) d(x, \hat{x}) \leq D. \tag{2.9}$$

2.1.2 Multiple Description Rate-Distortion Region

The SD rate-distortion definitions can be extended for MDC. The main difference lies in the definition of a quintuple (R_1,R_2,D_0,D_1,D_2) instead of a single pair (R, D).

Fig. 2.2 shows the generic MD scenario for two descriptions, originally named as the channel-splitting problem [Ozarow 1980]. The sequence of source symbols X^n is encoded into two redundant and mutually refinable descriptions represented by indices $f_{1n}(X^n) \in \{1, 2, ..., 2^{nR1}\}$ and $f_{2n}(X^n) \in \{1, 2, ..., 2^{nR2}\}$, which are transmitted separately over two different noiseless or error corrected channels: *Channel 1* and *Channel 2*. Three decoders are used at the receiver side. *Decoder 0* (*Central Decoder*) is the one that receives the information from both channels and outputs the estimate of the source X^n, $\hat{X_0}^n$. The remaining decoders (*Decoder*1 and *Decoder*2) receive information from each individual channel and are named as *Side Decoders*. The outputs are $\hat{X_1}^n$ and $\hat{X_2}^n$, respectively.

A $(2^{nR1}, 2^{nR2}, n)$ rate-distortion code consists of an encoding function,

$$f_{1n} = \chi^n \rightarrow \{2^0, 2^1, ..., 2^{nR1}\}; \tag{2.10}$$

$$f_{2n} = \chi^n \rightarrow \{2^0, 2^1, ..., 2^{nR2}\}; \tag{2.11}$$

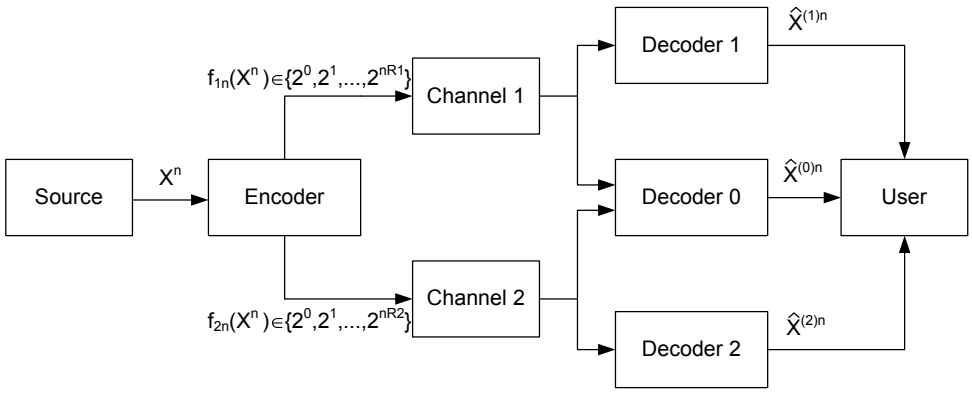

Figure 2.2: MD source coding model for $N = 2$ descriptions.

The decoding functions are defined as,

$$g_{0n} = \{\{1, 2, ..., 2^{nR1}\}, \{2^0, 2^1, ..., 2^{nR2}\}\} \rightarrow \hat{\chi_0}^n, \tag{2.12}$$

$$g_{1n} = \{2^0, 2^1, ..., 2^{nR1}\} \rightarrow \hat{\chi_1}^n, \tag{2.13}$$

$$g_{2n} = \{2^0, 2^1, ..., 2^{nR2}\} \rightarrow \hat{\chi_2}^n. \tag{2.14}$$

This MD model produces three different approximations of the sequence of source symbols X^n, which implicitly requires the definition of three distortion measures, one for each of the three decoded reconstructions $\hat{X_i}^n$ produced by decoder i, $i = 0, 1, 2$, in regard to the same sequence of source symbols X^n.

The distortion measure fuctions for each reconstruction $i = 0, 1, 2$ is defined as

$$d_i(x^n, \hat{x_i}^n) = \frac{1}{n} \sum_{k=1}^{n} d(x_k, \hat{x}_{ik}) \quad i = 0, 1, 2. \tag{2.15}$$

Consequently, the distortion associated to $(2^{nR1}, 2^{nR2}, n)$ codes are defined as the following set of distortions

$$D_0 = E[d_0(X^n, g_{0n}(f_{1n}(X^n)), f_{2n}(X^n)))], \tag{2.16}$$

$$D_1 = E[d_1(X^n, g_{1n}(f_{1n}(X^n)))], \tag{2.17}$$

$$D_2 = E[d_2(X^n, g_{2n}(f_{2n}(X^n)))], \tag{2.18}$$

where D_0 is defined as *central distortion* and $D_i, i = 1, 2$ are the *side distortions*. A rate pair (R_1, R_2) is said to be *achievable* for distortion $\mathbf{D} = (D_0, D_1, D_2)$, if there exists a sequence of $(2^{nR1}, 2^{nR2}, n)$ rate-distortion codes

$$\lim_{n\to\infty} E[d_0(X^n, g_{0n}(f_{1n}(X^n), f_{2n}(X^n)))] \leq D_0. \tag{2.19}$$

$$\lim_{n\to\infty} E[d_1(X^n, g_{1n}(f_{1n}(X^n)))] \leq D_1, \tag{2.20}$$

$$\lim_{n\to\infty} E[d_2(X^n, g_{1n}(f_{2n}(X^n)))] \leq D_2. \tag{2.21}$$

The *MD rate-distortion region* for a source is defined as the closure of the set of achievable rate pairs (R_1, R_2) whose reconstructions achieve expected distortions $\mathbf{D} = (D_0, D_1, D_2)$.

Inversely, the MD rate-distortion region is regarded as the achievable set of distortions $\mathbf{D} = (D_0, D_1, D_2)$ with rate pairs (R_1, R_2) that satisfies the following cumulative conditions,

$$D_0 \geq D(R_1 + R_2), \tag{2.22}$$

$$D_1 \geq D(R_1), \tag{2.23}$$

$$D_2 \geq D(R_2). \tag{2.24}$$

Achieving equality simultaneously in three equations (2.22), (2.23) and (2.24) would imply that an optimal rate $R1 + R2$ description can be partitioned into optimal rate R_1 descriptions and R_2 descriptions. Unfortunately, this is not possible because the individual descriptions with rates $R1$ and $R2$ are similar and redundant when combined. Thus, the MD rate-distortion problem differs from the SD classical rate-distortion since it is impossible to satisfy the limits $R_1 = R(D_1)$, $R_2 = R(D_2)$, and at the same time, obtain from the central decoder distortion D_0 such that $R_1 + R_2 = R(D_0)$. The main problem is to find appropriate descriptions sufficiently different, with reduced redundancy when combined together.

2.1.3 Multiple Description Bounds for a Gaussian Source

The definition of the MD rate-distortion region is in general a complex problem, highly dependent of the source statistics and number of descriptions. The characterization of the set of achievable distortions for a Gaussian source described by two descriptions was made by Ozarow in [Ozarow 1980]. For a memoryless Gaussian source of variance σ^2 and squared error-distortion measure, the set of achievable distortions $\mathbf{D}(\sigma^2, R_1, R_2)$ is the one that satisfies the equations [Ozarow 1980].

$$D_1 \geq \sigma^2 2^{-2R_1}, \tag{2.25}$$

$$D_2 \geq \sigma^2 2^{-2R_2}, \tag{2.26}$$

$$D_0 \geq \sigma^2 2^{-2(R_1+R_2)} \gamma_D, \tag{2.27}$$

where,

$$\gamma_D = \begin{cases} \frac{1}{1-(\sqrt{\Pi}-\sqrt{\Delta})^2}, & D_1 + D_2 < \sigma^2(1 + 2^{-2(R_1+R_2)}) \\ 1, & otherwise \end{cases} \tag{2.28}$$

Π and Δ are defined as,

$$\Pi = (1 - D_1/\sigma^2)(1 - D_2/\sigma^2), \tag{2.29}$$

and

$$\Delta = (D_1 D_2)/\sigma^4 - 2^{-2(R_1+R_2)}. \tag{2.30}$$

The bounds (2.25) and (2.26) are referring to the side-distortion and follows from equations (2.23) and (2.24). The inequality (2.27) means that the central distortion must exceed the rate-distortion bound by the factor γ_D.

The behavior of γ_D was clarified in [Goyal 2001b] for three different cases, considering $\sigma^2 = 1$:

- (a) Both descriptions are individually good;
- (b) Joint description is as good as possible;

- (c) Balanced descriptions, where $R_1 = R_2$ and $D_1 = D_2$.

(a) Both descriptions are individually good Assuming that both descriptions are individually good, distortions D_1 and D_2 are set by minimum distortion given by equations 2.23) and (2.24, which means that $D_1 = 2^{-2R_1}$ and $D_2 = 2^{-2R_2}$. Replacing D_1 and D_2 in equations 2.29 and 2.30, $\gamma_D = \frac{1}{1-(1-D_1)(1-D_2)}$ and $\Delta = 0$ respectively.

Consequently,

$$D_0 \geq D_1 D_2 \frac{1}{1-(1-D_1)(1-D_2)} = \frac{D_1 D_2}{D_1 + D_2 - D_1 D_2}. \tag{2.31}$$

A further chain of inequalities gives $2D_0 \geq min\{D_1, D_2\}$, which means that the joint description is only slightly better than the two individual descriptions. (**provar esta equação**)

(b) Joint description is as good as possible Considering that joint description is as good as possible, this means that $D_0 = 2^{-2(R_1+R_2)}$. In that case $\gamma_D = 1$, and thus

$$D_1 + D_2 \geq 1 + 2^{-2(R_1+R_2)}. \tag{2.32}$$

An unitary distortion value for a gaussian source means that $R(D) = 0$, which means that no information is sent, and the reconstruction is made using its mean [Cover 2006]. From the right side of inequality of equation (2.32), one can conclude that at least, one of the side decoders have a very poor reconstruction.

Balanced Descriptions The last case is considering balanced descriptions where $R_1 = R_2$ and $D_1 = D_2$. Two different cases are important to be mentioned in order to qualitatively understand the MD rate-distortion bounds. The first relates the exponential decay of the three distortions, while the second bounds the side distortion based on the gap between D_0 and $2^{-2(R_1+R_2)}$.

Assuming $R_1 = R_2 \gg 1$ and $D_1 = D_2 \ll 1$, then,

$$\gamma_D = \frac{1}{1-\left((1-D_1)-\sqrt{D_1^2-2^{-4R_1}}\right)^2} \approx \frac{1}{1-((1-D_1)-D_1)^2} = \frac{1}{4D_1-4D_1^2} \approx \frac{1}{4D_1}. \tag{2.33}$$

With $\gamma_D = (4D_1)^{-1}$ in (2.27) the central distortion is

$$D_0 \geq \frac{2^{-4(R_1)}}{4D_1} \tag{2.34}$$

and the product of central and side distortions D_0D_1 is approximately lower-bounded by $4^{-1}2^{-4R_1}$. With $D_1 = D_2 \approx 2^{-2(1-\alpha)R_1}$ and $0 \leq \alpha \leq 1$, the best central distortion exponential decay is

$$D_0 \approx \frac{2^{-2(1-\alpha)R_1}}{4} \tag{2.35}$$

This result shows that the penalty in the exponential decay of D_1 inserted by the parameter α corresponds to the increase of the rate decay of D_0, which means that increasing the side balanced distortions reduce the corresponding central one.

2.1.4 Representation of MD Rate-Distortion Region

Instead of represent MD rate-distortion region in function of the minimum distortion for a given rate as defined in equations (2.25), (2.26) and (2.27), this can be done describing the dual problem which is representing the rate-distortion region, by finding the minimum rates for each description a given a distortion value. Thus, given the quadratic distortions $\mathbf{D} = (D_0, D_1, D_2)$, the set of admissible rate pairs R_1, R_2 are defined as [Zamir 1999],

$$R_1 \geq \frac{1}{2}\log(\frac{\sigma^2}{D_1}), \tag{2.36}$$

$$R_2 \geq \frac{1}{2}\log(\frac{\sigma^2}{D_2}), \tag{2.37}$$

$$R_1 + R_2 \geq \frac{1}{2}\log(\frac{\sigma^4}{D_1D_2}) + \delta, \tag{2.38}$$

where $\delta = \delta(\sigma^2, D_0, D_1, D_2)$ is defined as

$$\delta = \begin{cases} \frac{1}{2}\log(\frac{1}{1-\rho^2}), & D_0 \leq D_{0max} \\ 0, & D_0 \geq D_{0max} \end{cases} \tag{2.39}$$

and represents the excess rate in each description comparing with a rate-distortion functions for a Gaussian source $X \sim N(0, \sigma^2)$ at distortions D_1 and D_2.

$$\rho = \begin{cases} -\frac{\sqrt{\pi\varepsilon_0^2+\gamma}-\sqrt{\pi\varepsilon_0^2}}{(1-\varepsilon_0)\sqrt{(\varepsilon_1\varepsilon_2)}} & D_0 \leq D_{0max} \\ -\sqrt{\frac{\pi}{\varepsilon_1\varepsilon_2}}, & otherwise \end{cases} \tag{2.40}$$

$$\gamma = (1-\varepsilon_0)[(\varepsilon_1-\varepsilon_0)(\varepsilon_2-\varepsilon_0)+\varepsilon_0\varepsilon_1\varepsilon_2-\varepsilon_0^2], \tag{2.41}$$

$$\pi = (1-\varepsilon_1)[(1-\varepsilon_2), \tag{2.42}$$

$$\varepsilon_i = \frac{D_i}{\sigma^2}, \quad for \quad i = 0, 1, 2, \tag{2.43}$$

and

$$D_{0max} = \frac{D_1 D_2}{D_1 + D_2 - \frac{D_1 D_2}{\sigma^2}}. \tag{2.44}$$

Fig. 2.3 shows the typical form of the achievable rates $R(\sigma^2, D_0, D_1, D_2)$ for $D_0 \leq D_{0max}$. δ, $\gamma \geq 0$ and $-1 \leq \rho \leq 0$, for all $D_1, D_2 \leq \sigma^2$ and $D_0 \leq D_{0max}$.

The quantity $\delta = \delta_1 + \delta_2 \geq 0$ is defined as *excess marginal rate* where

$$\delta_i = R_i - \frac{1}{2}\log(\frac{\sigma^2}{D_i}) \quad i = 1, 2. \tag{2.45}$$

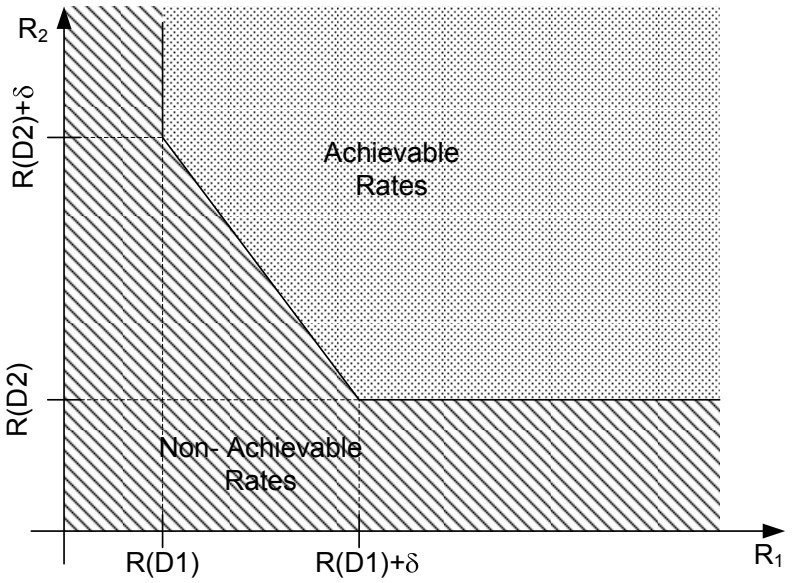

Figure 2.3: MD source coding model for $N = 2$ descriptions.

If D_1 and D_2 are held fixed, then the maximum central distortion D_{0max} can be achieved without excess marginal rate, *i.e* with $\delta = 0$, and

$$R_1 + R_2 = R(D_0). \tag{2.46}$$

2.1.5 Redundancy

Following [Goyal 2001b], the tradeoff between central and side distortions is interpreted by rearranging equation (2.27), considering that $D_1 = D_2$ and $\sigma^2 = 1$. Consequently,

$$D_1 \geq \min\{\frac{1}{2}[1 + D_0 - (1 - D_0)\sqrt{1 - 2^{-2(R_1+R_2)}/D_0}], 1 - \sqrt{1 - 2^{-2(R_1+R_2)/D_0}}\}. \tag{2.47}$$

The overall coding rate $R_1 + R_2$ is comprised of two components. The *base rate* r and *redundancy*, ρ where the base rate corresponds to the quality of the joint description and the redundancy is the additional rate necessary to make the side distortion low.

D_0=2^{-2r} and ρ=$R_1 + R2 - r$. Substituting in 2.47 gives,

$$D_1 \geq \begin{cases} \frac{1}{2}[1+2^{-2r}-(1-2^{-2r})\sqrt{1-2^{-2\rho}}], & for \quad \rho \leq r-1+log_2(1+2^{-2r}) \\ 1-\sqrt{1-2^{-2\rho}}, & for \quad \rho \geq r-1+log_2(1+2^{-2r}) \end{cases} \tag{2.48}$$

Fig.2.4 shows the graphical representation of equation 2.48. From this figure, one can see that, the curves merge because the distortion bound is independent of r at high redundancies.

From 2.48, the slope of the low-redundancy D_1 versus the ρ characteristic is defined as

$$\frac{\partial D_1}{\partial \rho} = -\frac{1-2^{-2r}}{2}\frac{2^{-2\rho}ln2}{\sqrt{1-2^{-2\rho}}}. \tag{2.49}$$

At $\rho = 0^+$, the slope is infinite. When designing a system one should look, not only at the central distortion but also at the side distortion. If side distortion is also to be considered, the infinite slope means that a small increase in the overall rate will have higher impact in reducing the side distortion itself than in reduction of the central distortion.

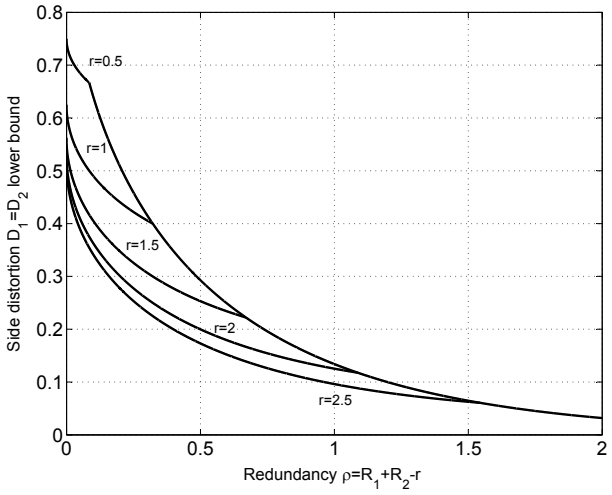

Figure 2.4: Lower Bounds on side distortions $D_1 = D2$ when the base rate is r and fixed distortion $D_0 = 2^{-2r}$ [Goyal 2001b].

As shown in Fig. 2.4, although the central distortion is independent from redundancy, since it depends only of r, the side distortion is reduced with redundancy insertion. Therefore, practical MDC redundancy allocation schemes are needed in order to adapt the coding parameters to the channel conditions, and consequently, reduce the overall distortion.

The rate-distortion bounds for non-Gaussian sources are not completely known. Nevertheless, an extension of Shannon rate-distortion bounds was proposed in [Zamir 1999] in order to define inner and outer bounds for MD coding of any continuous-valued memoryless source with squared-error distortion. The authors have shown that codebooks of Gaussian sources provide an efficient mechanism to encode non-gaussian sources.

2.2 Practical Implementation

The theoretical foundations described in section 2.1.2 do not answer the question about how to generate two or more redundant and mutually refinable descriptions, but only establish the optimal achievable rates for given joint central and side distortions.

The first practical approach to generate two descriptions was to split the source in odd and even samples, each one independently encoded. The distortion performance of the side decoder at high rates is shortened due to the interpolation error, but at low rates, its performance is competitive comparing with the central decoder [Goyal 2001a]. Nevertheless, such scheme adds implicit redundancy due to loss of coding efficiency but do not allow to dynamically adapt the redundancy level of MDC.

Other techniques were proposed in the literature, in order to improve the MDC features for source delivery with path diversity in error prone transmission scenarios. Such techniques need to solve several theoretical questions. On the other hand, the description must be independently decoded and needs to be refinable, which means that by joining both descriptions, the overall quality must be improved. Still, the redundancy among descriptions must be dynamically controllable, in order to obtain the best MD rate-distortion achievable points. Thus, the design of MD coders requires some structured approach for trading off source coding efficiency and MDC objectives.

2.2.1 Multiple Description with Correlating Transforms

Most of the source coding techniques are based on transforms, which are used to exploit the redundancy between samples. Transforms concentrate the energy of the source signal in a few set of coefficients, and this is one of the most important properties to use in efficient source coding. Each coefficient carries its own information which means that

each one is not correlated with the others. A transform should give an optimal energy concentration which is achieved when the coefficients are totally uncorrelated.

However, the objective of coding efficiency through the use of transforms has a interesting conflict with MDC, and it does not contribute to MDC performance. In fact, in error prone transmission, uncorrelated coefficients either cannot be used or are completely inefficient in recovering the missing information. This means that, for MDC, statistical dependencies among transform coefficients can be useful in order to estimate the lost coefficients on missing descriptions at the encoder.

MDCT is based on the principle that correlation between coefficients should be controlled for providing the necessary tradeoff between source coding efficiency and MDC goals. For this purpose, this technique uses a linear transform which inserts correlation between coefficients explicitly adding redundancy to each description. The MDCT method was introduced by *Wang et. al.* [Orchard 1997] [Wang 2001] and then the concept was generalized in [Goyal 2001b] for N descriptions.

Generally MDCT follows equation (2.50), where Transform $\hat{\mathbf{T}}$ takes two independent input variables A and B, and outputs two transform variables C and D. $\hat{\mathbf{T}}$ is a 2×2 matrix.

$$\begin{bmatrix} C \\ D \end{bmatrix} = \hat{\mathbf{T}} \begin{bmatrix} A \\ B \end{bmatrix}. \tag{2.50}$$

The transform $\hat{\mathbf{T}}$ controls the correlation between C and D, which in turn gives a measure of the redundancy introduced by this MDCT scheme and at the same time, preserves decorrelation among variables within each description.

The MDCT scheme is then represented in Fig. 2.5. Two independent Gaussian random independent variables (A,B) with variances σ_A^2 and σ_B^2 are considered. A linear transform $\hat{\mathbf{T}}$ is applied to the quantized indices $(\bar{A}, \bar{B})$ in order to obtain the correlated indices $(\bar{C}, \bar{D})$. The ($\bar{C}$,$\bar{D}$) indices are entropy encoded in independent streams, *i.e.* descriptions.

At the decoder side, if both descriptions are received, then the reconstructed pair $(\hat{A}^0, \hat{B}^0)$ is obtained by using the inverse transform $\hat{\mathbf{T}}^{-1}$. If only one description is received, then the $(\hat{A}, \hat{B})$ pair is estimated from the received coefficients by exploiting the correlation

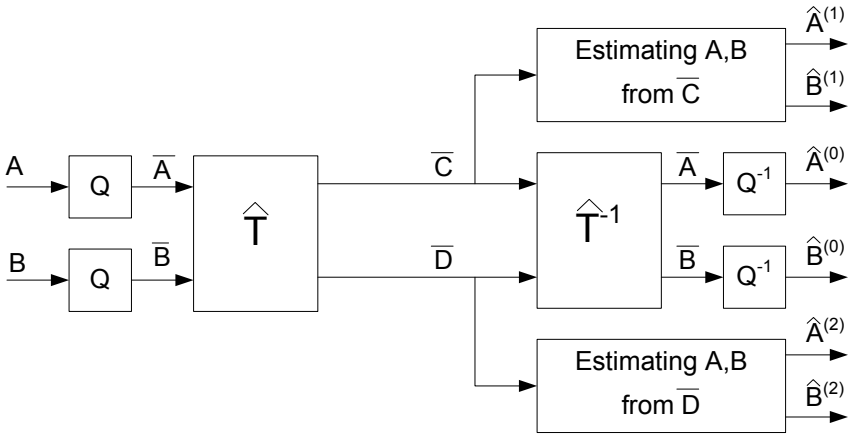

Figure 2.5: MDCT Generic Scheme.

added at encoder. If only one transform coefficient is available, it is possible to predict the original missing one. For instance, if $\bar{D}$ is missing, the estimate $\hat{D}$ from the available one $\bar{C}$ is given by

$$\hat{D}(\bar{C}) = \gamma_{D\bar{C}}\bar{C}, \tag{2.51}$$

where $\gamma_{D\widetilde{C}}$ depends on the variances of C and D and also on the quantization error energy of the two channel encoder. Similarly, if $\bar{C}$ is missing

$$\hat{C}(\bar{D}) = \gamma_{C\bar{D}}\bar{D}, \tag{2.52}$$

and the reconstructed values $\hat{A^{(i)}}, \hat{B^{(i)}}, i = 1, 2$ are obtained from inverse transform

$$\begin{bmatrix} \hat{A^{(1)}} \\ \hat{B^{(1)}} \end{bmatrix} = \mathbf{T}^{-1} \begin{bmatrix} \bar{C} \\ \hat{D} \end{bmatrix}, \tag{2.53}$$

and,

$$\begin{bmatrix} \hat{A^{(2)}} \\ \hat{B^{(2)}} \end{bmatrix} = \mathbf{T}^{-1} \begin{bmatrix} \hat{C} \\ \bar{D} \end{bmatrix}. \tag{2.54}$$

An arbitrary transform $\hat{\mathbf{T}}$ with determinant one is defined as

$$\hat{\mathbf{T}} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \tag{2.55}$$

where $\hat{\mathbf{T}}$ can be either orthogonal or nonorthogonal. If the transform is nonorthogonal,

the quantization in (C, D) transform is not geometrically coincident with quantization in (A, B) domain, increasing the quantization error. In order to overcome this problem the transform is not applied to variables (A, B) but to the integer indices $(\bar{A}, \bar{B})$. This is done using a discrete version of the transform in order to yield $(\bar{C},\bar{D})$ from $(\bar{A},\bar{B})$. The discretized transform should be designed to allow reversible integer-to-integer mapping, preserving the quantization cells in both domains [Li 1998]. $\hat{\mathbf{T}}$ must have determinant one in order to define one-to-one correspondence between the index pairs $(\bar{A}, \bar{B})$ and $(\bar{C}, \bar{B})$, thus, maintain the MDC coding efficiency, *i.e*,

$$\bar{A} = \left[\frac{A}{Q}\right], \bar{B} = \left[\frac{B}{Q}\right]. \tag{2.56}$$

From the method proposed in [Li 1998] based on LU decomposition, the output pair $(\bar{C}, \bar{D})$ is determined as *i.e*,

$$Z = \bar{B} + \left[\frac{1+c}{d}\bar{A}\right], \tag{2.57}$$

$$\bar{D} = [dZ] - \bar{A}, \tag{2.58}$$

$$\bar{C} = Z - \left[\frac{1-b}{d}\bar{D}\right], \tag{2.59}$$

where [.] is the rounding operation. If $(\bar{C}, \bar{D})$ values are received, then the corresponding inverse transform is implemented as,

$$Z = \bar{C} + \left[\frac{1-b}{d}\bar{D}\right], \tag{2.60}$$

$$\bar{A} = [dZ] - \bar{D}, \tag{2.61}$$

$$\bar{B} = Z - \left[\frac{1+c}{d}\bar{A}\right], \tag{2.62}$$

and

$$\hat{A} = \bar{A}Q, \hat{B} = \bar{B}Q. \tag{2.63}$$

The general transform $\hat{\mathbf{T}}$ can be parameterized as,

$$\hat{\mathbf{T}} = \begin{bmatrix} r_2\cos\theta_2 & -r_2\sin\theta_2 \\ -r_1\cos\theta_1 & r_1\sin\theta_1 \end{bmatrix} = \begin{bmatrix} \mathbf{v}1 & \mathbf{v}2 \end{bmatrix} \tag{2.64}$$

The transform coefficient is the representation of the original variables using the basis vectors $[\mathbf{v}_1\mathbf{v}_2]$. The transform parameters are defined in order to $\hat{\mathbf{T}}$ have determinant one. For that, $r_1 r_2 = 1/\sin(\Delta\theta)$ where $\Delta\theta = \theta_1 - \theta_2$, and $0 < \Delta\theta < \pi$. Considering these parameters, for an output pair (C, D) with variances σ_C and σ_D, the correlation between both of them is given by angle ϕ where,

$$\sigma_C^2 \sigma_D^2 \sin^2\phi = \sigma_A^2 \sigma_B^2 \tag{2.65}$$

Variables C and D are correlated and more bits are required to represent them in comparison with A and B. Using the rate-distortion functions for Gaussian variables, each pair of rates is given as

$$R = \tfrac{1}{2}\log_2 \tfrac{\sigma_C \sigma_D}{D_0} + K, \quad R^* = \tfrac{1}{2}\log_2 \tfrac{\sigma_A \sigma_B}{D_0} + K, \tag{2.66}$$

for some constant K. For Gaussian sources and entropy encoding, $K = (1/2)\log_2(\pi e/6)$. Using (2.67) and (2.66) the redundancy $\rho = R - R^*$ is given by

$$\rho = -\frac{1}{2}\log_2 \sin(\phi). \tag{2.67}$$

The optimal transform is the one that minimizes the *side distortion* for a given redundancy level ρ. The definition of the optimal $\hat{\mathbf{T}}$ was presented in [Wang 2001], as a non-orthogonal transform outperforming the one presented in [Orchard 1997]. On the other hand, the same authors have shown that, for orthogonal transforms, the quantization error is small if quantization is applied to the output transform. The orthogonal transforms have the disadvantage of showing smaller range of redundancies in comparison with nonorthogonal ones.

Practical implementations of $MDCT$ are found in image and video coding schemes based on block partitioning using transform coding, such as DCT. $MDCT$ is only used for pairs of variables and optimization algorithms are needed to find the best pair combinations. The optimization criterion is the one that minimizes the MD rate-distortion region, represented as the MDC redundancy and side distortions. In [Wang 2001] an optimization scheme is proposed based on three aspects: 1) optimal redundancy allocation assuming that are used M pairs; 2) optimal paring strategy and finally; 3) selection of the coefficients to be consider to transformed considering the quantization effect, by pairing only

coefficients having a large enough variance.

2.2.2 Multiple Description Sub-band Coding

MDC based on discrete wavelet transform (DWT) or sub-band decomposition was firstly proposed by Yang*et.al.* [Yang 2000] and Dragotti*et.al.* [Dragotti 2002] to be used on image and video coding. The proposed scheme is shown in Fig. 2.6 for the case of two bands.

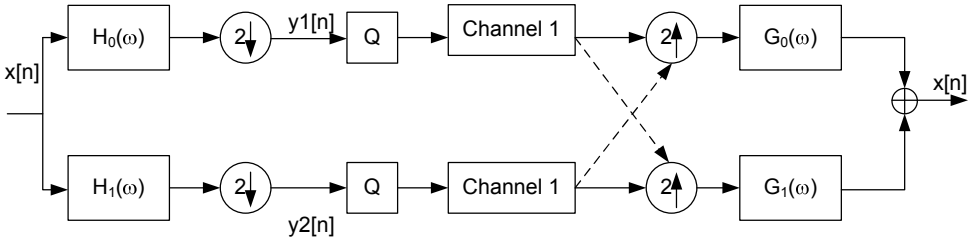

Figure 2.6: Multiple Description Subband Coding.

The input signal $x[n]$ is passed through low-pass and high-pass filters, with frequency responses $H_0(\omega)$ and $H_1(\omega)$, respectively, then each resulting component are sub-sampled to the new sampling rate thus yielding the sub-band signals $y_1[n]$ and $y_2[n]$. Following this step, each sub-band signal $y1[n]$ and $y2[n]$ are separately encoded and transmitted through different channels. To finally reconstruct the signal, each sub-band is up-sampled to the sampling rate of the input. All up-sampled components are passed through the synthesis filters $G_0(\omega)$ and $G_1(\omega)$. Then, each output component is interpolated and subsequently added all together to form the reconstructed signal $\hat{x}$. With separable filters, multi-dimensional analysis and synthesis can be carried out with stages of uni-directional filters. For example, an image might be first analyzed in its vertical direction followed by the same analysis in the horizontal direction. Missing samples are recovered using linear estimation based on received ones.

Sub-band filters are designed to optimize the rate-distortion performance of the MDC scheme. The optimization intends to finding the sub-band filters frequency response for a given redundancy level. The variation of the optimization parameters provides the ability to achieve all points on the theoretical redundancy rate-distortion curve, and associate the optimal filter banks with all redundancy rates. At zero redundancy, the optimization problem becomes the usual coding gain optimization, while for the maximum redundancy solutions, the two sudbband signals are maximaly correlated and have the same stochastic properties. In practice, Multiple Description Sub-band Coding needs one filter bank for

each redundancy level which limits the use of this technique is applications where MDC parameters have to be adapted in function of coding and transmission conditions.

2.2.3 Polyphase Decomposition and Selective Quantization

A particular case of the Multiple Description Sub-band Coding described in the section 2.2.2 was proposed by Ortega in [W. Jian 1999] by using the simplest polyphase decomposition filters [Vaidyanathan 1990], which consists in partitioning input samples into even and odd numbered subsets. Such scheme also adds redundancy by coding the original source sequence using different quality resolutions as redundant information. Fig. 2.7 shows the proposed MDC scheme where the source is decomposed into two sub-sources via polyphase transform.

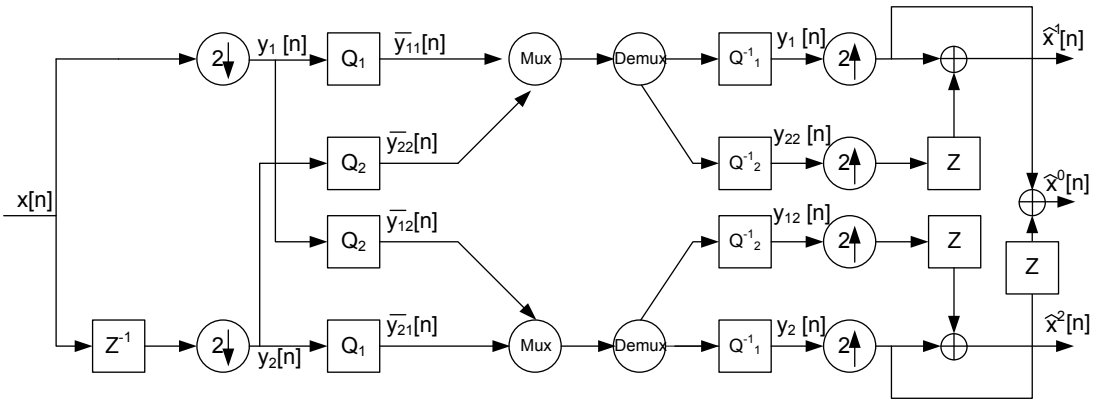

Figure 2.7: MDC with polyphase transform and selective quantisation.

The source x[n] is split in its polyphase components $y_1[n]$ and $y_2[n]$. Then each component is quantized with different quantization step-sizes, forming two types of information: primary, resultant of a finner quantization, and secondary, from a coarse quantization. Therefore, the primary information, $\bar{y}_{1,1}[n]$ and $\bar{y}_{2,1}[n]$ of each partition $y_1[n]$ and $y_2[n]$, is independently quantized with a finner quantization parameter Q_1. The secondary and totally redundant information is formed by a coarser quantized version of each partition with Q_2, $\bar{y}_{1,2}[n]$ and $\bar{y}_{2,2}[n]$. Primary information of each polyphase component are multiplexed with the complementary secondary partition, generating two descriptions, one with $\bar{y}_{1,1}[n]$ and $\bar{y}_{2,2}[n]$ and the other with $\bar{y}_{2,1}[n]$ and $\bar{y}_{1,2}[n]$.

At the decoder, each description is de-multiplexed, and its components are de-quantized and interpolated. If both descriptions are available, the primary information $\bar{y}_1[n]$ and

$\bar{y}_2[n]$ are combined in order to obtain $\hat{x}^0[n]$. The secondary components are fully redundant (they do not contribute with new information to decoding). When only one description is received the, primary and secondary components of the same description are combined, obtaining $\hat{x}^1[n]$ and $\hat{x}^2[n]$.

This MDC scheme was applied in [W. Jian 1999] on image coding using the Said-Pearlman wavelet coder [Said 1996]. The input image is first wavelet transformed and its polyphase components were extracted. Two different types of polyphase transforms were used. The first is obtained by simply group all the even coefficients into one description and all the odd coefficients into the other description. The other polyphase method is to use the zerotree structures separating the even blocks to one description and all the odd blocks to the other description. Such approach have inspired many proposals applied to image and video coding due to his flexibility, as will be described in chapter 3.

2.2.4 Multiple Description Scalar Quantisation

Multiple Description Scalar Quantization (MDSQ) was proposed by Vaishampayan in [Vaishampayan 1993]. MDSQ is usually designed to generate two independent descriptions from the same source signal and it is based on two operations: *scalar quantization* and *index assignment*. A generic MDSQ scheme is shown in Fig. 2.8.

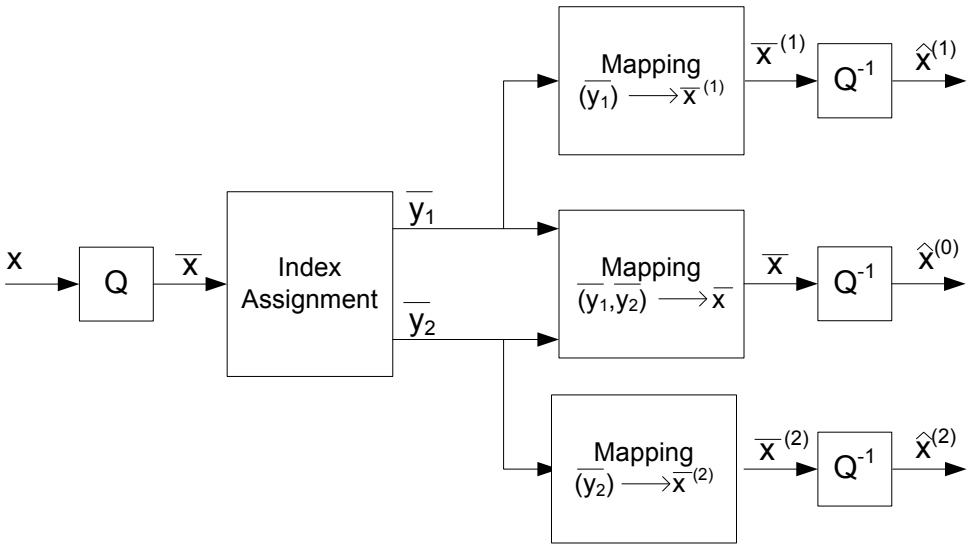

Figure 2.8: Generic MDSQ scheme.

A source x is scalar quantized, and for each index $\bar{x}$, an index pair $(\bar{y_1}, \bar{y_2})$ is assigned,

each one encoded using an entropy encoder. At the decoder, three types of mapping are used with three different reconstructions of x. When $(\bar{y_1}, \bar{y_2})$ are available, the output of the mapping function $(\bar{y_1}, \bar{y_2}) \rightarrow \bar{x}$, is inverse quantized and the central reconstruction $\hat{x}^{(0)}$ is obtained. When only $\bar{y_1}$ or $\bar{y_2}$ reach the decoder, a mapping $\bar{y_i} \rightarrow \bar{x}^{(i)}, i = 1, 2$ is used, and side reconstruction signals $\hat{x}^{(i)}, i = 1, 2$ are obtained after inverse quantization of $\bar{x}^{(i)}$.

Using as reference SD coding, in a M-level scalar quantization, each source sample x is mapped to a reconstruction level $\hat{x}$, which takes values form the codebook $\hat{\chi} = \{x_1, x_2, ..., x_M\}$. The M-level scalar quantization is usually regarded as composition of an encoder map $f : \mathbb{R} \rightarrow \{1, 2, ..., M\}$ whose output is a codeword index and a decoder map $f : \{1, 2, ..., M\} \rightarrow \mathbb{R}$. The encoder partitions $\mathbb{R}$ in M cells. Therefore, the SDC M-level scalar quantization is completely described by its partition and its codebook. Partition is represented by $A = \{A_1, A_2, ..., A_M\}$ where $A_i = \{x : f(x) = i\}$, $i = 1, 2, ..., M$.

The main principle of MDSQ is to construct two uniform quantisers with coarse reconstruction values such that separate inverse quantization is possible. Extending the SDC M-level scalar quantization to MDC, an $(M1, M2)$-level MDSQ maps the source sample x to reconstruction levels $\hat{x}^{(0)}$, $\hat{x}^{(1)}$ and $\hat{x}^{(2)}$ that take values in the codebook $\chi^0 = \{\hat{x}_{\bar{y_1},\bar{y_2}}, (\bar{y_1}, \bar{y_2})\varepsilon C\}$, $\chi^1 = \{\hat{x}^1_{\bar{y_1}}, \bar{y_1}\varepsilon I_1\}$ and $\chi^2 = \{\hat{x}^2_{\bar{y_2}}, \bar{y_2}\varepsilon I_2\}$ and C is a subset of $I_1 \times I_2$. An MDSQ can be broken up into two side encoders, $f_1 : \mathbb{R} \rightarrow I_1$ and $f_2 : \mathbb{R} \rightarrow I_2$, which select the indexes $\bar{y_1}$ and $\bar{y_2}$, respectively, and three decoders, $g_0 : C \rightarrow \mathbb{R}$ (central decoder) and $g_1 : I_1 \rightarrow \mathbb{R}$, $g_2 : I_2 \rightarrow \mathbb{R}$ (side decoders), whose outputs are the reconstruction levels with indexes $\bar{y_1}\bar{y_2}$, $\bar{y_1}$, and $\bar{y_2}$ from codebooks χ^0, χ^1 and χ^2 respectively. The two encoders impose a partition $A = \{A_{\bar{y_1},\bar{y_2}}, (\bar{y_1}, \bar{y_2})\varepsilon C\}$ on C, where $A_{\bar{y_1},\bar{y_2}} = \{x : f_1(x) = \bar{y_1}, f_2(x) = \bar{y_2}\}$. The MDSQ is completely described by A, χ^0, χ^1 and χ^2.

MDSQ formulation is illustrated in Fig. 2.9, using two distinct representations, (a) matrix representation and (b) using a linear representation. The source x is quantized according to the mapping function $Q_0 : \mathbb{R} \rightarrow 1, ..11$, representing the quantization function of the central decoder, with quantization stepsize Q_0, which means that are used $M = 11$ quantization cells. For each description, $M = 6$ and the MDSQ encoder output is $Q_i : \mathbb{R} \rightarrow 1, ..., 6, i = 1, 2$, representing the quantization function of the side decoders with quantization stepsise $Q_i, i = 1, 2$. In matrix representation, the inside values are the quantization indices $\bar{x}$, and each one have a correspondent index pair, $(\bar{y_1}, \bar{y_2})$, the indices of each description. In linear representation, each quantizer has one cell representation.

The reconstruction of $Q_i(x) = k_i, i = 0, 1, 2$ is the centroid of the cell $Q_i^{-1}(k_i)$. The central decoder results from the intersection cell $Q_1^{-1}(k_1) \cap Q_2^{-1}(k_2)$ between both side quantisers. The intersection cell is half size of the side quantizer cells, which means that an equivalent quantisation with a finner step size are obtained.

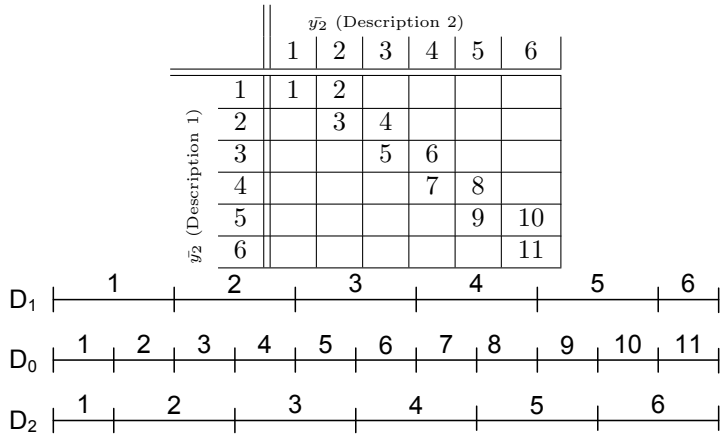

Figure 2.9: MDSQ example 1 - (a) Matrix representation ; (b) Linear representation

.

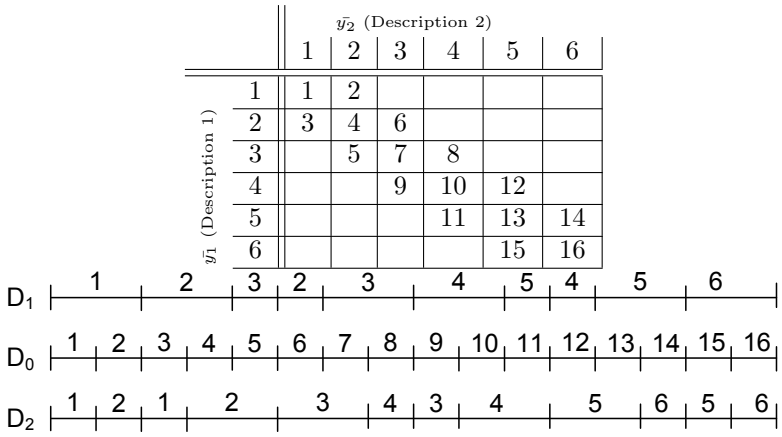

Figure 2.10: MDSQ example 2 -(a) Matrix representation; (b) Linear representation

.

Differently from single description scalar quantization, the MDSQ side quantizer levels are characterized by having discontinuous partition cells. Fig. 2.10 shows an example where the side quantizer cells are not continuous. This can be observed in linear representation where each side quantizer cells are interleaved between them. Although, the discontinuity of the quantiser cells of each description, they refine each other in a symmetric way, which means that central decoder have a continuous characteristic with also a finner step size.

In general, MDC algorithms use balanced descriptions where the respective rates and dis-

tortions are approximately the same. Considering two descriptions with rate-distortions pairs $(R1, D1)$ and $(R2, D2)$, these two descriptions are generated using side-quantisers from a central quantiser with rate-distortion pair $(R0, D0)$. For balanced MDSQ, $R1 \approx R2$ and $D1 \approx D2$. Given a certain partition of the equivalent central encoder (the one that would give the reconstruction $\hat{x}^{(0)}$), the main problem is to find side quantizers that produce the optimal combination between central distortion D_0 and side distortions $D1$ and $D2$.

The first MDSQ implementation was proposed in [Vaishampayan 1993] for fixed length binary encoders and is named as level-constrained MDSQ. In [Vaishampayan 1994] an optimal entropy-constrained MDSQ was proposed by using variable-length coding. The practical implementation is based on the MDSQ matrix representation as shown in Figs. 2.9 and 2.10, which is defined as *index assignment matrix*. The main problem is to design the index assignment function for individual quantizers, where the aim is to find the index assignment scheme that gives optimal combined exponential decay rates for the central and side decoders. Then, the index assignment problem is to find the best scanning sequence for a selected set of index pairs that results in small spread of each cell of side partition, *i.e.*, a small as possible interleaving between quantization cells of side encoders. For a given scanning scheme, the number of central cells is N, $n1 = n2$ is the number of index pairs, and $M = n_1 * n_2$. The central and side distortions are adjusted by the index spread k, that represents the number os diagonals above and below to the main diagonal of index assignment matrix. For instance in Fig. 2.10(a), $N = 16$, $M = 36$ and $k = 1$. Varying k and M, for a given fixed N, means that D_0 does not change, and side distortion D_1 (D_2) increases with higher k. Redundancy is reduced because M is lower. On the other hand, for fixed M and increasing k, the central distortion D_0 is reduced and side distortions $D_1 = D_2$ increase. Therefore, the index spread k is the most important parameter in order to tune both the central and side distortions and, consequently, the redundancy of the MDSQ scheme.

An interesting evaluation of the MDSQ scheme is the comparison with the optimal MD rate distortion bounds. The asymptotic analysis of MDSQ was made in [Vaishampayan 1998]. In asymptotic sense, the optimal MD rate distortion bound at large rates for a squared-error distortion measure, and unit-variance memoryless Gaussian source can be approximately given as

$$D_0 D_1 \approx \frac{1}{4} 2^{-2(2R)}. \tag{2.68}$$

The optimum level-constrained MDSQ for a unit-variance Gaussian source is given by

$$D_0 D_1 \approx \frac{2\pi^2}{16} 2^{-2(2R)}. \tag{2.69}$$

Finally, for the entropy-constrained MDSQ the RD bound is given by

$$D_0 D_1 \approx \frac{\pi^2 e^2}{144} 2^{-2(2R)}. \tag{2.70}$$

This results show that for entropy-constrained MDSQ, the RD bound is closer to the optimal level than the level-constrained MDSQ. A simpler MDSQ implementation based on two-stage index assignment with RD performances close to the entropy-constrained MDSQ scheme was proposed in [C. Tian 2004]. A continuum of tradeoff points between the central and side distortions are achieved by controlling quantization parameter of the first stage and redundancy parameter of the second stage. The index assignment is done based on these two steps and avoids extensive training used in entropy-constraint MDSQ on the quantizer thresholds and reconstruction values for a given index assignment matrix.

2.2.5 MDC Based on Error Correction Codes

Most of modern communication networks use Forward Error Correction (FEC) Codes as a channel coding scheme, specially in applications where retransmissions are not available or are not possible due to delay constraints. FEC coding consists in adding redundancy at channel coding level, where the redundant codes provide useful information to recover missing one at decoder. The redundancy level is dependent of the error probability of the channel. One important FEC codes finally are the Reed-Solomon (RS) block coding, which are characterized by the size (N, k), where N is the length of the block code and k is the length of the source symbols. The $N - k$ symbols are the channel coding redundancy added to the source symbols in order to detect and recover lost information. The main feature of a (N, k) RS code is that, all of the lost symbols can be recovered at the decoder if the number of erasures does not exceed $N - k$.

Extending the FEC properties to MDC problem, a new MDC scheme is defined as MDC with FEC (MD-FEC), which is suited to be used in MDC systems with $N > 2$ descriptions. The use of such scheme is tailored for benefiting of multi-resolution or scalable source coding. In fact, the main goal of MD-FEC is to give unequal error protection (UEP) to

different coded layers or quality levels, which means that the most important layers or high quality levels are assigned with stronger FEC codes in order to better protect those layers. Fig. 2.11 shows an example of the $MD-FEC$ scheme, where the source is divided in N layers or quality levels, $\{1, ...N\}$, each one with rates $\{R_1...R_N\}$, respectively.

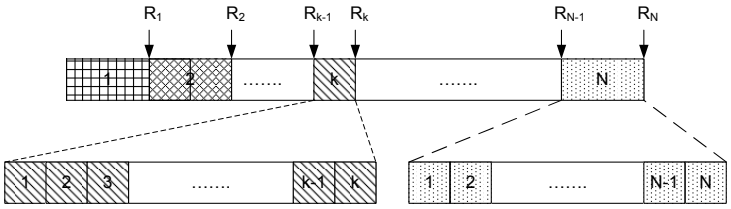

Figure 2.11: MDC FEC.

Fig. 2.12 shows the N-description generation. For each one of the N descriptions, description k have $N - k + 1$ source layer blocks. Description 1 contains groups of all layers. In Description 2, the first block is filled with FEC codes, and all the other layers are filled with the corresponding blocks, and so on. As a result of this process, N descriptions are encoded and N RS code blocks are formed. As shown in Fig. 2.12, the kth layer is protected by a (N, k) RS code, which means that layers $\{1, 2, ...N\}$ have $\{(N, 1), (N, 2), (N, N)\}$ RS codes. In this way, is assured that all the first N source layers are decoded by performing RS decoding from any N descriptions.

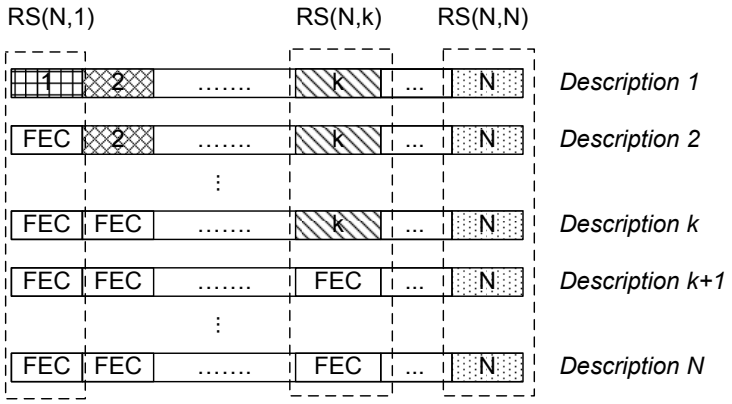

Figure 2.12: MDC FEC.

The MDC redundancy is tuned using the optimal layer partitioning $(R_1, R_2, ...R_N)$. The optimization criteria is the minimization of the overall distortion at receiver given the total rate constraint, based on loss probability of descriptions K, $K = 1, 2, ...N$ [Puri 1999]. Although MD-FEC add robustness to many layer coding schemes, the delay caused by FEC operations may affect the performance and applicability in systems where such parameter is critically constraint.

Chapter 3

Multiple Description Video Coding and Networking

Contents

3.1 Basic Video Coding Techniques **36**

3.2 Multiple Description Video Coding-an overview **40**

3.2.1 MDC with pre-processing 43

3.2.2 Embedded MDC 46

3.2.3 High Dimensional MDC 50

3.2.4 Multiple Description Scalable Coding 51

3.2.5 Stereo/Multiview MDC 52

3.2.6 Unbalanced MDC 54

3.2.7 Multiple Description Distributed Video Coding 55

3.3 Multiple Description Video Streaming **56**

3.3.1 MDC Video Streaming over Content Delivery Networks 57

3.3.2 MDC Streaming over Peer-to-peer networks 59

3.3.3 Scheduling in path diversity networks 62

This chapter presents a literature review of the most relevant research in the field of Multiple Description (MD) coding and transmission for video signals. The first part describes Multiple Description Coding (MDC) techniques and architectures that have been proposed in the literature. MDC video architectures are described for different video

coding schemes, based on techniques described in Chapter 2. These are applied in different video coding approaches such as, advanced video coding, scalable video coding, distributed video coding, multiview and stereo video coding, for several coded descriptions. The second part of this chapter describes MDC applications in networks with path diversity scenarios. MD video delivery network topologies, architectures and applications are described for content delivery networks and peer-to-peer networks.

3.1 Basic Video Coding Techniques

The main goal of video encoding schemes is to compress the amount of data used to represent visual information, controlling the coding distortion introduced in the original signals. Video coding and compression schemes exploit two main principles: *redundancy* and *irrelevancy* of video signal.

Redundancy is related with the source sample information that is correlated with other neighbor samples. In video sequences, the redundancy exists in both the spatial and temporal dimensions. In the spatial dimension, this is due to the fact that pixels within the same frame are highly correlated. In the temporal dimension, the use of high frame rates, capturing several images per second results in consecutive frames that have almost the same information, since temporal evolution of signal contents is generally smooth. In addition, source coding symbols are correlated, which means that entropy redundancy is also exploited in order to compress source information. The characteristic of irrelevancy is directly connected to the Human Visual System (HVS), which acts as a filter in regard to some spatio-temporal frequency components. The principle behind lossy coding schemes is to remove the irrelevant components from the video signals, in order to reduce the amount of information to be coded without introducing noticeable distortion.

Therefore, a variety of coding tools to efficiently compress visual information are used, which includes lossless and lossy schemes. These are based on Differential Pulse Coding Modulation (DPCM) schemes where the rate-distortion performance is mainly dependent on the prediction functions, that are responsible to minimize the residue energy of a given sample or block of samples. Associated to the best prediction, most of the video coding architectures are based on transform coding and quantization (in lossy schemes) that allows to concentrate the residue energy in few coefficients. Finally, source symbols are entropy encoded, ideally, close to the source entropy. Thus, video coding algorithms are

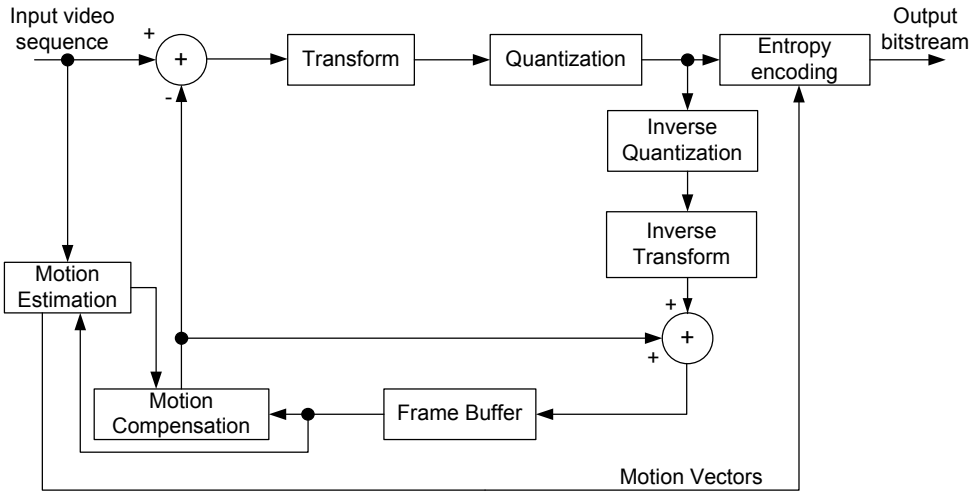

Figure 3.1: Hybrid video encoder.

grouped in three main approaches: 1) DCT coding, which is used in modern video coding standards; 2) 3D-suband coding, based on sub-band decompositon in temporal and spatial domains; 3) Distributed Video Coding, which transfer the computational complexity to the decoder, through simpler encoders and using side information to generate decoded video sequences.

DCT based video coding Fig. 3.1 shows the common usual DCT based video coding scheme which consists in a feedback loop of a block-based *motion estimation* (ME) and *motion compensation* (MC) to combat redundancy followed by a transform (usually DCT) of the prediction error. The transformed residue is then quantized and entropy encoded. One of the main blocks is the ME, which finds the motion information between the current frame and the reference ones. This is essential to reduce the error energy of the motion compensated prediction, and consequently, to improve coding efficiency. Irrelevancy is exploited by quantization which is characterized by the quantization step-size that maps each transform coefficient into a corresponding index. The quantization step-size is directly related to the distortion of decoded signal. Headers, motion information represented by *motion vectors* and prediction modes are entropy encoded. This encoding scheme is used in standard video codecs, such as, MPEG-2 [ISO/IEC13818-2 2000] and H.264/AVC [ISO/IEC14496-10 2012].

The DCT-based video architecture is also used in scalable video coding tools, which consists in encoding video sequences in several dependent layers. This means that higher layers need lower layers to be decoded. Different layers correspond to different resolutions

or quality that can be obtained by partial decoding of the scalable stream. In general, full resolution is obtained when all layers are decoded. Scalability can be classified in tree types: temporal, spatial and quality. In temporal scalability each layer correspondes to different temporal resolutions that are combined to reach the full resolution. In spatial scalability the spatial resolution of the original sequences is down-sampled in order to obtain several resolutions. The quality scalability is based on the refinement of the transform coefficients in order to obtain several quality levels at the decoder.

Sub-band video coding Sub-band video coding extends the spatial sub-band decomposition based on filter banks Discrete Wavelet Transform (DWT) used on image coding to the temporal dimension. Although good performance of DWT for image coding, which has been adopted in JPEG2000 standard [15444-1 2002], sub-band video coding has not been successful for current state of the art video coding standards. Nevertheless, sub-band video coding is inherently scalable which means that is still an interesting video coding paradigm for video coding applications in heterogeneous environments.

Fig. 3.2, shows the block diagram of sub-band video coding. The main functional block is the 3D-analysis, to the sequence in temporal and spatial frequency sub-bands. This is done using a filter banks, composed by temporal filtering in high and low frequencies and, spatial DWT, that divides spatial information in frequency sub-bands. Finally, each sub-band coefficient is quantized. The output bitstream results from efficient entropy coding of the temporal and spatial sub-bands. 3-D analysis filtering can be made without motion compensation [Said 1996] or with motion compensation [Ohm 1994,Kim 1997,Choi 1999], which is the Motion Compensated Temporal Filtering (MCTF) video encoder scheme.

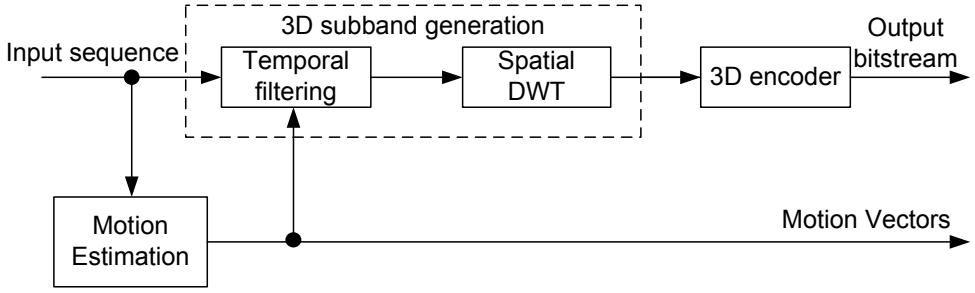

Figure 3.2: 3-D sub-band video encoding.

Fig. 3.3 illustrates an overall 3-D analysis example. The sub-band decomposition is based on a hierarchical tree-structured scheme. At the first stage, the input video is decomposed into temporal L- and H-sub-bands. Then, the L-sub-band is decomposed again into

temporal LL and LH-sub-bands at the second stage. Thus, the resultant four frame unit is composed by decompositions t(temporal)-LL, t-LH, and two t-H sub-bands. Finally, a spatial wavelet transform is done on each temporal sub-band. Image wavelet sub-band coding use embedded zero-tree wavelet coding (EZW) schemes based on [Shapiro 1992] which exploits the inter sub-band similarities. In video coding the extension to 3D-EZW is used in order to exploit also the temporal sub-bands [Kim 2000].

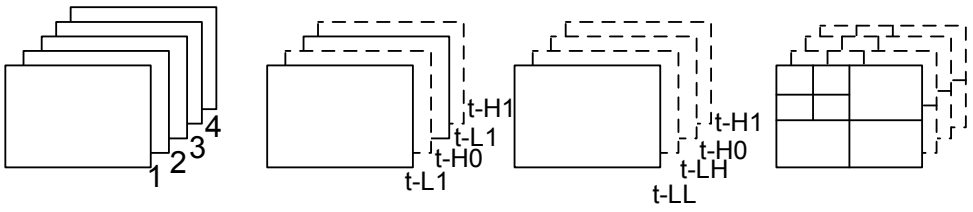

Figure 3.3: 3-D sub-band decomposition.

Distributed Video Coding Hybrid schemes are characterized by complex encoders and simpler decoders. Distributed video coding intends to change this paradigm by transferring the complexity from the encoder to decoder. This is a new approach in video coding paradigms which is based on two theorems from information theory: Slepian-Wolf [Slepian 1973] and Wyner-Ziv theorems [Wyner 1976]. The Slepian-wolf theorem states that two statically dependent signals X an Y can be compressed in a distributed way, i.e. separately encoded and jointly decoded, using approximately the same rate if encoded and decoded together. On the other hand, Wyner-Ziv theorem extends the earlier theorem for lossy encoding and also states that if the correlated signal Y is only available at decoder (the side information at decoder), there is no loss in coding efficiency in X, comparing with the lossy joint encoding of X and Y. These theorems together have provided the basis for the new video coding paradigm that allows to encode two statistically dependent signals independently and decoding them jointly with coding efficiency closer to the predictive encoding schemes.

Fig. 3.4 shows the distributed video coding (DVC) scheme. Two different kind of frames are considered, the Wyner-Ziv and Key frames, which are separately encoded. The Wyner-Ziv frames are intraframe encoded based on DCT transform followed by quantization. Each coefficient are bit-plane encoded and each one is encoded using channel encoding schemes. Key frames are conventionally intraframe encoded and are used as side information to decode the Wyner-Ziv frames. DVC transfer the computational complexity to the decoder, where Wyner-Ziv decoding is done using the key frames as side

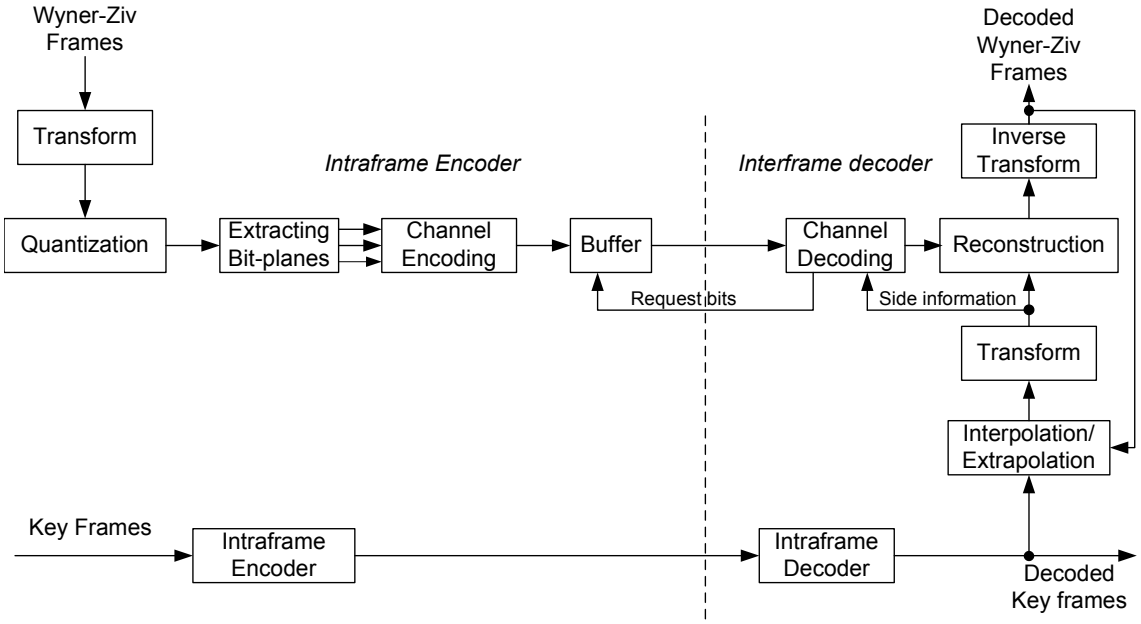

Figure 3.4: Distributed video encoding.

information. The decoding of Wyner-Ziv frames are based on interpolation/extrapolation of key frames, that combined with encoded bit-planes, reconstruction frames are obtained. The received channel coding information depends of the quality that is needed at decoder, which can interactively request bits to encoder. A review on DVC video coding techniques can be found in [Pereira 2009].

3.2 Multiple Description Video Coding-an overview

The video encoding schemes briefly explained in Section 3.1 are the starting point of the development of MD video encoders proposed in the literature by combining Single Description Coding with the MDC techniques referred in Chap. 2. Considering that a MD video encoder outputs M descriptions, these can be generated in different encoding stages: i) pre-processing stage; ii) encoding stage inside the prediction loops; iii) post-processing stages. This gives the MDC classification as pre-processing MDC, embedded MDC and post-processing MDC respectively.

Fig. 3.5 shows a generic framework where the different stages can be combined to achieve MD video encoding. In the pre-processing MDC, the video sequences are pre-processed in order to form M descriptions. The input video sequence is first sub-sampled to form

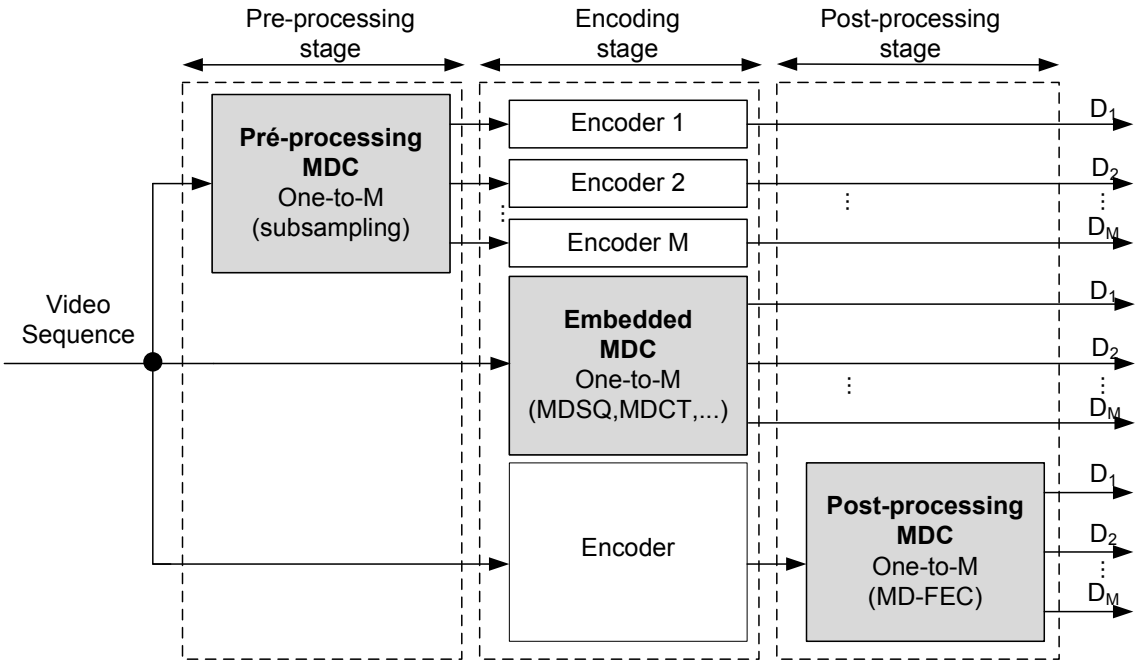

Figure 3.5: Generic MD encoder

different versions of the source signal and then, each one is independently encoded resulting in different descriptions. Down-sampling is done in spatial or temporal domains. In embedded MDC encoders the M descriptions are generated at the video encoder prediction loop. This MDC schemes includes $MDCT$ and $MDSQ$ techniques. Finally, in post-processing MDC, M descriptions are generated on the output bitstream. MD-FEC techniques are included in this group of MDC encoders.

Due to the predictive nature of most of video coding algorithms, most of the MD video coding schemes have to deal with the problem of drift introduced when some description does not reach the decoder, which originates mismatch between the decoder and encoder reference frames. The drift problem raises the need to review the encoder model because it is not known the decoding state, which means that the encoder does not know what description is available at each decoding time. This can only be solved if all predictions are combined at the encoder in order to prevent the mismatch occurrence at decoder. As the main drawback of MD video coding, the drift control is one of the issues that the MD video coders should solve. Along with the redundancy rate-distortion definition, this is one of the most important factors used to compare the overall performance in MD video coding.

Drift control can be done by introducing different prediction loops at the encoder to pre-

vent all the failure possibilities at decoder. Obviously, this approach is only practicable if a reduced number of descriptions is used, because MD video coding architecture complexity increases with the number of descriptions. The encoding process has to provide this ability without knowing the available descriptions at each decoding time. This means that MDC encoders have to be able to prevent error drift propagation at decoder, when transmission errors occur.

Usually, balanced schemes are used, but in error multiple description directly associated to error correction codes, namely, in FEC codes, unbalanced error protection is used where coding parameters are set as function of each channel condition. At source level, unbalanced descriptions can be generated in order to adapt MDC scheme to channel conditions, not only due to packet loss rates but also due to available bit-rates.

The redundancy allocations is dependent on the channel conditions and must define the MDC parameters in order to obtain the optimal allocation that minimize the overall distortion. The overall rate allocation includes the redundancy, that controls two fundamental characteristics of the MDC scheme: the side distortion (only one description is decoded) and the central distortion when all descriptions are available. The redundancy of MD video encoders has many sources, and depends on the specific MD scheme that is applied. The redundancy components includes: (i) coding the prediction error for each description (ρ_{md}); (ii) prediction less efficient compared with SDC (ρ_{ef}); (iii) explicit signal for drift reducing (ρ_{dr}) and finally, (iv) side information, for instance by duplicating motion or header information (ρ_{si}). The total redundancy is defined as

$$\rho = \rho_{md} + \rho_{ef} + \rho_{dr} + \rho_{si}. \tag{3.1}$$

The decoder must be able to decode each description individually, producing reconstructions with an acceptable quality, and combine them to form a higher quality reconstruction, knowing that one description might be missing. The decoding process is dependent on the MDC technique and includes, error concealment, received side information that is used only when some description fails and interpolation based on received information. Depending on the encoding process, the occurrence of mismatch and consequent error propagation needs to be controlled or mitigated. this effect might be limited with different prediction loops at encoding or using less efficient prediction schemes in rate-distortion sense but more effective in dealing with error propagation (*i.e.* drift distortion).

This review presented is based on the type of SDC encoding, which means that the proposed classification depends on the original SDC scheme that is in the genesis of the corresponding MDC architecture. The MDC video coding schemes are classified as: (a) MD embedded video encoding; The MD embedded video encoder schemes includes all the techniques where the MD functional block is embedded at encoder prediction loop. These include MDSQ, MDCT, splitting of transform coefficients and motion information. (b) MDC with pre-processing; In MDC with pre-processing include all the schemes that use spatio-temporal sub-sampling before encoding of each description, where redundant information can be explicitly inserted in order to improve the recovery of each description that is independently decoded. (c) Scalable MDC; Includes all MDC schemes based on 3D-subband and DCT based scalable video coding. (d) Multiple Description Distributed Video Coding (MDDV), which use MDC based on DVC (MDDVC). (e) Stereo/Multiview MDC; Stereo and multiview video coding is an inherently multiple description based approach which has been explored to generate MDC schemes.

3.2.1 MDC with pre-processing

In MDC with pre-processing the input video signal is downsampled in order to form two or more descriptions of the source which are independently encoded and decoded. The full reconstructed sequence is obtained by merging all decoded descriptions. Two types of architectures have been proposed: open-loop and, three-loop architectures. In open-loop architectures, each description is separately encoded and decoded. Some proposals send explicit redundancy of complementary descriptions with a coarser quality independently encoded. In three-loop architectures the complementary information of each description is sent as side information, that is generated using different prediction loops.

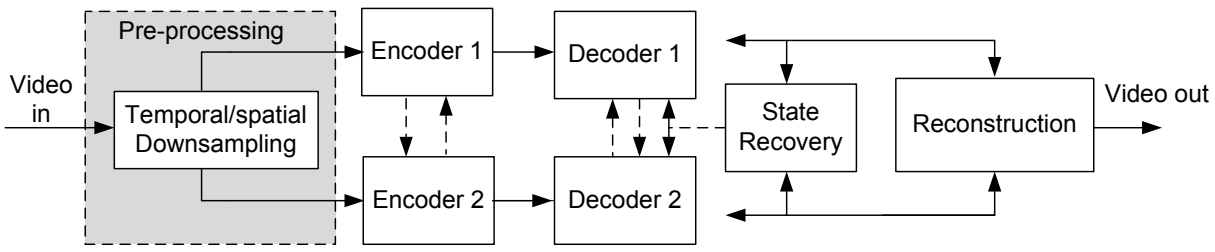

Figure 3.6: Pre-processing MDC.

Open-loop architectures Fig. 3.6 shows a generic MDC scheme with pre-processing for two descriptions. The whole chain between encoder and reconstructed sequence is shown. The video sequence is pre-processed using either i) spatial down-sampling, where each frame is divided, usually in macroblocks or slices in order to make two or more groups of frames with reduced resolution, or ii) temporal down-sampling, where the video frames are separated to form two or more descriptions with reduced temporal resolution comparing with original. Usually, each description is individually encoded. However, each encoder is able to use the information of the other description, and to encode the other group as redundant information.

Each description is decoded individually, and in function of the state recovery that depends of the received information of each description, use available information in order to merge both descriptions. The reconstruction is done by merging decoding frames of each description. The error recover at decoder is done by using the other description already decoded, redundant information of each description or both.

A simple MDC scheme based on these principles was proposed in [Apostolopoulos 2000] where an odd/even frame splitting is proposed with two independent encoders. The redundancy results from the loss of prediction efficiency and to the increase of the temporal distance between frames. The main problem of such schemes are the poor side reconstruction comparing with the SDC full rate scheme.

Several works exploit standard error resilience tools, such as slice reordering and redundant frames, where redundant information is set in function of error probability of the main information. MD schemes follow the same principle enumerated in section 2.2.3, where a coarser redundant information is coded to be used whenever the main information is not available. A redundant frame approach was proposed in [Radulovic 2010] in order to deal with missing frames at decoder when odd/even separation is used at encoder. The input video sequence is split into sequences of odd and even source frames. In the encoding process, each primary frame in the even/odd description is predicted only from other frames of the same description, typically the previous frame. In addition, redundant frames are included in the bitstream of each description, thus carrying the information from the alternate description. In the time domain, they are positioned such that they can replace a lost primary picture. Unlike the primary pictures, which use the previous primary frames from the same description as a reference, redundant pictures are predicted from the previous frame in the input sequence.

A similar approach was proposed in [Tillo 2008], where each frame is split in even/odd slices and for each description a complementary set of primary and redundant slices are defined. An optimal redundancy allocation at GOP level was proposed. The same scheme was improved by using the optimization method at slice level in [Peraldo 2010] and more recently at macroblock level [Lin 2011]. In later works, redundant slices use the same prediction modes as the primary slices. An independent encoding of the redundant slices are considered in [Schmidt 2011] where a joint optimization mode decision of redundant information is proposed. In pre-processing schemes where explicit redundant information is not encoded, concealment and interpolation techniques using the neighboring frames can be used. An example of such an approach can be found in [Ma 2008].

Three-loop Architectures A three-prediction loop MD video codec using temporal sub-sampling was proposed in [Wang 2002]. The video sequence is partitioned in multiple threads using temporal sub-sampling, and then an independent prediction loop to each set is employed. The original sequence is split in odd frames and even frames. The overall redundancy is controlled by the predictor coefficients and by the quantisation stepsize for encoding the mismatch signal. Depending on the network conditions, the amount of redundancy is controlled to achieve a good tradeoff between coding efficiency and error resilience.

Another type of approach is based on the slice reordering feature of H.264/AVC, namely the Flexible Macroblock Ordering (FMO) tool. A slice group scheme is presented in [D. Wang 2005] where three motion compensation loops are used. The video signal is encoded in the central encoder and, then, divided into two descriptions, each one corresponding to one slice group. Each slice group includes redundant information from the other one. Based on this scheme, a rate controlled redundancy-adaptive model that takes into account the effects of error propagation and concealment is proposed in [N. Kamonoonwatana 2010]. A similar approach is proposed in [Su 2008] where the temporal and spatial correlations between macroblocks are exploited to achieve efficient redundant coding.

Although they exhibit good performance in terms of redundancy allocation, these schemes are highly dependent on the specific features of the corresponding SDC coding schemes, and are not always available for all coding profiles.

3.2.2 Embedded MDC

In video encoding with embedded MDC schemes, the MD module is embedded into the prediction loop. These MD schemes include scalar quantisation (MDSQ), correlating transforms (MDTC), splitting of DCT coefficients, sub-sampling of residual and motion information. Embedded MDC schemes are further classified as in either open-loop architectures, where descriptions are generated independently from the prediction loop or three-loop architectures that encode the residual difference to reduce or even eliminate the drift at decoder.

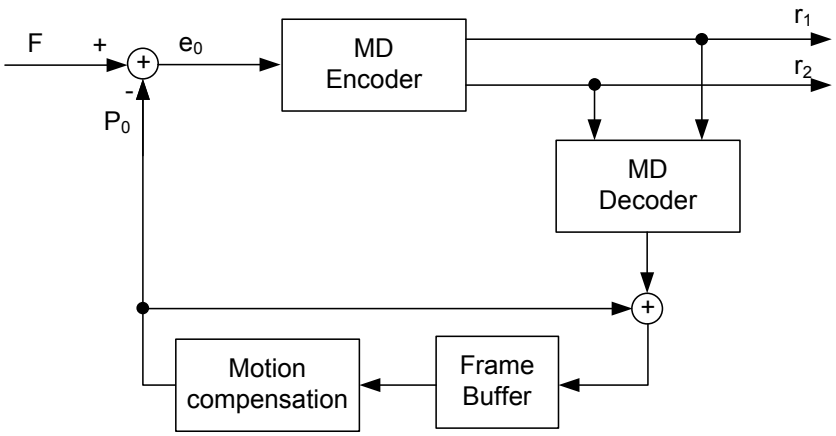

Figure 3.7: Open-loop multiple description video encoder.

Open-loop architectures Fig. 3.7 shows the generic open-loop MD video encoder using two descriptions. The MD encoder is embedded in the prediction generating two output independent streams r_1 and r_2. An MD decoder is used in the loop that merges both descriptions, which means that reference frames used for prediction are resulting from SDC decoding. This means that only with both descriptions the prediction loop at decoder is the same as correspondent encoder.

One of the first MD video encoders proposed in [Vaishampayan 1999] based on a mutually refining $DPCM$ scheme, with two independent prediction loops using two coarse quantization step-size, one for each single-channel reconstruction. In this scheme, if both descriptions are received by the decoder, then a specific method is needed to incorporate both predictions on the decoding loop to improve the joint quality. While avoiding prediction mismatch in the side decoders, this results in a poorer prediction when the outputs of both channels are available.

Other MDC methods have been proposed, where the prediction mismatch between encoded and decoder is not controlled. These include MDC video encoders based on rate-distortion splitting wherein the output of a standard video encoder is split into two correlated streams. In these schemes, the redundancy is relatively high because header information, motion vectors and DCT coefficients with magnitudes above a certain threshold are duplicated in the two descriptions, whereas the remaining coefficients are evenly distributed among the two streams. At the decoder, if both descriptions are received, then the redundant information is discarded [Reibman 2001] [Matty 2005]. An open-loop architecture for H.264/AVC based on MDSQ has been proposed in [Campana 2008]. In such architecture, drift control is not supported which results in distortion accumulation along the GOP.

Usually, motion information is duplicated in both descriptions, considering that motion information has not an expensive cost in total redundancy, which might not be exactly truth in low rate application environments. Fig.3.8 shows the MD motion coding method proposed [Kim 2001], where the original motion vector field is sub-sampled using quincunx lattices to form two different descriptions. In both descriptions, the macroblock and motion vectors are alternated. In order to avoid visual artifacts in the reconstructed signal, overlapped block motion compensation (OBMC) is used when only one description is present.

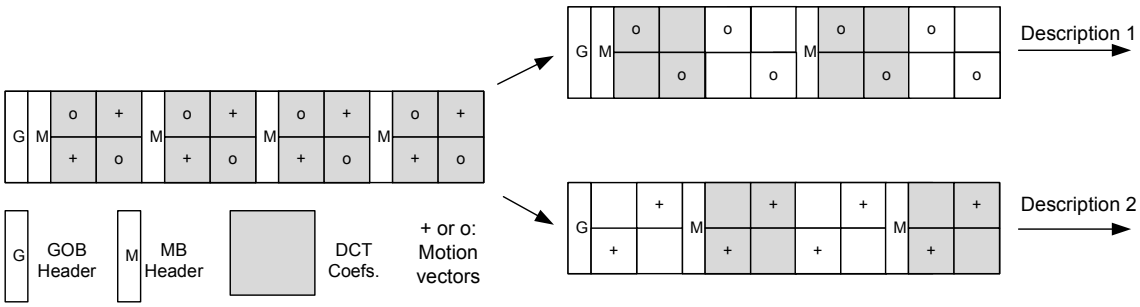

Figure 3.8: Data Partitioning among channels
[Kim 2001]

At decoder, the motion compensation is done with the available motion vectors. Since DCT coefficients are sub-sampled to form two descriptions, if one descriptions is not available, the missing *DCT* coefficients are replaced by the corresponding one in previous frame. This scheme does not use drift compensation, thus adding distortion accumulation at decoded sequences, whenever some description fails.

Three-loop architectures In order to overcome the problem of distortion in MDC, multi-loop architectures with additional side information have been proposed [Wang 2005b]. An excess rate is added in comparison with open loop architectures, but such degradation of overall rate distortion performance is expected to be compensated by improvement of the side distortion.

The general architecture of a three-loop architecture is presented in Fig. 3.9. The architecture is formed by three encoders: one central encoder and two side encoders which generate a controlled amount of redundancy, directly related with the side distortion. Signal $e_0 = F - P_0$, is the prediction error of the central decoder, where F is the current frame and P_0 is the central motion compensated or spatial prediction. MD function is applied to the signal e_0 and the two output residues r_1 and r_2 are generated and entropy encoded. Two additional prediction loops are used in order to mimic the decoder when only one description is available. The residue r_1 and r_2 are individually decoded in the MDC decoder functional block, resulting in $\hat{r}_1$ and $\hat{r}_2$. The prediction error e_1 is the difference between current frame F, the prediction P_1 and the residue $\hat{r}_1$ while e_2 is the difference between current frame F, and prediction P_2. e_1 and e_2 are encoded according to a redundancy allocation scheme generating side information s_1 and s_2. Description 1 is formed by the multiplexed signals r_1 and s_1 and description 2 is formed multiplexed signals r_2 and s_2.

Based on the generic architecture of Fig. 3.9, several MDC schemes have been proposed. The main differences are related to the central encoder and MDC schemes. For instance, an MDC architecture using $MDCT$ was proposed in [Reibman 2002]. The MD video coder is jointly optimized using motion vectors, DCT and side information levels. Only the DCT coefficients are considered to make the different descriptions, while motion vectors and side information are simply duplicated. A three prediction loop based on polyphase downsampling of sub-band images have been proposed in [Franchi 2005], replacing the MDCT module. Instead of using the original frame, in this case the current reconstructed frame is used as reference to compute the prediction error of the side information.

In [Tang 2002], an MDC scheme is proposed using Matching Pursuits (MP) video coding instead transform coding. Video coding based on Matching Pursuits does not use DCT, and the residue from motion compensation is decomposed onto a larger basis set than the complete set provided by DCT. At each stage of coding the residue, a search for the best basis function is performed by computing the inner product between the residual and the dictionary of basis function. The residue is subtracted from the best basis function and the

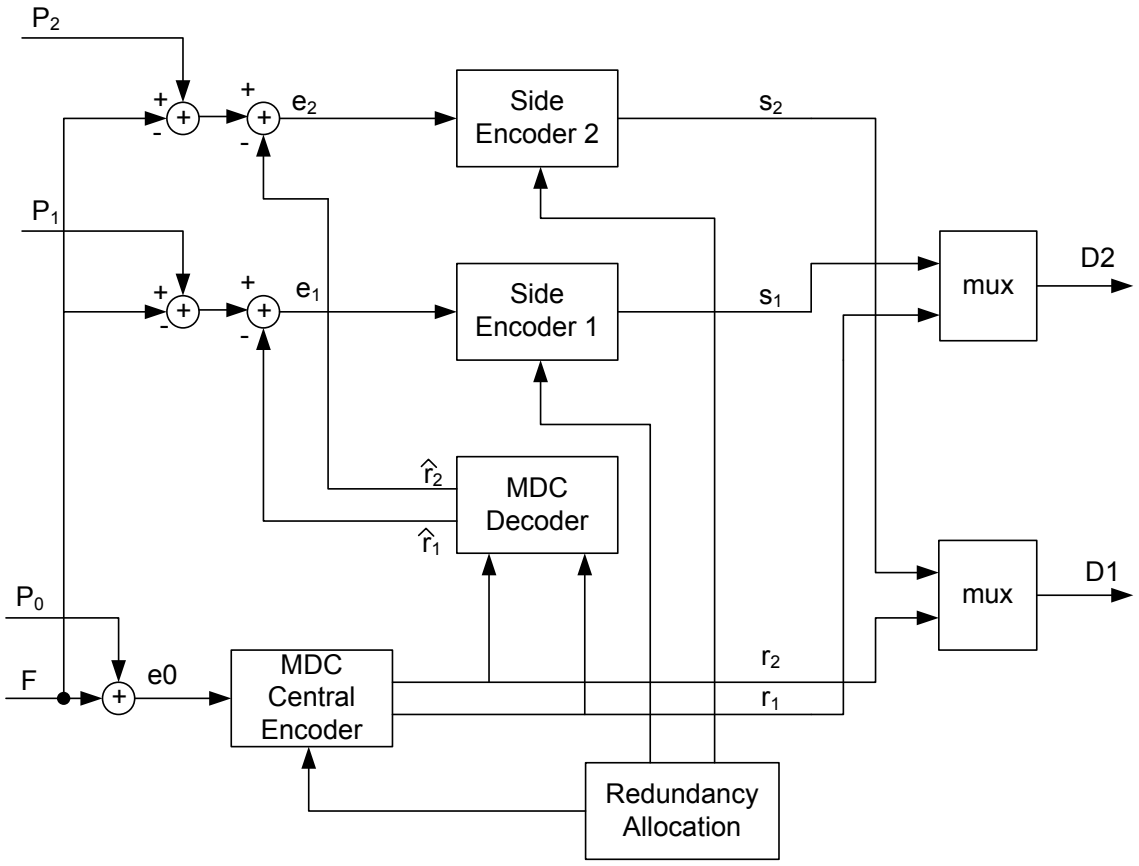

Figure 3.9: Tree-loop multiple description video encoder.

iteration repeats. This ensures that the most important features are coded first. The set of over-complete basis is called MP dictionary which consists in a general family of time-frequency atoms generated by scaling, translating and modulating a single time function. In MP-MDC, the residue is coded into two sets of atoms (one basis function). The first L atoms found during MP iterations are shared by both sets and subsequent atoms are alternatively put into the two sets. As a result, these sets are of approximately equal importance because atoms are found in decreasing order of magnitude in MP iterations. The correlation between these two sets of atoms is controlled by the number of shared atoms L. Another MP-MDC scheme is proposed in [Hsi-Tzeng Chan 2005] where MDSQ is used to encode the MP atom parameters into two descriptors.

Usually, the side information is useless when both descriptions are available, and in this case it leads to an important loss in coding efficiency. In order to improve the coding efficiency in cases where the error probability is not severe, [Y.-C. Lee 2004] have proposed an alternative three prediction loop. The redundancy is reduced by using the side

information to improve the quality of the central description in error-free conditions.

These schemes have two redundancy components. The implicit redundancy introduced by the multiple description technique, and redundancy introduced by the side information,and duplicated information, such as motion vectors and headers. Rate-distorion optimization methods are needed in order to optimally define the MDC parameters (for instance, in MDCT is the correlation between descriptions) and side information rate usually defined by quantization step-size [Liu 2005].

3.2.3 High Dimensional MDC

High Dimensional MDC is achieved by using pre-processing to increase the temporal and spatial down-sampling. Increasing the number of descriptions has the advantage of improving the concealment of missing descriptions, because the information lost in any single description does not have the same importance than if lower number of descriptions are used. Still, increasing the number of descriptions, leads to lower coding efficiency, which means that the overall redundancy increases. For this reason, the overall quality is not improved comparing with MDC schemes with two descriptions. Nevertheless, due to its high scalability level, MDC with high number of descriptions is an interesting approach in application scenarios where the source video is highly distributed.

In the recent past, several works have been proposed in order to increase the number of descriptions at source level. In [Franchi 2005] polyphase downsampled is used order to generate four descriptions. Recently, a hybrid spatial/frequency downsampling scheme is proposed where the source data is spatial downsampled followed by the splitting of *AC* DCT coefficients [Hsiao 2010]. Another approach is proposed in [Tsai 2010] where four descriptions are generated by using spatial and temporal dowsampling, including also error concealment tools in order to recover lost descriptions.In [Tillo 2010] a multiple description scheme based on multirate coding for JP2000 and H.264/AVC is proposed where the source data is divided into subsets and then each one is encoded at several rates to properly assign subsets at different rates to the descriptions. In [Zhu 2009], four temporal shifted versions descriptions are proposed using hierarchical B-frames.

3.2.4 Multiple Description Scalable Coding

The scalable nature of the motion compensated video codecs based on sub-band/wavelet transforms were also exploited in order to obtain multiple descriptions. A few proposals have been published in the literature joining open-loop scalable video coding with multiple description coding. Bajic and Woods [Bajic 2003], apply data-partitioning approaches to the wavelet coefficients and motion vectors produced by the Motion Compensated Temporal Filter (MCTF). Van der Schaar and Turaga [van der Schaar 2003] propose splitting the sequence of temporally filtered frames to create multiple representations of the input. Critical data, such as MVs and the low-pass temporal band are duplicated in all descriptions. Another proposal in [Tillier 2007] employs a redundant motion-compensated scheme derived from the Haar multiresolution analysis, to build temporally correlated descriptions in a t+2D video coder where the allocated redundancy decreases with the resolution level. This is achieved by additionally subsampling some of the wavelet temporal subbanbds. Verdiccio *et al.* [Verdicchio 2006] propose an Embedded Multiple Description Scalar Quantiser applied to the spatiotemporal decomposition of the input. This results into a number of descriptions for each subband that are independently decodable and mutualy refine the accuracy of the reconstruction. An MDC scheme based on 3D-scan based wavelet video coder is proposed in [M. Pereira 2003], using an optimal redundancy allocation method that defines the scalar quatisation parameters of each wavelet subbands in order to minimize the central distortion.

Other methods aim to combine the advantages of Layered Coding and MDC, defining the technique as Layered Multiple Description Coding. Using the scalable structure of base and enhanced layers, multiple descriptions are used either in the base or enhanced layers or even in both of them. For instance in [Wang 2003] the base layer is coded in two descriptions and three versions of the enhanced layer are encoded, one for each prediction at decoder.

In [Chen 2007a] FGS scalability scheme is used with an open loop MDC scheme for both layers. Other distinct approach was proposed in [Kondi 2005] where only the enhanced layers are multiple description encoded. Using the scalable extension of H.264/AVC (SVC) multiple description based on distinct quality resolution is proposed in [Abanoz 2009]. This defines two balanced descriptions, each one scalable encoded, where several combinations of video segments (GOP or frames) are coded at high and low rates. An SVC-based proposed in [Favalli 2011] to encode the polyphase subsampled versions of the source

video.

The use of hierarchical B-frames in MDC is exploited in [Zhu 2009], where two descriptions are generated by duplicating the original sequence and then coded in hierarchical B structures with staggered key frames in the two descriptions. By using different QPs at different levels, their approach enables each frame to have two different quality fidelities in different descriptions. Another approach using the same scalability coding structures to define different levels of redundancy according to the hierarchy level, was recently proposed in [Tsai 2012]. In this approach, the authors use unequal redundancy allocation to encode frames with different fidelity requirements.

3.2.5 Stereo/Multiview MDC

Considering the multi-view video coding, MDC is an efficient method for transmission over lossy channels. MDC applied to multi-view video coding exploits the additional dimension that is the interview redundancy. The simpler MDC scheme that can be considered is the simulcast coding, where each view is independently coded. The introduced redundancy is the loss of the interview prediction coded efficiency and also the stereo vision for the case where only one description (view) is available at decoder. An improvement of the simulcast scheme, was proposed in [Norkin 2006] where two MDC approaches have been introduced: spatial scaling stereo-MDC and, multi-state stereo MDC schemes. The first approach consists in encoding a redundant left (L_{red}) sub-sampled view predicted form the main right view(R_{main}) and vice-versa, *i.e.*, encode redundant right (R_{red}) sub-sampled view predicted form the main left view(L_{main}). The redundancy is added to this MDC scheme by two components. Firstly, each main view is predicted as monoscopic video losing the interview correlation and, secondly, the sub-sampled versions are totaly redundant if both descriptions are available. The second approach consists in temporally split the stereo pair whose frames are then independently encoded. The additional redundancy is due to the temporal distance increase with the corresponding loss of coding efficiency.

Another proposal is an 3-D even and odd MDC scheme with adaptive side information (redundancy) which is added to frames with motion activity higher than a threshold [H. Abdul Karim 2009]. The authors in [Milani 2010] combine multiple description coding for video+depth using spatial down-sampling and Wyner-Ziv coding according to the characteristics of the video sequence and to the channel conditions. In [Erhan Ekmek-

cioglu 2010], a MDC is presented using video+depth coding, with viewpoint synthesis. Scalable coding is used and each of the two viewpoints is encoded with two spatial layers and two temporal layers. In addition to the scalable layers, a redundant bitstream for the base layer is generated. The redundant bitstream is generated such that only foreground objects that are more significant are redundantly encoded.

An MDC scheme based on scalable coding was also proposed in [M.B.Dissanayake 2010], where a redundant version of the enhanced layer is encoded. Only disparity vectors, intra residues and headers are encoded as redundant information and the interview residue are set to zero. A View Interpolation Based Multiple Description Coding of Multiview Images scheme is proposed in [Xiu 2009]. Each description is made by the even and odd views, and for each view, redundant information is encoded, which corresponds to the residue between the interpolation of even(odd) views from odd(even) views and even(odd) original ones. A rate-distortion optimization model for stereoscopic video streaming with unequal error protection is presented in [A. Serdar Tan 2009]. The coding structure uses scalability where each layer have different priority and protection.

Multiview video coding is inherently multiple description coding, since each view can be seen as a different description on the same scene. The use of the MDC paradigm for multiview video delivery is a promising application not only as a error resilience tool, but also as an network adaptation scheme. In fact, the idea of having different views of one scene distributed in different servers that can be used by each user without having the scalability constraints is an usefully application of MDC in a multiview scenario. The proposed literature in this field is still incipient and usually use the available coding tools as scalable coding tools in order to propose MDC schemes basically based on stereo video. Actually, the complexity of the multiview video coding tools is a restriction of the MDC development, but also is a challenge and is definitely a research open issue. For instance, the combination of different multiview coding formats for the same source video that can be combined in function of the user constraints is an open issue that have not been explored in the literature.

3.2.6 Unbalanced MDC

In general, most of the existing MDC architectures use $N = 2$ descriptions with approximately the same rate, *i.e.* balanced. However, due to non-stationarity characteristics of channel conditions in terms of available bandwidth and packet error rates, MDC encoders should dynamically adapt the transmission rate of each description in order to minimize the distortion at decoder.

Thus an MDC scheme should be able to produce asymmetric bitrate for each description as necessary, in order to improve the overall video decoded quality. Considering an MDC scheme with two descriptions, in unbalanced MDC (UMDC), each description is accepted with different resolutions and qualities for individual decoding.

Unbalancing between descriptions is archived by using distinct approaches such as adapting the quantisation, temporal or spatial resolutions. In [Apostolopoulos 2001b] an UMDC scheme is proposed where two descriptions are encoded with different frame rates but still maintaining the quality of individual frames. Such scheme introduce higher redundancy in the lower rate description because the temporal distance between frames mode motion information increase. In [Comas 2003], an unbalanced MD video coding with rate-distortion optimisation is proposed to generate two descriptions, where one provides high resolution (HR) and the other one low resolution (LR). The main problem of such approach is that the overall distortion corresponds to the HR description distortion, so the LR description is totally redundant, which means that the LR description does not improve the decoded sequence when both of them are received. A similar approach is proposed in [Bin 2006], for transmission using only one physical path. However, such approaches cannot be considered true MDC schemes because the overall quality is not improved when both descriptions are available. In [Kim 2005], an unbalanced multiple description coding framework for wavelet-based coders is proposed in order to cope with the time-varying network conditions and in [Cardinal 2004] an UMDC method using MDSQ with an optimal rate-constrained index assignment (IA) for asymmetric descriptions was proposed for a generic gaussian source.

An asymmetric rate control scheme was recently proposed in [Kamnoonwatana 2011], where the MDC redundancy is adaptively allocated based on an end-to-end distortion model. The overall rate is controlled by jointly setting the central and side distortion. Although the rates of the encoded descriptions may vary, the corresponding redundancy according to channel conditions, this method does not allow changes in each description

rate without severe degradation of the central distortion.

3.2.7 Multiple Description Distributed Video Coding

Multiple description based on DVC intends to exploit the error robustness of Wyner-Ziv (WZ) coding which is inherent to its channel coding approach. Several works have been proposed in order to join both coding paradigms. In [Milani 2010] an MDDVC based on redundant slices scheme is proposed where WZ slices are encoded with different lossy syndroms. The performance is increased over MDC based on redundant slices for high loss rates.

Although such approach could not be considered as a true MDC scheme, in [Rane 2008] a redundant slices scheme was proposed where the redundant slices are encoded as WZ frames. At the decoder in order to correct the lost primary slice, the ones received without errors, act as side information for WZ redundant slices.

In [Fan 2011] an MDDVC scheme based on a pair of staged quantizers and, consequently, a bitplane extraction scheme is proposed in order to make two WZ descriptions. The intra frames are encoding with MDCT. Although such scheme achieve drift reduction, its performance is only better than the classic MDCT scheme for higher packet loss rate due to the high level of redundancy that is introduced. Another proposal using DVC is presented in [Crave 2010] where for both each key frame and WZ-frames two descriptions are produced using MDSQ. Finally, a DVC scheme based on three-band motion compensated temporal filtering video encoder (MCTF) is proposed in [Crave 2008]. The original subband layers split into two descriptions. Each frame is conventionally encoded in one description and WZ-encoded in the other description. Comparing with conventional MDC encoder this proposal has better side-distortion performance due to channel error recovery feature of the WZ frames used only at side decoding.

The DVC approach in MDC field improves the error robustness due to channel encoding of WZ frames. However, the redundancy that must be used does not compensate the loss of central encoder quality, which means that only for high packet loss rates such approach could be competitive. Nevertheless, the MDDVC scheme is an interesting solution in applications where most of the computational complexity must be transferred to the decoder.

3.3 Multiple Description Video Streaming

Video streaming is a generic concept that can be used in order to express the transmission of video signals through packet networks. Real-time video transmission schemes can be grouped in two main categories: bi-directional and unidirectional, which includes generation and transmission in real-time. This can be further divided in live and on-demand. In live streaming, video signals are seen by clients in real-time from one or more servers, usually from a unique multicast session. In on-demand streaming, each user have a different and asynchronous streaming session from one or more servers.

Generally, streaming video is characterized to have a strict delay constraint that makes streaming services very sensitive to packet loss and network outages. For example, when receiving a streaming video session, data packets that arrive late are useless and thus considered as lost. Also, streaming video over wireless mobile networks, impose significant challenges to video transmissions since node mobility and lack of infrastructure in the network can lead to frequent link failures and route changes. Furthermore, the link quality is affected by fading and interference in the wireless channels. Since traditional approaches for dealing with packet loss, such as retransmissions, may not be possible in real-time streaming context or in wireless networks due to its error-prone nature, and also the vulnerability of compressed video to packet losses, additional mechanisms are needed to provide streaming media delivery over packet networks.

Due to these issues, the deployment of mechanisms to overcome the effects that errors and failures have on the streaming quality, is a mandatory requirement. In this context, path diversity is a powerful solution to overcome this problem. Therefore, MDC appears as high competitive source/channel coding approach in order to be applied in networks where its infrastructures must be highly dynamic. Considering this, MDC video streaming is included in a wide range of applications where path diversity networks are used to deliver compressed video either from the source (e.g., video codec, streaming server) to the final destination (e.g., user terminals) or from a network node to mobile terminal (e.g., wireless link with path diversity). In particular, MDC is regarded as an efficient error resilience coding scheme to overcome the lack of centralized QoS control that might exist in more classic streaming application scenarios.

The usage of MDC in path diversity video streaming is also a competitive alternative with other network-adaptation schemes, such as scalable video coding for robust streaming

(SVC) over time-varying networking conditions. In general, the main difference between SVC and MDC lies in the inter-dependency of SVC layers in contrast with independent MDC descriptions. Therefore, MDC clearly has the advantage of higher flexibility and independence from a common base layer, which also contributes to simplify the network management mechanisms [Chakareski 2005]. The comparison between both techniques is greatly dependent on the application scenario, but, generally, MDC gives a stronger error resilience protection in applications where feedback channels are not available and also in networks with either high packet loss rate or high dynamic topologies like P2P frameworks [Xua 2012].

Different streaming video architectures have been proposed but in general they can be grouped in two main areas: content delivery networks (CDN), which is the traditional client-server streaming approach, and peer-to-peer networks (P2P), where each client acts simultaneously as a peer to deliver video contents to other users. Traditional client-server based technologies offer the possibility to provide good performance and high availability rates. However, those approaches usually need high deployment and maintenance costs. These costs could increase drastically if more videos continue to improve to higher qualities. Therefore, several P2P streaming systems have been proposed to enable live and Video-on-Demand (VoD) streaming to large audiences with better video quality at low deployment and server costs [Liu 2008b] [Liu 2008a]. Furthermore, utilization of the P2P paradigm yields several advantages, such as, inherent bandwidth and resource scalability, network path redundancy, and self organization [Jurca 2007].

3.3.1 MDC Video Streaming over Content Delivery Networks

Content Delivery Networks are point-to-point communications that were developed to overcome network congestions and server overload when the number of users highly increases. In order to enhance the end user performance, contents are distributed over networks, at different levels, not only at network core of the topology, but also using catching of the popular contents on edge servers located closer to the users. The main advantage is that it prevents the server overload, since the replicated contents can be delivered from the closest edge server and not from the origin server and, secondly, the use of a shortest path reduce the request response time, the packet loss rate and the network resources.

Video streaming in content delivery networks using path diversity and MDC was proposed in [Apostolopoulos 2002], which extended the traditional streaming CDN approach (single description)(SD-CDN) to the multiple description case (MD-CDN). In fact, the traditional scheme deals with the problem that consists of finding the best server or edge proxy to feed one specific user. The optimization criterion is usually the shortest client-server path.

Fig. 3.10 shows an MD-CDN topology which allows to achieve path diversity. Different servers feed several clients and video content is distributed between them. Different paths can be defined in order to achieve a true path diversity. In fact, despite the existence of different servers, complete paths or parts can be the same for all descriptions. In that case if path fails all descriptions are lost and path diversity is not explored. So, MD-CDN poses different problems compared with SD-CDN. In case of MD-streaming, to design different design issues are posed: i) is how to distribute the servers in the network, ii) is how to distribute different descriptions among the available servers or proxies and iii), how to select for each client multiple servers with complementary descriptions and paths, in order to achieve a true path diversity.

In this context, MDC streaming offers the advantage of either using a distributed architecture with several servers at different locations. In [Apostolopoulos 2002] it is shown that MD streaming performs better than SD streaming using the same CDN infrastructure without any MD optimized server placement, with fewer number of servers for the same expected distortions. If the benefits of path diversity are considered taking into account the definition of disjointness ratio of the paths, then the reduction of distortion is significantly improved. Also [Begen 2005] shows the benefits of using an overlay infrastructure in order to achieve an optimal multi-path routing using disjoint paths for MD source coding.

The overall quality improves significantly considering the routing and the overall end-to-end distortion as optimization criterion to select multiple paths instead of using only the shortest path or maximally link-disjoint. Another approach is to adapt the characteristics of multiple description coding to channels conditions using unbalanced MDC [Kim 2005], or by using a multipath routing strategy based on bandwidth availability estimation and an hybrid assessment tool for evaluating the quality of experience (QoE) [Ghareeb 2010]. In [Ahuja 2008] the problem of optimal server placement in CDN using MDC is addressed based on the minimization of delays between client and its associated servers.

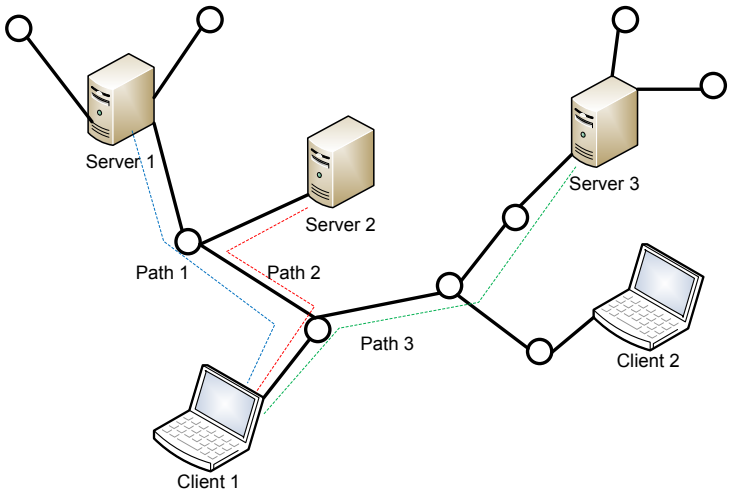

Figure 3.10: Path diversity in content delivery networks.

3.3.2 MDC Streaming over Peer-to-peer networks

Peer-to-peer (P2P) streaming networks are mostly characterized by having a fully decentralized architecture where all network nodes have equal functionality. The members are connected based on a specific construction policy forming a distributed network topology. On the other hand, streaming peers are usually heterogeneous in terms of their computing, storage capacity and network bandwidth. Also, the time and available resources that a peer node can contribute to the streaming topology are dynamic and unpredictable, which means that P2P members may join or leave the network at any time. Therefore, not only the streaming topology are very dynamic, but also no assumptions should be made about the availability of resources or network paths. Consequently, P2P streaming networks must be able to recover from the unexpected and ungraceful leave of any of its members at any time.

Different from other P2P systems like file sharing, for successful peer-to-peer video streaming, it must provide low end-to-end packet loss rate, delay, and delay jitter, similarly to point-to-point live streaming. In order to satisfy these requirements, the transmission bandwidths must be utilized effectively in bandwidth-scarce environments, and the system must adapt to changes in network conditions, such as increasing congestion in some of the peer-to-peer transmission paths.

In order to overcome the network layer limitations P2P streaming topologies are defined using an overlay layer, that is responsible for maintaining the network topology at every

instant, improving the dynamics of the overall topology. So, resilience in P2P networks is achieved by defining different redundant topologies to deal with peer nodes joining and leaving the overlay during the streaming session while still, maintaining the cohesion of the overlay structure, and minimizing the information loss due to node failures.

Path diversity can be achieved by different forms of exploiting the specific network topology, *i.e.* structured (tree-based) or unstructured (mesh-based). Usually in tree-based topology, path redundancy is achieved by setting a multi-tree structure in order to overcome the possible fail or congestion of one tree. In the mesh-based topology, several logical disjoint paths are defined in order to set distinct paths. In the case of P2P applications, the tree-based approach consistently exhibits an inferior performance over the mesh-based approach in their ability to cope with the instability of the peers, who join and leave the system, due to the static mapping of content to a particular overlay tree and diverse placement of peers in different overlay trees [Magharei 2007]. MDC-based streaming in a P2P network can be significantly better than in a CDN, despite the high degree of unreliability of the P2P systems [Khan 2004].

Tree-based streaming One of P2P topologies for video streaming is tree-based, where clients are assigned to a multicast session, usually with centralized management. One major drawback of tree-based streaming is their vulnerability to peer churn (peer leaves the network), because if one peer leaves the structure then, all the tree must be reconfigured in order to define a new topology. Multiple-trees are used in order to provide error resilience peer-to-peer tree-based streaming. Fig. 3.11 shows an example of redundant multiple-tree topology, using two redundant trees. Each description of MDC video is streamed from the server through different trees. For each peer, descriptions are delivered for different paths. This fact mitigates the problem of peer churn and reconfiguration because at least one description is available for each node.

This scheme was used in CoopNet system [Padmanabhan 2003], based on MD-FEC encoding scheme. In CoopNet, plays a central role in constructing and managing the distribution trees, trough a centralized management, which is responsible to construct and maintain a diverse set of trees. Also in [Castro 2003] a multiple tree based system was proposed and [Setton 2008] a P2P live multicast streaming base on multiple trees and rate distortion optimization scheduling. Also, in [Wei 2007] a multiple-tree transmission scheme for wireless ad-hoc networks were proposed using MDC as source coding.

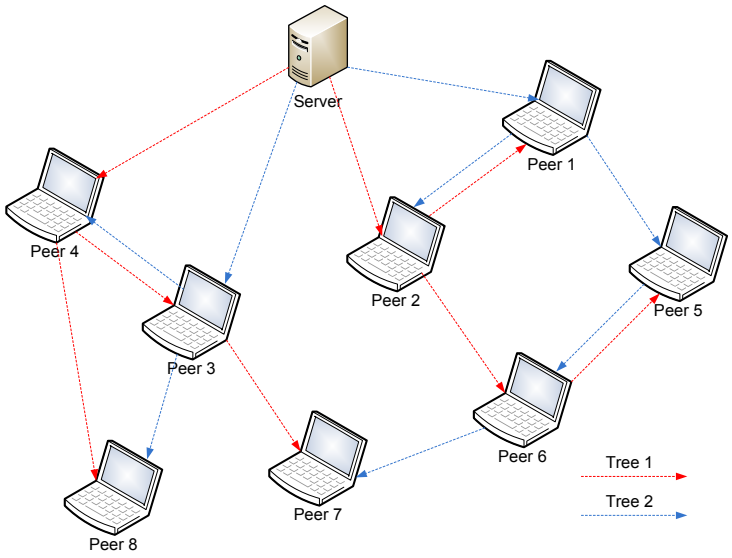

Figure 3.11: Path diversity by multiple-tree topology.

Mesh-based streaming In mesh-based P2P streaming system, peers are not confined to a static topology which improves the robustness to peer churns comparing with tree-based topology. Fig. 3.12, shows a mesh based P2P similar to that proposed in [Zha 2005], the CoolStreaming/Donet system. This is characterized by three key modules: 1) a membership manager, which helps the node to maintain information about other overlay nodes; 2) a partnership manager, responsible for establishing and maintaining the partnership with other nodes; 3) a scheduler which distributes video data among nodes. A P2P system using mesh based topology and using the Internet Protocol Television (IPTV) infrastructure for heterogeneous networks was proposed in [Lu 2007]. The system is based on a partnership formation to reduce disconnecting time and isolated peers and uses a dynamic MDC based on pre-processing downsampling which adapts description rates according bandwidth and device capability.

In [Mao 2005] the authors propose a meta-heuristic approach to solve the cross-layer optimization problem of routing, which minimizes the application layer video distortion, whilst in [Ma 2004], a multi-path selection scheme for overlay networks is proposed by defining a channel correlation metric as an additional QoS parameter to improve the final quality. More recently, in [Wei 2009] an MDC streaming scheme for multi-path wireless Ad-Hoc networks that minimize the concurrent packet drop probability is proposed.

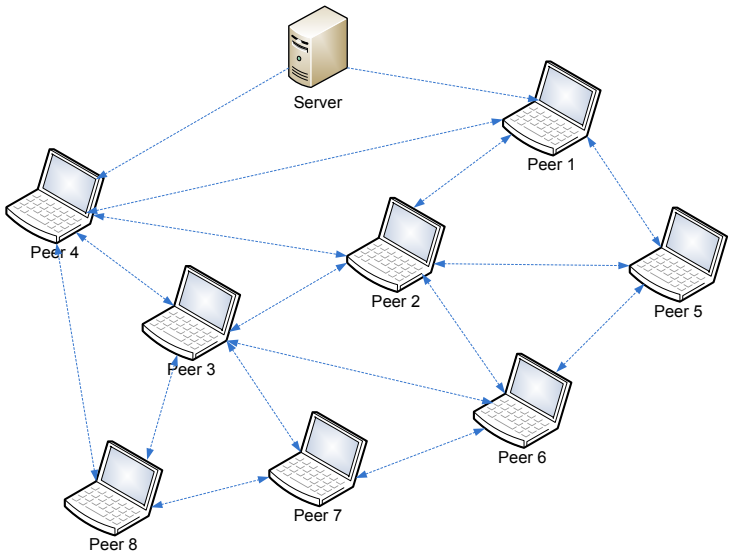

Figure 3.12: Path diversity using mesh topology.

3.3.3 Scheduling in path diversity networks

In path diversity networks, behind the streaming paths routing and average rate allocation definition, proper scheduling schemes of the media packets are need in order to have efficiency in video delivery. Neither all packets of video stream have the same importance which means that they do not equally contribute to the video quality at the receiving node. Furthermore, a packet is useful at decoder only if its delay is not higher than an acceptable limit, and also if previous packets have been correctly received. The unequal importance of video packets, along with timing constraints, requires the derivation of efficient packet scheduling algorithms that determine which packets should be transmitted at a given time instant on a given streaming path, in order to maximize the overall video quality.

One approach in to classify video data based on coding parameters and to add error correction capabilities according with their importance. For instance, in [Qu 2006] the FEC redundancy is set according to the motion activity classification of the source. Also in [Tillo 2011], an UEP scheme is used, where FEC redundancy allocation is a function of the relative importance of each source slice.

Considering packet scheduling solutions, these have been studied in client server architectures with a single channel where rate-distortion optimized packet scheduling strategies are used [Chou 2006]. These approaches have been extended for path diversity environment and using MDC. These were proposed in [Chakareski 2005], [Chakareski 2008],

[Chakareski 2006]. A quite different approach is the congestion distortion optimization scheduling for packet prioritization based on retransmissions in P2P streaming proposed in [Setton 2008] and [Xua 2012]. The importance of each packet is not only adjusted based on rate-distortion issues but also based on the numbers of descendent nodes in multicast tree that are affected by the loss or late arrival of this packet.

In [Milani 2008] a packet classification scheme for MDC video streaming, based on the number of zeros of each description, is proposed to match the QoS classes supported by the IEEE 802.11e MAC layer. A slightly different classification approach based on game theory is proposed in [Milani 2010] in order to cope with network congestion periods in a video streaming scenario where multiple descriptions are streamed from several servers. In [Rong 2010] a multipath routing over wireless mesh networks for multiple description video transmission is proposed, based on a guaranteed-rate packet scheduling algorithm with the aim of reducing the latency of video delivery.

Chapter 4

MDSQ for Advanced Video Coding

Contents

4.1 Balanced MDSQ **66**

4.2 Unbalanced MDSQ **67**

4.3 MD Video Architecture **69**

4.3.1 Open-loop MD Encoder 69

4.3.2 Open-loop MD Decoder 70

4.3.3 Drift Analysis 71

4.3.4 Multi-loop MD Encoder 74

4.3.5 Multi-loop MD Decoder 76

4.4 R-D Performance - Two Descriptions **77**

4.4.1 Balanced MDSQ 77

4.4.2 Unbalanced MDSQ 84

4.5 R-D Performance - Single Description **91**

4.5.1 Balanced MDSQ 91

4.5.2 Unbalanced MDSQ 101

4.6 Conclusions **106**

This chapter proposes a multiple description coding scheme for Advanced Video Coding based on a MDSQ mixed open-loop/multi-loop structure, which prevents drift distortion accumulation in both intra and motion compensated predicted slices. The drift is

compensated by generating a controlled amount of side information used by the decoder whenever any description is lost in the corresponding network path. In order to maintain a reduced redundancy cost, the side information is only applied to anchor slices, which allows to prevent drift propagation for single description decoding.

The proposed scheme is based MDSQ for both either balanced and unbalanced descriptions. New types of index assignment tables are proposed which can be used to change dynamically the rate of each description. The proposed method has also the ability to control drift distortion for side description decoding and also to control the amount of redundancy introduced by the side information. The proposed method allow to improve the MDSQ error robustness comparing with open-loop schemes by controlling a small amount of additional information.

4.1 Balanced MDSQ

Most of MD video encoders based on MDSQ are designed to generate two independent descriptions from the same source signal. The main principle is to construct two uniform quantisers with coarse reconstruction values such that separate inverse quantisation is possible to implement. If joint inverse quantisation of both quantised descriptions is performed, then finer reconstruction values are obtained. Considering only two different descriptions, the goal of MDSQ is to find coarse scalar quantisers for each side encoder such that after being combined together, a finer central quantiser is obtained, producing lower distortion than each individual side decoder. MDSQ is based on two different functions: central quantisation and index assignment. The index assignment is used for mapping each quantisation index of the original transform coefficients i_0 (i.e., central indices) into a pair of side indices (i_1, i_2) which are then entropy encoded. In general, MDC algorithms use balanced descriptions where the respective rates and distortions are approximately the same in all descriptions. Considering two descriptions $S1$ and $S2$, with rate-distortions pairs $(R1, D1)$ and $(R2, D2)$, the two descriptions are generated using side-quantisers from a central quantiser with rate-distortion pair $(R0, D0)$. For balanced MDSQ, $R1 \approx R2$ and $D1 \approx D2$.

The basis for proposed MDC for Advanced video coding follows the same approach as proposed in [Vaishampayan 1993] and explained in Chap. 2. An index assignment matrix is defined as shown in Table 4.1, whose elements are the SDC quantiser indices, i.e., central

indices, each one corresponding to a pair of side indices defined by the respective column and row. The amount of redundancy introduced by this MDC scheme is controlled by an index spread parameter k where $2k + 1$ is the number of diagonals of the index matrix. In Table 4.1, $k = 2$ which means that 5 diagonals are used.

Table 4.1: Balanced MDSQ,k=2

i1 (Description 1) \ i2 (Description 2)	-4	-3	-2	-1	0	1	2	3	4	5
-4	-20	-16	-14							
-3	-17	-15	-12	-8						
-2	-13	-11	-10	-6	-4					
-1		-9	-7	-5	-1					
0			-3	-2	0	1	4			
1					2	5	6	8		
2					3	7	10	12	14	
3						9	11	15	16	18
4							13	17	20	...
5								19	...	...

At the decoder, if both descriptions are available, then an inverse index assignment process is used to restore a unique central index i_0 to be inverse quantised and inverse transformed. In the case where any description is not available for decoding due to transmission errors/losses, the central index cannot be unambiguously determined because there are several possible values for each individual description index. For instance, if only description 1 is decoded and $i_1 = 2$, then the reconstructed index i'_0 can be any of the following values, $i'_0 = 3, 7, 10, 12, 14$. Using the choice of main diagonal index as reconstruction rule, then $i'_0 = 10$, which can be different from the original SDC value i_0. This leads to a decoding error in received indices and in the case of predictive encoding, as all current standard video coding schemes, such error originates a mismatch between the original SDC prediction loop and that of the decoder, producing drift.

4.2 Unbalanced MDSQ

Traditional index assignment tables are made in order to produce balanced descriptions as was shown in earlier sections. Encode unbalance descriptions using MDSQ for video is a new topic that has not been exploited in the literature,as much as the author know. The main purpose is, that joining two unbalanced descriptions, a good quality reconstruction is obtained, while a worst reconstruction is archived for one description decoding. In this

case a distinct side distortion is obtained for each description. Thus, unbalanced MDSQ has to generate two equivalent side encoders with shifted and different coarser step-sizes in order to produce different rates for each description and at the same time, generate an equivalent finer central step-size.

The specific characteristic of unbalanced MDC is to generate two descriptions $S1$ and $S2$, in which any rate-distortions pair $(R1, D1)$ and $(R2, D2)$ have $R1 \neq R2$ and $D1 \neq D2$. In MDSQ, the central quantiser is also characterized by a different rate-distortion pair $(R0, D0)$. In order to obtain unbalanced descriptions, a new type of index assignment table is proposed to allow dynamic changes in the number of central coefficients that are indexed as $i_n = 0, n = 1, 2$.

Table 4.2: Unbalanced MDSQ, k=1 and $Z = 0$

		i2 (Description 2)									
		-4	-3	-2	-1	0	1	2	3	4	5
i1 (Description 1)		-11									
	-3	-9	-7								
	-2	-8	-6	-5							
	-1		-4	-3							
	0			-2	-1	0	1	2			
	1							3	5		
	2							4	6	7	
	3								8	9	11
	4									10	12
	5									...	...

The proposed unbalanced index assignment method is defined by k, which in turn defines the number of diagonals $(2k + 1)$ and the central index spread variation parameter Z, respectively. The total spread is expressed by $S = 2 * (2k) + 1 + Z$. The spread S is the number of central indices that are coded as zero in a given description $i_n = 0, n = 1, 2$. Different values of S can be defined by varying Z and consequently several index assignment tables can be used to dynamically unbalance each description. Table 4.2 shows an example of the index assignment function with k=1 (3 diagonals) and $Z = 0; S = 5$ ($i_1 = 0$ for central indices $i_0 = -2, -1, 0, 1, 2$).

4.3 MD Video Architecture

Starting with the open-loop MDSQ video encoding, a novel MD architecture is proposed for Advanced Video Coding using MDSQ, based on the tree-loop scheme. A new mixed open-loop/multi-loop structure is proposed in order to best achieve a tradeoff between MDC coding efficiency at encoder and drift control at decoder. The proposed architecture use the three-loop scheme for anchor slices, namely, I and P slices, while B slices use the open-loop scheme without adding side information. This approach allows to reduce the side information redundancy, maintaining the drift control since B frames are not used as reference frames in most of the coding structures. In this way, the proposed architecture intends to deal with drift distortion that is propagated in MDSQ open-loop architectures, and at the same time, to reduce the redundancy introduced by the traditional three-loop architectures.

4.3.1 Open-loop MD Encoder

Fig. 4.1 shows the H.264/AVC open-loop MDSQ video encoding scheme. MDC function is implemented by inserting the MDSQ index assignment module in the encoding loop. The quantized DCT coefficients are mapped into two indices (i_1, i_2). These are entropy coding and the other encoded information, such as slice maps, prediction modes and motion vectors are encoded in each description, resulting in two independent decodable bitstreams.

The output rates of each description is dependent of the chosen MDSQ scheme. Tradicionally, these rates are balanced, i.e., the resulting rates have the identical rate and distortion levels. Nevertheless, new MDSQ schemes are proposed, where output rates can have distinct rates, i.e., unbalanced descriptions. The respective rates are defined using specific index assignment scheme, which poses the need to eventually to send additional information to decoder marking each index assignment table is use.

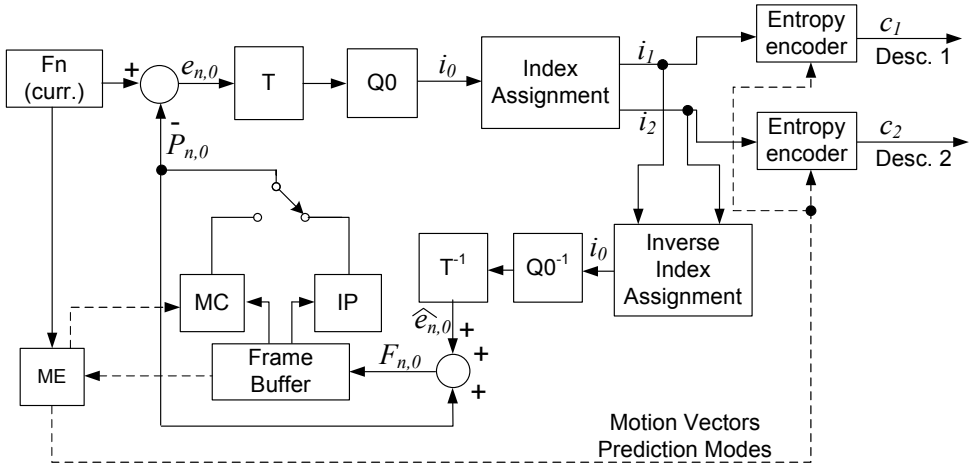

Figure 4.1: Open-loop MDSQ encoder.

4.3.2 Open-loop MD Decoder

The decoder architecture for MDSQ open-loop decoding is represented in figure 4.2. Descriptions 1 and 2 are composed by its main streams $c_i, i = 1, 2$. If both descriptions are available then an Inverse Index Assignment is made to restore the central description index, and then the inverse quantisation and inverse transform are applied. The decoded residue $\widehat{r}_0$, is added to the correspondent prediction. If some description is lost, the side decoder will decode $\widehat{r}_i, i = 1, 2$, and corresponding reconstructed frames $\widehat{F}_{n,i}$, with poorer quality than if both descriptions were received. The reconstructed frames $\widehat{F}_{n,i}$ can be defined as

$$\widehat{F}_{n,i} = \{P_{n,i} + \widehat{r}_i\}, i = 1, 2. \tag{4.1}$$

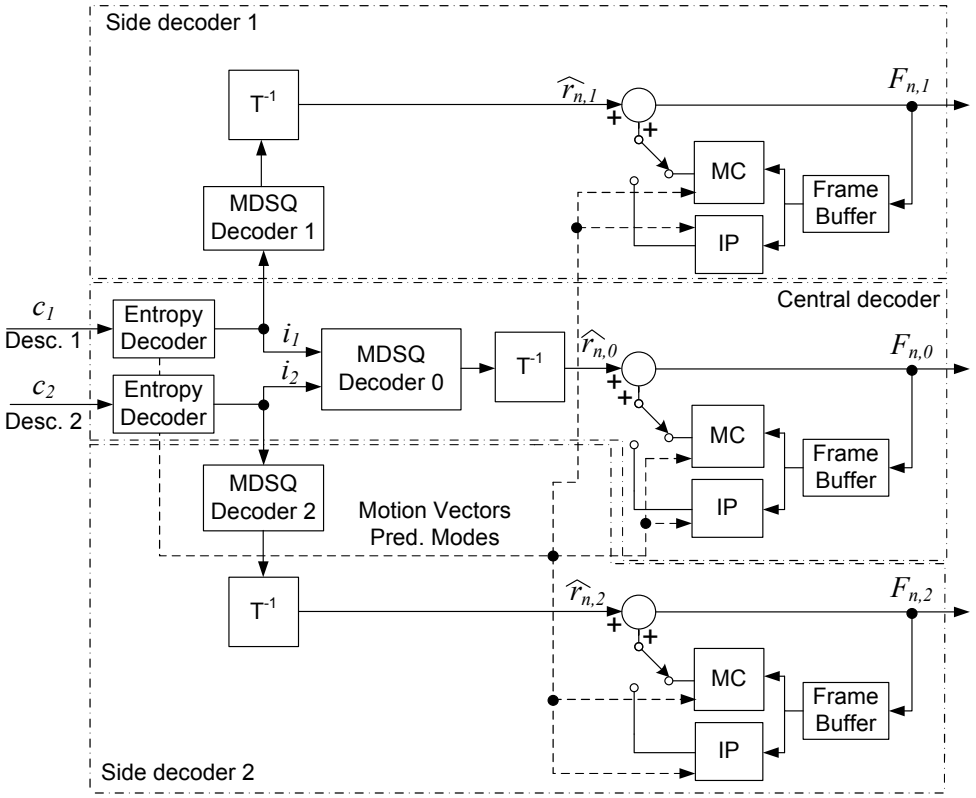

Figure 4.2: MDSQ open-loop video decoder.

4.3.3 Drift Analysis

If MDSQ open-loop encoder scheme of Fig. 4.1 is used for encode two descriptions of the same source video, then drift is introduced at the decoder whenever any description is lost. The drift distortion component can be determined from the relevant signals involved in encoding and decoding. When two descriptions are received, both side indices (i_1, i_2) are decoded and merged into the corresponding central index i_0. In this case, for each block n, the reconstructed central pixel values $F_{n,0}$ are given by,

$$F_{n,0} = r_{n,0} + P_{n,0} \tag{4.2}$$

where $r_{n,0}$ is the decoded residue and $P_{n,0}$ its associated prediction either from intra prediction or motion compensation, formed from decoding both descriptions. If only one description j (either $j = 1$ or $j = 2$) is decoded, then the reconstructed pixel values $F_{n,j}$ are given by,

$$F_{n,j} = r_{n,j} + P_{n,j} \tag{4.3}$$

where $r_{n,j}$ is the decoded residue and $P_{n,j}$ its prediction formed from decoding description j only. Since $r_{n,j}$ results from inverse index assignment using with only one description as input, the difference between the original SDC residue and that decoded from only one description produces a reconstruction error ε, i.e.,

$$r_{n,j} = r_{n,0} + \varepsilon \tag{4.4}$$

Substituting (4.4) in (4.3),

$$F_{n,j} = r_{n,0} + \varepsilon + P_{n,j} \tag{4.5}$$

and then using (4.2) in (4.5), $F_{n,j}$ becomes,

$$\begin{aligned} F_{n,j} &= F_{n,0} + P_{n,j} - P_{n,0} + \varepsilon \\ &= F_{n,0} + d_P + \varepsilon \end{aligned} \tag{4.6}$$

where

$$d_P = P_{n,j} - P_{n,0} \tag{4.7}$$

is the drift component due to mismatch between the predictions used in the encoder and those reconstructed at the final decoder from only one description. Note that the above analysis is valid for both the spatial and temporal drift components, though these can be identified as separate contributors to the overall drift distortion.

The actual impact of the drift component given by equation 4.7 in the objective video quality was experimentally evaluated for the open-loop scheme of Fig. 4.1. The results for intra predicted and motion compensated (MC) frames, are shown in Figs. 4.3 and 4.4, respectively.

4.3.3.1 Intra Prediction - I Frames

Fig. 4.3 shows the drift effect over one intra frame from the *coastguard* sequence, originally encoded using H.264/AVC with all intra prediction modes enabled. The Peak Signal Noise Ratio (PSNR) is shown for each macroblock decoded from only one description and also from both of them. The SDC trace provides a reference for comparison at the same rate as the single description (1.59 bpp). Fig. 4.3 shows that drift distortion introduced by open-loop MDSQ encoder of an I frame yields unacceptable quality when only one description is decoded. The more macroblocks are decoded, the higher is the accumulated distortion,

leading to continuous drop of PSNR along each row of macroblocks. The peaks in PSNR correspond to reset the accumulated drift distortion to zero at the beginning of each row of macroblocks because the first macroblock of each row is not predicted from the previous ones.

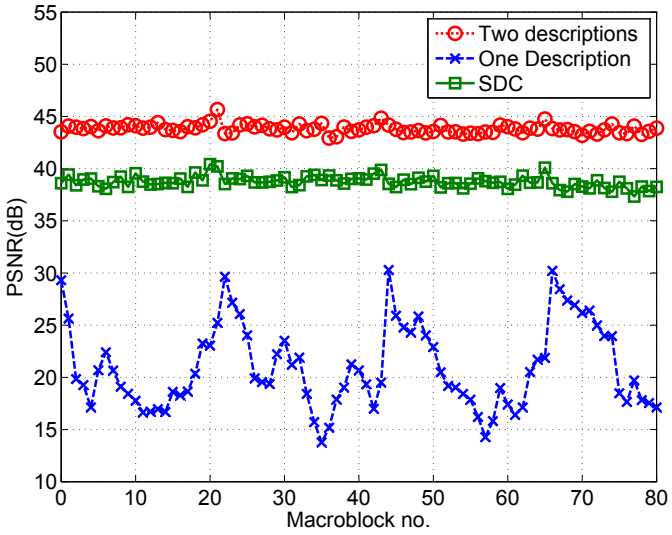

Figure 4.3: Distortion accumulation within an intra frame - *Coastguard*

4.3.3.2 MC prediction - P Frames

Fig. 4.4 shows the effect of drift accumulated over one GOP, comprised of one initial I frame followed by 20 P frames. Both descriptions of the initial I frame are fully decoded in order to not influence drift in subsequent P frames. The PSNR is shown for each frame decoded from only one description and also from both of them. For comparison, the PSNR of the SDC stream at the same bit rate as that of a single description (3.9Mbit/s) is also shown. Fig. 4.4 shows that drift distortion introduced by classic MDVS over one GOP also yields unacceptable quality when only one description is decoded. The effect of drift is quite evident from the rapid decrease of PSNR due to distortion accumulation along the GOP. Since the SDC stream has the same rate as the single description, the continuous decreasing of PSNR observed in the latter does not depend on the actual coding rate. Rather it is due to prediction mismatch between encoder and decoder.

The above analysis and results clearly show that open-loop SDC encoder cannot be used without drift compensation. An efficient solution is described in the next sections.

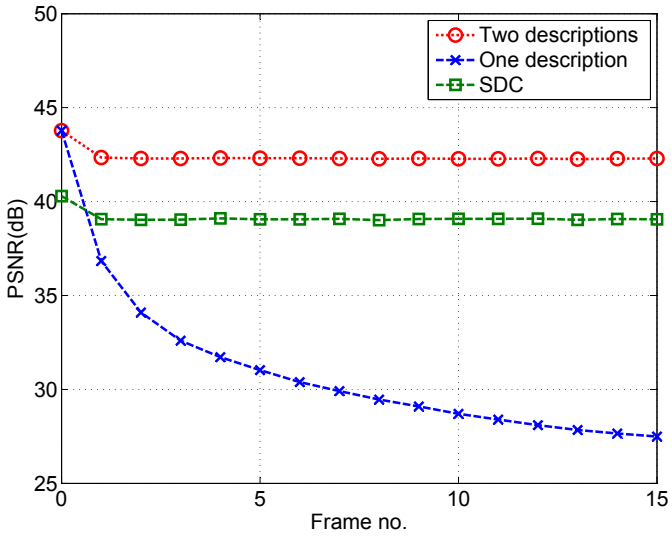

Figure 4.4: Distortion accumulation in MC predicted frames-*Coastguard*

4.3.4 Multi-loop MD Encoder

Figure 4.5 shows the proposed H.264/AVC drift-free MD video encoding scheme is based on a three prediction loop architecture. This MD architecture is based on MDSQ in order to generate two distinct independent decodable H.264/AVC bitstreams that can be combined in order to achieve different output quality levels. Each description is composed by correspondent MDSQ output and also by the side information that is generated for drift compensation for one description decoding. For I and P frames, the drift-free MD for advanced video encoder is composed by one central encoder and two side encoders. The central encoder includes the MDSQ module that produces two descriptions and the corresponding streams. These two distinct bitstreams are run-length and entropy encoded. All the syntax elements, namely headers and prediction modes are duplicated in both descriptions.

The two side encoders control the redundant information generated at the output in order to guarantee a drift free acceptable decoded frame. The side information s_i can be defined as,

$$s_i = Q_i\{T\{F_n - P_{n,i} - \widehat{r}_i\}\}, i = 1, 2 \tag{4.8}$$

Q_i is the quantisation operation with side quantiser QP_i that determines the amount of redundant information, T is the transform operation. F_n is the current frame and $P_{n,i}$

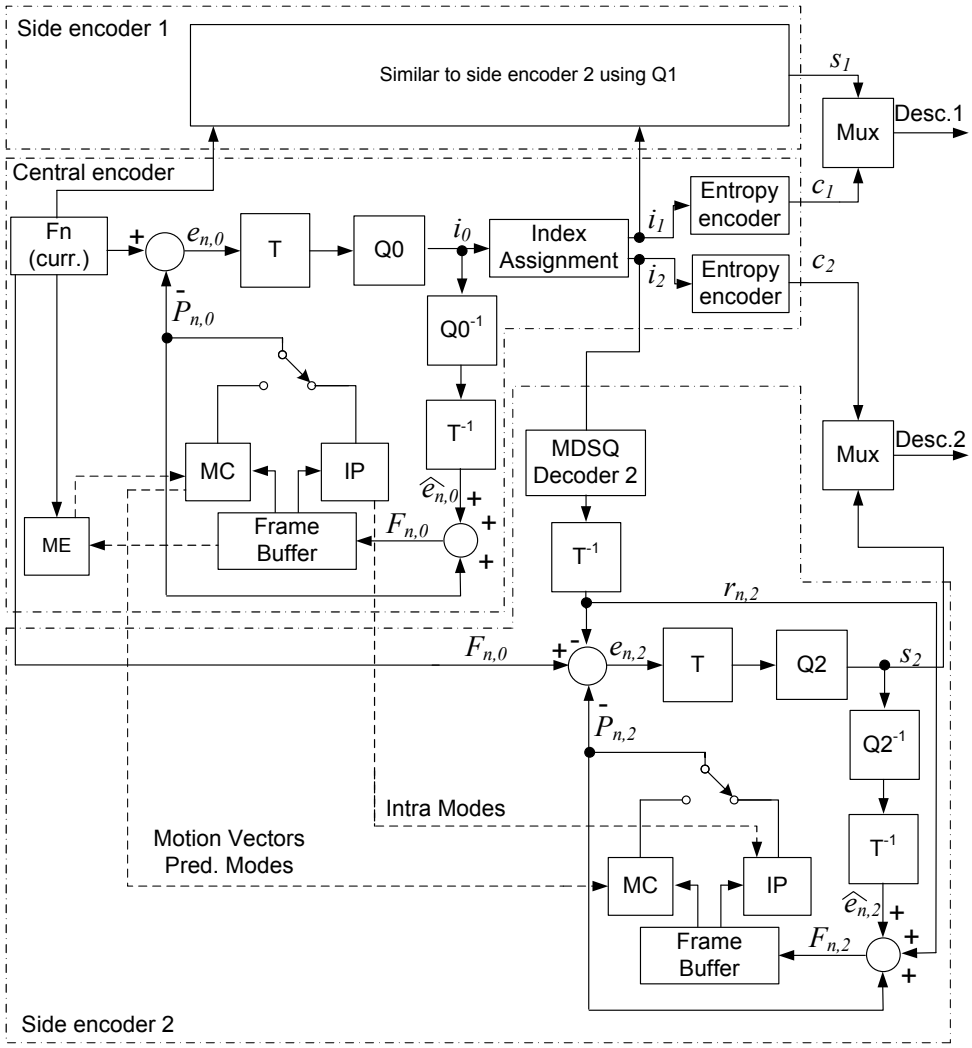

Figure 4.5: MDSQ encoder with drift compensation.

are the prediction value from each respective side encoders. $\widehat{r}_i$ is defined as

$$\widehat{r}_i = T^{-1}\{Q_0^{-1}\{A_i^{-1}(r_i)\}\}, i = 1, 2 \tag{4.9}$$

which represents the residue available at decoder if only one description is correctly received. A_i^{-1} represents the inverse index assignment operation if only one description exists. In this case it is necessary to define a reconstruction rule for each side index because each one corresponds to several possible central indices. In this work, the main diagonal central value is used as reconstruction rule. Q_0^{-1} is the inverse quantisation with the central encoder quantisation parameter QP_0.

4.3.5 Multi-loop MD Decoder

The decoder architecture is represented in figure 4.6. Descriptions 1 and 2 are composed by its main streams and correspondent side information $s_i, i = 1, 2$. If both descriptions are available then an Inverse Index Assignment is made to restore the central description index. Then inverse quantisation and inverse transform are applied. If some description is lost, for I and P frames, the side decoder will decode a drift free frame, though with poorer quality than if both descriptions were received. The reconstructed frame $\widehat{F}_{n,i}$ can be defined as

$$\widehat{F}_{n,i} = P_{n,i} + \widehat{r}_i + T^{-1}\{Q_i^{-1}\{s_i\}\}, i = 1, 2. \tag{4.10}$$

The drift free decoded frames are used to decode the subsequent central frames whenever both descriptions are correctly decoded. For B frames side decoding is used in open-loop, which means that for each decoded residue, $s_i = 0, i = 1, 2$ and mismatch between encoder and decoder is introduced. Drift propagation is mitigated by I and P frames, since usually B frames are not used as reference frames.

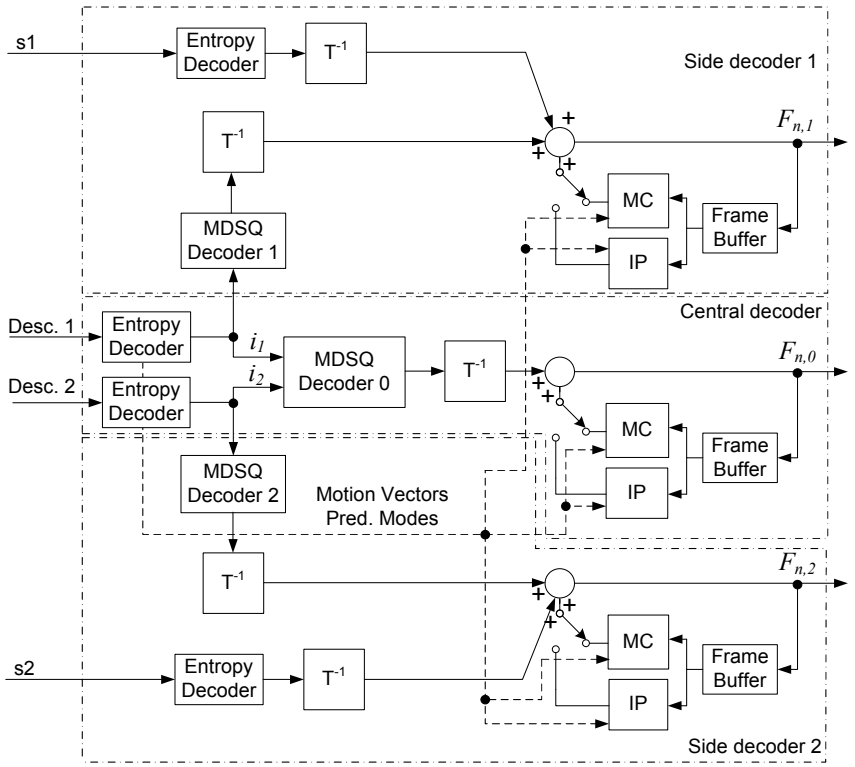

Figure 4.6: MDSQ video decoder.

4.4 R-D Performance - Two Descriptions

4.4.1 Balanced MDSQ

The R-D performance evaluation compares different balanced MDSQ index assignment tables with 2,3,5, and 7 diagonals for open-loop architecture described in section 4.3. MDC performance is compared with single description coding (SDC) and also with duplicated SDC. Headers, prediction modes and motion vectors are duplicated in both descriptions which are included in the overall rate. Evaluation is done using sequences *foreman* and *bus*, CIF format, 15Hz. Resulting rates are from coding with fixed quantization parameters QP between 10 and 40, without using rate-control schemes. Also, the redundancy evaluations is presented, where obtained rates of the MDSQ scheme is compared with SDC rates.

R-D for Intra Frames Figs. 4.7 and 4.8 shows the R-D performance for intra frames. Total rate is composed by duplicated headers and prediction modes and MDSQ indices for each description. From the graphical analysis, and taking as reference the duplicated SDC curve, the MDSQ redundancy is higher for lower rates compared with higher rate regime. This can be explained due to the fact that lower level coefficients mainly depend of quantization parameter instead of the MDSQ scheme. On the other hand, comparing the different index assignment tables some conclusions can be taken. First is that rates for index assignment matrices with more than 3 diagonals produce approximately the same rate. Index assignment with 2 diagonals has higher redundancy which is significant for sequence *bus* as is seen in Fig. 4.8. Independently of MDSQ scheme is noted that for higher rates lower redundancy is produced.

Redundancy is the excess rate introduced by MDSQ scheme, which means that in percentage,

$$\rho = (\frac{R_1 + R_2}{R_{SDC}} - 1) \times 100 \tag{4.11}$$

where $R_i, i = 1, 2$ is the rate of each description and R_{SDC} is the rate of the SDC stream. Other perspective of earlier results are represented in Figs. 4.9 and **??** where redundancy percentage levels are shown for different quantization parameters. Is shown that for QP lower than 25, overall redundnacy is maintained below 30% for 3,5 and 7 diagonals. For 2 diagonals, these values are around 45%. For higher QP values, are relatively low for lower

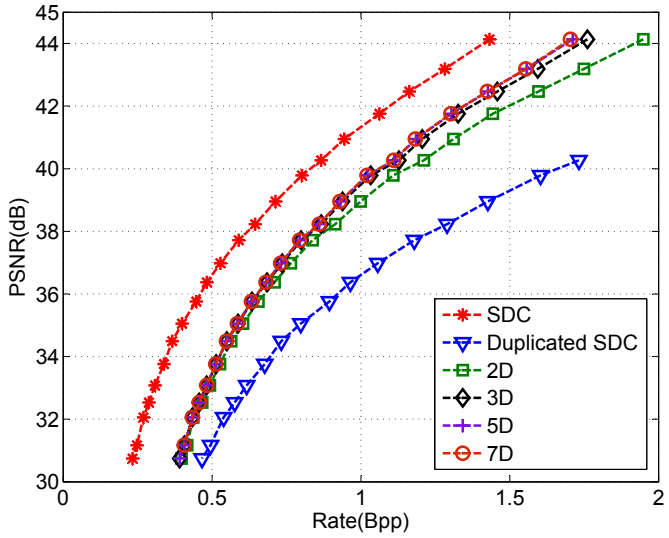

Figure 4.7: Central Distortion *Foreman* sequence, I Frame

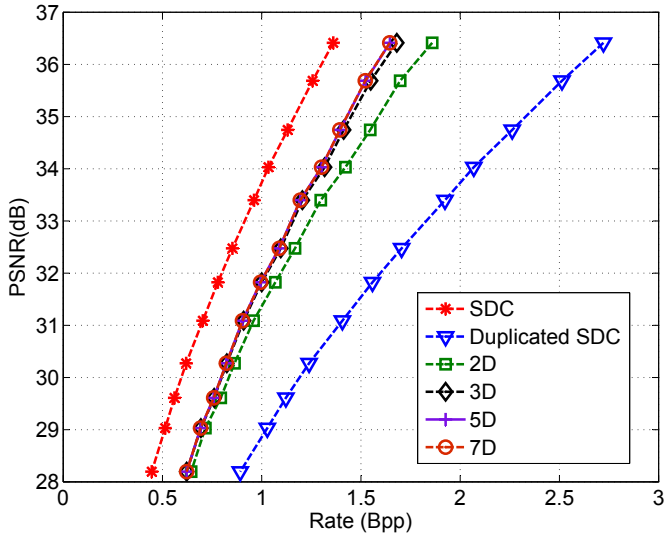

Figure 4.8: Central Distortion *bus* sequence, I Frame

quantization parameters. For higher rates, redundancy increases with lower number of diagonals. Also, is seen that for higher QP's the redundancy are around 50-60% and are independent of the number of diagonals that are used. These is due to the fact that for higher QP's, side quantizer does not change with different diagonals because central step-size is sufficiently larger that does not allows to improve the coding efficiency of MDSQ side encoder.

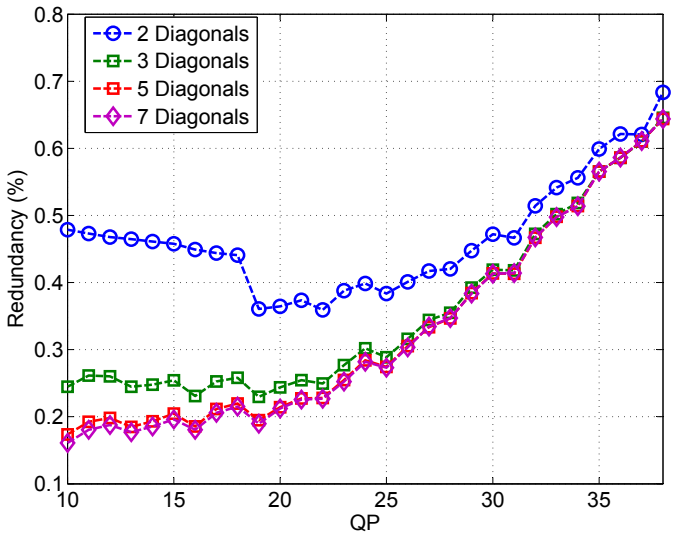

Figure 4.9: Redundancy of *Foreman* sequence, I Frame

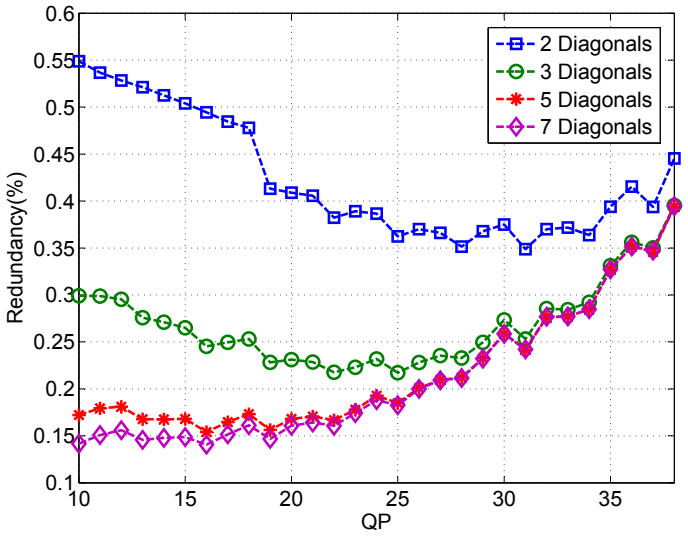

Figure 4.10: Redundancy of *bus* sequence, I Frame

R-D performanced for IPPB GOP In Figs. 4.11 and 4.12 R-D performance for IPBBP GOP structure. For this case overall rates include the motion information for both descriptions that are duplicated. Comparing the different MDSQ schemes, figures show that R-D performance for index assignment matrices with 3, 5 and 7 diagonals are similar. Using 2 diagonals, the redundancy increase comparing with the other MDSQ schemes. Therefore, MDSQ with 2 and 3 diagonals can be a good solution to use in advanced video coding, because is expected that have better side-distortion performance comparing with schemes with mode diagonals, as will be shown in the next section. On the other hand, for lower rates, the motion information contribution for the overall rate increases and resultant rates approximates to the duplicated SDC, which is limit for MDC R-D overall performance. Overall, MDSQ are more suited for high rate regimes, where finer step-sizes are used and more differentiation between methods and redundancy can be archived. Nevertheless MDSQ gives the tools in order to be used as a valid MDC technique for video coding for a wide rage of rates and redundancy.

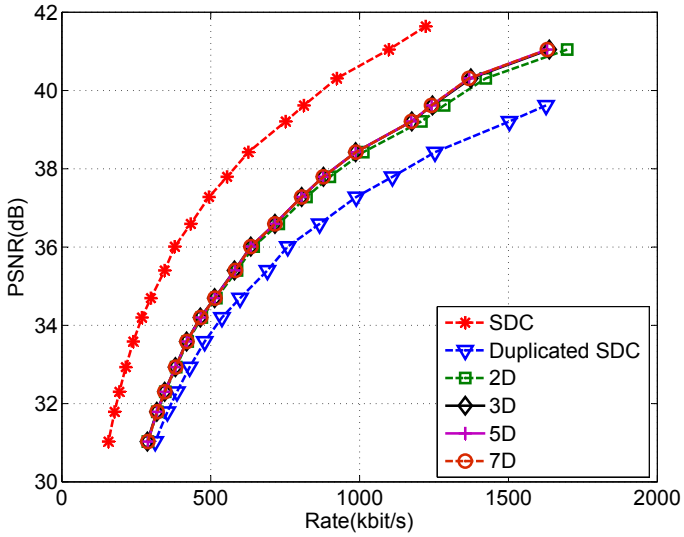

Figure 4.11: Central Distortion *Foreman* sequence, IPBB Frame

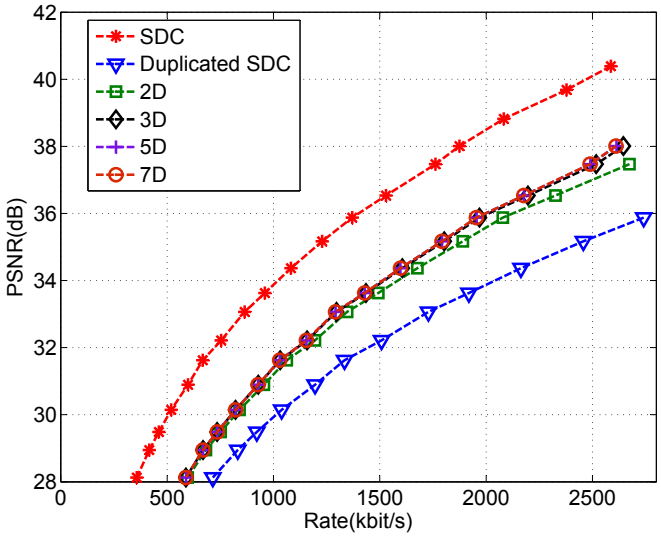

Figure 4.12: Central Distortion *bus* sequence, IPBB Frame

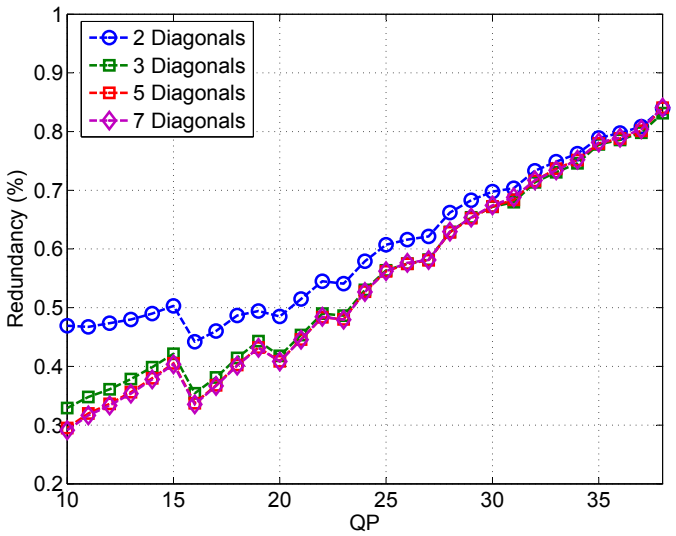

Figure 4.13: Redundancy of *Foreman* sequence, IPBB Frame

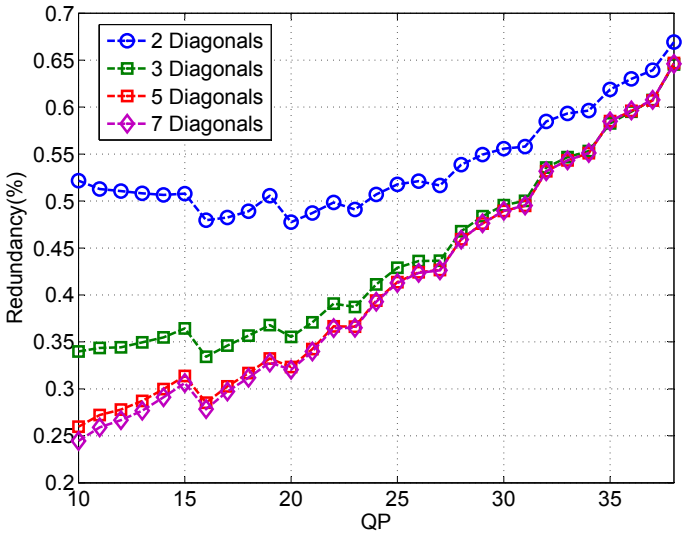

Figure 4.14: Redundancy of *bus* sequence, IPBB Frame

4.4.2 Unbalanced MDSQ

Figures 4.15 and 4.16 show the balancing rate percentages considering the total rate $R = R1 + R2$ and $R1 \neq R2$ as functions of Z parameter. The I,P and B frame types are analyzed and headers, motion vectors and macroblock modes are duplicated. It is noted that varying the Z parameter from -1 to 3, the rates of each description are set to different balances. Analyzing the results we can conclude that the balancing among descriptions can reach variations in about $10 - 15\%$ for several values of Z. Defining (π_1,π_2) as the balancing percentages of the total rate for description 1 and description 2, respectively, this means that the proposed set of index assignment tables are capable of producing unbalanced descriptions between $(50, 50)$ (balanced case) to $(25, 75)$ for different types of frames.

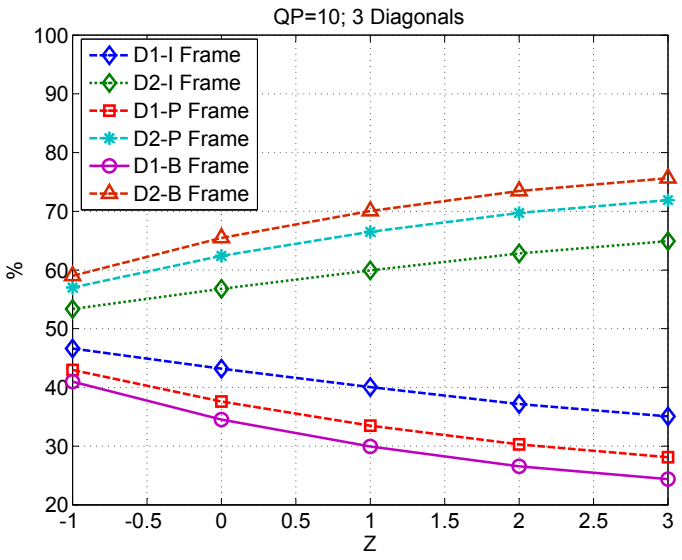

Figure 4.15: *Bus* sequence: QP=10;k=1

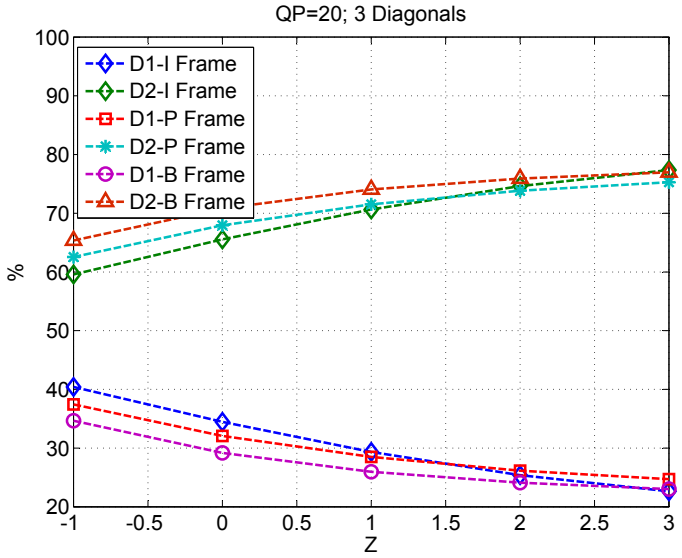

Figure 4.16: *Bus* sequence: QP=20;k=1

The R-D performance evaluation is done comparing different unbalance rates for MDSq with 3 diagonals. MDC performance for unbalanced MDC is compared with balanced MDC, single description coding (SDC) and duplicated SDC. Headers, prediction modes and motion vectors are duplicated in both descriptions which are included in the overall rate. Evaluation is done using sequences *foreman* and *bus*, CIF format, 15Hz. Resulting rates are from coding with fixed quantization parameters QP between 10 and 40, without using rate-control schemes.

Intra Frames Figs. 4.7 and 4.8 shows the R-D performance for only intra frames. Total rate is composed by duplicated headers and prediction modes and MDSQ indices for each description. Some conclusions can be taken from the graphical analysis. First is that taking as reference the duplicated SDC curve, the MDSQ redundancy is higher for lower rates compared with higher rate regime. This can be explained due to the fact that lower level coefficients mainly depend of quantization parameter instead of the MDSQ scheme. On the other hand, comparing the different index assignment tables some conclusions can be taken. First is that rates for index assignment matices with more than 3 diagonals produce approximately the same rate. Index assignment with 2 diagonals produce higher redundancy which is significant for sequence *bus* as is seen in Fig. 4.8. Independently of MDSQ scheme is noted that for higher rates redundancy redundancy is reduced. These is due to the fact that for higher QP's, side quantizers does not change with different diagonals because central step-size is sufficiently larger that does not allows to improve

the coding efficiency of MDSQ side encoder.

IPPB GOP Figs. 4.22 and 4.22 show the R-D performance for IPBBP GOP structure. For this case overall rates include the motion information for both descriptions that are duplicated. New unbalanced MDSQ R-D performance is evaluated for several central index spread variation parameters Z in index assignment matrices with 3 diagonals. $Z = 0, 1, 2, 3$ are used. Comparing with Balanced MDC, the lower rates are achieved comparing with balanced MDSQ. Comparing the unbalancing rate defined by Z, Figs. show that MDSQ redundancy decrease with the spread Z. For lower rates, as in balanced case, the motion information contribution for the overall rate increases and resultant rates approximates to the duplicated SDC, which is limit for MDC R-D overall performance.

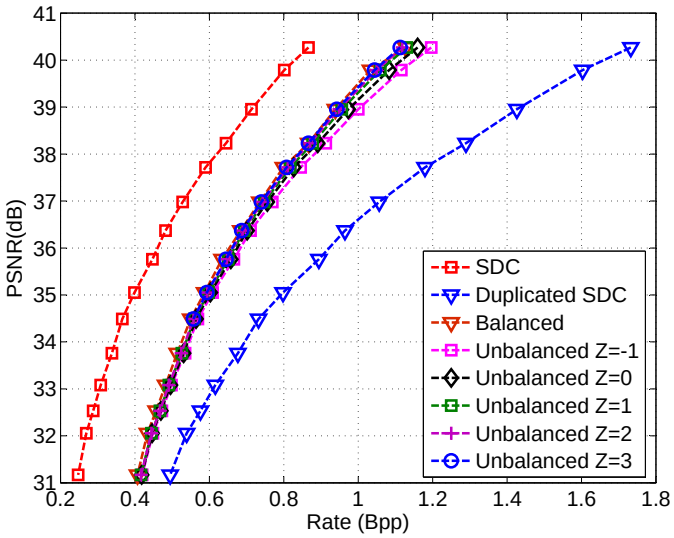

Figure 4.17: Central Distortion *foreman* sequence, Intra Frame

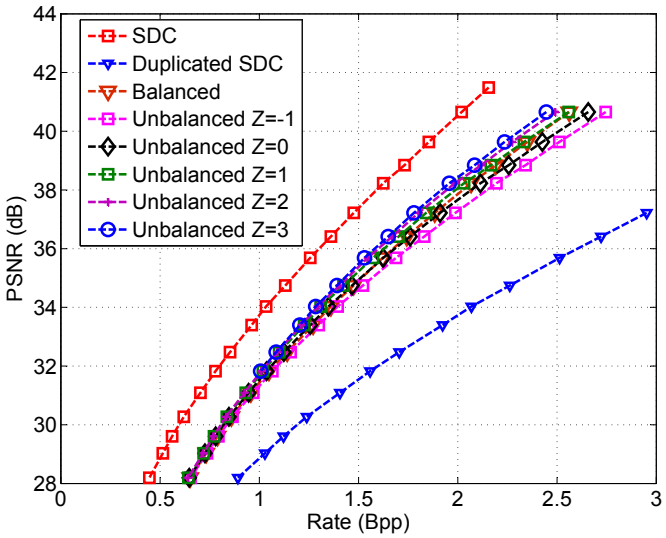

Figure 4.18: Central Distortion *bus* sequence, Intra Frame

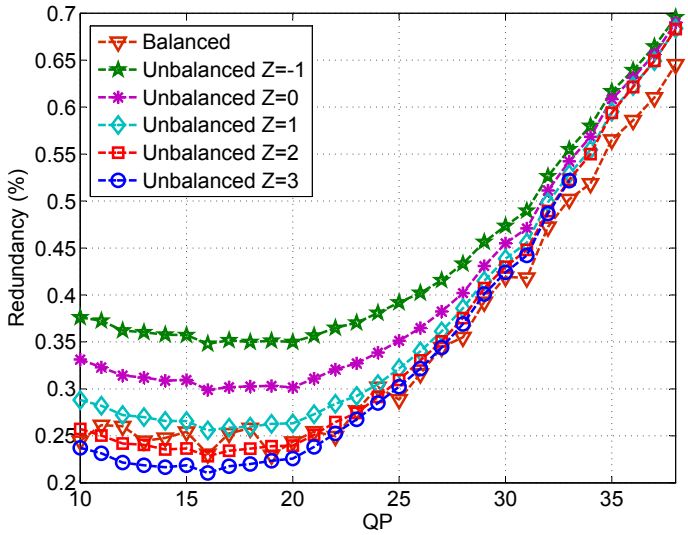

Figure 4.19: Central Distortion *foreman* sequence, Intra Frame

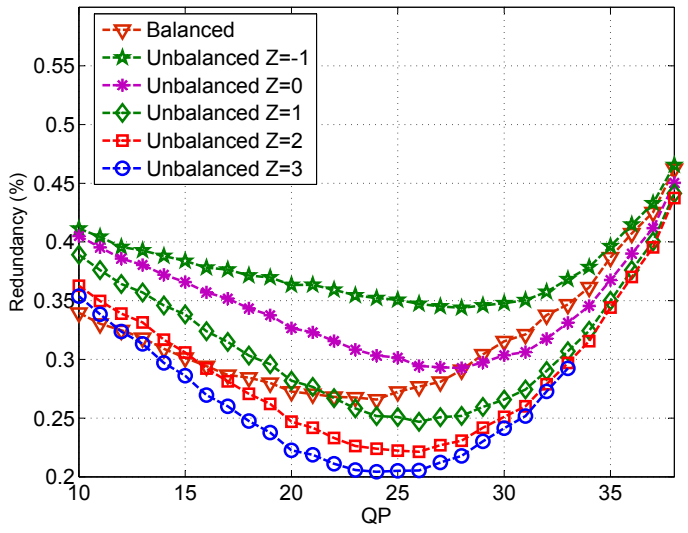

Figure 4.20: Central Distortion *bus* sequence, Intra Frame

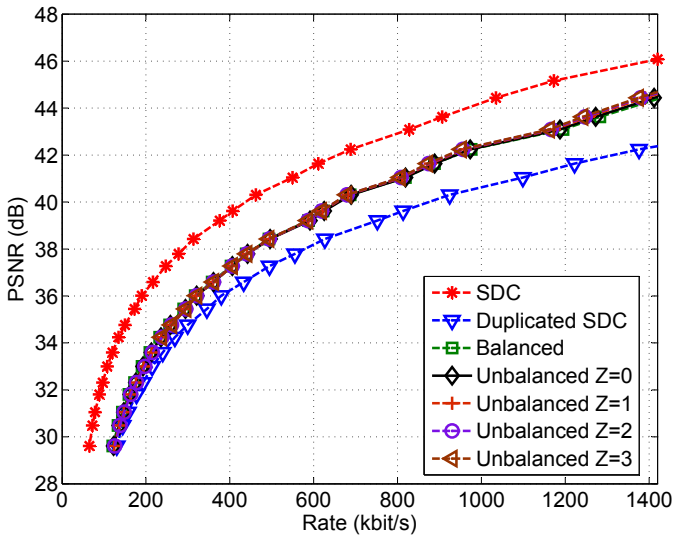

Figure 4.21: Central Distortion *foreman* sequence, IPBB Frame

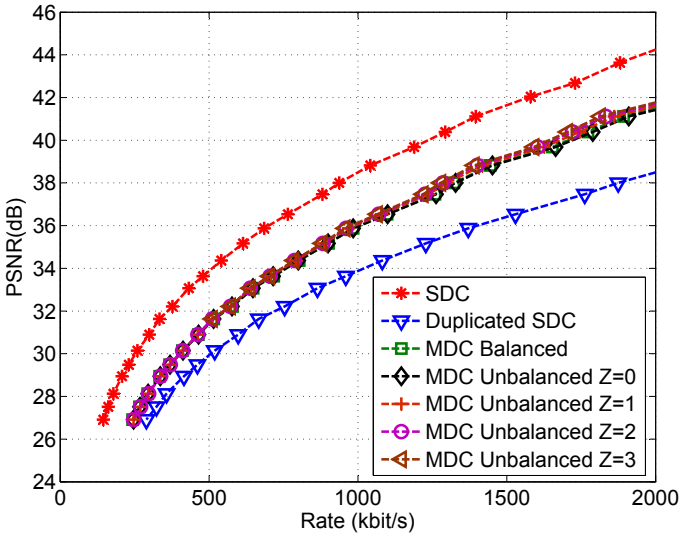

Figure 4.22: Central Distortion *bus* sequence, IPBB Frame

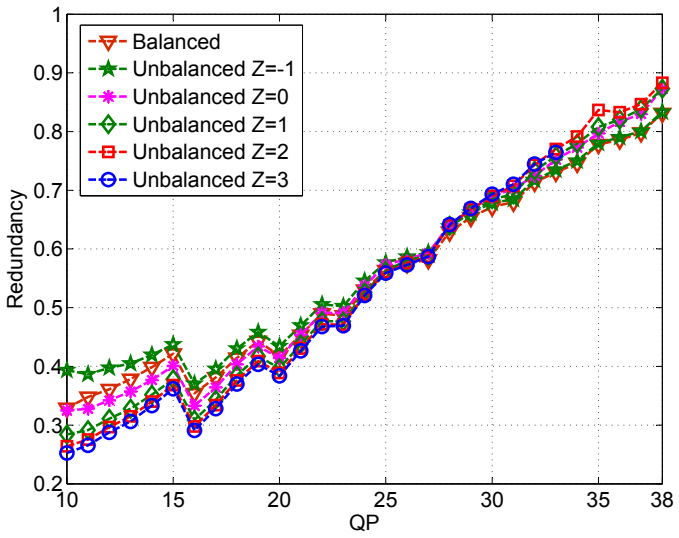

Figure 4.23: Redundancy of *foreman* sequence, IPBB Frame

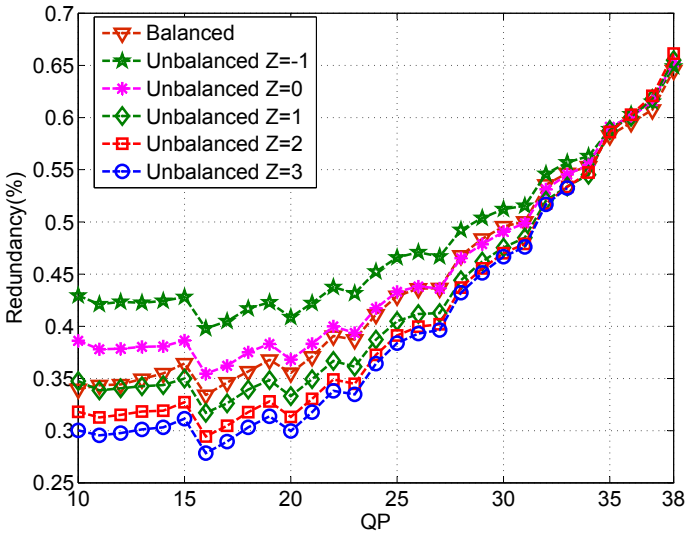

Figure 4.24: Redundancy of *bus* sequence, IPBB Frame

4.5 R-D Performance - Single Description

4.5.1 Balanced MDSQ

The side distortion evaluation define the performance of one MDC scheme and correponds to the PSNR that is obtained when only one description is decoded. The proposed scheme is compared with the open-loop MDSQ decoding. Excess rate introduced by side information is only used for side decoding, so it is important to study the gains that the additional redundancy in overall decoding quality. This study is done in Intra slices and also for IPPB GOP structures. Proposed MDSQ for advanced video coding scheme is compared with: MDSQ open loop, MDSQ open loop with I repetition and SDC. The methods comparison is made for the same rate.

4.5.1.1 Intra Slices

Intra prediction in Advanced Video Coding poses new challenges to MDC schemes, mainly because drift distortion is very significant as was shown in earlier section. The proposed MDSQ architecture for intra frames is an efficient solution in order to overcome this problem, by adding side information that allow to have an drift-free reconstruction with an acceptable quality at decoder when only description is available. A study of intra drift-free MDSQ architecture is done for different MDSQ schemes, with index assignment matrices with 2,3,5 diagonals, where the effectiveness of the solution is shown comparing with the open-loop architectures. Evaluation is done using sequences *foreman* and *bus*, CIF format.

MDSQ with 2 Diagonals Figures 4.25 and 4.26 show the proposed Intra MDSQ scheme compared with MDSQ Intra open loop scheme using 2 diagonals. The results of drift-free proposed solution are obtained using different levels side information for a given central distortion. These results show PSNR gains in around 10dB, depending of the side information level. One can see the open-loop results are very low quality, which suggests that are highly dependent of drift.

MDSQ with 3 Diagonals Figures 4.27 and 4.28 show the proposed Intra MDSQ scheme compared with MDSQ Intra open loop scheme for 3 diagonals. The results of

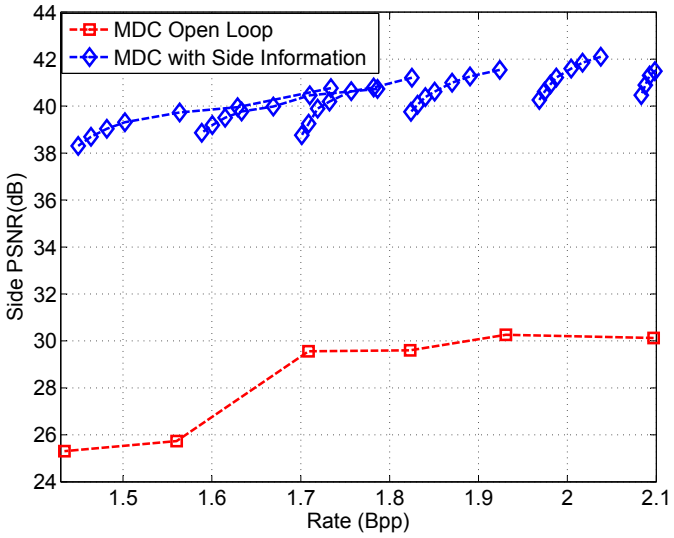

Figure 4.25: Side-distortion *Foreman* sequence, Intra Frame CIF, MDSQ with 2 diagonals

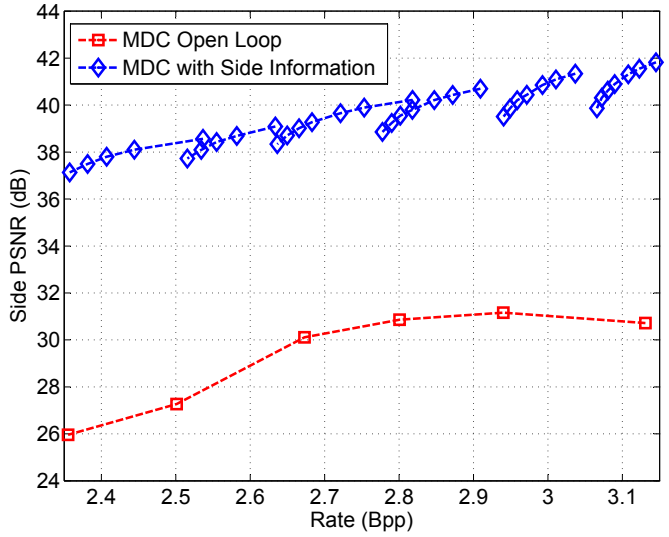

Figure 4.26: Side-distortion *bus* sequence, Intra Frame CIF, MDSQ with 2 diagonals

drift-free proposed solution are obtained using different levels side information for a given central distortion. The proposed solution show gains in around 8-10dB, depending of the side information level. One can see the open-loop results are very low quality, worse that using 2 diagonals. This is because due the fact that with higher number of diagonals, the open loop side-distortion increases, and the overall quality decreases.

Also, comparing the achieved side-PSNR with 3 diagonals with 2 diagonals, lower PSNR levels are obtained. This is due to the fact that with more diagonals, the prediction error that is encoded as side information increases, and consequently, a lost of coding efficiency is obtained. Also, the prediction modes of side-information are the same of the original ones, each means that also some coding efficiency is lost due to this fact. Nevertheless, the gains are sufficiently expressive, to conclude that proposed scheme is an effective solution in order to have side reconstructions with an acceptable quality, when MDSQ with 3 diagonals is used.

MDSQ with 5 Diagonals Figures 4.29 and 4.30 show the proposed Intra MDSQ scheme compared with MDSQ Intra open loop scheme for 5 diagonals. The results of drift-free proposed solution are obtained using different levels side information for a given central distortion. The proposed solution show gains in around 14-16dB, depending of the side information level. One can see the open-loop results are very low quality, highly below to 20dB, which is an no acceptable quality for only one description reconstruction. This is because due the fact that with higher number of diagonals, the open loop side-distortion increases, and the overall quality decreases, and one can conclude obtained side-distortion and also, low redundancy reduction comparing with lower number of diagonals, one can conclude that the use of 5 diagonals are not suited to be applied in MDSQ for advanced video coding in Intra slices.

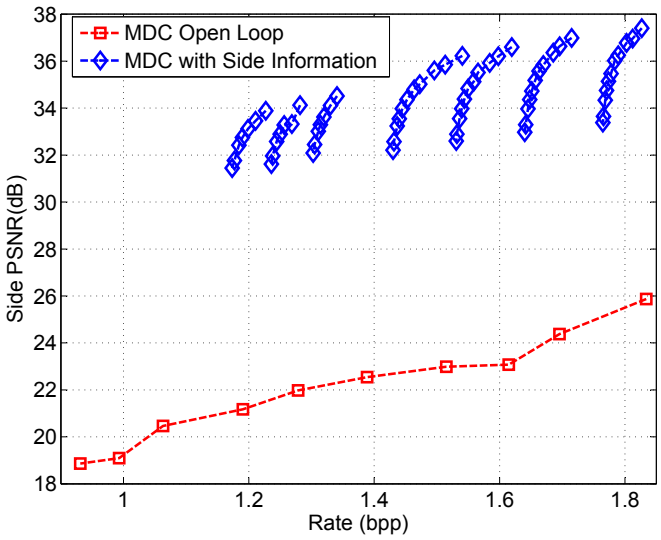

Figure 4.27: Side-distortion *Foreman* sequence, Intra Frame CIF, MDSQ with 3 diagonals

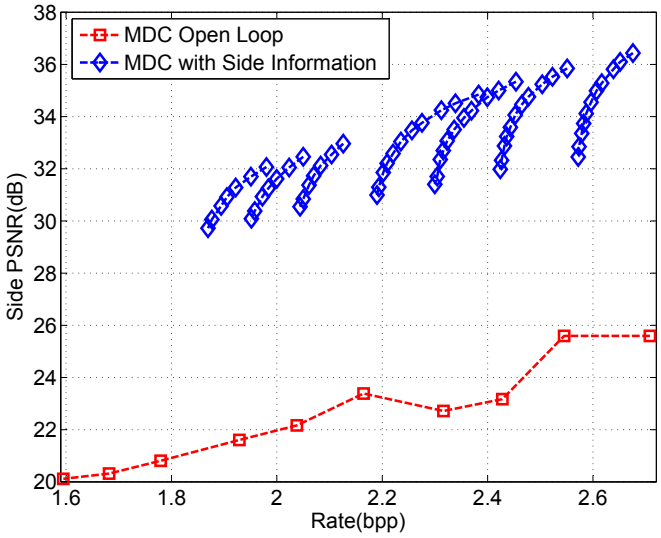

Figure 4.28: Side-distortion *Bus* sequence, Intra Frame CIF, MDSQ with 3 diagonals

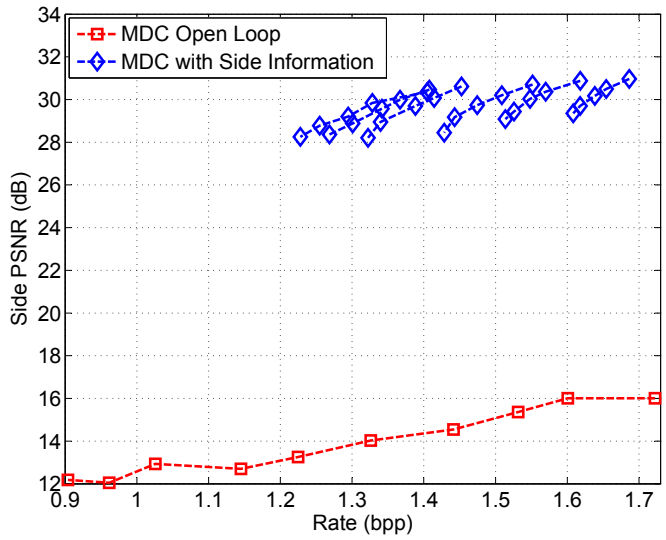

Figure 4.29: Side Distortion *Foreman* sequence,Intra Frame CIF, MDSQ with 5 diagonals

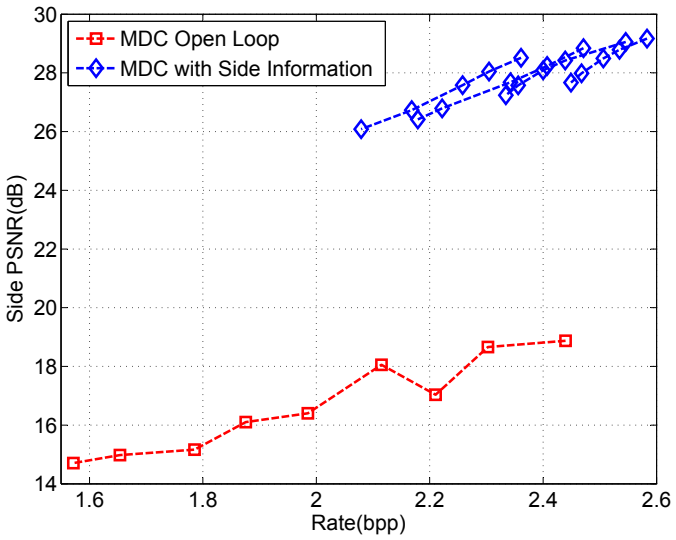

Figure 4.30: Side Distortion *Bus* sequence,Intra Frame CIF, MDSQ with 5 diagonals

4.5.1.2 Motion Compensation Predicted Slices

Performance evaluation of the proposed MDC architecture is done for Motion Compensation (MC) predicted slices. The main novelty of this proposal is the use drift compensation on only at anchor slices I and P that are used on prediction for other slices. This allows to reduce the amount of side-information and at the same time to control drift-distortion in the whole GOP. In this section an evaluation of the proposed scheme is done comparing MDSQ with 2, 3, and 5 diagonals. Evaluation is done using sequences *foreman* and *bus*, CIF format, 15Hz. Resulting rates are from coding with fixed quantization parameters QP_0 between 10 and 20, and $QP_i, i = 1, 2$ between 20 and 40 without using rate-control schemes. Proposed MDSQ for Advanced Video Coding scheme is compared with: MDSQ open-loop, MDSQ open-loop with I repetition and SDC where same rate is used.

MDSQ with 2 Diagonals Figs. 4.31 and 4.32 show the R-D performance for *foreman* and *bus* sequences, CIF format, 10Hz, with GOP type IPBBP and GOP size of 16 frames for MDSQ with 2 diagonals. In order to reduce the excess rate, only I and P frames use side information. The proposed MDC with side information have performance gains comparing with the other methods. For the same rate, Figs. show that open loop scheme have a poor side decoding quality. This is improved when Intra frames are used as SDC in both descriptions. Nevertheless, due to drift distortion that is generated, the overall quality is improved using the side informarion. Comparing proposed method with open loop MDC, the side PSNR gains are in the order of 10 dB. Comparing the proposed method with MDC open loop with SDC intra slices, the gains are in the order of 3-5 dB, depending of side information rate.

MDSQ with 3 Diagonals Figs. 4.33 and 4.34 show the R-D performance for *foreman* and *bus* sequences, CIF format, 10Hz, with GOP type IPBBP and GOP size of 16 frames with 3 diagonals. In order to reduce the excess rate, only I and P frames use side information. Obtained results are inline with the obtained with 2 diagonals. The proposed MDC with side information have performance gains comparing with the other methods. For the same rate, Figs. show that open loop scheme have a poor side decoding quality. This is improved when Intra frames are used as SDC in both descriptions. Nevertheless, due to drift distortion that is generated, the overall quality is improved using the side informarion. Comparing proposed method with open loop MDC, the side PSNR gains are

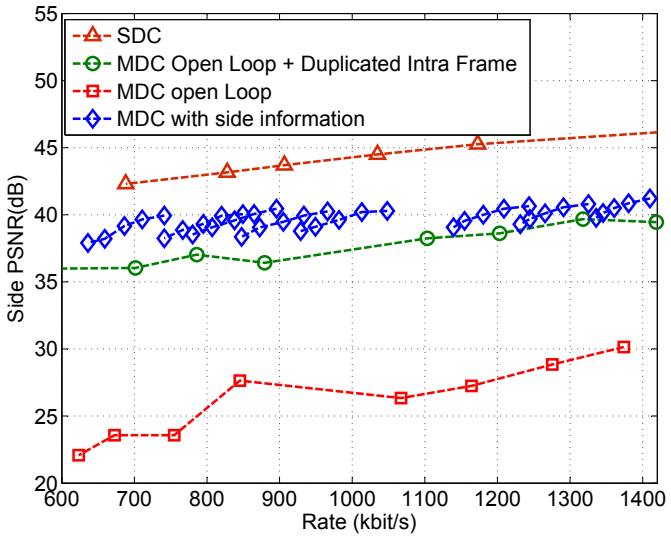

Figure 4.31: Side Distortion *Foreman* sequence, CIF 10HZ, MDSQ with 2 diagonals

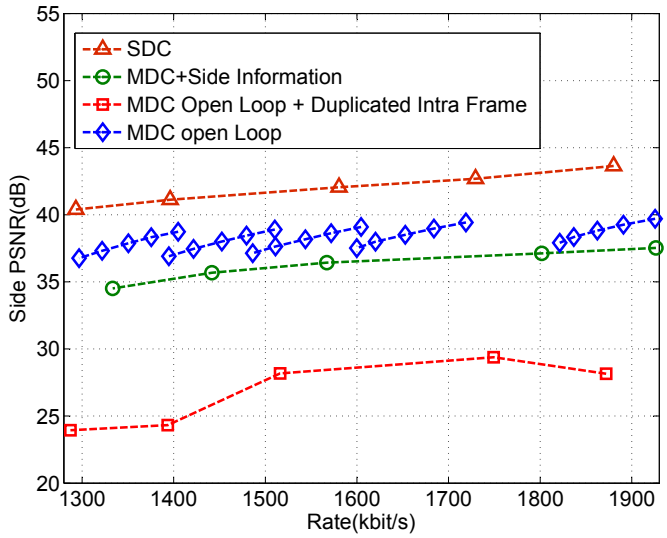

Figure 4.32: Side Distortion *bus* sequence, CIF 10HZ, MDSQ with 2 diagonals

in the order of 15-20 dB. Comparing the proposed method with MDC open loop with SDC intra slices, the gains are in the order of 3-8 dB, depending of side information rate. Note that Side PSNR are lower compared with MDSQ with 2 diagonals. Two main reasons are related to this fact. An higher MDSQ side distortion is obtained, that explain worse results in open loop schemes. On the other hand, the side information coding efficiency decrease because prediction mismatch error increase, resulting in a poor side-information R-D performance.

MDSQ with 5 Diagonals Figs. 4.35 and 4.36 show the R-D performance for *foreman* and *bus* sequences, CIF format, 10Hz, with GOP type IPBBP and GOP size of 16 frames with 5 diagonals. In order to reduce the excess rate, only I and P frames use side information. MDC with side information have performance gains comparing with the other methods. For the same rate, Figs. show that open loop scheme have a poor side decoding quality. In this case side distortion os around 15-20 dB which is not an acceptable quality. Comparing the proposed method with MDC open loop with SDC intra slices, the gains are in the order of 3-10 dB, depending of side information rate. Also increasing the number of diagonals, the average level of side PSNR decreases for the reasons earlier explained.

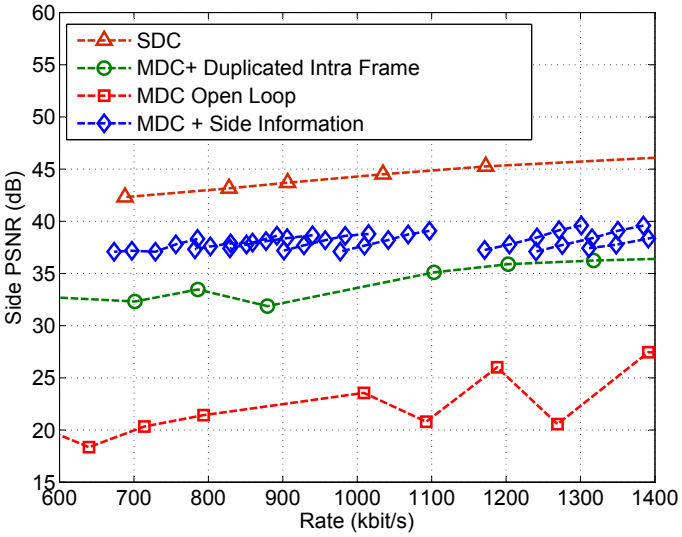

Figure 4.33: Side Distortion *Foreman* sequence, CIF 10HZ, MDSQ with 3 diagonals

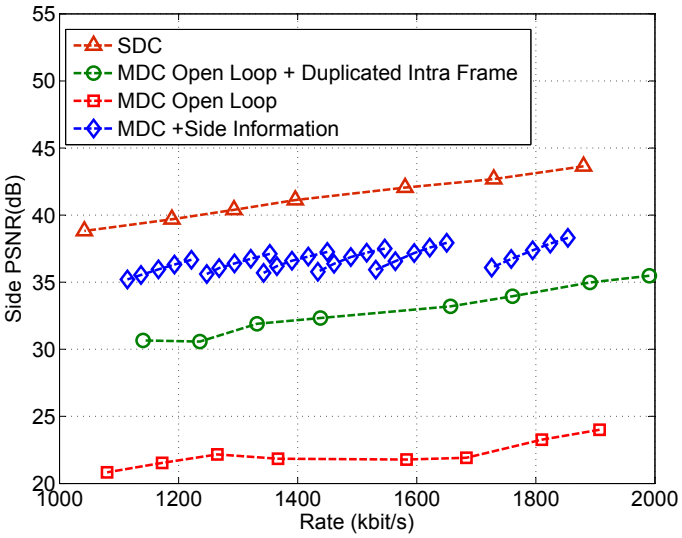

Figure 4.34: Side Distortion *Bus* sequence, CIF 10HZ, MDSQ with 3 diagonals

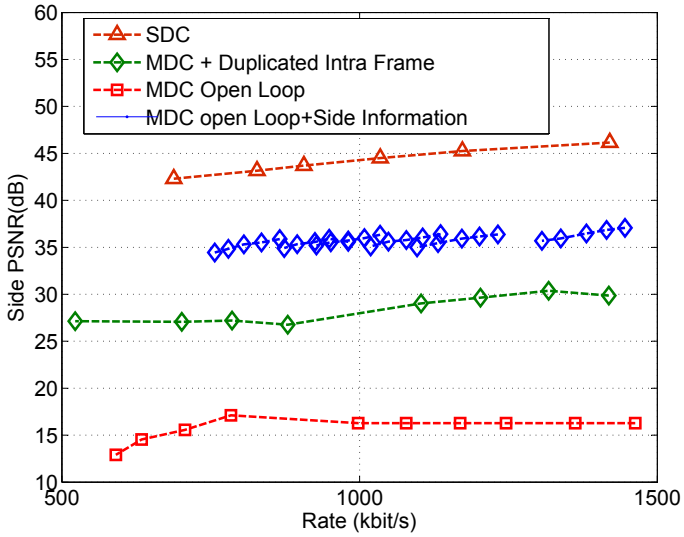

Figure 4.35: Side Distortion *Foreman* sequence, CIF 15HZ, MDSQ with 5 diagonals

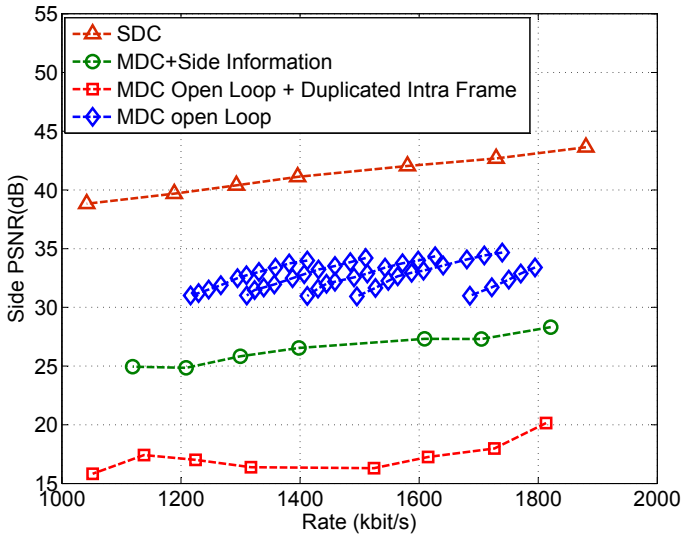

Figure 4.36: Side Distortion *Bus* sequence, CIF 15HZ, MDSQ with 5 diagonals

4.5.2 Unbalanced MDSQ

Unbalanced MDSQ intends to produce description with a certain amount of controlled balancing between each description rates. As it was shown in last section, balanced rates are achieved changing parameters of the index assignment matrix, namely, Z parameter that controls the central SDC indices rounding zero. It is important to note that, as is shown in Chapter 2, that in case of unbalanced descriptions, at least one of the side distortions is very poor, which in practice can makes such scheme unfeasible. Applying this scheme in highly predictive coding schemes such as Advanced Video Coding, its usage is critic in other to obtain an acceptable side-distortion performance.

Using the proposed MDC scheme, now for unbalanced descriptions a new problem is posed, that is to find the amount of the side-information that needs to be added in order to have side-distortions with acceptable quality. The problem is distinct from balanced case because now, the degree of freedom to find the best combination between side-information and central-information increases. This means that, to have unbalanced rates in each descriptions, this can be done by balancing MDSQ results and unbalancing the side-information, or unbalancing the MDSQ rates using balancing rates for side-information, thus maintaining the balancing rates among descriptions. The main goal is to assure that none of the descriptions has poor decoding quality.

Therefore, the performance evaluation is done through understanding the effect of the side-information in side-distortion of each description, given a ratio between the side-information and central-information. on the other hand in important to understand how to side-information affects the global balancing of each description and also the global redundancy that is introduced.

Figs.4.37, 4.38,4.39, 4.40 and 4.44, 4.42, 4.43 ,4.44 shows the side-distortion resulting from using fixed $Q_i, i = 1, 2$ in side encoders between 20 and 30 applied only to anchor slices. Redundant information is represented by the side-information/central information ratio, which is the relation between the rates produced by the side-encoder and for the MDSQ encoder for each description.

This results show that for Description 1 has a relatively high side-information percentage relatively to central information. Nevertheless, Description 1 has relatively less rate, explanting these results. On the other hand for description 2, the percentage of side information that is introduced is very low which means that the gains that are obtained.

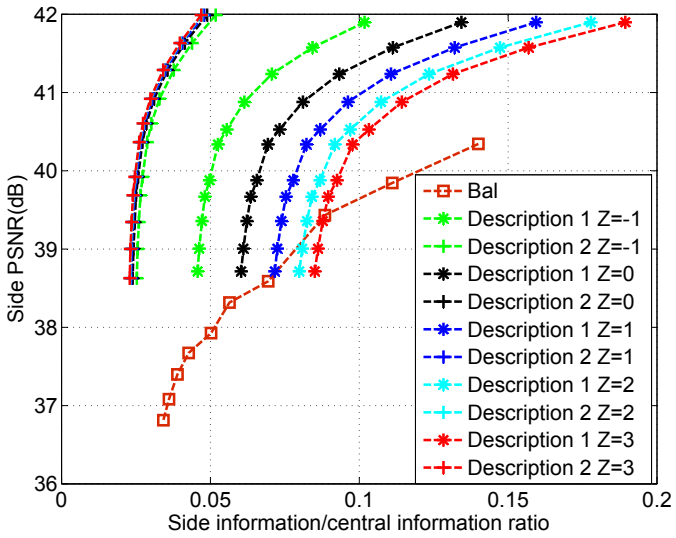

Figure 4.37: Side Distortion *foreman* sequence, IPBB Frame

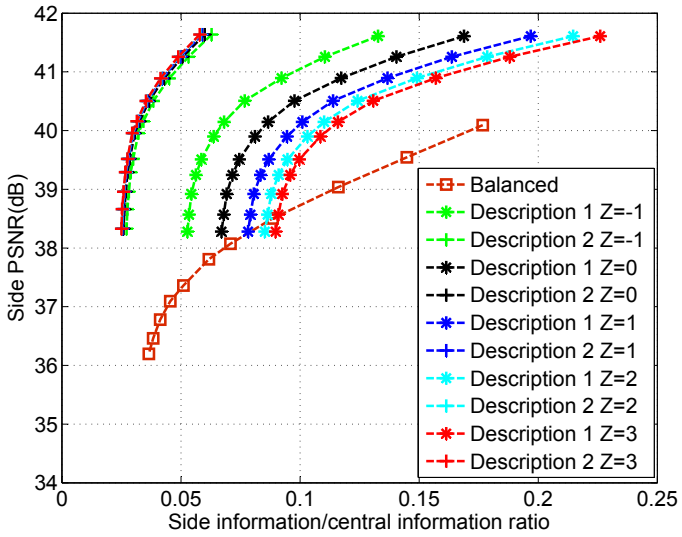

Figure 4.38: Side Distortion *foreman* sequence, IPBB Frame

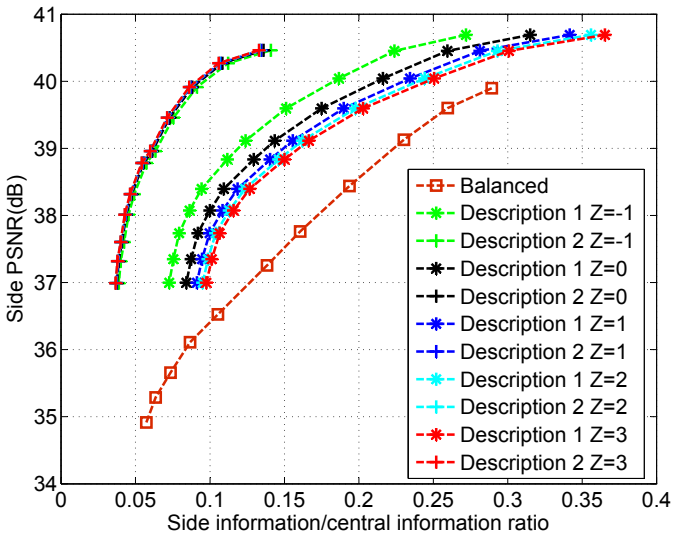

Figure 4.39: Side Distortion *foreman* sequence, IPBB Frame

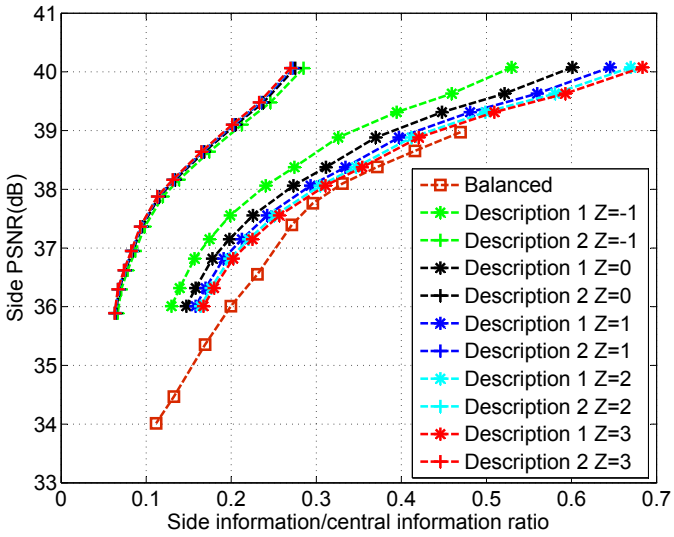

Figure 4.40: Side Distortion *foreman* sequence, IPBB Frame

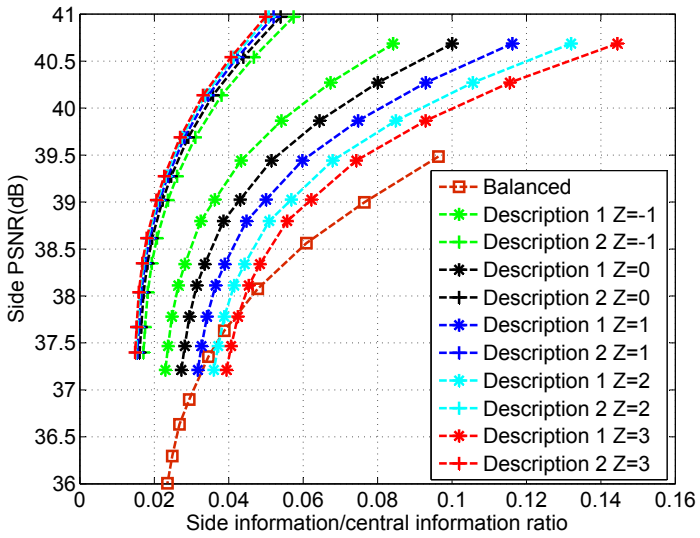

Figure 4.41: Side Distortion *Bus* sequence, IPBB Frame

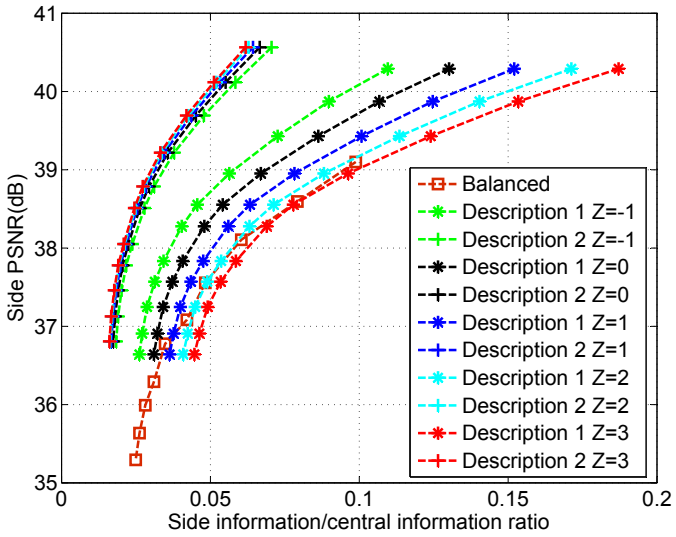

Figure 4.42: Side Distortion *Bus* sequence, IPBB Frame

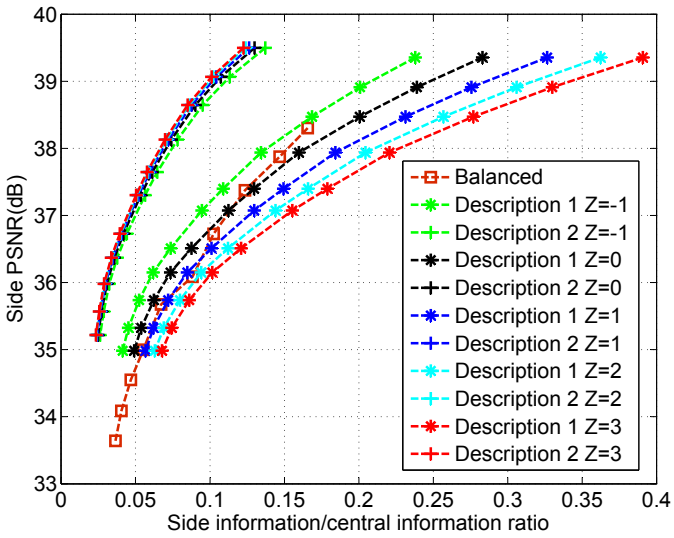

Figure 4.43: Side Distortion *Bus* sequence, IPBB Frame

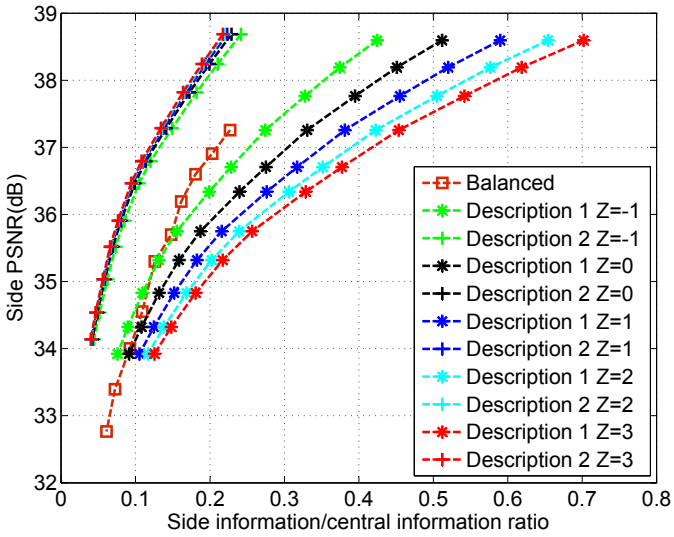

Figure 4.44: Side Distortion *Bus* sequence, IPBB Frame

4.6 Conclusions

This chapter proposes a new MDSQ architecture for H.264/AVC with drift-control for balanced and unbalanced descriptions. A study of the main issues related to MDSQ implementation is presented, where drift distortion, redundancy, and overall rate-distortion performance is evaluated. The proposed architecture allows to improve the open-loop MDSQ architecture in terms of side-distortion performance with a reduced amount of additional information. The novel unbalanced MDSQ, improves the overall performance comparing with open-loop architecture and also with the balanced close-loop MDSQ scheme. Is shown that proposed MDSQ unbalanced scheme increases the coding efficiency of the side information, which means that proposed method even using different rates at each description, the overall quality quality can be improved by correctly using the side information allocation and at the same time to mantain the unbalancing rate.

Chapter 5

Multiple Description Video Splitting of Coded Streams

Contents

5.1 MDVS Scenario **109**

5.2 Classic MDVS **111**

5.2.1 Drift Analysis 111

5.3 MDVS with drift compensation **113**

5.3.1 MDVS architecture 113

5.3.2 Simplified MDVS 115

5.4 MDVS drift performance **118**

5.4.1 Intra predicted frames 118

5.4.2 MC predicted frames 118

5.4.3 Generic regular GOP 119

5.4.4 The overall effect of side information 123

5.5 MDVS streaming with path diversity **124**

5.6 MDVS vs MDC: comparative discussion **128**

5.7 Conclusion **129**

This chapter extends the current concept of multiple description coding (MDC) to the compressed domain, by proposing efficient splitting of standard single description coded

(SDC) video into a multi-stream representation. A novel multiple description video splitting (MDVS) scheme is proposed to operate at network edges, for increased robustness in path diversity video streaming across heterogeneous communications chains. It is shown that poor performance of existing methods is mainly due to distortion accumulation, i.e., drift, when decoding is carried out with missing descriptions. The proposed scheme is able to effectively control drift distortion in both intra and inter predictive coding, even when only one description reaches the decoder. This is achieved by generating a controlled amount of relevant side information to compensate for drift accumulation, whenever any description is lost in its path. The simulation results show that any individual description can be decoded on its own without producing drift, achieving significant quality improvement at reduced redundancy cost.The overall performance evaluation, carried out by simulating video streaming over lossy networks with path diversity, also demonstrates that MDVS enables higher quality video in such heterogeneous networking environments, for a wide range of packet loss rates.

This chapter proposes a novel MDVS scheme based on multiple description scalar quantisation (MDSQ), using side information to control drift in both spatial and temporal prediction. The side information is generated from the original stream and its rate is controlled with an independent quantisation parameter which also controls redundancy. Then, a simplified architecture is devised to reduce the overall complexity in regard to the number of processing functions and memory requirements. No additional information is needed from the original SDC encoder in order to generate such side information at any MDVS-enabled network node. As highlighted in section 5.2.1, another novel aspect of this paper is to provide evidence about the catastrophic effect of drift, which drastically reduce performance if the splitting architecture does not adequately compensate for its accumulation.

In this chapter, the main novel aspects of the proposed MDVS scheme comprise: (i) a two-loop MDVS architecture with drift control in both intra and inter predictive coded slices; (ii) an equivalent single-loop architecture; (iii) a method to generate side information from SDC video; (iv) the capability of controlling the amount of side information according to the expected decoder drift; (v) an overall performance similar to MDC using uncompressed video.

5.1 MDVS Scenario

Fig. 5.1 shows a possible video streaming scenario where MDVS might be useful. A single video stream (i.e., SDC stream) is distributed from the streaming server to diverse user terminals over heterogeneous networks, some of them having disjoint separate paths from an intermediate node to the user terminal. Since the server storage capacity and streaming bandwidth required for SDC video is less than that of equivalent MDC (i.e., same quality), due to the inherent redundancy of MDC, SDC is a more efficient coded format for storage and distribution over single path networks. The advantage of MDVS is to introduce a further level of flexibility in SDC video streaming, in order to benefit from transmission over multiple paths where these are available along the delivery chain. Therefore, MDVS can be seen as a novel adaptation functionality of edge nodes in the heterogeneous video streaming environments of the future media internet. In comparison with the existing MD schemes previously cited, this chapter addresses a different MD problem, which consists in splitting compressed video streams rather than producing MD from uncompressed video signals. Moreover, since MDVS suffers from the same intrinsic problem of drift as MDC, i.e., accumulation of decoding distortion when any description is lost in the network, a novel aspect of the proposed scheme in this chapter is its capability for limiting such type of distortion, by generating a controlled amount of side information, specifically for this purpose.

Secondly, by using a reference picture selection (RPS) method based on automatic repeat request (ARQ). In [S. Lin 2001], the reference frames for motion compensated prediction are selected according to feedback information received from the transmission paths. A similar principle is used in [Liao 2011], where routing messages are used in order to estimate packet loss error rate, by dynamically selecting the best reference frame in order to alleviate error propagation. If the drift compensation process is considered as a bit stream switching problem (i.e., switching from decoding with two descriptions to decoding with a single description), then periodic switching frames (e.g., H.264/AVC SI/SP slices) might be used to enhance MDC error resilience [Wang 2005a]. However, this mechanism cannot be used directly with compressed streams because SI/SP frames would need to be dynamically computed by full MDC encoders at MDVS network nodes.

Since the multi-loop drift compensation methods, referred to above, lead to high complexity implementations, a different approach is followed in the proposed MDVS scheme, which is based on a single-loop MDSQ. This comprises a novel MDVS architecture with

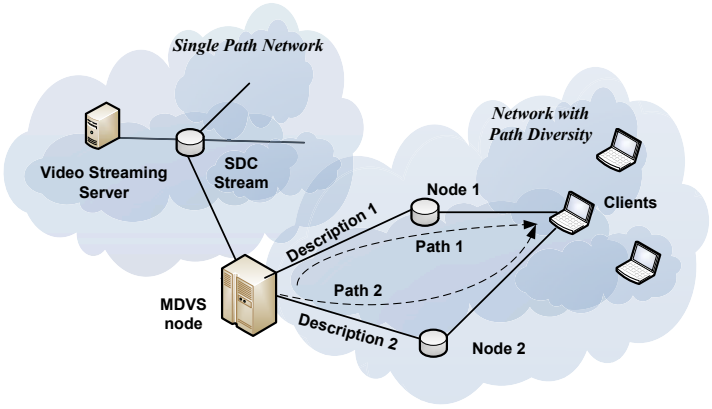

Figure 5.1: MDVS application scenario.

low processing complexity, achieved by reusing most coding parameters of the incoming SDC video streams.

Two rather different approaches, based either on channel coding or source coding, have been followed in video streaming applications using MDVS. In [Puri 2001] the channel coding approach is proposed for an end-to-end video communication system where MDVS, based on forward error correcting codes (FEC), is integrated in a congestion control framework for video streaming over the internet. Other examples of previous work in MDVS based on channel coding are reported in [Gan 2006] and [Essaili 2007]. MDVS based on the source coding approach was also addressed in the past, but mainly focussed on its application in some networking scenarios without taking in account the rate-distortion efficiency of actual MDVS processing architectures. In particular, drift free MDVS architectures cannot be found in the available literature. Related work can be found in [Kim 2003], where an MDVS scheme is proposed based on redundancy rate-distortion optimisation for splitting DCT coefficients of the incoming bitstream. In [Kim 2006] another MDVS scheme is proposed, based on replication and interleaving of DCT coefficients among all descriptions. However, these are open-loop schemes with no drift compensation, also resulting in higher levels of rate redundancy. As highlighted in section 5.2.1, another novel aspect of this proposed architecture is to provide evidence about the catastrophic effect of drift, which drastically reduce performance if the splitting architecture does not adequately compensate for its accumulation. A rather different approach of video adaptation using side information based on distributed source coding techniques is described in [Shanableh 2008]. Although this may not be classified exactly as an MDC scheme, it is worth to mention that such adaptation scheme generates redundant side information

taking drift into account. In [Noh 2012], a robust mobile video streaming in P2P system using a distributed transcoding between fixed nodes with more resources, namely, base stations and personal computers where mobile nodes are only used to receive packets.

5.2 Classic MDVS

MDVS can be regarded as a data partitioning scheme, capable of generating two descriptions from an SDC video stream. Since current coded video formats convey a great deal of the source information in transform coefficients, MDSQ is good candidate to design low complexity MDVS systems. Fig. 5.2 shows a classic MD video splitting scheme where each transform coefficient is represented by two different values, which result in dividing an SDC stream into two descriptions. For instance, this MDVS method was used in [Chen 2007b] where the coding information embedded in the original SDC stream, such as slice maps, prediction modes and motion vectors are duplicated in the two resultant descriptions. In the scheme shown in Fig. 5.2, an index assignment function is used for mapping each quantisation index of the original transform coefficients i_0 (i.e., central indices) into a pair of side indices (i_1, i_2) which are then entropy encoded.

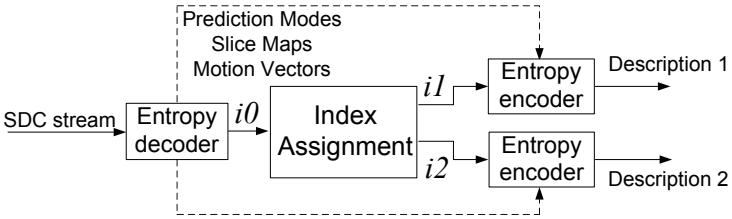

Figure 5.2: Classic MDVS scheme.

5.2.1 Drift Analysis

If the classic MDVS scheme of Fig. 5.2 is used for adapting coded video streaming to networks with path diversity, then drift is introduced at the decoder whenever any description is lost. The drift distortion component can be determined from the relevant signals involved in MD splitting and decoding. When two descriptions are received, both side indices (i_1, i_2) are decoded and merged into the corresponding central index i_0. In

this case, for each block n, the reconstructed central pixel values $F_{n,0}$ are given by,

$$F_{n,0} = r_{n,0} + P_{n,0} \tag{5.1}$$

where $r_{n,0}$ is the decoded residue and $P_{n,0}$ its associated prediction either from intra prediction or motion compensation, formed from decoding both descriptions. If only one description j (either $j = 1$ or $j = 2$) is decoded, then the reconstructed pixel values $F_{n,j}$ are given by,

$$F_{n,j} = r_{n,j} + P_{n,j} \tag{5.2}$$

where $r_{n,j}$ is the decoded residue and $P_{n,j}$ its prediction formed from decoding description j only. Since $r_{n,j}$ results from inverse index assignment using with only one description as input, the difference between the original SDC residue and that decoded from only one description produces a reconstruction error ε, i.e.,

$$r_{n,j} = r_{n,0} + \varepsilon \tag{5.3}$$

Substituting (5.3) in (5.2),

$$F_{n,j} = r_{n,0} + \varepsilon + P_{n,j} \tag{5.4}$$

and then using (5.1) in (5.4), $F_{n,j}$ becomes,

$$\begin{aligned} F_{n,j} &= F_{n,0} + P_{n,j} - P_{n,0} + \varepsilon \\ &= F_{n,0} + d_P + \varepsilon \end{aligned} \tag{5.5}$$

where

$$d_P = P_{n,j} - P_{n,0} \tag{5.6}$$

is the drift component due to mismatch between the SDC predictions used in the original encoder and those reconstructed at the final decoder from only one description. Note that the above analysis is valid for both the spatial and temporal drift components, though these can be identified as separate contributors to the overall drift distortion.

The actual impact of the drift component given by equation 5.6 in the objective video quality was experimentally evaluated for the MDVS scheme of Fig. 5.2. The results for intra predicted and motion compensated (MC) frames, are shown in Figs. **??** and **??**, respectively.

5.3 MDVS with drift compensation

A novel MDVS architecture and an equivalent simplified version are proposed to overcome the problem of spatial and temporal drift. In comparison with the classic scheme of Fig. 5.2, the proposed splitting architecture generates an additional amount of side information for the specific purpose of preventing drift when only one description is decoded. As explained in the next subsections, the side information is generated for each description by further encoding the difference between the original SDC signal and the one decoded from the description itself.

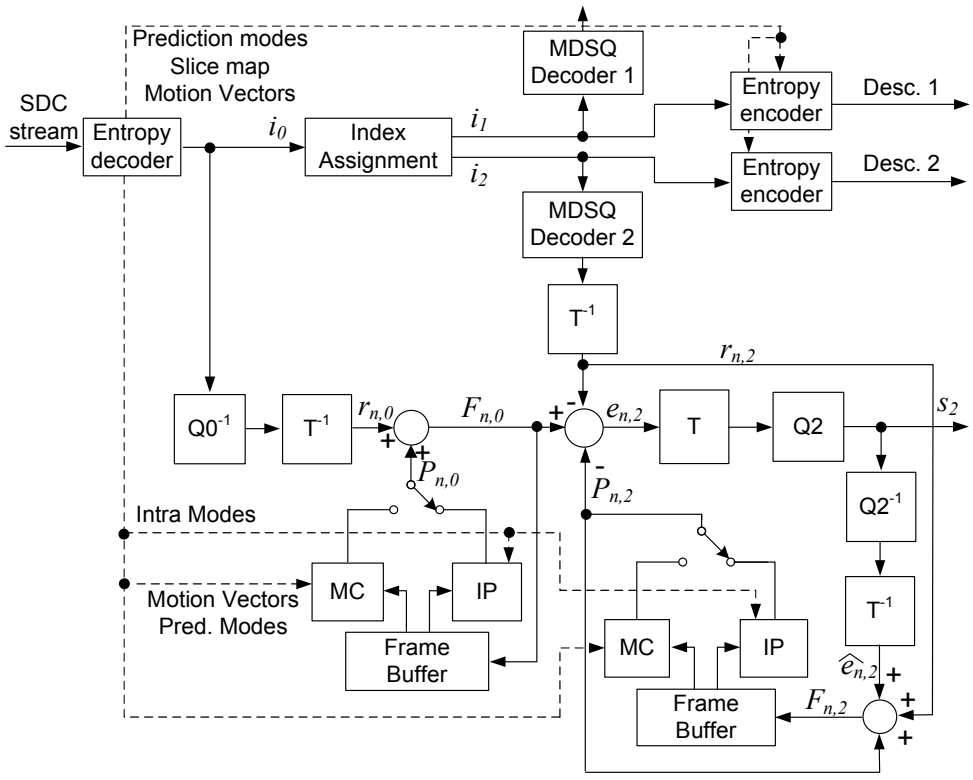

Figure 5.3: Two-loop MDVS architecture.

5.3.1 MDVS architecture

The proposed MDVS architecture using side information for drift compensation is shown in Fig. 5.3. Such an architecture is used for drift free MDVS of both intra predicted and motion compensated predicted frames. However, the B frames that are not used as references may be processed with the classic MDVS scheme of Fig. 5.2 because they do not

contribute for drift accumulation. Similarly to the classic MDVS, the headers, prediction modes and motion vectors are duplicated in both descriptions. The side information is encoded by also using the same prediction modes and motion vectors.

In the MDVS architecture of Fig. 5.3 the side information s_j for each description j, with $j = 1, 2$, is defined as,

$$s_j = Q_j\{T(e_{n,j})\} \tag{5.7}$$

where

$$e_{n,j} = F_{n,0} - r_{n,j} - P_{n,j}. \tag{5.8}$$

and Q_j is the quantisation with side quantiser Q_{Pj} that determines the amount of side information sent to the decoder, T is the transform operation, $F_{n,0}$ is the current frame and $P_{n,j}$ is the motion compensated prediction from side encoder j. The residue $r_{n,j}$, available at the decoder when only one description is correctly received, is given by,

$$r_{n,j} = T^{-1}(Q_0^{-1}\{A_j^{-1}(i_j)\}), \tag{5.9}$$

where $A_j^{-1}(.)$ denotes the inverse index assignment operation when only description j is available and Q_0^{-1} is the inverse quantisation using the SDC quantisation parameter Q_{P0}. The central index i_0' is assigned to the value located in the main diagonal of the index assignment matrix which corresponds to either the row or column of side index i_j.

At the decoder, if both descriptions are available, then the inverse index assignment operation is responsible to restore the exact value of the original SDC index, which is then inverse quantised and inverse transformed. However, if only one description is decoded, then frame $F_{n,j}$ is reconstructed without drift as follows,

$$F_{n,j} = P_{n,j} + r_{n,j} + \widehat{e}_{n,j}. \tag{5.10}$$

with

$$\widehat{e}_{n,j} = T^{-1}(Q_j^{-1}\{s_j\}). \tag{5.11}$$

By comparing (5.2) with (5.10), one concludes that $e_{n,j}$ is the drift compensation signal that is encoded as side information and sent to the decoder. Note that in the classic MDVS scheme of Fig. 5.2, only signal $r_{n,j}$ is sent for decoding which is not enough to prevent drift.

5.3.2 Simplified MDVS

The MDVS architecture of Fig. 5.3 can be simplified by assuming that prediction is a linear function (this is valid except for rounding and truncation arithmetic). Using equations (5.1) and (5.8), the following relation can be derived,

$$e_{n,j} \quad = \quad P_{n,0} - P_{n,j} + r_{n,0} - r_{n,j} \tag{5.12}$$

Considering equation (5.12) and the previous assumption of linearity, the simplified architecture of Fig. 5.4 can be derived. In this architecture there is only one loop for intra prediction (i.e., IP) and one motion compensation (i.e., MC) loop to accumulate the differential signal used for drift compensation.

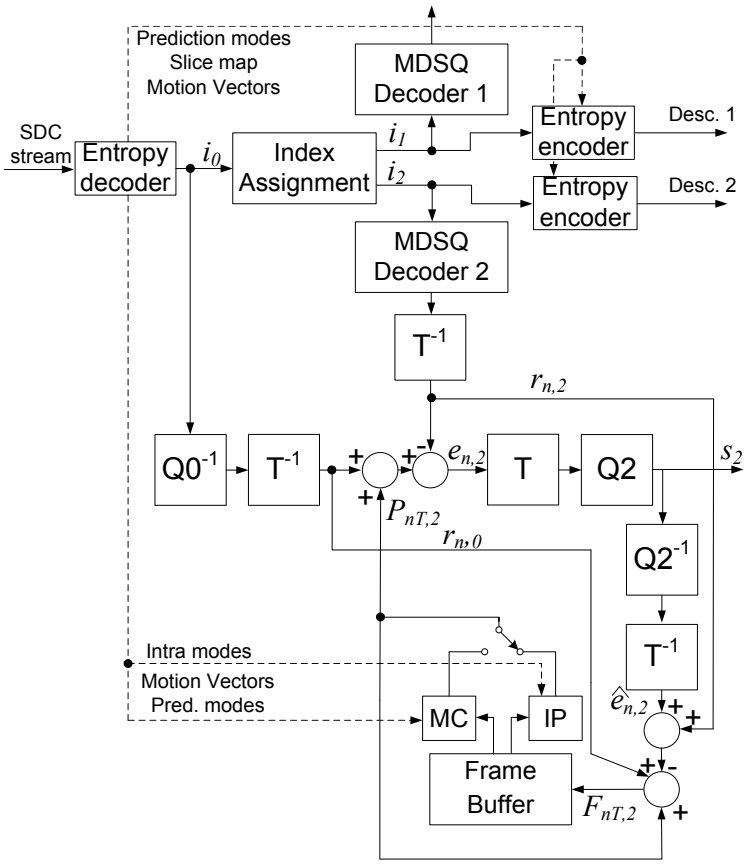

Figure 5.4: Single-loop MDVS architecture.

The equivalence between architectures of Figs. 5.3 and 5.4 can be demonstrated as follows. In the architecture of Fig. 5.4, the side information $e_{n,j}$ is given by,

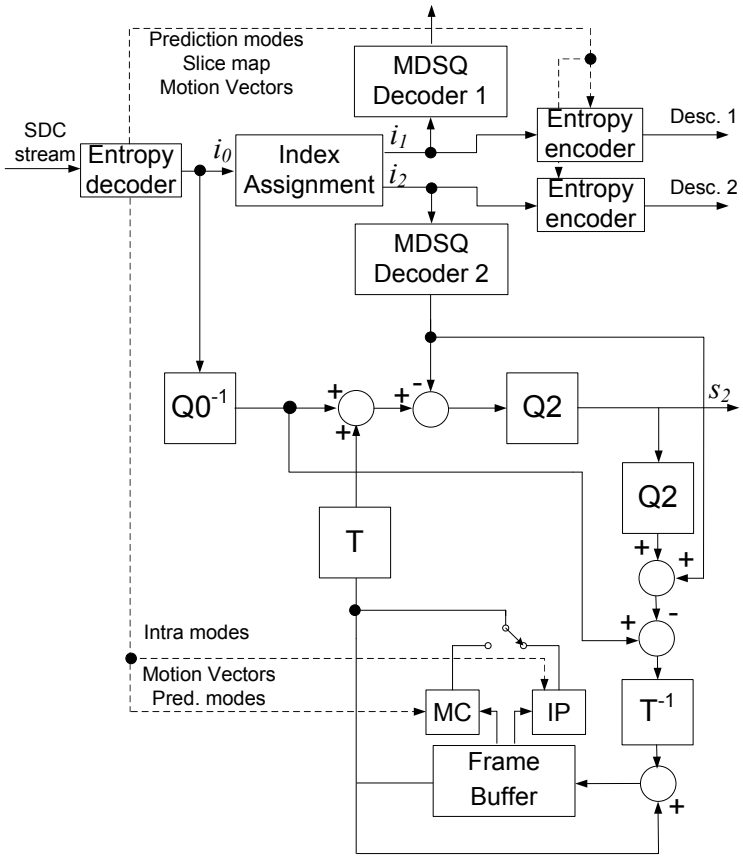

Figure 5.5: Equivalent single-loop MDVS architecture.

$$e_{n,j} = r_{n,0} - r_{n,j} + P_{nT,j} \tag{5.13}$$

and the signal accumulated in the local prediction loop for drift compensation is defined as follows,

$$F_{nT,j} = r_{n,0} - (r_{n,j} + \widehat{e}_{n,j}) + P_{nT,j} \tag{5.14}$$

then, by using equations (5.1), (5.10), (5.13) and (5.14) the resulting expression is,

$$\begin{aligned} F_{n_T,j} &= F_{n,0} - P_{n,0} + \\ &\quad - (r_{n,j} + F_{n,j} - P_{n,j} - r_{n,j}) + P_{n_T,j} \\ &= F_{n,0} - F_{n,i} - P_{n,0} + P_{n,j} + P_{nT,j}. \end{aligned} \tag{5.15}$$

Considering the linearity of prediction $P_{nT,j} = P_{n,0} - P_{n,j}$ then

$$F_{nT,j} = F_{n,0} - F_{n,j}. \tag{5.16}$$

Equation (5.16) represents the difference between frames reconstructed from the two prediction loops of Fig. 5.3. In the simplified architecture of Fig. 5.4 such difference is accumulated in only one loop and the result is used in the same manner as in Fig. 5.3, which demonstrates that both architectures are equivalent.

Moreover, since transform and quantisation can be implemented as independent operations, the architecture of Fig. 5.4 can be further simplified to that of Fig. 5.5. For each description, this scheme only uses one frame buffer and two transforms while that of Fig. 5.3 needs two frame buffers and four transforms. Note that in H.264/AVC, the scheme of Fig. 5.5 needs to use scaling coefficients in the quantisation and inverse quantisation functions in order to make them independent from the transform. This can be easily done as described in [Lefol 2006].

Although some operations may involve non-linear arithmetic, such as clipping functions in transform/quantisation, rounding in sub-pel MC interpolation and deblocking filtering, the actual effect on the drift performance of the simplified MDVS is mostly negligible. However, it might be slightly more significant in high motion sequences.

The proposed MDVS architectures were implemented on the reference software of H.264/AVC online available. Each coded description produced by MDVS is standard-compliant and the corresponding coded data is encapsulated into Video Coding Layer (VCL) Network Adaptation Layer (NAL) units. To include the side information in the standard syntax, a new type of VCL NAL unit must be defined for such coded data. This can be done by extending the existing NAL types using different approaches. For instance in H.264/SVC [Ye-Kui 2007], new NAL unit types were defined to accommodate several layers and associated information and in [Lamy-Bergot 2010], a new type of NAL unit is proposed for embedding redundant information inside standard video streams.

5.4 MDVS drift performance

In this section, the performance results were obtained to prove the advantages of using the proposed MDVS evaluating the impact of drift accumulation when one description is totally lost, either in intra or inter predicted coded frames. The drift performance of MDVS was evaluated in two different aspects: (i) distortion accumulation in decoded video when only one of the two descriptions reaches the decoding terminal and (ii) the extra redundancy of the side information produced by the proposed MDVS to compensate for drift, also in one of the two descriptions. The reference used for comparison is one description obtained from the classic MDVS scheme of Fig. 5.2. The spatial and temporal drift performance were evaluated by using coded streams with intra predicted and MC predicted frames (P and B), respectively. In order to obtain a comparable evaluation, all streams were encoded at the same rate.

The original headers, prediction modes, slice maps and motion vectors are duplicated into both descriptions. The side information is encoded using the same coding modes as the corresponding descriptions. Note that coding modes and motion vectors are not included in the side information because they are available from the respective coded description.

5.4.1 Intra predicted frames

The benefit of drift compensation in intra predictive MDVS is shown in Fig. 5.6 where the PSNR of each macroblock of one frame, in one of the two descriptions (*bus* sequence), is shown for classic MDVS and for two-loop MDVS (i.e., Fig. 5.3) at the same bit rate. The same rate is ensured by an average central quantiser $Q_{P0} = 15$ for Classic MDVS and $Q_{P0} = 18$, $Q_{P1,2} = 33$ for proposed MDVS. Fig. 5.6 clearly shows that the proposed MDVS produces much higher and smoother PSNR along the I frame than in the case where drift compensation is not done. In the case of no drift compensation, i.e., classic MDVS, the lowest PSNR is below 20dB, which is definitely not acceptable.

5.4.2 MC predicted frames

A different experiment was carried out to evaluate the performance of temporal drift compensation. A GOP structure with high number of predicted frames was used to

provide a worst case scenario in regard to temporal drift, i.e., a sequence of P frames using only one reference, i.e., IPPP... The GOP size was set to 20 frames. The entire loss of one description is simulated in the path diversity network, for the initial I frame and also for all subsequent P frames (i.e., only one description is decoded). In the error-free descriptions of all streams the initial I frame is sent with side information in order to not affect the quality of subsequent P frames. Both the two-loop and the single-loop MDVS architectures were used in the experiment in order to evaluate the effect of the non-linearity of motion compensation in the drift accumulation over a significant number of temporally predicted frames.

Fig. 5.7 shows the PSNR for *coastguard* and *foreman* sequences at the same bit rate, using an average central quantiser $Q_{P0} = 16$ for Classic MDVS and $Q_{P0} = 18$, $Q_{P1,2} = 30$ for proposed MDVS. It is quite evident that the proposed MDVS architectures are drift-free, while temporal drift accumulation is responsible for severe quality degradation in classic MDVS. At the end of the GOP, the PSNR obtained by using the proposed MDVS is about $6dB$ higher than classic MDVS. For classic MDVS, two-loop and single-loop, the average PSNR of *coastguard* is, respectively, $28.96dB$, $32.25dB$(+3.46) and $32.42dB$ (+3.29) and that of *foreman* is $31.41dB$, $34.34dB$ (+2.93) and $34.43dB$ (+3.02). These results also show that non-linearity of motion compensation is negligible, since the PSNR obtained from the single-loop MDVS architecture are quite similar to those obtained from two-loop MDVS. Therefore, full decoding of the incoming stream is not necessary to achieve drift free MDVS and these results validate the proposed single-loop architecture.

5.4.3 Generic regular GOP

The overall performance of the proposed MDVS using generic IBBP GOP structures was also evaluated (GOP size of 20 frames) using the same three sequences, i.e., *coastguard*, *foreman* and *bus*. In the proposed MDVS, drift compensation was used for both I and P frames but not for B frames because these are not used as references for prediction. Performance evaluation was carried out by splitting SDC streams of each sequence into two descriptions and then simulating that only one description reaches the decoder for all frames in the GOP. In all streams, the initial I frame is always sent with side information in order to not influence the quality of subsequent predicted frames.

Figs. 5.8, 5.9 and 5.10 shows the PSNR obtained from the proposed MDVS architectures

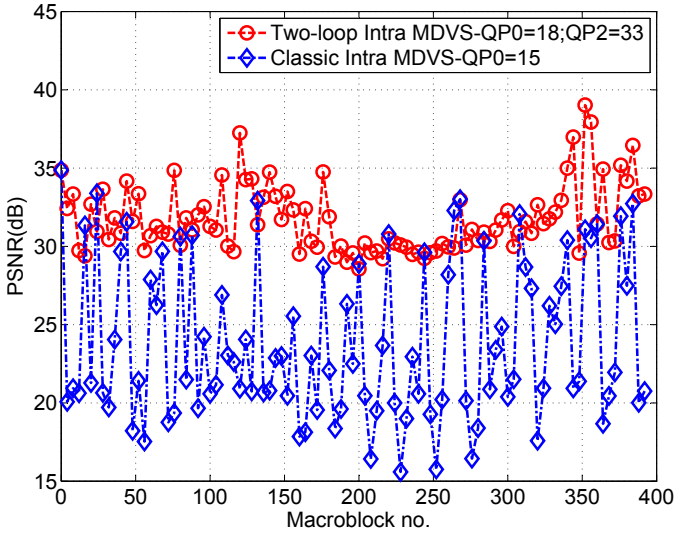

Figure 5.6: PSNR of Intra Frame macroblocks.

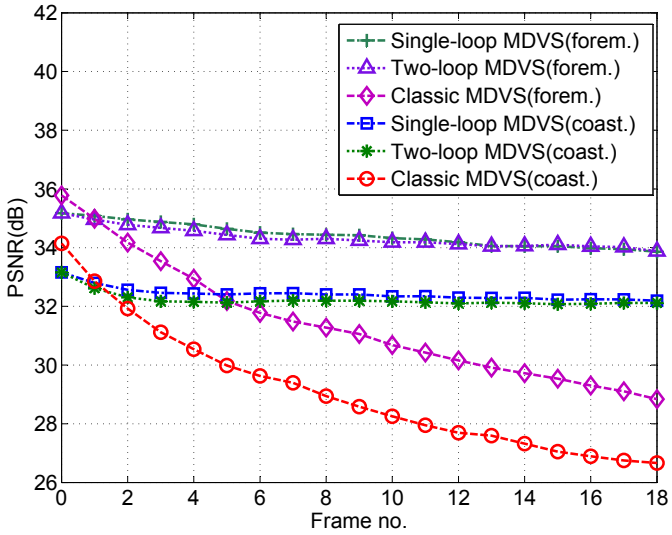

Figure 5.7: PNSR for MC predicted frames (IPPP...) for *coastguard* and *foreman*.

and Classic MDVS. The same bit rate was obtained for all streams using an average central quantiser $Q_{P0} = 17$ for classic MDVS and $Q_{P0} = 18$ for the proposed MDVS, which also used $Q_{P1,2} = 30$ for *coastguard* and *foreman* and using an average central quantiser $Q_{P0} = 17$ for classic MDVS and $Q_{P0} = 15$ and $Q_{P1,2} = 28$ for the proposed MDVS for *bus* sequence. The results in Figs. 5.8, 5.9 and 5.10 clearly confirm the effectiveness of the proposed architecture to eliminate the drift and consequently to achieve significant quality improvement in MD adaptation of coded video streams. The proposed architecture compensates for the drift in P frames which results in a significant overall quality improvement. Comparing PSNR of both the two-loop and single-loop MDVS architectures, these are very similar, which further validates the effectiveness of the simplified single-loop.

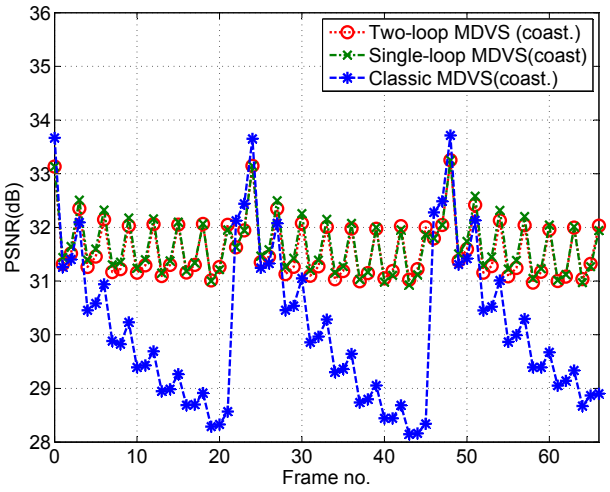

Figure 5.8: PNSR for generic regular GOP (IPBBP...) for *coastguard*.

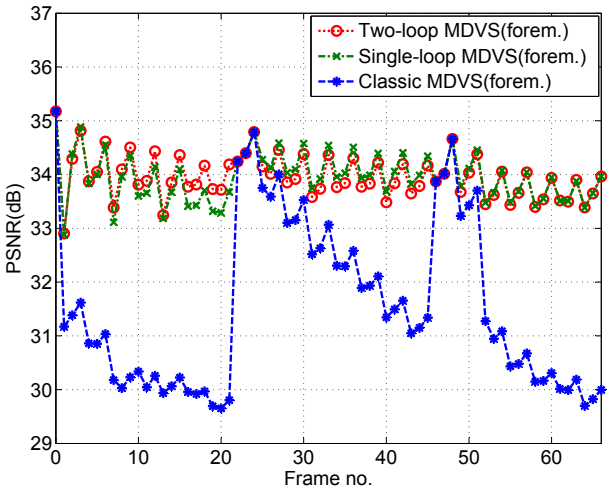

Figure 5.9: PNSR for generic regular GOP (IPBBP...) for *foreman*.

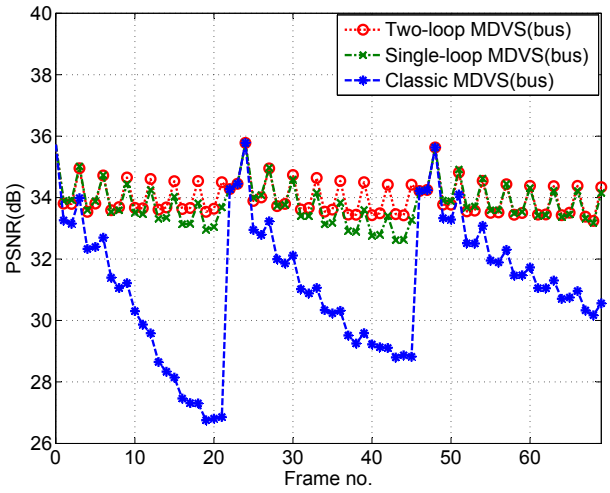

Figure 5.10: PNSR for generic regular GOP (IPBBP...) for *bus*.

Table 5.1: PSNR vs Side Information Redundancy (one description).

Seq.	QP_0	$QP_{1,2}$	Classic MDVS PSNR(dB)	Two-loop MDVS PSNR(dB)	Single-loop MDVS PSNR(dB)	Red. (%)
foreman	10	20	36.9	40.1(+3.2)	40.0(+3.1)	8.1
		25	36.2	38.0(+1.8)	38.0(+1.8)	2.9
		28	35.8	36.9(+1.1)	36.9(+1.1)	2.0
		30	35.5	36.3(+0.8)	36.3(+0.8)	1.8
	15	28	33.4	35.5(+2.1)	35.4(+2.0)	4.6
		30	33.1	34.6(+1.5)	34.6(+1.5)	3.4
		33	32.6	33.6(+1.0)	33.7(+1.1)	2.8
		35	32.2	32.9(+1.0)	32.9(+1.0)	2.6
	18	30	31.5	34.0(+2.5)	33.9(+2.4)	6.5
		33	31.1	32.8(+1.7)	32.9(+1.8)	4.9
		35	30.7	32.0(+1.3)	32.0(+1.3)	4.4
		38	30.0	30.9(+0.9)	30.8(+0.8)	4.1
	21	30	30.8	33.7(+2.9)	33.5(+2.7)	14.0
		33	30.3	32.5(+2.2)	32.3(+2.0)	8.4
		35	29.9	31.5(+1.6)	31.4(+1.5)	6.8
		38	29.2	30.2(+1.0)	30.2(+1.0)	6.0
coastguard	10	20	35.9	39.1(+3.2)	39.1(+3.2)	6.4
		25	35.4	36.9(+1.5)	37.0(+1.6)	2.1
		28	35.0	35.8(+0.8)	35.9(+0.9)	1.2
		30	34.6	35.2(+0.6)	35.2(+0.6)	1.0
	15	28	31.5	33.6(+2.1)	33.7(+2.2)	3.6
		30	31.3	32.7(+1.4)	32.8(+1.5)	2.3
		33	30.9	31.7(+0.8)	31.7(+0.8)	1.6
		35	30.4	31.0(+0.6)	31.0(+0.6)	1.3
	18	30	29.1	31.7(+2.6)	31.7(+2.6)	4.8
		33	28.8	30.4(+1.6)	30.5(+1.7)	2.8
		35	28.4	29.6(+1.2)	29.7(+1.3)	2.2
		38	27.8	28.6(+0.8)	28.6(+0.8)	1.8
	21	30	27.3	31.4(+4.1)	30.8(+3.5)	11.0
		33	26.7	29.7(+3.0)	29.4(+2.7)	5.9
		35	26.5	28.8(+2.3)	28.5(+2.0)	4.0
		38	26.0	27.5(+1.5)	27.4(+1.4)	2.9
bus	10	20	36.0	39.3(+3.3)	39.0(+3.0)	6.9
		25	35.4	37.2(+1.8)	37.1(+1.7)	2.3
		28	35.0	36.1(+1.1)	36.1(+1.1)	1.4
		30	34.7	35.5(+0.8)	35.5(+0.8)	1.2
	15	28	31.5	34.0(+2.5)	33.9(+2.4)	3.7
		30	31.2	33.1(+1.9)	33.0(+1.8)	2.5
		33	30.8	32.1(+1.3)	32.1(+1.3)	1.8
		35	30.4	31.6 (+1.2)	31.5(+1.1)	1.6
	18	30	29.3	32.1(+2.8)	32.0(+2.7)	5.1
		33	28.9	30.9(+2.0)	30.8(+1.9)	3.1
		35	28.6	30.1(+1.5)	30.1(+1.5)	2.5
		38	28.1	29.1(+1.0)	29.0(+1.0)	2.2
	21	30	28.0	31.9(+3.9)	31.4(+3.4)	11.5
		33	27.4	30.2(+2.8)	29.9(+2.5)	5.7
		35	27.0	29.3(+2.3)	29.1(+2.1)	4.1
		38	26.5	28.0(+1.5)	27.9(+1.4)	3.1

5.4.4 The overall effect of side information

The overall effect of the side information in the rate and video quality for different combinations of average Q_{P0} and $Q_{P1,2}$ is presented in Table 5.1, where the average PSNR and extra redundancy are shown. This extra redundancy is due to the side information and it is measured as the percentage of total bit rate increase in each description, using the SDC rate obtained at the same Q_{P0} as reference. Therefore, this is the actual cost of the side information for achieving drift compensation in MDVS. Without such extra redundancy, the overall redundancy is equal to that of classic MDVS and it is in line with various MDC schemes, as discussed in 5.6. Note that, as previously pointed out, the side information does not include coded motion vectors neither prediction modes.

When a single description is received, these results demonstrate that the proposed MDVS

can significantly improve the video quality at a small cost in additional redundancy. The table shows that extra redundancy due to side information ranges from 1% to 14%, while PSNR benefits from increases between 0.6dB and 4.1dB, in comparison with classic MDVS. As previously pointed out, such PSNR improvement is due to drift compensation. Although this is dependent on the type of sequence, it is worthwhile to note that PSNR obtained from the proposed MDVS is consistently better for acceptable levels of extra redundancy. For instance, for the *foreman* sequence with $Q_{P0} = 10$ and an excess rate of 2.9%, the mean PSNR improves $1.8dB$ and for *bus* sequence with $Q_{P0} = 15$ and with an excess rate of 3.7% the mean PSNR improves $2.5dB$. Better improvements can be achieved with higher redundancy values. This is the case, for example, of the *bus* sequence where the quality improves by 3.4dB if 11.5% of excess rate is used by side information.

The results in Table 5.1 also show that redundancy of side information increases with Q_{P0}. This is because the amount of side information to be encoded increases for higher values of Q_{P0}, which in turn is due to the larger differences between SDC and each description (i.e., difference between $F_{n,0}$ and $r_{n,j}$, $j = 1, 2$ in Fig. 5.3). Moreover, for each Q_{P0}, the extra redundancy decreases with $Q_{P1,2}$. This is due to the fact that $Q_{P1,2}$ is the quantiser used to encode the side information itself. Thus, the higher the value of $Q_{P1,2}$, the smaller the respective coded rate is. These results provide useful insight for future design of efficient MDVS rate control algorithms.

5.5 MDVS streaming with path diversity

The MDVS performance using the proposed architectures was evaluated in a simulated path diversity scenario where an SDC video stream is split into two descriptions at the network edge for streaming over different paths (e.g., Fig. 5.1). After splitting the SDC stream at the MDVS edge node, each description is streamed over independent paths subject to the same packet loss rates (PLR) and average burst error length (BEL). The side information is then multiplexed and packetised along with the corresponding description. Therefore when a packet is lost, the coded description and side information are both lost. In the simulations, the packet size was set to 1000 bytes. The reference used for performance comparison of the proposed MDVS is SDC streaming under the same networking conditions, i.e., using the same amount bandwidth and suffering from equal PLR and BEL.

Burst packet loss was simulated using a Gilbert-Elliott 2-State Markov Model in order to generate different average packet loss rates and mean burst duration [Li 2009]. In order to obtain statistically meaningful results, the transmission of each sequence was simulated 100 times under the same network conditions, i.e., average PLR values of 3%,5%, 7% and 10% and average BEL of 4 and 12 packets.

Five test sequences were used, *Bus*, *Foreman*, *Mother-daughter*, *News* and *City*, CIF@30Hz. The GOP structure was IBBPBBP... with GOP=15 frames. In all cases, an index assignment matrix of 3 diagonals ($k = 1$) was used to generate the two descriptions from the compressed SDC stream. In this case a single index encoded in a description represents 3 coefficients in the original SDC stream. Frame-copy error concealment was used whenever one packet is lost. Note that, for this type of performance evaluation, such low performance concealment method is preferable over more efficient ones, because the quality results do not include masking effects due to concealment.

Figures 5.11-5.16, show the average PSNR obtained for different PLR (3%,5%, 7% and 10%), BEL and rates (1.25, 1.8 and 2.16 Mbit/s) using *bus* CIF@30Hz sequence. For PLR higher than 3%, the simulation results show that proposed method achieve better average PSNR than classic MDVS and SDC. For longer burst length and higher PLR, gains are significantly increased, particularly for longer error burst lentghs. Considering BEL=12 and PLR=7% and 10% the gains comparing with SDC are 2-3dB and comparing with classic MDVS are 1-2dB, considering all rates. For BEL=4, the proposed MDVS architecture improves the decoded video quality, where it is most significantly for PLR=5% and PLR=7%. Note that for higher PLR, the probability of losing both descriptions simultaneously is also higher, which tends to increase the influence of the error concealment at the decoder and to reduce the advantages of MDC streaming. Figs 5.11-5.16 also show a critical point around PLR=3%, which indicates that for lower values of PLR it is better to keep SDC instead of MDVS. These results suggest that such switching point should be used for no serious loss network conditions. The optimal computation of such switching under various networking conditions conditions is an open issue that deserves further investigation. Note that similar switching points are referred to in the literature, (e.g., [Franchi 2005]). Other recent MDC schemes based on DVC (e.g., [Milani 2010,Fan 2011]) also exhibit similar behaviour at relatively higher PLR (e.g., 5-15%).

As expected, in the lossless case (i.e., PLR=0%), both the SDC and classic MDVS achieve better PSNR in comparison with the proposed MDVS. This is due to the overhead required to encode the side information, since the PSNR of the three streams is compared at exactly

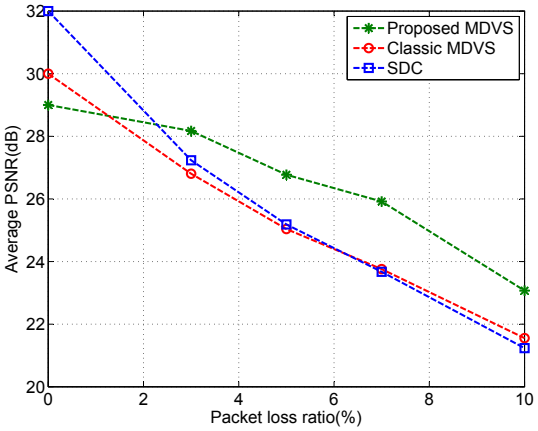

Figure 5.11: Average PNSR for *bus* at 1.25 Mbit/s (Burst length BEL=4).

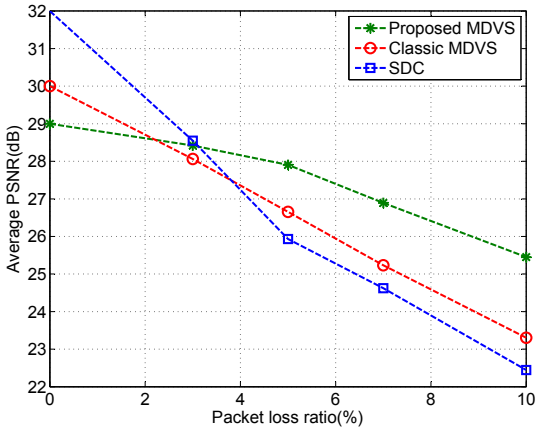

Figure 5.12: Average PNSR for *bus* at 1.25 Mbit/s (Burst length BEL=12).

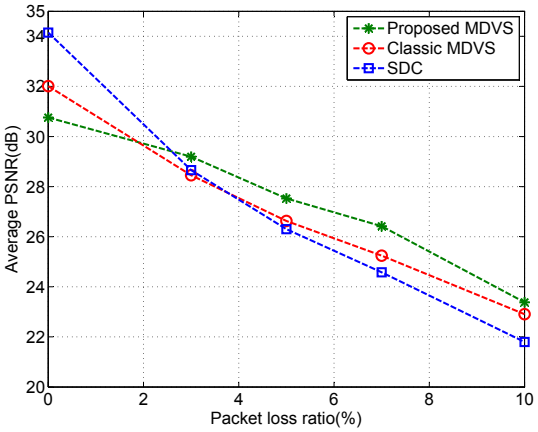

Figure 5.13: Average PNSR for *bus* at 1.8 Mbit/s (Burst length BEL=4).

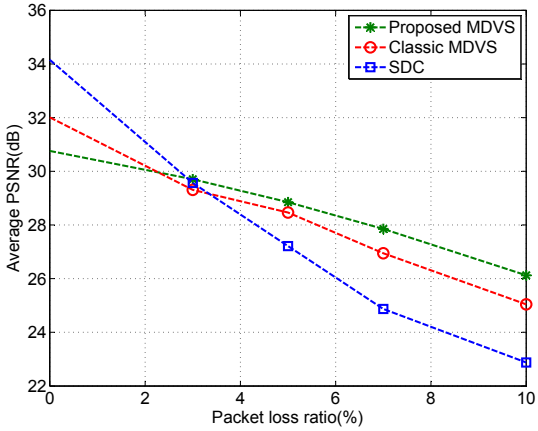

Figure 5.14: Average PNSR for *bus* at 1.8 Mbit/s (Burst length BEL=12).

the same overall bit rate. The difference of about 2-3 dB is in line with other MDC schemes available in the literature, as discussed in section 5.6.

Nevertheless, in the presence of packet loss, the perceptual quality is better because the variation of PSNR is much lower in the case of MDVS. This is shown in Fig. 5.17 for *bus* sequence, by comparing classic MDVS with proposed MDVS, affected by the same lost packets. During the period affected by packet loss, the proposed MDVS achieves PSNR gains of about 3-4dB and much lower quality variation, i.e., about 3dB variation in comparison with 7dB of classic MDVS.

Finally, Table 6.1, shows simulation results for different sequences for PLR=3%,5%,7%

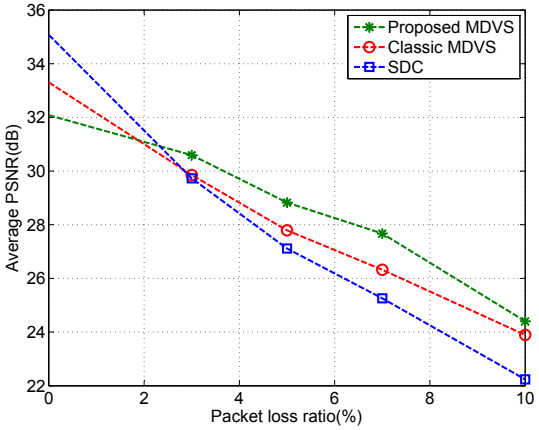

Figure 5.15: Average PNSR for *bus* at 2.16 Mbit/s (Burst length BEL=4).

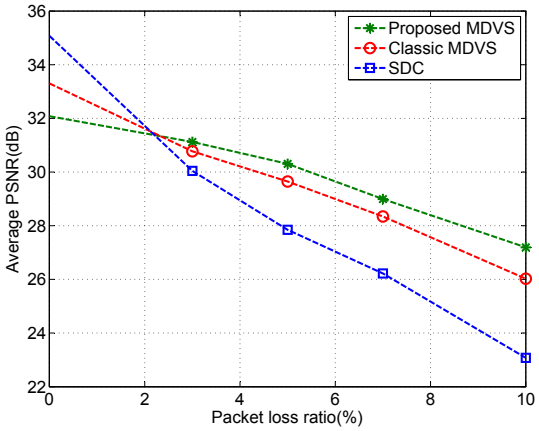

Figure 5.16: Average PNSR for *bus* at 2.16 Mbit/s (Burst length BEL=12).

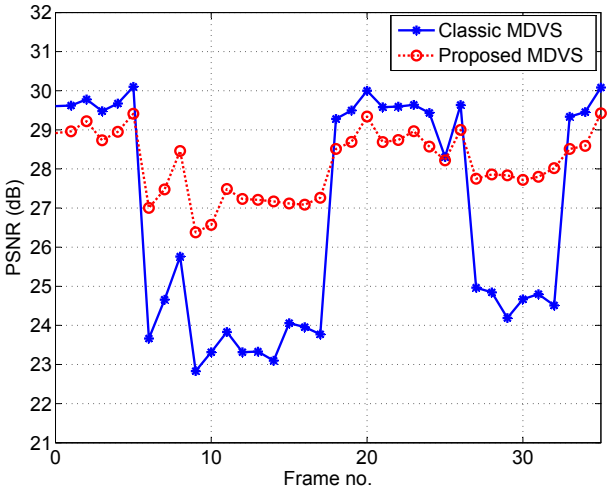

Figure 5.17: Frame by frame PSNR for *bus* sequence

and 10% and average BEL of 4 and 12 packets, at 1Mbit/s. Comparing average PSNR gains for BEL=4 packets of the proposed MDVS over classic MDVS, these are 1-2dB for PLR=3%, 1.8-2.6dB for PLR=5%, 2-2.5dB for PLR=7% and 1.1-2.6dB for PLR=10%. These results show that the proposed scheme achieves better quality while the small excess rate is beneficial by avoiding drift in case of packet loss. In comparison with SDC the average PSNR gains are -0.3-0.6 dB for PLR=3%, 0.4-1.5 dB for PLR=5%, 0.1-2.6 dB for PLR=7% and 0.1-2dB for PLR=10%. For average BEL=12 packets and comparing the proposed MDVS over classic MDVS, the gains are from 0.1-0.3dB for PLR=3%, 0.9-1dB for PLR=5%, 1.1-1.5dB for PLR=7% and 1.6-2.4dB for PLR=10%. In comparison with SDC the average PSNR gains are 0-1 dB for PLR=5%, 0.2-1.3 dB for PLR=7% and 0-

Table 5.2: Average PSNR in transmission with different packet loss rates for 1Mbps@30Hz

Sequence	BEL=4											
	Packet Loss Rate											
	3%			5%			7%			10%		
	SDC	Classic MDVS	Proposed MDVS	SDC	Classic MDVS	Proposed MDVS	SDC	Classic MDVS	Proposed MDVS	SDC	Classic MDVS	Proposed MDVS
	PSNR(dB)			PSNR(dB)			PSNR(dB)			PSNR(dB)		
Bus	26.9	25.7	26.9	24.9	24.1	25.9	22.2	22.5	24.7	21.0	20.8	23.0
Foreman	33.1	32.5	33.7	30.7	30.4	32.3	27.8	28.4	30.4	27.1	26.4	28.2
Mother	39.9	38.2	39.8	37.2	35.7	38.3	34.9	33.4	35.9	33.8	31.4	34.0
News	37.9	36.7	37.6	34.8	33.6	35.2	31.8	31.5	33.0	29.4	29.4	30.5
City	31.9	30.1	31.9	30.0	28.3	30.9	28.7	26.3	28.8	26.7	24.5	26.8
Sequence	BEL=12											
	Packet Loss Rate											
	3%			5%			7%			10%		
	SDC	Classic MDVS	Proposed MDVS	SDC	Classic MDVS	Proposed MDVS	SDC	Classic MDVS	Proposed MDVS	SDC	Classic MDVS	Proposed MDVS
	PSNR(dB)			PSNR(dB)			PSNR(dB)			PSNR(dB)		
Bus	28.3	27.1	27.3	25.9	25.6	26.6	24.8	24.8	26.1	21.9	23.2	25.1
Foreman	34.8	34.2	34.3	33.0	32.6	33.5	32.0	31.4	32.5	30.0	29.6	31.3
Mother	41.1	39.9	40.7	39.4	38	39.4	38.6	37.0	38.4	36.6	34.5	36.9
News	39.4	38.6	38.7	37.2	36.3	37.2	35.2	34.7	35.8	33.4	32.0	33.6
City	32.9	32.3	32.6	30.9	30.9	32.0	29.4	28.9	30.4	29.5	27.3	29.5

3dB for PLR=10%. Higher gains are obtained in sequences with high motion and texture complexity, like *bus* and *foreman* sequences. This means that the proposed architecture is more efficient to reduce the overall drift distortion in sequences with high motion and texture complexity.

Overall, these results show that network-adaptive MDVS for path diversity consistently improves the robustness of video streaming across networks with multiple paths. Furthermore, the proposed MDVS exhibits better performance than existing classic schemes because of its improved drift characteristics. During packet loss periods, significantly higher average PSNR is obtained at the expense of acceptable redundancy which is necessary for drift compensation.

5.6 MDVS vs MDC: comparative discussion

Despite the inherent differences between MDVS and MDC, due to the different nature of their input signals (i.e. compressed vs uncompressed video in MDVS and MDC, respectively), the rate-redundancy performance of the proposed MDVS can still be compared and discussed in the light of that obtained in previous MDC schemes at PLR=0% (same overall rate in both MDC and SDC), namely those based on multi-loop approaches that cope better with drift than classic MDVS. The overall performance of MDVS in comparison with SDC was found to have a PSNR drop of about 1.6dB and 3.1dB for rate-redundancies between 50% and 90% (including extra redundancy of side information), using the set of sequences referred to above. Under the same conditions, the multi-

loop MDC architectures proposed in [Reibman 2002], [Franchi 2005], and [Tang 2002] exhibit rate-redundancies from 40% to 100% for PSNR drops about 1.69dB to 4dB. Also in [Matty 2005], an open loop MDC scheme is proposed where the results show rate-redundancies between 45% and 100% for PSNR drops between 1.81dB and 3.35dB.

The MDC scheme proposed in [Su 2008], uses a multi-loop approach based on a spatial slice partitioning method previously used in [D. Wang 2005]. This scheme has a temporal partitioning counterpart, also using multiple-loops, proposed in [Wang 2002]. In these papers, the overall rate-redundancy distortion performance (at PLR=%0) was found to be better than that obtained in MDSQ based architectures. However, the coding approaches used by such MDC schemes cannot be used in MDVS without fully decoding the input SDC video followed by independent MDC encoding. Therefore, the use of these multi-loop MDC methods in the same networking scenarios as MDVS is highly complex, which is a significant disadvantage in comparison with MDVS. A wavelet-based MDC scheme was recently proposed in [M. Biswas 2009], but exhibits lower performance than MDVS. The results in [M. Biswas 2009] must be combined with those achieved in [Tillier 2007] in order to find out that 4dB quality drop is obtained at a relatively low rate-redundancy (i.e., 30%). Similar conclusions can be derived by comparing MDVS with another recent work in MDC [Fan 2011]. Overall, even though the rate-redundancy distortion performance of the proposed MDVS is affected by the coding distortion present in the input signal, the global rate-redundancy distortion of MDVS is inline with that of different MDC schemes.

5.7 Conclusion

This chapter demonstrates that splitting of compressed video streams into multiple descriptions using a classic MDC architecture leads to unacceptable drift accumulation, which severely affects the quality of decoded video when only one description reaches the decoder. Novel MDVS architectures were proposed to overcome the problem of drift. The proposed schemes are effective to prevent drift by using a controlled amount of side information. The experimental results provide evidence that the decoded video quality can be significantly improved at the expense of an acceptable redundancy increase in comparison with classic MDVS, for channels with distinct packet loss rates. Overall the proposed MDVS architecture finds application in multimedia networking heterogeneous environments, where lossy networks with single and multiple available paths co-exist along the same delivery chain.

Chapter 6

MDC Rate Control and Priority Video Streaming

Contents

6.1 MDC Rate Control **132**

6.1.1 The ρ model for MDC 133

6.1.2 Rate Control 134

6.1.3 R-D Performance 134

6.1.4 Rate Control Accuracy 136

6.2 Priority MDC video streaming **137**

6.2.1 MD priority streaming scenario 138

6.2.2 Packetization and Priority Streaming 139

6.2.3 Optimal binary picture classification 141

6.2.4 Simulation Results 144

6.3 Conclusions **147**

This chapter addresses the problem of rate control in unbalanced MDC video coding based on Multiple Description Scalar Quantisation (MDSQ). A new method is proposed to generate unbalanced descriptions from MDSQ and to ensure asymmetric target bit rates in each description. A new set of index assignment tables is devised to dynamically generate two unbalanced descriptions and an extension of the ρ model is developed based on the evidence that the linear relationship between the rate of each description and the corresponding percentage of zeros of transform coefficients is maintained when moving

from the single description domain (SDC) into the MDC domain. The simulation results show that the rate-distortion performance of the proposed unbalanced MDC rates is better than the equivalent balanced MDC. The rate control algorithm exhibits high accuracy for a given target bitrate and unbalanced ratio between two descriptions. The proposed method is useful in MDC video streaming over asymmetric transmission paths, where each description rate can be dynamically adapted to different channel conditions.

Also this chapter proposes a robust video streaming scheme for priority networks with path diversity, based on a combined approach of multiple description coding (MDC) with optimal picture classification into two priorities. A binary classification algorithm is proposed to define high (HP) and low (LP) priority network abstraction layer units (NALU), which in turn define the packet priorities. An optimisation algorithm is used to find HP pictures, based on dynamic programming and relying on minimisation of the packet loss concealment distortion. The paper shows that the proposed algorithm is able to effectively improve the decoded video without increasing the MDC stream redundancy. The overall performance evaluation, carried out by simulating MDC video streaming over lossy networks with path diversity, demonstrates that the proposed algorithm yields higher video quality for a wide range of packet loss rates (PLR).

6.1 MDC Rate Control

This paper proposes a rather new approach for MDSQ-based MDC. Taking the index assignment for asymmetric descriptions concept as reference, a new unbalanced MDC scheme based on MDSQ for video coding is devised. The balancing rate of each description is dynamically allocated by changing the index assignment tables of MDSQ.

Then the overall rate is controlled by an MDC rate control method capable of producing two descriptions with different rates based on the evidence that the linear dependency between the percentage of zeros of transform coefficients and the rate achieved for each description after MDSQ. From our knowledge this is the first practical implementation of unbalanced MDC using MDSQ applied to video coding. The results show that the proposed framework is able to choose the rate of each description without losing the overall rate distortion performance.

The paper is organized as follows. Section 2 introduces balanced MDSQ and section 3 describes the proposed unbalanced MDSQ method. In section 4 the MDC rate control is

described. Section 5 presents the simulation results and finally section 6 concludes the paper.

6.1.1 The ρ model for MDC

It has been shown in [He 2001] that in transform-based video coding there is a linear dependence between coding rate and the percentage of null transform coefficients ρ, after quantisation, for a given frame. This can be expressed by the following expression, using a single model parameter ϕ,

$$R(\rho) = \phi(1 - \rho). \tag{6.1}$$

We have further investigated this linear relation in order to find out whether this is maintained in each description after applying a cascade of quantisation and MDSQ to H.264/AVC transform coefficients. After an exhaustive experimental study using several different sequences, the conclusion is that there is also a linear relation between the percentage of zeros before MDSQ and the rate of each unbalanced description. Relevant results of these findings are shown in Figures 6.1 and 6.2 for I and P frames, respectively. These figures show that in both cases, the linear model can be used for unbalanced descriptions with the interesting characteristic of keeping the proportionality fairly constant among different unbalancing ratios.

Based on these findings, a ρ model was devised to determine the quantisation parameter (QP) for a given target rate for each description. For each frame, the model parameter is computed by using the source statistics of the last frame with the same type. Thus, the model parameter ϕ_i for frame i is computed by,

$$\phi_i = R_{i-1}/(1 - \rho_{i-1}). \tag{6.2}$$

After computing ϕ_i for a given frame and target bitrate, the corresponding ρ value (i.e., percentage of zeros) is determined using the specified target bit rate in equation 6.1. Then the corresponding QP value is found by using the histogram of the transform coefficients. For each quantisation step-size δ the corresponding percentage of zeros is determined from the histogram and a lookup table is built, establishing the relationship between ρ and δ. Then using the quantisation step-size previously found, the quantisation parameter QP

is determined from the following expression, already used in H.264/AVC,

$$\delta = 0.67.2^{(\frac{QP}{6})}, 0 \leq QP \leq 51. \tag{6.3}$$

6.1.2 Rate Control

In order to control the output rate, a target number of bits per frame is set according to the specified bitrate. The rate control module can be divided into two distinct levels: GOP and frame levels, respectively.

At GOP level two main operations are performed: firstly, the overall target bitrate for the whole GOP is defined. Secondly, the rate balancing among descriptions is set. The total number of bits allocated for one GOP is determined as defined in JVT-G012 rate control method [Li 2006]. The unbalance among descriptions is achieved by using the index assignment tables described in section 4.2. At each encoding time instant (i.e., frame), the target number of bits is computed following the approach described in [Li 2006], where the quadratic RD model is replaced by the proposed linear model described in section 6.1.1.

6.1.3 R-D Performance

The performance evaluation of the proposed unbalanced MDC rate control method is focused on two aspects: i) the R-D performance; ii) the rate control accuracy. The sequence *foreman*, CIF resolution, frame rate= 15Hz, GOP=IPPP... and 15 frames/GOP (baseline profile) was used in the simulation. The MDC video architecture described in [Correia 2012] was used to implement the proposed rate control method for unbalanced MDC. A buffer size equivalent to 2.4 sec delay was used in these simulations.

The coding efficiency of the proposed unbalanced MDC is compared with the balanced MDC and SDC. The target bitrate for one description and the target balancing percentages (π_1,π_2) are defined as setting parameters.

Fig. 6.3 shows the PSNR obtained from both unbalanced descriptions for differentb unbalancing ratios. The figure shows that unbalanced MDC has better overall RD performance in comparison with balanced MDC. Unbalanced MDC provides about 1dB to 1.3dB better

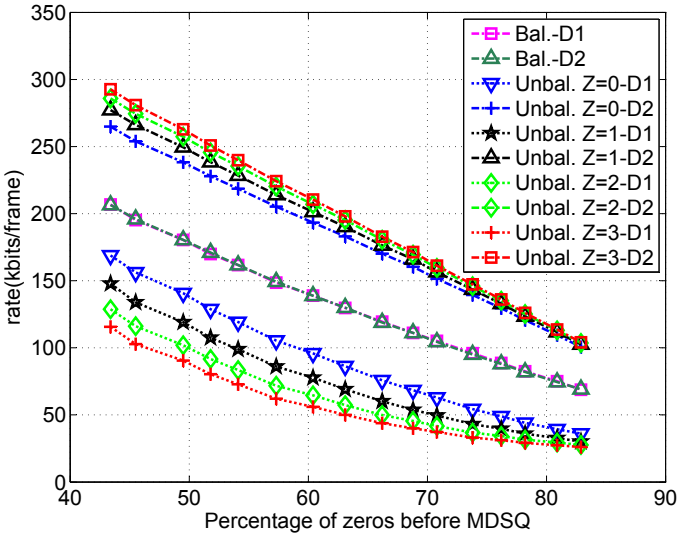

Figure 6.1: I frames

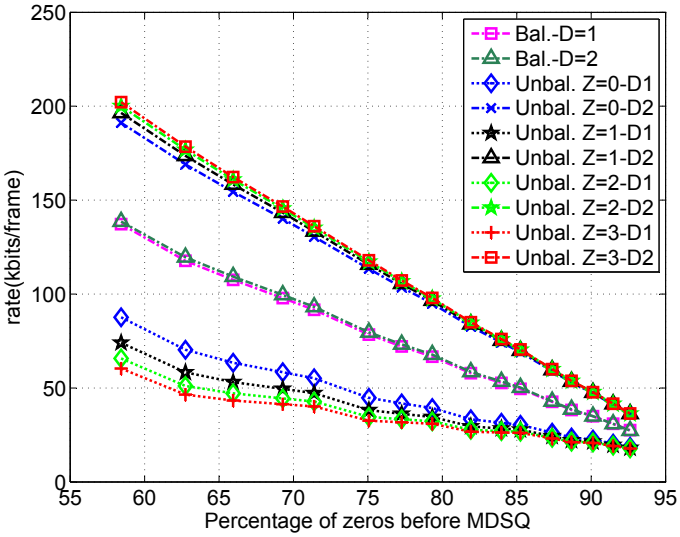

Figure 6.2: P frames

quality than balanced MDC and approximates the SDC mode as the unbalancing ratio increases. The unbalanced MDC improves the coding efficiency comparing with balance MDC, because the number of zero coefficients in one description is increased while the other carries the full value of a higher number of non-zero ones, i.e., more coefficients are not split between the two descriptions. The same behavior for unbalanced MDC was found in [Kim 2005], though in a different context. Overall these results show that unbalanced MDC achieves higher coding efficiency than balanced MDC, which proves its usefulness in transmission over channels with asymmetric available bandwidth.

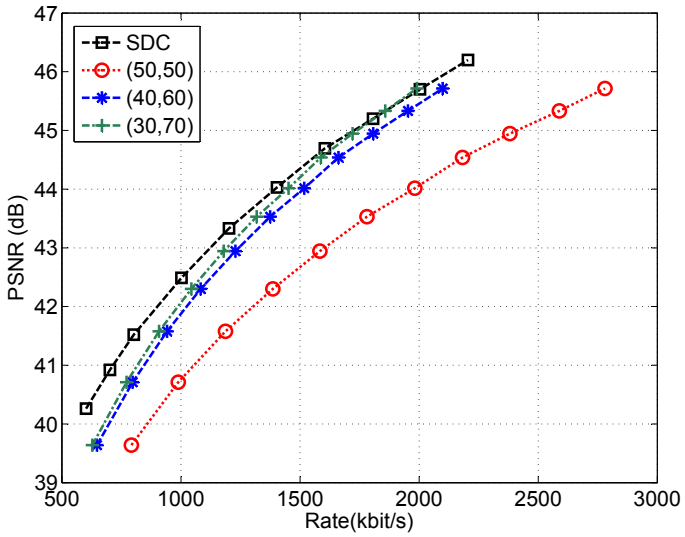

Figure 6.3: Unbalanced R-D performance

6.1.4 Rate Control Accuracy

The accuracy of the proposed rate control method was evaluated in order to find the magnitude of deviations from the specified target. Table 6.1 shows the rate obtained for description 1 R_{D1}, description 2 R_{D2}, and achieved balancing between descriptions π_1, π_2. The results show that achieved bitrate closely follows the specified target, as well as the unbalancing ratio defined for both descriptions. The small deviations obtained in the achieved balancing ratio are mainly absorbed by the rate of description 1, while description 2 does not have significant deviations. In these simulation study there was not any occurrence of buffer undeflow/overflow. Therefore the proposed rate control method achieves good overall performance to be used in unbalanced MDC applications.

Table 6.1: Rate control accuracy

Target(D2) (kbit/s)	Target Balancing (π_1,π_2)=(50,50)		
	R_{D1}	R_{D2}	(π_1,π_2)
400	390.3	401.4	(49.3, 50.7)
500	487.9	501.3	(49.3, 50.7)
600	585.4	601.5	(49.3,50.7)
700	684.2	701.7	(49.3, 50.7)
800	781.6	802.2	(49.3, 50.7)
(900)	879.8	901.5	(49.3, 50.7)
(1000)	979.7	1002.4	(49.4, 50.6)
Target(D2) (kbit/s)	Target Balancing (π_1,π_2)=(40,60)		
	R_{D1}	R_{D2}	(π_1,π_2)
400	245.7	401.4	(38.0, 62.0)
500	295.5	501.6	(37.1, 62.9)
600	340.0	602.2	(36.1, 63.9)
700	381.7	701.6	(35.1, 64.9)
800	427.0	801.5	(34.8, 65.2)
900	472.7	902.1	(34.4, 65.6)
1000	515.5	1002.4	(34.0, 66.0)
Target(D2) (kbit/s)	Target Balancing (π_1,π_2)=(30,70)		
	R_{D1}	R_{D2}	(π_1,π_2)
400	201.3	401.5	(34.7,65.3)
500	251.8	501.6	(33.4, 66.6)
600	283.5	601.9	(32.0, 68.0)
700	313.4	701.9	(31.0, 69.1)
800	342.8	802.0	(30.0, 70.0)
900	376.8	902.3	(29.5, 70.5)
1000	405.4	1002.0	(28.8, 71.2)

6.2 Priority MDC video streaming

Multiple Description Coding (MDC) is an efficient approach to improve the quality of multimedia streaming over path diversity channels. In this context, MDC video streaming is a promising communication framework particularly suited to networks with multiple available paths from the sender to the receiver (e.g., mesh and overlay networks) and multiple source coding representations (i.e., MDC). This goes beyond the classical paradigm of SDC (Single Description Coding), where each source is encoded into one single representation [Frossard 2008]. The combination of MDC with path diversity has been used in communication chains typically comprised of a source signal feeding an MDC encoder, followed by multiple transmission paths to the receiver or by streaming multiple complementary descriptions distributed across the edge servers of content delivery networks. Using MDC in video streaming over channels with multiple paths, the effect of packet loss in each path has distinct consequences in error propagation of decoded sequences in comparison with SDC streaming. For instance, burst packet losses in SDC streaming may be equivalent to isolated packet losses in multiple path channels [Apostolopoulos 2002]. Thus for higher packet loss rates, MDC schemes take advantage over SDC schemes despite the redundancy introduced by MDC [Goyal 2001a].

This paper proposes a method to improve the robustness of MDC video streaming over multiple path channels without increasing the MDC source redundancy, using a binary picture classification approach for prioritised transmission. In the past, several methods have been proposed in the literature to classify coded video units. For instance, in [Tillo 2011], an Unequal Error Protection (UEP) scheme is proposed, where the FEC redundancy allocation is a function of the relative importance of each source slice. In the context of congestion resilience, a packet scheduling framework of multiple streams over shared resources is proposed in [Chakareski 2006], based on rate-distortion information of each packet. In [Milani 2008] a packet classification scheme for MDC video streaming, based on the number of zeros of each description, is proposed to match the QoS classes supported by the IEEE 802.11e MAC layer.

The approach proposed in this paper is different from existing ones, by using pre-processing to classify pictures before encoding, based on minimization of a distortion measure that takes into account the loss of non-priority pictures. The proposed scheme uses an optimal packet priority classification method that identifies a limited set of high priority pictures within a temporal sliding window. These high priority packets carry those pictures which minimize the distortion of the decoded video affected by packet loss. The optimization criterion is based on the minimization of the picture loss concealment distortion assuming the conservative concealment model of picture-copy. Prioritised network abstraction layer units (NALU) are then packetised according to the NALU priority of each description and the corresponding packets are classified as either high-priority (HP) or low-priority (LP). This classification is used to set different levels of Quality of Experience (QoE) for HP and LP packets, which means that the most important NALU are always received at the decoder in order to maximise the decoded video quality. A novel aspect of this method is the independence from other source coding RD optimized error resilient tools, which allows its use regardless of the type of video encoder. The overall performance evaluation shows the effectiveness of such approach for different ratios of HP pictures and a wide range of packet loss ratios.

6.2.1 MD priority streaming scenario

MDC video streaming has inherent error/loss robustness, nevertheless, better QoE can be achieved when the multiple paths available for transmission comprise channels with different transmission priorities. In this case, the problem open for research is how to find

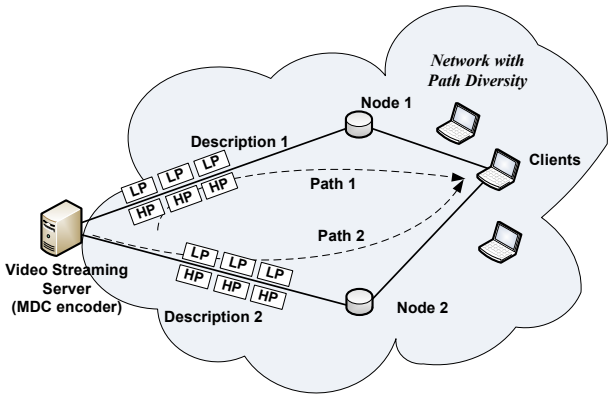

Figure 6.4: MD priority streaming scenario.

the best packet priority assignment at the MDC encoding/streaming server, based on the relevance of video content.

Figure 6.4 shows a video streaming scenario where the proposed priority scheme might be useful. An MDC streaming scenario with two disjoint paths is assumed, including an optimal priority picture classification scheme used to packetise each NALU according to its importance for the receiver. Then those packets classified as HP are delivered with no losses, while the LP ones are prone to transmission losses. Thus, packet priority classification is used to assign different levels of QoS to the source packets. Although only two channel paths are considered in this paper, the proposed method can be used in networks with high diversity, i.e., more available paths. Nevertheless, the use of two independently streams in MDC is regarded as a good compromise between capability for error recovery and compression efficiency [Apostolopoulos 2001a].

6.2.2 Packetization and Priority Streaming

The MDC scheme used in this work is tailored to be used with H.264/AVC coded video which is structured according to the network abstraction layer (NAL) concept combined with RTP encapsulation. H.264/AVC distinguishes between two different conceptual layers: the video coding layer (VCL) and the NAL. Both the VCL and NAL definitions are part of the H.264/AVC standard. The VCL specifies an efficient representation for the coded video signal, while the NAL defines the interface between the video codec itself and the outside world. NALU provides the appropriate support for packet-based transport such as current IP networks. Each slice encoded in the VLC layer corresponds to one NALU, which can be carried on a single RTP packet. The layered structure used

in this paper is shown in Figure 6.5. In order to include the side information in the standard syntax, a new type of video coding layer (VCL) NAL unit must be defined for such coded data. This can be done by extending the existing NAL types using different approaches. For instance in H.264/SVC [Ye-Kui 2007], new NAL unit types were defined to accommodate several layers and associated information and in [Lamy-Bergot 2010], a new type of NAL unit is proposed for embedding redundant information inside standard video streams.

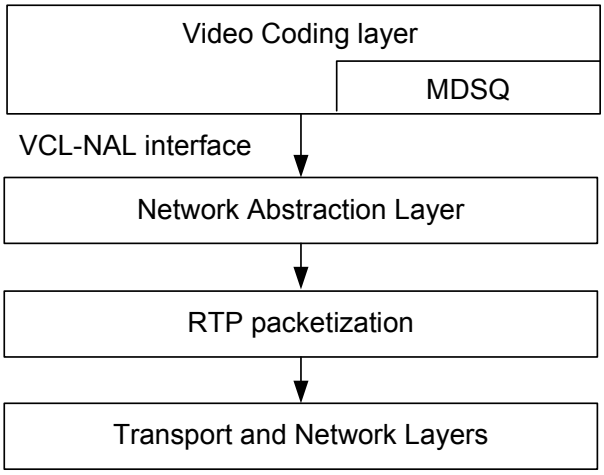

Figure 6.5: NAL packetization.

Figure 6.6 shows the proposed optimal MDC priority streaming scheme. The optimal binary picture classification described in section 6.2.3 is applied to a temporal window comprising several pictures. Prioritised NALU are then packetised according to the NALU priority of each description and the corresponding packets are classified as either high-priority (HP) or low-priority (LP). This is achieved by setting the first byte in the NAL unit packet, the *nal_ref_idc* (NRI) bits, that identifies the relative transport priority of each NAL unit [**?**].

After reordering the received packets, the decoder knows which packets are lost in any description and then decides how to decode the received stream. If both corresponding packets are received, then the MDSQ decoder merges the respective decoded indices in order to obtain the original transform coefficients. If only the packet of one description is received, then the decoder uses the side information in order to decode the corresponding slice with controlled mismatch. In the case where both packets are lost, then error concealment must be used (in this work the conservative picture-copy method is used).

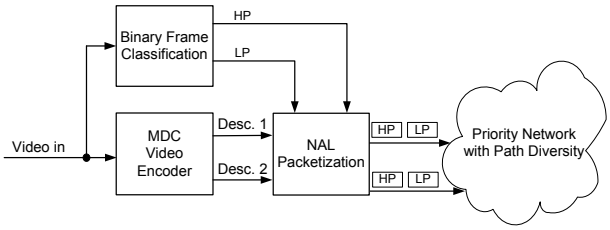

Figure 6.6: Priority MDC streaming scheme.

6.2.3 Optimal binary picture classification

The optimal binary picture classification algorithm identifies a set of m HP pictures and its complementary set of $n - m$ LP ones, from a temporal video segment with n pictures, defined as a sliding window over the original sequence. The classification of HP pictures is based on minimization of the picture loss concealment distortion at the decoder. Given the set of HP pictures, the whole temporal segment can be reconstructed using any interpolation algorithm for concealment. The picture-copy method was chosen for this work, i.e., zero-order interpolation. The algorithm finds the set of $m < n$ HP pictures which is the best representation of its corresponding temporal segment. This means that the temporal segment is fully reconstructed from the HP set of pictures with minimum concealment distortion, i.e., assuming that all LP pictures are missing and concealed by interpolation of the HP ones.

6.2.3.1 Definitions and Formulations

Let a temporal segment of n pictures be denoted by $F = \{f_0, f_1, ..., f_{n-1}\}$, where the subscripts represent the temporal order of pictures. The corresponding HP picture set with m pictures is denoted by $P = \{f_{l_0}, f_{l_1}, ..., f_{l_{m-1}}\}$, in which l_k is the picture index (referred to the source set F) of the k^{th} element of the set P. P is determined by the index picture selection $l_0, l_1, ..., l_{m-1}$, such that $l_0, < l_1, ..., < l_{m-1}$. Note that l_0, l_1, .., l_{m-1} are defined in the temporal segment F, thus the l_k values do not comprise an arithmetic progression. For example, given a temporal segment $F = \{f_0, f_1, f_2, f_3, f_4\}$ a possible HP picture set would be $P = \{f_0, f_3\}$, with $l_0 = 0$ and $l_{m-1} = 3$.

The reconstructed set $F' = \{f'_0, f'_1, ..., f'_{n-1}\}$ from the HP picture set P is obtained using picture-copy by substituting missing picture with the most recent picture that belongs to P, i.e., the same picture is repeated along the time where HP pictures do not exist, that

is

$$f'_k = f_{i=max(l):s.t.l\varepsilon\{l_0,l_1,...,l_{m-1}\},i\leq k} \tag{6.4}$$

The picture concealment distortion $d(f_k, f'_k)$ quantifies the difference between picture f_k of F and its corresponding picture f'_k from F'. In this paper, we use the mean squared error (MSE) between the two pictures as the metric for computing the concealment distortion, but other metric can be used, since the optimal solution is independent from the distortion metric. Note that if f_k is selected into P, then $d(f_k, f'_k) = 0$. The MSE metric can be calculated as:

$$d(f_k, f'_k)_{MSE} = \frac{1}{h * w} \sum_{y=0}^{h-1} \sum_{x=0}^{w-1} (f_k(x, y) - f'_k(x, y))^2 \tag{6.5}$$

The concealment distortion is computed as the average picture distortion between the temporal segment F and its approximation set F', i.e., the one concealed from the HP picture set P. This concealment distortion is given by:

$$D(P) = 1/n \sum_{k=0}^{n-1} d(f_k, f'_k) \tag{6.6}$$

The HP ratio $R(P)$ is the ratio between the number of pictures m belonging to the HP picture set P and the total number of pictures n of temporal segment F. Therefore,

$$R(P) = m/n \tag{6.7}$$

6.2.3.2 Problem Formulation

Using the definition presented in the previous subsection, we formulate the binary priority picture classification algorithm as a HP ratio-distortion concealment optimization problem where the objective function is to find a subset of pictures in each temporal segment that provides its best representation within a given maximum HP ratio R_{max}, i.e., without using more than $m = R_{max} * n$ HP pictures. Given the HP ratio constraint R_{max}, the optimal HP picture set P^* is the one that minimizes the concealment distortion given by,

$$P^* = \arg\min_P D(P), s.t. R(P) \leq R_{max} \tag{6.8}$$

where $R(P)$ and $D(P)$ are defined by the equation (6.7) and (**??**) respectively. For example, given a temporal segment F of $n = 10$ pictures and a priority ratio $R(P) = 0.2$, the proposed algorithm classifies at most 2 pictures as high priority, i.e. $m = 2$. Thus, the constraint R_{max} is a user-defined parameter to define the number of HP pictures in each temporal segment.

6.2.3.3 Dynamic Programming Solution

In general, there are $\binom{n-1}{m-1} = \frac{(n-1)!}{(m-1)!(n-m)!}$ feasible solutions to above problem, assuming the first picture of temporal segment always belongs to set P. When n and m are large, an exhaustive search solution would be prohibitive. Alternatively to this solution, the dynamic programming method was used to solve the priority ratio-distortion concealment optimization problem by breaking it down into simpler subproblems in a recursive manner [**?**]. In a previous work a similar method was used for different purpose [**?**]. The concealment distortion stage D_t^k is defined as the minimum total concealment distortion incurred by HP picture set that has t pictures and ends with the picture $f_k(l_{t-1} = k)$. Therefore,

$$D_t^k = \min_{l_1,l_2,...,l_{t-2}} \sum_{j=0}^{n-1} d(f_j, f_{i=max(l):s.t.l\varepsilon\{0,l_1,l_2,...,l_{t-2}\},i\leq j}) \tag{6.9}$$

Note, $l_0 = 0$ and $l_{t-1} = k$ are removed from the optimization process. Since $0 < l_1 < l_2 < ... < l_{t-2} < k$, and $i \leq j$. After some math manipulation, the current concealment distortion stage D_t^k can be broken into two parts, (6.10), where the first part is the previous concealment distortion stage $D_{t-1}^{l_{t-2}}$ already computed, and it represents the minimum total concealment distortion produced by HP picture set with $t-1$ pictures and ending with picture corresponding to index l_{t-2} and the second part, $e^{l_{t-2},k}$ represents the previous concealment distortion stage reduction, if picture k is selected into the HP picture set of $t-1$ pictures ending in picture with the l_{t-2} index. Therefore, we have

$$D_t^k = \min_{l_{t-2}} \{D_{t-1}^{l_{t-2}} - e^{l_{t-2},k}\} \tag{6.10}$$

where the distortion reduction was defined by,

$$e^{l_{t-2},k} = \sum_{j=k}^{n-1} [d(f_j, f_{l_{t-2}}) - d(f_j, f_k)] \tag{6.11}$$

Since it was assumed that the first picture is always selected into the hight priority set P, the initial concealment distortion stage D_1^0 is given as

$$D_1^0 = \frac{1}{n}\sum_{j=1}^{n-1} d(f_0, f_j) \tag{6.12}$$

At this point, we can compute the current concealment distortion stage D_t^k for any HP picture set of t pictures and ending with the picture k by the recursion in (6.10) with the initial concealment distortion stage given by (6.12).

Finally, the HP picture set P is given by the calculation of picture index $l_0, l_1, ..., l_{m-1}$ found by backtracking the calculations already performance of D_t^k (6.10). Therefore,

$$\begin{aligned} l_{m-1} &= \arg\min_k \{D_m^k\} \\ l_0 &= 0 \end{aligned} \tag{6.13}$$

where l_0 and l_{m-1} represent the first and last picture index selected into the high priority picture set $P = \{f_{l_0}, f_{l_1}, ..., f_{l_{m-1}}\}$, respectively. In this work, each temporal segment F was defined as being coincident with the group of pictures (GOP) used by the MDC encoder and each picture corresponds to one NALU.

6.2.4 Simulation Results

The performance of the MDC priority streaming scheme described in the previous sections, was evaluated assuming path diversity communication with several PLR. Two main scenarios were compared: MDC with binary picture classification and MDC without binary picture classification.

A random packet loss generator with uniform distribution is used to drop packets according to the required packet loss rate. Each description is streamed over two independent paths with the same PLR. The packets classified as HP are not lost. In order to obtain statistically sound results, the transmission of each sequence (*mother & daughter*, *foreman* and *bus*, CIF resolution, frame rate= 30Hz) was simulated 50 times under the same network conditions, i.e., average PLR of 3%, 5%, 7%, 10% and 15%. The GOP structure

is IPPPP... with GOP=15 pictures, i.e., constant temporal segment $F = 15$ pictures. A fixed QP set was used to encode each sequence, i.e., *mother & daughter* was encoded with $QP_I = 23$ and $QP_P = 24$, *foreman* was encoded with $QP_I = 30$ and $QP_P = 31$ and *bus* was encoded with $QP_I = 37$ and $QP_P = 38$. The resulting MDC rate for both descriptions added together is about 1 Mbps, including the inherent redundancy.

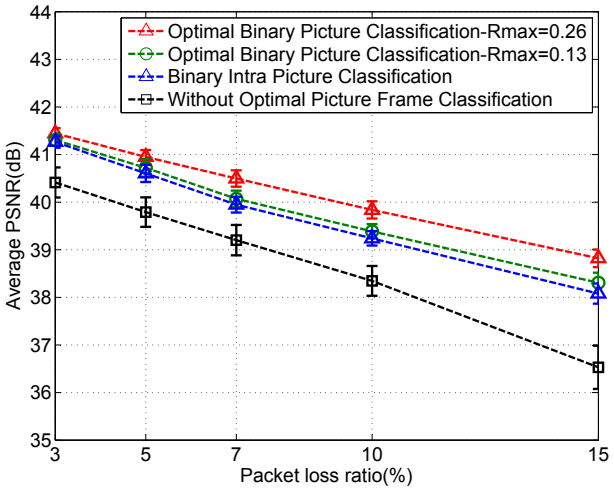

Figure 6.7: Simulation results for *Mother & daughter* sequence.

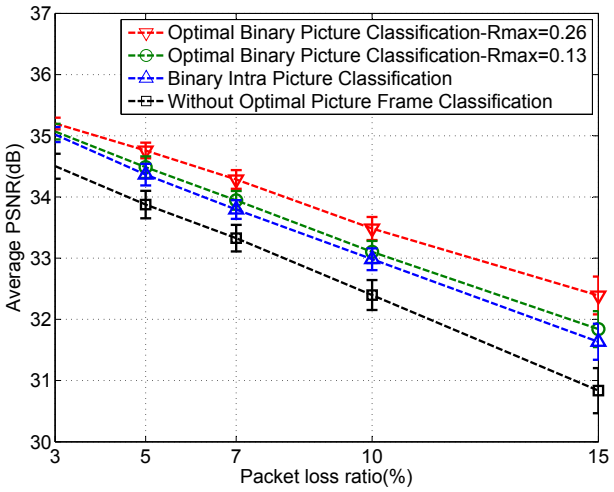

Figure 6.8: Simulation results for *Foreman* sequence.

The picture-copy error concealment was used whenever packets carrying the same NALU in both descriptions are simultaneously lost. In order to evaluate the impact of priority picture classification, the HP constraining ratio was set into two fixed levels, $R_{max} = 0.13$ and $R_{max} = 0.26$, which corresponds to 2 and 4 HP pictures per GOP respectively. In all experiments, the I pictures were always set as HP pictures. The proposed method is compared with i) MDC with Intra Picture Binary Classification: only I pictures are classified

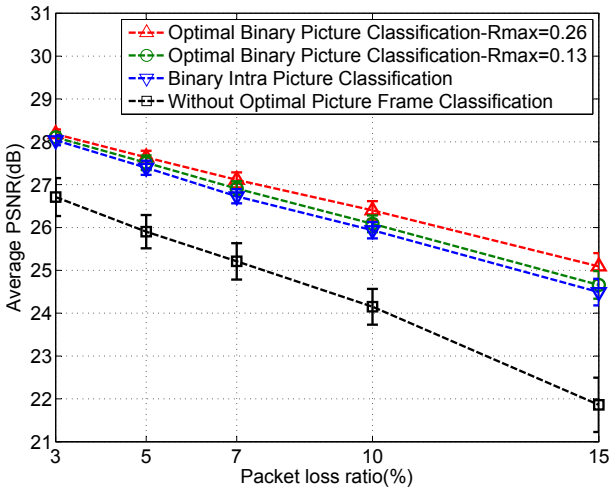

Figure 6.9: Simulation results for *Bus* sequence.

as HP and ii) MDC Without Binary Picture Classification methods: i.e., corresponding to use MDC without any picture classification method.

Figures 6.7, 6.8 and 6.9, show the video quality, i.e., average peak signal-to-noise ratio (PSNR) and 95% confidence intervals, obtained after simulation for the sequences *mother&daughter*, *foreman* and *bus* respectively, under random packet loss. The simulation results show that the average PSNR improves for all PLR when the proposed optimal binary picture classification is used. For instance, considering the *mother&daughter* sequence and comparing the proposed method, using $R_{max} = 0.26$, with MDC without HP pictures, the average PSNR gains are 1.03dB for PLR=3%, 1.16dB for PLR=5%, 1.30dB for PLR=7%, 1.49dB for PLR=10% and 2.3dB for PLR=15%. Comparing MDC where only I pictures are HP with the no priority scheme, the PSNR gains are 0.85dB for PLR=3%, 0.82dB for PLR=5%, 0.75dB PLR=7%, 0.89dB for PLR=10% and 1.55dB for PLR=15%. On the other hand, the confidence intervals show that average PSNR have smoother variations in the case where picture classification and priority is used.

These results show that classifying intra pictures as HP significantly improves the overall video quality. Also, comparing the simulation results of the proposed method with $R_{max} = 0.26$ and MDC with only I pictures as HP, the PSNR gains are 0.18dB for PLR=3%, 0.35dB for PLR=5%, 0.55dB for PLR=7%, 0.60dB for PLR=10% and 0.74dB for PLR=15%, which means that optimal choice of few priority pictures over a GOP, improves the average video quality received at the user end.

Note that in any case, packet loss always leads to distortion propagation along the affected

GOP in the video decoder because HP pictures are not intra-coded, i.e., they are P pictures. This means that, even with a very small amount of priority packets the overall quality of decoded video is significantly improved. Identical results are obtained for *foreman* and *bus* sequences in Figures 6.8 and 6.9.

Moreover MDC schemes usually have higher error resilience capabilities in comparison with other error resilience methods, such as adaptive intra refresh [Radulovic 2010]. Therefore, since the proposed scheme improves the performance of classic MDC, one might infer that better performance is also obtained in comparison with those methods. Overall, these results show that the optimal priority MDC streaming based on binary picture classification can be used to improve the quality of decoded sequences in MDC video streaming over multipath networks. In particular, the proposed scheme exhibits consistently better performance over a non-priority MDC scheme.

6.3 Conclusions

In this chapter a new rate control method for unbalance MDC video coding was proposed. The method was shown to be effective in unbalanced rate control of MDSQ descriptions by using a set of new index assignment tables. The experimental results show that overall RD performance is improved using unbalanced MDC in comparison with balanced MDC. Therefore, the proposed scheme finds application in video delivery over path diversity networks with asymmetric rate constraints.

Also this chapter demonstrates that improved quality can be achieved in video streaming over path diversity networks by using MDC combined with optimal binary picture classification for setting packet priority. The proposed scheme was shown to be effective in increasing the QoE by classifying NALUs into two levels of priority, according to their relevance for the concealment distortion. The experimental results show that the decoded video quality can be significantly improved under different channel loss conditions. Therefore, the proposed scheme finds application in video delivery over path diversity networks where different levels of priorities can be implemented.

Chapter 7

Conclusions and future work

Contents

7.1 Conclusions 149

7.2 Future Work 152

7.1 Conclusions

Throughout this thesis new MDC schemes for advanced video coding and MDC video adaptation have been proposed.

In Chapter 2, the theoretical definitions and optimal rate-distortion bounds were discussed and also the MDC techniques proposed in the literature were described, and have included Multiple Description with Correlating Transforms (MDCT), Multiple Description Sub-band Coding, based on Polyphase Decomposition and Selective Quantisation, Multiple Description Scalar Quantisation, and MDC based on the Error Correction Codes.

In Chapter 3 the main MDC video architectures were described for the different video coding schemes. The state of the art of MDC for video was described and discussed including applications using scalable coding, multi-view and stereo video coding, Wyner-Ziv coding and also for high-dimensional MDC. Different path diversity topologies and the MDC networking applications were also discussed.

Chapter 4 has described a multiple description coding scheme for Advanced Video Coding based on a MDSQ mixed open-loop/multi-loop structure. Drift distortion accumulation in both intra and motion compensated predicted slices was evaluated. The proposed

method use side information is only applied to anchor slices maintaining a reduced redundancy cost, which allows to prevent drift propagation for single description decoding. A new MDSQ scheme for unbalanced multiple description coding was proposed using new types of index assignment tables used to change dynamically the rate of each description and also controlling drift distortion at the decoder. Simulation results have shown that proposed method has the advantage to reduce the redundancy of MDSQ without loss of MDC error robustness. Also, the proposed method allows to add side information to unbalanced description, without losing the balancing achieved by MDSQ. Following the theory considerations, one of the descriptions would give a low quality reconstruction. Experimental results show us that with a small amount of side information, the side-distortion has a good performance comparing inclusively with balanced case. This fact is due to the increasing of coding efficiency in both central and side information in unbalanced MDSQ compared with balanced MDSQ. Further investigation is needed in order to characterize the overall distortion including the side-information and also to derive new index assignment tables that gives more granularity in balancing rates.

Chapter 5 has addressed a new MDVS scheme in order to adapt an SDC compressed stream into a MDC with two descriptions to act in any network edge node. The main novel aspects of the proposed MDVS scheme comprise a two-loop MDVS architecture with drift control in both intra and inter predictive coded slices. An equivalent single-loop architecture that reduce the operations that are needed. The proposed architecture generates side-information from the compressed streams without full decoding reusing the most of the original coding parameters and coding structure, such as slice maps. Such feature gives the capability of controlling the amount of side information according to the expected decoder drift which allow to outperforms the open-loop schemes available in the literature. A evaluation discussion is addressed comparing the proposed MDVS method with other results obtained in literature in MDC context, and is concluded that the overall performance similar to MDC using uncompressed video, which is possible to affirm that is a valid approach to be used in any network edge node with MDVS capability. Further investigation is needed in order to allow adapt the overall rate of MDVS scheme. Re-quantization schemes of the original SDC bitstream associated to MDSQ schemes are need to be added to the overall schemes. Current distributed processing schemes, like cloud computing can be explored in order to obtain more efficient MDVS schemes, tanking account the coding efficiency, drift control and the number of descriptions.

In Chapter 6 a new rate control method for MDSQ video coding has been proposed thats

ensure an asymmetric target bit rate in each description based on ρ model. The linear relationship between the rate of each description and the corresponding percentage of zeros of transform coefficients for MDSQ indices is evaluated which is maintained when moving from the single description domain (SDC) into the MDC domain. Based on this evaluation, a new rate control method was proposed, which is independent of the side information allocation. This new method overcomes the lack of using fixed index assignment tables and allows to dynamically balance each description rate and also to choose the coding parameters of the MDC framework. Several improvements can be done in the proposed method. The achieved results are very interesting in order to control the rate of the central information part of each description. However an optimized redundancy rate-distortion method is needed in order to combine each rate component, adapting the coding parameters to the channel conditions, which is the available rate and also the packet loss rate. Despite this drawback, the proposed method presents an effective evolution in rate control methods for MDC schemes, in particular using MDSQ.

Also in Chapter 6 has been addressed a robust video streaming scheme for priority networks with path diversity, based on a combined approach MDC with optimal picture classification into two priorities. An optimization algorithm finds high priority frames based on minimization of the packet loss concealment distortion. In this section was shown that quality improvement at the decoder can be achieved in video streaming over path diversity networks by using MDC combined with optimal binary picture classification for setting packet priority. The experimental results show that the decoded video quality can be significantly improved under different channel loss conditions. Therefore, the proposed scheme finds application in video delivery over path diversity networks where different levels of priorities can be implemented. An improved distortion model that takes into account the error propagation among descriptions will improve the overall performance of the proposed method. Nevertheless, simulation results show that the main idea that is prioritizing each packet of each description in order to be used in schedule systems whenever is possible allows to affirm that proposed scheme is an effective solution to be used in MDC video frameworks.

7.2 Future Work

Adding some contribution to further research in topics addressed in this thesis, several ideas must be discussed.

MDC for High Efficiency Video Coding In test model of next generation coding standard called high efficiency video coding (HEVC) has adopted new coding structures, such as the usage of larger coding unit and transform blocks dimensions. The coding unit which is larger than 16×16 and less than or equal to 64×64 provides great bit rate saving while the coding complexity increases dramatically. This new coding units will give new spacial and temporal prediction dependencies which means that new MDC schemes must be developed in order to improve the coding efficiency and error resilience capabilities. For instance, new schemes based on redundant information with different qualities adapted for HEVC structures can be evaluated.

MDC for 3D video Multi-view video coding is inherently MDC, where each view can be seen as a description of some video scene where each view has redundancy among them. Presently the main research effort has been done in coding efficiency and also in 3D rendering both as in video+deph representation, but also in view synthesis. Multi-view video transmission needs a research effort in order do give error resiliency this new video format characterized by its high complexity prediction dependencies. MDC can make an important issue in order provide error resilience in 3D video transmission. This can be done simplifying the prediction dependencies, introducing some loss of coding efficiency, but also to exploit disjoint different interview predictions, introducing some redundancy but adding several decoding combinations. Also several quality representations can be combined in a rate- distortion sense in order to improve quality at decoding exploiting also the specific subjective evaluation of 3D video that poses new problems is perceived quality of video contents.

MDC adaptation using distributed computing Distributed computing, namely cloud computing, have been largely disseminated with increasing of access network technologies. Adapting video contents at some network edge or mobile device can be done using distributed computing, where some tasks can be done in another device with different processing capabilities. The new concept of MDC video adaptation with drift control

introduced in this thesis can be evaluated in this new context, eventually exploiting more complex schemes in terms of rate-distortion optimization and coding dependencies, which implies full-decoding operations. Also several prediction representations provided by re-encoding operations improves the flexibility of the MDC scheme in terms of drift control and decoding descriptions.

Optimal Solutions in High Dimensional Spaces Multiple descriptions transmitted over multipath lossy networks gives rise to a multidimensional problem for which no complete solution is yet known. The problem that can be posed is to find the optimal number of descriptions, as well as the bit rates that should be allocated to each one under known network conditions (e.g., packet loss statistics, bandwidth) such that the total average distortion per user is minimized. New optimization methods are needed to evaluate for this multidimensional problem that can be find application in new schemes of streaming video where video contents can be divided in small chunks that can be distributed in several network nodes.

Appendix A

Published Papers

A.1 Journal Papers

J1. Pedro D. F. Correia, Pedro A. Amado Assunção, Vitor Silva, "Multiple Description of Coded Video for Path Diversity Streaming Adaptation", *IEEE Transactions On Multimedia*, vol. 14, no. 2-3, pp. 923-935, June 2012.

A.2 Conference Papers

C1. Pedro D. F. Correia, L.F. Ferreira, Pedro A. Amado Assunção, Luís Cruz, Vitor Silva, "Optimal priority MDC video streaming for networks with path diversity", *Proc. International Conference on Telecommunications and Multimedia (TEMU)*, Heraklion, Greece, pp.54-59, July 2012.

C2. Pedro D. F. Correia, Pedro A. Amado Assunção, Vitor Silva, "Enhanced H.264/AVC Video Streaming using Network-adaptive Multiple Description Coding", *Proc. IEEE Region 8 EUROCON 2011 International Conference on Computer as a Tool*, Lisbon, Portugal, April 2011.

C3. Pedro D. F. Correia, Pedro A. Amado Assunção, Vitor Silva, "Multiple Description Video Transcoding with Temporal Drift Control", *Proc. Picture Coding Symposium-PCS*, Nagoya, Japan, Vol. 1, pp. 558-561, December 2010.

C4. Pedro D. F. Correia, Pedro A. Amado Assunção, Vitor Silva, "Drift-free Multiple Description Intra Video Coding", *Proc. IEEE International Conference on Image Processing-ICIP*, Cairo, Egypt, Vol. 1, pp. 625-628, November 2009.

C5. Pedro D. F. Correia, Pedro A. Amado Assunção, Vitor Silva, "Multiple Description Coding Scheme for H.264/AVC Intra Slices", *Proc. Conference on Telecommunications - ConfTele*, Santa Maria da Feira, Portugal, May 2009.

References

[15444-1 2002] Rec. ITU-T T.802—ISO/IEC 15444-1. *Recommendation ITU-T T.802: JPEG 2000 image coding system*, 2002.

[A. Serdar Tan 2009] Godze Akar A. Serdar Tan Anil Aksay and Erdal Arikan. *Rate Distortion Optimization for Steroecopic Video Streaming with Unequal Error Protection.* Eurasip Journal On Advances in Signal Processing, 2009.

[Abanoz 2009] T. Berkin Abanoz and A. Murat Tekalp. *SVC-Based scalable multiple description video coding and optimization of encoding configuration.* Signal Processing: Image Communication, no. 24, pages 691–701, 2009.

[Ahuja 2008] S. Ahuja and M. Krunz. *Algorithms for Server Placement in Multiple-Description-Based Media Streaming.* IEEE Transactions on Multimedia, vol. 10, no. 7, pages 1382 –1392, Nov. 2008.

[Akyol 2007] E. Akyol, A. M. Tekalp and M. R. Civanlar. *A Flexible Multiple Description Coding Framework for Adaptive Peer-to-Peer Video Streaming.* IEEE Journal of Selected Topics in Signal Processing, vol. 1, no. 2, pages 231–245, 2007.

[Apostolopoulos 2000] J.G. Apostolopoulos. *Error-resilient video compression through the use of multiple states.* In Image Processing, 2000. Proceedings. 2000 International Conference on, volume 3, pages 352 –355 vol.3, 2000.

[Apostolopoulos 2001a] J. G. Apostolopoulos. *Reliable Video Communication over Lossy Packet Networks using Multiple State Encoding and Path Diversity.* Proceedings of Visual Communications and Image Processing, Jan. 2001.

[Apostolopoulos 2001b] J.G. Apostolopoulos and S.J. Wee. *Unbalanced multiple description video communication using path diversity.* In Proceedings. 2001 International Conference on Image Processing, volume 1, pages 966 –969 vol.1, 2001.

[Apostolopoulos 2002] J. Apostolopoulos, T. Wong, W. tian Tan and S. Wee. *On Multiple Description Streaming with Content Delivery Networks.* In IEEE INFOCOM, 2002.

[Bajic 2003] I. V. Bajic and J. W. Woods. *Domain based multiple description coding of images and video.* IEEE Transactions on Image Processing, vol. 12, no. 10, pages 1211–1225, Oct. 2003.

[Begen 2005] Ali C. Begen, Yucel Altunbasak, Ozlem Ergun and Mostafa H. Ammar. *Multi-path selection for multiple description video streaming over overlay networks.* Signal Processing: Image Communication, vol. 20, no. 1, pages 39 – 60, 2005.

[Bin 2006] Li Bin, Huang Feng, Sun Lifeng and Yang Shiqiang. *Optimized Rate Allocation for Unbalanced Multiple Description Video Coding Over Unreliable Packet Network.* In 2006 IEEE International Conference on Multimedia and Expo(ICME), pages 105 –108, july 2006.

[C. Tian 2004] S. S. Hemami C. Tian. *Universal Multiple Description Scalar Quantization:Analysis and Design.* IEEE Transactions On Information Theory, vol. 50, no. 9, pages 2089–2102, Sept. 2004.

[Campana 2008] O. Campana, R. Contiero and G. A. Mian. *An H.264/AVC Video Coder Based on a Multiple Description Scalar Quantizer.* IEEE Transactions On Circuits and Systems For Video Technology, vol. 18, no. 2, pages 268–272, Feb. 2008.

[Cardinal 2004] J. Cardinal. *Entropy-constrained index assignments for multiple description quantizers.* IEEE Transactions on Signal Processing, vol. 52, no. 1, pages 265 – 270, jan. 2004.

[Castro 2003] Miguel Castro, Peter Druschel, Anne-Marie Kermarrec, Animesh Nandi, Antony Rowstron and Atul Singh. *SplitStream: high-bandwidth multicast in cooperative environments.* In Proceedings of the nineteenth ACM symposium on Operating systems principles, SOSP '03, pages 298–313, New York, NY, USA, 2003. ACM.

[Chakareski 2005] J. Chakareski, S. Han and B. Girod. *Layered coding vs. multiple descriptions for video streaming over multiple paths.* Multimedia Systems, vol. 10, no. 4, pages 275–285, 2005.

[Chakareski 2006] J. Chakareski and P. Frossard. *Rate-distortion optimized distributed packet scheduling of multiple video streams over shared communication resources.* Multimedia, IEEE Transactions on, vol. 8, no. 2, pages 207 – 218, april 2006.

[Chakareski 2008] J. Chakareski and P. Frossard. *Distributed Collaboration for Enhanced Sender-Driven Video Streaming.* Multimedia, IEEE Transactions on, vol. 10, no. 5, pages 858 –870, aug. 2008.

[Chen 2007a] Chih-Ming Chen, Chien-Min Chen, Chia-Wen Lin and Yung-Chang Chen. *Error-Resilient video streaming over wireless networks using combined scalable coding and multiple-description coding.* Signal Processing: Image Communication, vol. 22, pages 403–420, 2007.

[Chen 2007b] Chih-Ming Chen, Chia-Wen Lin, Hsiao-Cheng Wei and Yung-Chang Chen. *Robust video streaming over wireless LANs using multiple description transcoding and prioritized retransmission.* Journal of Visual Communication & Image Representation, vol. 18, pages 191–206, 2007.

[Choi 1999] Seung-Jong Choi and J.W. Woods. *Motion-compensated 3-D subband coding of video.* IEEE Transactions on Image Processing, vol. 8, no. 2, pages 155 –167, feb 1999.

[Chou 2006] P.A. Chou and Zhourong Miao. *Rate-distortion optimized streaming of packetized media.* Multimedia, IEEE Transactions on, vol. 8, no. 2, pages 390 – 404, april 2006.

[Comas 2003] David Comas, Raghavendra Singh, Antonio Ortega and Ferran Marques. *Unbalanced Multiple Description Video Coding with Rate-Distortion Optimization.* In EURASIP Journal on Applied Signal Processing, pages 81–90, 2003.

[Correia 2012] Pedro D.F. Correia, P. Assuncao and V. Silva. *Multiple Description of Coded Video for Path Diversity Streaming Adaptation.* IEEE Transactions on Multimedia, vol. 14, no. 2-3, pages 923–935, June 2012.

[Cover 2006] Thomas M. Cover and Joy A. Thomas. Elements of information theory 2nd edition (wiley series in telecommunications and signal processing). Wiley-Interscience, July 2006.

[Crave 2008] Olivier Crave, Christine Guillemot, Béatrice Pesquet-Popescu and Christophe Tillier. *Distributed temporal multiple description coding for robust video transmission.* EURASIP J. Wirel. Commun. Netw., vol. 2008, January 2008.

[Crave 2010] O. Crave, B. Pesquet-Popescu and C. Guillemot. *Robust Video Coding Based on Multiple Description Scalar Quantization With Side Information.* IEEE Transactions on Circuits and Systems for Video Technology, vol. 20, no. 6, pages 769 –779, June 2010.

[D. Wang 2005] N. Canagarajah D. Wang and D. Bull. *Slice Group based multiple description video coding wth three motion compensation loops.* In IEEE International Sysmposium on Circuits and Systems, pages 960–963, May 2005.

[Dragotti 2002] P.L. Dragotti, S.D. Servetto and M. Vetterli. *Optimal filter banks for multiple description coding: analysis and synthesis.* Information Theory, IEEE Transactions on, vol. 48, no. 7, pages 2036 –2052, jul 2002.

[Erhan Ekmekcioglu 2010] Maheshi Dissanayake Erhan Ekmekcioglu Banu Günel, Stweart T. Worral and Ahmet M. Kondoz. *A Scalable Multi-view Audiovisual Entertainment Framework With Content-Aware Distribuition.* In IEEE International Conference On Image Processing, ICIP, 2010, pages 2401–2404, Sept. 2010.

[Essaili 2007] Ali El Essaili, Shaoib Khan, Wolfgang Kellerer and Eckehard Steinbach. *Multiple Description Video Transcoding.* In IEEE International Conference on Image Processing (ICIP), volume 6, pages 77–78, 2007.

[Fan 2011] Yuhua Fan, Jia Wang and Jun Sun. *Distributed Multiple Description Video Coding on Packet Loss Channels.* IEEE Transactions on Image Processing, vol. 20, no. 6, pages 1768 –1773, June 2011.

[Favalli 2011] Lorenzo Favalli and Marco Folli. *ILPS: a scalable multiple description coding scheme for H.264.* Multimedia Tools and Applications, vol. 54, pages 609–634, 2011. 10.1007/s11042-010-0577-0.

[Franchi 2005] N. Franchi, M. Fumagalli, R. Lancini and S. Tubaro. *Multiple Description Video Coding for Scalable and Robust Transmission Over* IP. IEEE Transactions On Circuits and Systems For Video Technology, vol. 14, no. 3, pages 321–334, March 2005.

[Frossard 2008] P. Frossard, J.C. de Martin and M. Reha Civanlar. *Media Streaming With Network Diversity.* Proceedings of the IEEE, vol. 96, no. 1, pages 39 –53, Jan. 2008.

[Gamal 1982] A.E. Gamal and T. Cover. *Achievable rates for multiple descriptions.* Information Theory, IEEE Transactions on, vol. 28, no. 6, pages 851 – 857, nov 1982.

[Gan 2006] Tong Gan, Lu Gan and Kai-Kuang Ma. *Reducing video-quality fluctuations for streaming scalable video using unequal error protection, retransmission, and*

interleaving. IEEE Transactions on Image Processing, vol. 15, no. 4, pages 819 –832, April 2006.

[Ghareeb 2010] M. Ghareeb and C. Viho. *A Multiple Description Coding Approach for Overlay Multipath Video Streaming Based on QoE Evaluations*. In Multimedia Information Networking and Security (MINES), 2010 International Conference on, pages 39–43, Nov. 2010.

[Goyal 2001a] V. Goyal. *Multiple Description Coding: Compression Meets the Network*. IEEE Signal Processing Magazine, vol. 18, no. 5, pages 74–93, Sept. 2001.

[Goyal 2001b] V. Goyal and J. Kovacevic. *Generalized Multiple Description Coding With Correlated Transforms*. IEEE Transactions on Information Theory, vol. 47, no. 6, pages 2199–2224, Sept. 2001.

[H. Abdul Karim 2009] S. Worrall H. Abdul Karim A. Sali, Abduk H. Sadka and A.M. Kondoz. *Multiple Description Video Coding for Stereoscopic 3D*. Consumer Electronics, IEEE Transactions on, vol. 55, no. 4, pages 2048–2056, November 2009.

[He 2001] Zhihai He and S.K. Mitra. *A unified rate-distortion analysis framework for transform coding*. Circuits and Systems for Video Technology, IEEE Transactions on, vol. 11, no. 12, pages 1221 –1236, dec 2001.

[Hsi-Tzeng Chan 2005] Chih-Ming Fu Hsi-Tzeng Chan and Chung lin Huang. *A New Error Resilient Video Coding Using Matching Pursuit and Multiple Description Coding*. IEEE Transactions On Circuits and Systems For Video Technology, vol. 15, no. 8, pages 1047–1052, Aug. 2005.

[Hsiao 2010] C.-W. Hsiao and W.-J. Tsai. *Hybrid Multiple Description Coding Based on* H.264. IEEE Transactions on Circuits and Systems for Video Technology, vol. 20, no. 1, pages 76 –87, Jan. 2010.

[ISO/IEC13818-2 2000] ISO/IEC13818-2. *Generic coding of moving pictures and associated audio information: Video*, 2000.

[ISO/IEC14496-10 2012] ISO/IEC14496-10. *Coding of audio-visual objects – Part 10: Advanced Video Coding*, 2012.

[Jurca 2007] Dan Jurca, Jakob Chakareski, Jean-Paul Wagner and Pascal Frossard. *Enabling Adaptive Video Streaming in P2P Systems*. IEEE Communications Magazine, vol. 45, no. 6, pages 108–114, 2007.

[Kamnoonwatana 2011] Nawat Kamnoonwatana, Dimitris Agraotis and Nishan Canagarajah. *Rate controlled redundancy-adaptive multiple description video coding.* Signal Processing Image Communication, vol. 26, no. 5, pages 205 – 219, 2011.

[Khan 2004] S. Khan, R. Schollmeier and E. Steinbach. *A performance comparison of multiple description video streaming in peer-to-peer and content delivery networks.* In Multimedia and Expo, 2004. ICME '04. 2004 IEEE International Conference on, volume 1, pages 503 –506 Vol.1, june 2004.

[Kim 1997] Beong-Jo Kim and W.A. Pearlman. *An embedded wavelet video coder using three-dimensional set partitioning in hierarchical trees (SPIHT).* In Data Compression Conference, 1997. DCC '97. Proceedings, pages 251 –260, mar 1997.

[Kim 2000] Beong-Jo Kim, Zixiang Xiong and W.A. Pearlman. *Low bit-rate scalable video coding with 3-D set partitioning in hierarchical trees (3-D SPIHT).* Circuits and Systems for Video Technology, IEEE Transactions on, vol. 10, no. 8, pages 1374 –1387, dec 2000.

[Kim 2001] C.-Su Kim and S.-Uk Lee. *Multiple Description Coding of Motion Fields for Robust Video Transmission.* IEEE Transactions On Circuits and Systems For Video Technology, vol. 11, no. 9, pages 999–1010, Sept. 2001.

[Kim 2003] Il Koo Kim, Nam Ik Cho and Jeho Nam. *Error Resilient Video Transcoding Dased on the Optimal Multiple Description of DCT Coefficients.* In International Workshop on Advanced Image Technology, 2003.

[Kim 2005] Joohee Kim, R.M. Mersereau and Y. Altunbasak. *Distributed video streaming using multiple description coding and unequal error protection.* IEEE Transactions on Image Processing, vol. 14, no. 7, pages 849 –861, july 2005.

[Kim 2006] Il Koo Kim and Nam Ik Cho. *Video Transcoding for Packet Loss Resilience Based on the Multiple Desciptions.* In International Conference on Multimedia and Expo(ICME), pages 109–112, 2006.

[Kondi 2005] L. P. Kondi. *A Rate-Distortion Optimal Hybrid Scalable/Multiple-Desciption Video Codec.* IEEE Transactions On Circuits and Systems For Video Technology, vol. 15, no. 7, pages 921–927, July 2005.

[Lamy-Bergot 2010] Catherine Lamy-Bergot and Benjamin Gadat. *Embedding Protection Inside H.264/AVC and SVC Streams.* EURASIP Journal on Wireless Communications and Networking, vol. 2010, 2010.

[Lefol 2006] D. Lefol, D. Bull and N. Canagarajah. *Performance Evaluation of Transcoding Algorithms for H.264.* IEEE Transactions on Consumer Electronics, vol. 52, no. 1, pages 215–221, 2006.

[Li 1998] Xin Li, Bo Tao and M.T. Orchard. *On implementing transforms from integers to integers.* In Image Processing, 1998. ICIP 98. Proceedings. 1998 International Conference on, pages 881 –885 vol.3, Oct. 1998.

[Li 2006] Z.G. Li, W. Gao, F. Pan, S.W. Ma, K.P. Lim, G.N. Feng, X. Lin, S. Rahardja, H.Q. Lu and Y. Lu. *Adaptive rate control for H.264.* Journal of Visual Communication and Image Representation, vol. 17, no. 2, pages 376 – 406, 2006.

[Li 2009] Zhicheng Li, J. Chakareski, Xiaodun Niu, Yongjun Zhang and Wanyi Gu. *Modeling of distortion caused by Markov-model burst packet losses in video transmission.* In IEEE International Workshop on Multimedia Signal Processing (MMSP), pages 1 –6, Oct. 2009.

[Liao 2011] Y. Liao and J. D. Gibson. *Routing-aware multiple description video coding over mobile ad-hoc networks.* IEEE Transactions On Multimedia, vol. 13, no. 1, pages 132–142, 2011.

[Lin 2011] Chunyu Lin, T. Tillo, Yao Zhao and Byeungwoo Jeon. *Multiple Description Coding for H.264/AVC With Redundancy Allocation at Macro Block Level.* IEEE Transactions on Circuits and Systems for Video Technology, vol. 21, no. 5, pages 589 –600, may 2011.

[Liu 2005] Yilong Liu and S. Oraintara. *Drift-free multiple description video coding with redundancy rate-distortion optimization.* In Circuits and Systems, 2005. ISCAS 2005. IEEE International Symposium on, pages 4034 – 4037 Vol. 4, may 2005.

[Liu 2008a] Jiangchuan Liu, S.G. Rao, Bo Li and Hui Zhang. *Opportunities and Challenges of Peer-to-Peer Internet Video Broadcast.* Proceedings of the IEEE, vol. 96, no. 1, pages 11 –24, jan. 2008.

[Liu 2008b] Yong Liu, Yang Guo and Chao Liang. *A survey on peer-to-peer video streaming systems.* Peer-to-Peer Networking and Applications, vol. 1, no. 1, pages 18–28, March 2008.

[Lu 2007] Meng-Ting Lu, Jui-Chieh Wu, Kuan-Jen Peng, P. Huang, J.J. Yao and H.H. Chen. *Design and Evaluation of a P2P IPTV System for Heterogeneous Networks.* Multimedia, IEEE Transactions on, vol. 9, no. 8, pages 1568 –1579, dec. 2007.

[M. Biswas 2009] J. Arnold M. Biswas M. Frater and M. Pickering. *Improved Resilience for Video Over Packet Loss Networks with MDC and Optimized Packetization.* IEEE Transactions On Circuits and Systems For Video Technology, vol. 19, no. 10, pages 1556–1560, Oct. 2009.

[M. Pereira 2003] M. Antonioni M. Pereira and M. Barlaud. *Multiple Description Image and Video Coding for Wireless Channels.* Elsevier Signal Prcessing: Image Communications, vol. 18, pages 925–945, 2003.

[Ma 2004] Zheng Ma, Huai-Rong Shao and Chia Shen. *A new multi-path selection scheme for video streaming on overlay networks.* In Communications, 2004 IEEE International Conference on, volume 3, pages 1330 – 1334, June 2004.

[Ma 2008] Mengyao Ma, O.C. Au, Liwei Guo, S.-H.G. Chan and P.H.W. Wong. *Error Concealment for Frame Losses in MDC.* IEEE Transactions on Multimedia, vol. 10, no. 8, pages 1638 –1647, dec. 2008.

[Magharei 2007] N. Magharei, R. Rejaie and Yang Guo. *Mesh or Multiple-Tree: A Comparative Study of Live P2P Streaming Approaches.* In INFOCOM 2007. 26th IEEE International Conference on Computer Communications. IEEE, pages 1424 –1432, May 2007.

[Mao 2005] Shiwen Mao, Y.T. Hou, Xiaolin Cheng, H.D. Sherali and S.F. Midkiff. *Multiple description video in wireless ad hoc networks.* In INFOCOM 2005. 24th Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings IEEE, volume 1, pages 740–750, March 2005.

[Matty 2005] K. R. Matty and L. P. Kondi. *Balanced Multiple Description Video Coding Using Optimal Partitioning of the DCT Coefficients.* IEEE Transactions On Circuits and Systems For Video Technology, vol. 15, no. 7, pages 928–934, July 2005.

[M.B.Dissanayake 2010] S.T.Worral M.B.Dissanayake D.V.S.X. De Silva and W.A.C. Fernando. *Error Resilience Technique for Multi-view Coding Using Redundant Disparity Vectors.* In IEEE International Conference On Multimedia and Expo, ICME, 2010, pages 1712–1717, 2010.

[Milani 2008] S. Milani, G. Calvagno, R. Bernardini and R. Rinaldo. *A low-complexity packet classification algorithm for multiple description video streaming over IEEE802.11E networks.* In Image Processing, 2008. ICIP 2008. 15th IEEE International Conference on, pages 3072 –3075, oct. 2008.

[Milani 2010] S. Milani and G. Calvagno. *Multiple Description Distributed Video Coding Using Redundant Slices and Lossy Syndromes.* IEEE Signal Processing Letters, vol. 17, no. 1, pages 51 –54, Jan. 2010.

[N. Kamonoonwatana 2010] D. Agrafiotis N. Kamonoonwatana and C.N. Canagarajah. *Rate-Controlled Redundancy Multiple Description Coding For Video Transmission Over MIMO Systems.* In IEEE 17th International Conference on Image Processing (ICIP), pages 1269–272, Sept. 2010.

[Noh 2012] Jeonghun Noh and Bernd Girod. *Robust mobile video streaming in a peer-to-peer system.* Signal Processing: Image Communication, vol. 27, no. 5, pages 532 – 544, 2012.

[Norkin 2006] Andrey Norkin, Anil Aksay, Cagdas Bilen, Gozde Akar, Atanas Gotchev and Jaakko Astola. *Schemes for Multiple Description Coding of Stereoscopic Video.* In Multimedia Content Representation, Classification and Security, volume 4105 of *Lecture Notes in Computer Science*, pages 730–737. Springer Berlin / Heidelberg, 2006.

[Ohm 1994] J.-R. Ohm. *Three-dimensional subband coding with motion compensation.* Image Processing, IEEE Transactions on, vol. 3, no. 5, pages 559 –571, sep 1994.

[Orchard 1997] M.T. Orchard, Y. Wang, V. Vaishampayan and A.R. Reibman. *Redundancy rate-distortion analysis of multiple description coding using pairwise correlating transforms.* In Image Processing, 1997. Proceedings., International Conference on, volume 1, pages 608 –611 vol.1, oct 1997.

[Ozarow 1980] L. Ozarow. *On a Source Coding Problem with Two Channels and Three Receivers.* The Bell System Technical Journal, vol. 59, no. 10, pages 1909 –1921, Dec. 1980.

[Padmanabhan 2003] V.N. Padmanabhan, H.J. Wang and P.A. Chou. *Resilient peer-to-peer streaming.* In Network Protocols, 2003. Proceedings. 11th IEEE International Conference on, pages 16 – 27, nov. 2003.

[Peraldo 2010] L. Peraldo, E. Baccaglini, E. Magli, G. Olmo, R. Ansari and Y. Yao. *Slice-level rate-distortion optimized multiple description coding for H.264/AVC.* In IEEE International Conference on Acoustics Speech and Signal Processing (ICASSP), pages 2330 –2333, March 2010.

[Pereira 2009] Fernando Pereira. *Distributed video coding: basics, main solutions and trends.* In Proceedings of the 2009 IEEE international conference on Multimedia and Expo, ICME'09, pages 1592–1595, Piscataway, NJ, USA, 2009. IEEE Press.

[Puri 1999] R. Puri and K. Ramchandran. *Multiple description source coding using forward error correction codes.* In Signals, Systems, and Computers, 1999. Conference Record of the Thirty-Third Asilomar Conference on, volume 1, pages 342 –346 vol.1, 1999.

[Puri 2001] R. Puri, K. won Lee, K. Ramchandran and V. Bharghavan. *An Integrated Source Transcoding and Congestion Control Paradigm for Video Streaming in the Internet.* IEEE Transactions on Multimedia, vol. 3, pages 18–32, 2001.

[Qu 2006] Qi Qu, Yong Pei and J.W. Modestino. *An Adaptive Motion-Based Unequal Error Protection Approach for Real-Time Video Transport Over Wireless IP Networks.* Multimedia, IEEE Transactions on, vol. 8, no. 5, pages 1033 –1044, oct. 2006.

[Radulovic 2010] I. Radulovic, P. Frossard, Y.-K. Wang, M. M. Hannuksela and A. Hallapuro. *Multiple Description Video Coding with H.264/AVC Redundant Pictures.* IEEE Transactions On Circuits and Systems For Video Technology, vol. 20, no. 1, pages 144–148, 2010.

[Rane 2008] S. Rane, P. Baccichet and B. Girod. *Systematic Lossy Error Protection of Video Signals.* Circuits and Systems for Video Technology, IEEE Transactions on, vol. 18, no. 10, pages 1347 –1360, oct. 2008.

[Reibman 2001] A. Reibman, H. Jafarkhani, Yao Wang and M. Orchard. *Multiple description video using rate-distortion splitting.* In Image Processing, 2001. Proceedings. 2001 International Conference on, volume 1, pages 978 –981 vol.1, 2001.

[Reibman 2002] A. R. Reibman, H. Jafarkhami, Y. Wang and M. T. Orchard. *Multiple-Description Video Coding Using Motion-Compensated Temporal Prediction.* IEEE Transactions On Circuits and Systems For Video Technology, vol. 12, no. 3, pages 193–204, March 2002.

[Rong 2010] Bo Rong, Yi Qian, Kejie Lu, R.Q. Hu and M. Kadoch. *Multipath routing over wireless mesh networks for multiple description video transmission.* Selected Areas in Communications, IEEE Journal on, vol. 28, no. 3, pages 321 –331, April 2010.

[S. Lin 2001] Y. Wang S. Lin S. Mao and S. Panwar. *A Reference Picture Selection Scheme for video Transmission over Ad-HOC Networks using multiple paths.* In IEEE International Conference on Multimedia and Expo (ICME), pages 96–99, Aug. 2001.

[Said 1996] A. Said and W.A. Pearlman. *A new, fast, and efficient image codec based on set partitioning in hierarchical trees.* Circuits and Systems for Video Technology, IEEE Transactions on, vol. 6, no. 3, pages 243 –250, Jun. 1996.

[Schmidt 2011] J.C. Schmidt and K. Rose. *Jointly Optimized Mode Decisions in Redundant Video Streaming.* IEEE Transactions on Circuits and Systems for Video Technology, vol. 21, no. 4, pages 513 –518, april 2011.

[Setton 2008] E. Setton, P. Baccichet and B. Girod. *Peer-to-Peer Live Multicast: A Video Perspective.* Proceedings of the IEEE, vol. 96, no. 1, pages 25 –38, jan. 2008.

[Shanableh 2008] T. Shanableh, T. May and F. Ishtiaq. *Error resiliency transcoding and decoding solutions using distributed video coding techiques.* Signal Processing:Image Communication, vol. 23, pages 610–623, 2008.

[Shapiro 1992] J.M. Shapiro. *An embedded wavelet hierarchical image coder.* In Acoustics, Speech, and Signal Processing, 1992. ICASSP-92., 1992 IEEE International Conference on, volume 4, pages 657 –660 vol.4, mar 1992.

[Slepian 1973] D. Slepian and J. Wolf. *Noiseless coding of correlated information sources.* Information Theory, IEEE Transactions on, vol. 19, no. 4, pages 471 – 480, jul 1973.

[Su 2008] C.-C. Su, Homer H. Chen, Jason J. Yao and P. Huang. H.264/AVC-*based multiple description video coding using dynamic slice groups*. Image Commun., vol. 23, no. 9, pages 677–691, 2008.

[Tang 2002] X. Tang and A. Zakhor. *Matching Pursuits Multiple Description Coding for Wireless Video*. IEEE Transactions On Circuits and Systems For Video Technology, vol. 12, no. 6, pages 566–575, June 2002.

[Tillier 2007] C. Tillier, T. Petrisor and B. P.-Popescu. *A Motion-Compensated Overcomplete Temporal Decomposition for Multiple Description Scalable Video Coding*. EURASIP Journal on Image and Video Processing, 2007.

[Tillo 2008] T. Tillo, M. Grangetto and G. Olmo. *Redundancy Slice optimal Allocation for H.264 Multiple Description Coding*. IEEE Transactions On Circuits and Systems For Video Technology, vol. 18, no. 1, pages 59–70, 2008.

[Tillo 2010] T. Tillo, E. Baccaglini and G. Olmo. *Multiple Descriptions Based on Multirate Coding for JPEG 2000 and H.264/AVC*. IEEE Transactions on Image Processing, vol. 19, no. 7, pages 1756 –1767, July 2010.

[Tillo 2011] T. Tillo, E. Baccaglini and G. Olmo. *Unequal Protection of Video Data According to Slice Relevance*. Image Processing, IEEE Transactions on, vol. 20, no. 6, pages 1572 –1582, June 2011.

[Tsai 2010] W.J. Tsai and J.Y. Chen. *Joint Temporal and Spatial Error Concealment for Multiple Description Video Coding*. IEEE Transactions on Circuits and Systems for Video Technology, vol. PP, no. 99, page 1, 2010.

[Tsai 2012] Wen-Jiin Tsai and Hao-Yu You. *Multiple Description Video Coding Based on Hierarchical B Pictures Using Unequal Redundancy*. Circuits and Systems for Video Technology, IEEE Transactions on, vol. 22, no. 2, pages 309 –320, feb. 2012.

[Vaidyanathan 1990] P.P. Vaidyanathan. *Multirate digital filters, filter banks, polyphase networks, and applications: a tutorial*. Proceedings of the IEEE, vol. 78, no. 1, pages 56 –93, Jan. 1990.

[Vaishampayan 1993] V. A. Vaishampayan. *Design of Multiple Description Scalar Quantizers*. IEEE Transactions On Information Theory, vol. 39, no. 3, pages 821–834, May 1993.

[Vaishampayan 1994] V.A. Vaishampayan and J. Domaszewicz. *Design of entropy-constrained multiple-description scalar quantizers.* Information Theory, IEEE Transactions on, vol. 40, no. 1, pages 245 –250, jan 1994.

[Vaishampayan 1998] V.A. Vaishampayan and J.-C. Batllo. *Asymptotic analysis of multiple description quantizers.* Information Theory, IEEE Transactions on, vol. 44, no. 1, pages 278 –284, jan 1998.

[Vaishampayan 1999] V.A. Vaishampayan and S. John. *Balanced interframe multiple description video compression.* In Image Processing, 1999. ICIP 99. Proceedings. 1999 International Conference on, volume 3, pages 812 –816 vol.3, 1999.

[van der Schaar 2003] M. van der Schaar and D.S. Turaga. *Multiple description scalable coding using wavelet-based motion compensated temporal filtering.* In Proceedings. 2003 International Conference on Image Processing, 2003. ICIP 2003., volume 3, pages III – 489–92 vol.2, sept. 2003.

[Verdicchio 2006] F. Verdicchio, A. Munteanu, A. I. Gavrilescu, J. Cornelis and P. Schelkens. *Embedded Multiple Description Coding of Video.* IEEE Transactions On Image Processing, vol. 15, no. 10, pages 3114–3130, Oct. 2006.

[W. Jian 1999] A. Ortega W. Jian. *Multiple Description Coding via Polyphase Transform and Selective Quantization.* In SPIE Visual Communications and Image Processing Conference (VCIP 99), 1999.

[Wang 2001] Y. Wang, M. T. Orchard, V. Vaishampayan and A. R. Reibman. *Multiple Description Coding Using Pairwise Correlating Transforms.* IEEE Transactions On Image Processing, vol. 10, no. 3, pages 351–366, March 2001.

[Wang 2002] Y. Wang and S. Lin. *Error-Resilient Video Coding Using Multiple Description Motion Compensation.* IEEE Transactions On Circuits and Systems For Video Technology, vol. 12, no. 6, pages 438–452, June 2002.

[Wang 2003] H. Wang and A. Ortega. *Robust Video Communication by Combining Scalability and Multiple Description Coding Techniques.* Proc. of SPIE Image and Video Communications and Processing, vol. 1, Jan. 2003.

[Wang 2005a] D. Wang, N. Canagarajah and D. Bull. *S frame design for multiple description video coding.* In IEEE International Symposium on Circuits and Systems, ISCAS'2005, pages 2719 – 2722 Vol. 3, May 2005.

[Wang 2005b] Y. Wang, A. R. Reibman and S. Lin. *Multiple Description Coding for Video Delivery.* Proceedings of the IEEE, vol. 93, no. 1, pages 57–70, Jan. 2005.

[Wei 2007] W. Wei and A. Zakhor. *Multiple Tree Video Multicast Over Wireless Ad Hoc Networks.* Circuits and Systems for Video Technology, IEEE Transactions on, vol. 17, no. 1, pages 2 –15, jan. 2007.

[Wei 2009] Wei Wei and A. Zakhor. *Interference Aware Multipath Selection for Video Streaming in Wireless Ad Hoc Networks.* Circuits and Systems for Video Technology, IEEE Transactions on, vol. 19, no. 2, pages 165 –178, Feb. 2009.

[Wyner 1976] A. Wyner and J. Ziv. *The rate-distortion function for source coding with side information at the decoder.* Information Theory, IEEE Transactions on, vol. 22, no. 1, pages 1 – 10, jan 1976.

[Xiu 2009] Xiaoyu Xiu and Jie Liang. *View Interpolation Based Multiple Description Coding of Multiview Images.* In Asilomar,2009, pages 876–880, 2009.

[Xua 2012] Yuanyuan Xua, Ce Zhua, Wenjun Zengb and Xue Jun Lia. *Multiple description coded video streaming in peer-to-peer networks.* Signal Processing: Image Communication, vol. 27, no. 5, pages 412 – 429, 2012.

[Y.-C. Lee 2004] Y. Altunbasak Y.-C. Lee and R. M. Mersereau. *An Enhanced Multiple Description Video Coder With Drift Reduction.* IEEE Transactions On Circuits and Systems For Video Technology, vol. 14, no. 1, pages 122–127, Jan. 2004.

[Yang 2000] Xuguang Yang and K. Ramchandran. *Optimal subband filter banks for multiple description coding.* Information Theory, IEEE Transactions on, vol. 46, no. 7, pages 2477 –2490, nov 2000.

[Ye-Kui 2007] Wang Ye-Kui, M.M. Hannuksela, S. Pateux, A. Eleftheriadis and S. Wenger. *System and Transport Interface of SVC.* IEEE Transactions on Circuits and Systems for Video Technology, vol. 17, no. 9, pages 1149 –1163, Sept. 2007.

[Zamir 1999] R. Zamir. *Gaussian codes and Shannon bounds for multiple descriptions.* Information Theory, IEEE Transactions on, vol. 45, no. 7, pages 2629 –2636, nov 1999.

[Zha 2005] CoolStreaming/DONet: a data-driven overlay network for peer-to-peer live media streaming, volume 3, March 2005.

[Zhu 2009] C. Zhu and M. Liu. *Multiple Description Video Coding Based on Hierarchical B Pictures*. IEEE Transactions On Circuits and Systems For Video Technology, vol. 19, no. 4, pages 511–521, April 2009.