



Multimedia Communication over Mobile IP Wireless Networks

Sajid Nazir

**Thesis for the Degree of PhD
Department of Electronic and Electrical Engineering
2012**

The copyright of this thesis belongs to the author under the terms of the United Kingdom Copyright Acts as qualified by the University of Strathclyde Regulation 3.49. Due acknowledgement must always be made of the use of any material contained in, or derived from, this thesis.

I, Sajid Nazir, hereby declare that this submission is my own work and that, to the best of my knowledge and belief, it contains no material previously published or written by another person nor material which has been accepted for the award of any other degree or diploma of the university or other institute of higher learning, except where due acknowledgment has been made in the text.

Abstract

The use of Internet Protocol (IP) based mobile wireless transmission is increasing as novel multimedia applications are being deployed. Mobile wireless channels and IP based communications are inherently prone to errors and packet losses. Error resilience features and Forward Error Correction (FEC) at the application layer (AL) are often used to protect the video data against losses. The amount of redundancy added by the FEC attempts to counter the worst channel Signal-to-noise-ratio (SNR) but the protection generally comes at a high complexity and overhead.

It is thus imperative to design FEC solutions which are adaptive to the varying wireless channel conditions, i.e., bandwidth and packet loss rate. This adaptive behaviour becomes even more important for transmission to heterogeneous receivers. Fountain codes are rateless codes which can be used to potentially generate an unlimited number of encoded packets from a limited number of source packets. The decoding is possible if the number of received encoded packets at the receiver is just a little more than the source packets. As each portion of encoded video data does not have equal importance for the video re-construction, this characteristic can also be advantageously exploited while designing FEC solutions by providing more protection to important portions.

Random linear codes (RLC) based schemes have been compared with Raptor codes, and RLC solution is proposed for the mobile television broadcasting standards like Digital Video Broadcasting-Handheld (DVB-H) and DVB-T2 (Second Generation Terrestrial).

A reliable unicast video communication solution based on Luby Transform (LT) codes is proposed by exploiting unequal error protection (UEP) for encoded video data partitioned with the Data partitioning (DP) and slicing feature of H264/AVC.

A comparison of layered video data transmission with Amplify-and-Forward (AF) and Decode-and-Forward (DF) relay collaboration strategies is provided. A novel scheme for Multiple description coding (MDC) has been proposed and its advantages highlighted through simulations over relay based multi hop channels, like Long Term Evolution-Advanced (LTE-A).

An algorithm has been proposed which takes into account the PSNR contribution and temporal significance of each slice to prioritize H.264/AVC sliced video data. The simulation results with systematic RLC show the usefulness of the proposed scheme for applications such as video-on-demand (VoD).

Acknowledgements

A major undertaking like PhD is a collaborative learning process shared and shaped by many people. I am extremely thankful for the support and understanding of my advisor Dr. Vladimir Stankovic. I was lucky to have Dr. Vladimir as my supervisor and appreciate his valuable suggestions, patience, and guidance through the last three years without which this work would not have been completed. Prof. Ivan Andonovic from day one has been a great motivator and has all along been very kind and considerate. I am highly indebted to Dr. Dejan Vukobratovic for the guidance I received from him during his tenure at University of Strathclyde. I acknowledge his help for providing me with the exposure and understanding of Fountain codes, and providing the code for optimizing probabilities. I would also like to thank Dr. Lina Stankovic for her support all along. Dr. Robert Atkinson provided much help related to the implementation aspects of network simulations for which I am thankful to him. I am thankful to Dr. Kristian Nybom for providing the DVB error trace files and guidance for DVB-H and DVB-T2 systems. I acknowledge the support of Dr Samuel Cheng who provided the dynamic programming solution for the optimization problem.

I would like to express my thanks for the staff and students of the CIDCOM group for all the cooperation that I had from them. My special thanks to two of my fellow PhD students, Bojana Begovic, and Hani Hassan Attar, for all the help and the things that we shared. Ms. Morag Aitken has been very helpful and considerate in addressing the various administrative problems as they arose. I am also thankful to Dr. Keith Bell, and other associates with whom I served as a TA and learnt a lot from their experience.

I cannot thank my wife enough for the support that I got at home, as she managed everything amicably. I am thankful to my children, daughter Maria, and sons Salman and Haider, who showed enormous sense of responsibility, and supported the family throughout this time of my engagement in PhD studies.

Contents

1. Introduction.....	1
1.1 Scope of the Thesis.....	5
1.2 Main Contributions.....	7
1.3 List of Publications.....	8
1.4 Outline of the Thesis.....	9
2. Background.....	11
2.1 Introduction	11
2.2 Video Compression Standards	11
2.2.1 H.264 Design.	13
2.2.2 H.264/Advanced Video Coding (AVC).....	14
2.2.3 H.264/Scalable Video Coding (SVC)	14
2.2.4 Error-resilience Features	17
2.3 Fountain Codes.....	21
2.3.1 LT Codes.....	21
2.3.2 Raptor Codes.	25
2.3.3 Random Linear Codes (RLC)	26
2.4 Windowing over Fountain Codes.....	31
2.4.1 Non-overlapping Windows (NOW).....	31
2.4.2 Expanding Windows (EW)	32
2.5 Multimedia Communication Standards.....	33
2.5.1 Digital Video Broadcasting – Handheld (DVB-H).....	33
2.5.2 DVB – Next Generation Handheld (DVB-NGH).....	36
2.5.3 Long Term Evolution- Advanced (LTE-A)	37
2.6 Multiple Description coding.....	39
2.7 Summary	40
3. DP H.264/AVC Video Broadcasting over DVB-H and DVB-T2 Networks	41
3.1 Introduction	41
3.2 Proposed Video Broadcasting Based on DP H.264/AVC.....	44
3.2.1 Video Encoding.....	45

3.2.2	AL-FEC Coding	45
3.2.3	Data Packetization.....	46
3.3	H.264/AVC Video Broadcasting with Raptor and RLC Codes	47
3.3.1	Symbol Size Selection.....	47
3.3.2	Simulated Performance of Raptor and RLC	50
3.4	DP H.264/AVC Video Broadcasting with UEP RLC	54
3.5	Performance Evaluation of EW-RLC Based Solution with DVB-T2	66
3.6	Slice Partitioned H.264/AVC Video Broadcasting with UEP RLC.....	68
3.6.1	Overall System and Measurement Setup	68
3.6.2	Results and Analysis	70
3.7	Summary	75
4.	Rate Adaptive Selective Segment Assignment (RASSA) for Reliable Wireless Video Transmission.....	76
4.1	Introduction	76
4.2	Proposed System	78
4.2.1	Protection of DP-based AVC Video	79
4.2.2	Protection of Sliced AVC Video.....	82
4.3	Rate Allocation.....	83
4.4	Results and Analysis	85
4.4.1	DP AVC Transmission	85
4.4.2	Sliced AVC Transmission.....	91
4.5	Summary	93
5.	Video Delivery over Wireless Relay Networks.....	95
5.1	Introduction	95
5.2	Multi hop Cooperative Relay Networks.....	98
5.2.1	System Description	99
5.2.2	Measurement Setup.....	100
5.2.3	Video Transmission with H.264/AVC.....	102
5.2.4	Video Transmission with H.264/SVC.....	104
5.2.5	Comparison and Analysis	107
5.3	Adaptive Layered MDC for Wireless Video.....	109
5.3.1	System Description	110

5.3.2	Results for Uniform Loss Model.....	113
5.3.3	Results for Fading Error Model	116
5.4	System Description for Relay-assisted Rateless Layered MDC	117
5.5	Resource Allocation	119
5.5.1	Global Optimization.....	121
5.5.2	Proposed Fast Solution.....	122
5.5.3	AL-FEC and Packet Scheduling	123
5.6	Simulation Setup	124
5.6.1	Video Coding and AL-FEC	124
5.6.2	Simulation Parameters	126
5.6.3	Simulation Results and Discussion	127
5.7	Summary	137
6.	Systematic RLC for Sliced H.264/AVC Video Streaming.....	138
6.1	Introduction	138
6.2	Proposed System	139
6.3	Results and Analysis	143
6.3.1	Comparison between Systematic and non-systematic Codes	143
6.3.2	Proposed Slice Prioritization Scheme	147
6.3.3	Optimized Results - Different T.....	149
6.3.4	Optimized Results – Different w.....	150
6.4	Summary	154
7.	Conclusion and Future Work	155
8.	Bibliography	159

Abbreviations and Acronyms

4G	4 th Generation
AF	Amplify-and-Forward
AL	Application Layer
AL-FEC	Application Layer Forward Error Correction
ARQ	Automatic Repeat Request
AVC	Advanced Video Coding
BER	Bit Error Rate
B-Frame	Bi-Directionally predicted Frame
BP	Belief Propagation
BPSK	Binary Phase Shift Keying
CIF	Common Interchange Format
CIP	Constrained Intra Prediction
DCT	Discrete Cosine Transform
DF	Decode-and-Forward
DP	Data Partitioning
DVB-H	Digital Video Broadcasting - Handheld
DVB-NGH	DVB - Next Generation Handheld
DVB-T	DVB - Terrestrial
DVB-T2	DVB-Second Generation Terrestrial
EEP	Equal Error Protection
ETSI	European Telecommunications Standards Institute
EW	Expanding Windows
EFW	Expanding Window Fountain Codes
EW-RLC	Expanding Window – Random Linear Codes
FEC	Forward Error Correction
FMO	Flexible Macroblock Ordering
FPS	Frame Per Second
FSMC	Finite State Markov Chain
GE	Gaussian Elimination
GF	Galois Field
GOF	Group of Frames

GOP	Group of Pictures
H.264/AVC	H.264/Advanced Video Coding
H.264/SVC	H.264/Scalable Video Coding
HARQ	Hybrid Automatic Repeat Request
HDTV	High Definition Television
HPL	High Priority Layer
HSDPA	High-Speed Packet Downlink Access
IDR	Instantaneous Decoder Refresh
I-Frame	Intra Coded Frame
IP	Internet Protocol
ITU	International Telecommunication Union
LDPC	Low Density Parity Check
LT	Luby Transform
LPL	Low Priority Layer
LTE-A	Long Term Evolution- Advanced
MB	Macroblock
MBIU	Macroblock Intra Update
MBMS	Multimedia Broadcast Multicast Service
MC	Motion Compensation
MCP	Motion Compensated Prediction
MD	Multiple Description
MDC	Multiple Description Coding
ME	Motion Estimation
MPE	Multi Protocol Encapsulation
MPEG	Moving Picture Experts Group
NAL	Network Abstraction Layer
NC	Network Coding
NOW	Non-overlapping Windows
OFDMA	Orthogonal Frequency Division Multiple Access
PET	Priority Encoded Transmission
P-Frame	Predictively Coded Frame
PLP	Physical Layer Pipes
PLR	Packet Loss Rate

PS	Probability of Selection
PSNR	Peak Signal-to-Noise Ratio
QCIF	Quarter Common Intermediate Format
QoS	Quality of Service
RASSA	Rate Adaptive Selective Segment Assignment
RCPC	Rate Compatible Punctured Convolutional Codes
RGB	Red-Green-Blue
RLC	Random Linear Codes
RNG	Random Number Generator
RS	Reed Solomon
RTP	Real Time Transport Protocol
SNR	Signal-to-Noise Ratio
SVC	Scalable Video Coding
TCP	Transmission Control Protocol
TDM	Time Division Multiplexing
UEP	Unequal Error Protection
UDP	User Datagram Protocol
VCL	Video Coding Layer
VOD	Video-on-Demand
WiMAX	Worldwide Interoperability for Microwave Access
XOR	Exclusive OR
YUV	Brightness/Luminance/Chrominance Space

Chapter 1

Introduction

The emergence of communication infrastructures such as the Internet and wireless mobile networks are the enabling technologies for the widespread use of multimedia applications. In terms of applications, these range from simple music download to watching TV over mobile devices. Some of these applications over Internet protocol (IP) are traditional, such as voice-over-IP (VoIP) to a conventional telephone, or sending television over IP to an apparently conventional set top box.

Multimedia traffic over IP based wireless networks has steadily been increasing over the recent past. With the further development of multimedia services and communication technologies, it is projected that more and more services will be interactive. In future, the convergence of traditional television broadcasting and cellular communication will take place. The radio receivers will be capable of broadcast reception in addition to the other wireless communication capabilities like Bluetooth, wireless Local area network (LAN), and cellular communication. Multimedia communications over best-effort packet networks such as the Internet is quite challenging because of the dynamic and unpredictable delay, loss rate, and available bandwidth. Video transmission therefore must adapt to the varying channel conditions.

Depending on the relationship between the sender and number of receivers, the communication can be classified as unicast, multicast or broadcast. In a unicast transmission one sender is connected to one receiver. The main advantage of such a system is that the transmission can be tailored to the receiver based on the feedback from it. In multicast, one sender serves multiple receivers. Although multicast is efficient in terms of resource utilization but it lacks the ability to target the transmission to a particular receiver. In broadcast a sender transmits to all the reachable receivers in the network.

The data transmission based on IP is in the form of packets where each packet may be independently routed. Some of the packets may be lost en-route e.g., due to buffer overflows or may be delayed beyond their display deadline for real-time data. The traditional solutions for data delivery with re-transmission of lost packets as in Transmission control protocol

(TCP) does not work well for real-time video transmission because of the tight delay constraints of each packet. As TCP is based on an acknowledgement by the receiver so it will give even poorer performance where distances between sender and receiver are large. One of the solutions is to use channel coding techniques which could recover the original data despite losses. The latest state-of-the-art solutions like those based on Reed Solomon (RS) codes are inflexible because the code rate has to be fixed in advance. Moreover, the encoding and decoding operations are quite complex especially for large Galois Field. For such codes the error characteristics of the channel must be known in advance in order to adjust the code rate to it. This solution does not extend well to multiple receivers as then only a worst-case erasure channel can be assumed for all receivers.

It is thus imperative to design Forward Error Correction (FEC) solutions which are adaptive to the varying wireless channel conditions, i.e., the bandwidth and the packet loss rates. This adaptive behaviour becomes even more important for heterogeneous receivers. Fountain codes are rateless codes which can potentially generate an unlimited number of encoded packets from a limited number of source packets. Each encoded packet is based on a combination of the source packets according to some distribution. It is thus not necessary as to which packets are received but rather that they are received in sufficient quantity. The decoding is possible if the number of received encoded packets at the receiver is just a little more than the source packets. The encoding thus eliminates the effect of independent losses at different receivers, and also there is little requirement to send feedback to the sender.

For the application of Fountain codes to video data, a Group of pictures (GOP) could be treated as a source block. The encoded packets for a particular GOP are generated based on the available bandwidth. The solutions based on Fountain codes adapt well to the varying bandwidth as the receiver or sender can terminate reception/transmission depending on the available bandwidth. Such codes are ideally suited for use in multicast scenarios because there is no requirement to target a particular receiver.

The FEC protection could be provided at different layers of the network protocol stack. However, providing the FEC solution at the application layer (AL) makes it more flexible. It can also be easily implemented in software. In addition, a video application knows best how to handle each packet and therefore it is better to leave such decisions to be taken at the

application layer. Error resilience features and FEC at the AL are thus often used to protect the data against losses.

The Third Generation Partnership Project (3GPP) recommends the use of FEC for Multimedia Broadcast and Multicast Services (MBMS) and, more specifically, adopts the use of Raptor FEC code in the AL. Digital Video Broadcasting-Handheld (DVB-H) uses Raptor codes and similar AL-FEC schemes are proposed for DVB-Next Generation Handheld (NGH) and Long Term Evolution-Advanced (LTE-A) at the AL. Due to backward compatibility with DVB-Terrestrial (T), DVB-H uses the same first generation FEC schemes with convolutional coding and Reed-Solomon (RS) block coding in addition to AL Raptor codes.

Other than Raptor codes, a class of rateless codes that is gaining increased popularity for applications in wireless broadcast/cellular networks are Random Linear Codes (RLC). With multi-hop and cooperative communications becoming popular in emerging wireless network architectures, introduction of AL RLC may serve as a step forward towards exploring the benefits of network coding. In the context of future DVB networks, this may be important for emerging concepts such as hybrid broadcast/cellular networks and device-to-device communications.

The amount of redundancy added by AL-FEC can be adapted to provide better protection where it is most needed. The whole of the encoded video data is not equally important for the video re-construction, rather different importance classes can be identified within it. The most important data for video reconstruction is termed as base layer and provides an acceptable quality. Any error in the base layer causes severe degradation. The remaining data resides in one or more enhancement layers which progressively improves the quality of video. The enhancement layer is useless by itself and it is required that the base layer is also received for it to be of some benefit. Therefore it is important to provide more protection to the base layer. Such prioritized protection of the data is termed as unequal error protection (UEP). It is thus advantageous to exploit this aspect while designing FEC solutions. In this thesis various classes of data have been identified and used in conjunction with popular fountain codes, e.g., Luby Transform (LT) codes, and RLC.

Realistic channel loss patterns consisting of real-world trace files, Gilbert model, and Finite State Markov Chain (FSMC) have been used in addition to the commonly used simple random packet erasure model. The performance of the proposed schemes has been evaluated at different loss rates and bandwidth. One of the proposed schemes, Rate adaptive selective segment assignment (RASSA) selects the data to be transmitted based on the available bandwidth.

The proposed schemes have been simulated for the mobile television broadcasting standards like DVB-H and DVB-NGH, and also for LTE-A. Proposed solutions for these standards consider Real-time transport protocol (RTP)/User datagram protocol (UDP)/IP transport as it is prescribed for DVB-H and DVB-NGH, and is proposed for LTE-A standard. However, in this thesis some technology independent solutions are also proposed which do not make an assumption of IP transport. All the proposed solutions except those relating to DVB-H are assumed to be for unicast scenarios. However, as the proposed unicast solutions are based on Fountain codes, they are also usable in multicast scenarios.

Novel algorithms have been proposed in this thesis for the protection of packetized video data with an objective to increase the overall quality at the receiver in spite of the varying channel conditions. It is by conducting simulations, and deducing analytical results that the feasibility of using the proposed schemes has been established.

1.1 Scope of Thesis

The multimedia data over IP based wireless networks has seen tremendous growth and this traffic is likely to increase as new platforms and applications evolve. It becomes ever more important to provide a good video quality perception to the users in multicast scenarios. One of the ways such adequate quality can be ensured for all the users in a wireless network is to provide adaptive solutions using FEC schemes based on Fountain codes. It is often not enough to protect the whole of the video data equally with rateless codes; rather it is also important to exploit the unequal importance of video data for re-construction.

This thesis focuses on the feasibility of using AL Fountain codes to provide adaptive solutions to the video streaming applications. In particular, Expanding Window (EW) RLC schemes have been used to provide UEP to the H.264/AVC video data which is partitioned with DP and slicing.

A class of error correction codes becoming popular recently is RLC. RLC based proposed schemes have been compared with Raptor codes for supporting video applications in DVB-H and DVB-T2 environment.

A solution based on LT codes is proposed for reliable unicast video communication by using unequal error protection (UEP) with the Data partitioning (DP) and slicing feature of H.264/Advanced Video Coding (AVC). The video data is segmented according to its importance for video reconstruction, and the segments are transmitted in the priority order matching the available bandwidth and packet loss rate. The proposed scheme is an adaptive low-complexity solution to varying channel conditions.

The transmission of H.264/AVC and H.264/SVC SNR scalable video data is compared for a relay network for different relay collaboration strategies. A multiple description coding (MDC) scheme based on slicing and DP is proposed. The application of EW RLC has been investigated with proposed MDC scheme for relay based LTE-A standard highlighting promising results for video applications.

Systematic RLC with fixed size slice partitions is considered for video streaming and a novel slice prioritization scheme is proposed.

1.2 Contributions of the Thesis

The major contributions of this thesis are:

1. Design of Application Layer Forward error correction schemes based on RLC and Raptor codes for DVB-H and DVB-T2 with DP/Slicing (Chapter 3)¹.
2. A rate-adaptive solution for bandwidth-limited wireless channels and limited resource devices (Chapter 4).
3. Analysis of Amplify-and-Forward (AF) and Decode-and-Forward (DF) schemes for layered video streaming with EW-RLC over relay networks (Chapter 5).
4. Design of a novel Multiple Description Coding (MDC) scheme for relay based communication (Chapter 5).
5. Multipath EW RLC as a robust solution to unequally protecting the data of each description (Chapter 5)².
6. Analysis of the EW approach with systematic RLC as component codes for UEP of the slice-partitioned H.264/AVC video (Chapter 6).

¹ Dr. Dejan Vukobratović helped with understanding of Fountain codes, and provided the code for optimizing probabilities.

² The dynamic programming based optimization was provided by Dr. Samuel Cheng. Hani Attar generated the error trace files based on a program for FSMC error model, provided by Dr. Lina Stankovic.

1.3 List of Publications

Conference Papers

1. S. Nazir, D. Vukobratović, and V. Stanković, “Adaptive Layered Multiple Description Coding for Wireless Video with Expanding Window Random Linear Codes”, *accepted for publication in ICASSP-2012 IEEE Intl. Conf. Acoustics Speech and Sig. Proc.*, Kyoto, Japan, Mar 2012.
2. S. Nazir, D. Vukobratović, and V. Stanković, “Expanding Window Random Linear Codes for Sliced H.264/AVC over DVB-H Network”, *in Proc. ICON-2011 IEEE Intl. Conf. Networks*, Singapore, Dec. 2011.
3. S. Nazir, D. Vukobratović, and V. Stanković, “Expanding Window Random Linear Codes for Data Partitioned H.264 Video transmission over DVB-H Network”, *in Proc. ICIP-2011 IEEE Intl. Conf. Image Proc.*, Brussels, Belgium, Sept. 2011.
4. S. Nazir, D. Vukobratović, and V. Stanković, “Unequal Error Protection for Data Partitioned H.264/AVC Video Streaming With Raptor and Random Linear Codes for DVB-H Networks”, *in Proc. ICME-2011 IEEE Intl. Conf. Multimedia and Expo*, Barcelona, Spain, July 2011.
5. S. Nazir, D. Vukobratović, and V. Stanković, “Performance Evaluation of Raptor and Random Linear Codes for H.264/AVC Video Transmission over DVB-H Networks”, *in Proc. ICASSP-2011 IEEE Intl. Conf. Acoustics Speech and Sig. Proc.*, Prague, Czech Republic, May 2011.

1.4 Outline of the Thesis

The thesis is organized in seven chapters including this one. The outline of the remaining chapters is as follows:

Chapter 2 covers the relevant topics for the multimedia transmission over wireless IP networks which are required for understanding the later chapters. The coverage of latest video coding standards along with the error-resilience features is included. A brief description of Fountain codes is provided, including LT codes, Raptor and Random linear codes. Multiple description coding along with its advantages to exploit path diversity are covered.

Chapter 3 compares the Raptor and RLC codes for their suitability as Application-layer forward error correction codes for DVB-H standard. The best of the selected code configurations are then used to extend the comparison to EEP and UEP schemes. The investigation of UEP schemes encompasses both the non-overlapping windows and expanding window random linear codes.

Chapter 4 covers the novel algorithm proposed for the rate adaptive selective segment assignment (RASSA). The scheme is based on LT codes and is suitable for any application requiring adaptive transmission. The optimized results are presented for different channel conditions.

Chapter 5 covers the cooperative relay based multi hop wireless network to symbolize the emerging 4G communication standard, LTE-A. The transmission of layered video data for H.264/AVC and H.264/SVC is compared over relay networks.

A multiple description coding (MDC) scheme based on the slicing and DP is proposed. Expanding window-random linear codes are proposed to simulate the transmission of the layered descriptions for LTE-A standard.

Chapter 6 proposes a method that prioritizes the slices based on their PSNR contribution to reconstruction as well as temporal significance. Systematic EW-RLC are proposed for protection of slice-partitioned H.264/AVC video data which provides a

prioritized data transmission. Simulation results demonstrate usefulness of using relative slice priority with systematic codes for multimedia applications.

Chapter 7 contains the conclusion and future work directions.

Chapter 2

Background

2.1 Introduction

This chapter covers the required background for the later chapters. Section 2.2 covers the video compression standards. Fountain codes are described in Section 2.3. Section 2.4 covers windowing over Fountain codes. Multimedia communication standards, that is, DVB-H, DVB-NGH and LTE-A are covered in Section 2.5. MDC is explained in Section 2.6. Finally, Section 2.7 provides a summary.

2.2 Video Compression Standards

Video is the most important media for communications. The video communication is now mostly over best-effort packet networks. Video over best-effort packet networks is complicated by a number of factors including unknown and time-varying bandwidth, delay, and losses [1]. Video communication is unimaginable without video compression due to the large data size.

Video compression is accomplished by removing the redundancies present in the source video data. In a video sequence consecutive frames have temporal redundancy, that is, frames captured together in time will have the same objects and background displaced due to motion. Also, there is spatial redundancy present in each frame, which means that the pixels which are closer together will be often correlated. In case of colour video the Red, Green, and Blue components of a pixel can be highly correlated.

In the case of a colour image, a colour space conversion is first applied to convert the Red-Green-Blue (RGB) image into a YUV (brightness/luminance/chrominance space) where these components can be assigned different weights depending on the human visual perception, which is strongest for brightness. In order to remove temporal redundancy, the similarity between the frames which are closer together in time can be exploited. Instead of coding each frame in isolation, the similarity between frames is exploited by, first predicting it based on a previously coded frame, and then coding the difference in this prediction. In

order to reduce the difference between frames, the motion is estimated by a process termed motion estimation (ME). Spatial redundancy is exploited by employing signal transformation, such as, Discrete Cosine Transform (DCT) which works on adjoining pixels and removes the redundancy.

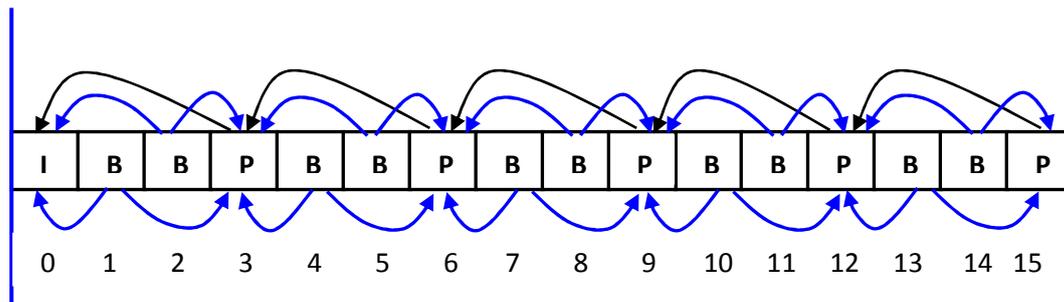


Figure 2.1: GOP structure of 16 frames.

A frame coded independently of other frames is termed as intra-frame or I-frame. Such frames typically exist at the start of video sequence or a group-of-pictures (GOP) and generally have larger sizes. A frame based on a prediction to a previous frame only is termed as a predictively-coded or P-frame. The prediction however could be done based on a previous and future frame as well, which gives even better compression, such frames are termed as bi-directionally predicted frames or B-frames. The different coded frames and their dependencies are shown in Figure 2.1. The selection of prediction dependencies between frames can have a significant effect on video streaming performance, e.g. in terms of compression efficiency and error resilience.

The video compression is normally lossy, which means that after compression of the video data, its reconstruction back to the original may not be exact. However, such loss is acceptable and its nature may be governed by a given application. The latest video compression standards achieve compression by applying the same basic principles as described above. The colour space redundancy is exploited by a colour space conversion. Then the temporal redundancy is removed by applying Motion Compensation (MC)-prediction, and the spatial redundancy is removed by applying the DCT. The resulting DCT coefficients are subsequently processed to generate the compressed bit stream. However, as

this compressed bit stream is highly susceptible to quality degradation even by bit losses hence the modern video coding standards employ quite advanced techniques to contain the adverse effect of such losses.

The latest video compression standard is H.264/AVC which works similarly to older standards, such as Moving Picture Experts Group (MPEG)-1 and MPEG-2 but adds many additional features to decrease data rate while maintaining quality. H.264/AVC has a scalable extension known as H.264/Scalable Video Coding (SVC).

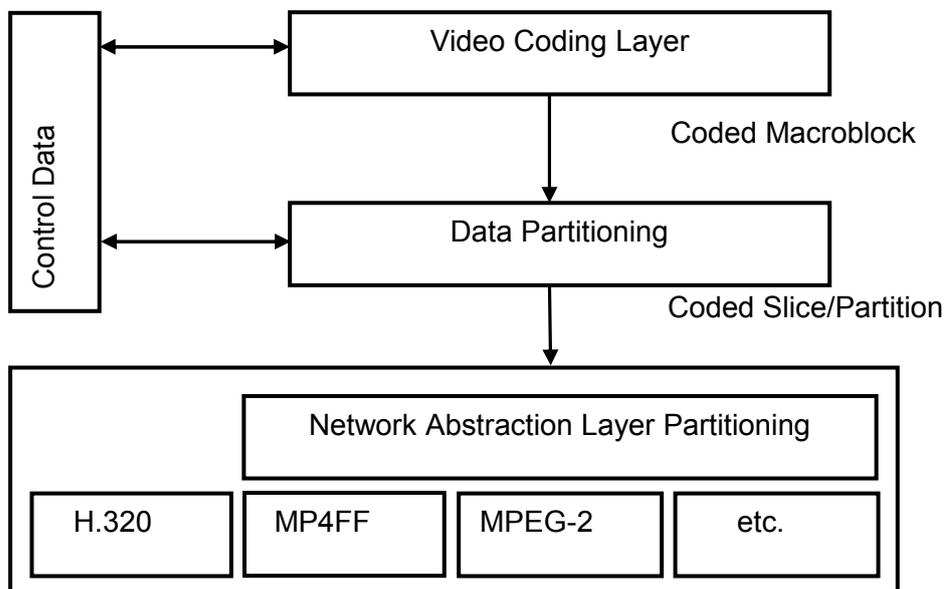


Figure 2.2: Structure of H.264/AVC encoder [2].

2.2.1 H.264 Design

The coded output bitstream of H.264 has two layers as shown in Figure 2.2, Video Coding Layer (VCL) and Network Abstraction Layer (NAL). In order to be suitable for a range of video applications, VCL is used to efficiently represent the video content, and NAL is used to represent VCL data for conveyance by different transport layers.

The VCL follows the traditional block-based hybrid video coding approach. The design is very similar to that of prior video coding standards, but H.264/AVC includes new features that enable it to achieve a significant improvement in compression efficiency relative to any prior video coding standard [3].

The coded video data is organized into NAL units, each of which is a packet which contains an integer number of bytes. NAL units could be VCL and non-VCL NAL units. The VCL NAL units contain the actual video data, and the non-VCL NAL units contain other information such as parameter sets and supplemental enhancement information which is important but not required for video decoding.

2.2.2 H.264/AVC

H.264/AVC [2] is the latest state of the art video compression standard. Similar to the prior video coding standards, it is based on Motion Compensated Prediction (MCP), which requires maintaining synchronisation between the encoding and decoding operations. This synchronisation may be lost due to packet loss, and the errors could then propagate to areas which may have been received correctly. Since MacroBlocks (MBs) are spatially and/or temporally dependent on neighbouring MBs, the errors can also propagate in time (in following frames) and in space (the same frame). This error propagation may continue until the next intra-coded frame. More intra-coded frames may restrict error propagation but are associated with loss in compression efficiency.

The main goals of the H.264/AVC standardization effort have been enhanced compression performance and provision of a “network-friendly” video representation addressing “conversational” (video telephony) and “non conversational” (storage, broadcast, or streaming) applications. H.264/AVC has achieved a significant improvement in rate-distortion efficiency relative to existing standards [3]. H.264/AVC has been adopted by various application standards and is increasingly used in most video applications.

2.2.3 H.264/SVC

Modern video transmission is typically characterized by a wide range of connection qualities and receiving devices. SVC is a highly attractive solution [4] to the problems posed

by the characteristics of modern video transmission systems. Scalability refers to removal of parts of the video bit stream in order to adapt it to the various needs or preferences of end users as well as to varying terminal capabilities or network conditions.

SVC is an extension of H.264/AVC standard and adds scalability features to AVC standard. A non-scalable video encoder generates a single compressed bit stream. A scalable video encoder compresses a raw video sequence into multiple layers. One of the compressed layers is the base layer. The base layer can be independently decoded and can provide a relatively low level of video quality. Additional compressed layers are enhancement layers that provide additional quality to the received video stream. Enhancement layers can be decoded only in conjunction with the base layer. On a Quality of Service (QoS)-enabled IP network it would even be possible to send the base layer with a higher priority than the other layers.

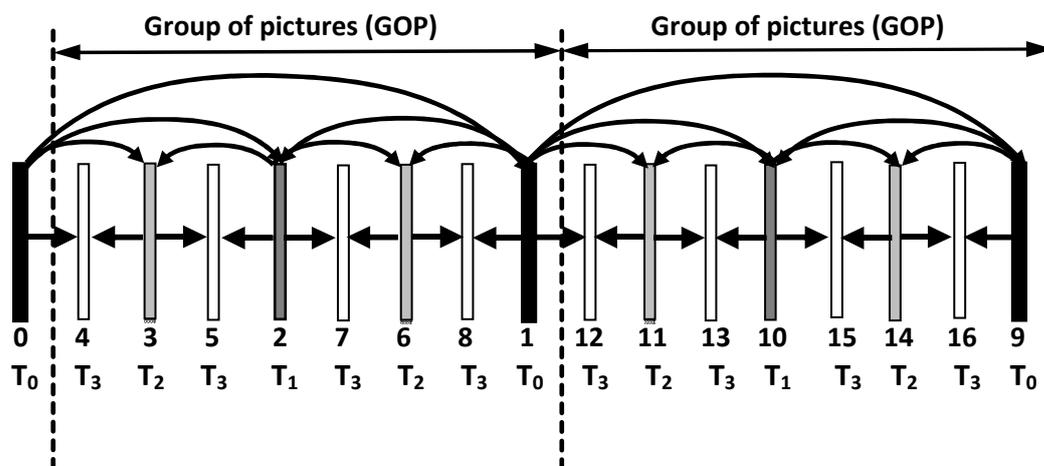


Figure 2.3: Hierarchical prediction structures for enabling temporal scalability [4].

The networks which support prioritization can make use of scalability by assigning a higher priority to the layers according to their importance. The base layer could thus be transported with the highest priority. However, Internet does not provide any such prioritization and all packets are equally likely to be lost. In such networks, the scalability

alone does not bring any advantage for video transport; however, channel coding can be used to make the base layer more tolerant to errors [6]. The layers can be combined to adapt to different frame rates, spatial resolutions, or quality, of the video content, giving rise to temporal, spatial, and quality scalability.

2.2.3.1 Temporal scalability

A bit stream provides temporal scalability when the video sequence can be partitioned into a temporal base layer and one or more temporal enhancement layers. For hybrid video codecs, temporal scalability can generally be enabled by restricting motion-compensated prediction to reference pictures with a temporal layer identifier that is less than or equal to the temporal layer identifier of the picture to be predicted, as shown in Figure 2.3.

2.2.3.2 Spatial Scalability

In each spatial layer, motion-compensated prediction and intra-prediction are employed as for single-layer coding. But in order to improve coding efficiency in comparison to simulcasting different spatial resolutions, additional so-called *inter-layer prediction* mechanisms are incorporated as illustrated in Figure 2.4.

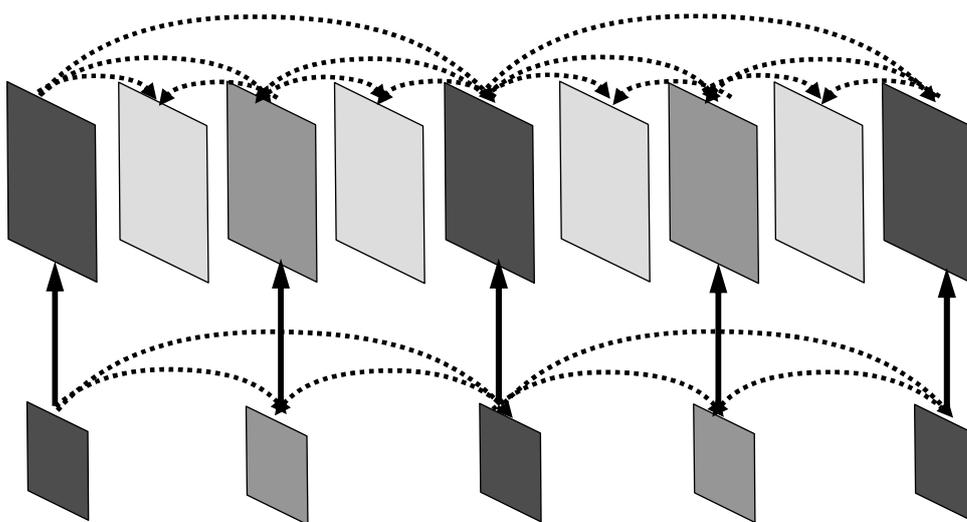


Figure 2.4: Multilayer structure with additional inter-layer prediction for enabling spatial scalable coding [4].

2.2.3.3 Quality Scalability

Quality scalability of H.264/SVC has a coarse-grain quality scalable coding (CGS) feature which can be considered as a special case of spatial scalability with identical picture sizes for base and enhancement layer. The CGS concept only allows for a few selected bit rates to be supported in a scalable bit stream. In general, the number of supported rate points is identical to the number of layers. When the relative rate difference between successive CGS layers gets smaller, the CGS concept becomes less efficient.

In a variation of the CGS approach, which is referred to as medium-grain quality scalability (MGS), any enhancement layer NAL unit can be discarded from a quality scalable bit stream.

2.2.3.4 Benefits in terms of applications

In case of a video transmission service with heterogeneous clients, where multiple bit streams of the same source content differing in coded picture size, frame rate, and bit rate are to be supported, the source content has to be encoded only once—for the highest required resolution and bit rate, resulting in a scalable bit stream. This stream can be used to obtain representations with lower resolution and/or quality by discarding selected data.

A scalable bit stream usually contains parts with different importance in terms of decoded video quality. This is especially useful in conjunction with unequal error protection for any transmission scenario with unpredictable throughput variations. The more important information can be protected by providing higher degree of protection errors.

2.2.4 Error Resilience Features

H.264 is aimed for packet-based networks and thus copes mainly with packet losses instead of bit errors. Thus, it is assumed that packets that contain bit errors are discarded and not fed into the decoder [2].

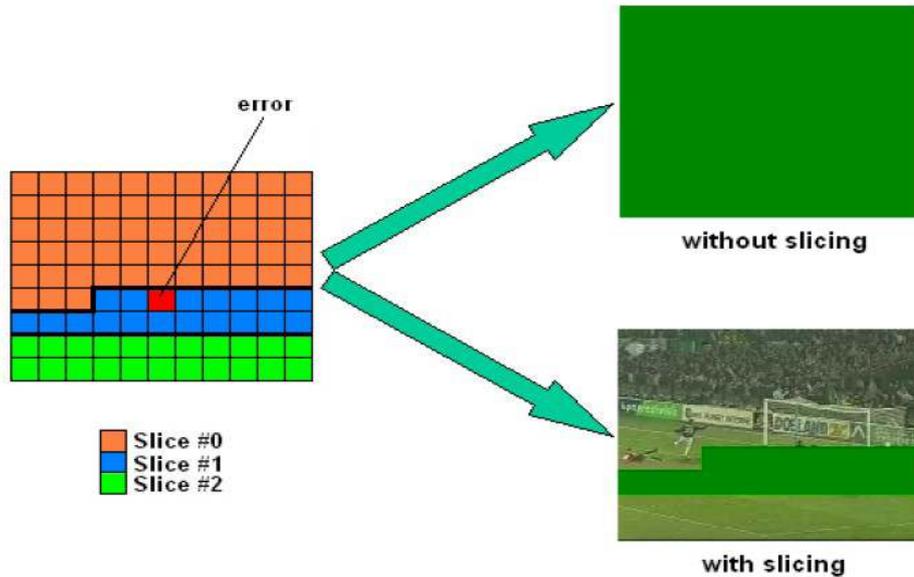


Figure 2.5: Error detection without and with slicing [5].

2.2.4.1 Slice- structured Coding

It is not suitable to transmit all the compressed data belonging to a complete coded frame in a single data packet. Each picture may be split into one or several slices as shown in Figure 2.5. The macroblocks will be organized in slices and this has the desirable effect of making transmitted packets smaller; making the overall recovered video frames better tolerant of errors. Slice structured coding also introduces slice headers to act as resynchronisation points to localize the errors and prevent error propagation to subsequent frames. Each slice can be correctly decoded without the use of data from other slices provided in the same frame. Each slice is encapsulated in a separate packet by H264/AVC encoder. Slices are self-contained, thus, if a coded slice is available to the decoder, all the MBs in that slice can be decoded. However, slices interrupt the in-picture prediction mechanisms; thus, the use of small slices reduces compression efficiency.

The advantage using slicing can be seen in Figure 2.5, when an error occurs, instead of concealing the whole frame, it is just the slice which is to be concealed.

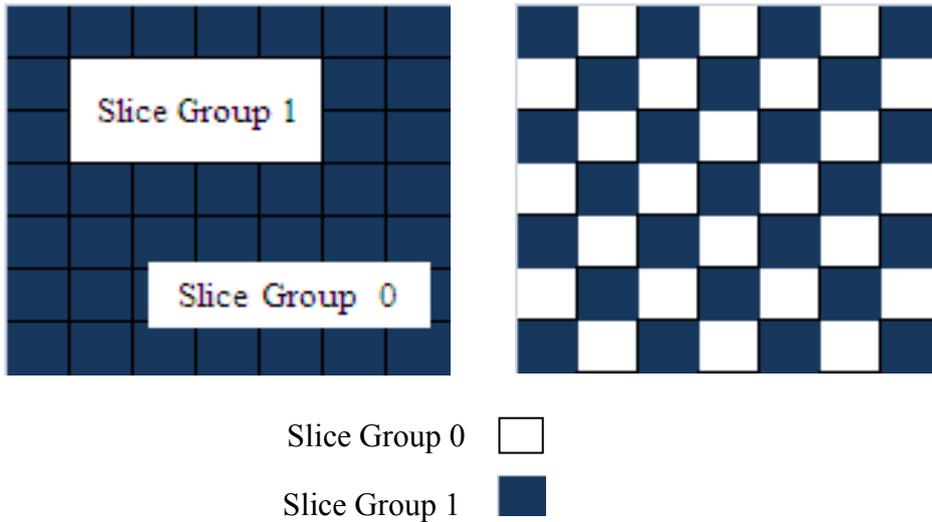


Figure 2.6: FMO with checkerboard pattern - MBs allocation to two slice groups.

2.2.4.2 Flexible Macroblock Ordering (FMO)

Another error-resilience technique exploiting partitioning of information is Flexible MacroBlock Ordering (FMO). Using FMO, spatially collocated image areas can be interleaved in different slice groups. FMO also allows for coding of a region of interest to improve coding loss. There are several pre-defined patterns, and one of them is the checkerboard pattern, shown in Figure 2.6.

In the FMO checkerboard pattern, if one slice group is lost and the other is correctly received, the MBs in the lost slice groups will have several neighbouring blocks that have been received correctly. Thus, error concealment will be easier.

2.2.4.3 Data Partitioning

DP is a feature available in the extended profile which supports the partitioning of a frame/slice in up to three partitions, based on the importance of the encoded video syntax elements for video reconstruction.

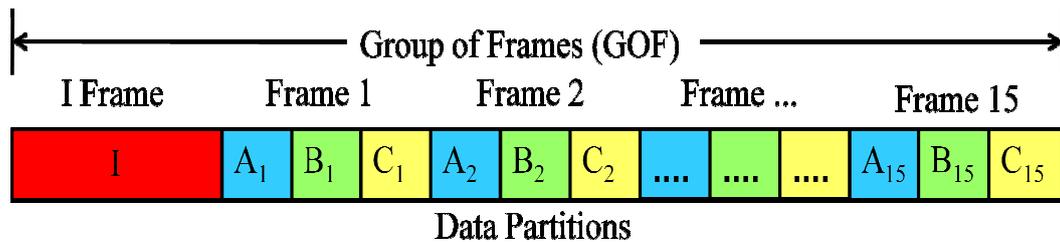


Figure 2.7: Relative position of DP A, B and C in a GOF of 16 frames.

A GOF with each frame split in three partitions is shown in Figure 2.7. DP A contains the most important data comprising slice header, quantization parameters, and motion vectors. DP B contains the intra-coded macroblocks (MB) residual data, and DP C contains inter-coded MB residual data. The decoding of DP A is always independent of DP B and C. At the decoder, if the Type B or Type C partition is missing, the header information in the Type A partition may be used in order to aid in error concealment.

However, if DP A is lost the remaining partitions cannot be utilized. The decoding of DP B is possible without DP C, but not other way around. To make DP B independent of DP C, Constrained Intra Prediction (CIP) parameter in H.264/AVC encoder must be set.

2.2.4.4 Redundant Slices

Redundant Slices allow for the encoding of one or more redundant representations of a slice, in addition to the original slice. The redundant slices are encoded using a coarser quantization parameter as compared to the original slice. Thus, the redundant slices will usually utilize fewer bits than the original representation. If the primary slice is available to the decoder, it will be used to reconstruct the macroblocks and the Redundant Slice will be discarded.

2.3 Fountain Codes

Fountain codes are erasure correction code, capable of correcting lost data. This ability to recover the lost data without requiring retransmission provides reliability and efficiency in data networks. These codes have transformed the conventional paradigm of ordered delivery for transmitting the packetized data over erasure channels.

With fountain codes approach a user can receive any packets (a little more than the source data) in any order to be able to decode the given data. In case of transmission to large number of users, it results in a much simplified data delivery [7].

Fountain codes are rateless codes, that is, given k packets of a file, potentially infinite encoded packets can be generated. If any $k(1 + \alpha)$ packets irrespective of the order are received, the receiver can reconstruct the original data. α is a non-negative small fraction less than 1, and is very small for large k . Fountain codes are application layer- forward error correction codes generally implemented in software, to protect packet losses.

2.3.1 LT Codes

Luby transform (LT) codes are the first class of fountain codes [8]. It works on a very simple algorithm based on exclusive OR (XOR) operation to encode and decode the message. LT codes are rateless which means that the encoder can generate an infinite number of encoded symbols. They are erasure correcting codes because they can be used to transmit digital data reliably over an erasure channel.

The LT encoder with a set of k source symbols can generate a potentially infinite sequence of encoded symbols. Each encoded symbol is computed independently of the other encoded symbols. Given a degree distribution d and source symbols k , an encoded symbol can be generated as per the following process:

- (1) Randomly select degree d of the encoding symbol from a degree distribution.
- (2) Select uniformly at random distinct input symbols equal to degree d .

- (3) The value of the encoding symbol is the exclusive-OR of the d neighbours.

By following the steps enumerated above, the encoder can generate potentially infinite encoded symbols. These encoded symbols can then be injected into the channel. Some of the symbols will be lost and assume that the symbols correctly received at the decoder is little more than k . The decoder needs to know the degree of each symbol and the connected symbols (exclusive-OR) before it can attempt to recover the original source symbols. There could be a variety of ways to pass this information to the decoder and a simple one is to pass a random number generator (RNG) seed along with each encoded symbol. The decoder follows the following process to recover the source symbols:

- (1) Select a symbol with degree 1. This is connected to only one symbol, that is, one symbol has been recovered.
- (2) Find other encoded symbols which contain this (decoded) symbol, exclusive – OR this symbol with those symbols, and reduce their degree by 1.
- (3) Repeat step (1) and (2) until all the symbols are recovered.

If there is no symbol in step (1) with degree 1 at the start of decoding or before all the symbols have been recovered, the decoding cannot succeed. In this case it would be necessary, if possible to receive more symbols. The step (2) is likely to continue generating symbols with degree 1.

The degree distribution used in the encoding process is a critical part of the design. Some of the encoded symbols should have enough high degrees to avoid the case where a source symbol is not connected to any encoded symbol at all. On the other hand, some of the encoded symbols should have lower degrees to start the decoding process. During the decoding process, ideally, at any time, there should be only one encoded symbol with degree one and after processing this degree-one encoded symbol, the degrees of the remaining encoded symbols are reduced in such a way that only one encoded symbol has degree one. In [8], the ideal Soliton distribution is specified. Also, a refinement over it termed as Robust Soliton distribution which provides better results is proposed.

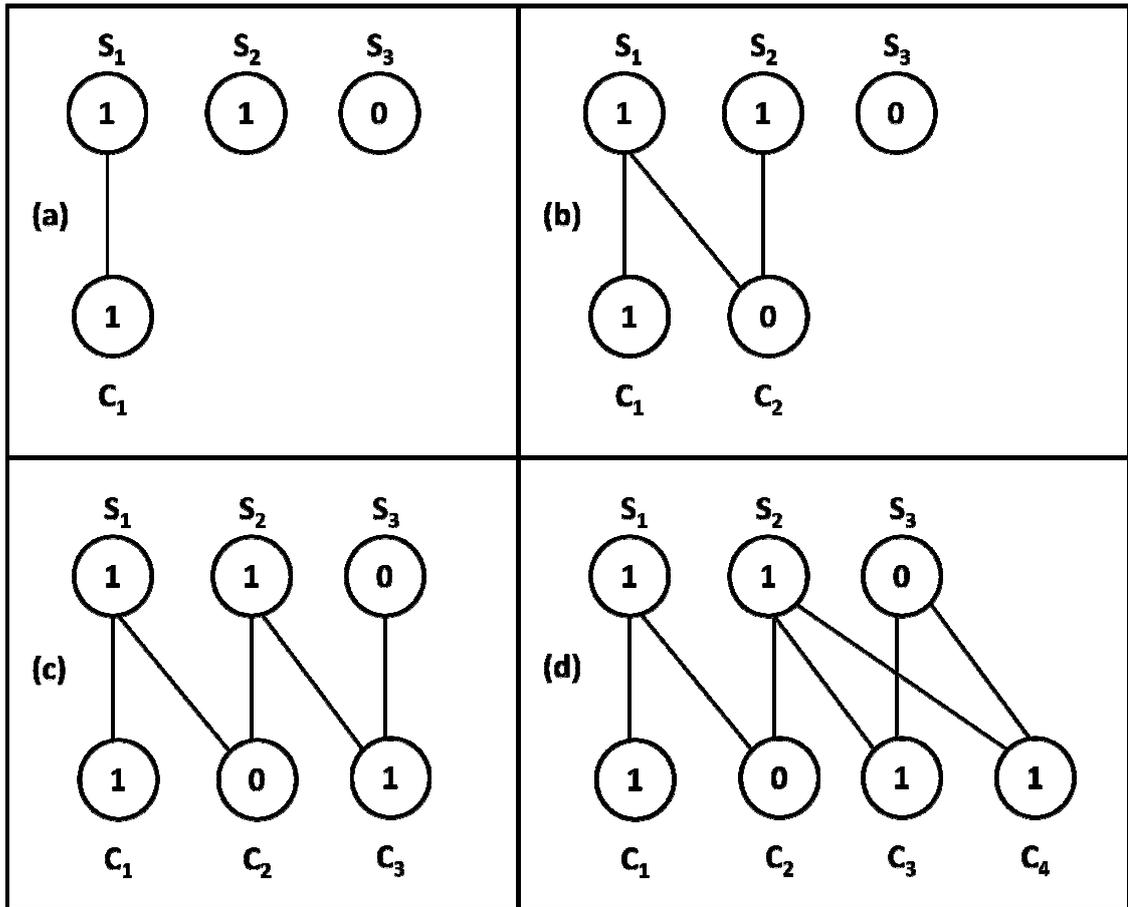


Figure 2.8: Encoding process for LT codes.

The encoding and decoding process for LT codes can be understood with the help of an example as given below:

Encoding

Assume an LT coder with the symbol size as 1 bit (for simplification). $S_1 = 1$, $S_2 = 1$, and $S_3 = 0$. The generation of an encoded symbol is depicted in each block of Figure 2.8, for the four symbols, C_i . The symbol \oplus represents Exclusive-OR operation.

Figure 2.8 (a) shows degree $d_1 = 1$, $C_1 = S_1$. (b) $d_2 = 2$, $C_2 = S_1 \oplus S_2$. (c) $d_3 = 2$, $C_3 = S_2 \oplus S_3$. (d) $d_3 = 2$, $C_3 = S_2 \oplus S_3$.

Assume further that during transmission symbol C_3 is lost. Hence, at the receiver the task is to recreate the original source symbols, given the three received encoded symbols.

Decoding

The first step would be to recreate the degree and the selected symbols for each encoded symbol. The starting position at the decoder is shown in Figure 2.9 (a) and each subsequent block shows decoding of one symbol.

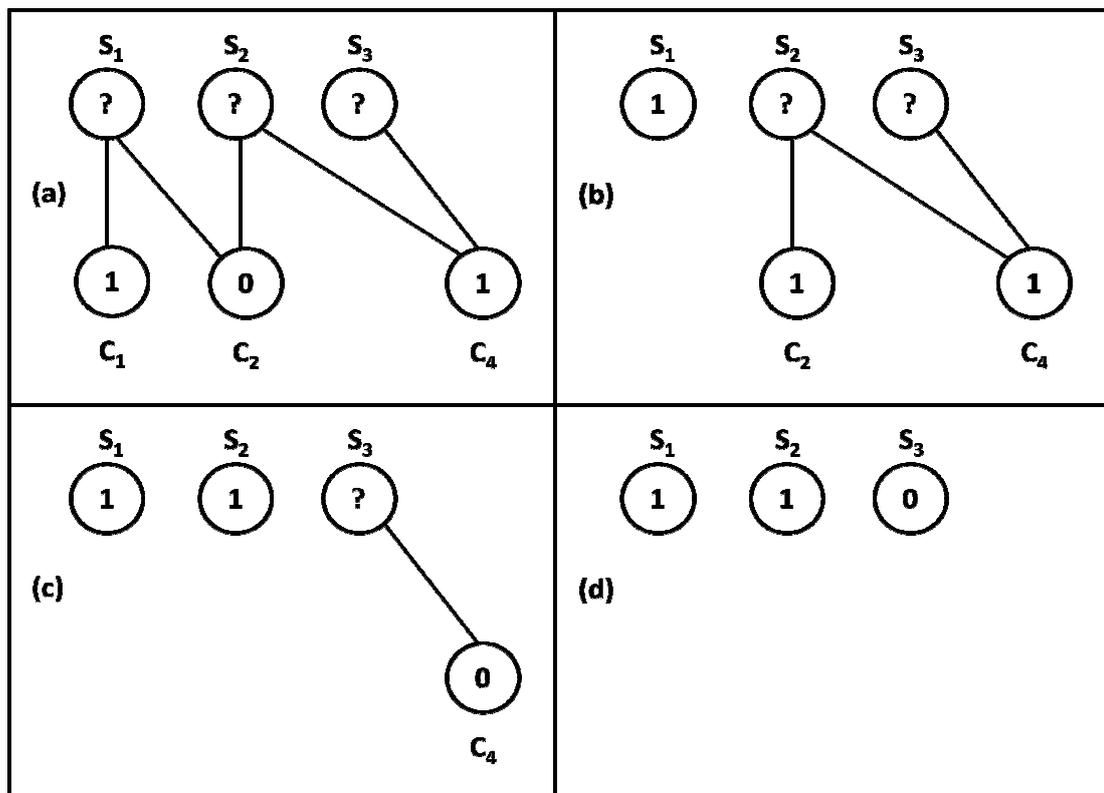


Figure 2.9: Decoding process for LT codes.

The decoding process starts at an encoded symbol with degree = 1, which is C_1 . (b) $S_1 = C_1 = 1$. S_1 is combined with connected symbol, $C_2 = C_2 \oplus S_2 = 1$ (c) degree $d = 1$ for C_2 . $S_2 = C_2 = 1$, $C_4 = C_4 \oplus S_2 = 0$ (d) degree $d = 1$ for C_4 . $S_3 = C_4 = 1$. The decoding process stops having recovered all symbols.

2.3.2 Raptor Codes

Raptor codes as initially described in [9] extend the idea of LT codes one important step further. Raptor codes use two encoding stages for encoding, consisting of a pre-code followed by reduced-complexity LT coding. Raptor codes achieve linear time encoding and decoding by pre-coding of the input symbols before the application of LT code.

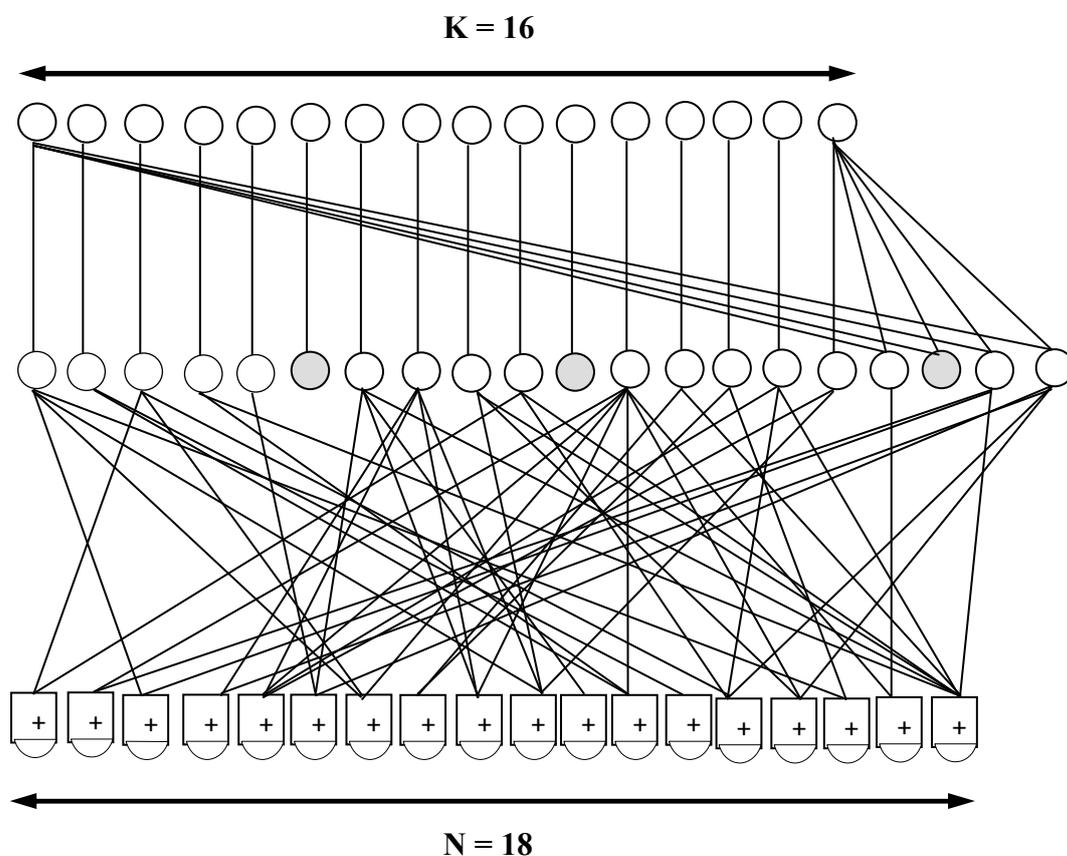


Figure 2.10: Diagram of a Raptor code [10].

An example [10] is shown in Figure 2.10. The source packets, $k = 16$, are shown to be encoded by an outer code (LDPC or such code) into 20 pre-coded packets. These pre-coded packets are encoded into 18 received packets with a weakened LT code. The average degree for LT code is 3. The weakened LT code fails to connect some of the pre-coded packets to any received packet. The lost packets are shown highlighted. The LT code however recovers

the other 17 pre-coded packets, and then the outer code is used to deduce the original 16 source packets.

Raptor codes [9] are rateless codes, i.e., they provide a flexibility to generate as many encoded symbols as desired from the source symbols. The Raptor decoder can recover the original source symbols from any set of encoded symbols, as long as their number is at least equal or slightly exceeds the number of source symbols.

These have recently been adopted for use in various standards including DVB-H. As an AL-FEC solution in DVB-H, systematic Raptor codes provide improved system reliability and a large degree of freedom in the choice of transmission parameters [11].

The paper by Shokrollahi [9] has many additional details, including useful descriptions of extremely effective practical constructions and analysis techniques for codes of finite length. Raptor codes currently give the best approximation to a digital fountain. A virtually limitless supply of packets can be generated on the fly after some small initial pre-processing, with each packet taking only constant time to produce. Decoding can be accomplished after receiving just a few percent more than the minimum of k encoding packets (with high probability), and requires space and time linear in the size of the original message. Moreover, very efficient implementations are possible.

For an explanation of the encoding and decoding algorithms and implementation guidelines, [12] is a detailed reference. The decoding process results in maximum-likelihood decoding performance and it can be considered successful if the received generator matrix at the decoder is invertible.

2.3.3 Random Linear Codes (RLC)

A class of rateless codes which has become popular recently are RLC [13]. RLC applied over a source message produces encoded symbols as random linear combinations of source symbols with coefficients randomly selected from a given finite field.

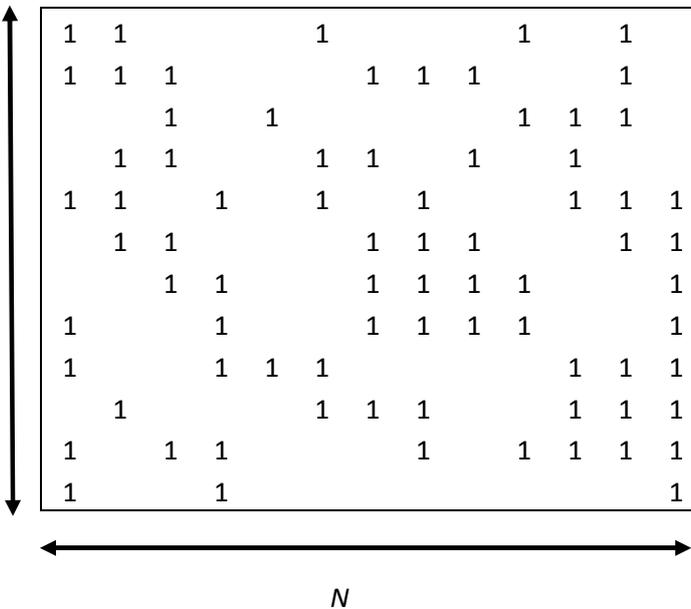
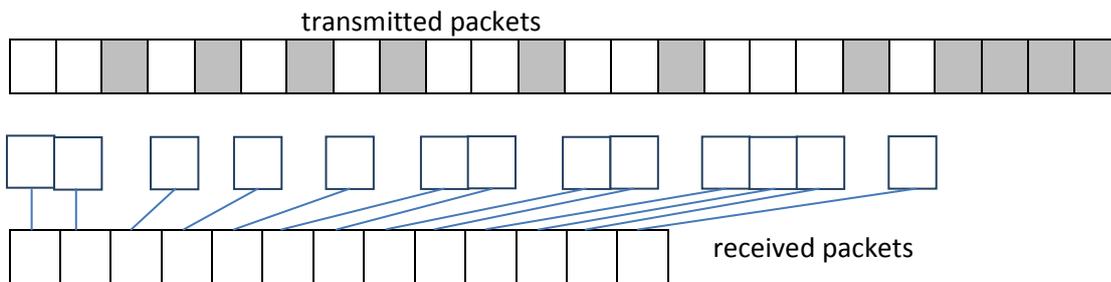
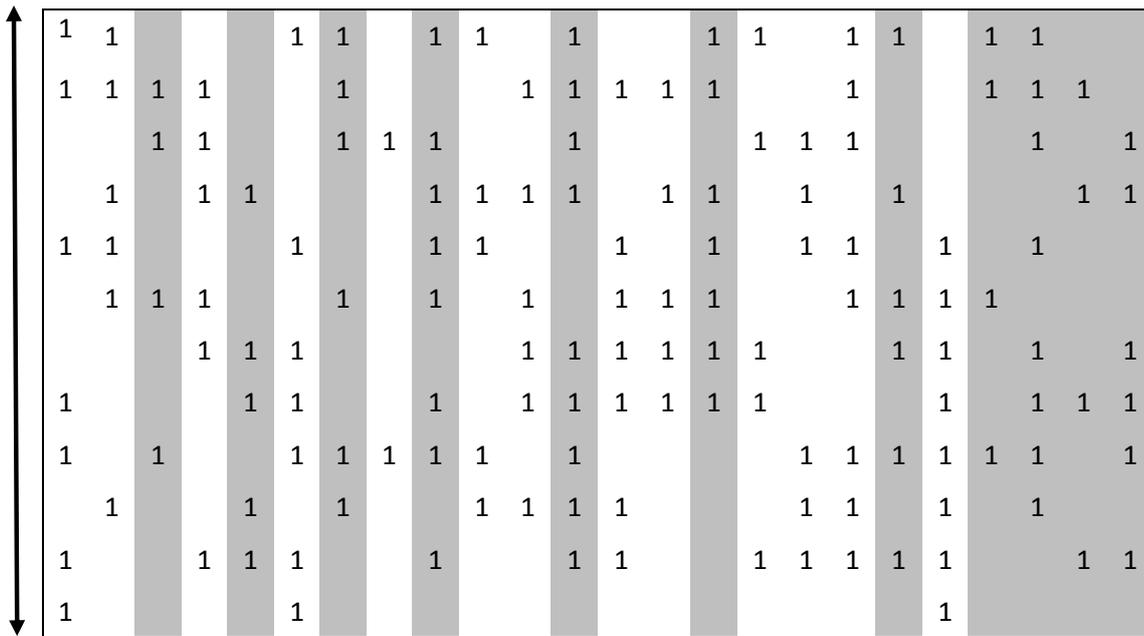


Figure 2.11: The generator matrix for random linear code. The bottom matrix after losses could be created at the decoder [10].

As a packet-level AL FEC solution, RLC is simple to implement and perform as near-optimal erasure codes [12][14] for sufficiently large finite field used for creating linear combinations of source symbols (one-byte Galois field GF(256) is usually sufficiently good). This makes RLC an attractive alternative to Raptor codes as a universal FEC/network coding solution for emerging wireless communication systems, such as LTE-A, Worldwide Interoperability for Microwave Access (WiMAX), and DVB-NGH [15][16][17][18][19].

If there are k source packets to be transmitted, then for each encoded packet, the source packets to be exclusive-OR are determined. That is each encoded packet is an exclusive-OR (in the selected field) of the selected source packets. This process could be likened to a generator matrix as shown in Figure 2.11. Each column of this generator matrix shows the source symbols which must be combined to yield an encoded symbol. The encoded symbols transmitted will be more than k depending on the channel characteristics.

Some of these encoded packets will be lost in the channel. The decoder receives the remaining packets which must be at least equal to k for the decoding to succeed. The decoder must be able to create the generator matrix in order to proceed with the decoding. Several methods could be used for this.

If the generator matrix at the decoder is invertible then the decoder can recover the original k source symbols. By performing Gaussian Elimination (GE) the inverse of the matrix can be computed.

Next the encoding and decoding process for RLC will be demonstrated using a simple example:

Encoding Operation

Consider a generator matrix, G_{nk} as shown below:

$$G_{nk} = \begin{bmatrix} 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 \end{bmatrix}$$

Assume the four source symbols (in GF(2)) to be given as: $S_1 = 1011$, $S_2 = 0011$, $S_3 = 1010$ and $S_4 = 0110$. The generation of encoded symbols can be understood as a matrix multiplication process in the GF. The first encoded symbol is obtained by multiplying each of the four symbols with the first column of the generator matrix. The symbol \oplus represents Exclusive-OR operation, and $(.)$ represents multiplication in GF.

$$C_1 = 1011.1 \oplus 0011.1 \oplus 1010.0 \oplus 0110.0$$

$$C_1 = 1011 \oplus 0011 = 1000$$

Through the same process, $C_2 = 0011$, $C_3 = 0111$, $C_4 = 1100$, and $C_5 = 1101$.

The coefficients of the generator matrix used to generate an encoded symbol are carried along with each symbol. Assume further that during transmission symbol C_3 is lost. The task at the receiver would be to recover the original source symbols from the encoded symbols and the associated information about the generator matrix.

Decoding Operation

The encoded symbols available to the decoder (with loss of C_3) would be, $C_1 = 1000$, $C_2 = 0011$, $C_4 = 1100$, and $C_5 = 1101$.

The first step would be to recreate the generator matrix from the received encoded symbols. Based on the four encoded symbols received the generator matrix at the receiver would be as:

$$G_{nk} = \begin{bmatrix} 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}$$

This matrix is inverted before the actual source symbols could be extracted. The Gaussian Elimination (GE) process is applied for matrix inversion. If the matrix inversion fails then the source symbols cannot be recovered. The inverted generator matrix (in GF(2)) is as shown below:

$$G_{nk}^{-1} = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}$$

The first source symbol is extracted by multiplying each of the four symbols (C_1 , C_2 , C_4 , and C_5 , with the first column of the generator matrix.

$$S_1 = 1011.1 \oplus 0011.1 \oplus 1010.0 \oplus 0110.0$$

$$S_1 = 1011 \oplus 0011 = 1000$$

Similarly, the other source symbols can be recovered.

The major limitation of the application of RLC is the decoding complexity of GE decoding, which is polynomial in the number of symbols. However, for short lengths of the source messages, the decoding complexity is acceptable (see [20] and references therein). Moreover, there is no performance penalty for using short-length codes as compared to standard rateless codes.

The short length RLC codes are meant to reduce the additional delay due to coding. However, it is assumed that any packet which arrives at the receiver after its display deadline will be considered as being lost. This situation could be improved somewhat by employing a buffer at the receiver for the incoming packets. If the additional delay due to coding is more than tolerable then it would result in more packets being lost which will degrade the video reconstruction quality.

A systematic code is any error-correcting code in which the input data is embedded in the encoded symbol. The advantage of such codes is that the receiver does not need to recover the original source symbol in case of correct reception.

When erasure rates are low, it is effective to use systematic RLC which further reduces the encoding and decoding complexity. In case of RLC the source symbol itself will be transmitted that simplifies the encoder operation. Also, systematic codes reduce the decoding complexity since with systematic codes the decoder operates with the matrix that has reduced number of rows (reduced, by the number of correctly received systematic packets). However, in general systematic codes do not provide improvement compared to non-systematic codes.

2.4 Windowing over Fountain Codes

The elements of compressed video data do not have same contribution for the video re-construction quality. It could thus be beneficial to design effective error protection schemes which exploit this inherent prioritization of video data for transmission. A scheme which treats all the data to be transmitted with Equal Error Protection (EEP) does not take into account the different priorities of video data into account for providing error protection. On the other hand, protection schemes can be designed which assign a degree of protection to a class of video data in accordance with its importance. Such a protection scheme is termed as UEP. In order to define a class/window over a subset of source symbols, all the high-priority data units belonging to a particular priority have to be aggregated together over the whole source data. The classification of video data could be based on DP, slicing, type of frames, or any other criteria.

The rateless codes can be used to provide UEP. The basic idea is to design the channel codes so that the more important data for video re-construction is better protected. Generally, two or three windows are deemed sufficient.

2.4.1 Non-overlapping Windows (NOW)

A simple UEP scheme could be designed with dividing the source data into non-overlapping sets, termed “windows”, which represent different classification of video data in accordance with its importance for video re-construction.

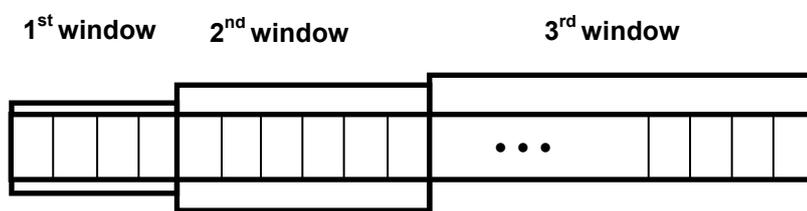


Figure 2.12: Non-overlapping window structure.

A general layout of the window structure with three priority classes is as shown in Figure 2.12. The most important data will reside in window W1, and the subsequent windows will hold data with a progressively decreasing priority. In order to generate channel symbols,

first a window has to be selected whose source symbols will be used. Therefore, in order to provide UEP all that is necessary is to assign a Probability of Selection (PS) to each of the windows. The window with most important data could be assigned a higher PS relative to the other windows. After a window has been selected then the coding procedure is same as described in Section 2.3 for the various codes.

With this layout of windows, the source symbols from each window are encoded and decoded independently of the other window. In order to decode the whole of video data in all windows, all the decoders must succeed. It is possible to alter the window structure such that the low-priority windows also aid the decoding process of the high-priority windows. Such a window structure is termed as Expanding Window (EW).

2.4.2 Expanding Windows

Expanding window fountain (EWF) codes [21] are a class of UEP fountain codes based on the idea of creating a set of “nested windows” over the source block. The rateless encoding process is then adapted to use this windowing information while producing encoded packets. To obtain source blocks amenable to UEP, the set of windows is defined over the groups of source symbols of unequal importance. The coding is then performed over progressively increasing source block subset windows aligned with this “most to least importance” subsets. The general layout of a window structure with three importance classes is shown in Figure 2.13.

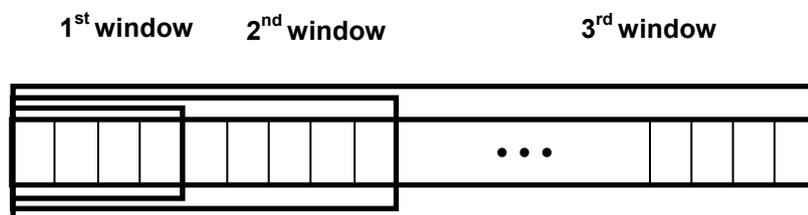


Figure 2.13: Expanding window structure.

The window with the most important subset of encoded data is the first window (W1) and the importance of data additionally included in windows progressively decreases on moving to the third window (W3). The subset data of W1 is contained in all the subsequent windows and is hence the best protected. Apart from W1, each window in addition to some of its own data also encloses the data of the higher importance windows. The size and structure of a window depends upon the elements meeting particular set criteria from a specific subset window. The number of windows is governed by the aggregation scheme employed to group encoded elements.

As in the case of NOW, the encoding process for EW has one important initial step that is to first select a window from which the encoded symbol is to be generated. This selection of a window is determined by PS of a window which is a pre-assigned parameter. PS is determined by the importance of different layers and the data rate available. After a window is selected, the encoding is the standard encoding performed over the source packets contained in that particular window only [22].

The decoding of a window is same as standard decoding, in that, a window is recoverable if the receiver collects at least the same amount of linearly independent encoded symbols obtained from the window (or the windows contained in it) as there were in the window [22].

2.5 Multimedia Communication Standards

2.5.1 Digital Video Broadcasting – Handheld (DVB-H)

DVB-H is the emerging digital broadcast standard for the transmission of broadcast content to handheld terminal devices [23]. DVB-H is based on the DVB-T standard for digital terrestrial television but designed specifically to address the requirements of the pocket-size mobile class of receivers.

The digitization of traditional broadcast systems has made significant progress in recent years. This development could be observed recently with respect to the standard for digital terrestrial television, DVB-T (Digital Video Broadcasting – Terrestrial), which is already in operation in many countries throughout the world. Currently, the system is being rolled out in

Germany and the UK (the Freeview DTT platform). DVB-T has also started in the Netherlands and Italy and was announced to start in France in early 2005; further countries have plans to start services in the near future. In many countries, the decision to select DVB-T as the terrestrial television system was based on the exceptional features of the DVB-T standard, among them the possibility to receive broadcast services also with portable devices and even in cars.

Meanwhile the benefits of a powerful terrestrial broadcast system like DVB-T have attracted the interest of the mobile communication industry. In particular, the ability to reach mobile terminals via a wireless point-to-multipoint link, in connection with wide geographical coverage and high transmission capacity that DVB-T can offer, are features which have sparked the interest of this industry. The international DVB Project has responded to the industry interest by specifying a new transmission standard: DVB-H (Digital Video Broadcasting - Transmission System for Handheld Terminals).

DVB-H is the latest development within the set of DVB transmission standards. Work on the technical specification started in autumn 2002 and was finalised in February 2004; the DVB-H standard was finally published by European Telecommunications Standards Institute (ETSI) as a European Norm in November 2004.

The DVB-H technology is a spin-off of the DVB-T standard. It is to a large extent compatible with DVB-T but takes into account the specific properties of typical terminals which are expected to be small, lightweight, portable and – very importantly – battery-powered. DVB-H can offer a downstream channel at a high data-rate which will be an enhancement to the mobile telecommunications network, accessible by most of the typical terminals. Therefore, DVB-H creates a bridge between the classical broadcast systems and the world of cellular radio networks. The broadband, high-capacity downstream channel provided by DVB-H will feature a total data-rate of several Mbps and may be used for audio and video streaming applications, file downloads and for many other kinds of services. The system thereby introduces new ways of distributing services to handheld terminals, offering greatly-extended possibilities for content providers and network operators.

Mobile multimedia broadcasting is still considered as an emerging technology with further service improvements and extensions expected over the following years. DVB-H is

the European standard for digital TV signals broadcast to handheld devices and is based on the DVB-T standard which mostly uses MPEG-2 packet streams. DVB-H is IP-based and has two important new features introduced in it. Firstly, the transmission takes place in intermittent bursts of maximum data rate of up to 2 Mbps. Secondly, to mitigate errors in mobile and wireless environments the DVB-H standard introduces an optional link layer FEC mechanism, called MPE-FEC.

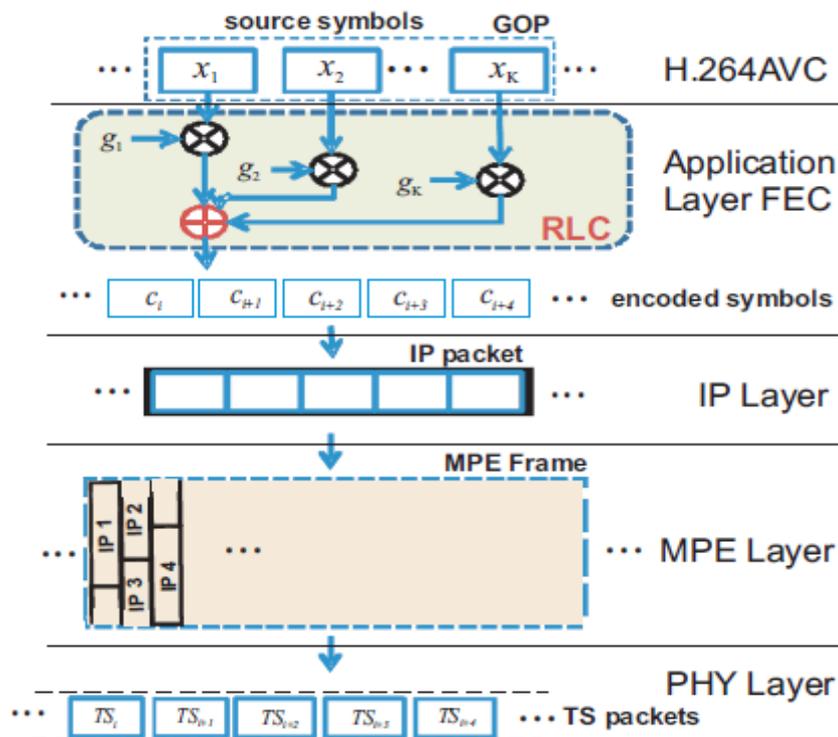


Figure 2.14: DVB-H Protocol Stack.

Assuming that the H.264/AVC standard is used for video encoding, the DVB-H protocol stack, which successively adds headers and prepares the video data for transmission, is shown in Figure 2.14. The H.264/AVC encoder generates NAL units which are then encapsulated into RTP/UDP packets. These packets are then placed into IP packets, which are inserted into MPE sections column by column. The MPE frame containing MPE sections is then divided into transport stream (TS) packets for transmission over the physical DVB-H layer. All TS packets belonging to one MPE frame are sent in one transmission burst. Note that, if AL-FEC is used, the link-layer FEC mechanism (based on RS codes) is switched off.

In general, one IP packet can be contained in multiple TS packets. Moreover, a single TS packet can contain data from two consecutive IP packets. Clearly, due to encapsulation the total header RTP/UDP/IP/MPE/TS overhead could be substantial especially for small IP packet sizes. The header overhead and unused bytes in a data packet are detrimental to the performance of the FEC scheme. Hence it is important to provide FEC configurations which can be adapted to the lower layers of the protocol structure. The same upper layer mechanisms/scheme is used for the DVB-T2 network.

As compared to DVB, DVB-T2 provides increased performance for HDTV broadcasting through the state-of-the-art technologies at the physical layer. An important feature of the DVB-T2 standard is the use of physical layer pipes (PLPs). The TS services are assigned to separate PLPs, and each PLP can have a code rate, modulation, and time interleaving length of its own, allowing for service specific robustness.

2.5.2 Digital Video Broadcasting – Next Generation Handheld (DVB-NGH)

The experience in establishing the DVB-H services have shown that there is a relationship between DVB-T and DVB-H services [24]. When specified, DVB-T and DVB-H targeted two different usage profiles. However, changing media consumption habits increasingly show that media applications require an ability to view the same content on different devices with varying screen resolutions (e.g. mobile and fixed receivers).

With the introduction of DVB-T2, there is a need to look at ways of leveraging the advantages of DVB-T2 in developing NGH in the new paradigm of rich media content delivery in the convergent era. DVB-NGH will be based on DVB-T2 and as DVB-H uses DVB-T physical layer, the physical layer of DVB-NGH will be DVB-T2 based [25].

Next Generation Handheld (NGH) system is needed to accompany digital switch over and convergence of fixed and mobile services as well as telecommunication services. NGH is expected to complement Telecom networks such as 3G and LTE. DVB-NGH in combination with mobile networks like LTE will require fully end-to-end IP based systems, including the full support of an IP transport layer enabling to deploy hybrid network topology to deliver the same audio/visual content over the two networks.

DVB-NGH is likely to be an integral part of the wireless Internet. The coexistence of DVB-NGH with the telecommunication networks like LTE will optimize the service convergence for the benefit of end user. It will be an advantage for the DVB services to be broadcast by LTE in areas without DVB coverage, and likewise, the high performance DVB-NGH in areas of its large customer base will facilitate the delivery of LTE services such as mobile TV/radio [24]. The broadcast approach has by nature considerable advantages in terms of efficiency. A broadcast delivery method is more efficient than unicast if two users require the delivery of the same multimedia content; also, such one-to-many delivery does not require management of the receiving ends.

As DVB-H shares the same physical layer of DVB-T this leads to some restrictions. These could be overcome by a better Time Division Multiplexing (TDM) -system design [28]. The TDM design will enable a more flexible system. This would support low bit rate services in a better manner. Also, in DVB, power saving efficiency and the burst size are linked together in a delicate manner which requires careful tuning of the parameters, which would be done in an independent manner in DVB-NGH.

2.5.3 Long Term Evolution- Advanced (LTE-A)

LTE is an Orthogonal Frequency Division Multiple Access (OFDMA)-based, mobile broadband technology developed by The Third Generation Partnership Project (3GPP). 3GPP initiated its LTE standardization work at the end of 2004, to meet the ever increasing requirements of better QoS and higher wireless data rate. This standardization was successfully completed by the end of 2007 [26]. Thereafter work was started on LTE-Advanced standardization process to address the requirements of Fourth Generation (4G) IMT-Advanced systems by considering new transmission technologies [27], like adaptive interference management, coordinated multiple point transmission and reception, and relay.

The next significant performance leap will come from heterogeneous networks, which bring the network closer to the user through small cells such as picocells and femtocells. This ensures an enhanced mobile experience with higher data rates to more users.

LTE-A was formally submitted as a candidate 4G system to International Telecommunication Union-Telecommunication (ITU-T) in late 2009, and was approved into ITU, IMT-Advanced and is expected to be finalized by 3GPP in 2011.

The possibility of a terminal communicating with the network, and the data rate that can be used, depends on several factors, such as the path loss between the terminal and the base station. The link performance of LTE is already quite close to the Shannon limit and from a pure link-budget perspective, the highest data rates supported by LTE require a relatively high signal-to-noise ratio. Unless the link budget can be improved, for example with different types of beam-forming solutions, a denser infrastructure is required to reduce the terminal-to-base-station distance and thereby improve the link budget [29].

A denser infrastructure is mainly a deployment aspect, but in later releases of LTE, various tools enhancing the support for low-power base stations are included. One of these tools is *relaying*, which can be used to reduce the distance between the terminal and the infrastructure, resulting in an improved link budget and an increased possibility for high data rates. In principle this reduction in terminal-to-infrastructure distance could be achieved by deploying traditional base stations with a wired connection to the rest of the network.

A wide range of relay types can be envisioned, some of which could already be deployed in release 8. *Amplify-and-Forward* relays, commonly referred to as *repeaters*, simply amplify and forward the received analog signals and are, on some markets, relatively common as a tool for handling coverage holes. *Decode-and-Forward* relays decode and re-encode the received signal prior to forwarding it to the served users. The decode-and-re-encode process results in this class of relays not amplifying noise and interference, as is the case with repeaters. They are therefore also useful in low-SNR environments.

LTE release 10 introduces support for a decode-and-forward relaying scheme (repeaters require no additional standardization support other than RF requirements and are available already in release 8). A basic requirement in the development of LTE relaying solutions was that the relay should be transparent to the terminal – that is, the terminal should not be aware of whether it is connected to a relay or to a conventional base station. This ensures that release-8/9 terminals can also be served by relays, despite relays being introduced in release 10.

2.6 Multiple Description Coding

Multiple description coding (MDC) is a coding technique that splits a single media stream into n substreams ($n \geq 2$) each of which is referred to as a *description*. These descriptions can then be routed over multiple paths to a destination. The paths may be partially or fully disjoint which ensures that even with adverse channel conditions on one path, some descriptions will make it through to the destination. It should be possible to recreate a low quality media stream using any one description. However, with reception of more descriptions the quality of media stream increases progressively. The idea of MDC is thus to provide error robustness to the transmitted data.

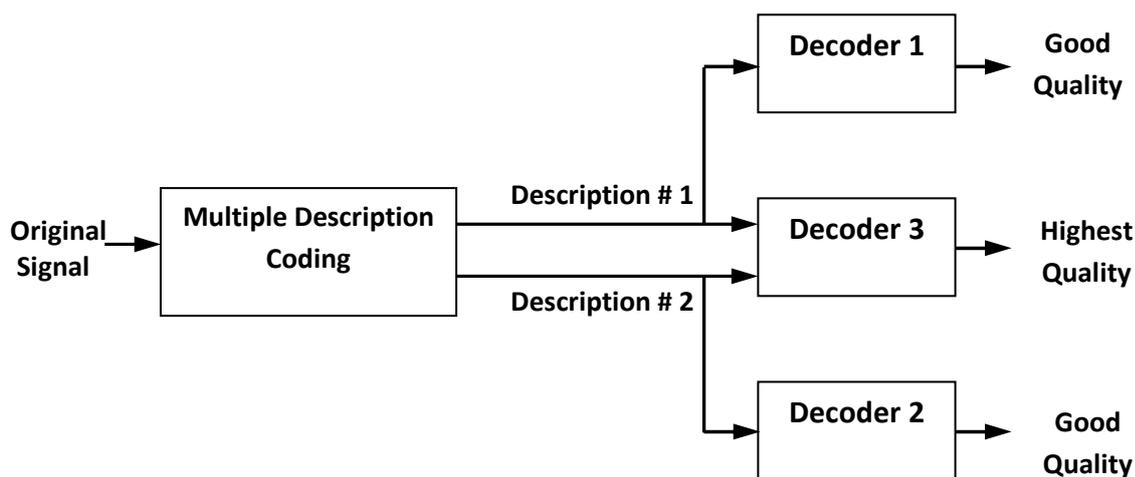


Figure 2.15: A MDC system generating two descriptions [6].

A very simplified way to create two descriptions could be to partition a video stream into odd and even frames and encode them separately, which could later be sent on different paths. Figure 2.15 shows that in a scheme with two descriptions, correct reception of just one description can result in an acceptable quality. In case, both descriptions are received correctly then the best quality can be obtained [6].

In case of scalable coding, the coding process generates a base layer and one or more enhancement layers. The enhancement layer will be rendered useless in case of loss of base layer. In contrast MDC creates multiple representations each of which is equally important for video re-construction. Therefore, in case of MDC it is possible to decode the original video data to a low-quality representation even with correct reception of any one description [6].

MDC must sacrifice some compression efficiency to gain robustness to the loss of descriptions [30]. Thus, there is a trade-off; MDC can be applied to great benefit if the disadvantage in compression is offset by the advantage of reducing transport failures.

2.7 Summary

This chapter covers the necessary background for the following chapters. A discussion of the latest state-of-the-art video compression standards is provided. The error-resilience features are described which are used in later chapters to design novel solutions.

The Fountain codes have been described. LT codes are the first realization of Fountain codes which are simple to implement but are not very efficient. Raptor codes use a two stage encoding process, with each of LDPC and LT codes in one stage. This novel design reduces the complexity and improves the decoding performance. RLC codes are recent codes which are simple to implement and provide good performance, but have higher decoding complexity. However, with short length codes and the increase in processing power this limitation of decoding complexity is not an issue. Raptor codes have been adopted for the emerging communication standard like DVB-H and may be used in DVB-NGH and LTE-A.

An improvement over DVB-H is being standardized as DVB-NGH. The traditional boundaries of broadcasting and cellular communication will diminish with the introduction of these upcoming standards.

A section on MDC describes the basic principles and architecture of the technique. These topics are used in later chapters along with their application to specific research problems.

Chapter 3

Data Partitioned H.264/AVC Video Broadcasting over DVB-H and DVB-T2 Networks

3.1 Introduction

Since each part of video data is not equally important for video reconstruction, it is beneficial to group it based on its importance, and then provide different degree of protection using UEP FEC, with the important data having more protection. DP is one such low-cost feature in H.264/AVC enabling partitioning of video data based on its importance. The proposed schemes exploit the DP H.264/AVC video transmission using Raptor and RLC and their performance is investigated as AL-FEC solutions in the DVB-H and DVB-T2 networks. For the analysis, the DVB-T2 network is configured according to the DVB-T2 Lite profile [31], which is mainly intended for mobile reception. The RLC results are extended to provide comparisons between NOW RLC and EW RLC, which are two effective UEP RLC strategies. The results obtained using realistic simulated DVB-H and DVB-T2 channel traces show viability of the EW RLC as a promising AL-FEC solution for multimedia broadcast applications due to its robustness to varying channel conditions. The results for DVB-H are also provided with slice-partitioned H.264/AVC.

Using FEC mechanisms is the most favoured approach as retransmissions in broadcasting applications are usually counter-productive [32], [33], [34]. Two key challenges of multimedia broadcasting over wireless networks are high and varying error characteristics of underlying wireless channels and large heterogeneity of users' equipment.

One of the major arenas where FEC codes are successfully applied for wireless streaming video protection in recent years are DVB networks, in particular, their extensions for handheld devices called DVB-H and next generation handheld (DVB-NGH) which will be based on DVB-T2. One of the video coding standards approved for DVB-H and DVB-NGH is H.264 Advanced Video Coding (AVC) [2] - a state-of-the-art video coding standard achieving significant compression efficiency and gaining widespread use in the emerging communication systems and applications.

For error protection of broadcasted video, DVB-H specifies a Multi-Protocol Encapsulation - Forward Error Correction (MPE-FEC) solution at the link layer designed for real-time services [23], which applies adaptive punctured Reed-Solomon (RS) FEC against packet losses. The DVB-H standard also provides a possibility of using AL FEC solution for IP datacasting services using a well-known class of rateless codes called Digital Fountain Raptor codes [9], [33], [35]. Though Raptor codes are currently proposed only for non real-time services, they have been investigated for multi-burst protection and compared to MPE-FEC in terms of performance and delay for real-time services over DVB-H [11]. As DVB-H uses DVB-T physical layer, the physical layer of DVB-NGH will be DVB-T2 based [25]. Motivated by the need to provide better high definition television (HDTV) services, DVB-T2 brings higher throughput than DVB-T.

Apart from DF Raptor codes, a class of rateless codes that have been gaining increased popularity recently for applications in wireless broadcast/cellular networks are RLC [13], [36]. In addition and for possible future extensions, as multi-hop and cooperative communications are becoming increasingly popular in emerging wireless network architectures, introduction of RLC may serve as a step forward towards exploring the benefits of network coding [37]. In the context of future DVB networks, this may be important for emerging concepts such as hybrid broadcast/cellular networks (with clients equipped with multiple wireless broadband interfaces) and device-to-device communications [38], [39].

The focus of this study is to analyse the use of the DP error resilience feature of H.264/AVC in combination with Raptor codes and RLC as an AL-FEC solution for video transmission in mobile DVB networks. Furthermore, the benefits of UEP RLC as an AL-FEC solution, tailored with the DP feature of H.264/AVC for video transmission in the DVB-H standard, is investigated in detail. The DP feature of H.264/AVC enables simple rate adaptation crucial in wireless broadcasting [23], [40], [41]. With DP the percentage of entirely lost frames can be reduced [42]. The use of slice partitioned H.264/AVC with UEP RLC is also covered. Note that the alternative is to use Scalable Video Coding (SVC) [4] (which by itself introduces performance penalty) possibly together with some type of error resilience [40]. Wireless video broadcasting of H.264/SVC, in combination with rateless codes, has been studied extensively (see [33], [40], [43] and references therein). The main advantage of DP H.264/AVC over SVC is its compatibility with H.264/AVC. The UEP using DP of H.264/AVC is proposed in [5]. DP H.264/AVC and UEP are later used for wireless

video delivery in [44] using rate compatible punctured convolutional (RCPC) codes, in [45] with hierarchical quadrature amplitude modulation, in [46] with Raptor codes for IPTV, and in [47] with growth codes. In [48], design of a Raptor generator matrix based on frame dependencies within a group of picture (GOP) has been proposed. The redundancy allocation process is tied to the knowledge of the channel loss which may not be suitable for real-time applications and video broadcasting. The proposed scheme also needs a mechanism to transport the generator matrix coefficients to decoder.

In [49], the Expanding Window Fountain (EWF) codes, as LT [8] UEP codes, are proposed and their asymptotic behavior analytically found; EWF is a layered scheme where the protection of lesser important layers also includes the more important layers. In [21], EWF codes are optimized for scalable video delivery. In [34], a layer-aware FEC mechanism has been proposed, similar to the EWF concept. However, in contrast to [21], [49], in [34] UEP low-density parity-check (LDPC) codes and Raptor codes are designed for protection of H.264/SVC video. The work also explores the suitability of the proposed codes to physical and application layer FEC protection for the latest video communication standards including DVB-H. Extending the UEP design methods of [21] from DF to RLC, UEP-based RLC strategies for the SVC delivery have been investigated in [14],[22] and [50].

While [22], [50] are more focused on performance analysis of UEP RLCs schemes over random erasure channel and SVC [4], this chapter addresses design challenges of UEP RLCs schemes with DP and slicing feature of H.264/AVC for real-world implementation in the mobile DVB networks and compares their performance to standard DF Raptor codes. While both DF Raptor codes and RLC have been individually well studied in literature, interestingly, no systematic performance/complexity comparison has been conducted yet. The key contribution of the chapter is comparison between Raptor and RLC codes for DVB-H and optimal code design guidelines for UEP RLC over DVBH and DVB-T2 configured according to the DVB-T2 Lite profile, which is a profile intended to allow simpler receiver implementations for very low capacity applications such as mobile broadcasting, although it may also be received by conventional stationary receivers. Thus, this chapter evaluates performance of the proposed schemes in a real-world environment and as a result highlights set of optimized design recommendations.

This chapter is organized as follows. The proposed video broadcasting system is described in Section 3.2. The video broadcasting with Raptor and RLC as AL-FEC is discussed in Section 3.3. Section 3.4 discusses the UEP AL-FEC for video broadcasting. Section 3.5 covers the performance evaluation of EW-RLC based solution with DVB-T2. The Slice Partitioned H.264/AVC Video Broadcasting with UEP RLC is covered in Section 3.6. Finally, Section 3.7 provides the chapter summary.

3.2 Proposed Video Broadcasting Based on DP H.264/AVC

In the following, a source-channel coding scheme is proposed for video broadcasting over DVB-H that exploits both state-of-the-art error resilient video coding and rateless AL-FEC to adaptively and optimally protect video against packet losses. The block diagram of the proposed system is shown in Figure 3.1.

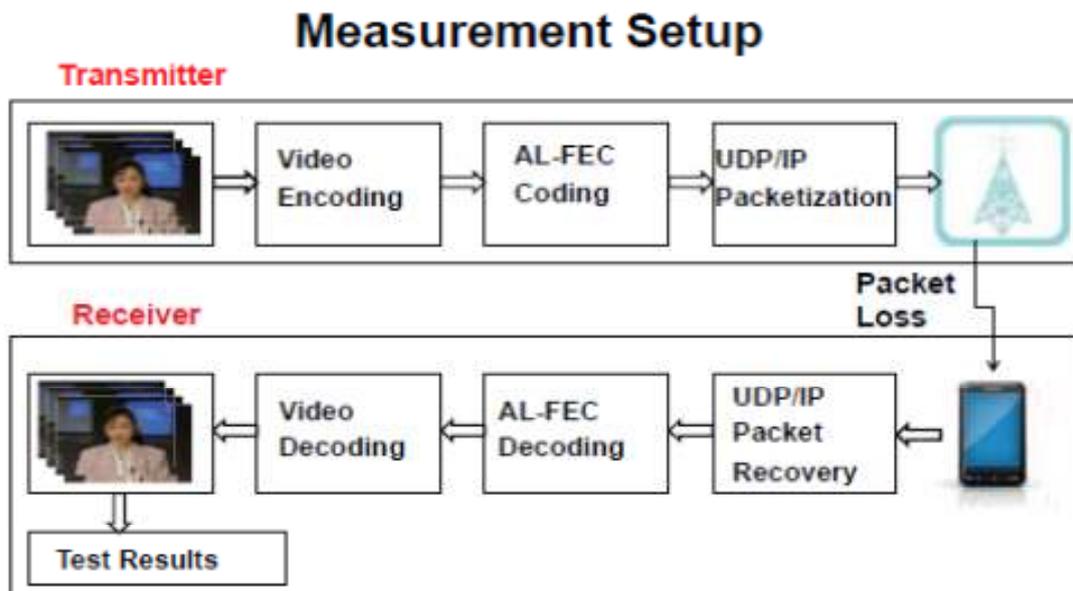


Figure 3.1: Block Diagram of the Proposed System.

3.2.1 Video Encoding

For each non-I frame, the video encoder output is split into three partitions: DP A, DP B, and DP C. The encoding is done with the Constrained Intra Prediction (CIP) flag set and the Macroblock Intra Update (MBIU) feature to limit the effect of error propagation. The cost of using CIP [51] can be neglected. All the partitions of a particular type from all frames within one GOP are extracted and aggregated together. That is, DPs A from all encoded frames of a GOP together with the encoded I frame are grouped generating the most important source class or layer. DPs B and C form second and third importance class/layers, respectively.

Each of the layers is packetized into equal-length source symbols/packets. Note that the length of the source symbols determines the number of source symbols contained in the source message, k , which directly influences FEC efficiency and decoding complexity. For Raptor codes, ideally, short source symbols should be selected resulting in higher values of source message length k , typically in the order of thousands of symbols, since decoding complexity increases only linearly with k , and on the other hand, performance is improved for larger k . In the case of RLC, however, large source symbols are favoured in order to yield small values of k , typically in the order of hundreds, because of high decoding complexity of RLC. The key issue to be investigated is the impact of different selection of k , imposed by performance and complexity issues, to these two codes.

3.2.2 AL-FEC Coding

A straightforward approach to AL-FEC is to apply one code over the entire source block (i.e., a GOP), that is, to equally protect the whole stream. This, equal error protection (EEP), however, does not capture adequately unequal importance of the three classes of the DP H.264/AVC code stream.

A preferable option is a UEP scheme which assigns redundancy to the source classes based on their importance to video reconstruction. Two different UEP design approaches are explored. The first one, NOW scheme uses a separate AL-FEC code, either Raptor or RLC, for each of the DP layers, and by using the codes of different code rate, UEP is realized. The

second approach uses the expanding window approach [21], [22] to encode jointly all the classes.

The encoding process for both NOW and EW codes has one important initial step, that is, to first select a window from which the encoded symbol is to be generated. This selection of a window is determined by selection probability of a window which is a pre-assigned parameter keeping in mind the importance of different layers and the data rate available.

3.2.3 Data Packetization

After AL-FEC coding, NOW or EW with Raptor or RLC codes, encoded symbols are grouped to fit the content of a single IP packet. This process also takes into consideration the header overhead. IP packets are placed into the MPE frames where each IP packet is encapsulated within a single MPE section with its own overhead containing the error-detection field. The option of using RS codes is not employed.

Each MPE frame is transmitted within a single transmission burst by mapping the MPE frame data onto 188-bytes long PHY TS packets. The IP packet size is set to be an integer multiple of 184 bytes (the content size of TS packet), i.e., more precisely, each MPE section is transmitted over the integer number of TS packets. It is important to underline that IP packets are either correctly received at the decoder or lost in the transmission process. An IP packet is considered lost, if at least one TS packet that forms it is lost, which is the standard operation of the DVB-H link layer [11], [52].

For simplicity, MPE sections and TS packets are assumed to be aligned, i.e., the borders between MPE sections at the MPE layer are borders between TS packets at the physical layer. Unfortunately, in practice this is not always possible, and that is why sometimes a single TS packet loss can cause loss of two consecutive IP packets. However, since this situation being rare is neglected and perfect alignment of MPE sections and TS packets is assumed.

3.3 H.264/AVC Video Broadcasting with Raptor and RLC Codes

In this section, we experimentally compare the performance of the standard Raptor codes with RLC for different source message sizes in order to select source symbol size, message size, and IP packet size that provide the best trade-off between coding efficiency and delay/complexity over a range of channel parameters. Then, after setting the message/packet/symbol size the UEP optimization can be done to find the optimal window selection probabilities.

3.3.1 Symbol Size Selection for Raptor and RLC Codes

The simulations are performed using DVB-H TS packet error traces obtained by realistic DVB-H link layer simulator [53] to determine adequate message/symbol/packet size for systematic Raptor codes (RTC) and RLC. Different RTC and RLC coding configurations are compared in terms of reconstruction performance as well as decoding complexity, where a configuration refers to one pair of source symbol and IP packet size used.

After AL-FEC coding (with either RTC or RLC), encoded symbols are grouped into sets of $N_s \geq 1$ symbols and packed into a single IP packet. The number of encoded symbols N_s carried by an IP packet depends on the symbol size l_s , the IP packet size l_{IP} and the set of protocol headers within each IP packet, described as follows. Each IP packet contains a four-byte random number generator seed used to recreate the RTC/RLC encoding coefficients for all the encoded symbols N_s contained within the packet. The remaining symbols derive their coefficients based on the first symbol's RNG seed. Then, another 60 header bytes comprising RTP/UDP/IP headers are added to form an IP packet (this number can be reduced using robust header compression). Therefore, IP packet size is equal to:

$$l_{IP} = l_{OH}^{(IP)} + N_s \cdot l_s, \quad (3.1)$$

where $l_{OH}^{(IP)} = 64$ bytes.

After IP packetization, at lower DVB-H layers, each IP packet is arranged into a separate link-layer MPE section containing MPE header of $l_{OH}^{(MPE)} = 16$ bytes. MPE section is

split into TS packets each with a four-byte header followed by 184 bytes of payload. Assuming that each IP packet with the corresponding MPE section overhead fits into exactly $N_{TS} = \frac{(l_{IP}+16)}{184} TS$ packets, the total overhead per IP packet equals:

$$l_{OH} = l_{OH}^{(IP)} + l_{OH}^{(MPE)} + N_{TS} \cdot 4 = 80 + \frac{l_{IP}+16}{184} \cdot 4 \quad (3.2)$$

Different IP packet sizes l_{IP} are tested ranging from 300 bytes to over 1450 bytes (the IP packet size limited to 1500 bytes). First, the IP packet size l_{IP} is picked so that, jointly with MPE overhead, it fits a whole number of TS packets, i.e., $l_{IP} = N_{TS} \cdot 184 - 16$, where N_{TS} is a positive integer. Then, the available capacity of IP packet is divided into a selected number of N_S equal-length encoded symbols, where the symbol size is determined as $l_s = (l_{IP} - l_{OH}^{(IP)})/N_S$.

Table 3.1: System Configurations.

FEC	l_{IP} [bytes]	N_{TS}	l_s [bytes]	N_S	$\frac{(l_{OH} + 16)}{N_{TS} \cdot 188}$ [%]
RTC656	352	2	144	2	23.40
RTC721	720	4	131	5	12.76
RTC1475	1088	6	62	16	9.22
RTC1086	1456	8	87	16	7.45
RTC136	1456	8	696	2	7.45
RLC144	720	4	656	1	12.76
RLC185	1088	6	512	2	9.22
RLC68	1456	8	1392	1	7.45
RLC136	1456	8	696	2	7.45

Let the total source bit budget be B bytes; for example, this is the size of a GOP. Then, the number of source symbols in the source message is $k = \text{ceil}(B/l_s)$. By varying l_s and l_{IP} , different coding configurations are obtained. Five different RTCs and four RLCs shown in Table 3.1 are tested. Four configurations use l_{IP} in the range of 1400-1500 bytes, two configurations use the IP size of roughly 1000 bytes, while three use very low IP packet sizes. This way low, medium and large IP packet sizes are covered.

The number within the FEC scheme's name denotes the source message length k , e.g., RTC656 is a Raptor code of source message length $k = 656$ symbols. The last column of the table describes the transmission efficiency in terms of percentage of overhead data within the total transmitted data. From (3.2), it is noted that with increasing N_{TS} , the overhead grows much slower than the total amount of transmitted data, which makes the overall transmission more efficient for larger IP packets.

The symbol sizes l_s are generally selected to be small for RTC, so as to yield larger source message size (in conventional FEC, the increase of message length generally improves performance). However, for a fair comparison, RTC with smaller source message sizes is included. The reverse is true for RLC, i.e., larger symbol sizes l_s are selected so as to yield shorter source message sizes k to make the decoding complexity feasible. The selection of different l_{IP} for RTC and RLC is motivated by the need to design and test schemes with different N_{TS} , so that the scheme's response to the TS error trace files could be evaluated over different possibilities. l_{IP} is always picked so that after adding a 16-byte MPE header, an integer multiple of 184 (payload of a TS packet) bytes is obtained.

Let $B = 94400$ bytes. Then, for example, in the RLC68 configuration, the source message length of $k = 68$ symbols results in the symbol size $l = 94,400/68 = 1392$ bytes (with data padding). In this case $N_s = 1$ encoded symbol is placed in the IP packet. This symbol gets added with 64 bytes of header data to yield an IP packet of size $l_{IP} = 1456$ bytes. During transmission, this IP packet occupies $N_{TS} = 8$ TS packets (after having added the MPE header to the IP packet) and includes the total overhead of l_{OH} bytes, resulting in the ratio of overhead data in the data stream equal $\frac{l_{OH}}{N_{TS} \cdot 188} = 7.45[\%]$.

3.3.2 Simulated EEP Performance of Raptor and RLC codes

The results over a range of acceptable code lengths for both codes are reported. Indeed, for RTC configurations with message lengths between $k = 136$ symbols and $k = 1475$ symbols are tested. For RLC, on the other hand, configurations between $k = 68$ symbols and $k = 185$ symbols are tested, since a higher k would result in unacceptable decoding complexity (see [20]).

It is important to underline that IP packets are either correctly received or lost in the transmission process. Therefore, using large N_s creates large increments in the number of encoded symbols received at the destination from each IP packet, which might compromise the FEC efficiency. On the other hand, small values of l_{IP} result in large transmission inefficiencies due to large fraction of overhead data in the transmitted data stream. Therefore, trade-off solutions point to the case of sufficiently large l_{IP} jointly with sufficiently low N_s (i.e., large l_s), which is exactly the case where the RLC solution is feasible.

Video sequence *Foreman* in CIF format encoded using the H.264/AVC software JM version 16.2 [54] is used without any error resilient option. All simulations have been performed using a GOP size of 64 frames, one slice per frame, and a frame rate of 25 frames per second. This configuration implies the source data rate of 295 kbps corresponding to 2.56 seconds of video data (one DVB-H burst). This rate ensures that sufficient amount of video data is transmitted during one burst of DVB-H transmission and is used also in [52].

The trace files with the losses of TS packets for different signal-to-noise ratio (SNR) in the DVB-H channel are obtained by realistic link-level simulator, which are confirmed to match very well the field measurements [53]. The link level simulation model assumes a TU6 DVB-H channel model with a constant Doppler frequency of 10 Hz, 8k mode, and guard interval 1/4. The simulated modulations are QPSK, 16-QAM, and 64-QAM encoded with a convolutional code of rate 1/2. The results presented in the rest of the chapter are for 16-QAM modulation; the results for other two modulations are of similar nature. The MPE-FEC error correction is not used. Since it includes complex physical layer simulation, the link-level simulator produces sample error traces that are further extended using appropriate

packet level DVB-H channel models [53]. Finally, the obtained TS packet error traces of sufficient lengths are obtained and used in the rest of the chapter.

Table 3.2: Raptor and RLC Configurations Results.

FEC	Error Free Data Rate [kbps]	Measured Data Rate SNR12 [kbps]	Measured Data Rate SNR14 [kbps]
RTC656	385.2	606.3	452.3
RTC721	341.5	559.3	406.5
RTC1475	334.5	560.5	394.8
RTC1086	325.5	554.6	394.8
RTC136	318.8	549.9	394.8
RLC144	337.3	549.9	401.8
RLC185	325.2	542.8	394.8
RLC68	318.8	564	390.1
RLC136	318.8	535.8	390.1

Each FEC configuration is simulated with at least 1000 runs for each channel SNR value. As in [52], performance criteria called ESR5(20) (Erroneous Seconds Ratio) is used which defines a limit of at most 1 sec of erroneous transmission within each 20 sec segment of video transmission. Thus, the results are presented as the minimum overall data rate (including all the headers), called measured data rate, that satisfies the ESR5(20) performance criteria vs. SNR in the DVB-H channel. The results for two different SNRs are shown in Table 3.2. The error free data rate is the one which is required by the FEC scheme to transport the GOP data if the channel were error free. It can be seen from the table that the best results are obtained with RLC136 and RTC136. However, note that RLC185, RLC68 and RTC1086 suffer only a negligible rate loss. Similar results for other SNRs and different rates can be found in [55].

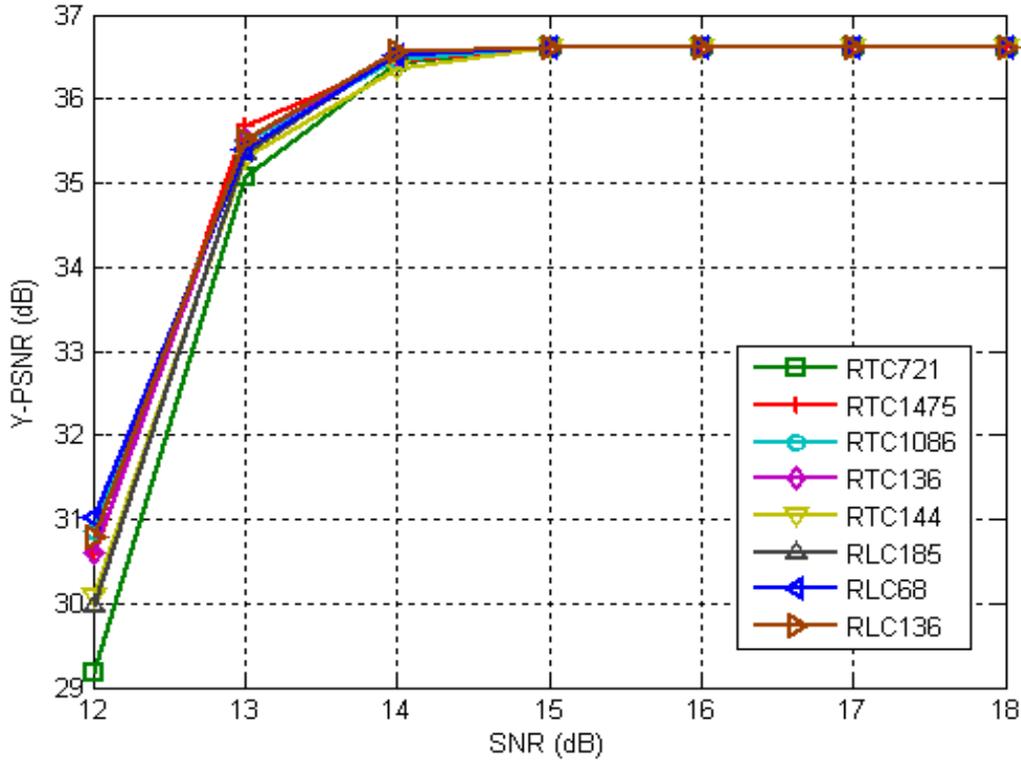


Figure 3.2: PSNR comparison between different RTC and RLC configurations: Y-PSNR vs. SNR in the DVB-H channel.

The expected peak signal-to-noise ratio (PSNR) averaged over all frames in the sequence as a criterion for quality comparison is also used. If an entire GOP cannot be decoded then the decoder applies a simple error concealment technique by replacing the lost GOP by the last frame of the previously successfully decoded GOP. To compare the quality for each of the used configurations, the overall data rate was fixed to be 452 kbps and each configuration is subjected to the same loss patterns. The results are shown in Figure 3.2. It can be seen that several configurations, RTC1475, RTC1086, RTC136, RLC68, and RLC136 provide the best PSNR performance. The worst performing sequence is RTC656 (not shown in the figure) and then RTC721 which have the heaviest data overhead. In particular, for SNR = 12 dB, RLC68 is the best performing with PSNR = 31.02 dB while the best performing RTC, RTC1086, is lagging behind for 0.16dB. At SNR = 13 dB, RTC1475 is the best performing with PSNR = 35.69 dB while RLC68 performs for only 0.09 dB worse, and this gap is similar at SNR = 14 dB. At SNR = 15 dB all eight schemes reach error free PSNR performance of 36.61 dB.

Besides performance comparison, it is important to conduct a complexity analysis as well. For source block size k and the number of received symbols n , the failure probability of an RTC for (n, k) decreases exponentially with increasing the number of received symbols. For roughly 12 additional symbols the failure probability is 0.1%, whereas for about 24 additional symbols the failure probability reduces to 0.0001% [32]. For the RLC case and $GF(2^8)$ field size that is used in the setup, the failure probability with zero additional received symbols is around 0.01%, and with one additional received symbol, it drops below 0.000001%.

The decoding complexity of the RTC is linear with the source message length k , i.e., it grows as $O(k)$, thus mobile phones can easily handle decoding of RTC codes of message in the order of thousand symbols. On the other hand, the GE decoder of RLC codes has complexity that grows as $O(k^3)$. This clearly makes it impossible to employ RLC for large source blocks. However, for short block lengths that are the focus, the asymptotic behavior does not say much about the decoding complexity. In several recent papers, the decoding complexity and possibilities of practical implementation of RLC is investigated [20], [56], [57]. In [57], the study focused on streaming server solutions, and using optimized algorithms and GPU it is demonstrated that RLC with 128 blocks and large symbol sizes can achieve encoding rates up to 294 Mbps and decoding rates up to 254 Mbps. It is also argued therein that with such high rates, deploying RLC as an encoding solution is feasible for streaming servers. In [20] and [56], the study focused on mobile user applications within smart phones, the symbol and block sizes for RLC similar to the one used in this study are practically implemented. It is concluded that the implementations of RLC decoding with short codes (source length $k = 64$ or 128) are feasible with standard video bit-rates of several hundreds of kbps.

From the above analysis, two best performing schemes are adopted from each code class (RTC and RLC), i.e., RTC136, RTC1086, RLC68, and RLC136, having in mind that shorter message sizes are preferable option if performance is not severely affected.

Table 3.3: Relative Partition Sizes for First GOP of the CIF Foreman Sequence.

Partition	Size [bytes]	Size [%]	Cumulative PSNR [dB]	Priority Class
IDR	3760	4.00	-	HPC
DP-A	47940	51.00	24.51	HPC
DP-B	15800	16.80	28.64	HPC/LPC
DP-C	26500	28.19	35.64	LPC
Total	94000	100	35.64	-

3.4 DP H.264/AVC Video Broadcasting with UEP RLC

In this section, the DPs are organized into classes of different importance. The partitions with their relative sizes, PSNR contributions and the window structure for two windows priority classification are shown in Table 3.3. Intuitively, Instantaneous Decoder Refresh (IDR) and DP A is always placed in so called High Priority Class (HPC), which is equivalent to the base layer in SVC [4], because the remaining data cannot be decoded unless IDR and DPs A are correctly received.

Furthermore, two schemes are created, RLC-IAB, where DP B together with DP A is placed in the HPC and DP C is treated as the less importance class called Low Priority Class (LPC). As can be seen from the table, the size of the first layer (HPC) is then about 71% of the total size. However, the second layer provides a significant performance improvement of over 7 dB. In the second scheme, RLC-IA, only IDR and DP A are placed into HPC, and DP B and DP C comprise LPC. The size of HPC is then about 55% of the total stream size.

In order to demonstrate the effect of removal of DP A and DP B, frame 45 for both RLC-IA and RLC-IAB has been extracted and is shown in Figure 3.3. The advantage of RLC-IA is a smaller base layer (with low quality) with less decoding failures though UEP.



(a) Original sequence, PSNR 35.64 dB.



(b) RLC-IAB, PSNR 28.64 dB.



(c) RLC-IA, PSNR 24.51 dB.

Figure 3.3: Frame 45 of the Foreman sequence for various configurations.

Table 3.4: Four Selected Coding Configurations for the UEP Analysis.

Scheme	No of Symbols		Symbol Size [bytes]	No. Symbols in IP Packets
	HPC	LPC		
RTC1086	776	305	87	16
RTC136	97	39	696	2
RLC68	49	19	1394	1
RLC136	97	39	696	2

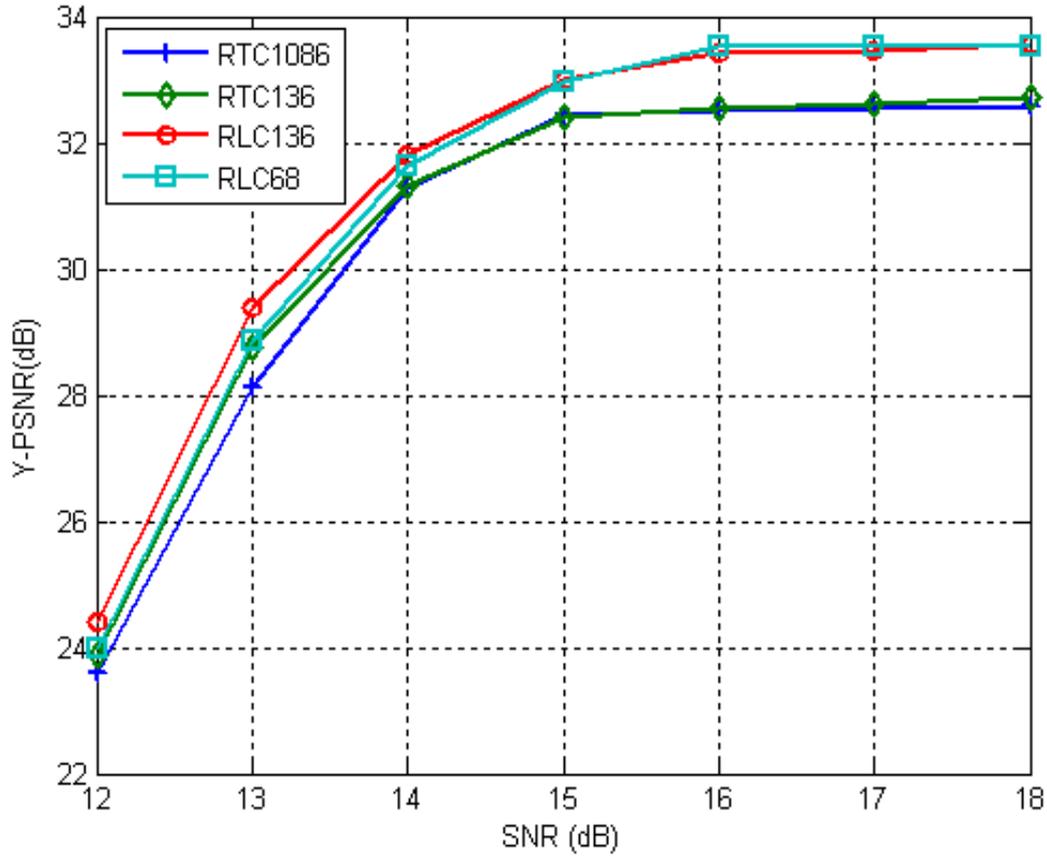


Figure 3.4: PSNR comparison between four selected RTC and RLC configurations: Y-PSNR vs. SNR in the DVB-H channel.

Figure 3.4 shows frame-average PSNR vs. channel SNR for the four selected RLC-IAB configurations, shown in Table 3.4, when transmission data rate is fixed at 383 kbps and with protection probability of HPC of 75%. The results with values of SNR that were available from the link-level simulator are presented. For simulating each scheme, the TS packets (as determined by the data rate) are transmitted and are subjected to the loss pattern. Loss of an IP packet is assumed if any of its enclosed TS packets is lost. Although the difference in performance is not very significant, however, RLC136 is the best performing configuration for all SNRs. The PSNR difference between RLC68 and RTC136 is marginal. However, RTC1086 provides the worst overall performance. Similar results were obtained for other UEP schemes (see [58]).

Having in mind the limited complexity of GE decoding for RLC68 (due to small code length), very small delay introduced, and competitive performance results, it is concluded

that it is a good candidate for real-world implementations. Thus, all further simulations are presented for different UEP configurations obtained with RLC68.

If channel characteristics are known, it is possible to provide optimal allocation of redundancy to different source priority classes. For this task, the expected PSNR is maximized at channel packet loss rate ϵ using analytically computed probabilities of decoding error performance. That is, $\max_{\pi} PSNR(\pi, \epsilon) = \sum_{i=0}^L P(i, \epsilon) psnr(i)$, where L is the number of classes, $P(i, \epsilon)$ is probability that class i will be the highest class recovered at channel packet loss rate ϵ , $P(0)$ is the probability that nothing is recovered, $psnr(i)$ is PSNR of the reconstruction if all classes up to and including class i are recovered ($i = 1, \dots, L$), π is an L -tuple vector of window selection probabilities that determines the UEP allocation scheme, and $PSNR(\pi, \epsilon)$ is the expected PSNR when UEP scheme π is used.

Analytical expressions [50] for probabilities $P(i, \epsilon)$ assuming a random channel loss model, for both NOW and EW schemes are as given below:

$$P(l) = \begin{cases} 1 - P_{d,N}(1), & l = 0 \\ \prod_{i=1}^l P_{d,N}(i) \cdot (1 - P_{d,N}(l+1)), & 1 \leq l \leq L-1 \\ \prod_{i=1}^L P_{d,N}(i), & l = L, \end{cases} \quad (3.3)$$

where $P_{d,N}(l)$ [50] is the probability of decoding of the l^{th} layer after N received symbols.

The optimization method is exhaustive search which has a low complexity because number of layers L being considered is 2 or 3. For each candidate UEP scheme $L+1$ computation are required. The proposed UEP schemes use an increment of 5 for selection probability, hence at the most only 20 schemes are considered.

Above, however, it is assumed that channel characteristics are known at the transmitter, which is unrealistic assumption in DVB-H. Thus, instead, the average performance over a range of expected packet loss rates is maximized:

$$\max_{\pi} PSNR(\pi) = \sum_{\epsilon_j \in [I]} \sum_{i=0}^L P(i, \epsilon_j) psnr(i), \quad (3.4)$$

where $[I]$ is the set of probable packet loss rates. This way, it is aimed to find a scheme that performs in average the best over a range of channel conditions.

Next, for RLC68, NOW and EW schemes are compared. The same source configuration shown in Table 3.3 is used, and employed EW schemes are shown in Table 3.5. The symbol size for all the schemes is 1394 bytes. EW1 refers to the single layer scheme (with all parameters set as in the EW2 scheme).

Table 3.5: Configurations and the Window Structure for RLC68.

Scheme	Window	No. Of Symbols	No. Of IP Packets	No. Of TS Packets
EW1	W1	68	68	544
EW2(RLC-IAB)	W1	49	49	392
	W2	19	19	152
EW(RLC-IA)	W1	37	37	296
	W2	31	31	248
EW3	W1	37	37	296
	W2	12	12	96
	W3	19	19	152

Results for four different UEP schemes and the EEP scheme of the RLC68-IA scheme for NOW and EW2 RLC-IA at data rate of 383 kbps are shown in Figure 3.5. UEP85N refers to a UEP scheme with 85% probability of selection for HPC, with N denoting NOW. Similarly, UEP75E is an EW scheme with 75% probability of selection for HPC. The frame averaged PSNR vs. SNR in the DVB-H channel is shown.

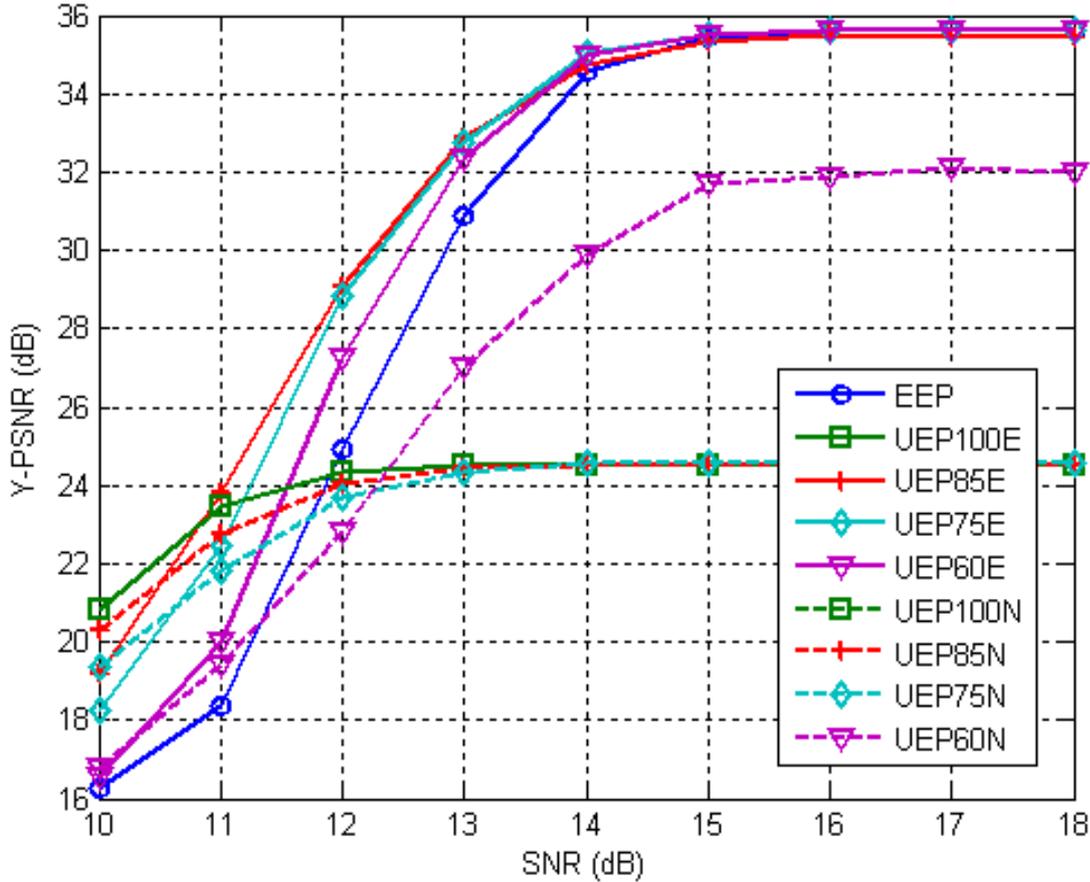


Figure 3.5: PSNR vs. SNR in the DVB-H channel for four different NOW and EW UEP schemes and the single-layer EEP scheme of RLC68 at data rate of 383 kbps for Foreman sequence.

As can be seen from the figure, the EW UEP schemes outperform EEP and NOW schemes for all SNRs. The EEP scheme outperforms the best NOW scheme for high SNRs, while for low SNR, the performance of the NOW UEP schemes is better. This is because at lower SNRs the contribution to PSNR by successful decoding of HPC is significant. As the SNR is increased, the UEP NOW curves get a performance penalty, with UEP schemes with higher protection of HPC being the ones which are severely affected. This is because at high SNRs the successful decoding of both classes is less probable due to higher selection probability of the HPC. Similar results for different parameters can be found in [58].

Superiority of the EW schemes over NOW schemes is clear. The main reason for performance gains of EW compared to NOW lies in the fact that NOW attempts to select

symbols from either class in a mutual exclusive manner, in the sense, that if the protection for HPC is increased, this directly detracts LPC. In case of the EW schemes the windows are nested and correctly received packets from smaller windows may be used in recovery of larger windows [22], [50].

The optimal EW scheme in terms of maximizing (3.4) was UEP(85,15), whereas the optimal NOW scheme was UEP(60,40). In both cases it is assumed that SNR values are integers between 10 dB and 18 dB, average packet loss rates for each SNR is determined, and maximization is performed as in (3.4).

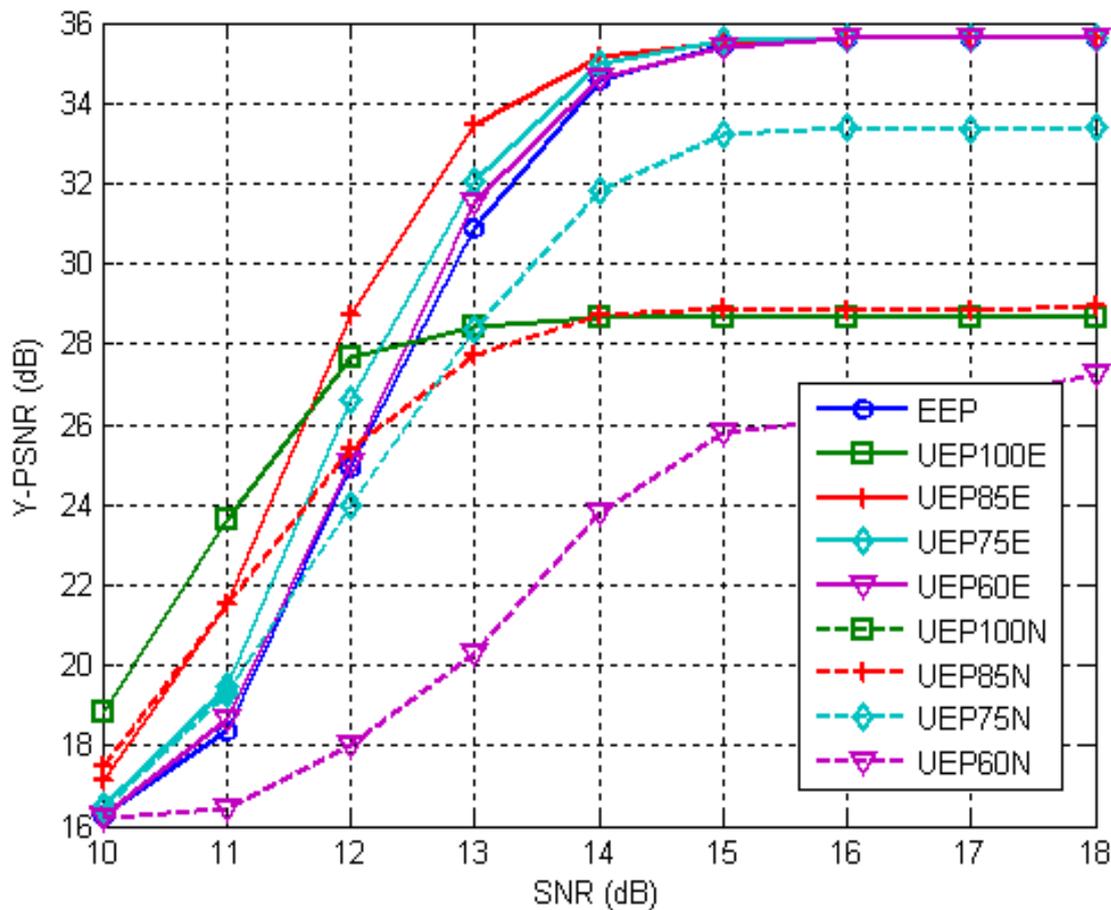


Figure 3.6: PSNR vs. SNR in the DVB-H channel for four different NOW and EW UEP schemes and the single-layer EEP scheme of RLC-IAB at data rate of 383 kbps for Foreman sequence.

The PSNR results obtained by varying the SNR over its entire available range and keeping the data rate fixed at 383 kbps are shown in Figure 3.6 for RLC-IAB. Similar conclusions as for the RLC-IA scheme can be taken. NOW schemes again lag behind EW for higher SNRs.

It can be seen from the two figures that the optimal EW2 UEP(85,15) scheme is superior to all the other schemes including EEP, except at low SNRs, when UEP(100,0) is better for only up to 1dB. This justifies the proposed optimization strategy and shows robustness of the EW RLC UEP schemes over large range of SNRs.

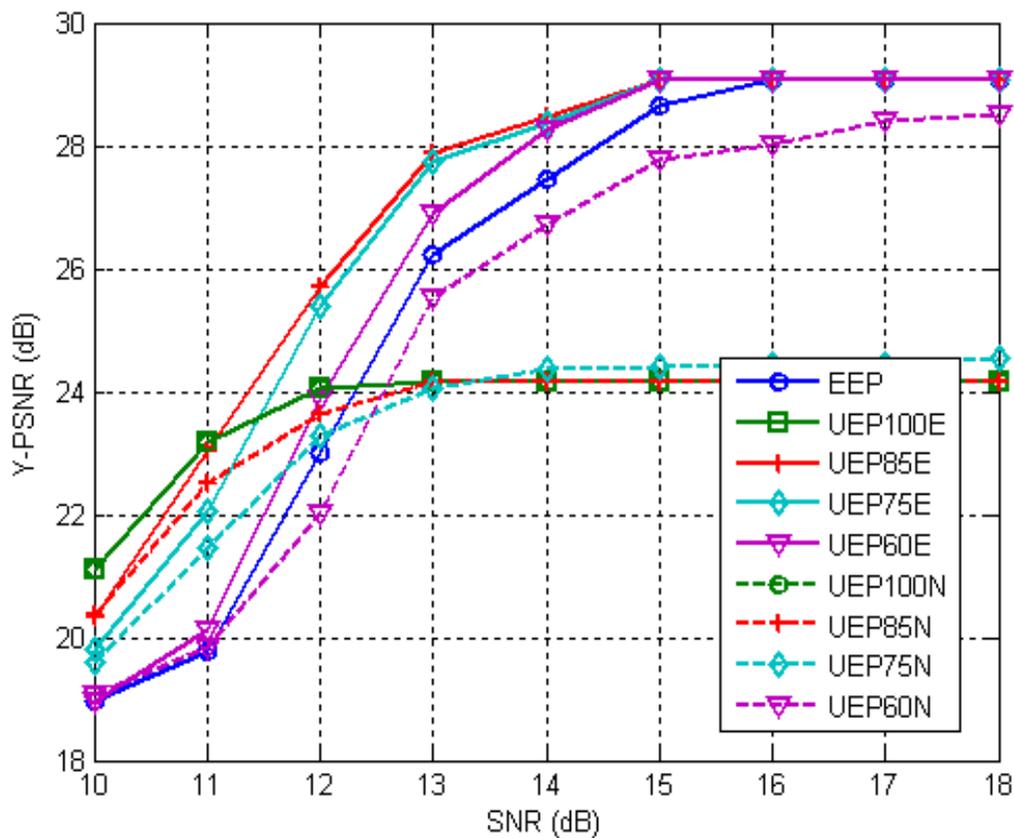


Figure 3.7: PSNR vs. SNR in the DVB-H channel for four different NOW and EW UEP schemes and the single-layer EEP scheme of RLC-IAB at data rate of 383 kbps for Coastguard sequence.

The Coastguard sequence is encoded with same configurations. The comparison of various schemes as in Figure 3.6 is next repeated for Coastguard sequence. The results are as shown in Figure 3.7. The results confirm the analysis as carried out for Figure 3.6.

Figure 3.8 shows frame-averaged PSNR vs. transmission data rate for EW1 and EW2 (RLC-IA) at two different SNRs. Single-layer EEP refers to EW1 scheme in Table 3.5, whereas EEPTS12 and EEPTS14 refer to an EEP scheme at SNR=12 and 14 dB, respectively. Similarly, UEP60TS14 represents a UEP scheme with 60% selection probability for HPC at SNR = 14 dB. The results for UEP75TS12/14 = UEP(75, 25) and UEP(85, 15) provide significant improvement over the single-layer case. As can be seen from the figure, UEP(85, 15) provides better results overall, whereas UEP(100, 0), in which only W1 is protected, provides good results only at low data rates. This is expected, because in this case, there is not enough bandwidth to transmit both classes, thus, it is better not to waste the bandwidth in trying to recover the LPC and use all the resources to ensure recovery of the HPC. Note also that UEP(100,0) reaches almost saturation for higher rates since then, HPC is always delivered and the remaining resources are not used.

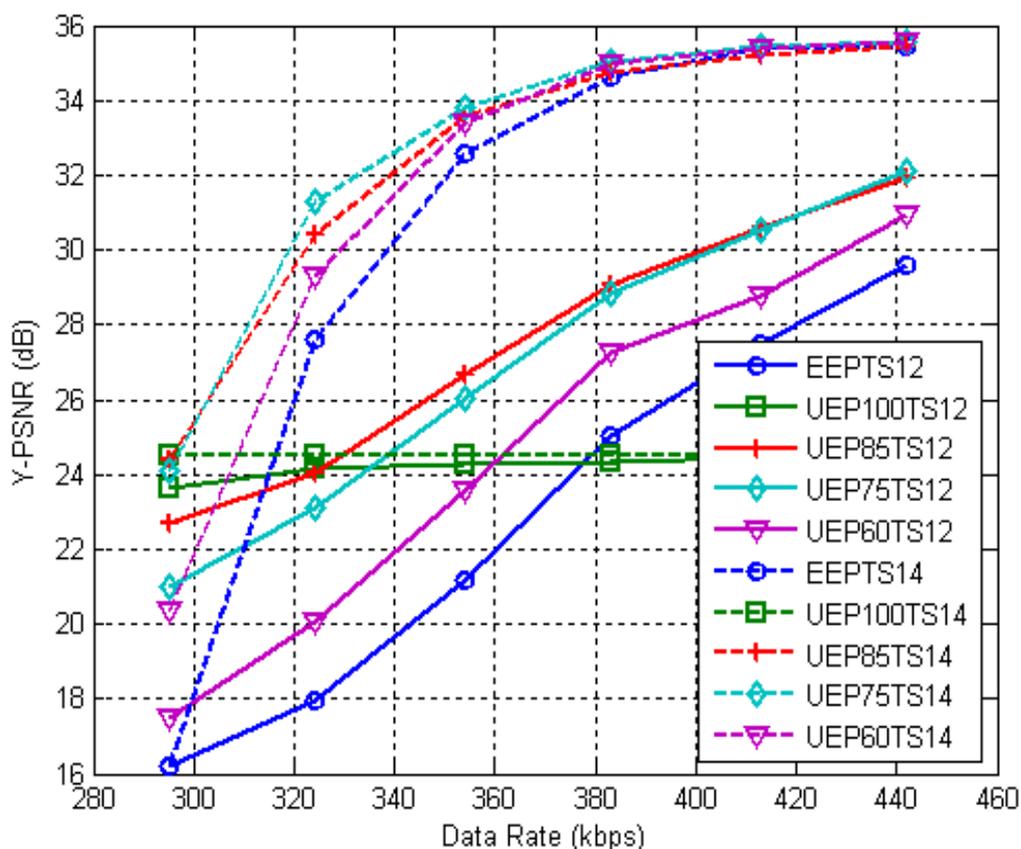


Figure 3.8: PSNR vs. data rate for EW2 (RLC-IA) and EW1 (single-layer) at SNR 12dB and 14dB for Foreman sequence.

Similar results were obtained for Coastguard sequence and are shown in Figure 3.9. Except for the fact that the overall PSNR for Coastguard sequence is lower as compared to the Foreman sequence, the general performance of the used configurations is the same. This confirms the analysis and results as given for Figure 3.8 for Foreman sequence.

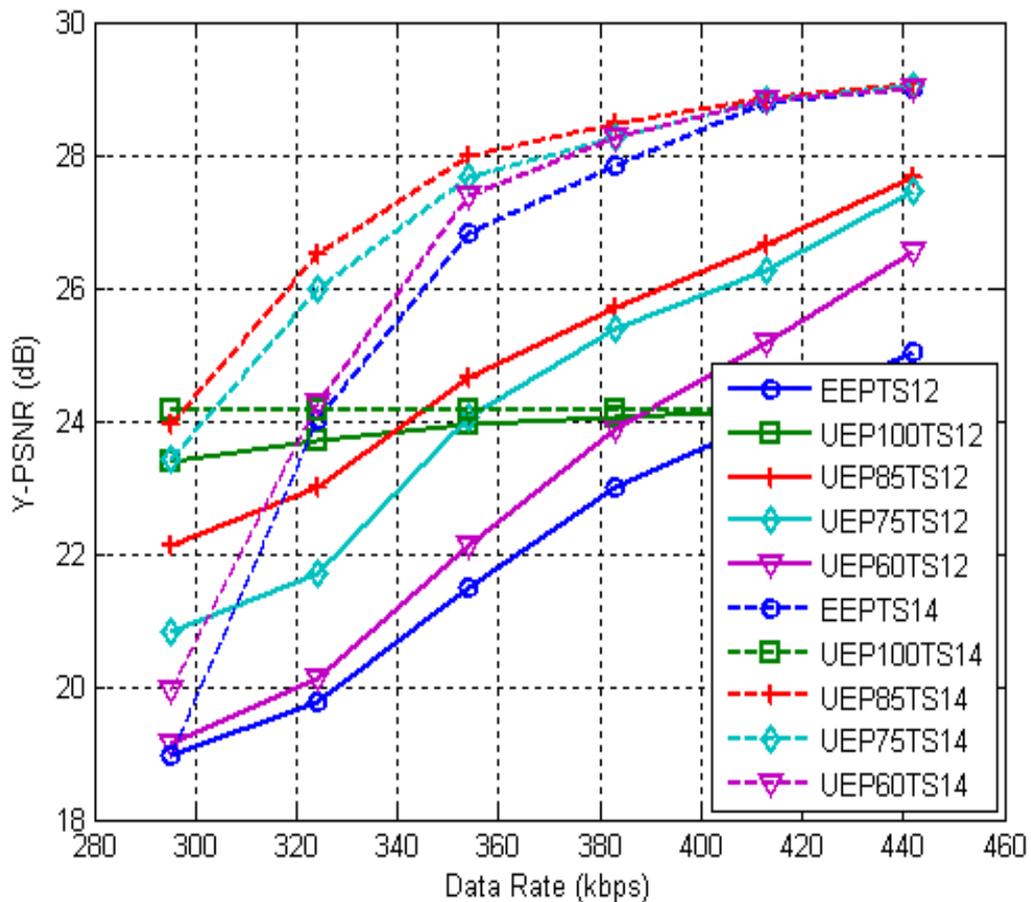


Figure 3.9: PSNR vs. data rate for EW2 (RLC-IA) and EW1 (single-layer) at SNR 12dB for Coastguard sequence.

The PSNR at various transmission data rates for EW3 and EW1 is shown in Figure 3.10. In the EW3 scheme, IDR and all DP A are placed in the first class, all DP B in the second, and all DP C in the third. As can be seen from the figure, UEP(58, 20, 22) performs the best overall. In this case, protecting W1 alone, which is smaller as compared to the EW2

case, assumes an almost horizontal alignment, which indicates that it is not affected adversely by decrease in data rate. At low data rates, the performance of UEP schemes is better in general than that of the single-layer scheme. Note that the PSNR performance of EW2 is slightly better than that of the EW3 schemes. However, EW3 provides fine grain control over the transmitted data rate.

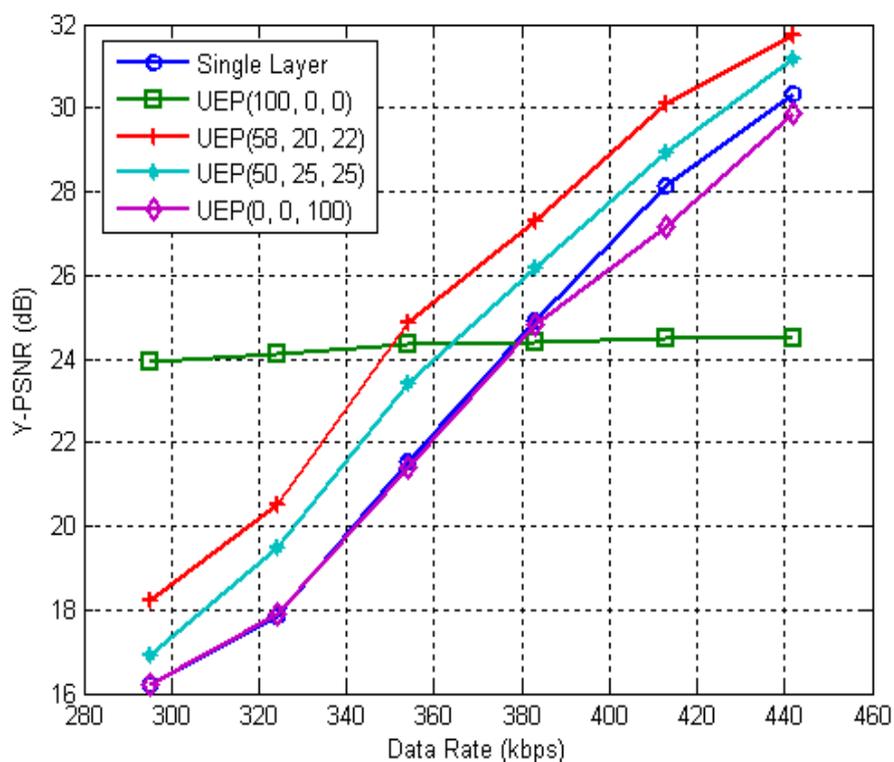


Figure 3.10: PSNR vs. data rate for EW3 and EW1 (single-layer) at SNR 12dB for Foreman sequence.

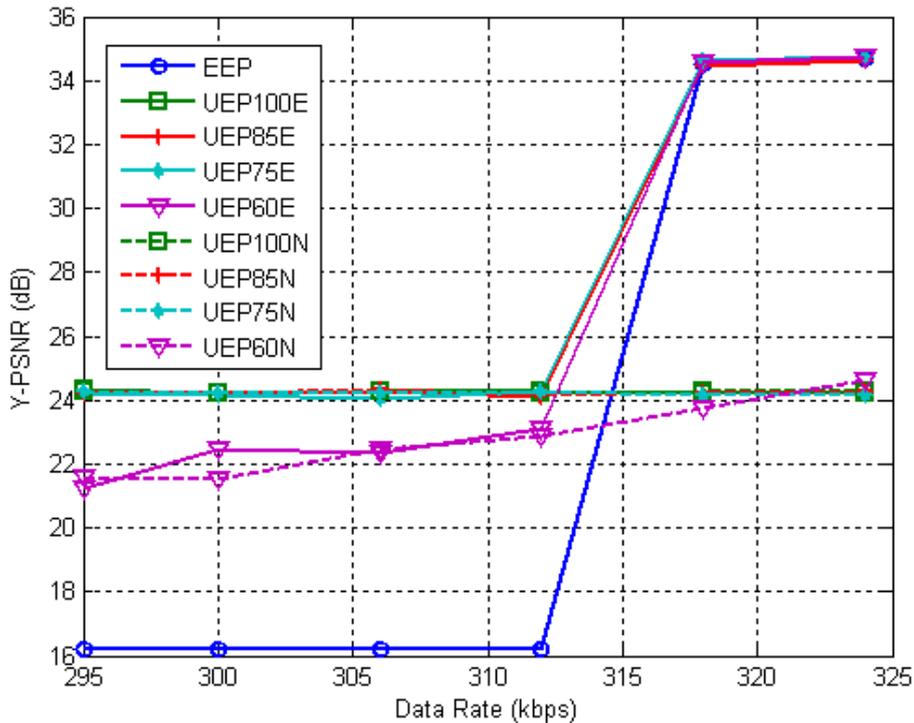


Figure 3.11: PSNR comparison between different RTC and RLC configurations: Y-PSNR vs. Data rate in the DVB-H channel for Foreman sequence.

3.5 Performance Evaluation of EW-RLC based Solution with DVB-T2

DVB-T2 Lite profile is intended to allow simpler receiver implementations for very low capacity applications such as mobile broadcasting, although it may also be received by conventional stationary receivers. The SNRs for the simulations have been obtained through DVB-T field measurements performed in Turku in August 2011, where each SNR value was measured with 250 ms intervals (corresponding to the maximal DVB-T2 frame duration). The length of the SNR series corresponds to 1040.5 s of reception of a DVB-T2 signal. The SNRs have been used for running DVB-T2 simulations with the following DVB-T2 Lite configurations: 16-QAM and 64-QAM modulations, 16200 bits LDPC codewords with rate 1/2, 8k FFT size, 1/4 guard interval, and TU-6 channel model with perfect channel estimation.

With these settings, each time interleaver (TI) frame contained 64 or 97 FEC frames for the 16-QAM and 64-QAM simulation scenarios, respectively. The FEC frames have been

decoded and a baseband frame error trace has been collected. The baseband frame error traces have then been converted to TS packet error traces, taking into account that the size of each baseband frame is 7032 bits.

Figure 3.11 shows the frame-averaged PSNR vs. data rate for EW1 and EW2 (RLC-IA) for 64-QAM. In case of NOW-RLC except for UEP60N, all schemes suffer an overprotection of HPC which is detrimental to overall performance. Hence, these schemes get constrained to a PSNR performance of about 24.51 dB, the best achievable with HPC alone. For EW-RLC schemes, the performance is poor at 295 kbps and thereafter all the schemes perform equally well at the increased data rate. Similar results were obtained for the Coastguard sequence and are shown in Figure 3.12.

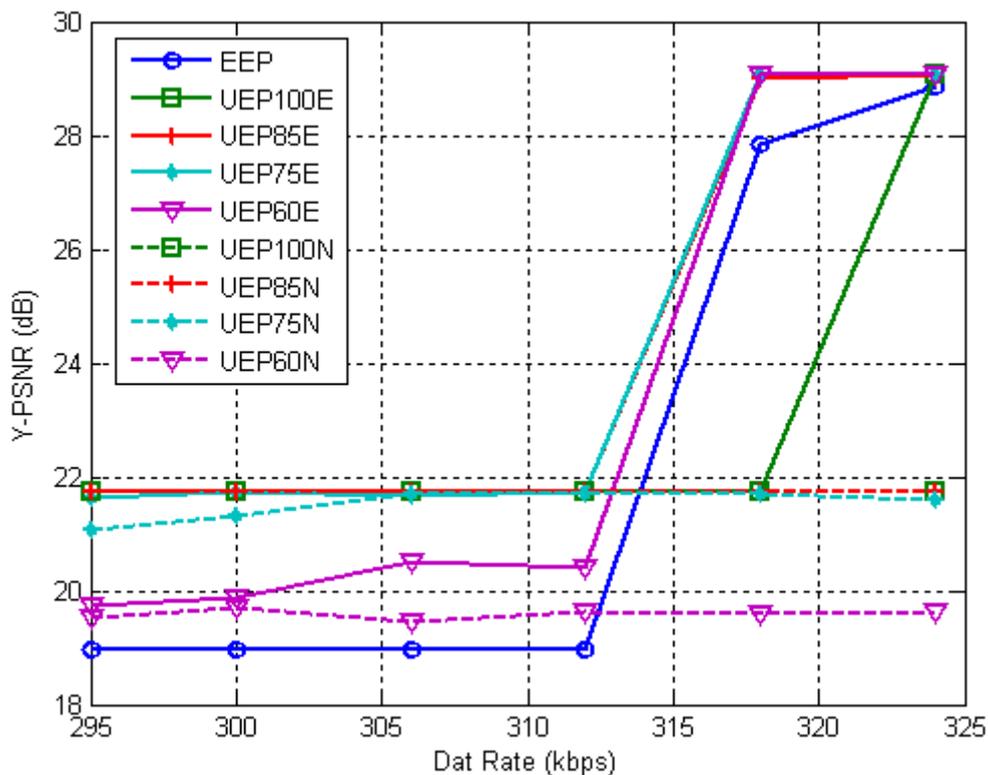


Figure 3.12: PSNR comparison between different RTC and RLC configurations: Y-PSNR vs. Data rate in the DVB-H channel for Coastguard sequence.

3.6 Slice Partitioned H.264/AVC Video Broadcasting with UEP RLC

In this section, the video data is partitioned based on the slicing feature of H.264/AVC in order to provide UEP with EW-RLC. The process of packetization etc is same as that for DP.

3.6.1 Overall System and Measurement Setup

In order to enable progressive recovery, the video sequence is encoded with the slicing feature with six slices per frame. The MBIU feature is used to additionally limit the effect of error propagation and each slice contains a fixed number of MBs. After the H.264/AVC encoding, the video data in which each frame including the Instantaneous decoder refresh (IDR) is divided into six slices is obtained. The priority of each slice is obtained by dropping it from the GOP data and measuring the resulting PSNR by actual decoding.

The slice sorting can also be done as an analytic process as described in [59], which explains the Slice sorting by relevance (SSR) algorithm and how it can be used to prioritize each slice. After the slice sorting in [59], different priority slice partitions are protected with RS codes.

The loss of a slice has an error propagation effect which adds to the GOP distortion, but if the previous frame is similar (low motion), such effect is minimal. This fact is used to create a prioritized and segmented video data. The cumulative PSNR of the GOP is measured with one slice dropped out, starting at the first P frame. The GOP structure is always IPPPP... After having obtained the cumulative PSNR values for each slice, the deviation from the full decoding PSNR of the GOP is measured as shown in Figure 3.13 for the first GOP (having 64 frames) of the standard CIF Foreman sequence. As can be seen from Figure 3.13, some slices have very limited contribution. It is such slices which can be assigned less protection in order to afford a higher degree of protection for the more important slices. On the other hand, loss of one slice causes a PSNR drop of 3.5 dB.

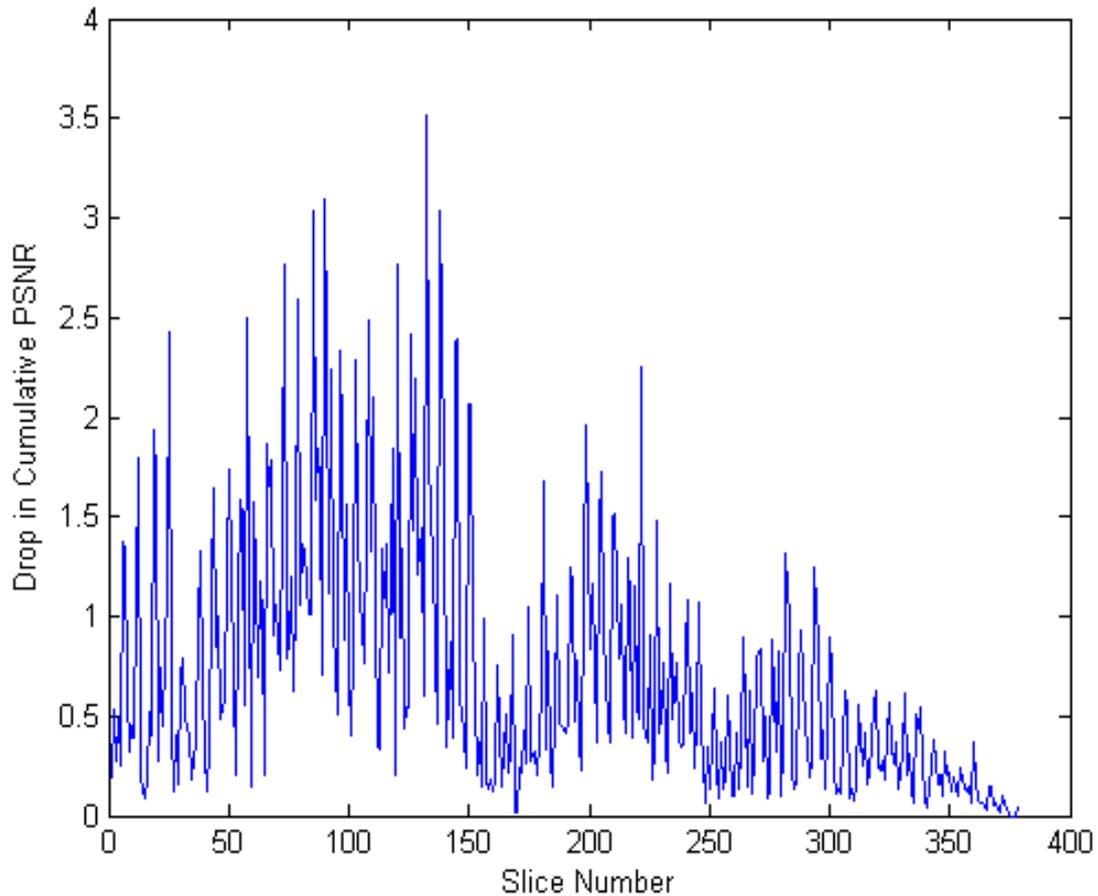


Figure 3.13: Drop in Cumulative PSNR vs a missing slice for the first GOP of Foreman sequence. The PSNR without missing slices is 35.63 dB.

The scheme that is proposed here is to choose a threshold T based on the PSNR contribution. All slices, whose contribution is above the threshold, get assigned to the higher priority segment/layer. Similarly, creation of N segments requires $N-1$ thresholds. Additionally, any frame which does not have any slice selected in the high priority layer based on the threshold alone will have at least one representative slice included to offset the effects of error propagation and to improve PSNR. Such frames occur mostly towards the end of GOP as can be seen from Figure 3.13. Note that the process of slice prioritization is of low complexity since the encoder already has all the information used for rate-distortion optimization.

In the proposed EW2 RLC scheme, with a single threshold T , two windows, $W1$ and $W2$ are created. In the proposed EW3 scheme, two threshold values are utilized to create three windows, $W1$, $W2$, and $W3$. Note that for both these schemes, all slices of IDR frame are placed in $W1$ being the most important data for video reconstruction.

The data in each window is divided into equal-length source symbols. The length of the source symbols determines the number of source symbols contained in the source message, k . The symbol size has been set to 1394 bytes because large source symbols are preferred as they yield small values of k which is necessary to keep the RLC decoding complexity acceptable. RLC configuration with 1394 symbol size achieves the best trade-off between performance and decoding complexity over a range of transmission rates.

After AL-FEC coding, one encoded symbol (together with RTP/UDP/IP headers) is placed in an IP packet. The IP packets are then placed into the MPE frames where each IP packet is encapsulated within a single MPE section with its own overhead containing the error-detection field. The MPE frame is transmitted within a single DVB-H physical layer transmission burst.

Table 3.6: Relative Partition Sizes for the First GOP of the Foreman Sequence.

Partition	Size (bytes)	Size (%)	Cumulative PSNR [dB]	EW3	EW2
SP1	52,980	55.65	22.09	W1	W1
SP2	15,320	16.16	26.84	W2	
SP3	26,480	28.20	35.63	W3	W2
Total	94,780	100	35.63	94,780 bytes	

3.6.2 Results and Analysis

In all the simulations the video sequence *Foreman* in the CIF format is encoded using the H.264/AVC software JM version 16.2 [54]. All the simulations are performed using a GOP size of 64 frames, with frame structure IPPPP..., six slices per frame with 66 MBs per

slice, and a frame rate of 25 frames per second (fps). This configuration implies that each GOP has an average size of 94,400 bytes corresponding to 2.56 sec of video data, and thus the source data rate is roughly 295 kbps. This ensures that sufficient amount of video data is transmitted in one burst of the DVB-H transmission. The partitions with their relative sizes and the PSNR contribution are shown in Table 3.6. The partitions SP1 to SP3 are obtained using threshold values on the PSNR difference from a full decoding (without any losses) PSNR.

Table 3.7: Configurations and the Window Structure.

Configuration	Sym. Size (bytes)	Window	No. of Symbols	No. of IP Packets	No. of TS Packets
EW2	1394	W1	49	49	392
		W2	19	19	152
EW3	1394	W1	38	38	304
		W2	11	11	88
		W3	19	19	152

Two windows (EW2) and three windows (EW3) are created by grouping the partitions as shown in the table. The GOF video data is divided into the source blocks containing 68 symbols, which are then assigned to various windows as shown in Table 3.7. The symbol size used and the number of symbols in each window for the different schemes are also shown.

The configurations shown in Table 3.7 are simulated with 1000 runs each for DVB-H channel SNR equal to 12 dB. To model the channel accurately, TS packet losses are represented based on trace files for different channel SNRs for the DVB-H channel with real-world measurements conducted at [52]. The simulations have been performed for different data rates to evaluate the EEP and UEP performance of EW2 and EW3. The base data rate is 295 kbps, and the successively higher data rates are obtained by successively adding 10% additional symbols. The highest rate obtained is 442 kbps, which is 1.5 times higher than the base rate.

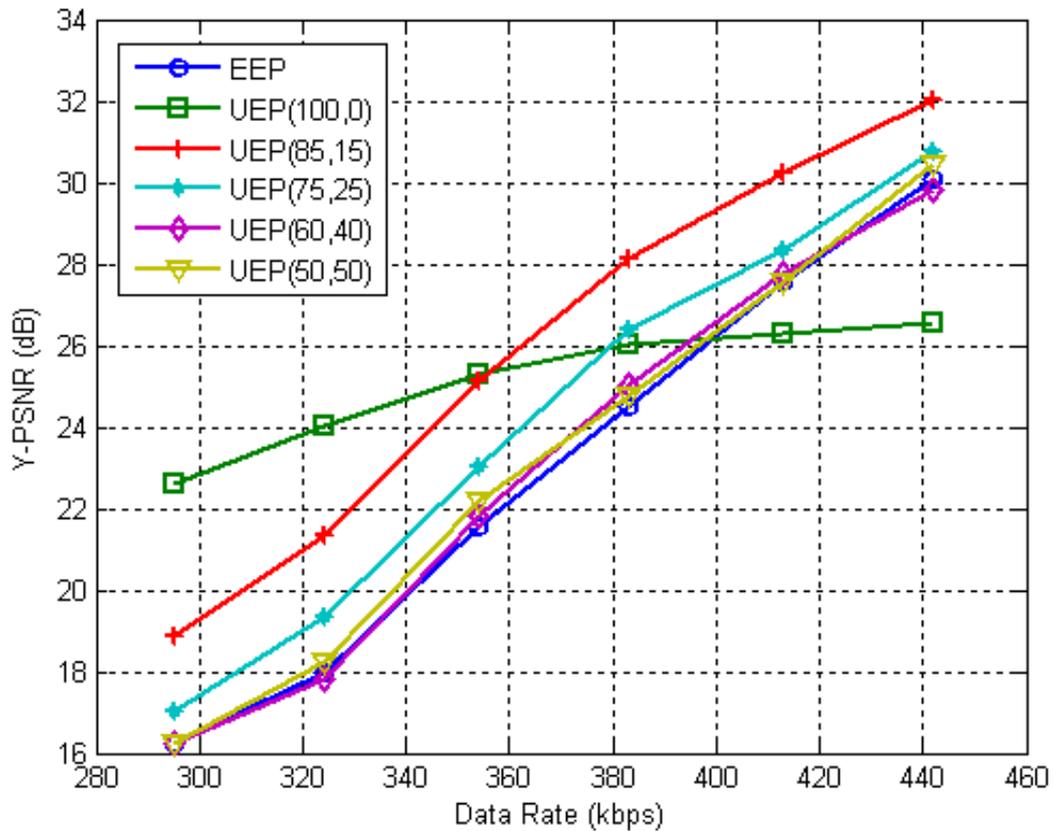


Figure 3.14: PSNR for EW2 at SNR 12 dB.

In case when not even the first window W1 gets decoded, the entire GOP is considered to be lost. The PSNR for such cases is obtained by using the last frame of the previously decoded GOP to replace all frames of the lost GOP. The various configurations are used to create different UEP schemes based on protecting the constituent windows with different protection, based on probabilistically selecting a window for each output symbol at the transmitter. An increase in the selection probability of W1 will improve its robustness at the cost of a decrease in robustness of the succeeding layer(s). The EEP scheme is the case where only the largest window is selected with 100% probability. This means that all of the data is protected with no preference for the data considered important, i.e., window W1.

Figure 3.14 shows PSNR vs. data rate for EW2. The numbers shown in brackets represent the selection probability of each of the two windows, e.g., UEP (85, 15) represents a code in which a symbol from W1 will be selected for transmission with probability 0.85.

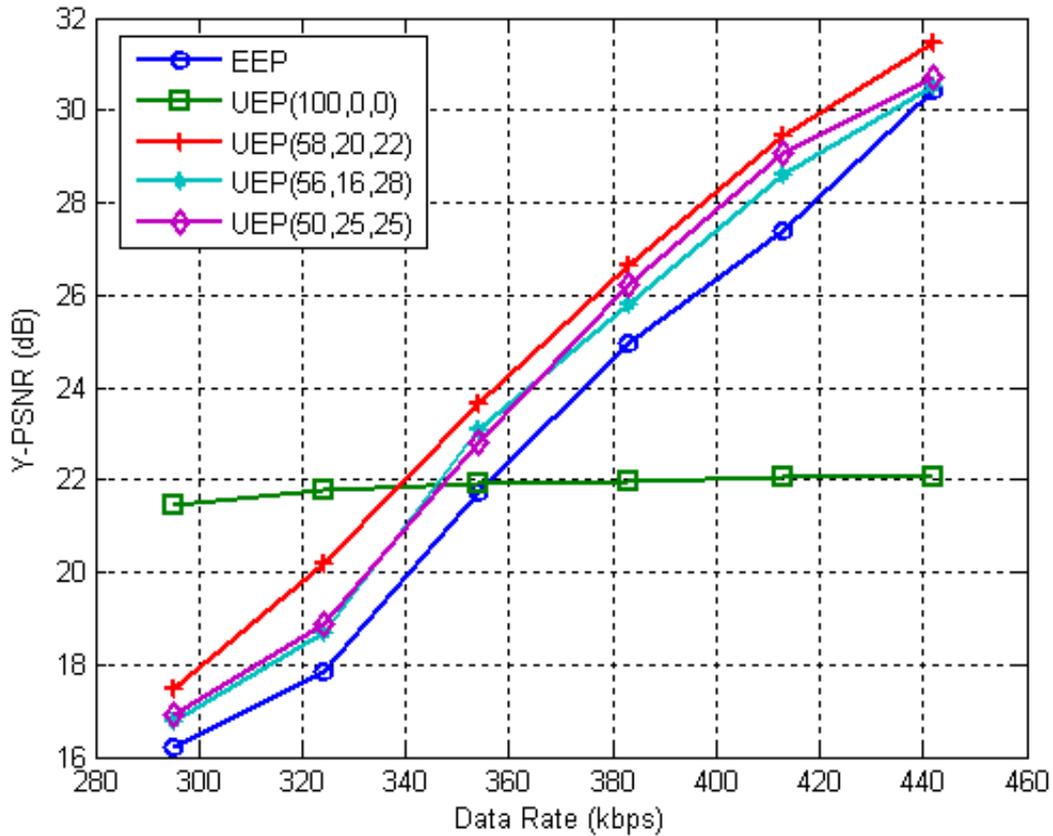


Figure 3.15: PSNR for EW3 at SNR 12 dB.

The performance of UEP (75, 25) and UEP (85, 15) is much better as compared to the EEP. As can be seen from the figure, UEP (85, 15) provides better results overall for higher data rates. UEP (100, 0) is a scheme in which only W1 is protected. The scheme is constrained in that it cannot achieve any better than 26.84 dB (see Table 3.6). The results are thus better than other schemes only at low data rates. However, the decoding failures, i.e., when the GOP data fails to be decoded, are much less for UEP (100, 0), which is good as this ensures that for each GOP, data is received, though at basic quality level.

The PSNR at various transmission data rates for EW3 is shown in Figure 3.15. As can be seen, UEP (58, 20, 22) performs the best overall. The UEP (100, 0, 0) wherein only W1 is protected at the expense of W2 and W3, is almost horizontal, which indicates that it is not affected very much by a decrease in data rate. The EEP scheme performs poorly as compared to the UEP schemes. The PSNR performance of EW2 is slightly better than the EW3 case. However, the EW3 case provides a fine grain control over the transmitted data rate.

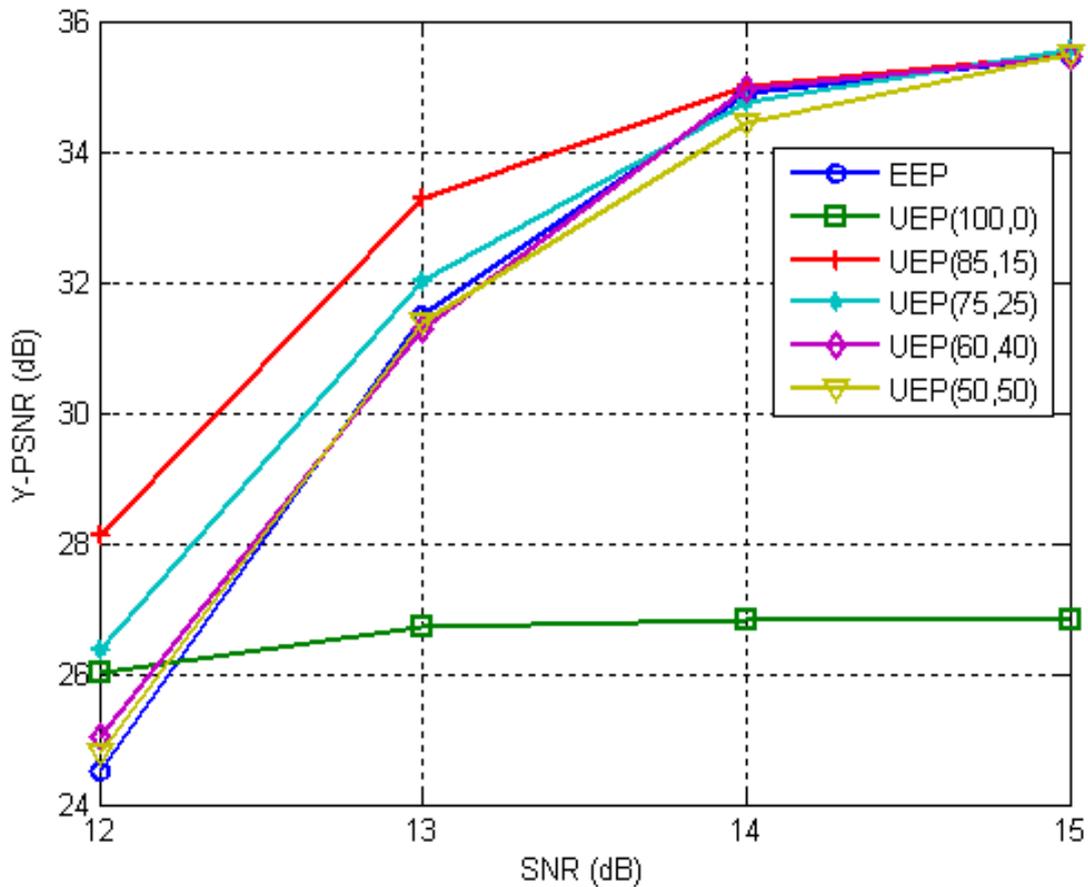


Figure 3.16: PSNR at different SNRs for 383 kbps.

Due to small code length for RLC68, the decoding complexity is minimal, which makes it a good candidate for real-world implementations such as the one described in [20]. The results of Figure 3.14 and 3.15 are similar to those in earlier sections where DP feature of H.264 encoder is used to create the windows.

The PSNR for EW2 at 383 kbps for various SNR values is shown in Figure 3.16. At SNR 12 dB, some schemes perform worse than UEP(100,0) which is only protecting W1, which signifies that with such low probabilities of selection, i.e., 50 and 60%, even W1 is not being protected adequately. On the other hand, UEP(85,15) has better performance even at such low SNR. All schemes reach highest PSNR of 35.63 dB at SNR 15dB. The degree of protection provided with selection probability in UEP schemes can thus be made dynamic to change on GOP basis.

3.7 Summary

The study has extensively compared the performance of RLC and Raptor codes by assigning different degree of protection to the data partitions of H.264/AVC video under different packet loss scenarios at different data rates. RLC are seen to perform very close to or somewhat better than Raptor codes. Within RLC configurations, it is seen that the scheme with less number of symbols results in better performance. This is promising result for the use of RLC for multimedia broadcast applications, as the decoding complexity for such configurations is manageable.

The average PSNR of UEP schemes with NOW configuration is not as good as that of single-layer AVC at high data rates. However, it is interesting to note that at lower data rates UEP schemes provide good results with smaller number of decoding failures.

The results and analysis are also substantiated by using another of error resilience feature, i.e., slicing. By choosing similar configurations as that of DP the results with slicing are provided which confirm the earlier results. For the slice-partitioned schemes the performance of RLC codes using EW-RLC has been analyzed and compared to the case of single-layer coding. EW2 and EW3 both perform up to 5 dB better compared to the single-layer case and show robustness to channel fluctuations. The EW3 configuration provides better degree of control to match a particular data rate and erasure channel characteristics, while EW2 provides slightly better performance overall.

The performance of EW-RLC schemes is significantly better than the corresponding NOW-RLC schemes. The number of symbols per codeword used was 68, which is a very promising result for the use of RLC for multimedia broadcast applications, as the decoding complexity for such configurations is manageable.

In the next chapter, the DP and slicing schemes described here are used to propose a rate-adaptive solution for reliable video transmission.

Chapter 4

Rate Adaptive Selective Segment Assignment (RASSA) for Reliable Wireless Video Transmission

4.1 Introduction

In this chapter, a reliable video communication system is proposed which uses DP and slicing feature of H264/AVC to create a layered stream, and LT codes [8] for erasure protection. The proposed system does not target any specific technology and is appropriate for unicast scenarios. The proposed scheme is an adaptive low-complexity solution to varying channel conditions. The comparison of the results of the proposed scheme is also provided for slice-partitioned H.264/AVC data. Simulation results show competitiveness of the proposed scheme compared to optimized unequal and equal error protection solutions. The simulation results also demonstrate that a high visual quality video transmission can be maintained despite the adverse effect of varying channel conditions and the number of decoding failures can be reduced.

Fountain codes [7], [9] are rateless and in non time-constraint applications can generate as many encoded packets as needed. The amount of additional packets transmitted is the redundancy which is necessary for decoding to succeed and can be adjusted to combat different channel conditions. In bandwidth-limited wireless networks it is important to keep the introduced redundancy to a minimum. Thus instead of targeting the worst possible channel conditions, the redundancy should be adaptively adjusted according to the varying channel conditions via dynamic source-channel coding. LT codes are used in this chapter due to their design and implementation simplicity. However, LT codes have a higher computational complexity, $O(k \log_e k)$ (k is the message length) than Raptor codes, $O(1)$.

Priority encoded transmission (PET) [60] is another FEC scheme which can be used to provide UEP for video data. The message gets encoded into packets depending on the laid down priorities and thus provides an effective solution to offset errors.

In this chapter four schemes are evaluated, namely fixed-source rate EEP, fixed source-rate UEP, PET, and source rate-adaptive schemes in the context of H.264/AVC data

prioritization. Different error resilience options are used to generate encoded segments that can be dropped to adjust the rate or provide UEP based on their importance. Error-resilience and concealment features are utilized, which are designed to make the video less vulnerable to the effects of lost data, and then the video is compressed at a higher source rate allowing for some decoding errors.

DP [2] is a low cost error resilience feature, supported by the extended AVC profile, which can be exploited to introduce a layered structure in H.264/AVC. Besides DP, it is possible to partition a frame into a fixed number of *slices*, which are of different importance to the video reconstruction. Thus, similar to DPs, the slices can be aggregated into different priority classes, with the higher priority classes containing slices that have higher contribution to the reconstruction. Such prioritization can make the sliced video data amenable to UEP and rate adaptation. DP has low overhead as its structure is determined in advance, whereas slicing generally requires a slice group map.

In this study the data in various partitions, based on DP or slicing, is treated as segments. Then, based on the available channel conditions, namely, the available bandwidth and the estimated wireless channel loss rate, the data from each of the segments is selected for FEC coding. The optimized EEP, UEP, and PET are compared with the source-adaptive scheme that is termed rate adaptive selective segment assignment (RASSA), and pros and cons of them are identified in channel mismatch scenarios.

A lot of work has been done on joint source-channel coding; see [61] for a review. In the domain of rateless source-channel coding, in [21], a class of UEP codes, called Expanding Window Fountain (EWF) codes is used for UEP of scalable video. The UEP with LT codes was first proposed in [65]. In [62], unequal protection has been proposed for video communications by duplicating the information symbols. In [63], unequal Growth codes have been proposed in which as the number of packets the receiver has increases, the degree for each new encoding symbol also needs to increase, hence the name Growth codes. The application of PET to UEP video data is provided in [66]. An adaptive rateless coding for DP AVC coded video has been proposed in [64]. The system uses intra-coded macroblocks (MBs) in each frame; some additional redundant data is piggybacked onto the on-going packet stream. In [67] channel coding is combined with additional duplicate packets.

The proposed method differs from the earlier work in that it increases the protection for the transmitted data by not selecting whole of the data for transmission. A scheme has been proposed to decode video even when the rateless decoding fails using packetization information. This is made possible by passing a video-table to the decoder containing the DP type and size information. Thus, the DPs with all or part of their data missing are discarded before the H.264 decoder tries to decode the data. It is important to note that without such information the decoding will fail on encountering such missing data.

The segmentation of video data facilitates a layered coded video that might be preferable to the H.264 Scalable Video Coding (SVC) extension [4] in some applications, since it complies with the AVC standard, provides scalability, and more robust output to packet losses than SVC. The proposed scheme can be applied in multicast scenarios with heterogeneous receivers, in which case a receiver can terminate reception and decoding of segments after having received data compatible with its processing power and memory.

The rest of the Chapter is organized as follows. Section 4.2 covers the proposed system. The proposed rate allocation algorithms are described in Section 4.3. The results and analysis are in Section 4.4. Finally, the summary is contained in Section 4.5.

4.2 Proposed System

In this section the proposed system is described that segments encoded video and provides equal or unequal error protection. It is to be noted that an LT packet size of 70 bytes is proposed, thus when an LT packet is lost, the corresponding DP/slice needs to be dropped as well. The DPs/slices which are received and hence decoded may have some intervening DP/slice missing. Although the results are not provided for comparison of the proposed prioritization scheme with other types of channel coding schemes, for example, using different types of frames in a GOP with more protection allocated to more important frames. However, it can be argued that such frame-based segregation into UEP classes will provide poor results as compared to DP/slice based classification. The reason for this is that with such frame based classification, loss of a single LT packet from within a frame will potentially make it useless. The probability of an LT packet being missing from a frame is more (at least three times) as compared to DP/slice based partitions. Moreover, DP/slice partitioning being

error-resilience schemes, they also mitigate the error propagation to subsequent frames, whereas comparably a lost frame will have more drastic degrading effect.

First, a system is described that forms layered outputs using the DP feature. Then, the system that exploits slicing instead of DP is presented.

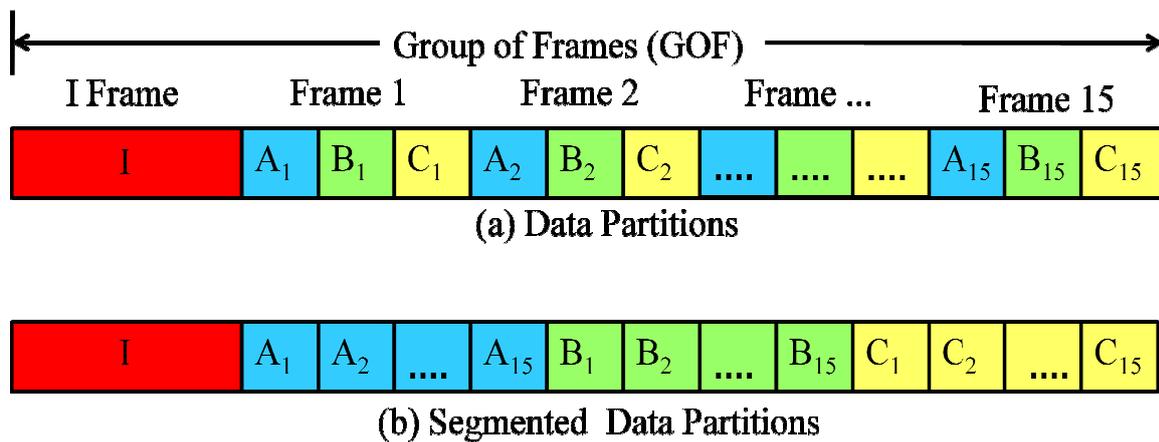


Figure 4.1: (a) Data Partitions. (b) Segmented Data Partitions.

4.2.1 Protection of DP-based AVC Video

The IDR frames are used without any partitioning. The video data of each non-IDR frame is divided into three data partitions by the H.264/AVC encoder. This partitioned data needs to be aggregated together to enable UEP. The structure of a segmented video is shown in Figure 4.1. The figure shows the DP A, B, and C together with IDR frame. (Note that the first non-IDR frame is denoted as A₁, B₁, C₁, and so forth.) Next, the partitions are prioritized and all DP As, Bs, and Cs are grouped together to form effectively three segments or layers as shown in Figure 4.1 (b). Note that by receiving only I/ IDR frame and DP A₁-A_n, the decoder will still be able to decode all *n* frames within the GOP, though at reduced quality. Further segmentation is not restricted to be done at the aggregate partition boundaries only. That is, if all IDR and DP A are sent as the first segment, then any number of DP B and DP C partitions can be selected for transmission in the second segment. This provides

flexibility that enables a fine-grained layered structure as a large number of reconstruction rate points become available, which can be matched to the channel statistics with a very fine control over video reconstruction quality. The DP B and DP C by virtue of having been aggregated are already in their priority order for reconstruction. The layer with important data (IDR and DP etc.) is termed as High priority layer (HPL), whereas the remaining data is placed in Low priority layer (LPL). In the proposed scheme, intra-refresh MBs are not used but instead periodic I frames are assumed.

The segmented data partitions are next protected by FEC codes applied on each GOP independently. To achieve UEP, each segment should be protected according to its importance using different amount of redundant symbols. To accomplish that, the FEC encoding process adds an important initial step, that is, to first select a segment from which the encoded symbol is to be generated determined by “selection probability” of a segment which is a pre-assigned parameter based on the importance of different segments and the data rate available. After a segment is selected, a conventional encoding is performed over the source packets contained in that particular segment only. Thus instead of defining a UEP scheme as a set of rates (one for each segment), it is equivalently defined by a set of selection probabilities. This resembles the method of [21]. For practical reasons the number of layers in the UEP is usually constrained to two or three.

Note that the UEP scheme allocates redundancies to the segments based on their importance. The optimal rate allocation depends not only on the channel characteristics but also on video data since the importance and sizes of the segments vary from one GOP to another GOP. Thus, the UEP has to be dynamically changed and optimal allocation needs to be found for each GOP, which is practically feasible only for a pre-recorded video. Note that in the extreme case when the bandwidth is very scarce or packet loss rate is high, which is often in mobile wireless scenarios, the optimal selection probability of low-priority segments would be zero and all redundancy would be allocated to the high-priority segments to ensure their successful decoding.

Motivated by this and targeting wireless applications with limited bandwidth available and high loss rates, another scheme is introduced called the RASSA scheme. The RASSA scheme is a special case of UEP that exploits the flexibility of layered coding of DP and slicing. First, given an estimated packet loss rate and total rate budget, the system calculates

the required overhead (and thus also the amount of source data) that will allow for error-free transfer with high probability (w.h.p.). Then, the data is filled starting from leftmost in Figure 4.1 (b), and remaining source data is discarded. This way, the scheme discards some of the lower-priority data by assigning it zero selection probability, to increase protection of the more important data.

Thus, this scheme is not constrained in having two or three segments/layers, and any number of DPs/slices can be selected enabling a very flexible rate control. For example, given channel statistics, enough redundancy can be provided for a segment containing DP A and B and part of DP C to be recovered at the decoder w.h.p. The unselected low-priority data (remaining DP Cs) are simply discarded. Note that the entire sent source block will be decoded, or decoding will fail, in which case the previous GOP is used for reconstruction.

RASSA can be seen as a UEP scheme since it protects only one part of the encoded data and discards the rest, but also as EEP since it provides equal protection of all sent source data. One immediate advantage of this scheme is reduced complexity since only one code is used, where UEP generally requires one code for each layer, and there is no need for complex rate optimization. Indeed, once the channel loss rate is estimated, the required code rate is set, and based on the available bandwidth (total budget) the decision to drop some of the NAL units that cannot fit the total budget is made.

UEP schemes require that the DPs of each type in LPL are aggregated together. To pass this information to the decoder, a video-table structure is proposed to be created at the encoder. The encoded video generated by the H.264/AVC encoder with DP is used to create a video-table with an entry for each NAL unit and its length. The number of NAL units per GOP is usually small (up to 64), and hence the table can conveniently be passed to the decoder within a header with negligible rate increase. The packet bearing the header will be transported with HPL. If HPL is lost then otherwise no video decoding is possible. At the receiver side, the video-table structure is used to rearrange the DPs to their original encoding order. The table is also used to discard NAL units with missing data. That is, since one DP/NAL can be sent in multiple packets, if one packet is missing the entire DP is dropped. Also, recovered DP B and DP C of a frame are dropped if DP A for that frame is not recovered properly.

There is no latency involved in bringing the DPs to their original order for decoding. The aggregation of DPs is only limited to a priority layer. For instance, if DP A and DP B both are in HPC then they will remain in their original encoding order. The number of DPs which are dropped in RASSA scheme are only from LPC and their number is limited. Thus, it is only limited LPC data which will need to be inserted in their encoding order. This is not likely to have any latency effect. If required, the data could simply be decoded in place by the decoder.

4.2.2 Protection of Sliced AVC Video

In previous work [68], a method for segmenting sliced-AVC output into multiple segments is proposed and tested based on importance of the slices for reconstruction. For example, two priority classes can be formed where more important slices that contribute to the PSNR level above a fixed threshold are put in the high priority layer (HPL), and all other in the low priority layer (LPL). Then, the protection methods described above (EEP, UEP, and RASSA), can be applied to such prioritized data without modification.

A video sequence is encoded using slicing with each frame divided into a fixed number of slices. The priority of each slice is obtained by dropping it from the GOP data and measuring the resulting PSNR, as a frame-by-frame average of the entire GOP, by actual decoding. Hence this scheme is meant for pre-encoded video. This also takes into account the error propagation effect to the subsequent frames due to loss of a slice in an earlier frame. That is, the cumulative PSNR of the GOP is measured by dropping each slice in turn starting at the first P frame. After having obtained the cumulative PSNR values for each slice (as dropped), the difference from the full-decoding PSNR of the GOP is measured. The importance of the slices on total frame-averaged PSNR generally decreases going towards the end of the GOP. Thus the slices can be sorted into multiple priority layers and assigned a higher degree of protection to the important layers as compared to the layers containing less significant slices. Such layering enables a prioritized data transmission with UEP schemes. Details of assigning slices to different layers can be found in [68].

4.3 Rate Allocation

In this section rate allocation optimization is discussed for the three proposed schemes. DP is assumed; however, in the same way, rate allocation can be done in case of slicing.

Let N be the given total rate budget expressed as the total number of packets/symbols that can be transmitted for each GOP. The video is encoded using DP H.264/AVC forming either four segments: IDR, DP-A, DP-B, and DP-C, or two classes of slices. It is assumed that each segment can be truncated arbitrarily. Let K be the total number of encoded source packets/symbols. Three schemes are considered: (i) an EEP scheme that generates N packets using all K source packets and transmit them over the network; (ii) a UEP scheme that groups source data into L importance layers starting from IDR; for example, it can have $L = 4$ where each of four segments forms one layer; (iii) an RASSA scheme that takes first $K_{RASSA} \leq K$ source packets to generate N transmission packets.

Assuming that video is pre-encoded, K is fixed and is not part of the optimization. Then, EEP scheme always uses an (N, K) code and thus does not require optimization.

An L -layer UEP scheme can be described by L -tuples $\pi = (p_1, p_2, \dots, p_L)$ and $k = (k_1, k_2, \dots, k_L)$, where p_i and k_i represent the selection probability and the size in packets, respectively, of layer i . Then the optimal rate allocation between the L layers can be found by maximizing the expected PSNR of the reconstruction given by

$$\hat{PSNR} = \sum_{i=0}^L P_i(\pi, k) PSNR_i$$

where P_0 is probability that no layer is recovered, and P_i is probability that first i layers can be recovered but not layer $i + 1$, and P_L is the probability that all layers can be recovered successfully. The task is to find L -tuples π^* and k^* that maximize the expected PSNR, over all possible L -tuples π and k . P_i can be obtained experimentally or for some FEC codes estimated analytically for each π , k , and each channel condition and are source independent. For simplicity, it is assumed that k is set a priori by the video encoder, which is often the case. Indeed, it is natural to group all packets from one segment together. For example, for L

= 3, IDR and DP A can be placed into one layer, DP B in another, and DP C in the last layer. Note that the sizes of each segment are determined by the video encoder, and are not subject of the optimization. The problem can further be simplified by maximizing the expected received rate instead of PSNR as:

$$\hat{R} = \sum_{i=0}^L P_i(\pi) k_i$$

where k_i is the number of packets in the first i layers and $k_0 = 0$. This way, the optimization is independent of the source content and depends only on the total rate, layer sizes, and channel loss rate. There are many methods proposed to efficiently accomplish the two optimization tasks (see [61], [21] and references therein).

For the RASSA scheme, recall that out of K generated source packets, only K_{RASSA} are selected that are protected by an (N, K_{RASSA}) channel code before transmission. The optimization problem is simplified to the following. Given a total number of transmission packets N and packet loss rate q , the task is to find the number of sent source packets $K_{RASSA} \leq k$, such that all K_{RASSA} source packets can be decoded w.h.p. Note that determining K_{RASSA} implies the used channel code (N, K_{RASSA}) . Again, the expected PSNR or the expected number of received source packets is maximized, given by:

$$\hat{PSNR} = (1 - P)PSNR_0 + P.PSNR_1$$

and

$$\hat{R} = P.K_1$$

respectively, where P is probability of successful decoding and $PSNR_0$ and $PSNR_1$ are reconstructed PSNR if decoding fails or is successful, respectively. K_1 denotes the number of source packets sent by the RASSA scheme. Note that P depends on q and K_1 and can be found experimentally or analytically. Indeed, for maximum distance separable codes, P is probability that the number received packets is at least K_1 , and then:

$$\hat{R} = \binom{N}{K_1} q^{N-K_1} (1 - q)^{K_1} K_1$$

which can be solved numerically.

In the next section, results of the rate and PSNR- optimized RASSA schemes are compared to that of EEP and optimized UEP schemes.

4.4 Results and Analysis

The robustness of the EEP, optimized UEP, PET, and RASSA schemes are tested when packet loss rates q and data rates N vary. For implementing PET, RS codes are used which do not suffer from overheads as in case of LT coding. The effectiveness of the proposed approach is shown using both DP and slicing features. Simulations have been performed using the H.264/AVC software JM 16.2 [54]. A GOP size of 16 frames is used with the IPPPP... structure.

4.4.1 DP AVC transmission

It is assumed that the video has been encoded at a fixed rate using DP into fixed length segments IDR, DP A, DP B, and DP C. The data in each segment is formed into source symbols/packets of size 70 bytes for the LT coding process, which is a good compromise between performance and complexity. I frame is put in the first NAL unit and it is not partitioned. CIP is used to make the decoding of DP B independent of DP C. Each non-I frame is partitioned into DP A, DP B, and DP C.

Table 4.1: Partitions for first GOP of “Paris” and “Football” sequence.

Partition	Paris		Football	
	Size	Cum. PSNR	Size	Cum. PSNR
IDR	22281	-	23374	-
DP A	12838	30.13	22823	-
DP B	97	30.32	2893	25.39
DP C	45732	39.16	31731	32.62
Total	80948	39.16	80821	32.62

The partitions with their relative sizes and the PSNR contribution for the first GOP of the CIF format “Paris” and “Football” video sequences are shown in Table 4.1. A two-layer UEP scheme is considered where the first, high-priority layer (HPL), contains the selected more important partitions, and the second low-priority layer (LPL) contains the remaining partitions. The UEP schemes are described by UEP(p_1, p_2), where p_1 and p_2 represent the selection probabilities of packets from HPL and LPL, respectively, and the optimal solution can be found as shown in Section 4.3.

Table 4.2: Priority classes and LT packetization for first GOP of “Paris” sequence.

Class	DP	PSNR	Number of LT Packets
HPC	IDR+ DP A	30.13	502
LPC	DP B+ DP C	39.16	655
Total	80,948	39.16	1157

Table 4.3: Priority classes and LT packetization for first GOP of “Football” sequence.

Class	DP	PSNR	Number of LT Packets
HPC	IDR+ DP A + DP B	25.39	702
LPC	DP C	32.62	453
Total	80,821	32.62	1155

The classification of the DPs and the resulting LT packets for the “Paris” and “Football” sequences is shown in Table 4.2 and 4.3 respectively.

After FEC coding, one encoded symbol (together with RTP/UDP/IP headers) is placed in an IP packet and is subject to a uniformly distributed error pattern with loss rates of 5, 10, 15 and 20%. Header compression is assumed, and thus a 4 byte header is considered. The base data rate is set to 1000 kbps, and the successively higher rates are obtained by adding roughly 10% additional symbols, up to a rate 1.5 times higher than the base rate.

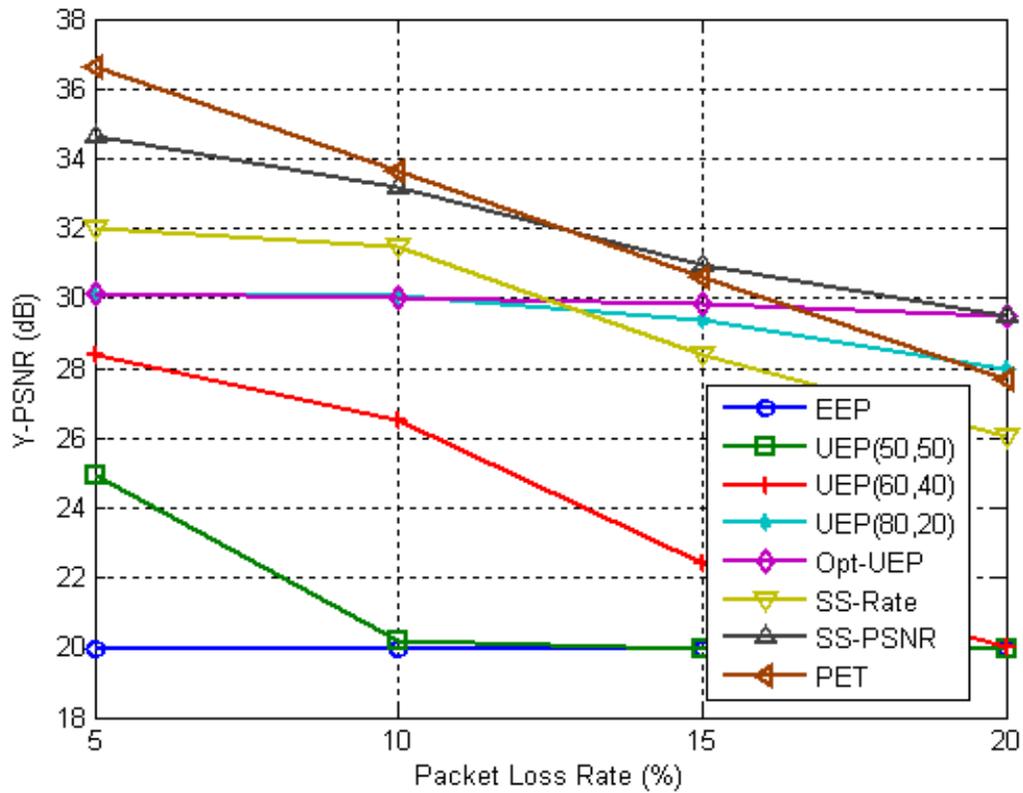


Figure 4.2: PSNR vs. PLR at overall data rate of 1.1 Mbps for Paris sequence.

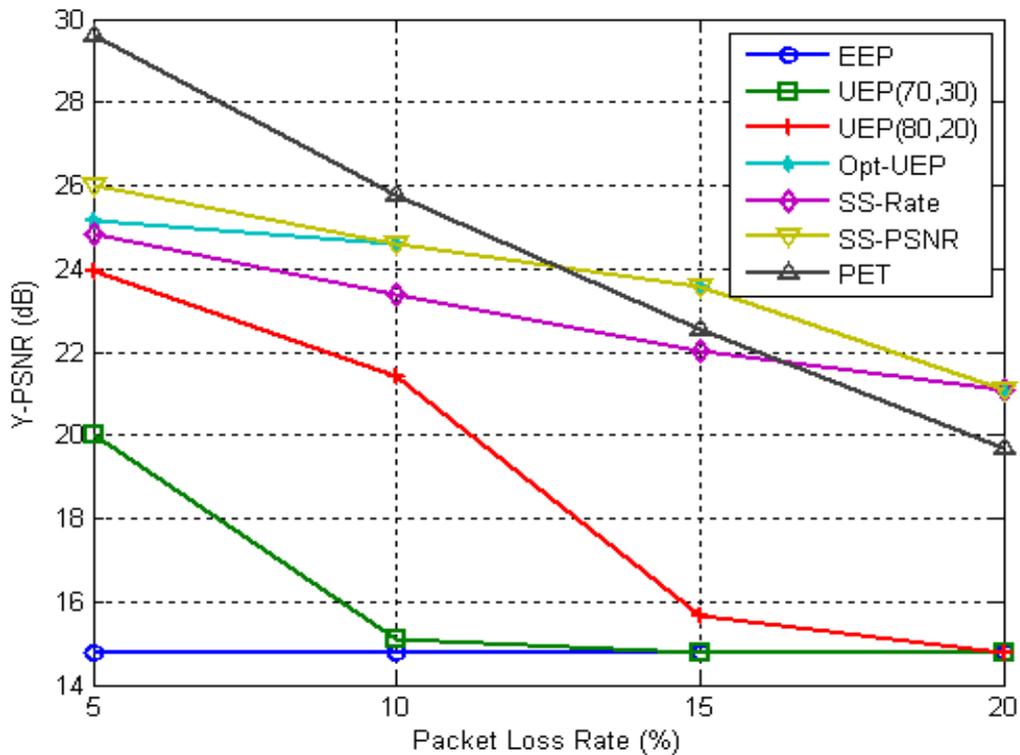


Figure 4.3: PSNR vs. PLR at overall data rate of 1.1 Mbps for Football sequence.

The simulations are performed using one slice per frame, and a frame rate of 25 frames per second (fps). The selected schemes are simulated with 100 runs for each GOP. In cases where the entire GOP is lost, the PSNR is obtained using the last frame of the previously decoded GOP to replace all frames of the lost GOP.

The results for the PET, EEP, and UEP schemes are provided and compared to the results obtained with two optimized RASSA schemes: SS-PSNR and SS-Rate scheme.

The results with frame-by-frame average PSNR performance of the six selected configurations at 1.1 Mbps with 10% PLR are shown in Figure 4.2 and Figure 4.3, for the "Paris" and "Football" sequence, respectively. "Opt-UEP" denotes the scheme that is optimized for each packet loss rate. As can be seen from Figure 4.3, the performance of the EEP scheme is the worst. The performance of the UEP schemes gets better with an increase in the protection of HPL. UEP(60, 40) performs worse as compared to UEP(80, 20) because the protection gets divided over both segments and none is protected enough. SS-PSNR performs the best of all the schemes except for PET scheme which has better result at low loss rates. The PET scheme has a better performance being based on RS codes which require less redundancy as compared to LT codes. However, beyond PLR of 12% SS-PSNR is better.

The performance of all the schemes for "Football" sequence in Figure 4.3 is similar to that of "Paris" sequence in Figure 4.2. However, an overall downward trend for all schemes is seen which is due to relatively low PSNR values of "Football" sequence. The performance of PET and SS-PSNR is also similar for the "Football" sequence. For the "Football" sequence performance of the optimized UEP scheme is very close to that of the SS-PSNR.

The results showing PSNR performance of the six selected configurations at 10% packet loss rates for different data rates are shown in Figs. 4.4 and 4.5, for the "Paris" and "Football" sequences, respectively. The performance of the EEP scheme gets progressively better at higher data rates. SS-PSNR and SS-Rate provide reliable and consistent performance at all the data rates. UEP(80, 20) is limited to 30 dB in Figure 4.4 even at higher rates because the DP C is not getting enough protection. Interestingly, at the highest rate the EEP scheme is better than the optimized UEP scheme, due to the absence of the performance penalty introduced by DP. PET is marginally better than SS-PSNR at low data rates.

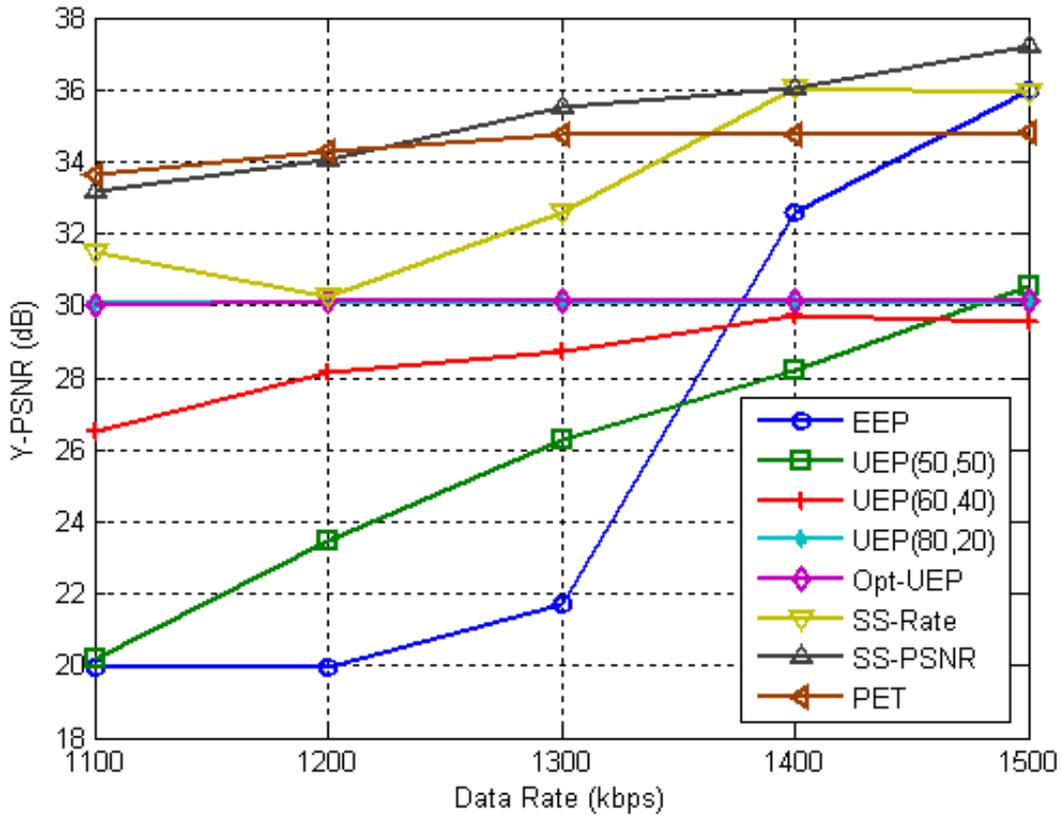


Figure 4.4: PSNR vs. data rate at packet loss rate of 10% for the “Paris” sequence.

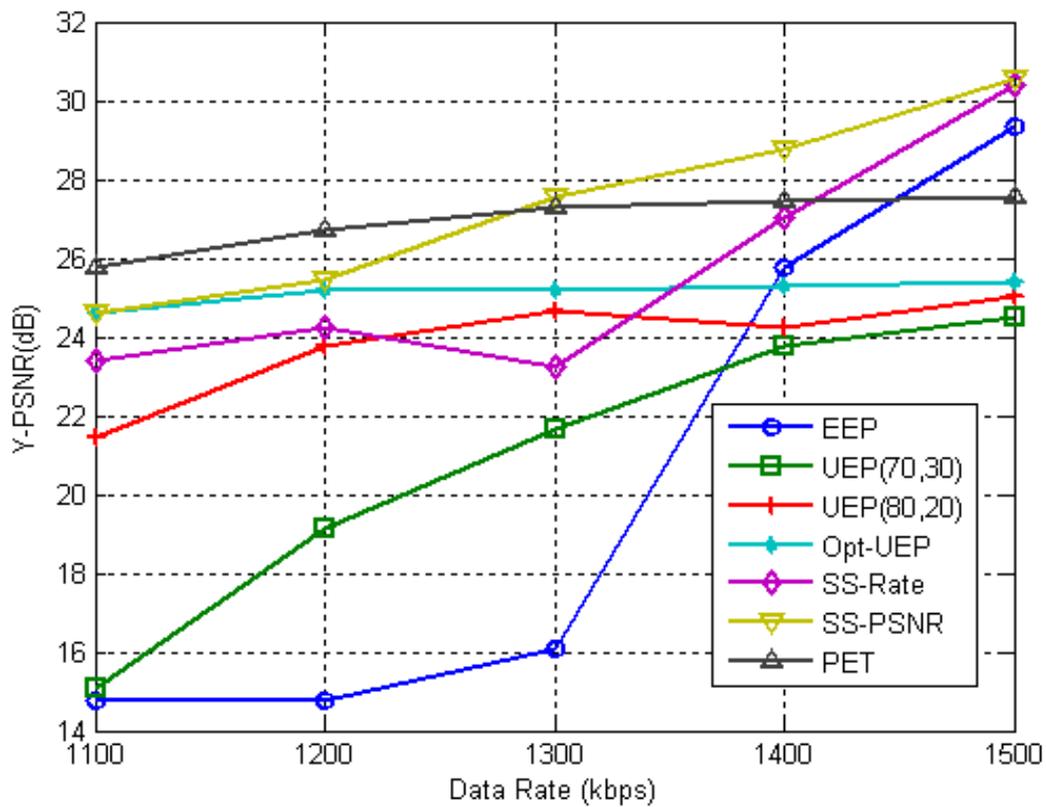


Figure 4.5: PSNR vs. data rate at packet loss rate of 10% for the “Football” sequence.



(a) Original Football Sequence.
PSNR 32.62 dB.



(b) Last 4 DP C dropped.
PSNR 31.36 dB.



(c) Last 8 DP C dropped.
PSNR 29.60 dB.



(d) HPC with 15 DP C dropped.
PSNR 25.39 dB.

Figure 4.6: Football sequence with some DPs dropped.

Frame 15 of Football sequence has been extracted for different configurations and displayed in Figure 4.6. It can be seen that the dropping of DP C as suggested in RASSA scheme does not bring much quality degradation.

Table 4.4: Priority classes and LT packetization for the first GOP of “Paris” sequence.

Class	Size(bytes)	Cum. PSNR	LT Pkts
HPC	34779	24.3	497
LPC	45536	39.08	651
Total	80315	39.08	1148

4.4.2 Sliced AVC transmission

In this section the simulation results with the slicing feature are presented. For simplicity, the case of $L = 2$ layers is considered: HPL that contains more important slices and LPL that contains less important slices [13]. The same video parameters are used as in the previous subsection. The sizes, number of packets and resulting PSNR values for the “Paris” video sequence are shown in Table 4.4.

The results are shown in Figures 4.6 and 4.7 and confirm the analysis carried out with the DP schemes. SS-PSNR is the best scheme overall. The UEP schemes, except UEP(45,55) in Figure 4.7 are around 24 dB as they suffer from an over protection of HPL. This is because the HPL size is only about 43% of the GOP size. This highlights the significance of considering the HPL size while designing UEP schemes. The EEP scheme becomes better than the UEP schemes at high data rates.

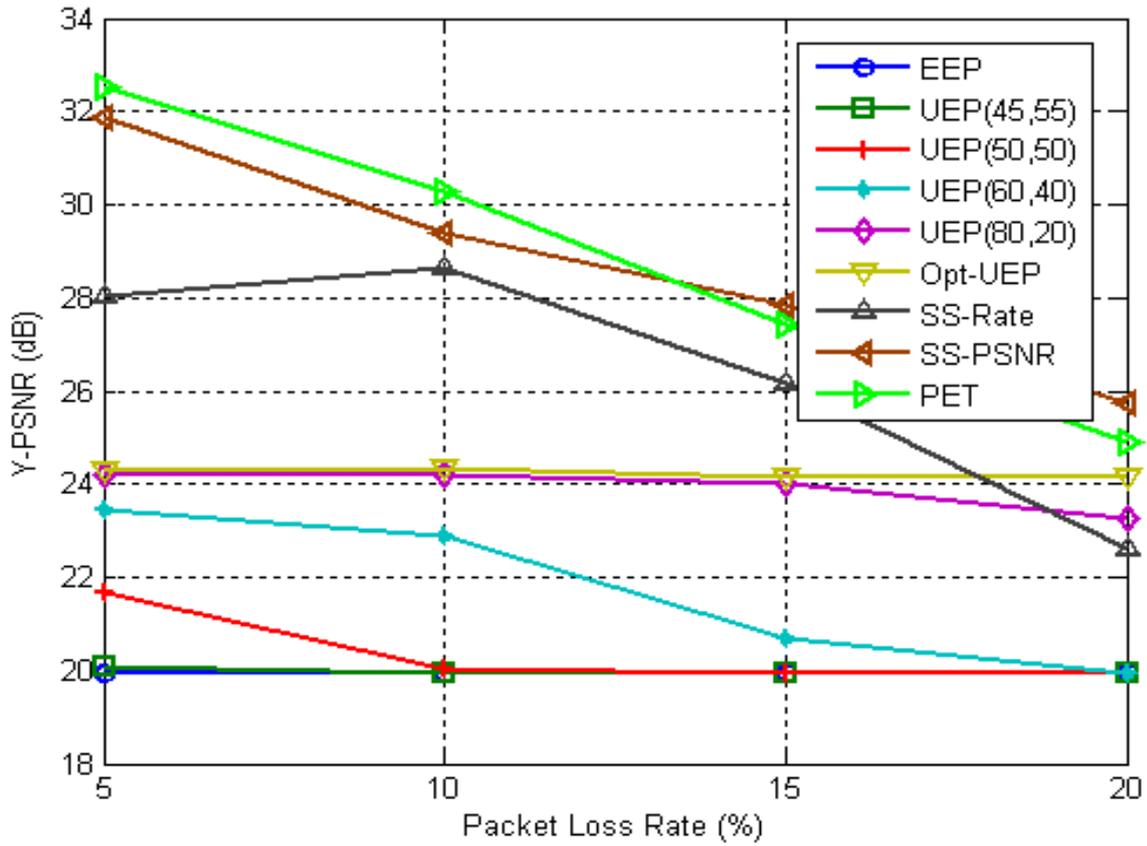


Figure 4.7: PSNR vs. PLR at overall data rate of 1.1 Mbps for "Paris" sequence.

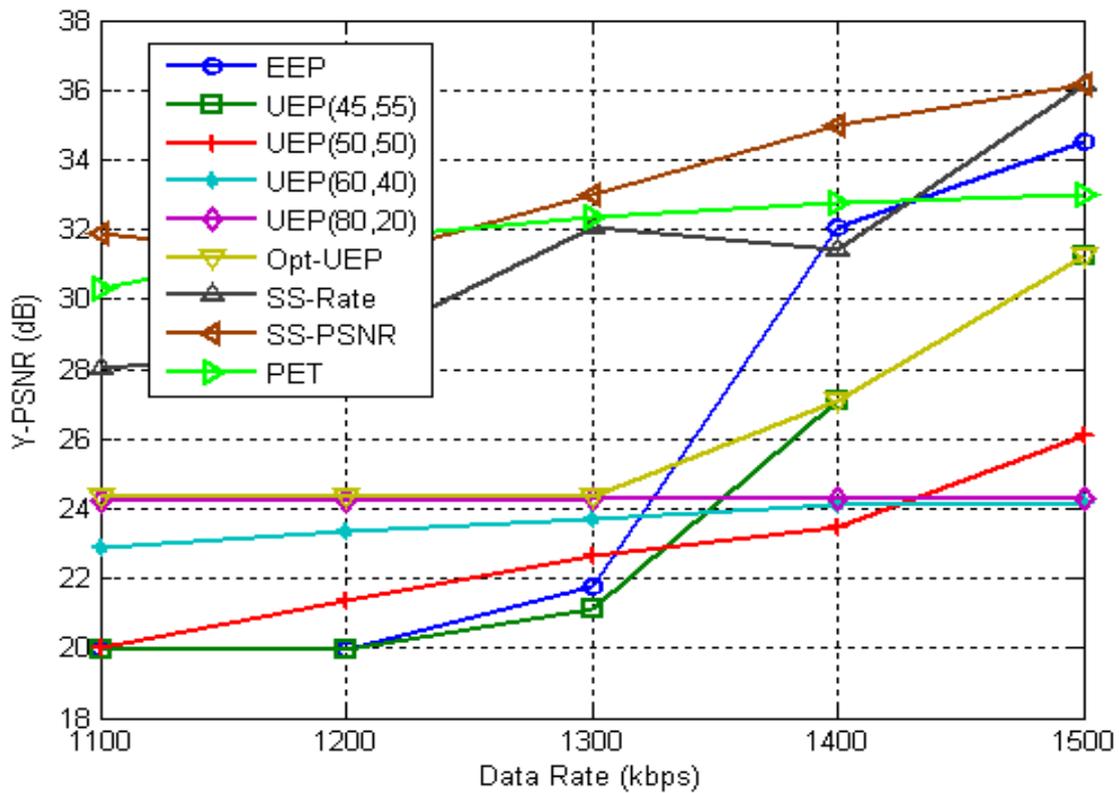


Figure 4.8: PSNR vs. data rate at packet loss rate of 10% for the "Paris" sequence.

4.5 Summary

Although both DP and slicing have been demonstrated to enable efficient layered video data transmission, the results with DP are seen to be better. The sizes and number of DP generated are as determined by the encoder subject to certain parameters as whether MB intra update (MBIU) or Random intra updates have been used. The prioritization of data into various partitions is thus optimum and can easily be used to create different rate points. Slicing, on the other hand, is more flexible as it allows for a finer layered structure. Moreover, in contrast to DP, slicing is available in the Baseline AVC profile. However, simulation results show small advantage of the DP-based scheme compared to the slicing-based, especially at high packet loss rates.

The performance of different coding schemes with segmented H.264/AVC video data has been analyzed. The segmented data can be selected to suit the available data rate and channel conditions. The UEP schemes provide better performance over EEP at some rates only. The RASSA scheme can be used to match the available transmission video data to the instantaneous channel conditions. It combines the best of both the EEP and UEP schemes to provide better and reliable video quality even in the worst channel conditions. The passing of the video-table to the decoder is a low-cost solution to an “all or nothing” decoding. Note that it is assumed that the video is pre-encoded, and thus the best way to match the source rate with the channel rate is to selectively drop some of the DPs, which is done in RASSA. Indeed, the results presented here, show that the pure UEP with fixed source rate suffers huge performance loss compared to the scheme that adjusts the source rate. The main advantage of the proposed scheme is a very simple adaptation of the source rate via DP AVC coding. Note that RASSA can be combined with UEP to better match source and channel characteristics. However, that would require multiple channel codes, increased complexity, UEP optimization algorithms, and reduction of the channel code length used could worsen channel codes' correction capabilities. This will be part of future work by incorporating expanding window codes [21].

The combined use of FEC and adaptively dropping some DP to maximize PSNR is thus shown as a practical method to ensure reliable delivery of multimedia data over wireless channels.

The DP and slicing features as described here are combined in Chapter 5 to split each of the slices into partitions (with DP) and hence design a novel MDC scheme.

Chapter 5

Video Delivery over Wireless Relay Networks

5.1 Introduction

In cooperative communications, different nodes in a network collaborate with each other to transmit information from a source node to destination node over multiple parallel paths [76]. This can not only result in reduction in power consumption by the low power mobile devices, but at the same time, it also increases the reliability.

The cooperative communications can also increase the capacity of wireless networks. The intermediate nodes in a transmission path could act as relays of the transmitted data from the source increasing the coverage and throughput of the network. These cooperative relay networks enable a direct, single-hop and multi-hop communication to take place simultaneously from a source to the destination. RLC are ideally suited for such multiple path transmissions as the aggregated received symbols over multiple paths helps decoding at the destination. In this study EW RLC is proposed for transmitting unequal error protected layered video data over a cooperative relay network. The results can be used for visual sensor and other similar networks.

For seamless and uninterrupted real-time multimedia transmission to heterogeneous mobile clients, it is necessary to exploit different diversity techniques to rectify severe noise and fading effects of wireless channels. Multipath diversity together with MDC [69],[70] is an efficient technique to combat noise, fading and shadowing in mobile channels.

Emerging wireless communications systems, such as LTE-A Release 10 [71], [72] and mobile WiMax, offer the possibility of increasing system capacity and extending coverage via relaying and cooperation [26]. In LTE-A, for example, the relays receive and retransmit the signals between the base station and mobiles, resulting in effective increase in the throughput and extending the coverage of cellular networks [73]. In order to autonomously and dynamically adapt to network bandwidth, delay, and channel noise fluctuations, a promising solution is to combine multipath relaying with MDC [69], [74]. MDC generates multiple descriptions of the same source data that can be independently decoded and is very

well suited to emerging relay-assisted networks, where each description can be routed via a different relay exploiting network diversity. However, since mobile channels are very prone to packet losses, FEC is still necessary to ensure that enough descriptions are correctly received.

For relay based configurations the paths have been assumed to be symmetric with reference to the delay encountered by packets traversing each path. Thus, it is assumed that packets transported over multiple paths which converge at a destination, each of them arrive so as to meet the delay deadline, unless lost based on the channel model used.

To provide adaptive resilience to packet losses, which are frequent in mobile channels, AL-FEC has become very popular. Indeed, AL-FEC via Raptor codes [9] has become standard for DVB-H and MBMS services [35]. Another class of erasure codes recently developed is RLC [13], [36], [20]. Like Raptor codes, RLC are rateless, capacity approaching, and of low encoding complexity. However, rooted in the network coding principles (see [13] and references therein), RLC enable simple and efficient relaying and network cooperation.

A novel and simple MDC scheme has been designed. In the proposed scheme independent descriptions are created with least possible duplication using slicing and DP [2] and used in conjunction with EW-RLC, originally developed in [22], enabling flexible rate adaptation and competitive performance [55]. Although slicing and data partitioned video data are used to apply EW-RLC, the proposed scheme is general enough to be applied to other layering MDC schemes e.g., as in [77], redundant slices are used for creating multiple descriptions.

A system is proposed for real-time video streaming to mobile users, via LTE-A and similar relaying wireless systems, using H.264 AVC-based layered MDC [2] together with RLC. The source data is encoded in multiple descriptions, which are sent to a receiver over non-cooperative relays. The receiver can reconstruct the encoded data from any subset of the descriptions received. To maximize benefits of relaying and MDC, two algorithms are proposed for optimal and sub-optimal, but fast, relay selection and optimal resource allocation that minimizes reconstructed source distortion.

Some work related to resource selection in relay- assisted communication is covered next. In [82], an amplify-and-forward wireless relay system is considered, and novel power allocation strategies are proposed, based on geometric programming, that optimize the maximum transmit power and the network throughput. In [83] a dynamic resource allocation of transmission powers and sub-channels under power constraints based on instantaneous channel states that maximize the instantaneous total transmission rate is investigated and a greedy approach proposed. In [84], a globally optimal Pareto-based solution and sequential optimization algorithm using channel state information are proposed for the OFDMA-based half-duplex single-relay channel. In contrast to the above papers which provide “lower layers” (physical, link, and/or network layer) parameters selection, this work focuses on the application-aware resource allocation, where upper-layer parameters (such as AL FEC) are considered as well, and the reconstructed source fidelity is used as performance measure.

Some earlier related work covers the path selection and resource allocation for multipath video. In [85], optimal and polynomial-time suboptimal algorithms for selecting the network path and scheduling packets of encoded scalable video are proposed that minimize video distortion and time delay at the receiver. In [86], an MDC relaying is considered, and an adaptive compression scheme is proposed at the source and relay based on the received feedback. In [78] a joint source-channel scheme is proposed that exploits information about packet losses to adaptively select reference frames in MDC. In [79] FEC and routing are jointly optimized to maximize reconstructed quality. The main difference between this study and the above contributions is that the proposed resource allocation strategy considers both lower and upper layers parameters and proposes a dynamic programming and a suboptimal greedy solution. In contrast to [78], [79], video encoding is not altered, making it compatible with the standard, but only the coding rate is adapted, and relay selection is done based on the channel characteristics

Targeting applications such as [80], LTE-A network is modelled using a Finite-state Markov chain (FSMC) [81] path loss channel model using LTE-A Release 10 parameters suggested in [75].

The rest of the chapter is structured as follows. Multi hop cooperative relay networks comparing results for AF and DF schemes are described for a generic system in Section 5.2. The design of an adaptive layered MDC coding scheme for wireless video is described and

the scheme is simulated for a technology independent multi-path model in Section 5.3. The next three sections specifically address relay based communication as envisaged in LTE-A. The proposed system for relay-assisted rateless layered MDC is described in Section 5.4. The resource allocation optimization is presented in Section 5.5. The simulation setup and results are provided in Section 5.6. Finally, Section 5.7 summarizes the chapter.

5.2 Multi hop Cooperative Relay Networks

The focus of this section is to compare the performance of layered video data over multi-hop relay networks using the EW approach [21] with RLC for video streaming. The layered video data is obtained with DP feature of H.264/AVC and quality scalability with H.264/SVC. In the simulations the video data is partitioned into two windows/layers, based on its importance for video reconstruction and each layer is assigned different selection probability to afford to it a selected degree of protection.

The simulation results show that EW RLC can be used to effectively protect the prioritized video data for reliable video transmission over the emerging relay networks. The results can easily be extended to more than one relay node. Although schemes exist which allow for passing back of an acknowledgement [87] of video decoding at the relay or the destination, this analysis is restricted to the case without provision of a feedback. Furthermore, for video broadcast/multicast applications, it can be advantageous to relate the degree of protection to various DP segments depending on the varying channel conditions, i.e., the available data rate and the packet loss rate.

In this section Amplify-and-forward (AF) [88] and Decode-and-forward (DF) schemes are analyzed for streaming of layered video with EW-RLC over relay networks. An adaptive scheme is proposed to vary the degree of protection to the most important layer to improve the overall quality of received video.

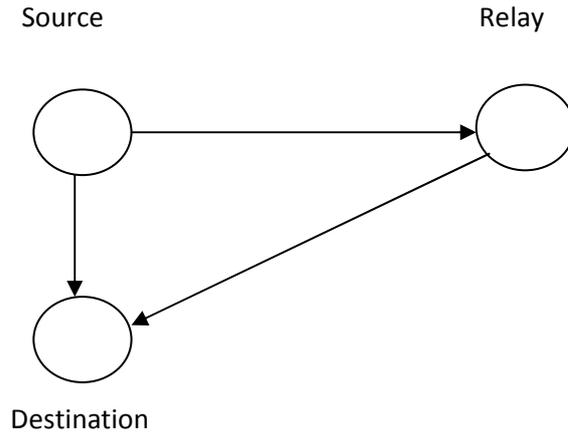


Figure 5.1: 3-relay node model.

5.2.1 System Description

The nodes in multi-hop relay networks communicate over unreliable wireless channels and have limited transmission power. The analysis is performed using 3-node relay model which comprises of a source (S), relay (R), and a destination (D) node as shown in Figure 5.1. The channel from source-to-relay (SR) and relay-to-destination (RD) is assumed to be better as compared to the direct channel between source and destination (SD). The transmissions on channel SD is received at relay R. If the channel SD is good then there may be no requirement for relay collaboration. However, if the direct channel SD is bad then the relay can enter the collaboration state, where it tries to aid the decoding operation at destination node. This could also result in other advantages such as power consumption, if the feedback channel from destination to relay is available. It is assumed that the relay is half-duplex that is; it cannot transmit and receive at the same time. However, the destination can receive transmissions from S and R simultaneously.

The relay collaboration could be as an AF scheme [88]. The relay acts as a repeater, simply re-transmitting each of the received packets towards destination node. Note that the same code is being transmitted on the channel SD; hence if both packets are received by the destination node, it does not bring any decoding advantage. At the destination the decoding is attempted at the end of transmission.

For the other relay collaboration scheme, that is DF, the relay keeps accumulating the packets till such time that the received data is decodable. As soon as the decoding at relay node succeeds, it re-encodes the video data and starts transmitting to the destination. In this case, if there is an early decoding at the relay, it can be quite advantageous especially once the channel conditions on direct link from source to destination are bad.

5.2.2 Measurement Setup

The video sequence Paris and Football have been used for all the simulations. The choice is dictated by Paris sequence having a high temporal complexity and the Football sequence having a high spatial complexity.

Both of the sequences used are in CIF format at 25 frames per second (fps). The GOP size is 16 frames with an IPPP... structure. To make it possible to provide UEP, two layers of video data are created which could then be protected by providing unequal protection by using EW-RLC.

The number of total video packets for different video configurations is kept to be about same so as to arrive at correct analysis of the proposed schemes.

The video data is packetized into packets of size 1024 bytes. We assume a packet header of 6 bytes only with header compression. Each simulation is repeated 1000 times and the results are then averaged.

The simulations are performed using the EW-RLC to compare the performance of AF and DF for transmission over the 3-node relay model. The Finite-state Markov chain (FSMC) [81] path loss channel model is used to model the network.

The equal error protection (EEP) scheme, with no priority for HPL is compared against UEP schemes. For the UEP schemes, the probability of selection (PS) for HPL is varied as 0.50, 0.60, 0.80 and 1.0 to create four UEP schemes. With a PS of 0.60, the symbols are generated from HPL with a 60% probability, hence, affording better protection to HPL. It is also to be noted that the size of HPL also governs selection of an appropriate PS. For UEP, PS should be higher as compared to the size (percentage) of HPL in comparison to LPL.

The packet loss rate (PLR) for both the channels SR and RD is kept constant at 0.025. Thus, for our simulations we assume better channel conditions for the relay. The PLR for the direct channel SD is varied as 0.05, 0.10, 0.20, 0.30, 0.40, and 0.50.

If the entire GOP is lost, we resort to a simple error concealment method based on using the last frame of the previously decoded GOP to replace all frames of the lost GOP.

Table 5.1: Relative Partition Sizes – Paris Sequence

Partition	Size (bytes)	Size (%)	Cumulative PSNR
IDR	22,281	27.52	-
DP-A	12,838	15.86	-
DP-B	97	0.12	30.32
DP-C	45,732	56.50	39.16
Total	80,948	100	39.16

Table 5.2: Relative Partition Sizes – Football Sequence

Partition	Size (bytes)	Size (%)	Cumulative PSNR
IDR	23,374	28.92	-
DP-A	22,823	28.24	-
DP-B	2893	3.58	25.39
DP-C	31,731	39.26	32.62
Total	80,821	100	32.62

Table 5.3: Packetization of Layers.

Sequence	Number of Packets		
	HPL	LPL	Total
Paris	35	45	80
Football	49	31	80

5.2.3 Video Transmission with H.264/AVC

5.2.3.1 System Configuration

DP feature is used to create encoded data with partitions. These partitions are then used to create two layers of video data which is then protected by providing unequal protection by using EW-RLC.

The breakdown of video data into constituent partitions for the Paris sequence and Football sequence are shown in Table 5.1 and 5.2 respectively. After the encoding process, the IDR frame together with DP A and B are placed in High priority layer (HPL), and the remaining portion, that is, DP C is placed in Low priority layer (LPL).

The packetization details for each of the two sequences are as given in Table 5.3. It can be seen from the table that the number of encoded packets for HPL of Football sequence are substantially higher relative to Paris sequence. This implies that same UEP strategy will not work equally well for both of these sequences.

5.2.3.2 Simulation Results

The results for the Paris sequence are shown in Figure 5.2. The scheme AF50 means an AF scheme with PS as 0.5. AF-EEP scheme applies the protection over the whole of the video data treating all the data to be at the same priority. The results show that the performance of EEP scheme is much worse than the UEP schemes. DF-EEP is worst of all as at the relay the decoding does not result, and hence it is almost a case of no relay collaboration.

The two interesting cases are for AF100 and DF100 which overlap exactly but provide a consistent PSNR contribution of 30.32 dB irrespective of the PLR. With an over-protection for HPL the performance is limited to PSNR achievable with receiving HPL (see Table 5.1) alone.

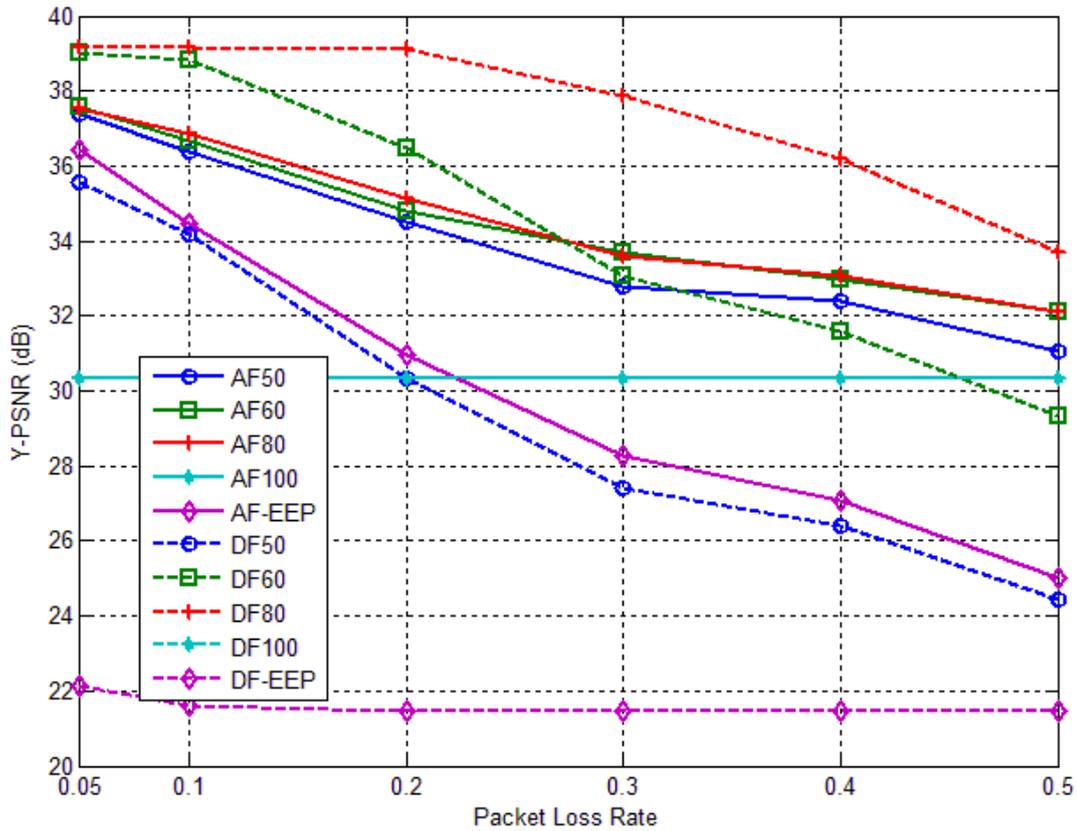


Figure 5.2: PSNR vs. PLR for Paris sequence (AVC).

The performance of DF becomes better with an increase in PS of HPL, such that DF80 provides significantly better results. This is attributable to the fact that with more protection to HPL, the probability of the HPL being decoded at the relay node increases. This results in the early start of the relay collaboration phase which aids the video decoding process at the destination.

The results for the Football sequence are shown in Figure 5.3. The performance of DF schemes is generally worse at low PS. The reason is that for Football sequence, the size of HPL is significantly large as compared to the LPL; hence the successful decoding of HPL comes very late for the DF scheme. At high PLR, with no relay participation, therefore there is a sharp decrease in the performance especially for the DF schemes with low PS. The AF100 and DF100 manage to maintain performance across the full range of PLR.

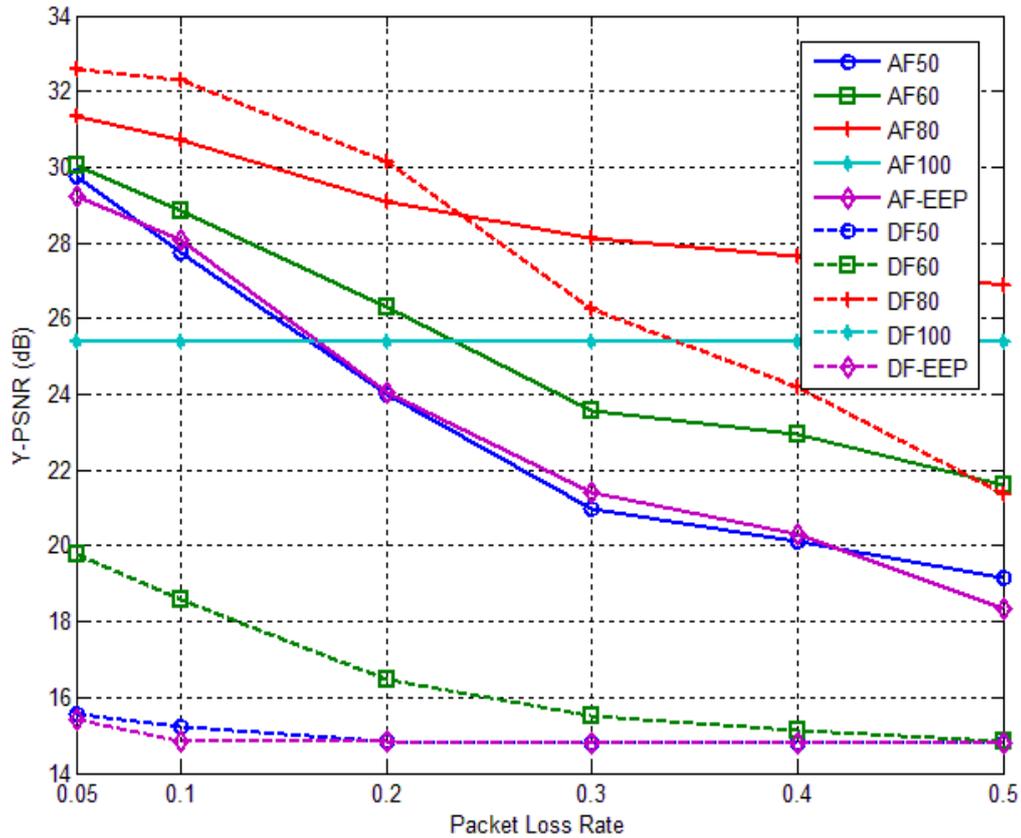


Figure 5.3: PSNR vs. PLR for Football sequence (AVC).

5.2.4 Video Transmission with H.264/SVC

5.2.4.1 System Configuration

Coarse Grain Scalability (CGS) has been used to create scalable video. Quantization parameter (QP) can be used to tailor the size of the base layer and enhancement layers. This also affects the resulting achieved PSNR for each layer.

Two layers of video data are created which could then be protected by providing unequal protection by using EW-RLC.

The breakdown of video data into two layers for the Paris sequence and Football sequence are shown in Table 5.4 and 5.5 respectively. The packetization details are shown in Table 5.6.

Table 5.4 : Relative Layer Sizes - Paris sequence.

Layer	Size (bytes)	Size (%)	Cumulative PSNR
HPL	29,738	36.64	32.34
LPL	51,441	63.36	32.34
Total	81,179	100	37.14

Table 5.5 : Relative Partition Sizes - Football sequence.

Layer	Size (bytes)	Size (%)	Cumulative PSNR
HPL	34,814	43.44	27.91
LPL	45,337	56.56	30.32
Total	80,151	100	30.32

Table 5.6 : Packetization of Layers.

Sequence	Number of Packets		
	HPL	LPL	Total
Paris	30	51	81
Football	34	45	79

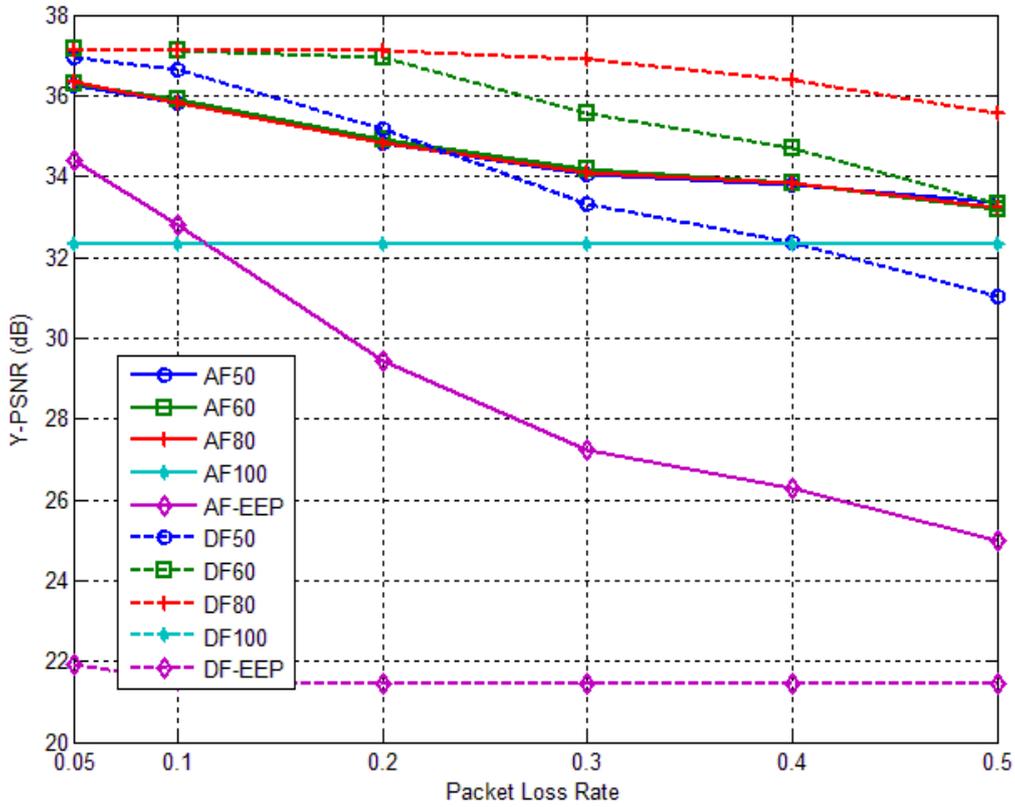


Figure 5.4: PSNR vs. PLR for Paris sequence (SVC).

5.2.4.2 Simulation Results

The results for the Paris sequence are shown in Figure 5.4. The results show that the performance of EEP scheme is much worse than the UEP schemes. The schemes AF100 and DF100 maintain a consistent PSNR contribution of 32.34 dB irrespective of the PLR, which is the PSNR achievable with HPL alone.

The performance of DF schemes is generally better than AF. However, as noted previously the performance of DF schemes gets better with higher PS of HPL. Again DF80 provides the best results overall.

The results for the Football sequence are shown in Figure 5.5. The results confirm the results obtained with Paris sequence. However, the performance of DF schemes is a bit worse than that of Paris sequence. This degradation is due to a larger HPL.

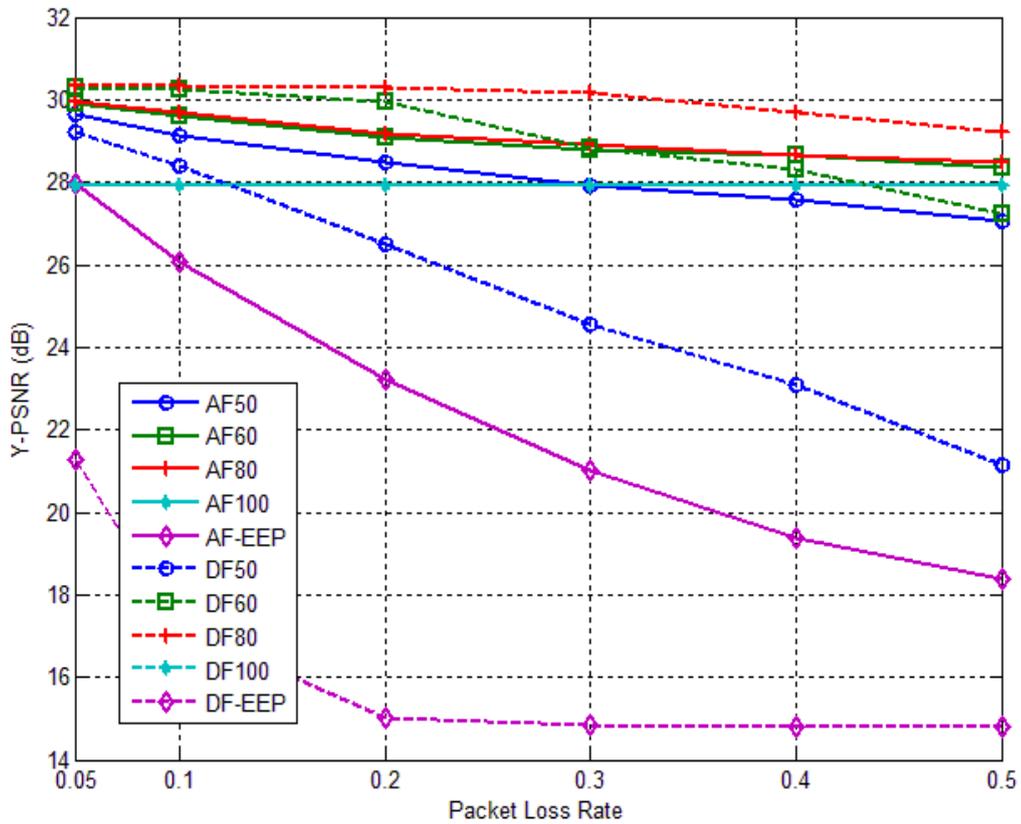


Figure 5.5: PSNR vs. PLR for Football sequence (SVC).

5.2.5 Comparison and Analysis

The performance of H264/AVC and H.264/SVC can be compared based on Figure 5.2 and 5.4. However, the difference in HPL sizes needs to be considered as well. Referring to Table 5.1 and 5.4, it can be seen that the size of HPL in case of H.264/AVC is larger with a lower PSNR (HPL) as compared to H.264/SVC. On the other hand, H.264/SVC imposes a loss in coding efficiency, so the total PSNR for H.264/AVC is slightly better than that of H.264/SVC.

An interesting effect is due to the increase in packet loss rate. With an increase in PLR, more and more PSNR contribution comes from HPL alone relative to combined decoding of

both HPL and LPL. This is seen in AF60 and DF60 of Figure 5.1, where DF60 drops below AF60 at higher PLR.

The effect is also seen by comparing DF80 results as shown in Figure 5.2 and 5.4, i.e., the results for DF80 for H.264/AVC are better at low PLR but for higher PLR the results of H.264/SVC become better. Similar analysis holds for AF schemes. The analysis is also substantiated with the results for Football sequence.

Next, an investigation into the effect of HPL size on the performance of AF and DF schemes is provided. The HPL size is chosen to be 30 and 70 (out of total 100) video packets for the two considered configurations. In order to quantify the effect, PE is kept as 0.3; PS is increased in steps and the resulting performance of AF and DF schemes is measured. The results are shown in Table 5.7. The decoding (%) implies a count of case when for instance only HPL gets decoded, or GE fails to recover anything (the column titled failure). As seen in Figure 5.5, a larger HPL size has an adverse effect on the performance of DF scheme.

Table 5.7: Effect of PS on Layer Decoding.

HPL Size (%)	PS (%)	Decoding (%), (PE = 0.3)					
		HPL		HPL+LPL		Failure	
		AF	DF	AF	DF	AF	DF
30	30	43	38	21	4	36	58
	40	78	60	21	37	1	3
	50	81	17	19	83	0	0
70	70	37	2	22	0	41	98
	80	76	47	23	1	1	52
	90	77	91	23	7	0	2

As can be seen from Table 5.7, the performance of DF schemes gets better with increase in PS. Hence, the PS of HPL must relate to the size of HPL in the GOP data being transmitted. The effect can similarly be seen for AF scheme. The change in PS can be easily accomplished by having an adaptive scheme which varies PS in relation to size of HPL (on GOP basis) ensuring a better video quality throughout. In a scenario where no feedback mechanism is available, DF80 would be the best choice over the entire range of PSNR.

The advantage of H.264/AVC is that once the DP have been created, it is possible to selectively drop selected partitions to yield a base layer of required size as described in Chapter 4 for RASSA.

For further study of relay based communication H.264/AVC standard is used. The DP feature as used in this section is used in combination with slicing to yield multiple descriptions. The design of the layered MDC scheme and simulations over a generic two-path model are provided in the next section.

5.3 Adaptive Layered MDC Coding for Wireless Video

In this section an MDC scheme is proposed based on DP and slicing feature of H.264/AVC. In the proposed scheme independent descriptions are created with least duplication and additionally EW-RLC [52] is proposed for FEC. The advantage being that the degree of protection can be adapted to suit a particular wireless channel.

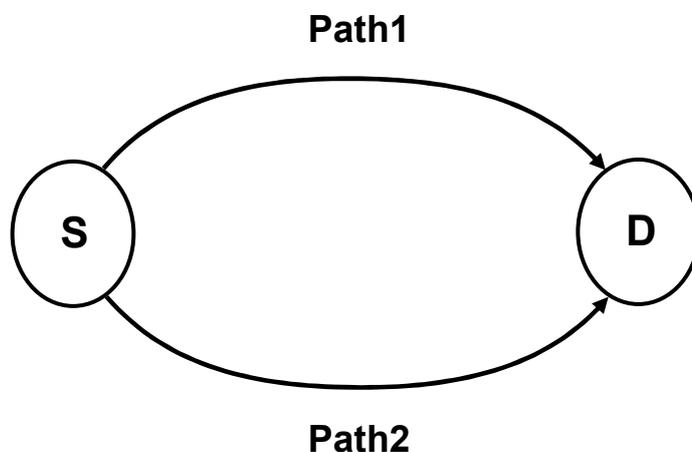


Figure 5.6: The system model.

The focus of this section is to simulate the transmission of layered descriptions over a two-path wireless network as shown in Figure 5.6 and find the best rate distribution for the descriptions that maximizes PSNR. In this section it is assumed that two independent wireless channels connect a source S to the destination D, which could exist in a multi-interface receiver in a multi-hop wireless network. The system parameters relating to the channels are the data rate and packet erasure rates. The possible applications are in multi-interface networks and VoD applications.

5.3.1 System Description

Two independent descriptions with equal size and importance can be created by dividing the slices of a frame into two independent sets. As shown in Figure 5.7, each frame has been divided into six slices, and each description comprises of three slices from each frame. The decoder is thus able to create the missing information in case of losing any description.

In order to make use of EW-RLC codes over each description, the DP could be used to partition each slice into two layers and these layers could then be provided with unequal protection.

The simulations have been performed using Foreman sequence in CIF format. The CIP flag has been set and MBIU are used to contain the error propagation. The GOP size is 8 and 6 slices per frame are used. The video data rate is 1000 kbps for 25 frames per sec. Each simulation is run 1000 times for the results to converge. Two descriptions using slice distribution are created for use in all the simulations. Each of the description contains 3 slices from each frame as shown in Figure 5.7. Note that each of the descriptions has the IDR frame to make possible the decoding of each description.

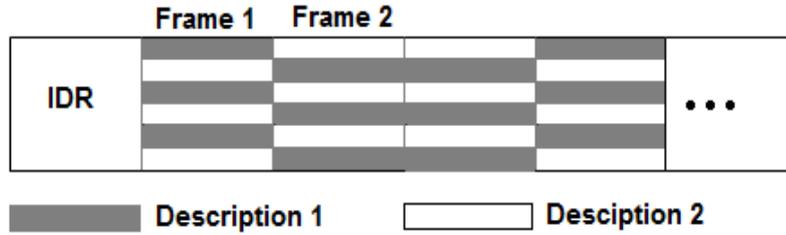


Figure 5.7: Slice layout for creating descriptions. Each rectangle represents a slice.

In order to enable unequal protection to each of the description's data, two importance layers are identified. The data within each of the slices is already partitioned into DP A, B and C. The most important data for video re-construction, consisting of IDR, DP A and DP B is termed as high-priority layer (HPL). The remaining data i.e. DP C constitutes low-priority layer (LPL).

Table 5.8: System configuration.

Cat.	GOP		MDC1		MDC2	
	Size (bytes)	Cum. PSNR	Size (bytes)	Cum. PSNR	Size (bytes)	Cum. PSNR
IDR	11515	-	11515	-	11515	-
DP A	9922	32.45	4892	28.37	5030	27.92
DP B	4352	33.91	1744	28.76	2608	28.49
DP C	14828	40.1	7538	29.55	7290	28.92
Total	40617	40.1	25689	29.55	26443	28.92

The sizes of the different partitions for the first GOP of the Foreman sequence together with the assignment of the partitions to MDC1 and MDC2 are shown in Table 5.8.

Table 5.9: PSNR Contribution.

Categories Received		PSNR (dB)
MDC1	MDC2	40.1
MDC1	HPL2	36.27
MDC1	Fail2	29.55
HPL1	Fail2	28.76
HPL1	MDC2	36.6
Fail1	MDC2	28.92
Fail1	HPL2	28.49
HPL1	HPL2	33.91
Fail1	Fail2	20.71

Table 5.10: Packet allocation to different layers.

Layer	MDC1			MDC2		
	Size (bytes)	PSNR	Pkts	Size (bytes)	PSNR	Pkts
HPL	18151	28.76	36	19153	28.49	38
LPL	7538	29.55	15	7290	28.92	15
Total	25689	29.55	51	26443	28.92	53

The PSNR contribution for each of the possible combinations has been calculated and is listed in Table 5.9. As can be seen from the table, the best video quality is achievable when both the descriptions are received correctly. However, the novel way that the descriptions have been designed, the video can be decoded with just the HPL of any one description only. This layering of video data within a description makes this scheme adaptive and practical for varying channel conditions. In case where both descriptions are completely lost and GOP is not decidable, then the decoder applies simple error concealment technique by replacing the lost GOP by the last frame of the previously successfully decoded GOP, which comes to 20.71 dB.

The video data belonging to each description is divided into packets of 512 bytes. The details of the packetization are as shown in Table 5.10. The simulations are carried out with both uniform loss and fading channel model to represent wireless channels.

For the video transmission, depending on the bandwidth and packet erasure rate, probability of selection, PS1 could be varied to maximize the PSNR. An interesting scenario is with PS1 as 100%, thus effectively sending HPL only.

In the absence of any errors the PSNR achieved with both the descriptions is the same as that of standard coding, considering slicing and data partitioning.

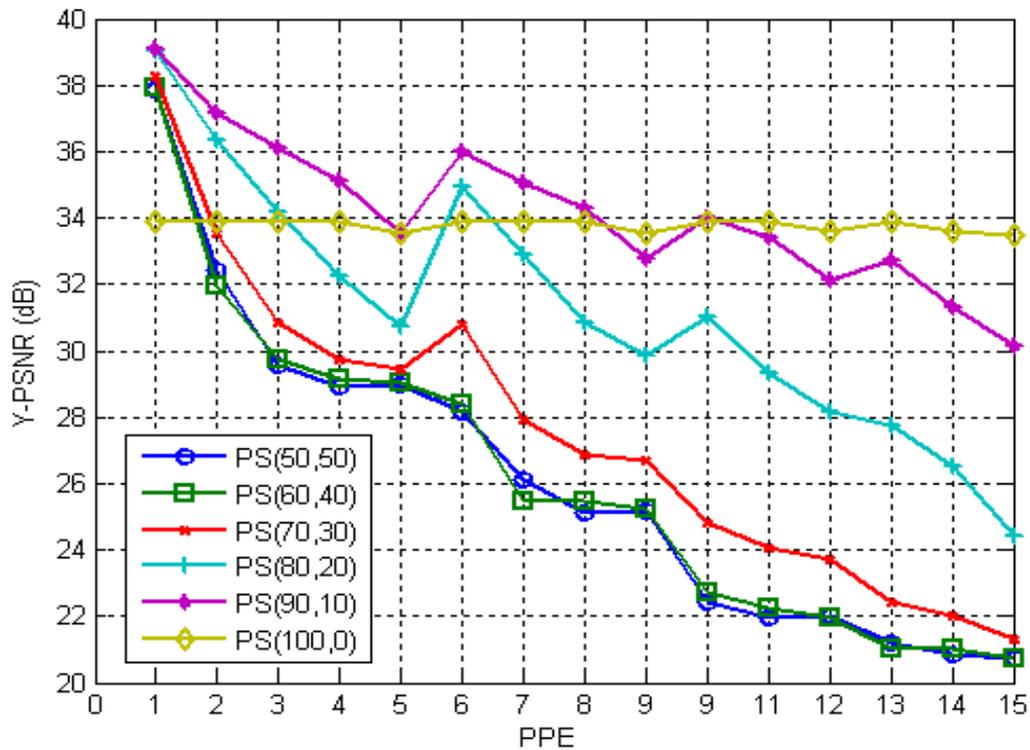


Figure 5.8: PSNR at different PPE values of Section 5.3.2.

5.3.2 Results for Uniform Loss Model

In this section the analysis is performed with uniform loss model. The probabilities PS1 and PS2 are kept as same for both paths. The data rate for each of the paths is kept as 10% over the source rate. The probability of error (PE) on paths are varied in increments of 0.05

from 0.05 to 0.25, to yield 15 combinations (0.05, 0.05), (0.05, 0.10), (0.05, 0.15), (0.05, 0.20), (0.05, 0.25), (0.10, 0.10), (0.10, 0.15), (0.10, 0.20), (0.10, 0.25), (0.15, 0.15), (0.15, 0.20), (0.15, 0.25), (0.20, 0.20), (0.20, 0.25), and (0.25, 0.25). This pair of PE is termed as PPE1 through PPE15.

The results for each set of PE with different PS1 are shown in Figure 5.8. Each curve (from left to right) depicts a PPE from (0.05, 0.05) to (0.25, 0.25) and shows a general downward trend subject to cumulative PE on both channels. The scheme with PS (100, 0) is seen to be virtually independent of the increasing packet erasures. The raised points on the plot in Figure 5.8, e.g., for PS (80, 20) at PPE6, occur because the PPE (pair of PE) at that point provides lower cumulative loss as compared to PPE the immediate left (PPE5) and right (PPE7) of that point (PPE6). For PPE 6, the pair of PPE is (0.10, 0.10), which is better overall than PPE5 as (0.05,0.25) and PPE7 as (0.10, 0.15). The other spikes in the plot can be explained similarly.

The PS(90,10) is providing the best result up to PPE8 because it is able to protect LPL also which contributes to PSNR as compared to PS(100,0) which is constrained to 33.91 dB. However, beyond PPE8, PS(90,10) cannot protect HPL with its 90% selection probability.

The result of PS(100,0) is otherwise impressive as over the entire range of PPE, it maintains a very high PSNR, i.e., 33.91 dB which is attributable to the way the MDC scheme has been designed.

Thus, by increasing the protection of HPL for a description it is possible to offset the effect of higher packet loss rates, and hence, a consistent reconstruction quality can be obtained.

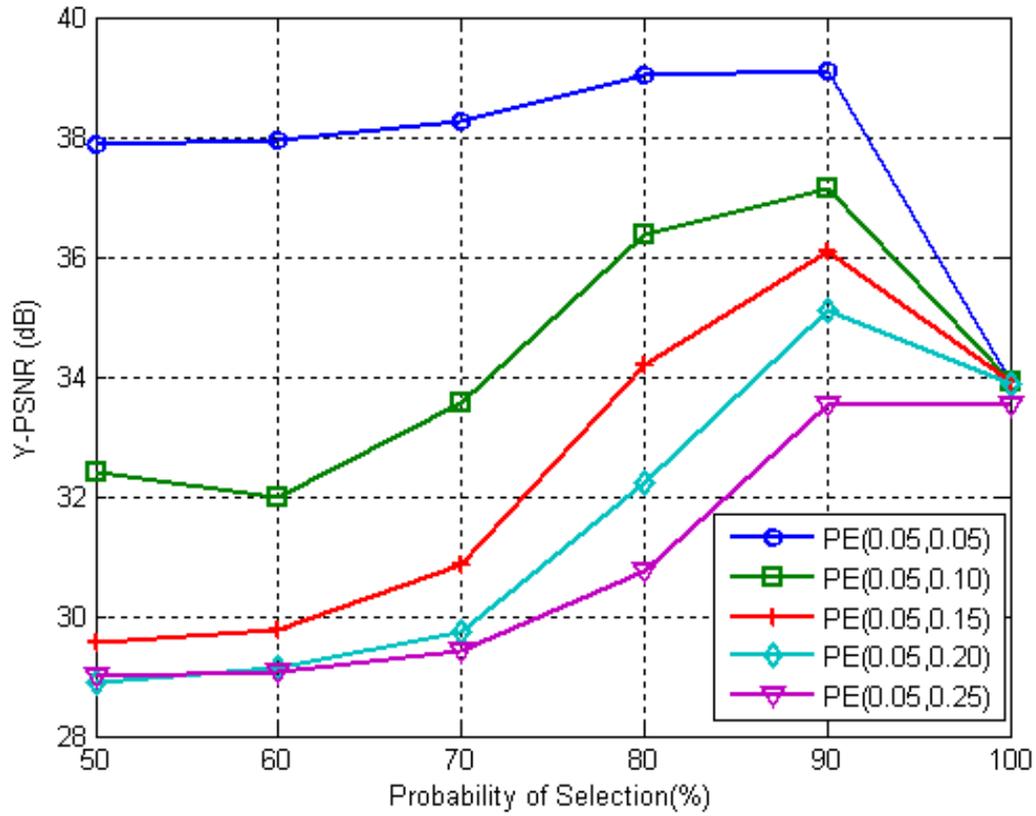


Figure 5.9: PSNR at different PE.

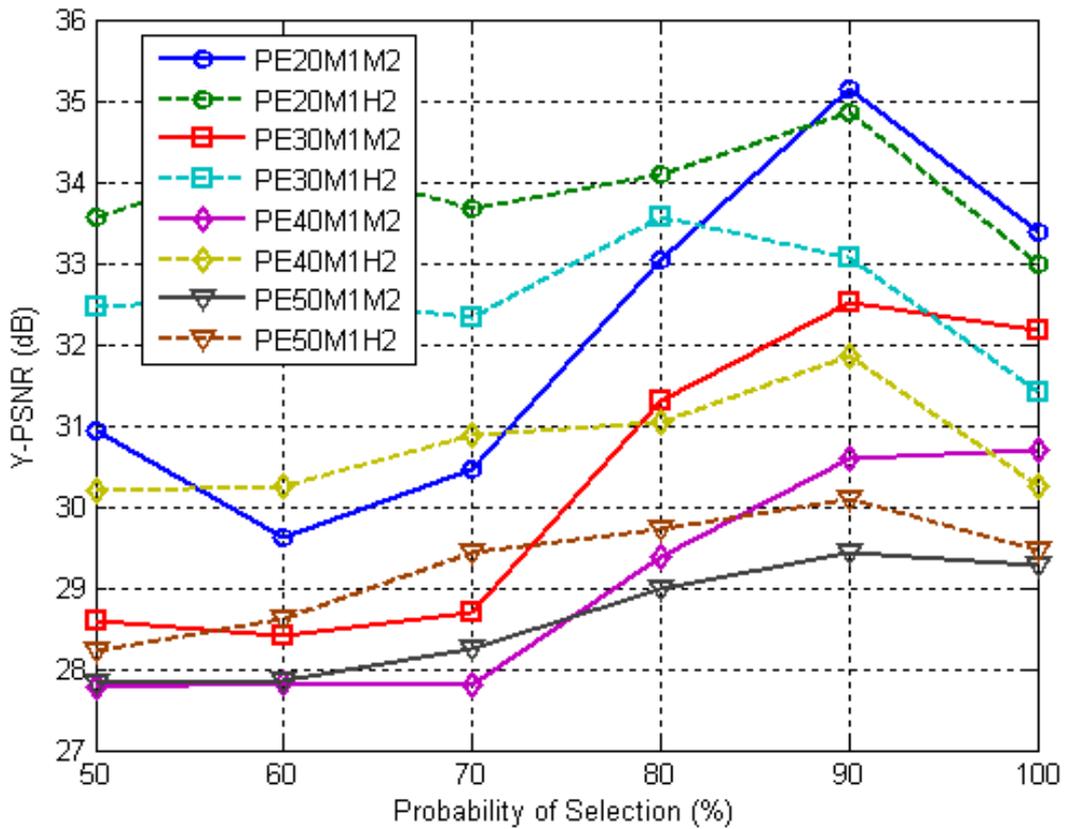


Figure 5.10: Comparison of sending (MDC1, MDC2) against (MDC1, HPL2).

The results for probability of selection vs. PSNR are shown in Figure 5.9. The scheme PE(0.05, 0.15) means PE= 0.05 on one path and PE=0.15 on the other. With an increase in PS from 50% to 90% there occurs a corresponding increase in PSNR. Within this range the PSNR contribution comes with successful decoding of both HPL and LPL. As the PS is increased to 100%, all the schemes converge at around 33.91 dB which is the PSNR achieved with receiving HPL only for both MDC1 and MDC2.

5.3.3 Results for Fading Error Model

In this section, one of the paths is assumed to be better than the other. The PE1 is kept as 0.05, whereas PE2 is varied as 0.2, 0.3, 0.4 and 0.5. Gilbert model is used to simulate the burst losses with an average burst length of 5 packets for PE1 and PE2. The PS is varied from 0.5 to 1.0 as in the case of uniform loss. The data rate for each of the paths is kept as 10% over the source rate.

Corresponding to each PE value for Path2, two configurations (based on descriptions) are transmitted. In one configuration, MDC1 is transmitted over path1 and MDC2 is transmitted over path2. Also, to establish the importance of adapting the description to the channel bandwidth and error characteristics, in the other configuration MDC1 is transmitted over path1 whereas only HPL2 is transmitted over path2. This is because for a low bandwidth path like path2 here, the protection of whole of the description MDC2 may not be possible. Results for the adaptive scheme are presented by varying the coding parameters, i.e., PS, and the information (complete description or HPL).

The results for transmission comparing the PSNR with burst loss, for transmitting each of the descriptions, MDC1 and MDC2 in full, with transmitting MDC1 with only HPL2 at different PS are shown in Figure 5.10. The configuration PE20M1H2 is a scheme with 20% burst errors wherein MDC1 and HPL2 are being transmitted. It can be seen that the schemes with only HPL2 being transmitted have better results for the entire range of PS. Thus, in any such scenario, where the channel is incapable of successfully communicating the whole of source data, it is advantageous to reduce the source data itself. This reduction brings a drop in

PSNR which is offset by gain in successful reception (from reduced information), hence providing an overall improved result.

The design of the two-path MDC scheme has been evaluated under various channel conditions. The scheme is fully compatible with the standard H.264/AVC encoder. The proposed scheme of UEP with slicing and DP of H.264/AVC constructs layered video data amenable to FEC with EW-RLC. The transmission of HPL layer only has been shown to offset the worst channel conditions. This makes such adaptation both beneficial and necessary for low bandwidth channels. This optimal scheme for rate adaptation over multiple paths for wireless video transmission can adaptively operate at GOP level.

The layered MDC scheme designed in this section with H.264/AVC is used in the next section to create layered MDC video for transmission over relay based channels simulating the emerging LTE-A standard.

5.4 System Description for Relay-assisted Rateless Layered MDC

In this section the system model used in the following sections is described. A source node generates source blocks that need to be streamed in real time to a mobile destination. The source node independently compresses each data block of length B symbols (e.g., bytes) into D descriptions, $D \ll B$, using an MDC scheme, and each description consists of l layers (e.g., quality, temporal, etc). After source coding the descriptions are fed into a rateless AL-FEC encoder, which for each description, generates potentially infinite stream of coded packets of L symbols each. Each encoded packet is a random linear combination of the source packets over sufficiently large field. Source can send the encoded packets to the destination via a direct link, or via $N-1$ relays available. The j^{th} channel ($j = 1, \dots, N$) is characterized by the total number of packets that can be transmitted until the deadline, which is denoted by R_j and depends on the channel bandwidth, and the packet loss probability, p_j . Note that R_j and p_j can in general be functions of time, and p_j also depends on the relay/source transmit power.

Note that though the source is exploiting rateless AL-FEC coding, due to real-time playout constraints, only a finite number of encoded packets R_j can be sent through channel j

for each encoded source block. The destination collects rateless packets from all available links, until it either receives enough to decode all descriptions or until playout deadline is reached. Note that each description can be decoded independently, descriptions are of unequal importance to final recovery, and each encoded description can only improve the reconstruction.

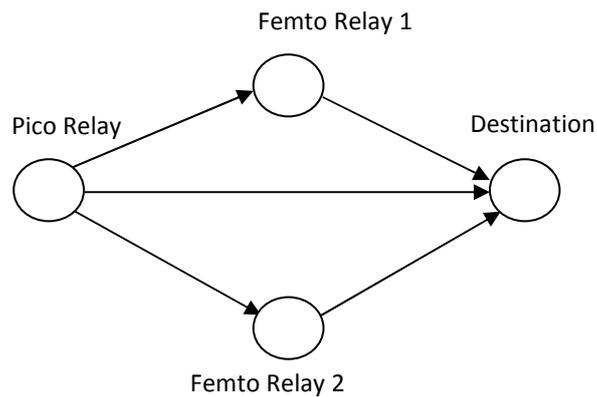


Figure 5.11: The block diagram of the system with two relays.

In particular, the above system can be used to model video streaming over LTE-A networks, where a base-station (source) in the pico-cell streams video to a mobile client via direct link and also via femto-cell relays as shown in Figure 5.11 for the case of two femto-cell relays. LTE-A transmission enables a more efficient operation of the Multimedia Broadcast and Multicast Service (MBMS) [12], allowing over-the-air combining of multi-cell transmissions towards the User Equipment (UE). Therefore a handheld mobile device can be served by multiple relays.

Due to strict time constraints of real-time streaming, it is assumed that the employed relays do not decode the received packets but operate in the amplify-and-forward mode. As in LTE, the pico-relay or the base station (BS) uses separate orthogonal links to send packets to the relays and destination. The relays also forward the packets to the client using orthogonal channels. Since the BS and the relays are static (and these connections are often optical), in our analysis, it is assumed that the source-relay channels are always error-free.

5.5 Resource Allocation

Given statistical knowledge of the source content and channel characteristics, the task is to find source-channel-relay resource allocation that maximizes the reconstructed video quality.

Let $\delta \subseteq \{1, \dots, D\}$ be a subset of descriptions. Let $\pi = \{r_{ij}\}$ be a $D \times N$ matrix of the rate allocation strategy over N channels, where r_{ij} , $i = 1, \dots, D$ and $j = 1, \dots, N$, is the rate (in terms of the number of encoded packets) of description i that will be sent over channel j . Note that the total number of encoded packets of description i is $\sum_j r_{ij} = \left\lfloor \frac{q_{s_i} q_{c_j} B}{L} \right\rfloor$, where q_{s_i} and q_{c_j} are source and channel coding rate for description i , respectively.

Matrix π is thus used at the source to select how many encoded packets from each description should be routed over each relay. However, since each description consists of l quality layers, encoded packets are not of equal importance. Thus, furthermore, for each description, the source probabilistically schedules encoded packets for transmission using channel coding probability matrix Q_{EW} (which will be explained later, see also [22]). Thus, Q_{EW} is used by the source to decide which particular packets for each description to send. This decision is based on the importance of each packet to recovery of the description or on the time playout information and channel characteristics.

Let $\rho = [\rho_1, \dots, \rho_N]$ be the set of transmitter powers at the source (ρ_1) and $N - 1$ relays. Then, the resource allocation problem that maximizes the total reward w is:

$$\max_{\pi, \rho, q_{s_i}, Q_{EW}} w = \max_{\pi, \rho, q_{s_i}, Q_{EW}} \sum_{\delta} P(\delta | \pi, \rho, q_{s_i}, Q_{EW}) w(\delta) \quad (5.1)$$

where $P(\delta | \pi, \rho, q_{s_i}, Q_{EW})$ and $w(\delta)$ are the probability of receiving δ and the reward in that case, respectively, given that strategy $(\pi, \rho, q_{s_i}, Q_{EW})$ is used. The above optimization is done under N channel constraints, namely $\sum_i r_{ij} \leq R_j$, D source coding constraints: $\sum_j r_{ij} \leq B/L$, and N power constraints $\rho_k \leq P_k$, where P_k is power constraint of the k -th transmitter.

Note that the optimization in (5.1) includes source optimization (q_{s_i}), relay power allocation (ρ), relay selection (π), and scheduling/channel coding (Q_{EW}). The above problem is a hard combinatorial problem that cannot be solved in real time. Thus, the following simplifications are introduced. First, assuming that the source block is pre-encoded, all q_{s_i} 's are fixed, and ρ, π, Q_{EW} are considered one at a time.

Power allocation ρ is done for each relay by looking only at the single relay-destination channel and selecting the lowest power level that will provide average packet loss rate below a certain threshold. If that is not possible, then $\rho_k \leq P_k$ is set. This strategy is justified by the fact that transmitter power is handled at the physical layer, in order to minimize the bit error rate for each link, and without any knowledge of the source.

Next, optimal π is found, assuming that no scheduling is done, i.e., all encoded packets of a description are treated equally. Then description i is decoded if at least $s_i \geq q_{s_i} B/L$ packets from this description are received. Note that for maximum separable codes, $s_i = \lceil q_{s_i} B/L \rceil$, with high probability, while for RLC codes s_i very slightly exceeds $\lceil q_{s_i} B/L \rceil$ (the number of source packets to be sent). In the analysis it is further assumed that R_j 's and p_j 's are constant during the transmission of one encoded source block for all $j = 1, \dots, N$. Even with fixed q_{s_i} and ρ , optimizing π is a difficult optimization problem. It is simplified by assuming that if the destination decodes description i , it gets reward w_i , where w_i is estimated based on the source recovery using only description i . The main motivation for this simplification is that it will reward descriptions that contribute highly to the reconstruction while significantly reducing optimization complexity. The total reward is the sum of the rewards for all decoded descriptions. Thus (5.1) is optimized restricting δ to be of cardinality 1. With a slight change of notation, for clarity (5.1) is written as:

$$\max_{\pi} w(\pi) = \max_{\pi} \sum_{i=1}^D P_i(\pi) w_i \quad (5.2)$$

where $P_i(\pi)$ and w_i are the probability that description i is recovered if π is used and the reward in that case, respectively. Next the algorithm is used that efficiently solves optimization problem in (5.2).

5.5.1 Global Optimization – Algorithm 1

Without loss of generality, it is assumed that all available resources are allocated. Therefore,

$\sum_{i=1}^D r_{ij} = R_j$. Then, the resource allocation problem is

$$RA(R_1, \dots, R_N, p_1, \dots, p_N; s_1, \dots, s_D, w_1, \dots, w_D)$$

whose optimal solution π^* is the one that maximizes (5.2).

Brute force complexity. For a channel with R_j , there are $(R_j + 1)^D / D!$ ways of allocating the D descriptions. Therefore, the brute force complexity is $\sim \left(\prod_{j=1}^N (R_j + 1) \right)^D \sim O(R^{ND})$, where R is the geometric mean of $R_j, j = 1, \dots, N$.

Dynamic programming solution. Note that if π^* is the optimal solution of

$$RA(R_1, \dots, R_N, p_1, \dots, p_N; s_1, \dots, s_D, w_1, \dots, w_D),$$

then the sub-solution defined by sub-matrix $\pi^* (1 : D-1, :)$ has to be the optimal solution of $RA(R_1 - r_{D,1}, \dots, R_N - r_{D,N}, p_1, \dots, p_N; s_1, \dots, s_{D-1}, w_1, \dots, w_{D-1})$.

The advantage can be taken of this fact to decrease optimization complexity by dynamic programming. In a nutshell, the algorithm will first seek the optimal solutions for all combinations of $1 \leq R'_j \leq R_j$ and all $1 \leq d \leq D$. At $d = 1$, the problem is trivial as the optimal allocations for $RA(R'_1, \dots, R'_N, p_1, \dots, p_N; s_1, w_1)$ is simply $\pi^* = [R'_1, R'_2, \dots, R'_N]$ (i.e., allocate all resources to the first description).

Assume now that solutions for d' are available. For $d = d' + 1$, computing optimal solutions of $RA(R'_1, \dots, R'_N, p_1, \dots, p_N; s_1, \dots, s_{d'+1}, w_1, \dots, w_{d'+1})$ will only involve scanning through the $\Pi_i R'_i$ solutions of the previous level. Therefore finding all optimal solutions for this level ($d' + 1$), involves approximately $\left(\prod_i (R'_i + 1) \right)^2 / 2$ computations.

Note that, to compute the optimal solution of $RA(R_1, \dots, R_N, p_1, \dots, p_N; s_1, \dots, s_D, w_1, \dots, w_D)$ will require going through all D levels and thus have complexity $(D - 1) \left(\prod_j (R_j + 1) \right)^2 / 2$ or approximately $D \left(\prod_{j=1}^N (R_j + 1) \right)^2 \sim O(DR^{2N})$ instead of $O(DR^{ND})$, needed for the brute force method.

5.5.2 Proposed Fast Solution

Algorithm 1 finds the optimal solution to the allocation problem in (5.2). However, its complexity and memory requirements might still be too large for real-time applications, especially when N or R_j 's are large. That is why; a simple and fast ad hoc strategy is also proposed. It will be shown in Section 5.6 that this strategy works well for the video MDC, where the descriptions do not vary largely in importance. First, the channels are arranged in the decreasing order of quality, i.e., such that $p_1 \leq p_2 \leq \dots \leq p_N$.

That is, Channel 1 (Ch1) is the best channel and so forth. Next, the descriptions are arranged such that $w_1 \geq w_2 \geq \dots \geq w_N$. That is, D_1 is the most important description, and so forth.

Without loss of generality it is assumed that $D \leq N$. If $D > N$ the descriptions could be merged to end up with total of N .

Algorithm 2 proceeds as follows: Put D_i into Ch_i either until (1) Ch_i is filled; or (2) $\frac{(1+\epsilon)s_i}{1-p_i}$ packets have been put, where ϵ is heuristically set to 10%. In the next step, channels $1, \dots, D$ that have not been filled are filled equally with the descriptions for which condition (1) occurred, or if there are no such descriptions, Ch_i is filled with description i . Channels $D + 1, \dots, N$ are filled equally with all descriptions.

The algorithm can be made blind of the source, by arranging descriptions in arbitrary order. Note that, Algorithm 2 is applicable to any channel model by properly selecting ϵ .

5.5.3 AL-FEC and Packet Scheduling

AL-FEC is a flexible software-based solution that can effectively combat fading and easily be tied to source coding via source-channel coding algorithms (see [55] and references therein). Digital Fountain Raptor codes [9], as most efficient rateless packet loss codes, have been adopted for AL-FEC in DVB-H and MBMS [35]. The main problem of Raptor codes is the fact that they equally protect the entire stream and have a sharp avalanche effect. Thus, they are not very suitable to applications where importance of the sent packets significantly varies, which is the case in video coding. EWF codes are a recent solution to this problem proposed in [21].

RLC codes provide similar performance as Raptor codes for point-to-point communications [55] at the expense of increased decoding complexity. However, it is shown in [55] that if the used source block is small, RLC performs similarly to Raptor codes, and have acceptable decoding complexity via progressive GE [20] [36].

The main advantage of RLC compared to Raptor codes comes in the multipath scenarios, as RLC effectively apply network coding concepts in the multi-node settings [13].

The encoding process for EW-RLC starts by selecting a window from which the RLC encoded symbol is to be generated. These probabilities of window selections are organized into a vector Q_{EW} and are equivalent to channel coding rates for each source layer/window. After a window is selected, the encoding is the standard RLC encoding performed over the source packets contained in that particular window only [22]. In NOW scheme, the windows are non-overlapping and each window is protected with a separate AL-FEC code. The performance of EW-RLC schemes is concluded to be better than the corresponding NOW-RLC schemes, and this gain comes from the increased coding flexibility introduced by overlapping the windows.

In this section EW-RLC over l windows are employed from consecutive source blocks that are of unequal importance, i.e., each description is organized into l layers of unequal importance, which form l windows of the EW-RLC scheme.

To provide optimal allocation of redundancy to different source priority classes, for each description i one can maximize the expected reward using analytically computed probabilities of decoding error performance. That is,

$$\max_{Q_{EW}^i} w_i(Q_{EW}^i) = \sum_{k=0}^l P_i(k|Q_{EW}^i) w_i(k) \quad (5.3)$$

where Q_{EW}^i is an l -length vector of window selection probabilities that determines the UEP allocation scheme, $P_i(k|Q_{EW}^i)$ is probability that layer k of description i will be the highest layer recovered if Q_{EW}^i is used, $P_i(0)$ is the probability that nothing is recovered, $w_i(k)$ is reward if all layers up to and including layer k are recovered, $k = 1, \dots, l$. Analytical expressions for probabilities that a layer is recovered assuming a random channel loss model can be found in [22].

5.6 Simulation Setup

In this section, the specific source-channel-network scheme is described along with the parameters used.

5.6.1 Video Coding and AL-FEC

In H.264/AVC error resilience via slicing splits the frame into multiple slices that are independently encoded and separated by resynchronization points. Since slices can be made independently decodable, the partitioning of a frame into slices can be used to create multiple descriptions with fine granularity. In this section, $D = 2$ descriptions, namely, MDC1 and MDC2, have been created by assigning three slices from each frame to each description as described in Section 5.3.1.

Firstly, both descriptions have the intra-coded IDR [2] as the first entry. Thereafter, starting from the first following frame, alternate slices are copied to each description. This way, the overlap in the source content between two descriptions is minimal.

Table 5.11: Compression results (size in bytes and PSNR in dB) for the first GOP of the Foreman sequence.

Layer	MDC1			MDC2		
	Size	PSNR	Packets	Size	PSNR	Packets
BL	27,960	25.76	39	31,871	25.58	44
EL	23,980	26.79	33	23,539	26.2	33
Total	51,940	26.79	72	55,410	26.2	72

After video encoding, two independently decodable descriptions of possibly different sizes are generated. Each description contains $l = 2$ quality layers obtained by grouping IDR + DP A + DP B into one layer (i.e., the base layer (BL)) and DP C into the second quality enhancement layer (EL).

Table 5.11 shows the peak signal-to-noise ratio (PSNR) and the rate (in bytes) for the first GOP of size 16 frames for the standard CIF Foreman video sequence. The video is encoded at 25 frames per sec and the total compression rate is 1.2 Mbps.

The best video quality of 41.07 dB is achieved when both MDCs are received completely. However, it can be seen from Table 5.11 that due to the way the descriptions have been designed, the video can be decoded with the BL of any one description only. This layering of video data within a description makes this scheme adaptive and practical for varying channel conditions. In the case where both descriptions are lost entirely and nothing is decodable for the GOP then the decoder applies a simple error concealment technique by replacing the lost GOP by the last frame of the previously successfully decoded GOP. Each description is independently protected against packet losses by an EW-RLC. The BL forms the first window, W_1 , and the (BL+EL) forms the second window, W_2 .

5.6.2 Simulation Parameters

The BS compresses the video, performs RLC protection, relay selection based on the feedback from the relay regarding the quality of the relay-destination link and the relay transmit power, and schedules encoded packets for transmission.

The setup with two relays and a direct link from the BS to the destination is considered, thus $N = 3$, as shown in Figure 5.11, with pico-relay being a BS. The distances between the destination and BS, femto relay 1 and femto relay 2 are 70m, 25m, and 20m, respectively [75]. To model fading channel with two femto-relays, FSMC is used that reflects the change in the bit error rate depending on Markov chain state-space resulting in different received data rates. In the model each wireless Rayleigh block-fading channel is represented by a packet-level FSMC channel model [81] that adapts with the mobility of the user and transmit power.

For proof of concept, 2.2 GHz carrier frequency and uncoded BPSK are used. In the first set of experiments, the speed of the mobile is 3 km/hr. The data transmission rate is 1000 kbps, 600 kbps and 550 kbps for the BS, Femto 1 and Femto 2 relays, respectively. This corresponds to the maximum number of encoded packets that these paths can take during one source block (one encoded GOP) of 100, 60, and 55, respectively.

Symbol sizes of 734 bytes are used for RLC. A header overhead of 60 bytes per packet is assumed to cater for the headers added at the various protocol layers, e.g., RTP/UDP/IP (note however that robust header compression can significantly reduce the size of the header). Only one symbol is placed in the packet, thus transmission packets are of size $734+60=794$ bytes. GE decoding is performed on source block sizes of 72 and 77 for the first and the second description, respectively. These low block sizes enable fast decoding on a smartphone without sacrificing performance [20]. The reconstructed quality is assessed using PSNR and optimization (Algorithms 1 and 2) is performed with reconstructed frame-average PSNR as reward w .

5.6.3 Simulation Results and Discussion

The simulations have been done for two different power constraints of the two femto relays: in Configuration 1 the required average packet loss rates are set in the femto relay 1 and 2 links to 0.25 and 0.3, respectively. This results in physical transmitter powers of 30 dBm and 35 dBm, respectively, which translates into the averaged received signal-to-noise ratio (SNR) of 23 dB and 33 dB, respectively.

In Configuration 2 the required average packet loss rates are set in the femto relay 1 and 2 links to be 0.05 and 0.1, respectively. This results in the physical transmitter power of 47 dBm, for both relays, which translates into the average received SNR of 40 dB and 44 dB, respectively.

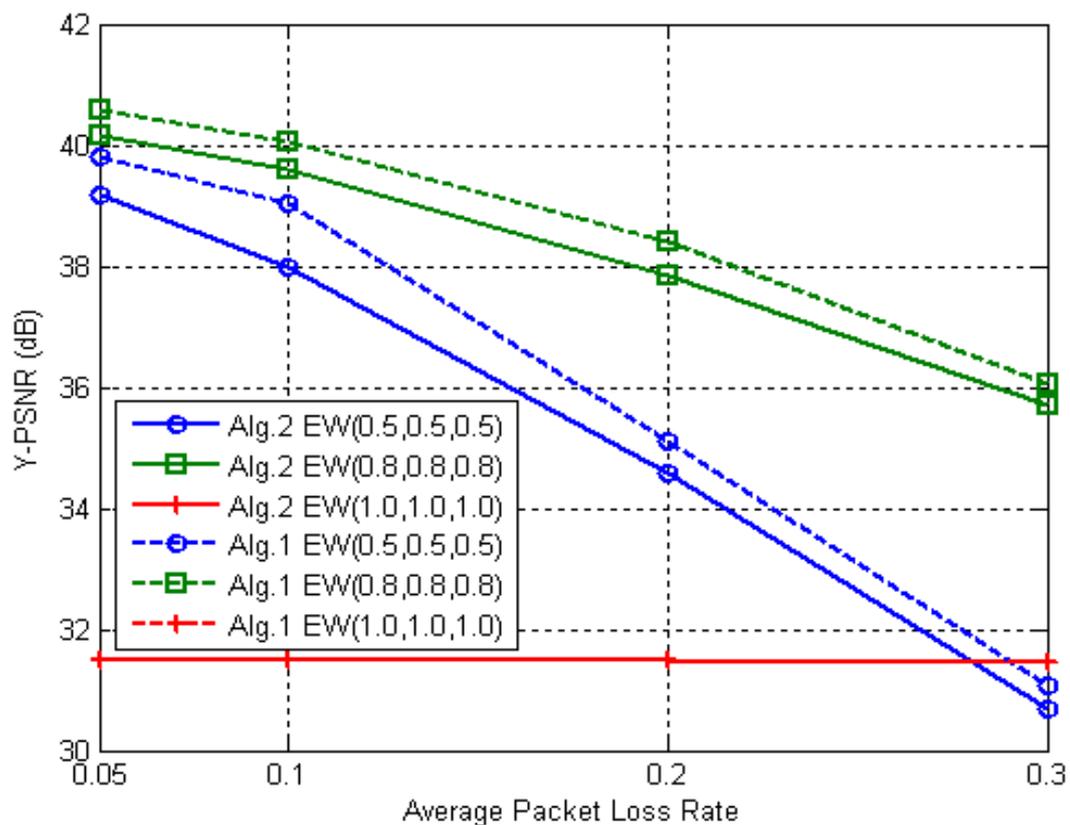


Figure 5.12: Frame average Y-PSNR vs. average packet loss rate for Configuration (1) for the FSMC channel – Foreman sequence.

For the above femto relay transmit power constraints, the BS power level is varied in the range of 40 to 70 dBm, which results in the average received SNR between 12 dB and 43 dB. Then, the average packet loss rate in the direct link varies from 0.05 to 0.3.

Several UEP EW-RLC schemes are selected, obtained by changing Q_{EW} to test their robustness to varying channel conditions. Note that the same Q_{EW} is always used for the two descriptions.

In this section simulation results are presented for the CIF Foreman and Container sequences compressed with H.264/AVC into two descriptions using parameters from the previous section.

The results for Configuration (1), comparing Algorithms 1 and 2 are shown in Figure 5.12. The term $EW(x, y, z)$ refers to a set of selection probabilities for windows W1 (BL), such that the selection probability of the BL for both descriptions for the direct link, link via Relay 1 and link via Relay 2 are x , y , and z , respectively. It can be seen from the figure that the two algorithms show similar performance. Indeed, Algorithm 1 is for only 0.5-1 dB better. Note that, in the $EW(1.0,1.0,1.0)$ scheme only window W1 is selected, hence the PSNR cannot go beyond that achievable with the two BLs.

Setting the probability of selection of the BL to 0.8 ($Q_{EW} = (0.8, 0.2)$) provides the best results for all source power levels. Note that in this case, 80% of rate is allocated to the BL, but still even at the loss rate as high as 30%, ELs are often recovered showing a PSNR gain of 4 dB compared to the case when only BL is sent alone.

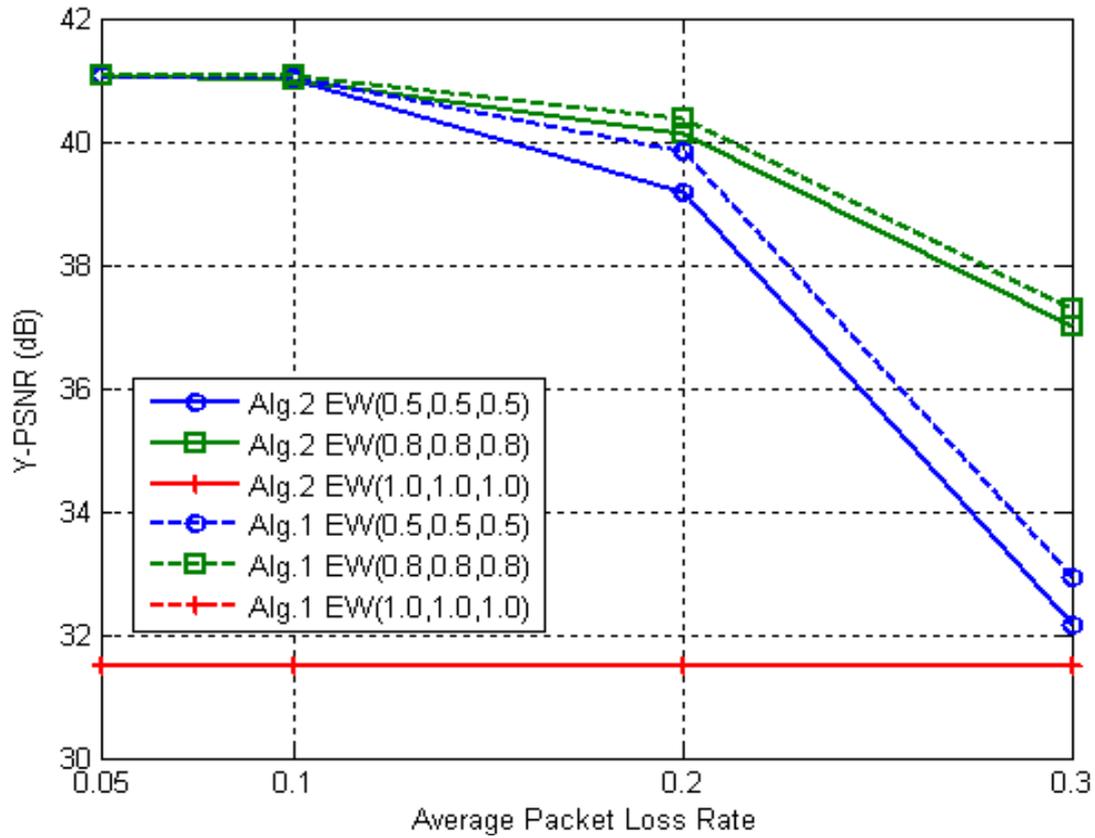


Figure 5.13: Frame average Y-PSNR vs. average packet loss rate for Configuration (1) for the random loss channel model- Foreman sequence.

The results for the random loss channel model are shown in Figure 5.13. Note that, the same average channel loss rates are used as those given by the FSMC model. As expected, the performance of Algorithm 1 is consistently better than that of Algorithm 2. However, the difference is very small (under 1dB). The EW(1.0,1.0,1.0) scheme again performs poorly, since it does not exploit the available bandwidth to the fullest. Indeed, a gain of 7-10 dB can be observed by sending ELs.

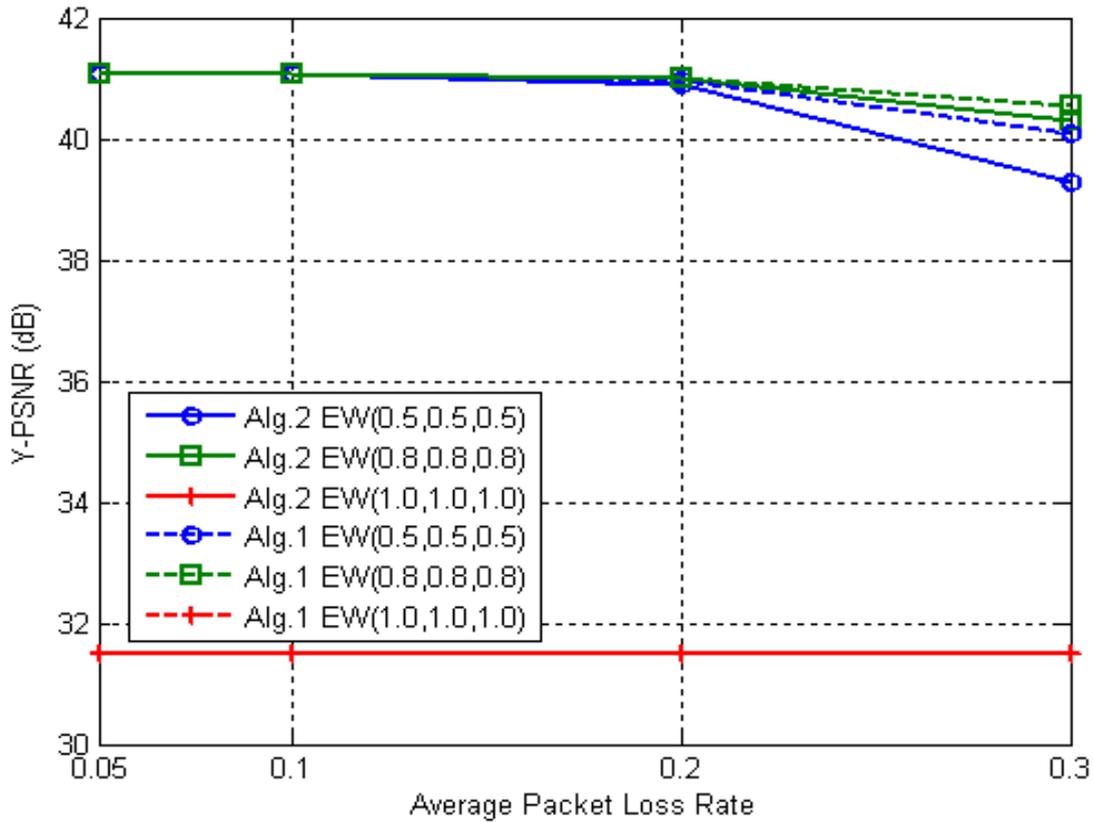


Figure 5.14: Frame average Y-PSNR vs. average packet loss rate for Configuration (2) for the FSMC channel – Foreman sequence.

The results for the random channel are for up to 0.5dB better than those for the FSMC channel, which demonstrates that the proposed schemes are robust to the bursty nature of the FSMC channel.

The results for Configuration (2) for the FSMC channel are shown in Figure 5.14. Setting the probability of selection of the BL to 0.5 or 0.8 provides the best results. Algorithm 1 provides a negligible performance gain over Algorithm 2 only at the highest packet loss rates. Similar results are obtained for the random loss channel.

In the next two figures results of Algorithm 2 are compared to the benchmark scheme - a scheme without relays that transmits a single-layer AVC (without MDC). For the benchmark scheme, three different source rates are used, 500 kbps, 700 kbps, and 900 kbps, which in error-free case result in the frame average Y-PSNR of 36.75 dB, 39.13 dB, and

40.13 dB, respectively. Note that this scheme equally protects all source packets using the same RLC codes as the proposed schemes.

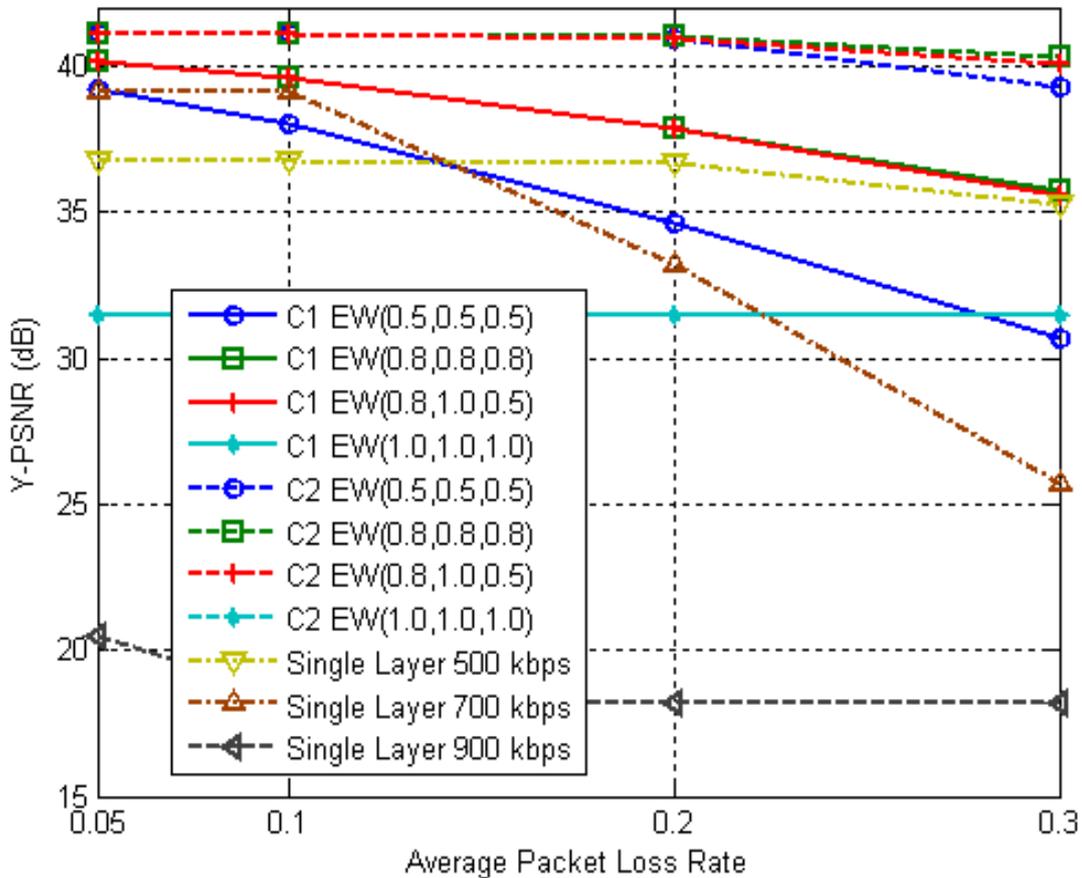


Figure 5.15: Frame average Y-PSNR vs. average packet loss rate for both configurations for the proposed Algorithm 2 and the single layer schemes for the FSMC channel for the mobile speed of 3 km/hr – Foreman sequence.

Figure 5.15 shows results for the FSMC channel and both configurations and four different scheduling strategies. It can be seen that Algorithm 2 significantly outperforms the single layer no relay schemes. Indeed, the gain to the best performing benchmark schemes ranges from 1 dB, at the lowest channel loss rates, to 5 dB at the highest channel loss rate. It can also be observed that by increasing the power levels of the two relays for 9 dB in total, a gain of roughly 5 dB at the highest channel loss rate and 0.5 dB for the lowest channel loss rate is obtained.

It can be seen from the figure that 900 kbps is too high source rate for this FSMC channel. Indeed, this benchmark scheme is under-protected resulting in very poor performance. The source rate 500 kbps benchmark scheme, on the other hand, is over-protected at the lowest channel loss rates. Thus, the benchmark scheme would have to adapt the source rate to the channel conditions.

Similar results to Figure 5.15 were obtained with the Container sequence which is as shown in Figure 5.16. The plots have similar trends except that in Figure 5.16 the PSNR values are comparatively higher. This serves to validate the analysis for Figure 5.15 with the Foreman sequence.

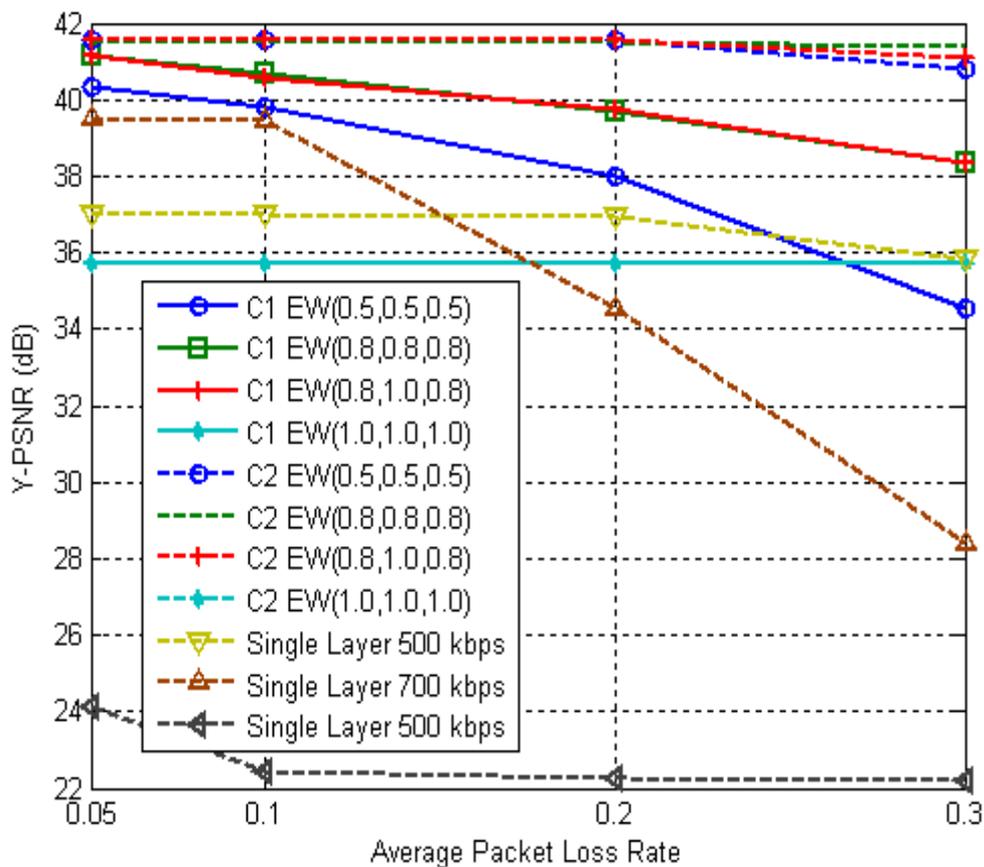


Figure 5.16: Frame average Y-PSNR vs. average packet loss rate for both configurations for the proposed Algorithm 2 and the single layer schemes for the FSMC channel for the mobile speed of 3km/hr - Container sequence.

Another single-layer no-relay benchmark scheme was tested that uses the same source power level as the sum of the powers of both relays and the BS and the source rate of 900 kbps showing worse performance than the best proposed one for up to 1 dB. Note that this scheme experiences the average packet loss rate less than 0.05, and thus is not shown in the figure. The very high BS power levels used (significantly exceeding typical power levels in LTE-A [75]) lead to SNR in the range of 80-110 dB which with channel coding rate of 9/10 results almost always in error-free performance. Note that the benchmark schemes do not use MDC, and hence provide better performance in the error-free scenarios. Using relaying, power is effectively distributed among three nodes providing increased performance compared to the single-layer no-relay schemes.

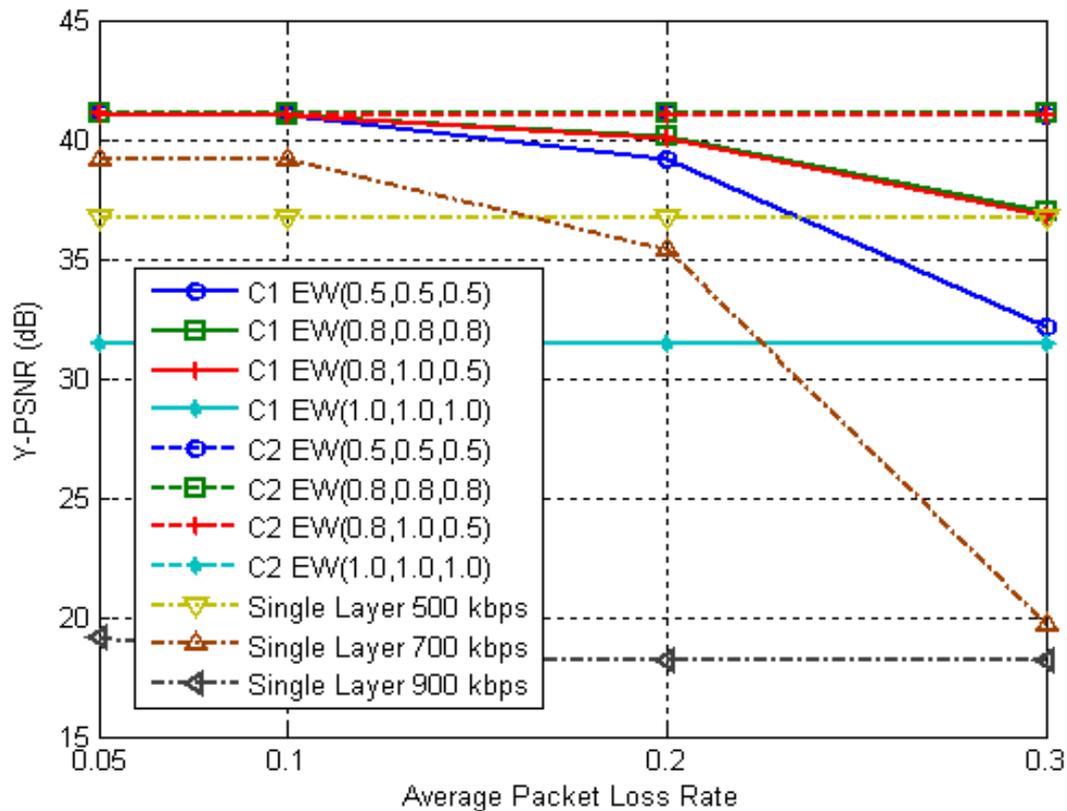


Figure 5.17: Frame average Y-PSNR vs. average packet loss rate for both configurations for the proposed Algorithm 2 and the single layer schemes for the random loss channel model for the mobile speed of 3km/hr – Foreman sequence.

The results for the random loss channel model are shown in Figure 5.17. The optimized relay-assisted scheme performs for 2 dB better than the single layer scheme at the lowest packet loss rate and 5 dB better for the highest packet loss rate. Results for Configuration 2 are for up to 5 dB better than those for Configuration 1. Note that for both type of channels,

Configuration 1 scheme at 0.3 packet loss rate shows roughly the same performance as the benchmark 500 kbps scheme, because in this case, there is only a little use of relaying due to high loss rates at the two relay links.

From these results, it is obvious that the relay-assisted schemes are much more robust to the change in the channel conditions of the direct link. Sub-optimal and fast Algorithm 2 shows close performance to Algorithm 1. The proposed schemes are robust to the change in channel conditions at the relay links. It is important to send both BL and EL for both descriptions to obtain high performance. Scheduling packets at the source (varying Q_{EW}) is important at the high packet loss rates and can provide a gain of over 10 dB.

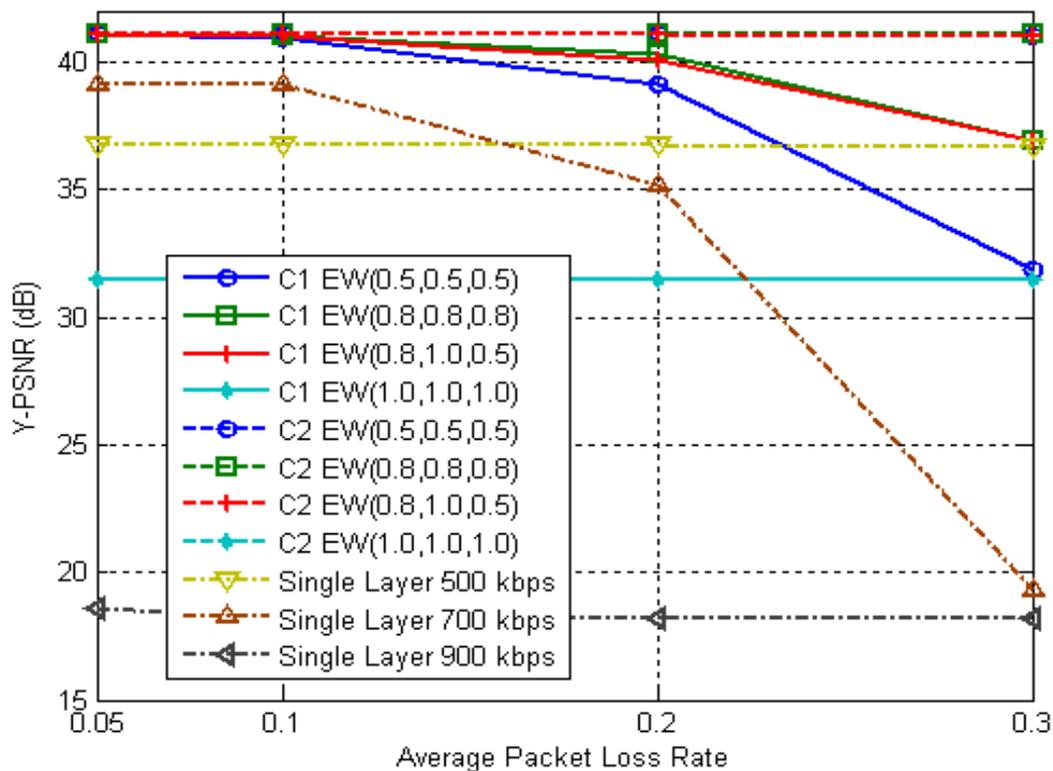


Figure 5.18: Frame average Y-PSNR vs. average packet loss rate for both configurations for the proposed Algorithm 2 and the single layer schemes for the FSMC channel for the mobile speed of 30 km/hr – Foreman sequence.



(a) M1M2, PSNR =41.56dB.



(b) H1H2, PSNR = 35.69dB.



(c) M1H2, PSNR =38.22dB.



(d) H1M2, PSNR =37.68dB.



(e) M1, PSNR = 33.77dB.

(f) M2, PSNR = 32.9dB.



(g) H1, PSNR = 32.9dB.

(h) H1, PSNR = 32.41dB.

Figure 5.19: Frame 13 for Container sequence for different combinations.

Figure 5.18 shows results for the FSMC channel and the mobile speed of 30km/hr. Comparing Figures 5.17 and 5.18; one can see that the proposed solutions are very robust to the increase in mobile speed.

Figure 5.19 shows frame 13 for the Container sequence. Each of these 8 combinations is sufficient in itself to re-create the whole 16 frames of the original GOP. The two-letter captions such as M_1H_2 implies that MDC1 has been received in full whereas for MDC2 only its HPC has been received. Similarly, one-letter caption such as H_1 means that only HPC of MDC1 has been received, whereas MDC2 has been completely lost. It is interesting to note that despite no perceptible visual difference between the images, the PSNR values are significantly different.

5.7 Summary

The real-time transmission of layered MDC video over relay-assisted paths is considered. MDC scheme is designed using slicing and data partitioning features of H.264/AVC and the resulting packets are fed into the EW-RLC encoder for erasure protection. The encoded packets were streamed over direct link and over two relay-assisted channels. A resource optimization framework was developed, using dynamic programming and the proposed fast algorithm that sub-optimally select relays and schedule packets for transmission.

There would be impact on battery life for multi-interface solution however it is small as compared to video decoding and display itself. Also, the solution proposed here is based on EW RLC which is comparable to Raptor codes in complexity. Raptor codes are already employed in phones so a short code length EW RLC should not have a large impact on such solutions.

In order to reduce the encoding/decoding complexity of RLC it can be advantageous to use systematic codes, which are treated in Chapter 6 together with a novel algorithm for sorting slices.

Chapter 6

Systematic RLC for Sliced H.264/AVC Video Streaming

6.1 Introduction

In this chapter, the use of systematic RLC is demonstrated as an application layer FEC solution to protect the transmission of slice-partitioned H.264/AVC data. The proposed solution is generic and can be used for a unicast or multicast solution. A system is proposed which exploits systematic EW RLC for H.264/AVC slice-partitioned data. The proposed system prioritizes slices based on their PSNR contribution to reconstruction as well as temporal significance. It is demonstrated by the simulation results that using relative slice priority with systematic codes can be useful for multimedia broadcast applications.

As compared to the H.264/AVC data partitioning feature [2], slicing has an advantage that the size of slices can be tailored to the application. The slicing feature of H.264/AVC can be used to partition video stream into classes of decreasing importance (for video reconstruction) with a very small decrease in overall performance. A scheme has been proposed in [89] based on macroblock classification into three slice groups and UEP of H.264/AVC streams. The ordering of macroblocks into three slice groups is done by examining their contribution to the video quality. The three slice groups are then protected with UEP using Reed-Solomon (RS) coding. In [59], a slice sorting by relevance (SSR) algorithm for prioritizing slices based on their contribution to the reconstruction is used together with RS coding. The work in [59] is later extended in [90] and proposes an algorithm termed Concealment Driven Slice Ordering with RS codes. The ordering of slices is based on error propagation effect and the rate devoted to each slice. The work in this chapter differs from the earlier work in the choice of method of prioritizing slices and use of systematic rateless codes for channel coding. The slice-partitioned video stream can provide an advantage with respect to H.264 Scalable Video Coding (SVC) [4] of better coding efficiency and compliance with the AVC standard. The layered video can be protected by EW RLC codes that can provide a different degree of protection to each layer/window.

The focus of this chapter is to analyse the use of the EW approach with systematic RLC as component codes for UEP of the slice-partitioned H.264/AVC video. Systematic RLC

have the advantage of supporting more efficient encoding and decoding procedures compared to non-systematic RLC.

In contrast to [59], where priority layers are built based purely on distortion information, in this chapter a new cost function is proposed that takes into account the frame play out deadline and temporal error propagation to better prioritize slices into quality layers. The simulation results show that EW RLC can be used to effectively protect the different priority windows for reliable video transmission over packet erasure channels. Significant performance gains are obtained compared to the equal error protection scheme and the benchmark scheme that prioritizes the sliced stream in an ad hoc fashion.

The rest of the chapter is structured as follows. The proposed system is described in Section 6.2. The experimental results and analysis are presented in Section 6.3. Finally, Section 6.4 summarizes the chapter.

6.2. The Proposed System

In this section, a system is proposed for optimally protecting the slice-partitioned H.264/AVC video data with systematic EW RLC. It is assumed that the encoded video stream is transmitted over a packet loss channel. That is, all packets that arrive at the application-layer RLC decoder are correct, while those with bit errors are discarded by error detection codes, such as Cyclic Redundancy Check (CRC) codes present at the lower layer in the protocol stack (e.g., physical or link layer). It is further assumed that error detection capability of the employed CRC codes is perfect, which is usual assumption [22], [55], [91],[59]. Thus, the application layer-to-application layer channel is modelled as packet erasure channel with random packet drop statistics.

In order to increase error resilience, a video sequence is encoded using slicing with a fixed slice size of 600 bytes. That is, after the H.264/AVC encoding, the video data is obtained in which each frame including the IDR is divided into slices of 600 ± 3 bytes, except for the last slice of each frame which may have a lesser size. The size of 600 bytes is chosen here to keep the number of RLC symbols per codeword low in order to reduce the decoding complexity of GE. See [20] and [37] for discussion about acceptable block lengths for real-

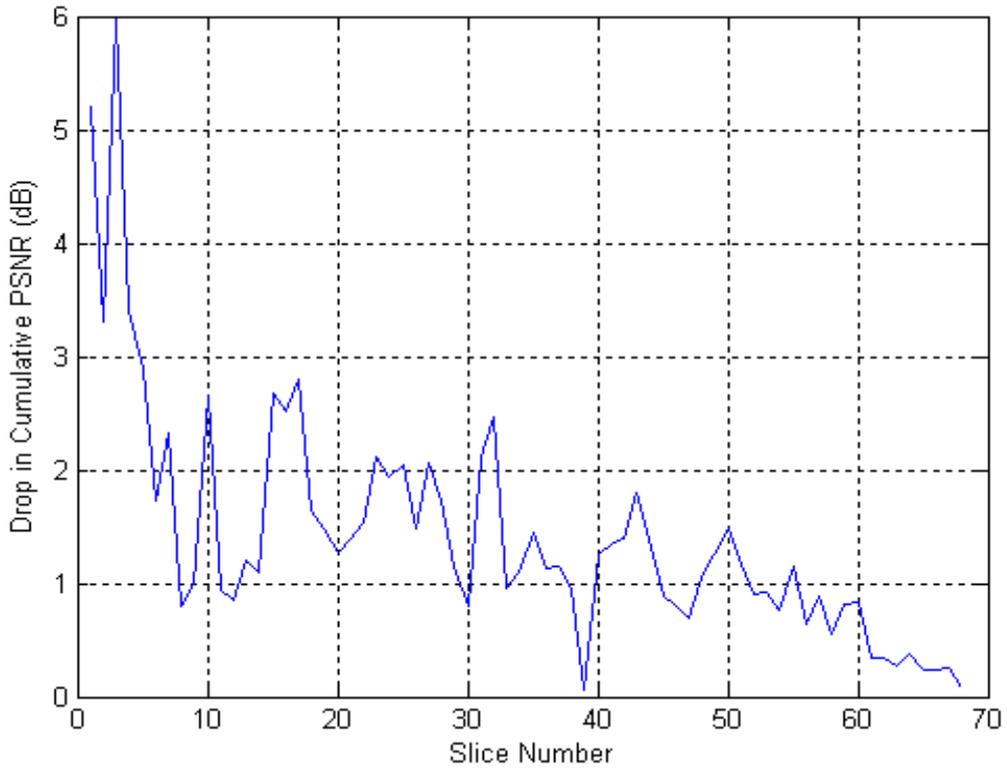


Figure 6.1: Drop in PSNR for non-IDR slices – Foreman sequence GOP16.

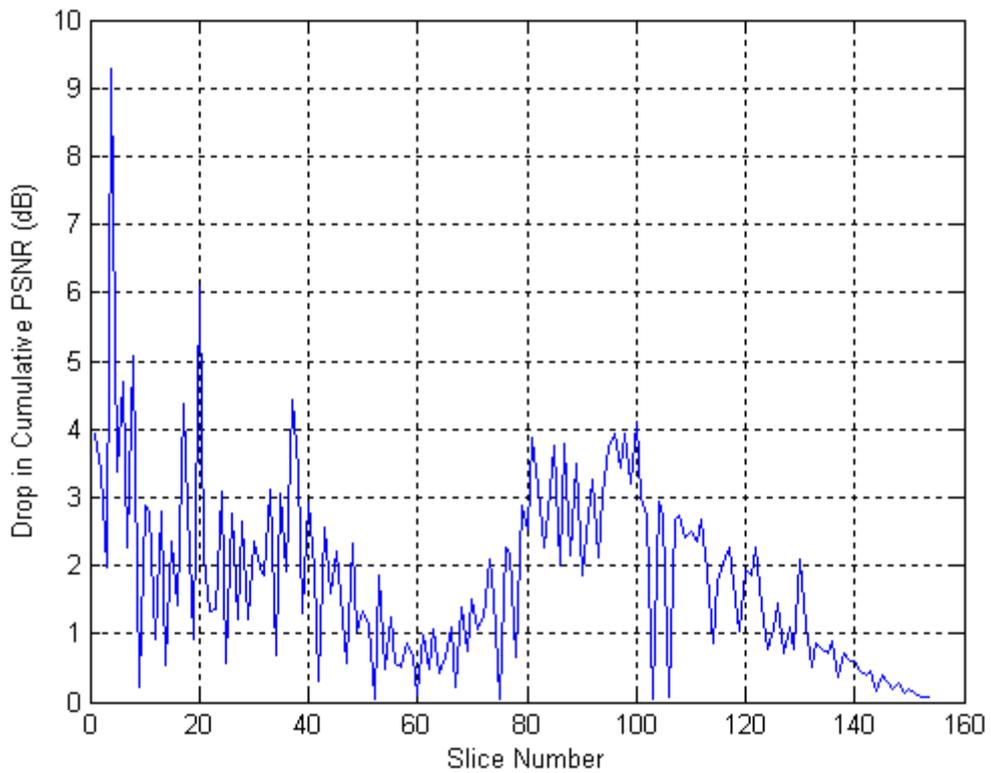


Figure 6.2: Drop in PSNR for non-IDR slices – Paris sequence GOP64.

time RLC decoding. The resulting slices carry different importance to reconstruction which has been used to achieve UEP (see [59], [68] and references therein).

After source coding, EW RLC coding takes place. Since systematic RLC are used, first all encoded symbols (from all the slices) are transmitted without any coding. Because of possible errors/erasures in the channels some packets will be missing at the decoder. To correct these erasures, RLC redundancy packets are generated next.

Before RLC, the priority of each slice is obtained by dropping it from the GOP data and measuring the resulting PSNR, as a frame-by-frame average of the entire GOP, by actual decoding. Determining PSNR drop can easily be done during the encoding process with negligible added complexity [59].

This also takes into account the error propagation effect to the subsequent frames due to loss of a slice in an earlier frame. That is, the cumulative PSNR of the GOP is measured by dropping each slice in turn starting at the first P frame. After having obtained the cumulative PSNR values for each slice (as dropped), the difference from the full-decoding PSNR of the GOP is measured. The results are shown in Figure 6.1 for the first GOP (having 16 frames and the encoding structure IPPPP...) of the standard CIF Foreman sequence. It can be seen from Figure 6.1 that the importance of the slices on total frame-averaged PSNR generally decreases on moving towards the end of the GOP. Similar results are shown in Figure 6.2 for the first GOP (having 64 frames and the encoding structure IPPPP...) for the Paris sequence. As can be seen from the figures, the PSNR drop values for Paris sequence are larger due to large GOP size. Thus the slices can be sorted into multiple priority layers and a higher degree of protection can be assigned to the important layers as compared to the layers containing less significant slices. Such layering enables a prioritized data transmission with UEP schemes and was used before in [59],[68].

Purely grouping the slices into priority classes based on the PSNR decrease shown in Figure 6.1, as done in [59],[68] does not take into account the real-time frame playout deadline (frames coming sooner should be given a higher priority).

Motivated by this, the cost function used in [68] is redefined, to take into account not only the drop in cumulative PSNR for each slice, but also the temporal importance of a slice:

$$W(\text{slice}) = D(\text{slice}) - w * \tau(\text{slice}), \quad (6.1)$$

where $D(\text{slice})$ represents the drop in the cumulative PSNR (see y-axis of Figure 6.1). The value of $\tau(\text{slice})$ represents the playout time deadline of the slice (frame) relative to the playout time of the first IDR in the GOP. That is, $\tau(\text{slice})$ of the IDR frame is set to zero, and each subsequent frame adds its playout time duration to this value. w is a constant that trades off the distortion $D(\text{slice})$ and remaining playout time.

In this way, a system is created to assign a priority to all the slices trading off importance of the slices to reconstruction and playout time deadline. After computing $W(\text{slice})$ for all the slices in a GOP, and selecting a threshold values $T_1 > T_2 > \dots > T_{L-1}$, the slices are grouped into L layers.

The first layer includes IDR and slices with $W(\text{slice}) \geq T_1$, the second layer includes all remaining slices with $W(\text{slice}) \geq T_2$, and so forth. In addition, our algorithm also puts at least one slice per frame to the first layer, if none is selected (from a frame) based on the above criteria alone. This helps to stop the error propagation effect further and thus improves resulting PSNR. Such selections may be needed for frames which occur towards the end of GOP as can be seen from Figure 6.1.

In the proposed scheme, L windows are created using a threshold $(L-1)$ -tuple T_1, \dots, T_{L-1} , and allocate different protection to each window. Note that the slices would already be in their decoding order within each layer. However, within each window the slices will need to be restored to the original order to enable decoding by the AVC decoder.

After determining thresholds and assigning slices to the L windows, the task is to find the optimal allocation of redundancy to each layer, or equivalently probability of window selection. For this task, the expected PSNR is maximized using analytically computed probabilities of decoding error performance. That is,

$$\max_{\pi} PSNR(\pi) = \sum_{i=0}^L P(i) psnr(i), \quad (6.2)$$

where $P(i)$ is probability that layer i will be the highest layer recovered, $P(0)$ is the probability that nothing is recovered, $psnr(i)$ is the PSNR of the reconstruction if all layers up to and including layer i are recovered, π is an L -tuple vector of window selection probabilities that determines the UEP allocation scheme, and $PSNR(\pi)$ is the expected PSNR when UEP scheme π is used. In the above maximization, it is assumed that if decoding of window i fails, none of the packets from window $j \geq i$ can be used for reconstruction. This is true for non-systematic EW RLC and approximation for systematic EW RLC.

The analytical expression [22] for $P(l)$ for a random channel loss model for EW RLC is as given below:

$$P(l) = \begin{cases} 1 - P_{d,N}(1), & l = 0 \\ \prod_{i=1}^l P_{d,N}(i) \cdot (1 - P_{d,N}(l+1)), & 1 \leq l \leq L-1 \\ \prod_{i=1}^L P_{d,N}(i), & i = L, \end{cases} \quad (6.3)$$

where $P_{d,N}(l)$ [22] is the probability of decoding of the l^{th} layer after N received symbols.

The optimization method is of linear complexity but requires rate-distortion values for each slice that can be obtained during the video encoding process. The optimization method is exhaustive search and scales linearly with the number of UEP schemes being used.

6.3. Results and Analysis

6.3.1 Comparison between Systematic and Non-systematic Codes

In this section the simulation results are presented. For simplicity the case of $L=2$ layers is considered: High priority layer (HPL) that contains more important slices, whose $W(\text{slice}) \geq T$, and Low priority layer (LPL) that contains less important slices for which $W(\text{slice}) < T$, where T is the chosen threshold.

The thresholds determine the source rate for each layer. For example, a lower T would result in a higher source rate (and hence, error-free performance) for the base layer. Thus, T_i 's are set based on available clients' bandwidths as well as desired error-free performance levels. In practice, transmitter can dynamically adapt the source rate per layer to varying channel conditions of different clients by changing T_i 's.

Table 6.1: Layer sizes and PSNR contributions for $T=0.78$ and $w=2.5$ for Foreman sequence.

Layer	Proposed			Benchmark		
	Size (bytes)	Pkt	PSNR (dB)	Size (bytes)	Pkt	PSNR (dB)
HPL	21818	42	27.6	13042	24	23.14
LPL	19218	35	36.39	27994	53	36.39
Total	41036	77	36.39	41036	77	36.39

The video sequence *Foreman* in the CIF format is encoded using the H.264/AVC software JM version 16.2 [54]. First, GOP size of 16 frames is used with a frame structure IPPP...., with a fixed slice size of 600 bytes. Three schemes are compared: one is the proposed UEP scheme optimized using (6.1). The second scheme is the benchmark scheme, where all the slices of IDR and the first slice of each frame are placed in HPL, and all other slices in LPL. The third is the equal error protection (EEP) scheme that protects all slices equally. Note that the benchmark scheme is a low-complexity scheme where prioritization is done in an ad hoc manner; it still uses the same systematic EW RLC for protection of the two layers.

The proposed scheme is designed in accordance with the algorithm described in section 6.2 with $T = 0.78$ and $w = 2.5$. The sizes, number of packets (same as the number of slices in a layer) and resulting PSNR values for both configurations are shown in Table 6.1. For this selection of T , the proposed UEP scheme has larger HPL than the benchmark.

Note however, that a smaller HPL for the proposed scheme could be obtained by suitably selecting parameter T and w in (6.1).

The proposed schemes are simulated with transmission of EW RLC for 1000 runs and the results averaged. All schemes are compared at the same transmission bitrate. For an L-layer scheme the overhead cost needed to describe a UEP solution is: $7 \times L$ plus $(L-1) \times 8 \times 2$ to convey T_i 's and w_i 's. For $L=2$, used in this study this number is only 30bits and has not been taken into consideration.

The total number of packets to be transmitted for each run is 100. Because of the employed systematic RLC the transmission takes place in two phases. In Phase I, 77 packets are transmitted consisting of the source symbols. In Phase II, additional packets are transmitted in accordance with EW RLC. Note that Phase I will be the same for all the three schemes, whereas in Phase II, the probability of selection can govern a prioritized transmission of HPL. The important phenomenon seen here is that since each slice is independently decodable, the PSNR obtained in the case when RLC decoding of LPL fails and decoding of HPL succeeds, is higher than the PSNR of successfully decoded HPL due to useful packets that are received from LPL during Phase I. This gain comes from the correct reception of additional LPL symbols from Phase I even with failure of LPL decoding. The simulations have been performed for different packet loss rates (PLR) and different probabilities of window selection to evaluate the performance of the slicing feature to overcome losses.

In the case of non-systematic codes, if the first window W1 (or W2) does not get decoded, the entire GOP is considered to be lost. However, in case of systematic codes, it is still possible for the H.264/AVC decoder to decode the GOP as long as its IDR frame has been received correctly. In case of loss of IDR with systematic codes the entire GOP is lost. The PSNR for such cases is obtained by using the last frame of the previously decoded GOP to replace all frames of the lost GOP.

The various configurations are used to create different UEP schemes based on protecting the constituent windows with different protection, based on probabilistically selecting a window for each output symbol at the transmitter. An increase in the selection probability of window 1 (W1) will improve its robustness at a cost of a decrease in robustness of the succeeding layer(s). The EEP scheme is the case where only the largest window is selected with 100% probability. This means that all of the data is protected with no preference for the data considered important, i.e., window W1.

In Figure 6.3, the results are presented for comparison between the systematic codes and non-systematic codes. The scheme PS60S is a scheme with probability of W1 selection equal 0.6 (i.e., probability to select a symbol from HPL is 0.6), and the suffix S indicates systematic codes. Similarly scheme PS80N, has probability of W1 selection of 0.8 with non-systematic codes. It can be seen from the figure that at low loss rates the performance of systematic and non-systematic codes are very similar. EEP scheme has the lowest performance because it has to suffer more decoding failures trying to protect whole of the data as compared to UEP schemes which succeed more often in decoding HPL. The systematic codes generally have better results than the non-systematic codes for the error range and data rates shown. The main advantage of systematic RLC codes over non-systematic codes however comes from reduced decoding complexity. This can be very helpful for receiver devices with limited processing capability.

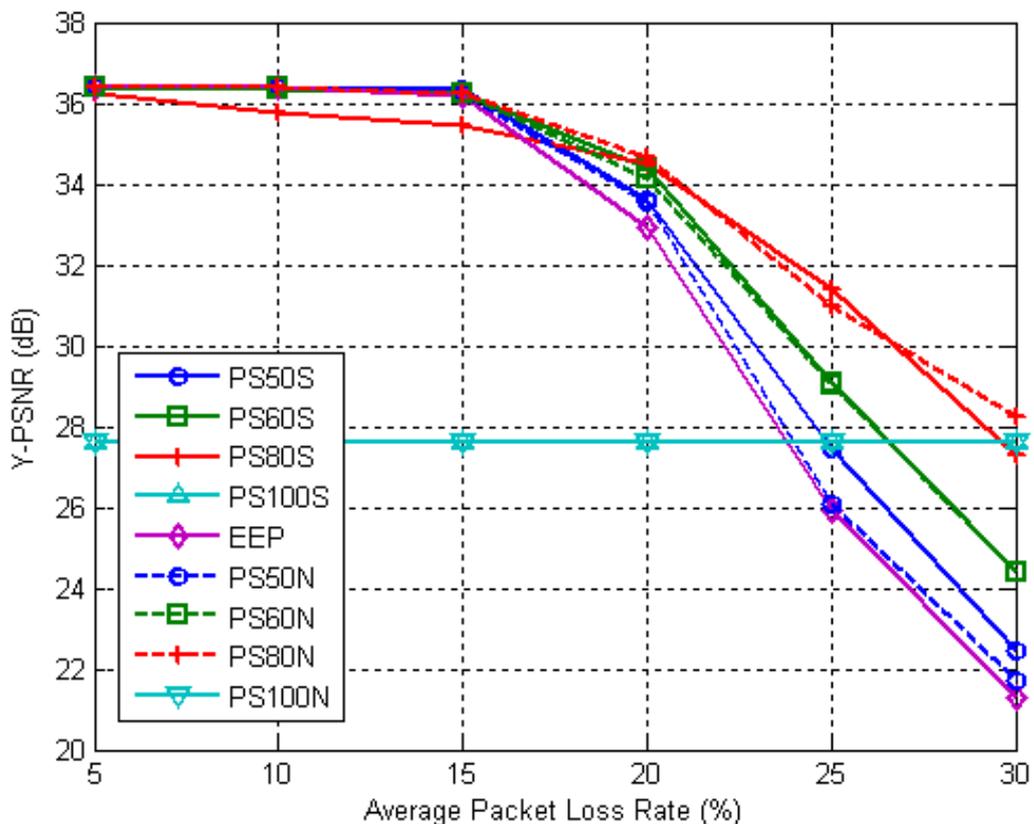


Figure 6.3: Comparison between systematic and non-systematic EW RLC codes.

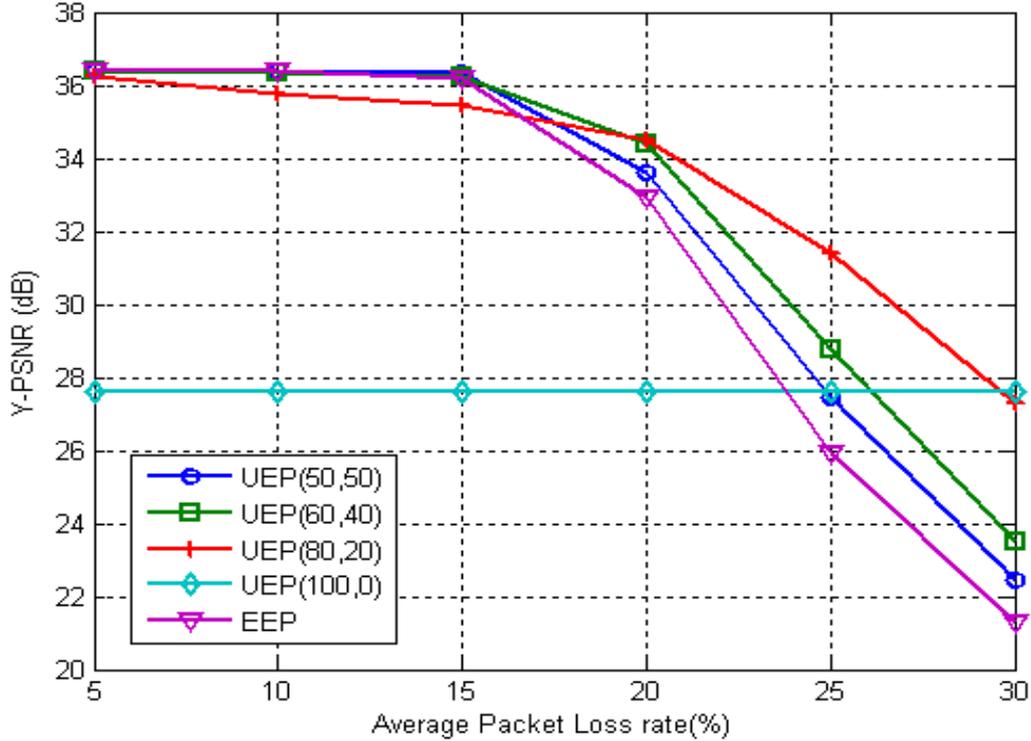


Figure 6.4: PSNR vs. average PLR for the proposed scheme and the EEP scheme.

6.3.2 Proposed Slice Prioritization Scheme

In this Section, the proposed systematic EW RLC scheme is used to compare the performance of different UEP and EEP schemes. Figure 6.4 shows PSNR vs. PLR for the proposed systematic EW RLC scheme. The numbers shown in brackets represent the selection probability of each of the two windows, e.g., UEP (60, 40) represents a code in which a symbol from W1 will be selected for transmission with a probability of 0.60. As can be seen from the figure, the results of UEP schemes are significantly better than the EEP schemes for high loss rates.

UEP (100, 0) is a scheme in which only W1 is protected and sent. The scheme is constrained in that it cannot achieve higher PSNR than 27.6 dB (see Table 6.1). However, the decoding failures, i.e., when the entire GOP data fails to be decoded, will be much less for UEP (100, 0), since HPL is protected strongest which facilitates each GOP to be received with high probability, though at basic quality level. This scheme could thus prove useful in higher PLR. Also, note that for this scheme, in Phase I of transmission, only the systematic

codes in the HPL will be transmitted and in Phase II the encoded symbols come from only HPL.

Table 6.2: Layer sizes and PSNR contributions for configurations with different values of T – Foreman sequence.

Layer	T1= 0.78			T2 = 0.58			T3 = 0.44		
	Size (bytes)	No of Pkts	PSNR (dB)	Size (bytes)	No of Pkts	PSNR (dB)	Size (bytes)	No of Pkts	PSNR (dB)
HPL	21818	42	27.6	23598	45	28.25	25366	48	29.55
LPL	19218	35	36.39	17438	32	36.39	15670	29	36.39
Total	41036	77	36.39	41036	77	36.39	41036	77	36.39

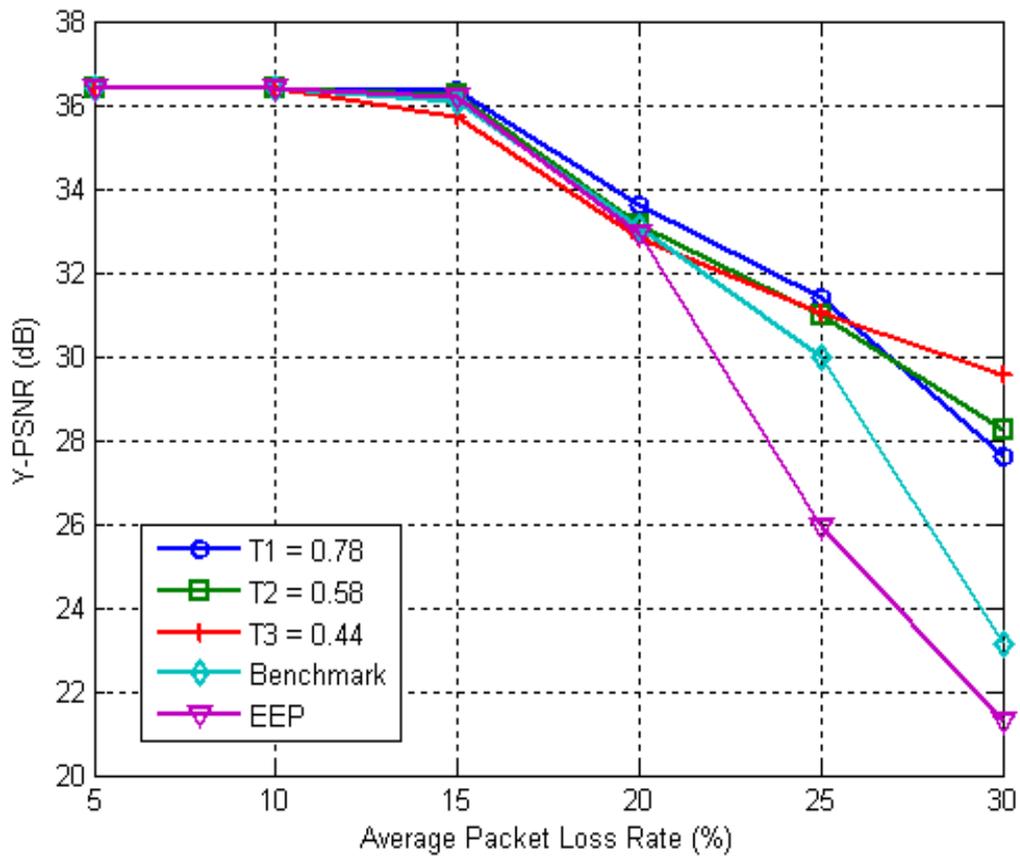


Figure 6.5: Optimized results for three different values of T-Foreman sequence.

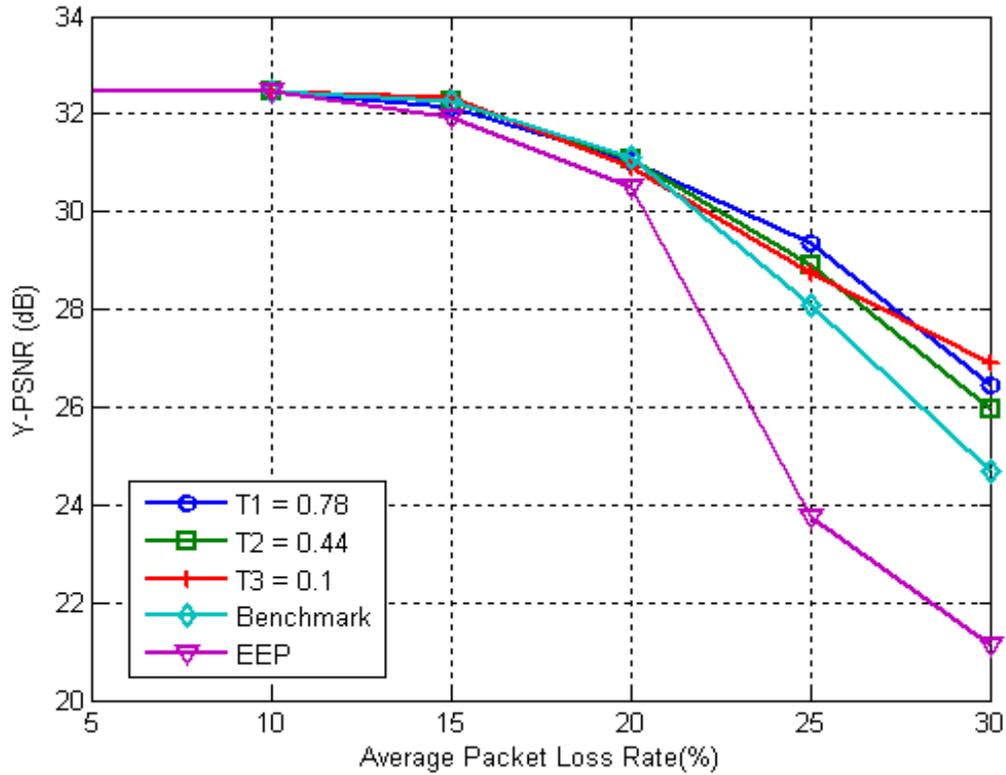


Figure 6.6: Optimized results for three different values of T-Paris sequence.

The PSNR results are improving with an increase in probability of selection of W1 because at higher probabilities of selection of W1, the decoding of HPL has high chance to be successful. As described earlier, the PSNR with decoding of HPL is enhanced by systematic LPL packets.

6.3.3 Optimized Results - Different T

In Table 6.2 the details of HPL size and PSNR contributions for the three schemes created with selecting three different values of T are shown. Intuitively, when the threshold T is lowered, the number of packets selected for HPL is higher. In Figure 6.5 the optimized results are presented for the schemes created in Table 6.2. The results for the EEP scheme and benchmark are also shown for comparison. For each PLR, the optimal proposed UEP and the optimal benchmark UEP are found using (6.2). It can be seen from the figure that the proposed method leads to significant gains for high PLRs compared to the EEP and the benchmark scheme. Lower T leading to a larger HPL, is better for higher PLRs, which is

expected since larger HPL (with higher PSNR) is better protected, and for LPL anyway there is not enough bandwidth. For lower PLR, larger HPL can be afforded.

Similar results obtained for Paris sequence are shown in Figure 6.6. Note that for high PLR, it is better to reduce T resulting in large HPL. In any case, varying T , one can effectively design HPL/LPL sizes for different PLR.

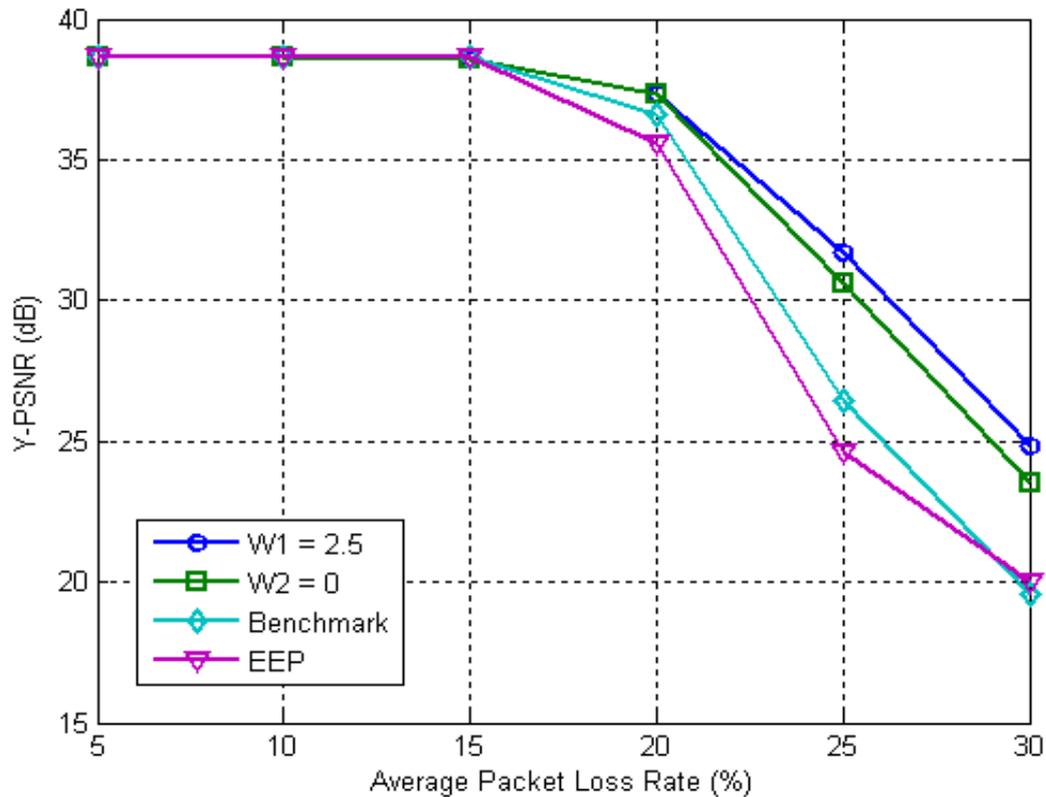


Figure 6.7: Optimized results for two different values of w – Foreman sequence.

6.3.4 Optimized Results – Different w

A larger GOP size may be required for applications such as DVB-H [91]. Foreman sequence is encoded with a GOP size of 64 frames. For this configuration, the total source packets are 161. The total number of sent packets is kept as 209 packets. In Figure 6.7, the optimized results are presented for the schemes created using two different values of w as shown. Both schemes have the value of $T = 3.1$, however, based on different value of w , different slices are selected for HPL of each scheme. The scheme $w_1 = 2.5$ has better

performance than $w_2 = 0$, especially at high packet loss, this comes from the fact that the former scheme prioritizes slices taking into account frame position in the sequence, which reduces error propagation. The benchmark scheme is created according to the selection criteria as used previously. EEP scheme performs the worst of all the schemes. The results for $w_1 = 2.5$ and $w_2 = 0$ are close at lower PLR. The reason for this is that with systematic codes, if the HPL is decodable then the packets received correctly (which could be from HPL or LPL) in Phase-I also contribute to improve the PSNR.

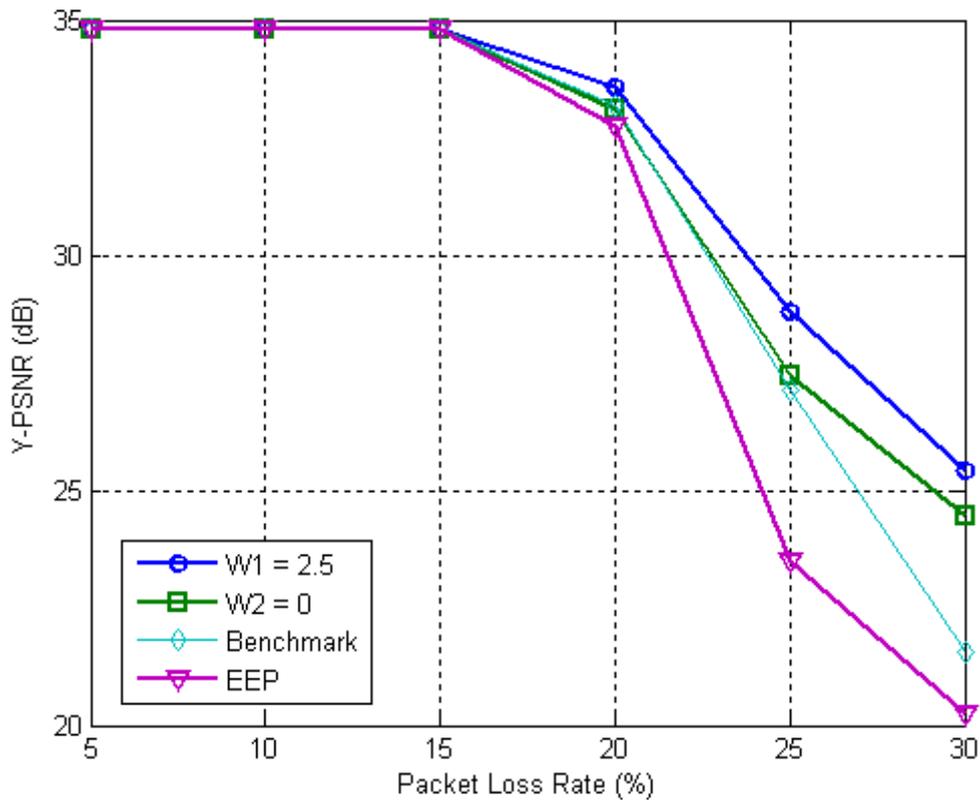


Figure 6.8: Optimized results for two different values of w – Paris sequence.

Paris sequence encoded with similar parameters is used to investigate the effect of w on performance. The optimized results are presented in Figure 6.8 for the schemes created using two different values of w along with Benchmark and EEP scheme. The results are similar to those in Figure 6.7 for Foreman sequence, which confirms the analysis carried out earlier.

The selection of w thus improves source packet allocation. Several different values of w , were tested and results for the typical cases that show achievable performance boundaries by varying w are shown. One can see from the figures that effect of w is small – up to 1db.



(a) Original sequence.



(b) Benchmark scheme.



(c) $w = 0$.



(d) $w = 2.5$.

Figure 6.9: Representative frames from Paris sequence.

The frame 15 for Paris sequence has been extracted for the HPL of three schemes (as shown in Figure 6.8) and are displayed in Figure 6.9. The benchmark scheme has a noticeable distortion on persons' faces and shoulders. The scheme $w = 0$ has distortion around the female person's hands. It can be seen that the picture with $w = 2.5$ is much more pleasing as compared to the other schemes. It is important to note here that with $w = 0$ and $w = 2.5$, the video data size is same, whereas the selected slices in each case have different temporal significance.

6.4. Summary

This chapter covers the systematic EW RLC scheme to protect the sliced-partitioned video data under various channel conditions at different probabilities of window selection. A novel slice prioritization method is proposed that takes into account PSNR contribution of a slice as well as position of its frame within GOP. The simulations for two layers case shows that UEP schemes perform better as compared to the EEP scheme and ad hoc prioritization, achievable with a minimal selection (one slice) of video data from each frame. Such reduced selections may be advantageously used for video-on-demand applications.

It is shown that incorporating the temporal position of a slice into account for slice prioritization gives better results as compared to slice prioritization based on distortion and error propagation effect alone. The proposed scheme is very flexible as by selecting a different threshold (for partitioning slices into HPL/LPL) an HPL of appropriate size can be obtained. Also, the slice size can be chosen to generate packets of appropriate size as required by different applications. Hence, the decoding complexity of RLC can be easily managed in the proposed scheme by an adaptive scheme which dynamically selects the slice size. The proposed schemes are thus suitable for real-time multimedia mobile applications.

Chapter 7

Conclusion and Future Work

The problem of supporting the multimedia communications over the emerging mobile wireless networks has been considered in this thesis. Multimedia communication over such wireless channels is difficult due to the fact that the channel conditions are generally poor and also the channel characteristics can change very rapidly. Most of the latest communication standards are IP based, whereas Internet provides only a best-effort service model and the priority-based service models are gradually being realized for real-time data. Considering the importance of the issues highlighted above, this thesis focuses on designing adaptive solutions to supporting multimedia traffic over wireless channels.

Application Layer FEC has been extensively used to provide reliable multimedia communication over wireless channels. Rateless codes can prove to be advantageous in such scenarios as potentially unlimited number of encoded symbols can be generated, from a given set of source symbols. Such codes can bring additional benefit if the unequal importance of different portions of encoded video data is also taken into account. The DP and slicing feature of H.264/AVC video have been extensively explored to design schemes for providing unequal protection to different video data elements. The focus of this work is on the recent rateless codes such as LT, RLC and Raptor codes to propose solutions for H.264/AVC video communication. Some proposed solutions are specifically targeted to DVB-H broadcasting and relay based video streaming applications such as LTE-A, whereas the remaining are generic solutions for unicast transmission scenarios.

In Chapter 3, the DVB system is the focus and novel solutions are proposed for AL-FEC for video broadcast. The use of RLC as an alternative to Raptor codes has been proposed. The priority based ordering of partitions using DP feature has been exploited to design schemes which enable use of UEP. The NOW and EW based RLC schemes have been extensively covered and analyzed. It is concluded that EW-RLC can be used as an AL FEC solution for video broadcast in DVB-H. Some results based on the emerging DVB-T2 standard have also been included. The proposed solution could be used in the emerging DVB standards, like DVB-NGH. The solutions have much wider applicability being designed

around H.264/AVC. The issue of decoding complexity for RLC codes is addressed by limiting the number of source symbols.

Chapter 4 proposes a novel scheme, called RASSA, which is an adaptive scheme for unicast video transmission. The scheme requires no changes to the LT and H.264/AVC decoder but still provides significant improvements over other comparative schemes. The limitation of “all or nothing” decoding of rateless codes is addressed by proposing a data structure, termed video-table which makes it possible for the video decoding at the receiver to succeed even if the LT decoding has failed. However, this requires that the layered video data is fine-grained so as to minimize the number of lost packets. The RASSA scheme is applied to video data partitioned with DP and slicing features. This is a low-cost solution which can easily be extended to other Fountain codes and can prove to be very beneficial for wireless communication with little or no feedback. The major application of the proposed solution could be in applications which need to adapt the video data to the varying channel conditions, as the data only needs to be encoded once and any rate adaptation is possible. The research in this direction could combine RASSA with a modified LT coding process to further increase the gains. Also, EW RLC codes which are very promising could be utilized with RASSA.

The emerging relay based networks such as LTE-A are addressed in Chapter 5. A comparison between H.264/AVC and H.264/SVC data is given for relay collaboration strategies, namely AF and DF. The possible applications and advantages of each scheme are highlighted. A novel solution for MDC is proposed based on DP and slicing features of H.264/AVC. The proposed solution is distinct because within each description created by slicing feature; provision of exploiting UEP is also available by utilizing the DP feature. Thus many possibilities emerge to selectively transmit the partitions of each description to exploit the multi path diversity. A resource optimization frame-work is provided and a fast algorithm is proposed that sub-optimally selects relays and schedules packets for transmission. The simulation results of the proposed solution for the emerging LTE-A standard are provided. The study is based on priority based layering of H.264/AVC video and it would be interesting to design similar schemes using H.264/SVC.

In Chapter 6, a novel solution for video streaming based on systematic codes is proposed. The work of Chapter 3 with slicing as used for DVB-H is extended and an algorithm is proposed for slice sorting based on PSNR contribution of a slice together with its

temporal significance. The systematic EW RLC codes are proposed due to their reduced encoding and decoding complexity. It is shown through simulations that the proposed algorithm results in significant gains over slice sorting schemes which do not exploit the temporal significance of each slice.

The general significance of the work contained in this thesis is that it provides solutions which do not require changes to the video and FEC encoders. The proposed solutions for video transmission are of low-complexity and can be deployed at the application layer through software. Some of the proposed schemes have immediate potential for commercial systems as described in the following paragraphs.

RASSA scheme arranges the encoded video into segments which can be truncated starting at the least significant partitions. The best part of designing such a scheme around DP is that it has no performance loss as DP just re-orders the encoded video data into prioritized partitions. The importance of having such a system is that similar to H.264/SVC, encoding the data once can serve different scenarios. The RASSA scheme is compatible with the H.264/AVC standard. In this manner, an application server can adaptively select the important video data to be transmitted. The possibility of sending a video table (a list of DP sizes) to the decoder opens a new possibility wherein the video decoding could succeed even when the channel decoding has failed to recover all symbols. Moreover, it enjoys a distinct advantage over H.264/SVC in that RASSA would tolerate some loss in the designed base layer whereas H.264/SVC will fail in such case. This work was limited to designing a data ordering and its associated data truncation system. This work could be extended by incorporating progressive transmission of the video segments with rateless codes.

The design of the proposed MDC scheme is based on a novel idea of splitting the encoded video based on slicing to create two independent descriptions, and then partitioning the data of each description with the DP feature. The scheme brings distinct advantages of both slicing and DP together. The scheme is designed in a manner that both descriptions support each other. A major benefit is that it is possible to obtain a low quality video by receiving the base layer of only one of the descriptions. The proposed MDC scheme is a generic solution which is suitable for any system with multiple paths. The coding of MDC scheme in this work uses EW RLC for error protection which makes it suitable for the emerging standards such as DVB-NGH and LTE-A. Potential use is in video streaming to a

single destination through multiple short range relays as envisaged in LTE-A. The limitation of the proposed scheme is that it is presently restricted to two descriptions only. This work can be extended by investigating techniques to make it possible to create more than two descriptions. This technique can have acceptance in the video communication industry as it not only makes the video transmission more error resilient but also that it is possible to transmit base layer of either or both descriptions to suit a particular transmission scenario.

The systematic codes as proposed with slicing provide the benefit of adapting the size of a slice to that of a network packet. As each slice is an independently decodable unit, therefore loss of a packet would not have adverse effect on other received packets. The advantage of systematic codes lies in simplified encoding/decoding which could be exploited for the delivery of time-critical data. Such schemes are also promising for mobile devices with limited processing power and where conservation of battery life is important. The significance of this work with slicing is that encoded video could be arranged into a base layer (and associated quality) of choice to suit a particular system. The system as proposed is limited to pre-encoded video only. The gain with slice ordering based on temporal significance can be increased further by employing other factors into prioritization decision. A limitation of the proposed system is that the video reconstruction fails with loss of any of the IDR slices.

The field of multimedia communication has seen tremendous growth. With the deployment of high capacity communication networks such as 4G and DVB-NGH it is likely that novel applications would follow. Hence there would always be a need for efficient solutions which address the challenges of multimedia communication over the emerging wireless networks. With the confluence of cellular and broadcast communication, RLC would assume added importance. It could be interesting future work to explore the advantages of RLC in such scenarios.

The manner in which DP and slicing features have been used to design novel solutions is extensible to other features as well. As part of future work the investigation of other error-resilience features beyond DP and slicing will be undertaken. This work did not consider any changes to the channel encoders. As part of future work the potential of designing video streaming solutions around such adaptation of channel encoders will also be considered.

Bibliography

- [1] “Video Streaming: Concepts, Algorithms, and Systems”, J.G. Apostolopoulos, W. Tan, and S.J. Wee, *Handbook of Video Databases*, CRC Press, March 2003.
- [2] T. Wiegand, G. J. Sullivan, G Bjøntegaard, and A. Luthra, “Overview of the H.264/AVC Video Coding Standard”, *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 13, no. 7, pp. 560-576, Jul. 2003.
- [3] T. Wiegand, H. Schwarz, A. Joch, F. Kossentini, and G. J. Sullivan, “Rate-constrained coder control and comparison of video coding standards,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, no. 7, pp. 688–703, Jul. 2003.
- [4] H. Schwarz, D. Marpe, and T. Wiegand, “Overview of the Scalable Video Coding Extension of the H.264/AVC Standard,” *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 17, no. 9, pp. 1103-1120, Sept 2007.
- [5] S. Wenger, “H.264/AVC over IP,” *IEEE Trans. Cir. Syst. Video Technol.*, vol. 13, pp. 645–656, July 2003.
- [6] M. van der Schaar, *Multimedia Over IP and Wireless Networks*, Academic Press, 2007.
- [7] M. Mitzenmacher, “Digital Fountains: A Survey and Look Forward,” *IEEE Information Theory Workshop*, pp. 271-276, 2004.
- [8] M. Luby, “LT codes”, *Proc. 43rd Annual IEEE Symposium on Foundations of Computer Science*, pp. 271-280, 2002.
- [9] A. Shokrollahi, “Raptor codes,” *IEEE Trans. Inform. Theory*, vol. 52, no. 6, pp. 2251-2567, 2006.
- [10] D. J. C. Mackay, “Fountain Codes,” *IEE Proceedings on Communications*, vol. 152, pp. 1062-1068, 2005.
- [11] D. Gomez-Barquero, D. Gozalvez, and N. Cardona, “Application layer FEC for mobile TV delivery in IP datacast over DVB-H systems,” *IEEE Trans. Broadcasting*, vol. 55, no. 6, pp. 396-406, 2009.
- [12] “ETSI TS 126 346, universal mobile telecommunications system (umts); multimedia broadcast/multicast service (mbms); protocols and codecs,” *ETSI Tech. Spec.*, 2005.
- [13] D. Lun, M. Medard, R. Koetter, and M. Effros, “On coding for reliable communication over packet networks,” *Physical Communication*, vol. 1, no. 1, pp. 22–30, 2008.
- [14] D. Vukobratović and V. Stanković, “Unequal Error Protection Random Linear Coding Strategies for Erasure Channels”, accepted for *IEEE Trans in Comm.*

- [15] Y. Lin, B. Li, and B. Liang, "Priority random linear codes in distributed storage systems," *IEEE Trans. Parallel Distributed Syst.*, vol. 20, no. 11, pp. 1653–1667, Nov. 2009.
- [16] K. Nguyen, T. Nguyen, and S. Cheung, "Video streaming with network coding," *Springer - J. Signal Process. Syst.*, vol. 59, no. 3, pp. 319–333, 2010.
- [17] M. Halloush and H. Radha, "Network coding with multigeneration mixing: Analysis and applications for video communication," *Proc. IEEE ICC 2008*, Beijing, China, May 2008.
- [18] X. Liu, G. Cheung, and C-N. Chuah, "Structured network coding and cooperative wireless ad-hoc peer-to-peer repair for WWAN video broadcast," *IEEE Trans. Multimedia*, vol. 11, no. 4, pp. 730–741, 2009.
- [19] N. Thomos, J. Chakareski and P. Frossard, "Randomized network coding for UEP video delivery in overlay networks," *Proc. IEEE ICME 2009*, New York, NY, USA, July 2009.
- [20] H. Shojania and B. Li, "Random network coding on the iphone: Fact or fiction?," in *NOSSDAV 2009, Williamsburg, USA*, June 2009.
- [21] D. Vukobratović, V. Stanković, D. Sejdinović, L. Stanković, and Z. Xiong, "Expanding window fountain codes for scalable video multicast," *IEEE Trans. Multimedia*, vol. 11, no. 6, pp. 1094–1104, 2009.
- [22] D. Vukobratović and V. Stanković, "Unequal error protection random linear coding for multimedia communications," in *Proc. MMSP-2010 IEEE Multimedia and Signal Processing Workshop*, Saint Malo, France, October 2010.
- [23] G. Faria, J. A. Henriksson, E. Stare, and P. Talmola, "DVB-H: Digital broadcast services to handheld devices," *Proc. of the IEEE*, vol. 94, no. 1, pp. 194-209, Jan. 2006.
- [24] DVB-CM-NGH015R1 Version 1.01, "Commercial Requirements for DVB-NGH", 15 June 2009.
- [25] L. Vangelista et al., "Key technologies for next-generation terrestrial digital television standard DVB-T2", *IEEE Commun. Magazine*, vol 47. no. 10, pp. 146-153, Oct. 2009.
- [26] Y. Yang, H. Hiu, J. Xu, and G. Mao, "Relay technologies for WIMAX and LTE-A mobile system," *IEEE Commun. Magazine*, Oct. 2009.
- [27] 3GPP TR 36.814 V1.2.1, "Further Advancements for EUTRA: Physical Layer Aspects," Tech. Spec. Group Radio Access Network Rel. 9, June 2009.
- [28] TM-NGHSM055r1 TM-H NGH Study Mission Report.

- [29] E. Dahlman, S. Parkvall, and J. Skold, “4G LTE/LTE-Advanced for mobile broadband” Academic Press, 2011.
- [30] V.K. Goyal, “Multiple description coding: Compression meets the network,” *IEEE Signal Processing Magazine*, vol. 18, pp. 74-93, 2001.
- [31] ETSI EN 302 755, “Digital Video Broadcasting (DVB); Frame structure channel coding and modulation for a second generation digital terrestrial television broadcasting system (DVB-T2)”, *V1.3.1, DRAFT*, Oct. 2011.
- [32] T. Stockhammer, A. Shokrollahi, M. Watson, M. Luby, and T. Gasiba, *Application Layer FEC for Mobile Multimedia Broadcasting (in Handbook of Mobile Broadcasting)*, CRC Press (Eds: B. Furht, and S. Ahson), 2008.
- [33] M. Luby, T. Gasiba, T. Stockhammer, and M. Watson, “Reliable multimedia download delivery in cellular broadcast networks,” *IEEE Trans. Broadcast.*, vol. 53, no. 1, pp. 235-246, Mar. 2007.
- [34] C. Hellge, D.Gomez-Barquero, T. Schierl, and T.Wiegand, “Layer-aware forward error correction for mobile broadcast of layered media,” *IEEE Trans. Multimedia*, vol. 13, pp. 551-562, Jun 2011.
- [35] “ETSI TR 102 591, digital video broadcast (DVB); IP datacast: Content delivery protocols (CDP) implementation guidelines part 1: IP datacast over DVB,” ETSI Tech. Spec., 2010.
- [36] J. Jin, B. Li, and T. Kong, “Is random network coding helpful in wimax?,” in *Proc. IEEE INFOCOM 2008*, Phoenix, AZ, April 2008.
- [37] R. Ahlswede, N. Cai, S. R. Li, and R. W. Yeung, “Network information flow,” *IEEE Trans. Inform. Theory*, vol. 46, no. 4, pp. 1204-1216, 2000.
- [38] Q. Zhang, F. Fitzek, and M. Katz, “On the energy saving potential in DVB-H networks exploiting cooperation among mobile devices,” in *Proc. Cognitive Wireless Networks*, 2007, pp. 473–484.
- [39] L. Benacem and S.D. Blostein, “Raptor-network coding enabled strategies for energy saving in DVB-H multimedia communications,” in *Proc. 1st Int. Conf. Green Circuits and Systems*, pp. 527–532, June 2010.
- [40] T. Schierl, T. Stockhammer, and T. Wiegand, “Mobile Video Transmission using Scalable Video Coding,” *IEEE Trans. Circuits and Sys. Video Techn.*, vol. 17, pp. 1204-1217, Sept. 2007.

- [41] A. T. Connie, P. Nasiopoulos, Y.P. Fallah, and V. C.M. Leung, "SCTP-based transmission of data-partitioned H.264 video," *In Proc. of the 4th ACM workshop on Wireless multimedia networking and performance modeling*, pp 32-36. ACM, 2008.
- [42] T. Stockhammer, and M. Bystrom, "H.264/AVC data partitioning for mobile video communication," *ICIP-2004 IEEE International Conference on Image Processing*, Brussels, Belgium, September 2004.
- [43] D. Gomez-Barquero, K. Nybom, D. Vukobratović, and V. Stanković, "Scalable video coding for mobile broadcasting DVB system," *in Proc. ICME-2010 IEEE Intl. Conf. Multimedia and Expo*, Singapore, July 2010.
- [44] W. Chung, S. Paluri, S. Kumar, S. Nagaraj, "Unequal Error Protection for H.264 Video using RCPC Codes and Hierarchical QAM," *in Proc. ICC-2010 IEEE Intl. Conf. on Comm.*, 2010.
- [45] B. Barmada, M.M. Ghandi, E.V. Jones, and M. Ghanbari, "Prioritized transmission of data partitioned H.264 video with hierarchical QAM," *IEEE Signal Processing Letters*, vol. 12. no. 8, pp. 577-580, Aug. 2005.
- [46] L. Al-Jobouri, M. Fleury, and M. Ghanbari, "Error-resilient IPTV for an IEEE 802.16e channel," *Wireless Engineering and Technology*, pp. 70-79, April 2011.
- [47] R. Razavi, M. Fleury, M. Altaf, H. Sammak, and M. Ghanbari, "H.264 video streaming with data-partitioning and Growth codes," *in Proc. ICIP-2009 Intl. Conf. Image Proc.*, Cairo, Egypt, Oct. 2009.
- [48] A. Bouabdallah and J. Lacan, "Dependency-Aware Unequal Erasure Protection Codes," *J. Zhejiang Univ. Science A*, vol. 7, pp. 27-33, Apr. 2006.
- [49] D. Sejdinović, D. Vukobratović, A. Doufexi, V. Šenk, and R. Piechocki, "Expanding Window Fountain codes for unequal error protection," *IEEE Trans. Commun.*, vol. 57, pp. 2510-2516, Sept. 2009.
- [50] D. Vukobratović and V. Stanković, "On unequal error protection random linear coding for scalable video broadcasting," *in Proc. Packet-Video Workshop-2010*, Hong Kong, Dec. 2010.
- [51] Y. Dhondt, S. Mys, K. Vermeirsch and R. Van de Walle, "Constrained Inter Prediction: Removing Dependencies between Different Data Partitions," *Proceedings of Advanced Concepts for Intelligent Visual Systems*, Delft, 28-31 August 2007, pp. 720-731.
- [52] K. Nybom, S. Gronroos, and J. Bjorkqvist, "Expanding window fountain coded scalable video in broadcasting," *in Proc. ICME-2010 IEEE Int. Conf. Multimedia and Expo*, Singapore, July 2010.

- [53] J. Poikonen and J. Paavola, "Error models for the transport stream packet channel in the DVB-H link layer", in *Proc. ICC-2006 IEEE Int. Conf. Commun.*, pp. 1861–1866, Istanbul, Turkey, 2006.
- [54] H.264/AVC Reference Software, <http://iphome.hhi.de>.
- [55] S. Nazir, D. Vukobratović, and V. Stanković, "Performance evaluation of raptor and random linear codes for H.264/AVC video transmission over DVB-H networks," in *Proc. ICASSP-2011 IEEE Intl. Conf. Acoustics Speech and Sig. Proc.*, Prague, Czech Republic, May 2011.
- [56] P. Vingelmann, F.H.P. Fitzek, M.V. Pedersen, J. Heide and H. Charaf, "Synchronized multimedia streaming on the iPhone platform with network coding," in *Proc. CCNC IEEE Consumer Commun. and Networking Conf.*, Las Vegas, USA, Jan. 2011.
- [57] H. Shojania and B. Li, "Parallelized progressive network coding with hardware acceleration," in *Proc. IWQoS-2007 the 15th Intl. Workshop Quality of Service*, Chicago, IL, July 2007.
- [58] S. Nazir, D. Vukobratović, and V. Stanković, "Unequal error protection for data partitioned H.264/AVC video streaming with Raptor and Random Linear Codes for DVB-H networks," in *Proc. ICME-2011 IEEE Intl. Conf. Multimedia and Expo*, Barcelona, Spain, July 2011.
- [59] E. Baccaglioni, T.Tillo, and G. Olmo, "Slice Sorting for Unequal Loss Protection of Video Streams," *IEEE Signal Proc. Letters*, vol. 15, pp. 581-584, June 2008.
- [60] A. Albanese, J. Blomer, J. Edmonds, M. Luby, M. Sudan, "Priority encoding transmission," *IEEE Trans. On Info. Theory*, vol. 42, pp1737-1744, Nov 1996.
- [61] R. Hamzaoui, V. Stanković, and Z. Xiong, "Optimized error protection of scalable image bitstreams," *IEEE Signal Proc. Magazine*, vol. 22, pp. 91-107, Nov. 2005.
- [62] S. Ahmad, R. Hamzaoui, M. Al-Akaidi, "Unequal error protection using LT codes and block duplication," *Proc. MESM 08, Middle Eastern Multiconference on Simulation and Modelling*, Amman, Aug. 2008.
- [63] A.G. Dimakis, J. Wang, and K. Ramchandran, "Unequal Growth codes: Intermediate performance and unequal error protection for video streaming", in *Proc. MMSP-2007 IEEE Multimedia Signal Proc.*, pp. 107-110, Oct. 2007.
- [64] L. Al-Jobouri, M. Fleury, and M. Ghanbari, "Adaptive rateless coding for data-partitioned video streaming over a broadband wireless channel", *6th Conference on Wireless Advanced (WiAD)*, pp. 1–6, June 2010.

- [65] N. Rahnavard, B. N. Vellambi, and F. Fekri, "Rateless codes with unequal error protection property", *IEEE Trans. Inf. Theory*, vol. 53, no. 53, pp. 1521–1532, Apr. 2007.
- [66] J. Goshi, A. E. Mohr, R.E. Ladner, E. A. Riskin, "Unequal Loss Protection for H.263 Compressed Video," in *Proc. DCC-2003 IEEE Data Compression Conference*, 2003.
- [67] L. Al-Jobouri, M. Fleury, and M. Ghanbari, "Channel Coding with Packet Duplication for Video Streaming over a Broadband Wireless Link," wireless advanced 2011.
- [68] S. Nazir, D. Vukobratović, and V. Stanković, "Scalable broadcasting of sliced H.264/AVC over DVB-H network," in *Proc. ICON-2011 IEEE Intl. Conference on Networks, Special Session on Robust and Scalable Multimedia Networking*, Singapore, Dec. 2011.
- [69] J. G. Apostolopoulos, "Reliable video communication over lossy packet networks using multiple state encoding and path diversity," in *Proc. SPIE Proc. Visual Commun. and Image Proc.*, vol. 4310, no. 1, 2001, pp. 392-409.
- [70] V. Stankovic, R. Hamzaoui, and Z. Xiong, "Robust layered multiple description coding of scalable media data for multicast," *IEEE Signal Processing Letters*, vol. 12, pp. 154-157, Feb. 2005.
- [71] "LTE-Advanced," 3GPP Global Initiative, <http://www.3gpp.org/LTE-Advanced>, 2010.
- [72] S. Parkvall, A. Furuskar, and E. Dahlman, "Evolution of LTE towards IMT-advanced," *IEEE Commun. Magazine*, pp. 84-91, Feb. 2011.
- [73] H. Luo, S. Ci, D. Wu, J. Wu, and H. Tang, "Quality-driven cross-layer optimized video delivery over LTE," *IEEE Commun. Magazine*, pp. 102-109, Feb. 2010.
- [74] S. Mao, S. Lin, S. Panwar, Y. Wang, and E. Celebi, "Video transport over ad hoc networks: Multistream coding with multipath transport," *IEEE J. Sel. Areas in Commun.*, vol. 21, no. 10, pp. 1721-1737, 2003.
- [75] A. Agrawal, "Heterogeneous networks. A new paradigm for increasing cellular capacity," http://netseminar.stanford.edu/seminars/01_29_09.pdf, Jan. 2009.
- [76] A.F. Molisch, N.B. Mehtra, J.S. Yedidia, J. Zhang, "Cooperative relay networks using fountain codes," in *Proc. IEEE GLOBECOM 2006*, pp. 1-6, 2006.
- [77] T. Tillo, M. Grangetto, and G. Olmo, "Redundant slice optimal allocation for H.264 multiple description coding," *IEEE Trans. Circuits and Sys. for Video Techn.*, vol. 18, no. 1, pp. 59-70, Jan. 2008.

- [78] Y. Liao and J.D. Gibson, "Routing-aware multiple description coding with multipath transport for video delivered over mobile ad-hoc networks," in *Proc. IEEE GLOBECOM-2010*, Miami, Dec. 2010.
- [79] L. Zhou, B. Geller, B. Zheng, A. Wei, and J. Cui, "System scheduling for multi-description video streaming over wireless multi-hop networks," *IEEE Trans. Broadcasting*, vol. 55, no. 4, pp. 731-741, Dec. 2009.
- [80] HD Video LTE-A Berlin Testbed:
<http://www.hhi.fraunhofer.de/en/departments/wireless-communication-and-networks/video/cooperative-systems-hd/>
- [81] Q. Zhang and S. Kassam, "Finite-State Markov model for Rayleigh fading channels," *IEEE Trans. Commun.*, vol. 47, no. 11, pp. 1688-1692, Nov. 1999.
- [82] K.T. Phan, L.-N. Tho, S.A. Vorobyov, and C. Tellambura, "Power allocation in wireless multi-user relay networks," *IEEE Trans. Wireless Commun.*, vol. 8, pp. 2535-2545, May 2009.
- [83] K. Bakanoglu, S. Tomasin, and E. Erkip, "Resource allocation in wireless network with multiple relays," in *Proc. Asilomar 2008*, Pacific Grove, CA, 2008.
- [84] E. Calvo, J. Vidal, and J.R. Fonollosa, "Optimal resource allocation in relay-assisted cellular networks with partial CSI," *IEEE Trans. Signal Proc.*, vol. 57, pp. 2809-2823, July 2009.
- [85] D. Jurca and P. Frossard, "Video packet selection and scheduling for multipath streaming," *IEEE Trans. Multimedia*, vol. 9, pp. 629-641, April 2007.
- [86] J.H. Sorensen, J. Astergaard, P. Popovski, and J. Chakareski, "Multiple description coding with feedback based network compression," in *Proc. IEEE GLOBECOM-2010*, Miami, FL, Dec. 2010.
- [87] Y. Liu, "A Low Complexity Protocol for Relay Channels Employing Rateless Codes and Acknowledgement," *IEEE Intl Symposium on Info Theory*, pp. 1244-1248, 2006.
- [88] J. Li, Z. Hu, Y. Wang, "DF/AF cooperative relay in LTE-A," *IEEE Vehicular Tech. Conf.*, pp. 1-5, 2010.
- [89] N. Thomos, S. Argyropoulos, N. V. Boulgouris, and M. G. Strintzis, "Robust Transmission of H.264/AVC Video Using Adaptive Slice Grouping and Unequal Error Protection," *EURASIP Journal on Applied Signal Processing*, vol 2006, pp 1-13, Feb 2006.
- [90] T. Tillo, E. Baccaglini, and G. Olmo, "Unequal protection of video data according to slice relevance," *IEEE Trans. Signal Proc.* vol. 20, pp. 1572-1582, June 2011.

- [91] S. Nazir, D. Vukobratović, and V. Stanković, “Expanding window random linear codes for data partitioned H.264 video transmission over DVB-H network,” in *Proc. ICIP-2011 IEEE Intl. Conf. Image Proc.*, Brussels, Belgium, Sept. 2011.