

# ENHANCEMENT OF WYNER-ZIV VIDEO CODING FOR WIRELESS CHANNELS

by

Kuganeswaran Thambu

B.Sc. Eng., University of Peradeniya, 2001 Nov.

M. Eng., Asian Institute of Technology, 2007 May.

A Dissertation

presented to Ryerson University

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in the

Program of Electrical and Computer Engineering

Toronto, Ontario, Canada, 2013

©Kuganeswaran Thambu 2013

I hereby declare that I am the sole author of this dissertation. This is a true copy of the dissertation, including any required final revisions, as accepted by my examiners.

I authorize Ryerson University to lend this dissertation to other institutions or individuals for the purpose of scholarly research.

---

I further authorize Ryerson University to reproduce this dissertation by photocopying or by other means, in total or in part, at the request of other institutions or individuals for the purpose of scholarly research.

I understand that my dissertation may be made electronically available to the public.

---

# Enhancement of Wyner-Ziv Video Coding for Wireless Channels

Doctor of Philosophy 2013

Kuganeswaran Thambu

Electrical and Computer Engineering

Ryerson University

## Abstract

Wyner-Ziv video coding (WZVC) is a fast emerging video coding technique for wireless video sensor networks. WZVC moves the complexity from the encoder (sensor) to the decoder (receiver). This thesis proposes few enhancements to solve challenging problems in WZVC, namely 1) handling impairments of a *fading wireless* channel in WZVC environment, 2) investigating rate penalty of WZVC in wireless fading channels, 3) adaptive encoder rate control using inter-frame cross-correlation properties, and 4) better side information estimation with a bit-based noise variance computation technique.

In case (1), the decoder metric values in WZVC are calculated with respect to the correlation noise, channel noise, and *fading*. Multiple input multiple output (MIMO) diversity scheme is studied with WZVC for improving the reconstructed video output. Simulation results show that the average peak signal to noise ratio (PSNR) of Foreman video is improved by  $\approx 3$  dB with the configuration of 2I4O compared to single input single output (SISO) channel at  $SNR = 2$  dB.

In case (2), expressions for the rate penalty are analytically derived under different wireless channel conditions. WZVC scheme with receiver diversity (WZVC-RD) is proposed

and also demonstrated that it alleviates the channel-induced rate penalty. Simulation results show that with adequate diversity, the channel induced rate penalty can be almost completely eliminated, i.e., the average rate penalty is reduced to be less than 20 Kbps in WZVC-RD for Foreman video at  $SNR = 2$  dB.

In case (3), theoretical rate-distortion behavior of conditional decoding is compared with that of the practical WZVC. An encoder rate control algorithm is proposed with respect to the cross-correlation threshold (CCTH) at the decoder, where the definition of the threshold is derived based on cross-correlation statistics of the key frames. Additionally, an adaptive threshold algorithm (ATHA) is proposed to improve the PSNR versus rate relationship.

In case (4), correlation noise variance estimation is proposed with respect to the bit pattern of each pixel; it is named as “bit-based noise variance”. PSNR improvement of  $\approx 4$  dB is observed for the Foreman video frames with higher correlation noise.

## Acknowledgement

This dissertation would not be possible without the help and support of many people. A few words mentioned here cannot adequately express my appreciation.

I would like to express my sincere gratitude to Professor Xavier Fernando and Professor Ling Guan for giving me this opportunity. I am extraordinarily grateful to their criticism, patience, invaluable support and guidance throughout my research program at the Ryerson University. Their helpful, detailed criticism has sharpened my scientific and communication abilities, and their professional discipline and diligence inspire me to the same. I am also very grateful for their positive encouragement at numerous times during my graduate studies.

I would like to extend my gratitude to Professors S. Krishnan , L. Zhao, J. Misic, and M. Kyan from the Ryerson University, and Professor Ha H. Nguyen from University of Saskatchewan for serving in my doctoral dissertation exam committee. Their insightful advice and comments have improved the quality of this dissertation. My sincere thanks to Professor Y. He for giving valuable comments on Chapter 5. Special thanks to my masters program advisor Professor N. Rajatheva for general discussion on wireless communication.

I gratefully acknowledge Ryerson University for the continuous support for my studies. I would like to acknowledge the support of my lab mates, who have helped me during each stage of my graduate studies. Finally, I am deeply grateful to my family, for their patient encouragement throughout my every adventure, including graduate school. Their unconditional love reminds me that my value as a person does not hinge upon my completion of this degree or upon professional accolades. Additionally, this would not have been possible without the love, laughter, and prayers of many friends.

# Contents

Abstract . . . . .	iii
Acknowledgement . . . . .	v
Table of Contents . . . . .	vi
List of Tables . . . . .	ix
List of Figures . . . . .	x
<b>1 Introduction</b>	<b>1</b>
1.1 Wyner-Ziv Video Coding . . . . .	4
1.1.1 Wyner-Ziv Video Coding in Practice . . . . .	5
1.1.2 Wyner- Ziv Video Coding and Correlation Noise . . . . .	6
1.1.3 Encoder Rate Control in WZVC . . . . .	7
1.2 Thesis Contribution . . . . .	8
1.2.1 WZVC in Wireless Channel Environment . . . . .	9
1.2.2 Rate Penalty Due to the Wireless Channel and Proposed Improvement	10
1.2.3 Encoder Rate Control in WZVC with Cross-Correlation Threshold (CCTH) . . . . .	10
1.2.4 Correlation Noise Estimation in WZVC . . . . .	11
1.3 Thesis Organization . . . . .	12
<b>2 Basics of Wyner-Ziv Video Coding</b>	<b>14</b>
2.1 Slepian-Wolf Theorem . . . . .	16
2.1.1 Practical Slepian-Wolf Model . . . . .	17
2.2 Wyner-Ziv Coding . . . . .	17
2.3 Turbo Based WZVC . . . . .	18
2.3.1 Quantization . . . . .	20
2.3.2 Turbo Encoder-Decoder . . . . .	20
2.4 Side Information Generation . . . . .	22

2.5	Reconstruction . . . . .	24
2.6	Video Quality Assessment . . . . .	25
2.7	Feedback Channel and Rate Control . . . . .	26
2.8	Correlation Noise Between $X$ and $Y$ . . . . .	28
2.8.1	Encoder Correlation Noise Models . . . . .	30
2.8.2	Decoder Correlation Noise Models . . . . .	32
2.9	Transmission Channel . . . . .	33
2.9.1	Additive White Gaussian Noise Channel (AWGN) . . . . .	34
2.9.2	Fading Channel . . . . .	35
2.10	Multipath Fading . . . . .	37
2.10.1	Wireless Channel Model . . . . .	38
<b>3</b>	<b>Wyner-Ziv Video Coding with Wireless Channels</b>	<b>43</b>
3.1	Introduction . . . . .	43
3.2	Theoretical Motivations and Contributions . . . . .	45
3.3	Theory . . . . .	48
3.3.1	Bit-plane Extractor . . . . .	48
3.3.2	Turbo Coder . . . . .	49
3.3.3	Wireless Channel . . . . .	52
3.3.4	MIMO Wireless Channel . . . . .	53
3.4	WZVC over a Single Wireless Channel . . . . .	56
3.5	WZVC over MIMO Wireless Channel . . . . .	59
3.6	Simulations . . . . .	61
3.7	Simulation Results . . . . .	64
3.7.1	Reconstructed WZ Frame of the WZVC with Different Channel Con- ditions . . . . .	64
3.7.2	Improvement of WZVC with Wireless Channel Environment . . . . .	69
3.8	Summary . . . . .	72
<b>4</b>	<b>Rate Penalty in Wireless Wyner-Ziv Video Coding</b>	<b>75</b>
4.1	Introduction . . . . .	75
4.2	Average Rate of WZ Coding in Wireless Channel . . . . .	77
4.2.1	WZC in Noiseless Channel . . . . .	77
4.2.2	WZC in Noisy Fading Channel . . . . .	78
4.2.2.1	WZC in AWGN Channel . . . . .	78

4.2.2.2	WZC in Noisy Fading Channel . . . . .	79
4.3	Proposed WZVC with Receiver Diversity (WZVC-RD) . . . . .	82
4.4	Rate at WZVC-RD . . . . .	84
4.5	Results . . . . .	85
4.6	Summary . . . . .	88
<b>5</b>	<b>Cross-Correlation based Rate Reduction Technique for Wyner-Ziv Video Coding</b>	<b>90</b>
5.1	Introduction . . . . .	90
5.2	Conditional Rate Distortion . . . . .	93
5.2.1	Relationship Between $\rho_{XY}$ and $R_{X/Y}$ . . . . .	97
5.3	Practical WZVC . . . . .	98
5.3.1	Observation of Cross-Correlations . . . . .	99
5.3.2	Observations of Rate-Distortion . . . . .	101
5.4	Cross-correlation Thresholds and Rate Control Algorithm . . . . .	105
5.5	Cross-correlation Thresholds (CCTHs) . . . . .	106
5.5.1	Mean as The Threshold . . . . .	106
5.5.2	Positive and Negative Thresholds . . . . .	107
5.5.3	Upper Limit of the Thresholds . . . . .	108
5.5.4	Adaptive Threshold (ATH) . . . . .	109
5.6	Rate Control Algorithm . . . . .	110
5.6.1	Rate Control Based on Mean (E) . . . . .	112
5.6.2	Rate Control Based on PTH and NTH . . . . .	112
5.6.3	Rate Control Based on Adaptive Thresholds APTH and ANTH . . . . .	113
5.7	Performance Analysis . . . . .	114
5.8	Summary . . . . .	121
<b>6</b>	<b>Bit Based Correlation Noise Estimation at the Decoder</b>	<b>124</b>
6.1	Introduction . . . . .	124
6.2	Correlation Noise Variance Estimation . . . . .	126
6.3	Results and Discussion . . . . .	128
6.4	Summary . . . . .	130
<b>7</b>	<b>Conclusions and Future Work</b>	<b>136</b>
7.1	Conclusions . . . . .	136

7.2 Future Work . . . . .	138
Bibliography . . . . .	140

# List of Tables

5.1	Cross-correlation Coefficients and Rate Regions with Threshold E . . . . .	112
5.2	Cross-correlation Coefficients and Rate Regions with Thresholds: PTH, NTH and E . . . . .	112
5.3	Cross-correlation Coefficients and Rate Regions with Thresholds: APTH, ANTH, PTH and E . . . . .	113
5.4	Observation of Rate Variation and PSNR Variation for ANTH-PTH-APTH Compared to 50% Compressed WZVC . . . . .	117

# List of Figures

1.1	Video Streaming Consists of Uplink and Downlink . . . . .	2
1.2	Video Sensors with Uplink . . . . .	3
1.3	Thesis Contribution . . . . .	13
2.1	Slepian-Wolf Model . . . . .	16
2.2	Practical Slepian-Wolf Model . . . . .	17
2.3	WZVC of Stanford Model . . . . .	18
2.4	Turbo Encoder . . . . .	21
2.5	AWGN Channel Model . . . . .	34
2.6	Fading Channel Model . . . . .	39
2.7	Frequency Selective Fading Channel Model . . . . .	40
3.1	Video Transmission over a Single Wireless Channel . . . . .	46
3.2	Video Transmission over Wireless Channel with Transmit-Receive Diversity .	47
3.3	Block Diagram of Turbo Decoder . . . . .	49
3.4	Antenna Configuration with Two Transmit Four Receive Antennas . . . . .	54
3.5	WZVC over a Single Wireless Channel . . . . .	56
3.6	WZVC over Multiple Input Multiple Output (MIMO) Wireless Channel . . .	59
3.7	Simulation Model . . . . .	62
3.8	Reconstructed Foreman Video Frames with Lossless, AWGN, and Fading Channels at $SNR = 2$ dB . . . . .	65
3.9	Reconstructed Carphone Video Frames with Lossless, AGWN, and Fading Channels at $SNR = 2$ dB . . . . .	66
3.10	Reconstructed Akiyo Video Frames with Lossless, AWGN, and Fading Chan- nels at $SNR = 2$ dB . . . . .	67
3.11	PSNR Performance for AWGN Channel with Different Channel SNRs for Foreman Video Sequence . . . . .	68

3.12	Some Selected Reconstructed Wyner-Ziv Frames from Fig. 3.11 . . . . .	69
3.13	PSNR Performance for AWGN Channel with Different channel SNRs for Car- phone Video Sequence . . . . .	70
3.14	Comparison of Average PSNR Values for AWGN Channel with the Different Channel SNRs . . . . .	71
3.15	PSNR Performance for Foreman and Carphone in Wireless Fading Channels with Different SNRs at Rayleigh Variance 0.5 . . . . .	72
3.16	PSNR Performance of the Reconstructed WZ Frames over 2IMO Wireless Channel for Foreman Video at Rayleigh Variance 0.5 . . . . .	72
3.17	PSNR Performance of the Reconstructed WZ Frames over 2IMO Wireless Channel for Foreman Video at Rayleigh Variance 0.5 . . . . .	73
4.1	Standard WZ Codec with Noiseless Channel . . . . .	77
4.2	Standard WZ Codec with AWGN Noise and Wireless Channel Fading . . . .	78
4.3	WZVC with an Improved Decoder . . . . .	82
4.4	Average PSNR Variation with Rate and SNR for Foreman Video . . . . .	86
4.5	Average PSNR Variation with Rate for Foreman Video . . . . .	87
4.6	Average PSNR Variation with Rate for Foreman Video . . . . .	87
5.1	Wyner-Ziv Codec . . . . .	94
5.2	Theoretical Relationship of Rate-Distortion with Different Cross-correlation Values . . . . .	98
5.3	A Block diagram of WZVC . . . . .	99
5.4	Cross-Correlations of Ten Different Video Sequences . . . . .	100
5.5	Coast Guard Video Sequence . . . . .	102
5.6	Suzie Video Sequence . . . . .	102
5.7	Mother and Daughter video Sequence . . . . .	102
5.8	Average Rate-Distortion for Coast Guard, Suzie, and Mother & Daughter Video Sequences . . . . .	104
5.9	Theoretical Rate-Distortion Comparison with Coast Guard, Suzie, and Mother & Daughter Video Sequences . . . . .	105
5.10	Cross-correlations of Foreman Sequences with Mean . . . . .	106
5.11	Cross-correlations of Foreman Sequences with Mean, PTH1 and NTH1 . . .	107
5.12	Cross-correlations of Foreman Sequences with Mean, PTH2, PTH1, NTH1, and NTH2 . . . . .	108

5.13	Block Diagram of Rate Control Process . . . . .	110
5.14	Flow Chart for Deciding Feedback Signal . . . . .	111
5.15	State Diagram for Deciding the Encoder Rate . . . . .	114
5.16	Rate and PSNR Comparison of Four CCTHs: Mean, NTH1PTH1, NTH2PTH2, and ANTH-PTH-APTH . . . . .	115
5.17	Rate and PSNR Comparison of CCTHs: NTH2PTH1 and NTH2PTH2 . . .	116
5.18	PSNR Variations vs Reconstructed WZ frames for Coast Guard Video . . . .	119
5.19	PSNR Variations vs Reconstructed WZ frames for Foreman Video . . . . .	120
5.20	PSNR Variations vs Reconstructed WZ frames for Mother & Daughter Video	121
5.21	Best Rate-Distortion (RD) Points with Theoretical Rate-Distortion Curves .	122
5.22	Best RD Points with Practical and Theoretical Rate-Distortion curves . . . .	123
6.1	Bit-planes of Two Key Frames . . . . .	131
6.2	Bit-plane of Side Information . . . . .	132
6.3	(a) Input Wyner-Ziv Frame , (b) Reconstructed Frame of Foreman . . . . .	132
6.4	(a) Input Wyner-Ziv Frame , (b) Reconstructed Frame of Carphone . . . . .	133
6.5	(a) Previous Frame of Wyner-Ziv Frame (Key Frame 1) , (b) Next Frame of Wyner-Ziv Frame (Key Frame 2) . . . . .	133
6.6	(a) Previous Frame of Wyner-Ziv Frame (Key Frame 1), (b) Next Frame of Wyner-Ziv Frame (Key Frame 2) . . . . .	133
6.7	Probability Distribution of Correlation Noise for Foreman Video Sequence .	134
6.8	Probability Distribution of Correlation Noise for Carphone Video Sequence .	134
6.9	PSNR Performance of Foreman Video Sequence with Dynamic Bit-based Vari- ance Estimation Compared to Fixed Variance Methods . . . . .	135

# List of Abbreviations

Abbreviation	Description
ACF	Autocorrelation function
AWGN	Additive white Gaussian noise
ATH	Adaptive threshold
ATHA	Adaptive threshold algorithm
BPSK	Binary phase shift keying
BER	Bit error rate
CATV	Cable television
CCTH	Cross-correlation threshold
DSC	Distributed source coding
DVC	Distributed video coding
DCT	Discrete cosine transform
DISCUSS	Distributed source coding using syndromes
ECC	Error control coding
FEC	Forward error correction
HDTV	High definition television
IEC	International electrotechnical commission
ISO	International organization for standardization
ITU	International telecommunication union
LLR	Log-likelihood ratio
LDPC	Low-density parity check coding
LSB	Least significant bit
MAP	Maximum a posteriori
MC	Motion compensation
MIMO	Multiple input multiple output

MPEG	Moving picture experts group
MSB	Most significant bit
NTH	Negative threshold
PDF	Probability density function
PTH	Positive threshold
PSNR	Peak signal to noise ratio
PSD	Power spectral density
QCIF	Quarter common intermediate format
RCPT	Rate compatible punctured Turbo code
SI	Side information
SISO	Single input single output
SIMO	Single input multiple output
SNR	Signal to noise ratio
SLEP	Systematic lossy error protection
STC	Space time code
VCEG	Video coding experts group
VQEG	Video Quality Expert Group
WMSE	Weighted mean squared error
WZ	Wyner-Ziv
WZVC	Wyner-Ziv video coding
WZVC-RD	WZVC with receiver diversity

---

# Chapter 1

## Introduction

These days, a large proportion of the world's population are forced to use image, audio, and video coding technologies in their day to day lives. Users in developed and developing countries are very familiar with some of the most popular devices such as; mobile phones, tablets, digital cameras, digital television and MPx players, where audio-video compression is a key technology.

The purpose of audio-video compression is to reduce the amount of the transmitting bits for a certain target decoded signal quality. The current video coding paradigm is mostly based on the well known techniques such as motion compensated temporal prediction between video frames, transform coding, quantization of the transform coefficients (It is essential to the human visual system limitations), and entropy coding. The current video coding considers both temporal and spatial domains, which is also known as predictive, or hybrid (time and frequency) coding.

Predictive coding has been used as the solution for the most of the available video coding standards; e.g. ITU-T H.26x and ISO/IEC MPEGx. In this video coding architecture, the correlation between the video frames (temporal coding) and within the video frames (spatial coding) is exploited at the encoder; this leads to rather complex encoders with simpler decoders. The aforementioned predictive coding adequately fits some application scenarios such as broadcasting and video storage.

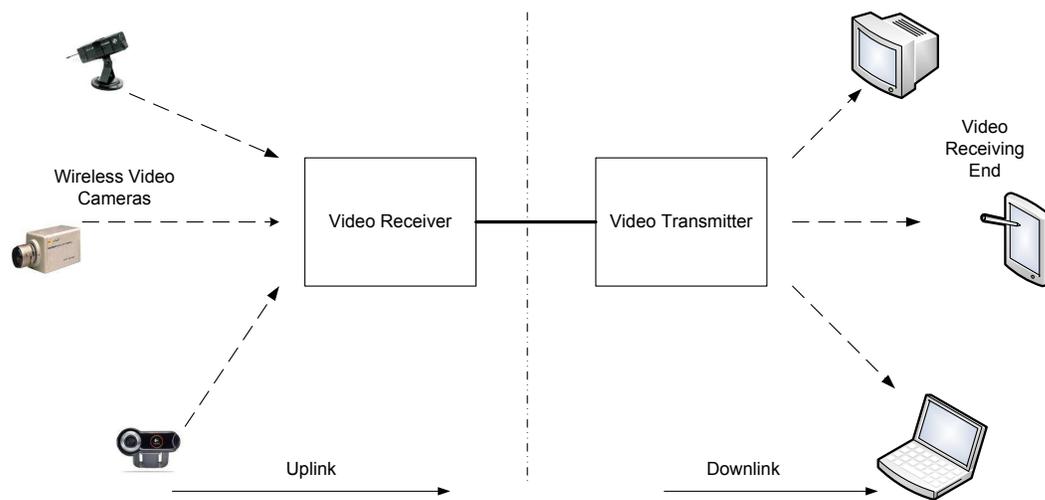


Figure 1.1: Video Streaming Consists of Uplink and Downlink

Fig. 1.1 shows the video streaming from portable video cameras to fixed and portable receiving end. It is divided as an uplink portion and downlink portion. In general, the downlink portion is composed of encoders, transmits to millions of decoders; here the con-

ventional predictive video coding is used as the most suitable technique. One of the most efficient video coding standard, H.264/ AVC, is applied to mobile video telephony, internet streaming to high definition television (HDTV), and Blu-ray discs [1]. The uplink portion in Fig. 1.1 shows the streaming from the video sensors, mobile devices, and internet to the base station. There are numerous applications prefer to use up-link model only, e.g., multiple number of sensors transmit data to a centralized receiver.

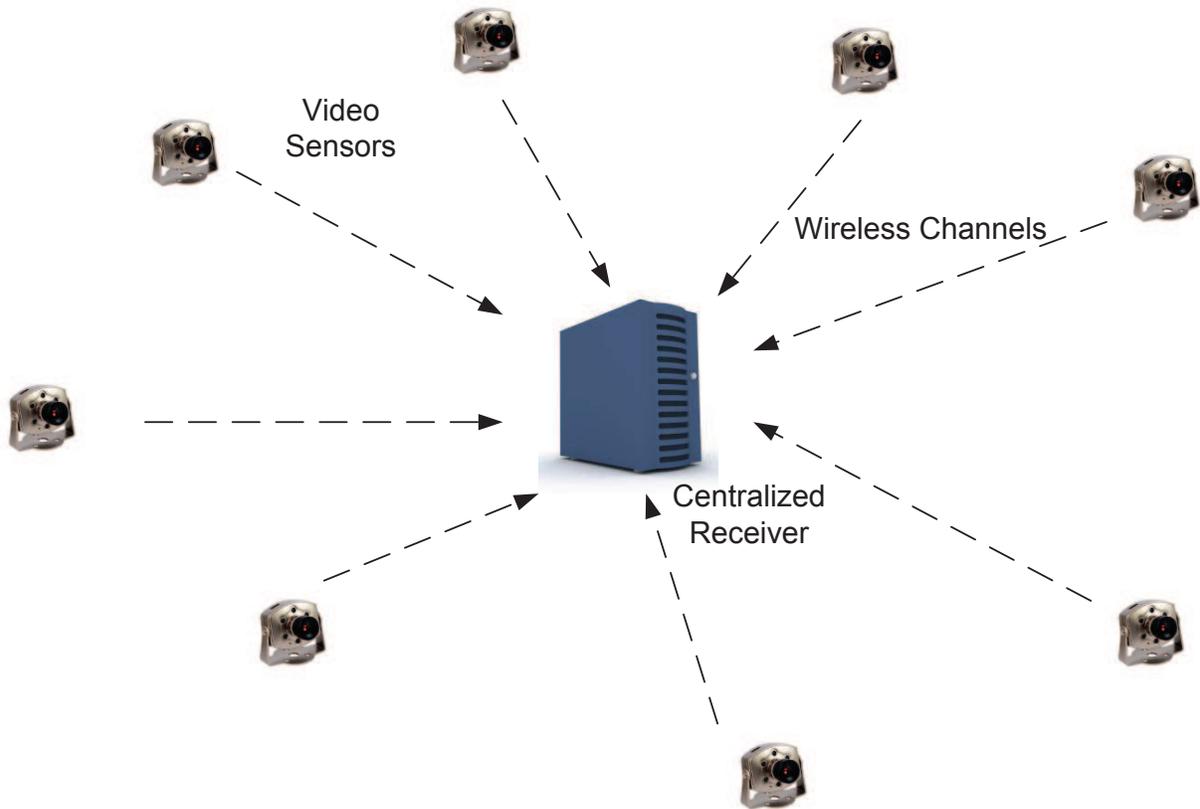


Figure 1.2: Video Sensors with Uplink

Fig. 1.2 shows an explicit diagram of the uplink portion from Fig. 1.1. Here, video sensors

transmit the video information to the centralized receiver/server. This type of video sensor networks require light encoding, robustness to channel losses, high compression efficiency, and low latency, e.g., wireless digital cameras, low power surveillance, and video sensor networks. Requirements of these emerging applications stimulated the development of the so-called distributed video coding (DVC)/Wyner-Ziv video coding (WZVC) paradigm.

## 1.1 Wyner-Ziv Video Coding

The WZVC is developed from the Slepian-Wolf [2] and the Wyner-Ziv [3] theorems. The Slepian-Wolf theorem refers to lossless compression, while the Wyner-Ziv theorem refers to lossy compression with side information available at the decoder. The Slepian-Wolf theorem states that two statistically dependent signals,  $X$  and  $Y$ , can be separately encoded and jointly decoded with an arbitrarily small error probability, if the achievable rate region satisfies the following conditions [2];

$$R_X \geq H(X/Y) \tag{1.1}$$

$$R_Y \geq H(Y/X) \tag{1.2}$$

$$R_X + R_Y \geq H(X, Y) \tag{1.3}$$

where,  $R_X$  and  $R_Y$  are achievable rates for  $X$  and  $Y$  respectively, and  $H(X|Y)$  and  $H(Y|X)$  are conditional entropies.

The Wyner-Ziv theorem states that by assuming  $X$  and  $Y$  to be statistically dependent Gaussian random processes, and  $Y$  as the side information for encoding  $X$ , the rate-mean squared error distortion function for  $X$  is the same, whether the side information  $Y$  is available only at the decoder, or both at the encoder and the decoder [3]. In WZVC, if  $X$  and  $Y$  are correlated video frames,  $Y$  can be considered as a corrupted or noisy version of  $X$ , and can be modeled as;

$$Y = X + N \tag{1.4}$$

where,  $N$  is defined as the correlation noise between  $X$  and  $Y$  with a Laplacian probability distribution. The practical designs of WZVC started around 2002 based on Slepian-Wolf and Wyner-Ziv theorems, where error correction codes (channel codes) are incorporated such as Turbo codes and low density parity check (LDPC) codes. The literature of WZVC is summarized in Chapter 2.

### 1.1.1 Wyner-Ziv Video Coding in Practice

WZVC is based on two pioneer research works; research team at Stanford University proposed WZVC with Turbo coder as channel coder [4], [5], and research team at University of California, Berkley, who proposed WZVC with sending syndromes [6], [7]. Both of these researchers considered hard-value of parities / syndromes from the encoder to decoder. Later

on, a few researchers have taken into account the impact of channel impairment. S. Rane. et al considered an error prone channel where the systematic lossy error protection (SLEP) is compared with forward error correction (FEC) [8]. SLEP is a joint algorithm of the hybrid video coding and WZVC [6]. The PRISM2007 [9] discussed the channel noise impact on the side information and outlined requirements for increasing predictors at the decoder side. However, many research works have not discussed wireless channel fading impairment in WZVC. Recently, a few research papers related to WZVC focus in hybrid distributed video coding and its applications [10], [11] and joint source coding and routing in distributed video coding [12].

### 1.1.2 Wyner- Ziv Video Coding and Correlation Noise

The side information  $Y$  is illustrated in (1.4), where  $X$  is the input frame and  $N$  is the correlation noise [4]. Estimation of  $N$  due to the residual difference between side information and input video frame is the exciting part in DVC/WZVC research. In previous works, researchers used a Laplacian distribution as a noise distribution  $f(Y - X) = \frac{1}{2}\beta \exp(-\beta|Y - X|)$ , where variance is known as  $\frac{2}{\beta^2}$ .  $\beta$  has been computed over the whole video sequence off line at the encoder before the WZVC procedure starts, and kept constant for the decoding of all Wyner-Ziv frames [4], [5], [13], [14].

Note that the value of  $\beta$  is essential to convert the side information into soft-value in-

formation which improves the performance of Turbo decoding. The soft-value of the side information can be computed with respect to the bit-metric value or log likelihood ratio (LLR). The noise variance or  $\beta$  decides LLR value; therefore, the noise variance  $\sigma^2$  needs to be estimated for each bit/symbol of side information.

Perira et.al. proposed the correlation noise estimation between WZ frame and side information using the statistical relationship between key frames [6], [15]. The weighted mean squared error (WMSE) between the the two key frames is given by,  $WMSE(x; y) = (\frac{1}{2})^2 \frac{\sum_{(x,y) \in SI} (X_B(x,y) - X_F(x,y))^2}{L}$ , where L denotes frame size , and  $X_B$  and  $X_F$  are key frames.  $WMSE(x, y)$  is considered as variance of the correlation noise [15], [16] and it varies from pixel to pixel; this correlation noise determines the soft-value of the side information. The decoder considers the side information as the systematic information. However, the pixel based variance will not be perfect when we consider the iterative decoding based on the systematic received bits. Therefore it is required to compute the variance with respect to the bit values.

### 1.1.3 Encoder Rate Control in WZVC

In WZVC, error control codes are applied to generate parity bits. Compression is achieved by sending the fraction of the parity bits (punctured from the original parity) to the decoder [17], [6]. The decoder uses the received parity to correct errors in side information for

reconstructing WZ frame; here, the side information is referred in (1.4). The most challenging task is to control the transmission of parity bits at the encoder.

Feedback channel is employed in WZVC to control the number of parity bits requested by the decoder [4], [18]. The encoder has a buffer to store the parity bits and the feedback channel requests the additional bits for successful decoding with reduced distortion, where the decoder uses a error-probability threshold to determine whether the available parity information is enough for decoding. If the decoder finds that the error probability is above the threshold, it requests additional parity bits from the buffer through the feedback channel; this process is repeated until the error probability falls below the threshold. The major drawback of this arrangement is latency, which arises due to the continuous error probability calculation and the process of reporting to the encoder. Some WZVC research works propose the rate allocation algorithms from the encoder side which eliminates the requirement of implementing the feedback channel [19], [20], [21]. However, the encoder complexity is increased considerably in the above setup, which deviates from the concept of complexity balance in WZVC.

## 1.2 Thesis Contribution

From the discussion in the previous section, it is clear that there are more challenges to be addressed on the decoder side of a WZVC than on the encoder side in video transmission

over wireless channels. Therefore, in this thesis, we studied some of the challenges on the decoder side and proposed solutions that effectively improve the performance of WZVC in a wireless environment.

### 1.2.1 WZVC in Wireless Channel Environment

Most of the previous work on WZVC assumed the ideal transmitting channel. Recently, a few researchers have started considering channel impairments. The robustness of the transmission scheme and their effects on the transmission errors is investigated by Puri et al. in [9], where the side information is influenced by the channel noise. We first investigate the reconstructed WZ video quality in the *wireless fading* channel. Since the wireless channel randomly fluctuates, utilization of more than one uncorrelated wireless links simultaneously improve the overall performance [22], [23], [24]. Hence, with MIMO-WZVC, we have incorporated the transmit-receive diversity technique/Multiple-Input-Multiple-Output (MIMO) with WZVC. On the otherhand, in Turbo decoding, systematic and parity information are fed as soft-values. Therefore, we have defined two parameters, *channel-metric value* and *bit-metric value* for parity and side information, respectively. The channel-metric value depends on the channel noise and fading gains of the wireless channel, and the bit-metric value is determined with respect to the correlation noise between WZ frame and side information.

### 1.2.2 Rate Penalty Due to the Wireless Channel and Proposed Improvement

In previous research [9], it was shown that the transmission rate of WZVC depends on both the correlation noise and the channel noise. However, in wireless channels, in addition to the channel noise, channel fading also influences the transmission rate. The combined effect of channel noise and fading on residual information ( $Z$ ) is very detrimental, which requires a higher rate  $Z$  to achieve the same quality. In other words, a noisy fading channel can effectively reduce the transmission rate; this is termed rate penalty.

For the first time we investigate the required transmission rate of WZVC in noisy fading wireless channels. We analytically show that the AWGN (additive white Gaussian noise) channel has a higher rate penalty than the noiseless channel, and a fading channel has even higher rate penalty than the AWGN channel. Thus, a WZVC with receiver diversity (WZVC-RD) setup is proposed in order to reduce the rate penalty in WZVC over the wireless fading channel.

### 1.2.3 Encoder Rate Control in WZVC with Cross-Correlation Threshold (CCTH)

We introduce an encoder rate control mechanism with respect to the cross correlation threshold (CCTH) of the available key frames at the decoder. We investigate the behavior of the

cross-correlation statistics of the available frames (key frames) at the decoder to analyze the rate distortion behavior. The reconstructed function with respect to the side information  $Y$  and residual information  $Z$  is analyzed, and  $Z$  is derived in terms of cross-correlation of  $X$  and  $Y$ , denoted by  $\rho_{XY}$ . In addition, the theoretical conditional rate is also derived with respect to  $\rho_{XY}$ . Furthermore, we justify the behavior of rate-distortion with respect to the cross correlations of the key frames which are closer to  $\rho_{XY}$ .

From the above fact, we propose an encoder rate control algorithm with respect to the cross-correlations of the key frames. We need to define thresholds with respect to the statistics of cross-correlation points; this determines the required number of parity bits from the encoder. Thus, we have also proposed an adaptive threshold algorithm (ATHA), which classifies number of the higher/lower threshold points.

#### 1.2.4 Correlation Noise Estimation in WZVC

Soft-value of side information is considered as the received systematic value at the decoder of WZVC. Log-likelihood ratio (LLR) has to be calculated with respect to each bit at the decoder; here LLR value depends on *correlation noise* for each bit/symbol of side information, and thus LLR supports the computation of the soft-value of the side information. *Correlation noise* is computed with respect to frame based and pixel based estimation. Instead of the static Laplacian model, as previously used, we propose a dynamic online estimation

technique of correlation noise with respect to bit pattern of each pixel, denoted as *bit-based variance* estimation. The bit-based variance of correlation noise is important due to the fact that an iterative algorithm at the Turbo decoder depends on the received bits per symbol.

### 1.3 Thesis Organization

The rest of the thesis is organized as follows; Chapter 2 briefly introduces the basics of WZVC theories and explains the wireless communication principles. Chapter 3 proposes WZVC with wireless channel impairments, where WZVC over a single wireless channel is considered initially. Thereafter, WZVC is incorporated with multiple input multiple output (MIMO) channel. Chapter 4 investigates the rate penalty in WZVC with wireless channel and the improvement of rate penalty with WZVC-RD. In Chapter 5, the encoder rate control of WZVC is determined with respect to the cross-correlation coefficients of the key frames at the decoder. Chapter 6 presents the correlation noise estimation at the decoder, where the soft-value of side information is determined by the bit-based noise variance. The contribution of this dissertation is summarized in Chapter 7, which also summarizes the future research directions relevant to this thesis. A snap shot of the thesis contribution work is shown in Fig. 1.3.

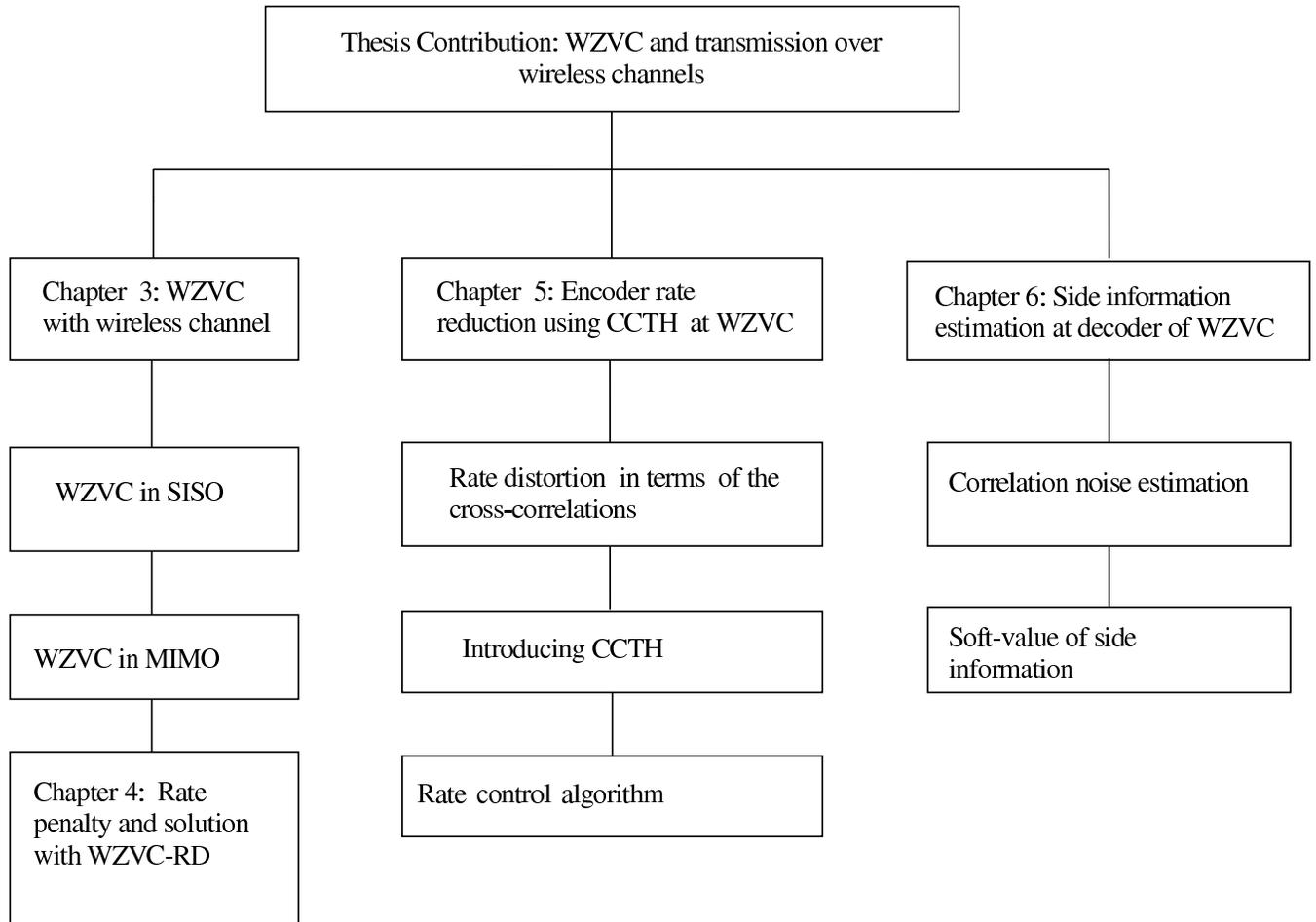


Figure 1.3: Thesis Contribution

## Chapter 2

# Basics of Wyner-Ziv Video Coding

Traditionally, video transmission focuses on broadcasting services where one complex high power transmitter serves multiple low cost receivers that originate from cable television (CATV) systems [25]. For example, the digital video coding solutions that are available in the market rely on hybrid block-based motion compensation/ discrete cosine transform (MC/DCT) architecture. Both the ITU-T VCEG and ISO/IEC MPEG standards follow this approach for applications where the video content is encoded once and decoded multiple times, such as in broadcasting or video streaming [26], [27]. In such applications, the video codec architecture is primarily a one-to-many model, where a single complex encoder and multiple light decoders are utilized.

The complexity burden of the encoder is mainly associated with the motion estimation and compensation tasks. However, some emerging systems such as wireless video surveillance [28], multimedia sensor networks [29], wireless PC cameras, and mobile camera phones

require a different arrangement. These systems are required to be implemented using low-complexity encoders at the expense of a high complexity decoder.

The WZVC is proposed to achieve low-complexity encoding with an asymmetric video compression, where the individual frames are encoded independently (intraframe encoding), but decoded conditionally (interframe decoding). This means that video statistics are used (partially or fully) in the decoder only, which result in low encoding complexity. For instance, if multiple cameras sense partially overlapped geographical areas in a wireless video surveillance scenarios, WZVC explores the correlation between the multiple video sequences at the decoder. Therefore it is possible to achieve a low encoder complexity, reducing the total cost of the system (in terms of hardware), since there will be multiple encoders and a single decoder.

The theoretical foundations of WZVC have been developed in 1970s from two main information theories described Slepian-Wolf and Wyner-Ziv theorems. The Slepian-Wolf theorem refers to lossless compression, while the Wyner-Ziv theorem refers to lossy compression with side information available at the decoder.

## 2.1 Slepian-Wolf Theorem

Slepian-Wolf theorem states that two statistically correlated discrete random sequences  $X$  and  $Y$  are encoded independently and then jointly decoded. Here,  $X$  and  $Y$  are independently and identically distributed (i.i.d) random variables.

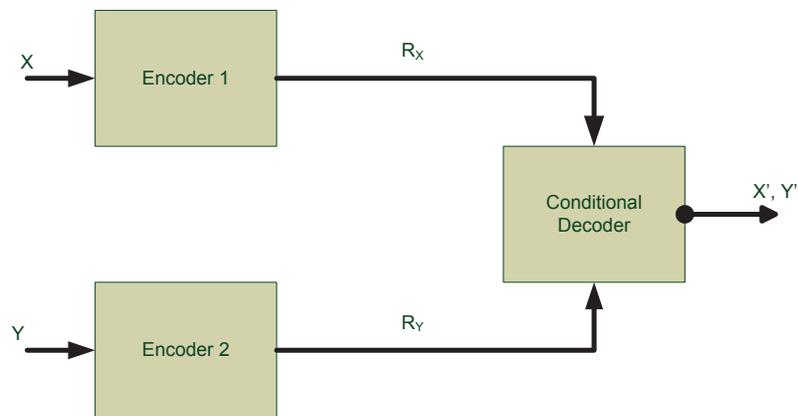


Figure 2.1: Slepian-Wolf Model

Fig. 2.1 shows the Slepian-Wolf model; here encoders 1 and 2 perform independent encoding, and joint decoding at the decoder. The Slepian-Wolf theorem states that the minimum rate for joint encoding-decoding is the same as the minimum rate for independent encoding- joint decoding, with an arbitrarily small error probability [2], such as  $R_X \geq H(X/Y)$ ,  $R_Y \geq H(Y/X)$  and  $R_X + R_Y \geq H(X, Y)$ .

### 2.1.1 Practical Slepian-Wolf Model

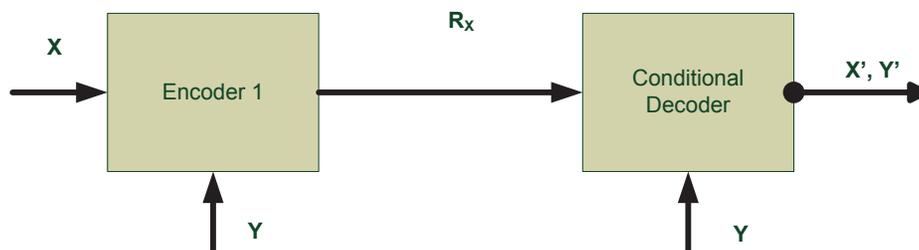


Figure 2.2: Practical Slepian-Wolf Model

The Slepian-Wolf coding is referred to in literature as lossless distributed source coding (DSC); two statistically dependent sequences are perfectly reconstructed at a joint decoder with a small and negligible decoding error probability. Fig. 2.2 shows the practical Slepian-Wolf codec, where  $X$  and  $Y$  are correlated sources; here,  $Y$  is the noisy version of the original uncorrupted signal  $X$ . The errors between  $X$  and  $Y$  can be corrected with a help of channel coding technique. [4].

## 2.2 Wyner-Ziv Coding

In 1976, Wyner and Ziv studied the case which deals with lossy compression of source  $X$  with the help of source  $Y$  at the decoder, where  $Y$  is known as side information. Wyner and Ziv proved that rate-distortion is preserved when the side information is not available at the encoder. This concept is also known as asymmetric coding.

Early solutions of practical WZVC were developed in 2002, especially the efforts from Stanford University [4], [17], [5], [30] and University of California, Berkeley [7], [6] with respect to the advance channel coding techniques, i.e., error control coding (ECC). The Stanford model is mainly categorized by frame based with respect to the Turbo codes, where a feedback channel is used to perform rate control. On the other hand, the Berkeley model is characterized by block based channel coding.

## 2.3 Turbo Based WZVC

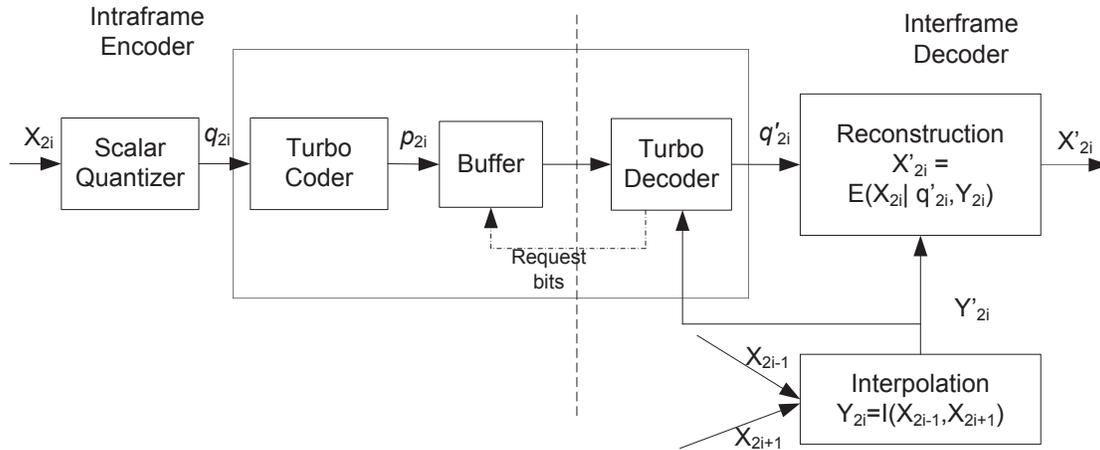


Figure 2.3: WZVC of Stanford Model

In this model, Wyner-Ziv coding is applied with a video signal.  $X$  denotes the even frames and  $Y$  is obtained from the odd frames of the video sequence.  $X$  is compressed by an intraframe encoder that is unable to access and utilize  $Y$ . The compressed stream is sent to

a decoder which uses  $Y$  as side information to conditionally decode  $X$ . This model does not use a correlation between  $X$  and  $Y$ , but the temporal similarities between adjacent video frames are considered [4], [17].

The basic structure is composed of an inner Turbo code-based Slepian-Wolf codec and an outer quantization-reconstruction pair. Both the Slepian-Wolf decoder and the reconstruction block make use of the side information available at the decoder. In Fig. 2.3, let  $X_1, X_2, \dots, X_N$  be the frames of a video sequence. The odd frames,  $X_{2i+1}$  and  $X_{2i-1}$ , are the key frames which are available at the decoder, where  $i = 0, 1, \frac{N}{2} - 1$ , and  $N$  is even number. For simplicity, compression of the key frames are not considered and is assumed to be known perfectly at the decoder. Each even frame,  $X_{2i}$ , is encoded independently without the knowledge of the key frames and the other even frames.

$X_{2i}$  is scanned row by row and each pixel value is quantized using  $2^M$  levels to obtain symbol stream  $q_{2i}$ . The symbols are fed into a Turbo encoder and the parity bit sequence  $p_{2i}$  is produced and stored in a buffer at the encoder side. The buffer transmits a subset of the parity bits to the decoder, upon request from the feedback channel.

For each frame  $X_{2i}$ , the decoder takes the adjacent key frames  $X_{2i-1}$  and  $X_{2i+1}$  and

performs temporal interpolation;  $Y_{2i} = I(X_{2i-1}, X_{2i+1})$ . The Turbo decoder uses the side information  $Y_{2i}$  and the received subset of  $p_{2i}$  to form the decoded symbol stream  $q'_{2i}$ . If the decoder is unable to reliably decode the symbols, it requests additional parity bits from the encoder buffer. The request and decode process is repeated until an acceptable probability of symbol error is guaranteed. For instance, the decoder needs to request  $k(\leq M)$  bits for decoding the pixel belongs to  $2^M$  bins; here the compression is achieved with respect to the amount of  $k$ . After the receiver decodes  $q'_{2i}$ , it calculates a reconstruction of the frame  $X'_{2i}$ , where  $X'_{2i} = E(X_{2i}|q'_{2i}, Y_{2i})$  [4].

### 2.3.1 Quantization

A uniform scalar quantizer is used with  $2^M$  levels to quantize the pixels of  $X_{2i}$ . Each quantizer bin is assigned a unique symbol. For a given frame, quantized symbols form a block of length  $L$ , which is then fed into Turbo encoder. In the above model, the task of grouping the code words into cosets is left to the Turbo coder, which operates in a space of much higher dimensionality.

### 2.3.2 Turbo Encoder-Decoder

The Turbo encoder-decoder system is implemented for the Slepian-Wolf coding of the quantized symbol stream  $q_{2i}$ . The Turbo encoder assigns a specific sequence of parity bits to a

given input block of symbols. Rate compatible punctured Turbo code (RCPT) structure is borrowed from channel coding [31], [32]. The rate flexibility of the RCPT enables it to adapt to the changing statistics between the side information and the frame to be encoded. This ensures that the encoder only sends the minimum number of parity bits required for the receiver to correctly decode  $q_{2i}$ . Feedback channel is used to support the rate compatibility, where the decoder requests additional parity bits until it can correctly decode the sequence. A simple encoder allows the decoder to control the bit allocation through the feedback channel. This way the encoder does not need to perform any statistical estimation that is necessary for proper rate control.

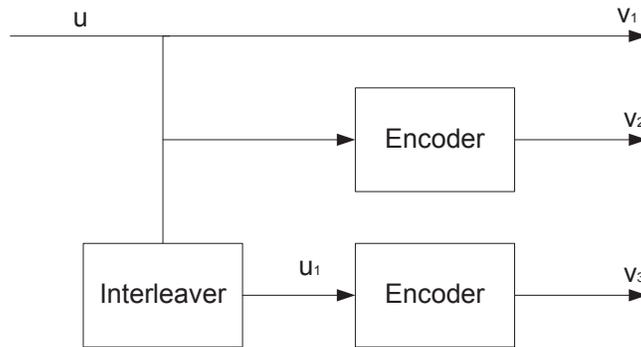


Figure 2.4: Turbo Encoder

The term Turbo code usually refers to the error correcting codes proposed by Berrou et al. in [33]. Fig. 2.4 shows the general structure of the Turbo encoder. In a standard

Turbo encoder, two convolutional coders (CC) operate on the same input bit, but there is an interleaver placed between these CC [34], [35], [36]. In general, Recursive Systematic Convolutional (RSC) codes are used as CC. Furthermore, output of the CC has one systematic bit and two parity bits [37]. By puncturing the output, Turbo encoder gives half rate, giving one parity bit and one systematic bit for every input bit; thus the overall coding rate of one-half. Generally only the parity bits are punctured, and not the systematic bits. However, performance of the code degrades if the systematic bits are punctured at the encoder.

Another key element that delivers substantial performance gain is the use of *soft* information rather than hard decisions, i.e., real numbers whose values are measures of confidence in the associated data symbol estimates. Soft-input decoding alone is known to deliver significant performance improvements [38], and exchanging soft information during the iterative decoding process yields larger gains.

## 2.4 Side Information Generation

Two adjacent key frames are taken, and temporal interpolation is performed, to generate the side information  $Y_{2i}$ , which is an estimate of  $X_{2i}$ . The first interpolation technique was performed simply averaging the pixel values at the same location from the two key frames. Let  $X_{2i-1,j}$  and  $X_{2i+1,j}$  be the pixel values at location  $j$  from the key frames, then

$$Y_{2i,j} = \frac{1}{2}(X_{2i-1,j} + X_{2i+1,j}).$$

Side information is also generated with more sophisticated techniques based on motion compensated (MC) interpolation at the decoder. One such system is block-based motion compensated interpolation with respect to the symmetric motion vectors (SMV) interpolation. Here, the motion vector of a given block in  $X_{2i}$  from time  $2i - 1$  to  $2i$  matches the motion vector from time  $2i$  to  $2i + 1$ . Block matching finds the best symmetric motion vector for a given block and then takes the average of the motion compensated blocks from the two adjacent frames.

The complexity of motion compensation at the decoder-side is increased but not at the encoder-side. However, the encoder of conventional video compression systems sends the motion information, which increases its rate. Improving the interpolation reduces the decoder's bit requirement. Thus, it is possible to improve compression performance by solely improving the decoder.

The decoder needs a model for the statistical dependency between  $X_{2i}$  and  $Y_{2i}$ ; it is required by the Turbo decoder and reconstruction block in order to determine the conditional probability and conditional expectation respectively. It has been observed that residual value between the current frame and the corresponding side information is very close to that of

a Laplacian random variable. Let  $X_{2i,j}$  and  $Y_{2i,j}$  be the given frame and the corresponding side information, respectively, then the distribution of the residual can be approximated by  $f(n) = \frac{1}{2} \exp(-\beta|n|)$ , where  $n = X_{2i,j} - Y_{2i,j}$ . Here,  $\beta$  can be estimated at the decoder using the residual between the key frame  $X_{2i-1}$  and  $Y_{2i-1}$ . It has been observed that an estimated  $\beta$  show better performance than a value with the closest fit.

## 2.5 Reconstruction

The reconstruction for each pixel with respect to the decoded symbols and the side information is given as  $X'_{2i,j} = E(X_{2i,j}|q'_{2i,j}, Y_{2i,j})$ . From the Stanford work [4], pixels are reconstructed independently regardless of the adjacent pixels so that spatial correlation is not exploited.

The reconstruction function is taken as a Laplacian model with  $\beta = 0.5$  and four quantization levels, where the  $X'_{2i,j}$  obtains the value of  $Y_{2i,j}$ , if  $Y_{2i,j}$  falls within the reconstructed bin  $q'_{2i,j}$ . However, if  $Y_{2i,j}$  is outside the bin, the function clips the reconstruction towards the boundary of the bin closest to  $Y_{2i,j}$ . This limits the magnitude of the reconstruction distortion to a maximum value, which is determined by quantizer coarseness. This is desirable because it eliminates large positive or negative errors which may be disruptive to the viewer. If the source video contains higher motion frames, then the side information deviate

significantly from the original signal. Thus, the side information would not lie within the reconstructed bin and the reconstruction scheme could only rely on the quantized symbol for reconstruction, which are quantized towards the bin boundary. If the quantization is coarse, contouring forms, which could be visually unpleasing. However, subtractive dithering, which is done by shifting the quantizer partitions for each pixel using a pseudo-random pattern, helps improve subjective quality in the reconstruction.

## 2.6 Video Quality Assessment

Video quality assessment methods are categorized as either subjective [39] and objective [40]. Subjective methods require human viewers to compare two video clips. However, objective test methods are used to measure and analyze the video signal. The peak signal to noise ratio (PSNR) has traditionally been used to evaluate the relative quality of reconstructed image [41];

$$PSNR = \frac{(2^n - 1)^2}{MSE} \quad (2.1)$$

where,  $n$  is the number of bits required to represent each pixel, MSE is the mean squared error between the distorted frame relative to the original frame, and MSE is given as;

$$MSE = \frac{\sum_{i,j} [X(i,j) - X'(i,j)]^2}{R \times C} \quad (2.2)$$

where,  $X(i, j)$  is the original frame,  $X(i, j)'$  is the reconstructed frame,  $R$  and  $C$  are the number of row and column elements respectively, and  $R \times C$  is the total number of pixels.

If each pixel is assigned 8 bits, PSNR is given by [42];

$$PSNR(dB) = 10 \log_{10} \left( \frac{255^2}{MSE} \right) \quad (2.3)$$

Video Quality Expert Group (VQEG) has used PSNR as the reference model in their study for video quality assessment methods. Typical values for the PSNR in image compression are between 30 and 40 dB.

## 2.7 Feedback Channel and Rate Control

Feedback channel between the encoder and the decoder is naturally employed in most practical WZVC [4], [16] schemes. Here, the decoder is able to enhance the quality of the decoded video by using additional refinement bits from the encoder. The channel encoder generates the parity bits for the WZ frame, which are then stored in a buffer. Iteration is used to determine the accurate amount of parity bits for decoding. Initially, the encoder transmits a portion of the parity bits from the buffer. Thereafter, the decoder uses a threshold to de-

termine whether the available information is sufficient for decoding. If the error probability at the decoder is above the threshold, the decoder requests additional syndrome bits from the buffer through the feedback channel. This transmission-request process is repeated until the error probability falls below the threshold.

The decoder must perform the calculations to determine the error probability in each block/ pixel and organize the parity bits in the buffer in order to get in correct order. In practice, computation of the error probability in real time results latency. Furthermore, the feedback channel is not available in numerous applications, that are performed between the decoder to the encoder; for example, a feedback channel is not available in broadcasting, as well as some unidirectional streaming applications, where the feedback rate-allocation solution is inapplicable.

In some researches, a feedback channel is not accounted for the encoder side rate allocation algorithms [19], [20]. Here, the minimum required number of parity bits are estimated at the encoder for each Wyner-Ziv frame under a given error probability, and parity bits are delivered to the decoder. Here, although the latency is minimized, the encoder complexity is increased.

There are some visible issues in no-feedback channel systems. The first issue is the correlation estimation at the encoder using the available  $X$  from the encoder and  $Y$  from the decoder;  $X$  and  $Y$  are unavailable at the decoder and encoder, respectively. A correlation estimation model (CEM) is required to exploit the statistical dependencies between  $X$  and an approximate version of  $Y$  ( $Y'$ ) estimated at the encoder using low-complexity algorithms. The CEM is proposed for pixel and transform domain WZVC in [43]. Another issue arises in allocating the appropriate number of bits for each frame  $X$  at the encoder, without feedback channel, so that it approaches the same rate distortion performance as when feedback is present. The total number of parity bits of  $X$  must be estimated to reconstruct  $X'$  with target quality from side information  $Y$  at the decoder. The research in [19] proposes a rate-allocation algorithm for pixel-domain WZVC with no-feedback channel; feedback channel is avoided with the certain loss of rate distortion performance. Overall, no-feedback channel increases encoder complexity and is not appropriate with the concepts of initial distributed video coding that utilize an encoder with the lowest possible complexity at the expense of higher decoder complexity.

## 2.8 Correlation Noise Between $X$ and $Y$

The side information is usually interpreted as an estimation of the original frame in the decoder. Error correcting codes (ECC) are used to improve the quality of the side information

until a target quality for the final decoded frame is obtained.

In [4], the decoder is responsible to explore all the source statistics, and compression is achieved by following the Wyner-Ziv paradigm. Therefore, the coding efficiency of Wyner-Ziv coding depends critically on the capability to model the statistical dependency between the original information at the encoder and the side information computed at the decoder. The original information is not available at the decoder. The side information quality also varies along with the video sequence and therefore correlation noise distribution between  $X$  and  $Y$  is not constant in the video sequence. Thus, it is difficult to maintain coding efficiency. For instance, in high motion sequences, it is more challenging to predict the errors in the side information where motion increases significantly.

In [16], estimation of correlation noise model is proposed using temporal correlation information, and in this case the motion compensated residual is obtained at the decoder. Initially, frame level correlation noise model is proposed at the decoder [16]. Furthermore, by considering the spatial stationary of the frames, the correlation noise statistics are estimated based on a block-by-block basis and at the pixel level.

### 2.8.1 Encoder Correlation Noise Models

Decoder needs to have reliable knowledge of the model that characterizes the statistical relation between the side information (SI) frame and the original WZ frame. The statistical dependency between these two frames corresponds to an error pattern characterized by the  $WZ - SI$  relationship [44].

The encoder computes offline WZ-SI relationship using Laplacian distribution;

$$f(WZ - SI) = \frac{\beta}{2} \exp(-\beta|WZ - SI|) \quad (2.4)$$

as the statistical correlation model between the WZ and SI [4], [17], [15], [45].

The Laplacian distribution converts pixel values of side information into soft-input information needed for Turbo decoding [15], [45]. Here  $\beta^2 = \frac{2}{\sigma^2}$  is computed offline, at the encoder, over the whole video sequence. The Laplacian parameter  $\beta$  is sent to the decoder before the WZ coding procedure starts and is kept constant throughout the decoding of all WZ frames. Here,  $\sigma^2$  is the variance of the residual between the WZ and the SI frames. However,  $\beta$  does not exploit the variability of the correlation model with respect to time and space [4], [15], [45].

Correlation model with respect to time is possible with frame level correlation noise estimation, where the variance of the residual between a WZ frame and the corresponding SI frame is calculated to compute the  $\beta$ . Subsequently, the decoder uses  $\beta$  in the decoding process of the WZ frame.

On other hand, the spatial variability of the correlation noise model is exploited by considering block level calculation of the Laplacian distribution parameter. The  $n \times n$  block variance  $\sigma_k^2$  is calculated from;

$$\sigma_k^2 = E((WZ_k - SI_k)^2) - (E(WZ_k - SI_k))^2 \quad (2.5)$$

where,  $WZ_k$  and  $SI_k$  represent the  $k^{th}$  sample block of the  $WZ$  and  $SI$  frames, respectively, and  $E(.)$  is the expectation operator.

For the encoder-side correlation noise calculation, the encoder needs to know the side information. However, generating the side information is an additional burden for the low complexity encoder, which violates the concept of the distributed video coding.

## 2.8.2 Decoder Correlation Noise Models

In this perspective, the realistic approach was considered for dynamic estimation of the Laplacian distribution parameter at the frame, block, and even at the pixel level. Here, it is not necessary to send the Laplacian distribution parameter from the encoder to the decoder. In [16], the frame, the block and the pixel level correlation noise estimation techniques are discussed in further detail.

Frame level correlation noise estimation was established with the help of residual frame generation from the frames  $X_B$  and  $X_F$ . The residual frame pixel  $R(x, y)$  is given as;

$$R(x, y) = ABS(X_B(x + dx_b, y + dy_b) - X_F(x + dx_f, y + dy_f)) \quad (2.6)$$

where,  $X_B(x + dx_b, y + dy_b)$  and  $X_F(x + dx_f, y + dy_f)$  are the backward and forward motion compensated frames, respectively. The  $(x, y)$  corresponds to the pixel location in the residual frame, and  $(dx_b, dy_b)$  and  $(dx_f, dy_f)$  represent the motion vectors for the  $X_B$  and  $X_F$  frames, respectively. The variance of the residual frame approximates the variance of the differences between WZ and SI frames. Variance of  $R$ ,  $\sigma_R^2$ , is given as;

$$\sigma_R^2 = E(R(x, y)^2) - (E(R(x, y)))^2. \quad (2.7)$$

Block level correlation noise estimation was performed using the initial step of residual

frame ( $R$ ) generation, as shown in (2.6). The  $k^{th}$  block variance of the residual frame,  $\sigma_{R_k}^2$ , can be obtained from (2.7). The  $\beta$  parameter of the  $k^{th}$  block of the WZ frame,  $\beta_k$ , is given as;

$$\beta_k^2 = \frac{2}{\sigma_k^2} \frac{E(R_k(x, y))}{E(R(x, y))} \quad (2.8)$$

where,  $E(R_k(x, y))$  and  $E(R(x, y))$  are the expectation operation over the  $k^{th}$  block and over the frame  $R$ , respectively.

## 2.9 Transmission Channel

Communication channels exist between transmitters and receivers. In wired communication, these channels are illustrated by twisted pairs, cables, wave guides, optical fiber and point-to-point microwave radio channels [46]. Regardless of the channel type used, the received signal differs from the input signal; here the differences are due to the channel impairments, which are known as dispersion, nonlinear distortions, delay and random noise.

In wireless WZVC, wireless channel imposes two major impairments such as channel noise and fading. Channel noise is modeled as additive white Gaussian noise (AWGN), and fading is modeled based on the channel randomness.

## 2.9.1 Additive White Gaussian Noise Channel (AWGN)

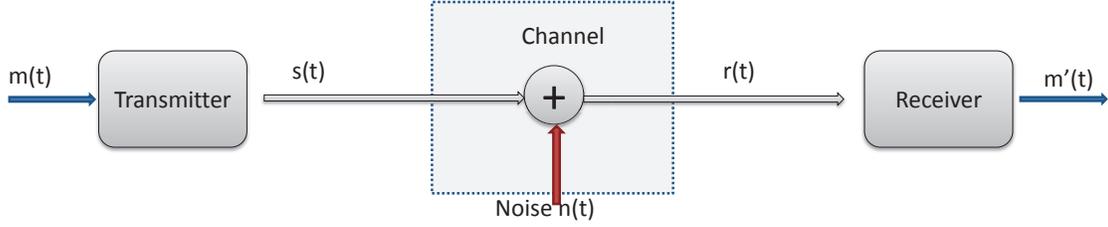


Figure 2.5: AWGN Channel Model

Thermal noise contains practically all frequency components up to some  $10^{13} Hz$  with the same power. Therefore, it is often referred to as white noise, analogous to white light that contains all colors with equal intensity. This white noise process can be characterized by its uniform power spectral density (PSD)  $N(\omega) = \frac{N_0}{2}$ . The autocorrelation function (ACF) of white noise is given as  $R(\tau) = \frac{N_0}{2}\delta(\tau)$ . In band-limited communication systems, PSD and ACF are given as  $N(\omega) = \frac{N_0}{2}; -B \leq \omega \leq B$ , and  $R(\tau) = \frac{N_0 B \times \sin(2\pi B\tau)}{2\pi B\tau} = N_0 B \times \text{sinc}(2\pi B\tau)$ , respectively [47], [48].

In time domain, the amplitude distribution of the white thermal noise has a normal or Gaussian distribution. The probability density function (PDF) is the bell-shaped Gaussian distribution given by;

$$p(n) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(\frac{-(n - \mu)^2}{2\sigma^2}\right) \quad (2.9)$$

where,  $\mu$  is the mean and  $\sigma^2$  is the variance. The effects of channel noise in received signal

can be diminished by increasing the power of transmitted signal. The signal to noise ratio (SNR) at the receiver is a measure of the received signal quality.

Fig. 2.5 shows a communication system with additive white Gaussian noise channel [47], [37]. Here the received signal  $r(t)$  is given as  $r(t) = s(t) + n(t)$ , where  $s(t)$  is the transmitted signal, and  $n(t)$  is the channel noise.

## 2.9.2 Fading Channel

Fading channels were carefully studied by information theoretical tools for a long time. Information theory provided, in many instances, the right guidance to a specific design for developing efficient communications systems. Multipath and shadowing are the two major complicated phenomena in radio wave propagation through wireless channels. A precise mathematical description of these phenomena is too complex for the analysis of communication systems. There were considerable efforts dedicated to statistical modeling and characterization of these effects, whereas simple and accurate statistical models were employed to analyze for fading channels in a particular propagation environment and basic communication scenario.

Fading in transmission produces fluctuations in the received signal's envelope and phase over time. If the fading is constant over the duration of a symbol's time, then it is called

slow fading. The slow fading has the symbol time duration  $T_s$ , which is smaller than the channel's coherence time  $T_c$ , otherwise it is characterized as fast fading. Coherence time measures the period of time over which the fading process is correlated. It is also related to channel Doppler spread;  $f_d$  by  $T_c \simeq \frac{1}{f_d}$  [47]. In slow fading, many successive symbols will be affected by a certain fade level, which accommodates burst errors, whereas fast fading reduces the correlation from symbol to symbol.

Frequency flat fading is defined based on all the spectral components of the transmitted signal that affect the signal in a similar manner [49]. In narrowband systems, fading is characterized as frequency flat fading. Here, the transmitted signal bandwidth is much smaller than the channel's coherence bandwidth  $f_c$ . The coherence bandwidth is defined as frequency range over which the fading process is correlated. Also the frequency selective fading occurs in wideband systems where the spectral components of the transmitted signal are affected by different amplitude gains and phase shifts. In the wideband, the transmitted bandwidth is broader than the channel's coherence bandwidth.

In narrowband systems, although fading occurs, the received carrier amplitude is modified by the fading amplitude  $\alpha$ , where  $\alpha$  is a random variable characterized by probability density function (PDF)  $p_\alpha(\alpha)$ ;  $\alpha$  is dependent on the nature of the radio propagation envi-

ronment [48].

## 2.10 Multipath Fading

Multipath fading is induced by the combination of constructive and destructive components of randomly delayed, reflected, scattered, and diffracted signal. This type of fading creates short-term signal variations that depend on the nature of the radio propagation environment.

Statistical behavior of the multipath fading is described by different models.

One such model is the Rayleigh distribution, which is a well known multipath fading channel model [50];

$$P_{\alpha}(\alpha) = \frac{2\alpha}{E(\alpha^2)} \exp\left(-\frac{\alpha^2}{E(\alpha^2)}\right) \quad (2.10)$$

where,  $\alpha$  is the channel fading amplitude,  $\alpha \geq 0$ .

The Nakagami- $q$  distribution or the Hoyt distribution is shown as [51];

$$P_{\alpha}(\alpha) = \frac{(1+q^2)\alpha}{qE(\alpha^2)} \exp\left(-\frac{(1+q^2)^2\alpha^2}{4q^2E(\alpha^2)}\right) I_0\left(\frac{(1-q^4)\alpha^2}{4q^2E(\alpha^2)}\right) \quad (2.11)$$

where,  $\alpha \geq 0$ ,  $q$  is the Nakagami- $q$  fading parameter, and  $I_0(\cdot)$  is the zero-order Bessel function of the first kind.

In the Nakagami- $n$  /Rice distribution [52] is given as;

$$P_{\alpha}(\alpha) = \frac{2(1+n^2)\exp(-n^2)\alpha}{E(\alpha^2)} \exp\left(-\frac{(1+n^2)\alpha^2}{E(\alpha^2)}\right) I_0\left(2n\alpha\sqrt{\frac{(1+n^2)}{E(\alpha^2)}}\right) \quad (2.12)$$

where,  $\alpha \geq 0$ , and  $n$  is the Nakagami- $n$  fading parameter;  $0 \leq n \leq \infty$ . Now the Rician factor ( $K$ ), is related to  $n$  as  $K = n^2$ . The Nakagami- $n$  distribution spans the range from Rayleigh fading (when  $n = 0$ ) to no fading (when constant amplitude, i.e.,  $n = \infty$ ).

The Nakagami- $m$  PDF is a central chi-square distribution, shown as [53];

$$P_{\alpha}(\alpha) = \frac{2m^m\alpha^{2m-1}}{E(\alpha^2)^m\Gamma(m)} \exp\left(\frac{-(m\alpha^2)}{E(\alpha^2)}\right) \quad (2.13)$$

where,  $\alpha \geq 0$ , and  $m$  is the Nakagami- $m$  fading parameter;  $\frac{1}{2} \leq m \leq \infty$ . For  $m = \frac{1}{2}$ , one-sided Gaussian distribution is observed, whereas the Rayleigh distribution is at  $m = 1$ . The Nakagami- $m$  fading channel converges to a nonfading AWGN channel when  $m \rightarrow +\infty$ .

### 2.10.1 Wireless Channel Model

Fading multipath channels are discussed with statistical models in [47] [54]. A fading multipath channel is generally categorized as a linear, time-varying system with an impulse response of  $c(t; \tau)$ . Time variations of the signal transmitted through the channel are gener-

ally called as Doppler spreading in the channel impulse response. A fading multipath channel may be commonly considered as a doubly spread channel in both time and frequency.

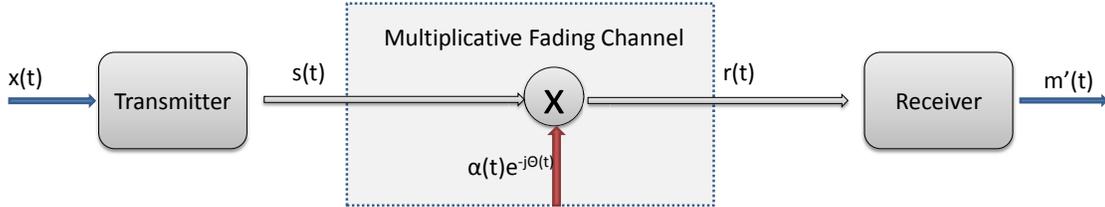


Figure 2.6: Fading Channel Model

Fig. 2.6 shows frequency non-selective wireless fading channel model [55]. If the bandwidth of transmitted low pass signal is much smaller than the coherence bandwidth of the channel, then all the frequency components in the transmitted signal undergo the same attenuation and phase shift over the transmission channel. This means that the time-variant transfer function  $C(t; f)$  of the channel is constant within that bandwidth of the signal. Therefore, that channel is called frequency-nonselctive, or flat-fading; it is also known as multiplicative channel model. The received signal,  $r(t)$ , is shown below [56];

$$r(t) = \int_{-\infty}^{\infty} c(t; \tau) s(t - \tau) d\tau \quad (2.14)$$

where,  $s(t)$  is the transmitted signal and  $c(t; \tau)$  is the impulse response.

A simplified form of  $r(t)$  is shown below;

$$r(t) = \alpha(t) \exp(j\theta(t)) s(t) \quad (2.15)$$

where,  $\alpha(t)$  represents the envelope, and  $\theta(t)$  represents the phase of the equivalent low pass channel response.

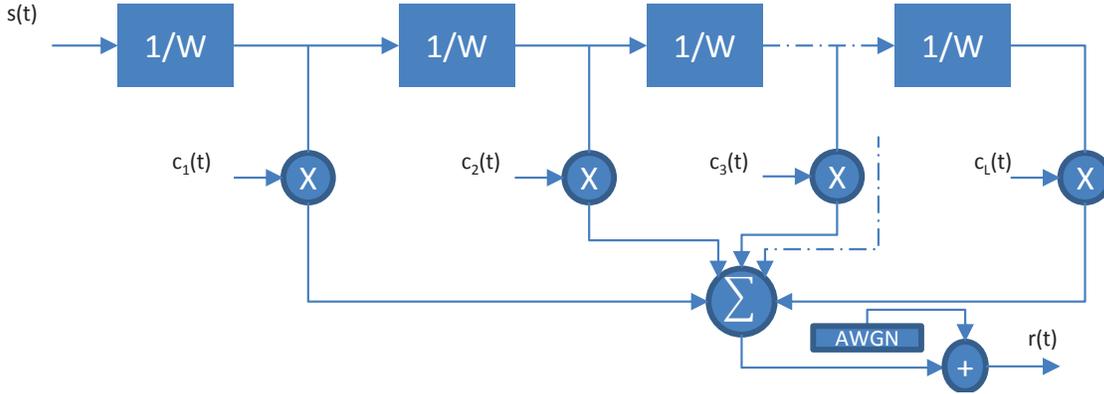


Figure 2.7: Frequency Selective Fading Channel Model

Fig. 2.7 shows a frequency selective channel model. If the transmitted signal's bandwidth is greater than the coherence bandwidth, then the frequency components of the transmitted signal are exposed to different gains and phase shifts. Thus, the channel is known as a frequency selective channel, or delay line channel model, where the time varying impulse response,  $c(t; \tau)$ , is given as [57];

$$c(t; \tau) = \sum_{n=1}^L c_n(t) \delta(\tau - \frac{n}{W}) \quad (2.16)$$

where,  $W$  is the channel bandwidth,  $c_n(t)$  is the complex-valued channel gain of the  $n^{th}$  multipath component, and  $L$  is the number of resolvable multipath components.

Time varying channel transfer function is given as;

$$C(t; f) = \sum_{n=1}^L c_n(t) \exp\left(j2\pi f \frac{n}{W}\right). \quad (2.17)$$

The randomly time-varying tap gains  $c_n(t)$  may also be represented by;

$$c_n(t) = \alpha_n(t) \exp(j\theta_n(t)). \quad (2.18)$$

where,  $n = 1, 2, \dots, L$ ,  $\alpha_n(t)$  represents the amplitudes, and  $\theta_n(t)$  represents the corresponding phases. Autocorrelation function,  $\Phi_n$ , of the tap gains  $c_n(t)$  are given based on the behavior of stationary (wide-sense) mutually uncorrelated random processes;

$$\Phi_n(\tau) = E\left(\frac{1}{2}c_n^*(t)c_n(t+\tau)\right). \quad (2.19)$$

Doppler power spectra is given as;

$$S_n(\lambda) = \int_{-\infty}^{\infty} \Phi_n(\tau) \exp(-j2\pi\lambda\tau) d\tau \quad (2.20)$$

where,  $\lambda$  is the frequency offset.

The envelope of the channel impulse response at any time instant will be represented by statistical models such as probability distribution (e.g. Rayleigh distribution), where phase

is uniformly distributed with the interval  $(0, 2\pi)$ . This envelop is given as;

$$R = |c(t; \tau)|. \quad (2.21)$$

For the frequency-nonselective channel, the envelope is the magnitude of the channel multiplicative as shown below;

$$R(t) = \alpha(t). \quad (2.22)$$

For the frequency-selective (tapped-delay-line) channel model, magnitude of the each tap gain is modeled with appropriate probability distribution (e.g. Rayleigh fading).

# Chapter 3

## Wyner-Ziv Video Coding with Wireless Channels

### 3.1 Introduction

In the past, many algorithms have been proposed to implement Wyner-Ziv video coding (WZVC). Pradhan and Ramachandran presented a practical framework of WZVC based on sending the syndrome of the codeword to compress the source [7], and they also provided a constructive practical framework based on algebraic trellis codes in distributed source coding using syndromes (DISCUS) [6]. Since then similar concepts have been extended to more advanced channel codes. Garcia-Frias and Zhao [58], [59], [60], Bajcsy and Mitran [61], [62] and Aaron *et al.* [4] showed the compression ratio can come close to the Slepian-Wolf bound using Turbo codes. Liveries *et al.* argued that low-density parity check codes (LDPC) are also suitable for this problem [63], [64]. Qian Xu *et al.* addressed the distributed joint source-channel coding using Raptor codes [65]. PRISM represented a radical departure from

conventional state-of-the-art video coding architectures [66]. In all these cases the channel was considered as *noiseless*.

In recent years, a few researchers have realized the impact of channel impairment on WZVC. S. Rane. *et. al* considered an error prone channel where the Systematic Lossy Error Protection (SLEP) is compared with Forward Error Correction (FEC). SLEP is a joint algorithm of the hybrid video coding and the WZVC [67], error resilience in current distributed video coding is discussed in [68]. The PRISM2007 proposed in [9] discusses the channel noise impact on the side information and outlined requirements for increasing predictors (received parity information) at the decoder side. However, until this point, previous researchers did not consider the wireless channel fading impairment (physical layer impacts) in WZ video transmission. Although the previous work had substantially contributed on WVC in application layer, the consideration towards implementing multipath wireless channel has become important with the increased deployment of wireless video sensors. In this chapter, we investigate WZVC over the *time varying fading* wireless channel.

This chapter is organized as follows: Section 3.2 briefs the theoretical motivations and contributions. Section 3.3 describes the background theory, section 3.4 investigates WZVC over a single input single output (SISO) wireless channel. Section 3.5 proposes the WZVC

over the MIMO wireless channel model. In Section 3.6, a simulation model is explained and Section 3.7 describes the results and discussions of our work. Finally summary is presented in Section 3.8.

## 3.2 Theoretical Motivations and Contributions

Channel noise and fading are the major impairments in the multipath wireless channel. Multipath fading requires special attention as it has different probabilistic and time/frequency domain characteristics.

Fig. 3.1 shows a block diagram of transmitting parity of the WZ video frame over a single wireless channel. Quantized pixel values are converted into bit-plane, and bit sequence from the bit-plane is coded with the channel encoder. For example, a pixel  $P_i$  of the video frame is converted into bit values in the bit-plane column. The channel coding generates the parity bits. Subsequently the modulated signals of the parity bits are transmitted via a single wireless channel, where the parity signals are influenced by the wireless channel noise and fading. Due to the uncertainty of the wireless channel, the error probability is high in the received pixel. Therefore, the reconstruction of the WZ frame contains significant distortions.

Fig. 3.2 shows a pixel transmission over the multiple wireless channels; the wireless channel with  $M$  inputs and  $N$  outputs. Therefore, the pixels in the video frame are able to

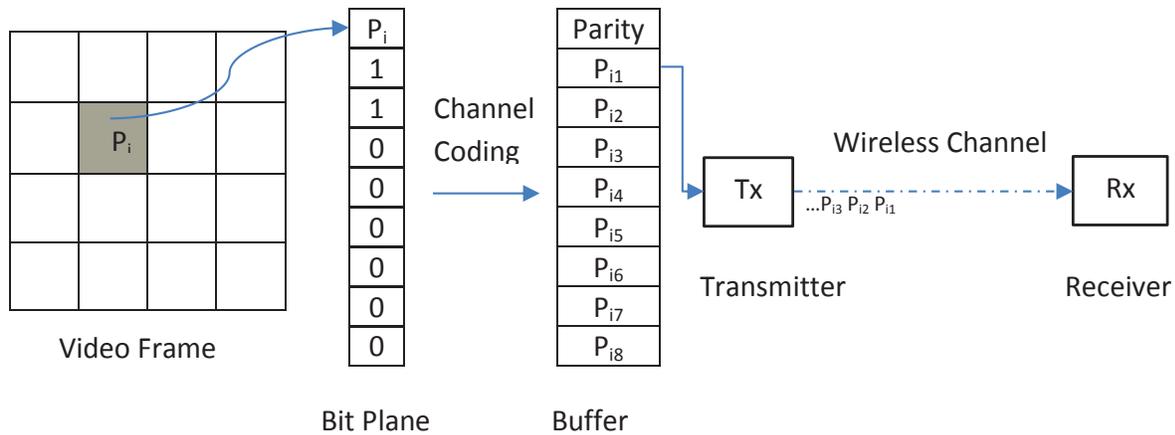


Figure 3.1: Video Transmission over a Single Wireless Channel

reach the receiver via multiple channels. In wireless video stream, video coding with channel diversity facilitates the transmission of each video frame through multiple wireless channels, whereas video coding in a single wireless channel transmits all the contents of video frame through a single channel. The bit-plane concept in WZVC has more potential of converting each pixel into bits, which can be transmitted via multiple channels. Therefore each pixel information has the potential to reach the receiver by multiple wireless channel, which increases the importance of MIMO in wireless video coding.

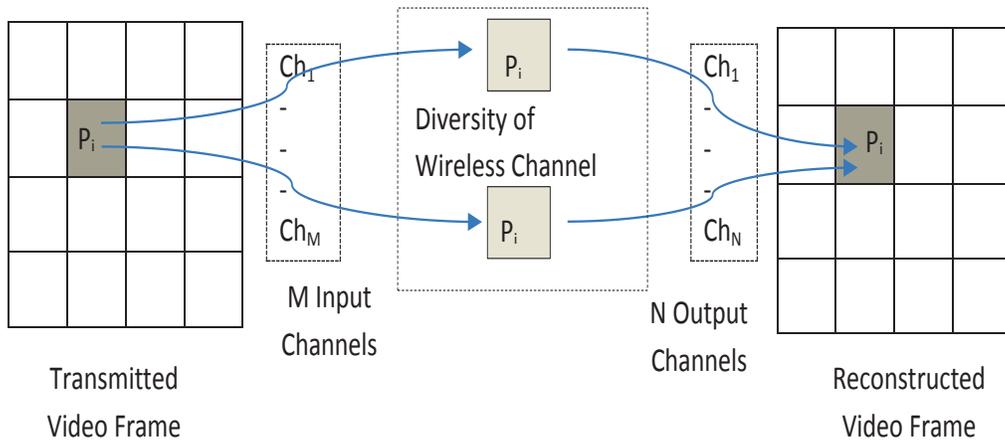


Figure 3.2: Video Transmission over Wireless Channel with Transmit-Receive Diversity

Furthermore, the channel coherence bandwidth of a typical WZVC is greater than the transmission bandwidth in low bit rate transmission. Therefore, the multipath fading will result in *flat* fading, which means that the entire signal will experience a single time varying gain,  $\alpha$  with Rayleigh probabilistic distribution.

In this chapter, we first investigate reconstructed WZ video quality under the noise

and random variations of the wireless channel fading gain ( $\alpha$ ). Since, the fading gain of the wireless channel randomly fluctuates, the received soft-values of parity influence the reconstructed video quality. Hence, we incorporate the transmit-receive diversity technique, called MIMO approach to the wireless WZVC. We also investigate *channel-metric value* and *bit-metric value* in wireless WZVC with respect to transmit-receive diversity scheme. The *channel-metric value* depends on both the channel noise and fading gains of the wireless channel, and the *bit-metric value* is determined with respect to the correlation noise between WZ frame and side information.

### 3.3 Theory

In this section, the WZVC structure and its mathematical derivations are briefed. A WZVC codec consists of four main components: a quantizer, channel encoder, channel decoder, and reconstruction module. The quantizer decides bit-depths of the pixels. For instance, in an eight-bit quantizer, the maximum and minimum possible pixel values are 255 and 0, respectively. Additionally, this section describes the wireless channel and its transmission diversity.

#### 3.3.1 Bit-plane Extractor

The bit-plane extractor organizes the quantized bits in an order from most significant bit (MSB) to least significant bit (LSB) for each pixel. By taking the YUV color space into con-

sideration, where  $Y$  contains the luminance information, and  $U$  and  $V$  contains the chrominance information, and letting  $h$  and  $w$  denote the height and width of the luminance video frame; we can express the total number of pixels of the frame to be  $\frac{3wh}{2}$  ( $= wh + \frac{wh}{4} + \frac{wh}{4}$ ). This predicts the dimensions of the bit-plane to be  $8 \times \frac{3wh}{2}$ . The bit stream is obtained from the bit-plane calculation, and is divided into blocks. For a given block size of  $K$ , the bit sequence in a block is written as  $u_l = u_0, u_1, \dots, u_K$ .

### 3.3.2 Turbo Coder

Turbo encoder is shown in Fig. 2.4, where input block length is taken as  $K$  and the number of memory slots is three. In WZVC, systematic bits are discarded at the encoder and parity bits are transmitted to the decoder. Parity bits are punctured in order to reduce the number of transmitting parities. Puncturing is done with some specific pattern in simulation.

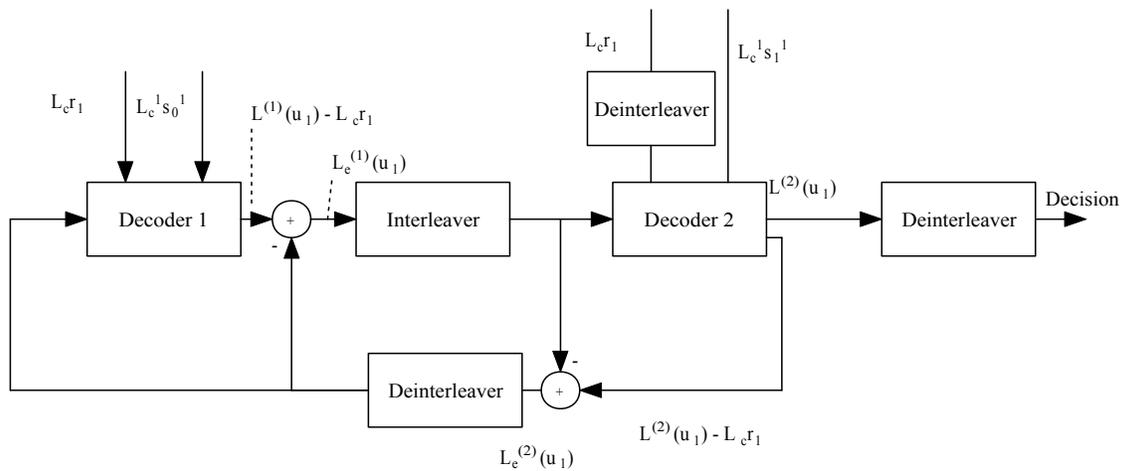


Figure 3.3: Block Diagram of Turbo Decoder

Fig. 3.3 depicts the block diagram of a Turbo decoder: here,  $r_l$  is the side information, which is used as the systematic information;  $s_0^1$  and  $s_1^1$  are the received parity information;  $L_c^1$  is the channel-metric value; and  $L_c$  is the bit-metric value due to the correlation noise between the WZ frame and side information. The log-likelihood ratio (LLR or L-value) [69] of a transmitted bit  $u_l$  of a video frame given the corresponding received systematic signal  $r_l$  is defined as  $L(u_l|r_l)$ ;

$$\begin{aligned}
L(u_l|r_l) &= \ln \frac{P(u_l = +1|r_l)}{P(u_l = -1|r_l)}, \\
&= \ln \frac{P(r_l|u_l = +1) P(u_l = +1)}{P(r_l|u_l = -1) P(u_l = -1)}, \\
&= \ln \frac{P(r_l|u_l = +1)}{P(r_l|u_l = -1)} + \ln \frac{P(u_l = +1)}{P(u_l = -1)}. \tag{3.1}
\end{aligned}$$

Now,  $r_l$  is given as  $r_l = u_l + n_0$ , where  $r_l$  is obtained from bit-plane of  $Y:Y\varepsilon\{r_1, r_2, \dots, r_l, \dots, r_n\}$ ,  $u_l$  is received from bit-plane of  $X:X\varepsilon\{u_1, u_2, \dots, u_l, \dots, u_n\}$ ; here,  $r_l$  is used as the received systematic signal, and  $n_0$  is the correlation noise between the WZ frame ( $X$ ) and side information ( $Y$ ). Now the equation (3.1) is expanded with respect to the Gaussian distribution of  $n_0$  and it is given as;

$$L(u_l|r_l) = \ln \frac{\exp(-(E_s/N_0)(r_l - 1)^2)}{\exp(-(E_s/N_0)(r_l + 1)^2)} + \ln \frac{P(u_l = +1)}{P(u_l = -1)}, \tag{3.2}$$

where,  $u_l$  and  $r_l$  have both been normalized by a factor of  $\sqrt{E_s}$ ,  $E_s$  is the symbol energy,

$E_s/N_0$  is the signal to noise ratio. The LLR can be simplified to;

$$L(u_l|r_l) = 4\frac{E_s}{N_0}r_l + \ln \frac{P(u_l = +1)}{P(u_l = -1)} \quad (3.3)$$

$$L(u_l|r_l) = L_c r_l + L_a(u_l), \quad (3.4)$$

where,  $L_c = 4(E_s/N_0) = \frac{2E_s}{\sigma_{pic}^2}$ , is the bit-metric factor for the side information,  $L_a(u_l)$  is the ‘a priori’ L - value of the bit  $u_l$ , and  $\sigma_{pic}^2$  is the variance of correlation noise in between the WZ frame and the side information. Referring back to Fig. 3.3,  $L^{(1)}(u_l)$  is the ‘a posteriori’ L -value of each information bit produced by decoder 1, the received signal is  $r_l$ , and the ‘a priori’ input is  $L_a^{(1)}$ . The extrinsic ‘a posteriori’ L -value associated with each information bit produced by decoder 1 is shown below;

$$L_e^{(1)}(u_l) = L^{(1)}(u_l) - [L_c r_l + L_a^{(1)}(u_l)]. \quad (3.5)$$

$L_e^{(1)}(u_l)$  is passed to the input of decoder 2 after interleaving as the ‘a priori’ value  $L_a^{(2)}(u_l)$ . Similarly, output of decoder 2 contains two terms  $L^{(2)}(u_l)$  and  $L_e^{(2)}(u_l)$ , where  $L_e^{(2)}(u_l) = L_a^{(1)}(u_l)$ .

Initially, equally likely information bits are assumed, therefore,  $L_a^{(1)}(u_l)$  is set to zero. The extrinsic L - values passed from the first decoder to the second decoder during the iterative decoding process are treated like new sets of ‘a priori’ probabilities by the maximum a posteriori (MAP) algorithm. After a sufficient number of iterations, the decoded information

bits are determined from the ‘a posteriori’ L - values  $L^{(2)}(u_i)$ .

The soft-values of the systematic information is computed of using bit-metric value ( $L_c$ ) with respect to the video contents. Soft-value of parity information is determined using channel-metric value ( $L_c^1$ ). The received parities  $s_0^1$  and  $s_1^1$  are given as  $s_i^1 = s_i + n_c$ , and  $i = 0, 1$ . From (3.4), LLR for the received parity information is given as  $L(s_i|s_i^1) = L_c^1 s_i + L_a(s_i)$ , where  $L_c^1 = 4(E_{s-p}/N_c)$ ;  $E_{s-p}$  is the energy of parity symbol; and  $n_c$  is the channel noise; and  $N_c$  is the noise power. Moreover,  $N_c = 2 \times \sigma_{ch}^2$ , where  $\sigma_{ch}^2$  is the channel noise variance.

### 3.3.3 Wireless Channel

The wireless channel is generally influenced by the multi-path behavior and fading. The multipath signals arrive with different delays and attenuations in different environment known as linear time variant channels. The received signal is given as  $r(t) = s(t) * c(\tau; t) + n(t)$ , where  $c(\tau; t)$  is the impulse response of the channel at time t for an impulse applied at time  $t - \tau$ . A radio multipath channel is characterized by its impulse response as,  $c(\tau; t) = \sum_{k=1}^L a_k(t) \delta(\tau - \tau_k)$ , where  $a_k(t)$  represents time varying attenuation factor for L multipaths, and  $\tau_k$  is the corresponding delays. Hence the received signal,  $r(t) = \sum_{k=1}^L a_k(t) s(t - \tau_k) + n(t)$ , consists of L multipath components each characterized by an attenuation and a delay. If we consider the frequency non selective channel (flat fading), then the received signal  $r(t)$  is given as  $r(t) = c(t)s(t) + n(t)$  [55], [70]. Thus the symbol-spaced samples of the matched filter output

of the received signal is given as [70];

$$r(k) = h(k)s(k) + n(k) \quad (3.6)$$

where,  $r(k)$  is the received signal,  $h(k)$  is the complex multiplicative distortion between the transmitter and receiver, and  $|h(k)| (= \alpha(k))$  is the Rayleigh fading coefficient,  $n(k)$  is the independent and identically distributed (i.i.d) complex AWGN, and  $s(k)$  is the transmitted signal.

Transmit-receive diversity is proposed to reduce the fading effect which improves signal quality at the receiver [22]. Subsequently, it also improves the error performance, data rate, and the capacity of a wireless communication system. We provide the multiple transmit multiple receive antenna configurations, as well as the equations of the received signals for a MIMO wireless channel in the next section.

### 3.3.4 MIMO Wireless Channel

In this section, we provide the derivations and formulas with respect to the multiple transmit multiple receive antenna configurations.

Let  $M$  and  $N$  be the numbers of transmit and receive antennas. Fig. 3.4 shows the configuration of  $M = 2$  and  $N = 4$ . where  $s_0$  and  $s_1$  are the coded symbols at time  $t$ , and  $-s_1^*$  and  $s_0^*$  are at time  $t + T$  from transmitting antennas  $\text{Tx}_0$  and  $\text{Tx}_1$  respectively [23]. The

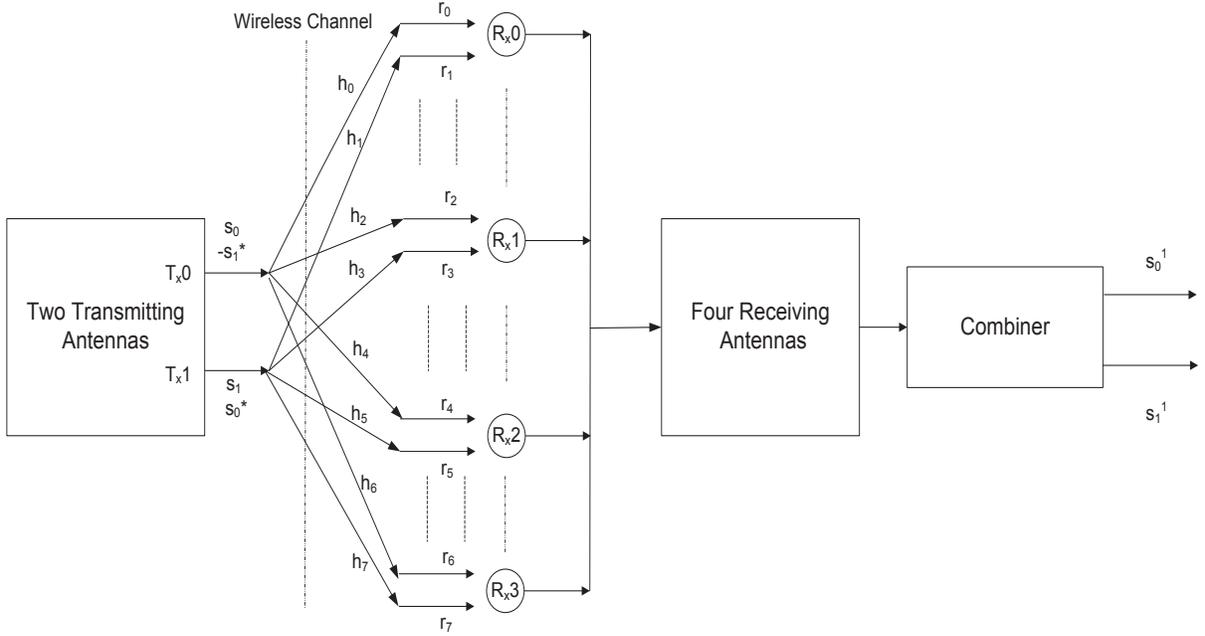


Figure 3.4: Antenna Configuration with Two Transmit Four Receive Antennas

$i^{th}$  receiving antenna;  $Rx_i$  receives signals  $r_{2i}$  and  $r_{2i+1}$  at time  $t$  and  $t + T$ , respectively ( $i=0,1,2,3$ ). At time  $t$ , the complex multiplicative distortion between the transmitting and receiving antenna is  $h_i(t) = h_i(t + T) = h_i = \alpha_i e^{j\theta_i}$ , where  $\alpha_i$  is the fading coefficient of the  $i^{th}$  channel and  $\alpha_i = |h_i|$  [22]. The distorted signals at the receiving end are shown below;

$$r_{2i} = h_{2i}s_0 + h_{2i+1}s_1 + n_{2i} \quad (3.7)$$

$$r_{2i+1} = -h_{2i}s_1^* + h_{2i+1}s_0^* + n_{2i+1} \quad (3.8)$$

where, the value of  $i$  depends on the number of receiving antennas, i.e., if  $N = 2$ , then  $i = 0, 1$ , and if  $N = 4$ , then  $i = 0, 1, 2, 3$ . The  $n_{2i}$  and the  $n_{2i+1}$  are the AWGN noise at  $t$

and  $t + T$  for the  $i^{th}$  receiver. The combiner computes the received parity sequence of WZ video frames. The output of the combiner for  $M = 2$  and  $N = 2$  can be given as;

$$s_0^1 = h_0^* r_0 + h_1 r_1^* + h_2^* r_2 + h_3 r_3^*, \quad (3.9)$$

$$s_1^1 = h_1^* r_0 - h_0 r_1^* + h_3^* r_2 - h_2 r_3^*. \quad (3.10)$$

Received signals  $r_0, r_1, r_2,$  and  $r_3$  can be obtained by substituting  $i = 0, 1$  in (3.7) and (3.8). Then  $r_0, r_1, r_2$  and  $r_3$  are substituted in (3.9) and (3.10);

$$s_0^1 = (\alpha_0^2 + \alpha_1^2 + \alpha_2^2 + \alpha_3^2) s_0 + h_0^* n_0 + h_1 n_1^* + h_2^* n_2 + h_3 n_3^* \quad (3.11)$$

$$s_1^1 = (\alpha_0^2 + \alpha_1^2 + \alpha_2^2 + \alpha_3^2) s_1 - h_0 n_1^* + h_1^* n_0 - h_2 n_3^* + h_3^* n_2 \quad (3.12)$$

where, the symbols  $s_0^1$  and  $s_1^1$  are the recombined signals for two-input two-output (2I2O) channel. Similarly, the received signals  $(r_0, \dots, r_i, \dots, r_7)$  of two-input four-output (2I4O) channel are derived by substituting  $i = 0, 1, 2, 3$  in (3.7) and (3.8). Then the recombined symbols  $s_0^1$  and  $s_1^1$  are obtained by substituting  $r_0, \dots, r_7$  in (3.9) and (3.10);

$$\begin{aligned} s_0^1 = & (\alpha_0^2 + \alpha_1^2 + \alpha_2^2 + \alpha_3^2 + \alpha_4^2 + \alpha_5^2 + \alpha_6^2 + \alpha_7^2) s_0 \\ & + h_0^* n_0 + h_1 n_1^* + h_2^* n_2 + h_3 n_3^* + h_4^* n_4 + h_5 n_5^* + h_6^* n_6 + h_7 n_7^*, \end{aligned} \quad (3.13)$$

$$s_1^1 = (\alpha_0^2 + \alpha_1^2 + \alpha_2^2 + \alpha_3^2 + \alpha_0^2 + \alpha_1^2 + \alpha_2^2 + \alpha_3^2)s_1 - h_0n_1^* + h_1^*n_0 - h_2n_3^* + h_3^*n_2 - h_4n_5^* + h_5^*n_4 - h_6n_7^* + h_7^*n_6. \quad (3.14)$$

### 3.4 WZVC over a Single Wireless Channel

The schematic block diagram of WZVC over a single wireless channel is shown in Fig. 3.5. If we let  $X_1, X_2, \dots, X_N$  be the frames of a video sequence, then WZ frame  $X_{2i}$  is sent through the quantizer; here, the quantization level depends on the expected quality of output and the available channel bandwidth.

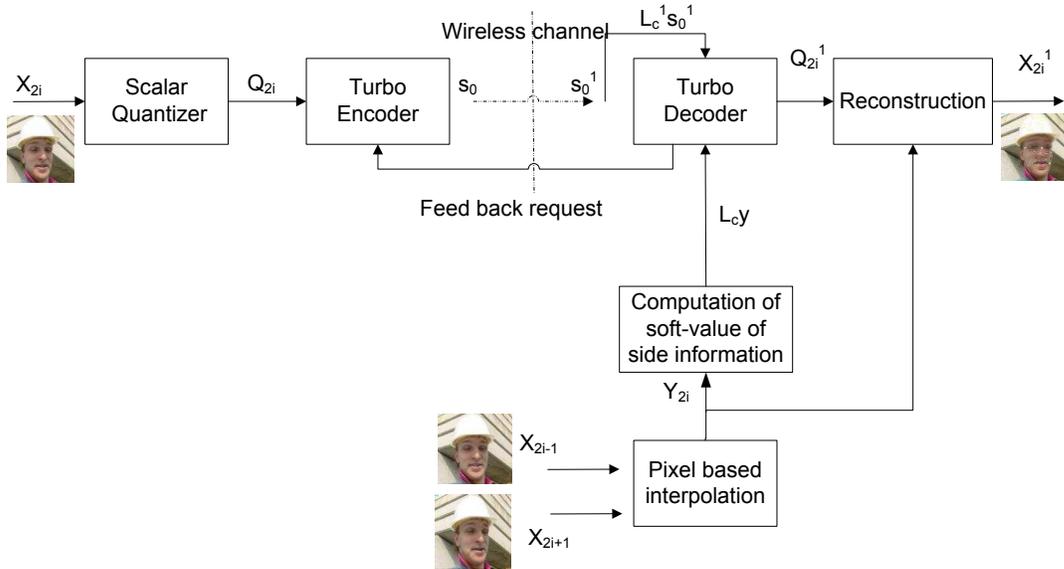


Figure 3.5: WZVC over a Single Wireless Channel

The quantized stream  $q_{2i}$  is obtained from each pixel of the WZ frame, and quantized bits are fed into the Turbo encoder.  $s_0$  is the coded symbol of parity bit sequence. Puncturing is used to control the number of transmitted parity bits. The punctured parity bits are stored

in the buffer, whereas the systematic bits are discarded from Turbo encoder. At the decoder, side information  $Y_{2i}$  is computed using pixel interpolation of the available key frames:  $X_{2i+1}$  and  $X_{2i-1}$

The side information is fed into the Turbo decoder as the systematic information of the WZ frame. Since  $Y = X + N$ , soft-value of side information is  $\frac{2E_s y}{\sigma_{pic}^2}$ , where the bit-metric value is  $L_c = \frac{2E_s}{\sigma_{pic}^2}$ ; here  $\sigma_{pic}^2$  is the variance of the correlation noise between the WZ frame and SI. At the decoder, the block “Computation of soft-value of side information” extracts bits from each pixels of side information (*bit-plane extraction*), and computes  $L_c y$ .

The parity bits of the WZ video frames are modulated before transmitting to the wireless channel; BPSK is used in our simulations. We have considered the Rayleigh flat fading channel. The received parity via the wireless channel is;

$$s_0^1 = \alpha s_0 + n_c \quad (3.15)$$

where,  $s_0^1$  and  $s_0$  are the received and transmitted symbols respectively,  $\alpha$  is the fading coefficient, and  $n_c$  is the AWGN noise. The soft-value of the received parity information can be computed with respect to the AWGN noise and the fading coefficient of the wireless

channel. The derivation of the channel-metric value  $L_c^1$  is obtained from equation (3.1).

Additionally;

$$L(s_0|s_0^1) = \ln \frac{\exp(-(E_{s-p}/N_c)(s_0^1 - \alpha)^2)}{\exp(-(E_{s-p}/N_c)(s_0^1 + \alpha)^2)} + \ln \frac{P(s_0 = +1)}{P(s_0 = -1)}, \quad (3.16)$$

$$= 4 \frac{\alpha E_{s-p}}{N_c} s_0^1 + \ln \frac{P(u_l = +1)}{P(u_l = -1)}. \quad (3.17)$$

The above equation is compared with equation (3.4), and then the channel-metric value  $L_c^1$  is derived as;

$$L_c^1 = \alpha \frac{4E_{s-p}}{N_c} = \alpha \frac{2E_{s-p}}{\sigma_{ch}^2} \quad (3.18)$$

where,  $E_{s-p}$  is the energy of parity symbol,  $\sigma_{ch}^2 (= \frac{N_\epsilon}{2})$  is the variance of AWGN, and  $\alpha$  is the fading coefficient of the channel, which varies randomly for each WZ frame. There has been much research done on the estimation of the wireless channel's characteristics. This can be easily done by transmitting a training sequence, or a pilot symbol, periodically. Since the pilot is known a priori, the unknown channel can be estimated by the receiver [70], [55].

The Turbo decoder receives the soft-values of parity signals as  $L_c^1 s_0^1$  and  $L_c^1 s_1^1$ , and the soft-value of systematic information from the side information as  $L_c y$ . Additionally, it generates the decoded symbol  $q_{2i}^1$  using the MAP algorithm. The reconstruction is completed with respect to the statistical relationship between the WZ frame and the side information in

order to suppress some of the quantization noise; i.e., the reconstructed frame is given as  $f(Y, Z)$  in [3], where  $Z$  is the residual information between the WZ frame and the side information. In a noisy, fading channel, the reconstructed video frame depends on  $L_c^1 s_0^1$ ,  $L_c^1 s_1^1$  and  $L_c y$ .

### 3.5 WZVC over MIMO Wireless Channel

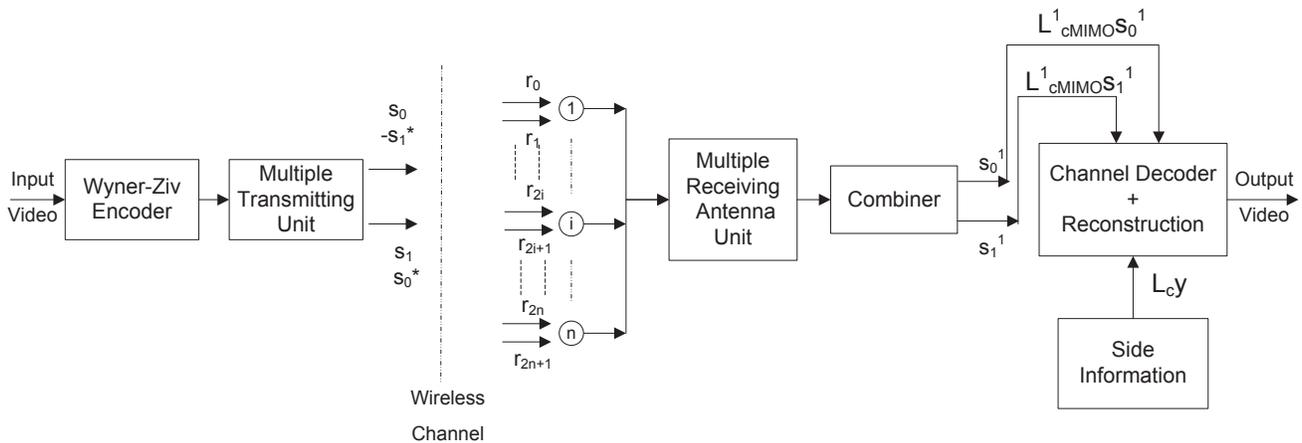


Figure 3.6: WZVC over Multiple Input Multiple Output (MIMO) Wireless Channel

The block diagram of WZVC over the MIMO wireless channel is shown in Fig. 3.6. Here, the punctured parity bits are placed on the multiple input wireless channel after the space time coding (STC). We have chosen the two input and multiple output diversity scheme (2IMO), where each channel is considered to be a Rayleigh flat fading channel.

Combiner outputs of 2I2O,  $s_0^1$  and  $s_1^1$ , are shown in (3.11) and (3.12) respectively; similarly (3.13) and (3.14) show for 2I4O. Below the simplified form of the combiner outputs are given;

$$s_0^1 = \left( \sum_{j=0}^{k-1} |\alpha_j|^2 \right) s_0 + \sum_{j=0}^{k-1} (h_{2j}^* n_{2j} + h_{2j+1} n_{2j+1}^*), \quad (3.19)$$

$$s_1^1 = \left( \sum_{j=0}^{k-1} |\alpha_j|^2 \right) s_1 + \sum_{j=0}^{k-1} (-h_{2j} n_{2j}^* + h_{2j+1}^* n_{2j}) \quad (3.20)$$

where,  $k$  is the number of channels;  $k = 4$  for 2I2O and  $k = 8$  for 2I4O. From 3.19,  $s_0^1$  can be written as  $s_0^1 = \alpha s_0 + n_c$ , where,

$$\alpha = \sum_{j=0}^{k-1} |\alpha_j|^2, \quad (3.21)$$

$$n_c = \sum_{j=0}^{k-1} (h_{2k}^* n_{2k} + h_{2k-1} n_{2k-1}^*). \quad (3.22)$$

From (3.17), LLR value with respect to the parity information of WZ frame via the MIMO wireless channel is obtained as;

$$L(s_0|s_0^1) = 4 \frac{\sum_{j=0}^{k-1} |\alpha_j|^2 E_{s-p}}{N_c} s_0^1 + \ln \frac{P(s_0 = +1)}{P(s_0 = -1)}. \quad (3.23)$$

By comparing  $L(s_0|s_0^1)$  with (3.18), the channel-metric value of the MIMO wireless chan-

nel  $L_{cMIMO}^1$  can be obtained as;

$$L_{cMIMO}^1 = \sum_{j=0}^{k-1} |\alpha_j|^2 \frac{4E_{s-p}}{N_c} = \sum_{j=0}^{k-1} |\alpha_j|^2 \frac{2E_{s-p}}{\sigma_{ch}^2} \quad (3.24)$$

where,  $\sigma_{ch}^2$  is the variance of  $n_c$ . The new channel-metric value  $L_{cMIMO}^1$  for the WZVC over MIMO wireless channel varies according to the number of channels  $k$ . The Turbo decoder receives the soft-values of parity signals as  $L_{cMIMO}^1 s_0^1$  and  $L_{cMIMO}^1 s_1^1$ ; and the soft-value of the systematic information from the side information  $L_c y$ , as shown in the Fig. 3.6. Since the MAP algorithm of the Turbo decoder depends on  $L_{cMIMO}^1$  and  $L_c$ , the reconstructed video quality depends on  $L_c y$ ,  $L_{cMIMO}^1 s_0^1$  and  $L_{cMIMO}^1 s_1^1$ .

## 3.6 Simulations

This section explains the simulation model of WZVC with the wireless fading channel and AWGN. Fig. 3.7 shows the schematic diagram of our simulation model. Quantized pixel values were converted into a bit stream by the bit-plane extractor. The bit stream is fed into the Turbo encoder, and the Turbo encoder gives the output as systematic bits and parity bits. The Turbo encoder was assigned with a block size of 400 bits, 3 memory slots. In the coded output, the systematic bits are ignored, and the parity bits are punctured to control the transmission rate. The BPSK modulation scheme is used to generate the symbols  $s_0$  and  $s_1$  from the parity bits. The STC is used for transmitting symbols in transmit-diversity

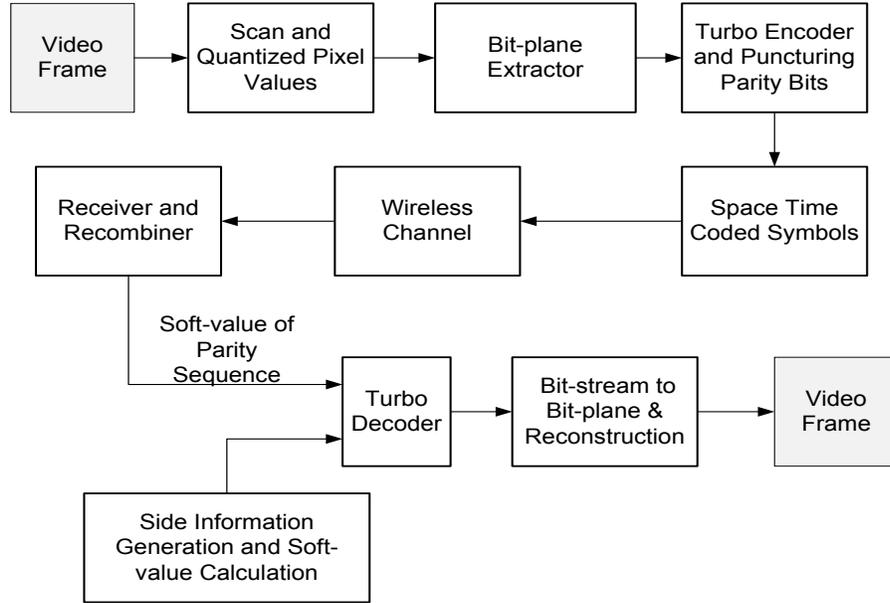


Figure 3.7: Simulation Model

scheme. For the simulation purpose, it is considered that the fading coefficient  $\alpha$  changes with frequency  $\frac{1}{n}$ , where the frame rate of the video sequence is  $n$  frames per second (fps). The receiver computes the resultant received parity information from the combiner. The Turbo decoder requires the soft-value of systematic information  $L_c y$  and the soft-value of the parity information  $L_c^1 s_0^1$  and  $L_c^1 s_1^1$ , as shown in Fig. 3.3. The systematic information  $y$  was derived from the side information. The simulation is done for the bit rate of 190Kbps. WZVC simulation is done with respect to the channel as shown below;

- WZVC Simulation for Noiseless Channel

The parity bits are transmitted as hard values without the influence of channel, the bit error rate (BER) is equal to zero. Side information is computed as soft-value:  $L_c y$ . Here, the reconstructed frame depends on  $L_c y$  and the hard-value of parity bits

- WZVC Simulation with Noisy Channel

The parity symbols are transmitted via Gaussian noisy channel. The noise variance  $\sigma_{ch}^2$  was calculated based on SNR of the Gaussian channel. The channel-metric value of the channel was estimated using equation  $L_c^1 = \frac{4E_{s-p}}{N_0} = \frac{2E_{s-p}}{\sigma_{ch}^2}$ , and  $L_c$  is the bit metric value; here, the reconstructed frame proportional to  $L_c y$ ,  $L_c^1 s_0^1$  and  $L_c^1 s_1^1$

- WZVC over Wireless Channel

The channel fading influences the transmitted parity symbols. We have considered the variance of the Rayleigh channel is 0.5. Rayleigh fading coefficient is determined in terms of the characteristics of standard deviation of Rayleigh distribution and the normal distribution  $N(0,1)$ . The channel-metric value is determined by  $L_c^1 = \alpha \frac{4E_{s-p}}{N_c} = \alpha \frac{2E_{s-p}}{\sigma_{ch}^2}$  or  $L_{cMIMO}^1 = \sum_{j=1}^k |\alpha_j|^2 \frac{4E_{s-p}}{N_0} = \sum_{j=1}^k |\alpha_j|^2 \frac{2E_{s-p}}{\sigma_{ch}^2}$ , and the  $L_c$  is the bit metric value.

## 3.7 Simulation Results

This section discusses the simulation results of WZVC with noiseless (also known as lossless), AWGN, and wireless fading channels; here the reconstructed WZ video quality is compared with different channel conditions.

### 3.7.1 Reconstructed WZ Frame of the WZVC with Different Channel Conditions

The received video frames are investigated in different channel conditions such as noiseless channel, AWGN and fading channel, here Foreman, Carphone and Akiyo video sequences are considered in the simulations.

Fig. 3.8 illustrates the reconstructed Foreman video frames of the WZVC with different channel conditions. Based on the output video quality, we have selected 1<sup>st</sup>, 9<sup>th</sup>, 17<sup>th</sup> and 21<sup>st</sup> frames from the reconstructed sequence. The reconstructed video quality decreases with AWGN and fading channel in above selected frames.

The reconstructed video quality of the 1<sup>st</sup> frame shows minor changes from lossless channel to lossy channel, whereas frames 9, 17 and 21 show significant loss of quality in the lossy channels.

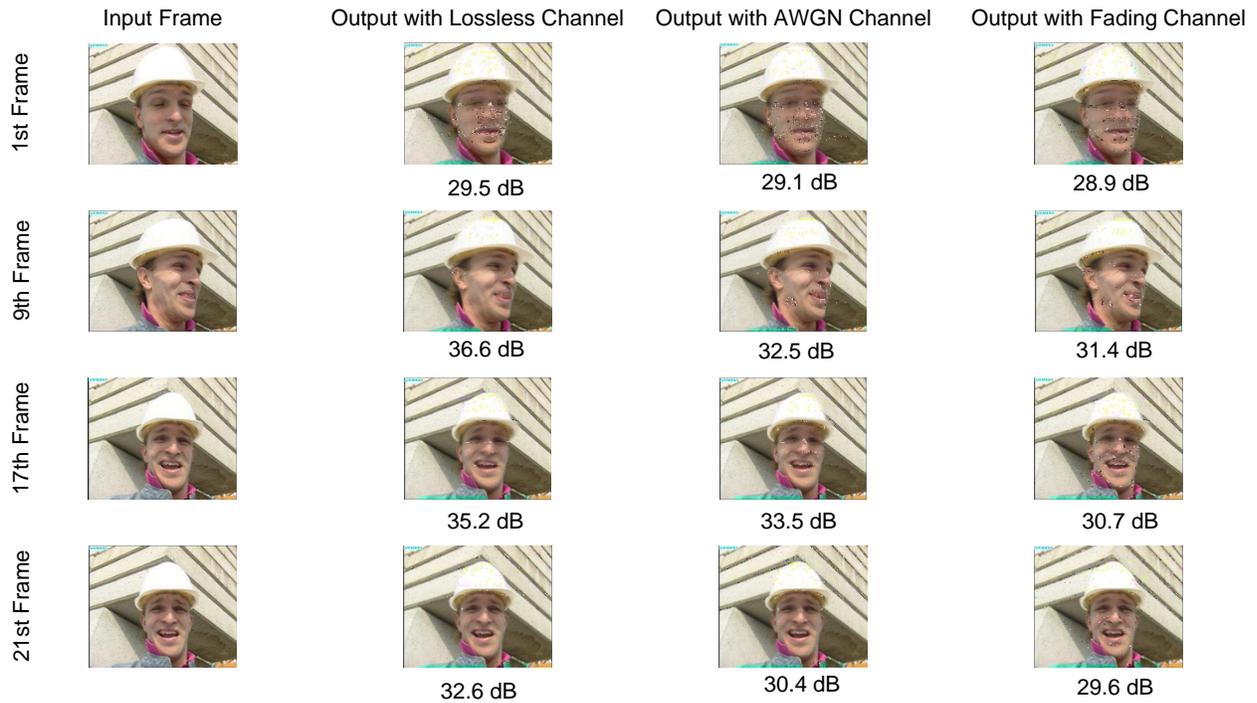


Figure 3.8: Reconstructed Foreman Video Frames with Lossless, AWGN, and Fading Channels at  $SNR = 2$  dB

Fig. 3.9 shows the reconstructed Carphone video frames. The reconstructed video frame quality varies with the channel property, i.e., the PSNR value is reduced for AWGN and then fading channel compared to noiseless channel. For example, Carphone frames 11, 15 and 17 show 1.6 dB, 1.1 dB and 1.2 dB reduction in PSNR, respectively.

Reconstructed video frames of Akiyo sequence are shown in Fig. 3.10. Akiyo frames show better quality due to its higher correlation among all the video sequences.

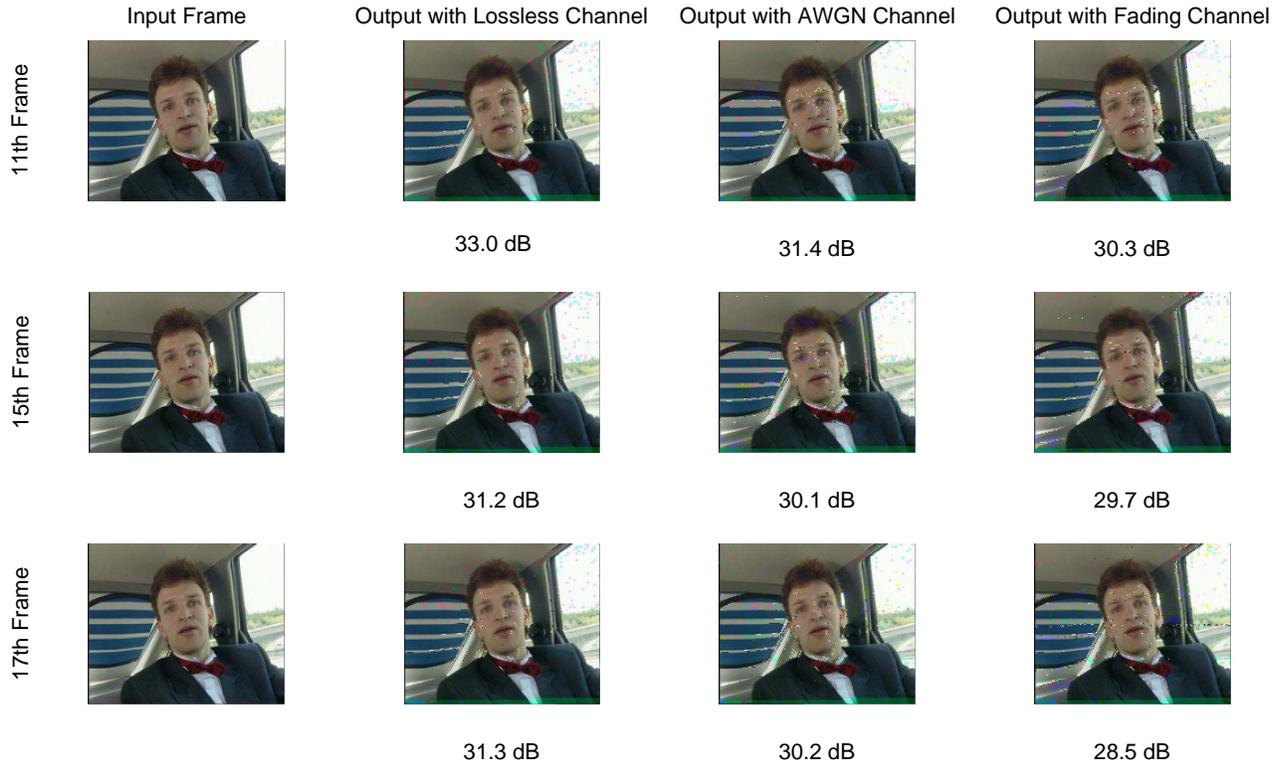


Figure 3.9: Reconstructed Carphone Video Frames with Lossless, AGWN, and Fading Channels at  $SNR = 2$  dB

Generally, the overall quality of the video frame in the AWGN channel and fading channel will be low compared to that of the noiseless channel. The reconstruction of the WZ frame depends on  $Y$  and  $Z$ , i.e.,  $f(Y, Z)$ . Akio video sequence performs better than the other two due to  $Y$  of Akio shows less distortion, i.e., correlation of Akio frames are high.

Fig. 3.11 shows the PSNR performance of the reconstructed WZ frames of the Foreman video sequence with different channel SNR values; PSNR values are improved for higher SNR values. We have selected three frames from the reconstructed Foreman sequence, as shown

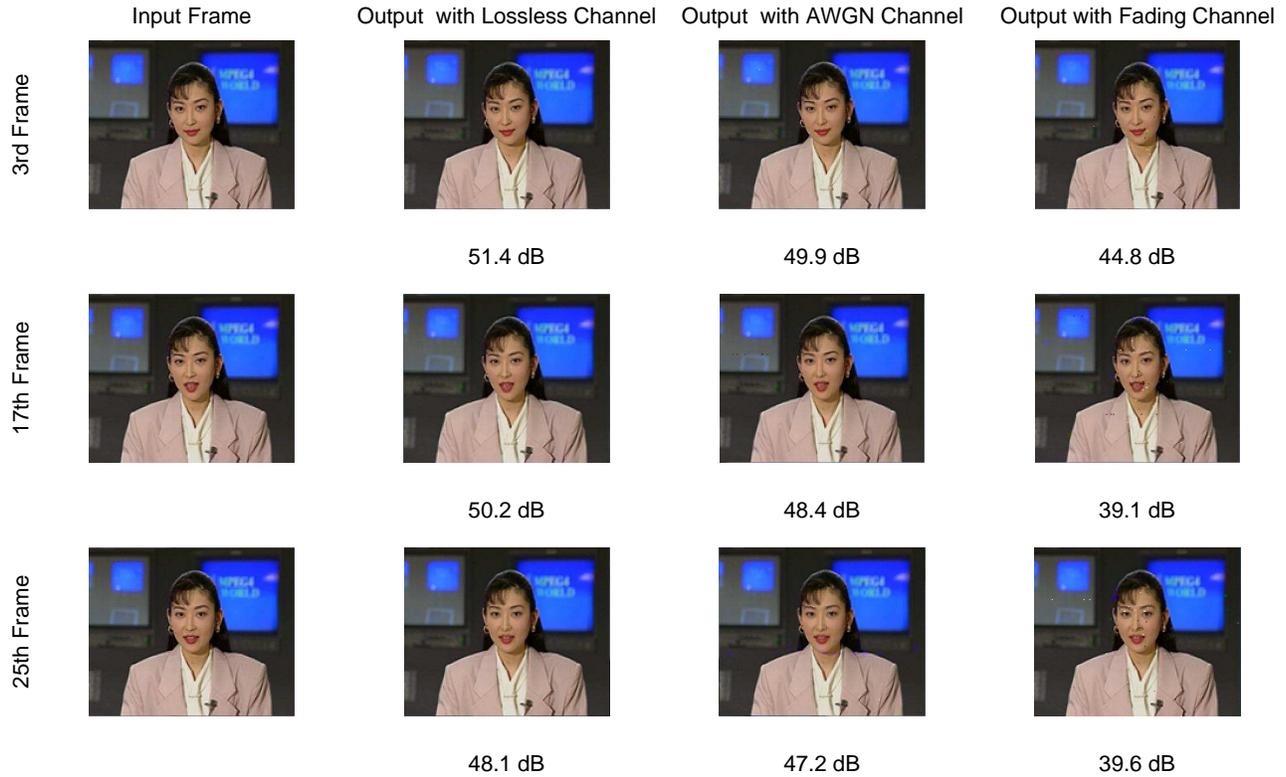


Figure 3.10: Reconstructed Akiyo Video Frames with Lossless, AWGN, and Fading Channels at  $SNR = 2$  dB

in Fig. 3.12, where they show an improvement when the channel SNR was increased from 2 dB to 10 dB. However, 9<sup>th</sup> and 67<sup>th</sup> frames showed better quality compared to 23<sup>rd</sup> frame. The frames with lower cross correlations will not exhibit improvement in quality with the increasing channel SNR due to the faster motion which requires more parity bits in order to reconstruct the WZ video frame at the decoder. In this simulation, the transmitting data rate is fixed as 190 kbps. In Chapter 5, we investigate the PSNR-rate relationship in terms of cross-correlation behavior of the video frames.

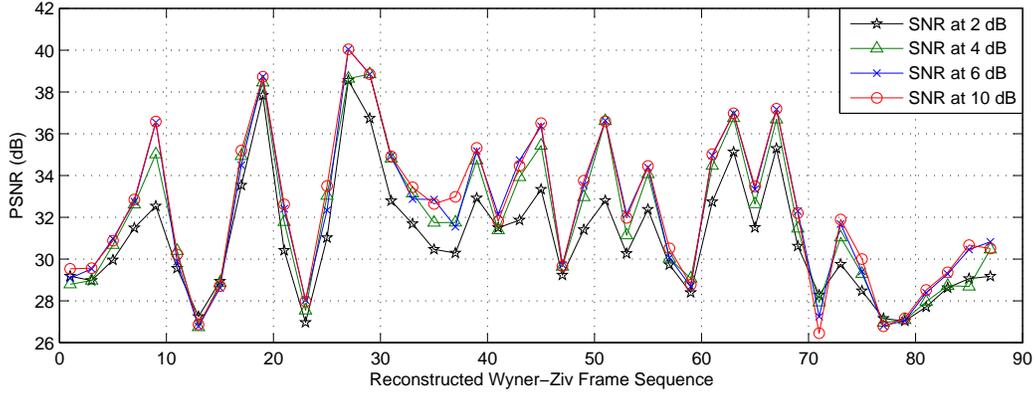


Figure 3.11: PSNR Performance for AWGN Channel with Different Channel SNRs for Foreman Video Sequence

Fig. 3.13 shows the PSNR performance of the reconstructed Carphone video sequence with the different channel SNR values. The PSNRs of the Carphone frames have improved with the channel SNR values. In the Carphone sequence, frames show reasonable cross-correlations, therefore it also shows improved PSNR in the simulated sequence.

The comparison of the reconstructed Akio, Carphone and Foreman video sequences are shown in Fig. 3.14. Here, we consider the average PSNR value of eighty seven WZ frames. In all three video sequences, average PSNR exhibits similar performance with increased SNR values; here, it is due to  $L_c^1$  which is improved with higher channel SNR. The PSNR value becomes saturated after a certain channel SNR. In Fig. 3.14(a),  $PSNR = 32.4$  dB is at  $SNR = 12$  dB for the Foreman video sequence, whereas  $PSNR = 32.6$  dB is at  $SNR = 10$  dB in Carphone video sequence. The reconstructed Akio video sequence shows  $PSNR = 48.7$

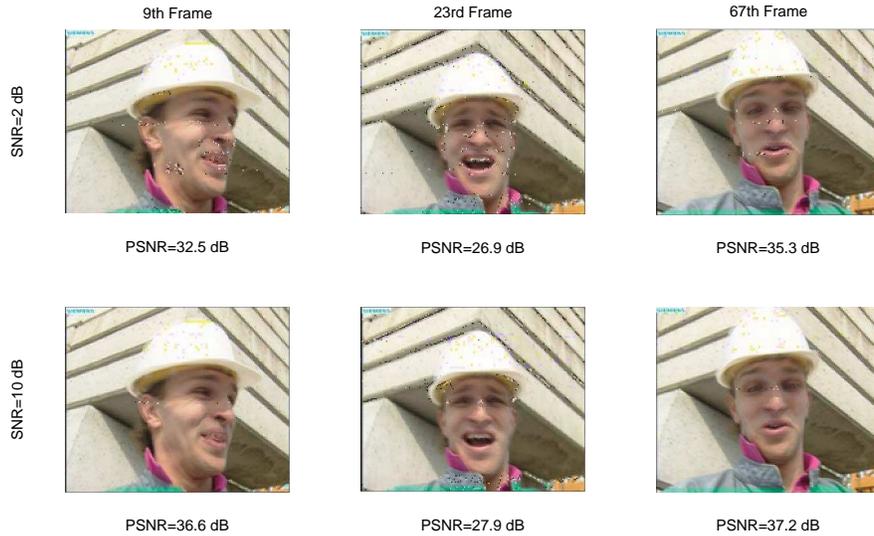


Figure 3.12: Some Selected Reconstructed Wyner-Ziv Frames from Fig. 3.11

dB after  $SNR = 6$  dB as shown in Fig. 3.14(b).

### 3.7.2 Improvement of WZVC with Wireless Channel Environment

Fig. 3.15 shows the average PSNR performance of the reconstructed WZ frames of Foreman and Carphone in wireless fading channel. Reconstructed video frames over the wireless fading channel show heavier distortion than that with only AWGN channel alone, in both video sequences. This is due to the fading impact on the transmitted parity of WZ frame. In wireless environment, PSNR will be improved only with the higher channel  $SNR(= 20$  dB). For instance, at  $SNR = 2$  dB, PSNR is decreased by 3 dB in the fading channel compared to lossless channel for the Foreman video, 2 dB for the Carphone. However in Foreman video sequence, 2.7 dB improvement in PSNR is observed only by increasing SNR by 18 dB.

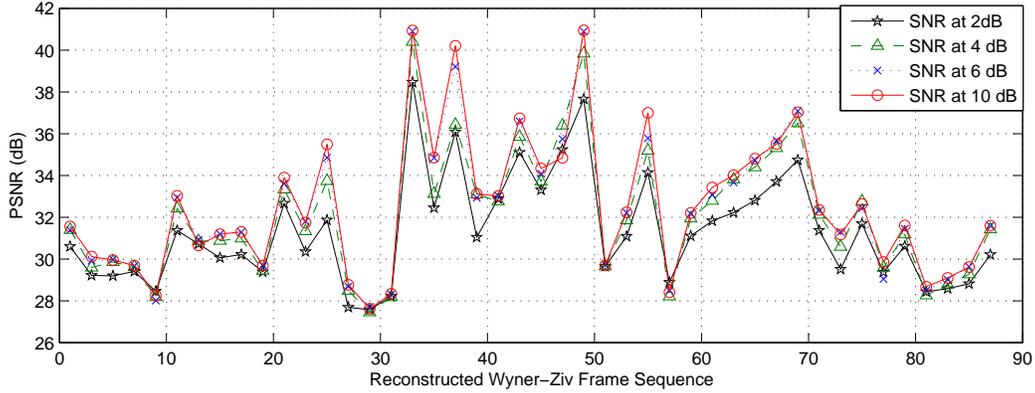
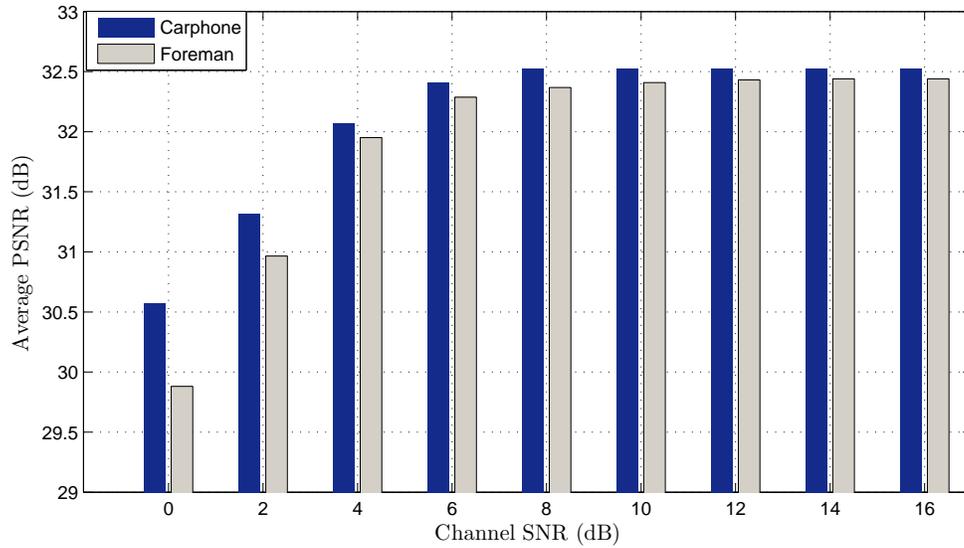
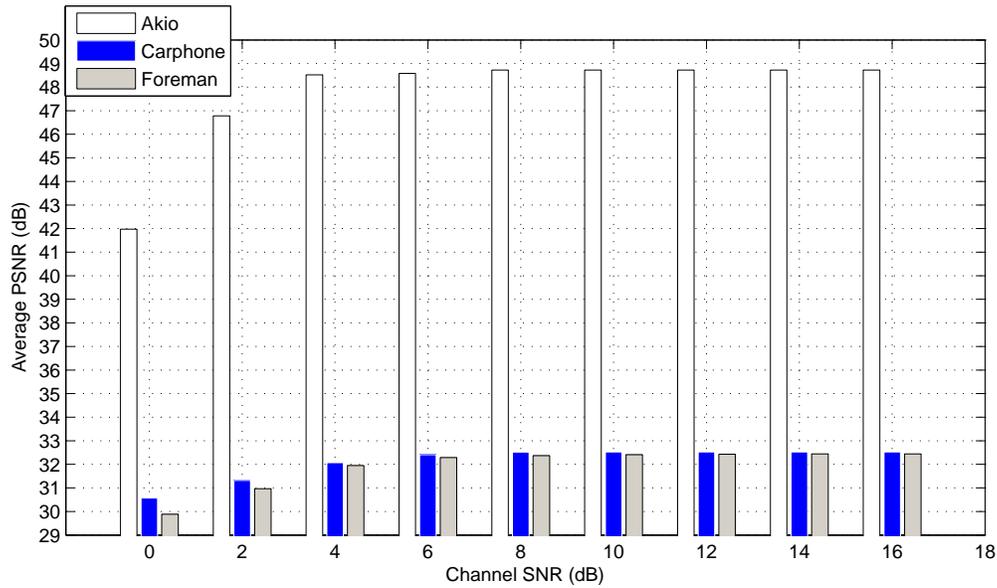


Figure 3.13: PSNR Performance for AWGN Channel with Different channel SNRs for Car-phone Video Sequence

PSNR performance vs. channel SNR is shown in Fig. 3.17; here, a single wireless channel is compared to 2I2O and 2I4O wireless channel for the Foreman video sequence. The simulation result shows that the PSNR performance was significantly improved for the configurations that uses two transmitter (2 Tx) two receiver (2 Rx) antennas and the two transmitter (2 Tx) four receiver (4 Rx) antennas compared to the single input single output wireless channel (SIS0). For  $SNR \geq 5$  dB, PSNR performance of the 2I4O channel reaches to the noiseless channel. Furthermore, the 2I4O channel saves 1.1 dB in PSNR compared to the 2I2O channel at  $SNR = 2$  dB. At  $PSNR \approx 32.3$  dB, the 2I20 and 2I4O save SNR of  $\approx 15$  dB and  $\approx 19$  dB, respectively over a single fading channel.



(a) For Carphone and Foreman



(b) For Akio, Carphone and Foreman

Figure 3.14: Comparison of Average PSNR Values for AWGN Channel with the Different Channel SNRs

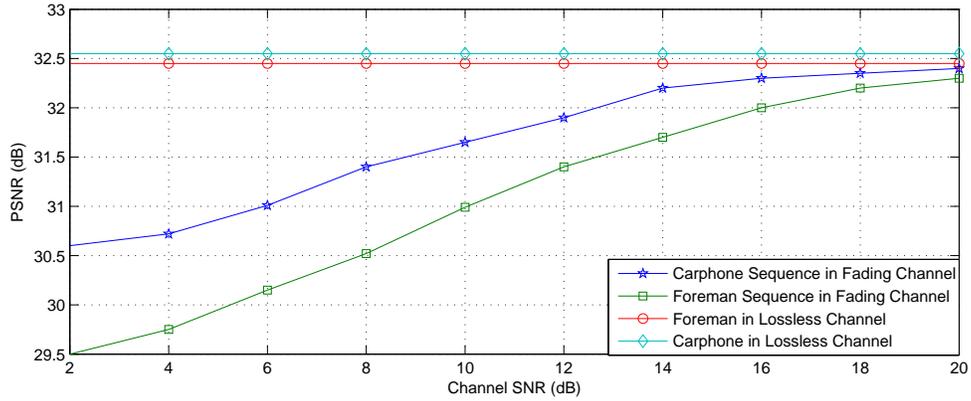


Figure 3.15: PSNR Performance for Foreman and Carphone in Wireless Fading Channels with Different SNRs at Rayleigh Variance 0.5

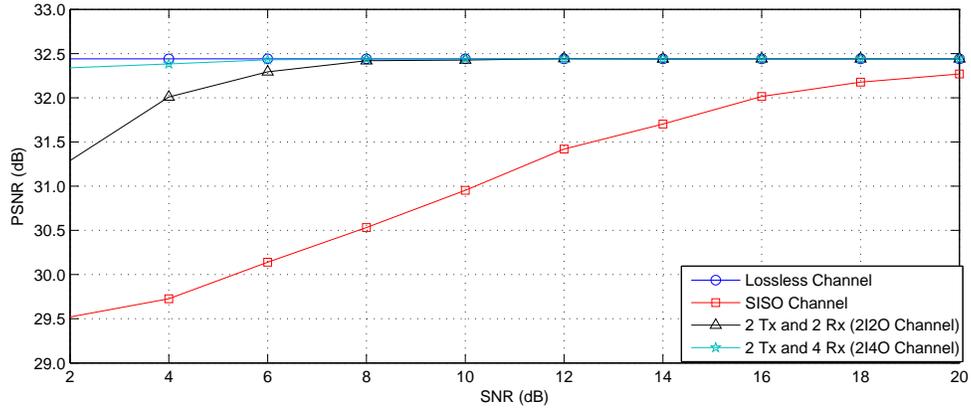


Figure 3.16: PSNR Performance of the Reconstructed WZ Frames over 2IMO Wireless Channel for Foreman Video at Rayleigh Variance 0.5

### 3.8 Summary

In this chapter, we have considered a practical WZVC with the effects of the wireless fading channel and AWGN. At the decoder, channel-metric values  $L_c^1$  and  $L_{cMIMO}^1$  are investigated with respect to the type of channel. The bit-metric ( $L_c$ ) value is computed based on the variance of correlation noise between the WZ frame and side information.

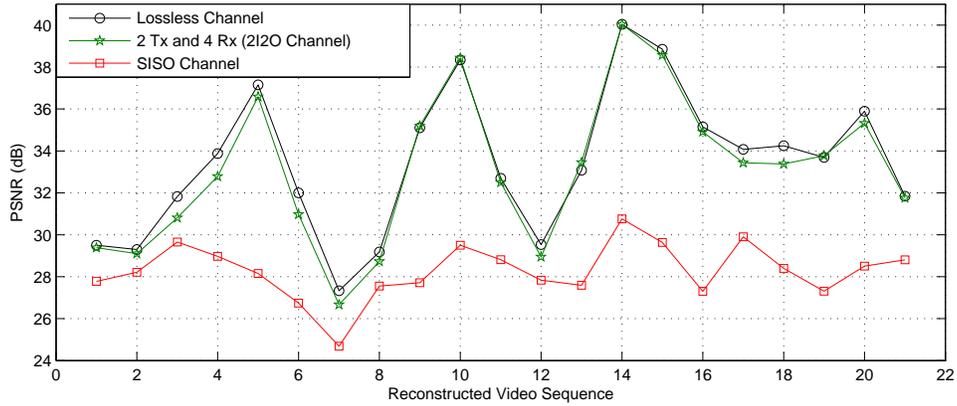


Figure 3.17: PSNR Performance of the Reconstructed WZ Frames over 2IMO Wireless Channel for Foreman Video at Rayleigh Variance 0.5

We thoroughly analyzed the quality of the reconstructed/received video with the *channel* SNR. The simulation results show the relationship between the channel SNR and PSNR, which will be significant to future research, especially for practical implementations of WZVC in wireless channels. Since PSNR values in a wireless and AWGN channel are lower than the noiseless channel, this chapter has further discussed the improvement of PSNR in a wireless fading channel environment.

Higher decoder complexity with a low encoder complexity is the primary motivation behind the development of WZVC. We have investigated WZVC over a multiple transmit-receive diversity in a wireless system, which improves the reconstructed video frames compared to single wireless channel. In 2 Tx and 4 Rx antennas, PSNR value reach to PSNR

of noiseless channel (32.3 dB) for the lower SNR values. Hence, WZVC over 2IMO channel performs better compared to SISO, i.e., 2IMO saves energy at the encoder side because it performs good at lower SNR. Therefore WZVC has the potential to use the transmit-receive diversity technique to overcome wireless channel impairment.

# Chapter 4

## Rate Penalty in Wireless Wyner-Ziv Video Coding

### 4.1 Introduction

Several algorithms have been proposed for video transmission based on the Wyner-Ziv and Slepian-Wolf theorems, which were discussed in Chapter 3 (Section 3.1). Let the average rate in Wyner-Ziv coding via noiseless channel is given as [71];

$$\hat{R}(d) = I(X; Z) - I(Y; Z) \quad (4.1)$$

where,  $Z$  is residual information between WZ signal  $X$  and side information  $Y$ . In Chapter 3, the encoder of WZVC was set to transmit parity bits to the decoder to reconstruct WZ video frame  $X$  ( $X'$ ) from the side information video frame  $Y$ , i.e., parity bits are transmitted from a code book of  $Z$ . Higher compression is obtained by sending smaller number of bits from  $Z$ . The compression ratio can approach the Slepian-Wolf bound when Turbo codes [13]

or low-density parity check coding (LDPC) are used [63].

Most of the previous research in this field assumed an ideal channel setup in WZVC; only a few researchers have started taking into account the channel impairments. The robustness of transmission errors was investigated by Puri *et al.* in [9] for the side information that is influenced by channel noise. They have shown that the transmission rate depends on both the correlation noise and the channel noise. However, *fading* also occurs in wireless channels due to multi-path propagation. The combined effect of channel noise and fading on  $Z$  is very detrimental; and results in a requirement of a higher rate of  $Z$  to achieve the same quality. In other words, a noisy fading channel can effectively reduce the transmission rate; this is termed *rate penalty*.

In this chapter, we investigate the required transmission rate of WZVC in noisy fading wireless channels. We analytically derived rate penalty of WZVC in AWGN channel and fading channel. Additionally, WZVC in a fading channel has higher rate penalty than even the AWGN channel. This chapter proposes WZVC with receiver diversity (WZVC-RD) as a provision to reduce the rate penalty in WZVC over the wireless fading channel.

This chapter is organized as follows: Section 4.2 investigates the average rate in WZC, and

derives rate penalty in WZC over the wireless channel. In Section 4.3, the simulation model of WZVC-RD is proposed, and Section 4.4 exhibits rate analysis in WZVC-RD. Finally, Section 4.5 explains the results, and the summary is drawn in Section 4.6.

## 4.2 Average Rate of WZ Coding in Wireless Channel

### 4.2.1 WZC in Noiseless Channel

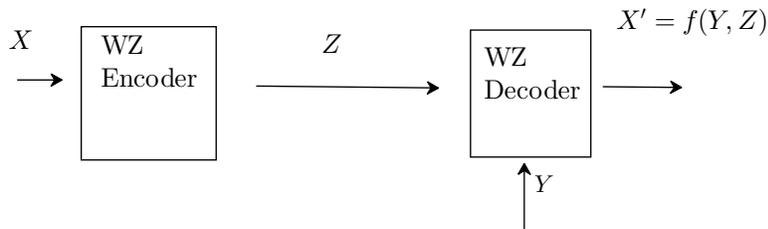


Figure 4.1: Standard WZ Codec with Noiseless Channel

Fig. 4.1 shows the traditional WZ Codec with a noise-free channel. The reconstructed vector  $X'$  depends on  $Y$  and  $Z$ , i.e.,  $X' = f(Y, Z) = aY + Z$ , as discussed in [71]; here the average rate-distortion is shown as  $\hat{R}(d) = I(X; Z) - I(Y; Z)$ , where distortion  $d$  is given as  $d = E(D(X, X'))$ .

## 4.2.2 WZC in Noisy Fading Channel

In Fig. 4.2, we incorporated the impact of noise and fading on  $Z$ . Noise and fading are considered as white Gaussian and Rayleigh channels, respectively.

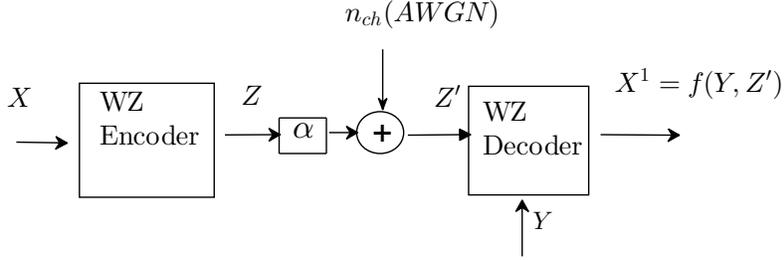


Figure 4.2: Standard WZ Codec with AWGN Noise and Wireless Channel Fading

### 4.2.2.1 WZC in AWGN Channel

First we consider channel noise  $n_{ch}$ . The received residual information is given as  $Z'_n = Z + n_{ch}$ , and the average rate  $\hat{R}_{AWGN}(d)$ , is rewritten by referring to (4.1) as;

$$\hat{R}_{AWGN}(d) = I(X; Z) - I(Y; Z'_n). \quad (4.2)$$

From (4.1) and (4.2), we reach at;

$$\hat{R}_{AWGN}(d) - \hat{R}(d) = I(Y, Z) - I(Y, Z'_n), \quad (4.3)$$

$$= (H(Y) - H(Y/Z)) - (H(Y) - H(Y/Z'_n)),$$

$$= H(Y/Z'_n) - H(Y/Z). \quad (4.4)$$

In an AWGN channel;

$$\begin{aligned} I(Z, Z'_n) &= H(Z) - H(Z'_n/Z) = H(Z) - H(n_{ch}), \\ &= \frac{1}{2} \log_2 \left( 1 + \frac{\sigma_Z^2}{\sigma_n^2} \right) < H(Z). \end{aligned} \quad (4.5)$$

Therefore from (4.4), we can see that;  $H(Y/Z'_n) > H(Y/Z)$  and,

$$H(Y/Z'_n) = H(Y/Z) + \epsilon_{AWGN} \quad (4.6)$$

where,  $\epsilon_{AWGN}$  defines the *rate penalty* in AWGN channel, and  $\epsilon_{AWGN} = \hat{R}_{AWGN}(d) - \hat{R}(d) > 0$ . On the other hand  $\hat{R}_{AWGN}(d)$  is shown as;

$$\hat{R}_{AWGN}(d) = \hat{R}(d) + \epsilon_{AWGN}. \quad (4.7)$$

#### 4.2.2.2 WZC in Noisy Fading Channel

Now we consider the fading channel with Gaussian noise; here the received residual information can be written as;

$$Z'_F = \alpha Z + n_{ch} \quad (4.8)$$

where,  $|\alpha| \leq 1$  is the normalized fading coefficient. The average rate over the fading channel is set by referring to (4.1);

$$\hat{R}_{Fading}(d) = I(X; Z) - I(Y; Z'_F) \quad (4.9)$$

From (4.1) and (4.9);

$$\hat{R}_{Fading}(d) - \hat{R}(d) = I(Y, Z) - I(Y, Z'_F) = H(Y/Z'_F) - H(Y/Z). \quad (4.10)$$

However;

$$I(Z, Z'_F) = \frac{1}{2} \log_2 \left( 1 + \frac{\alpha^2 \sigma_Z^2}{\sigma_n^2} \right). \quad (4.11)$$

Therefore from (4.5) and (4.11);

$$I(Z, Z'_F) \leq I(Z, Z'_n) < H(Z). \quad (4.12)$$

Since  $H(Y/Z'_F) > H(Y/Z'_n) > H(Y/Z)$ ;

$$H(Y/Z'_F) = H(Y/Z) + \epsilon_{Fading}. \quad (4.13)$$

Thus from (4.10) and (4.13);

$$\hat{R}_{Fading}(d) - \hat{R}(d) = \epsilon_{Fading} \quad (4.14)$$

where, the rate penalty in fading channel is defined as  $\epsilon_{Fading}$ .

Let us compare  $\epsilon_{Fading}$  and  $\epsilon_{AWGN}$  from (4.2) and (4.9);

$$\begin{aligned}\hat{R}_{Fading}(d) - \hat{R}_{AWGN}(d) &= I(Y, Z'_n) - I(Y, Z'_F) \\ &= H(Y/Z'_F) - H(Y/Z'_n)\end{aligned}\tag{4.15}$$

$$\begin{aligned}&= [H(Y/Z'_F) - H(Y/Z)] - [H(Y/Z'_n) - H(Y/Z)] \\ &= \epsilon_{Fading} - \epsilon_{AWGN} = \epsilon\end{aligned}\tag{4.16}$$

$$\hat{R}_{Fading}(d) = \hat{R}_{AWGN}(d) + \epsilon\tag{4.17}$$

where,

$$\epsilon = \begin{cases} > 0, & \text{if } |\alpha| < 1 \\ 0, & |\alpha| = 1. \end{cases}$$

If  $\epsilon = H(Y/Z'_F) - H(Y/Z'_n) > 0$ , then

$$\epsilon_{Fading} = H(Y/Z'_F) - H(Y/Z) > \epsilon_{AWGN}.\tag{4.18}$$

Therefore, from the equations (4.7), (4.14) and (4.17),

$$\hat{R}_{Fading}(d) \geq \hat{R}_{AWGN}(d) > \hat{R}(d).\tag{4.19}$$

This shows that the rate penalty of WZVC in fading channel,  $\epsilon_{Fading}$ , is higher than that in the AWGN channel;  $\epsilon_{AWGN}$ , which in turn is higher than that of the noiseless channel.

Next, Section 4.3 explains our proposed receiver diversity model which improves the rate

penalty in fading channels.

### 4.3 Proposed WZVC with Receiver Diversity (WZVC-RD)

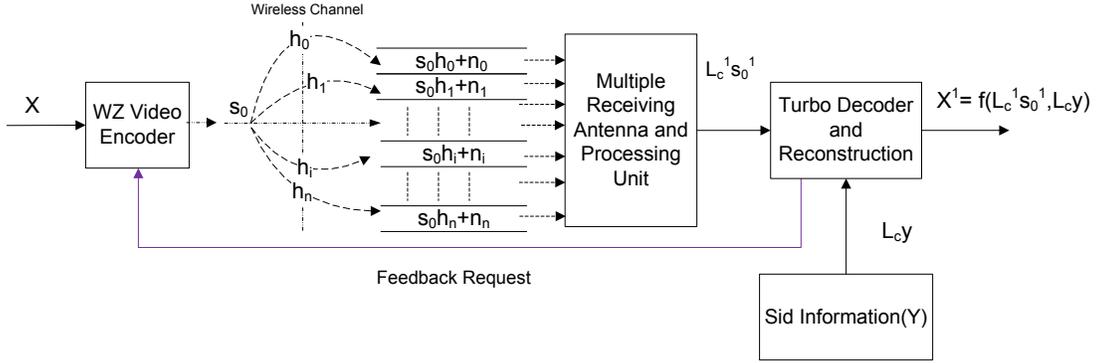


Figure 4.3: WZVC with an Improved Decoder

This section explains the proposed WZVC-RD model. Here, an encoder transmits parity signal  $s_0$ , and multiple uncorrelated receiving points receive the multipath signals at the receiver, as shown in Fig. 4.3. The received signal in  $i^{th}$  receiving point is denoted as  $r_i = s_0 h_i + n_i$ , where  $i = 0, 1, 2, \dots, n$ .

The received signals can be shown as below, for compactness;

$$\begin{pmatrix} r_0 \\ \dots \\ r_n \end{pmatrix} = \begin{pmatrix} h_0 \\ \dots \\ h_n \end{pmatrix} s_0 + \begin{pmatrix} n_0 \\ \dots \\ n_n \end{pmatrix} \quad (4.20)$$

where,  $n$  is the number of uncorrelated received channels,  $n_i$  is AWGN channel noise,

$h_i (= |h_i|e^{j\theta})$  is the channel transfer function, and  $|h_i|$  is the Rayleigh fading coefficient.

The receiver in Fig. 4.3 computes optimum parity ( $s_0^1$ ), channel-metric value ( $L_c^1$ ) and bit-metric value ( $L_c$ ). The  $s_0^1$  is computed using the matrix multiplication of the received vector  $(r_0 \ r_1 \ \dots \ r_n)^T$  and the row vector  $(h_0^* \ h_1^* \ \dots \ h_n^*)$ , where  $h_i^*$  is the conjugate of the  $i^{th}$  transpose element.

$$s_0^1 = \sum_{i=0}^{n-1} (h_i^2) s_0 + \sum_{i=0}^{n-1} (h_i^* n_i) \quad (4.21)$$

If  $\alpha_D = \sum_{i=0}^{n-1} h_i^2$  and  $n_{ch} = \sum_{i=0}^{n-1} (h_i^* n_i)$ , then (4.21) can be simplified as;

$$s_0^1 = \alpha_D s_0 + n_{ch}. \quad (4.22)$$

Maximum a posteriori probability (MAP) algorithm in the Turbo decoder depends on the  $s_0^1$ ,  $y$ ,  $L_c^1$  and  $L_c$ , i.e., the reconstructed video frame  $X^1$ ,  $X^1 \varepsilon \{y, L_c, s_0^1, L_c^1\}$ , is comparable to  $f(Y, Z) = aY + Z$  in [71], where  $Y \varepsilon \{L_c, y\}$ ,  $Z \varepsilon \{L_c^1, s_0^1\}$ .

The channel-metric ( $L_c^1$ ) and bit-metric ( $L_c$ ) values are derived in Chapter 3, which are used to calculate the received soft-values of parity and systematic information in the Turbo decoder respectively.

$L_c^1$  is given below;

$$L_c^1 = \alpha \frac{4E_s^p}{N_{ch}} = \alpha_D \frac{2E_s^p}{\sigma_{ch}^2} = \sum_{i=0}^{n-1} (h_i^2) \frac{2E_s^p}{\sigma_{ch}^2} \quad (4.23)$$

where,  $\sigma_{ch}^2$  is the variance of  $n_{ch}$ . Since the channel coefficient  $h_i$  and noise  $n_i$  are independent random variables in (4.21), noise variance  $\sigma_{ch}^2$  is given as;

$$\sigma_{ch}^2 = 2 \sum_{i=0}^{n-1} \sigma_{h_i}^2 \sigma_{n_i}^2. \quad (4.24)$$

This enables equation (4.23) to be rewritten as;

$$L_c^1 = \alpha_D \frac{E_s^p}{\sum_{i=0}^{n-1} \sigma_{h_i}^2 \sigma_{n_i}^2}. \quad (4.25)$$

The bit-metric value  $L_c$  can be shown as  $L_c = \frac{2E_s}{\sigma_{pic}^2}$ .  $E_s$  denotes the energy of symbols from the side information, and  $N_{pic}$  is the power spectral density of  $n_{pic}$ , where  $\sigma_{pic}^2$  is the pixel variance that is estimated from the key frames at the decoder [72]. The soft-value of side information can be derived as  $\frac{2yE_s}{\sigma_{pic}^2}$ ; it is used as the received systematic information at the decoder. Here,  $y$  is the bit sequence from the bit-plane of the side information [73]

## 4.4 Rate at WZVC-RD

In the receiver diversity scheme, the received residual information is given as  $Z'_D = \alpha_D Z + n_{ch}$ .

The resultant gain,  $\alpha_D^2$ , is given as  $\alpha_D^2 = \sum_{i=0}^n (\alpha_i^2)$ , where  $\alpha_D^2 \geq \alpha_i^2$ .

Therefore,

$$\frac{1}{2} \log_2 \left( 1 + \frac{\alpha_D^2 \sigma_Z^2}{\sigma_n^2} \right) \geq \frac{1}{2} \log_2 \left( 1 + \frac{\alpha^2 \sigma_Z^2}{\sigma_n^2} \right)$$

$$I(Z, Z'_D) \geq I(Z, Z'_F), \quad (4.26)$$

$$H(Y/Z'_D) - H(Y/Z) < H(Y/Z'_F) - H(Y/Z). \quad (4.27)$$

Hence, the required rate in the receiver diversity scheme  $\hat{R}_{Diversity}(d)$  is lower than the required rate  $\hat{R}_{Fading}(d)$  in single fading channel;

$$\hat{R}_{Diversity}(d) < \hat{R}_{Fading}(d). \quad (4.28)$$

Also the rate penalty in wireless channel is reduced by the diversity scheme;

$$\hat{R}_{Diversity}(d) - \hat{R}(d) < \hat{R}_{Fading}(d) - \hat{R}(d). \quad (4.29)$$

## 4.5 Results

Fig. 4.4 shows performance of PSNR vs. number of receiving points; here, different channel SNRs are used, and PSNR is averaged over 44 Foreman video frames. It is clear that the PSNR performance can be improved by increasing the number of receiving points for the tested SNR values: 2 dB, 4 dB, 6 dB, 8 dB and 10 dB. An improvement in PSNR is observed with higher receiving points at lower SNRs. Thus the receiver diversity is especially significant for enhancing the output video quality for a poor quality channel. For example,

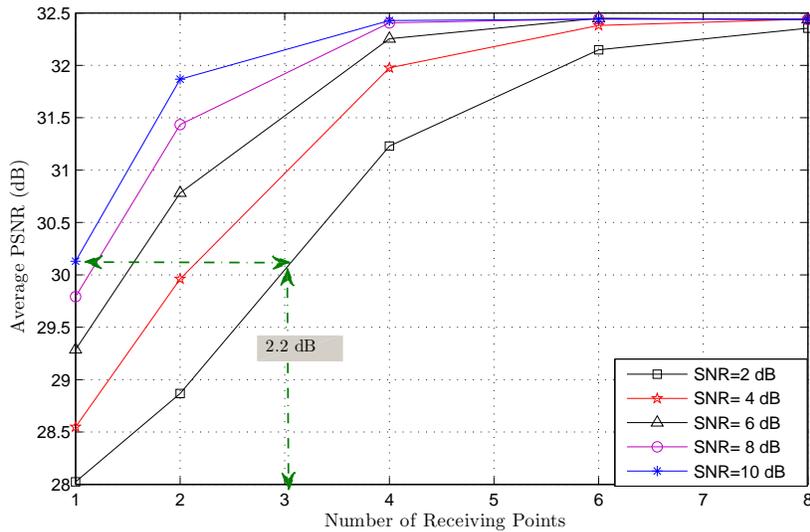


Figure 4.4: Average PSNR Variation with Rate and SNR for Foreman Video

three receiving points at  $SNR = 2$  dB perform at  $PSNR = 30.2$  dB; this same PSNR value can be reached with one receiving point, only at SNR of 10 dB. In addition, a 4.4 dB improvement in PSNR is observed for eight receiving points compared to the single receiving point at  $SNR = 2$  dB.

Fig. 4.5 shows PSNR performance with respect to the bit rate; three different types of channels are used in this simulation: noiseless, AWGN, and fading. PSNR shows an improvement for higher bit rates. However, the PSNR values decrease for the fading channel compared to the AWGN channel, which in turn, is lower than the noiseless channel. For instance, the rate penalty for AWGN channel over noiseless channel is 200 Kbps at  $PSNR = 32$  dB and the rate penalty for fading channel over the AWGN channel is 180 Kbps at  $PSNR = 31$  dB. The rate penalty for fading channel compared to the noiseless

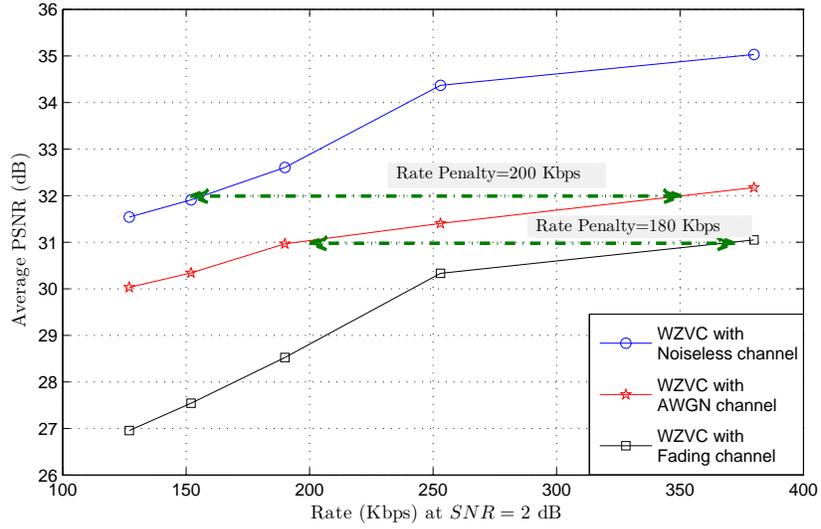


Figure 4.5: Average PSNR Variation with Rate for Foreman Video

channel is more than 200 Kbps at  $PSNR = 32$  dB.

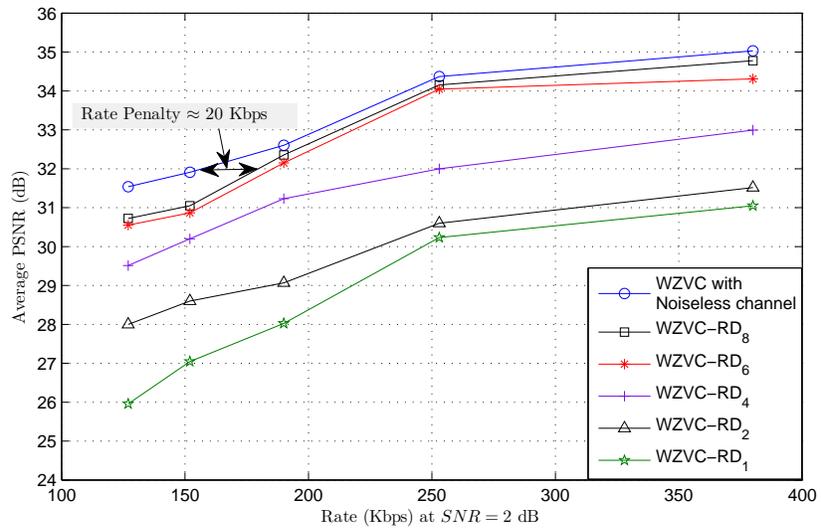


Figure 4.6: Average PSNR Variation with Rate for Foreman Video

Fig. 4.6 shows PSNR vs. bit rate for different numbers ( $n$ ) of receiver diversity (WZVC-RD $_n$ ,  $n = 1, 2, \dots, 8$ ) at channel  $SNR = 2$  dB. As expected, PSNR-rate relationship is

improved in WZVC-RD<sub>*n*</sub> for  $n > 1$ . For example, 127 Kbps shows a 4.9 dB PSNR gain for eight receiving points (WZVC-RD<sub>8</sub>) compared to a single receiving point (WZVC-RD<sub>1</sub>), whereas 380 Kbps shows 4.0 dB PSNR gain.

The single receiving point (WZVC-RD<sub>1</sub>) shows  $PSNR = 31$  dB at a bit rate of 380 Kbps. The same PSNR is obtained with two (WZVC-RD<sub>2</sub>), four (WZVC-RD<sub>4</sub>), six (WZVC-RD<sub>6</sub>), and eight (WZVC-RD<sub>8</sub>) receiving points at the rates of 200 Kbps, 185 Kbps, 155 Kbps and 150 Kbps, respectively. By comparing with WZVC-RD<sub>1</sub>, encoder preserves 230 Kbps with the help of WZVC-RD<sub>8</sub>, at  $PSNR = 31$  dB. Moreover, the rate penalty is decreased in WZVC-RD<sub>*n*</sub>, for  $n > 1$ . For example, the rate penalty in WZVC-RD<sub>8</sub> is reduced to 20 Kbps at  $PSNR = 32$  dB, and PSNR in WZVC-RD<sub>8</sub> also observed to approach that of WZVC with noiseless channel.

## 4.6 Summary

In this chapter, we have analytically derived expressions for the rate penalty of WZVC in AWGN ( $\epsilon_{AWGN}$ ) and fading channels ( $\epsilon_{Fading}$ ), and showed that it increases with the channel inefficiency. Furthermore, we have proposed WZVC with receiver diversity (WZVC-RD) which can reduce the impact of rate penalty. Simulation results show that the rate penalty decreases with an increase in the number of receiving points for a given channel

SNR. Moreover, our proposed scheme gives a higher PSNR at lower bit rates of the encoder. This reduces the encoder's energy consumption, thus the WZVC-RD setup is an appropriate technique for low power video encoders with higher power decoders.

# Chapter 5

## Cross-Correlation based Rate Reduction Technique for Wyner-Ziv Video Coding

### 5.1 Introduction

Several algorithms are proposed for sending syndrome or parity bits to the receiver to reconstruct the video information by moving the complexity from encoder to decoder side [7], [66], [62], [4], [63], [61]. The compression ratio can be approximate to the Slepian-Wolf bound using Turbo codes shown in [4], also the low-density parity check code (LDPC) is considered to be another form of iterative channel coding [63]. In [66], Majumdar and Ramachandran represented a radical departure from state-of-the-art video coding architectures.

WZVC can be classified as transform domain coding or pixel domain coding. In many WZVC schemes, error control codes are applied to each frame to generate syndromes or

parity bits. The compression is achieved by sending a fraction of the parity bits (punctured from the original parity) to the decoder. The decoder uses the received parity to correct errors in the side information ( $Y$ ) for reconstructing the WZ frame, where side information is known as the noisy version of the WZ frame ( $X$ ):  $Y = X + N$ . Since the most challenging task in encoding is the rate allocation, a feedback channel is employed in WZVC to control the number of parity bits requested by the decoder [17], [18]. The encoder contains a buffer that stores parity bits; the decoder uses an error-probability threshold to determine whether the available parity is enough for decoding, and requests additional bits through feedback channel for successful decoding with low distortion. This process is repeated until the error probability falls below the threshold.

The major drawback of implementing a feedback channel is its' latency, which is due to the continual error-probability calculation and the process of reporting to the encoder. A few WZVC research papers suggest using rate allocation algorithms from the encoder side, which eliminates the requirement of a feedback channel [19], [20], [21]. However, in such a setup, the encoder complexity is increased considerably, deviating from the concept of complexity balance in WZVC. The transmission rate depends on the sum of the correlation noise and channel noise [9]. Previously, we have discussed rate control with respect to the channel impairment in [74].

This chapter provides an encoder rate control mechanism at the decoder with respect to the cross-correlation threshold (CCTH) of the available key frames. In the first portion of this chapter, we investigate the behavior of the cross-correlation statistics of available frames (key frames) at the decoder to analyze the rate-distortion behavior. We also analyze the reconstructed function with respect to the side information ( $Y$ ) and residual information ( $Z$ ).  $Z$  is derived with the relation to the cross-correlation coefficient of  $X$  and  $Y$ ;  $\rho_{XY}$ . In addition, the theoretical conditional rate is derived with respect to  $\rho_{XY}$  for Gaussian vectors  $X$  and  $Y$ . Furthermore, we compare the theoretical conditional rate vs.  $\rho_{XY}$  with rate-distortion vs. cross-correlation values of the key frames.

The rest of the chapter proposes an encoder rate control algorithm with respect to the cross-correlation values of the key frames. The cross-correlation value of the frame index determines the required rate of parity bits of the expected WZ frame. If the cross-correlation value of the expected index of WZ frame is higher, then lower bit rates are sufficient to decode, and vice versa. Therefore, cross-correlation thresholds in terms of the statistics of cross-correlation points must be defined. These thresholds determine the required number of parity bits from the encoder. We also propose an adaptive threshold algorithm (ATHA), which classifies the number of higher/lower threshold points. So that the decoder can de-

termine the required rate based on ATHA. The determination of the rate can be used for a specific number of WZ frames according to the cross-correlation behavior of the available frames at the decoder. Therefore, the proposed ATHA method reduces the frequent usage of the feedback channel.

This chapter is organized as follows: Section 5.2 analyzes conditional rate of WZVC with respect to  $\rho_{XY}$ . Section 5.3 explains how rate corresponds to the cross-correlation values of the key frames in practical WZVC. Significance of CCTHs related to rate control is briefed in Section 5.4. In Section 5.5, the definition of threshold and its derivations are elaborated, Section 5.6 presents the rate control algorithm with respect to the proposed CCTHs. Finally, Section 5.7 and 5.8 show the results of our work and conclusions, respectively.

## 5.2 Conditional Rate Distortion

Both the Slepian-Wolf and Wyner-Ziv theorems are the key principles of the Wyner-Ziv video coding (WZVC), which allows for lower complex encoders at the expense of higher decoder complexity. The Slepian-Wolf theorem refers to lossless compression while the Wyner-Ziv theorem refers to lossy compression with side information available at the decoder. The Slepian-Wolf theorem states that two statistically dependent signals,  $X$  and  $Y$ , can be separately encoded and then jointly decoded with an arbitrarily small error probability. The achievable rate region satisfies the conditions;  $R_X \geq H_{X|Y}$ ,  $R_Y \geq H_{Y|X}$ , and

$R_X + R_Y \geq H(X, Y)$ , where  $R_X$  and  $R_Y$  are the rate-distortions of  $X$  and  $Y$ ,  $H(X, Y)$  is the joint entropy of  $X$  and  $Y$ , and  $H_{X|Y}$  and  $H_{Y|X}$  are conditional entropies [3]. The Wyner-Ziv theorem states that by assuming  $X$  and  $Y$  to be statistically dependent Gaussian random processes, and  $Y$  as the side information for encoding  $X$ , the rate-mean squared error distortion function for  $X$  is the same whether the side information ( $Y$ ) is available only at the decoder, or both at the encoder and the decoder.

This section briefly discusses the definition of the rate in the Wyner-Ziv coding. Here, we show the relationship between minimum required rate,  $R_{X/Y}$ , and cross-correlation coefficient of  $X$  and  $Y$ ;  $\rho_{XY}$ . Further we have derived the minimum distortion with respect to the  $\rho_{XY}$  value and verified with the rate  $R_{X/Y}$ .

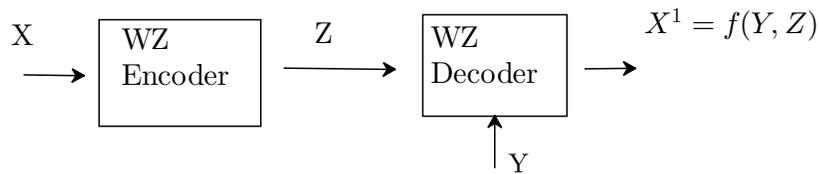


Figure 5.1: Wyner-Ziv Codec

Fig. 5.1 shows the block diagram of Wyner-Ziv codec; here,  $X$  is the input vector,

$Y$  is the side information, and  $Z$  is the residual information to be known to the decoder. The reconstructed frame  $X'$  is given as  $X' = f(Y, Z)$  [71]. The minimum required rate to reconstruct  $X$  with the knowledge of  $Y$  is known as the conditional rate,  $R_{X/Y}$ , where  $R_{X/Y} \leq R_X$  [3], [71].  $R_X$  is the conventional rate and is given as  $R_X = \frac{1}{2} \log_2 \left( \frac{\sigma_X^2}{d} \right)$  [75], where  $\sigma_X^2$  is the variance of  $X$  and  $d$  is the distortion.

Now,  $R_{X/Y}$  can be obtained from conventional rate-distortion by referring to equation (3.3) from [71], which is shown as,  $R_{X/Y} = \frac{1}{2} \log_2 \left( \frac{\sigma_{X/Y}^2}{d} \right)$ . The joint bivariate Gaussian probability density function  $f_{X,Y}(x, y)$  is given below [76];

$$f_{X,Y}(x, y) = \frac{\exp \left( \frac{-\left(\frac{x-\mu_X}{\sigma_X}\right)^2 - \frac{2\rho(x-\mu_X)(y-\mu_Y)}{\sigma_X\sigma_Y} + \left(\frac{y-\mu_Y}{\sigma_Y}\right)^2}{2(1-\rho^2)} \right)}{2\pi\sigma_X\sigma_Y\sqrt{(1-\rho^2)}} \quad (5.1)$$

where  $|\rho| \leq 1$ . From Theorems 5.18 and 5.19 in [76], the conditional variance  $Var(X/Y)$  and correlation coefficient  $\rho_{XY}$  ( $\rho_{X,Y} = \frac{Cov[X,Y]}{\sigma_X\sigma_Y}$ ,  $|\rho_{X,Y}| \leq 1$ ) are shown as  $\sigma_{X/Y}^2 = \sigma_X^2(1 - \rho_{XY}^2)$  and  $\rho_{XY} = \rho$ , respectively. Thus the  $R_{X/Y}$  can be written as;

$$R_{X/Y} = \begin{cases} \frac{1}{2} \log_2 \left( \frac{\sigma_X^2(1-\rho_{XY}^2)}{d} \right), & \text{if } d \leq \sigma_X^2(1 - \rho_{XY}^2) \\ 0, & \text{otherwise} \end{cases} \quad (5.2)$$

where,  $d$  is the distortion and it is defined as;

$$d = E((X - X')^2) \quad (5.3)$$

where, the reconstructed symbol  $X'$  depends on both the side information  $Y$  and the residual information  $Z$ . From Fig. 4 in [71],  $X'$  is denoted as;  $X' = f(Y, Z) = aY + Z$ . Equation (5.3) is expanded as;

$$d = E(X^2) + a^2E(Y^2) + Z^2 - 2aE(XY) + 2aZ(Y) - 2ZE(X) \quad (5.4)$$

where, variables,  $a$  and  $Z$ , determine the distortion level. Furthermore, the minimum distortion is derived in terms of  $a$  and  $Z$  as;

$$\left. \begin{aligned} \frac{\partial d}{\partial a} &= 2aE(Y^2) - 2E(XY) + 2ZE(Y) \\ \frac{\partial d}{\partial Z} &= 2Z + 2aE(Y) - 2E(X) \end{aligned} \right\} = 0 \quad (5.5)$$

From (5.5),  $a$  and  $Z$  can be solved as;

$$\begin{aligned} a &= \frac{E(XY) - E(X)E(Y)}{E(Y^2) - E(Y)^2}, \\ &= \frac{Cov[X, Y]}{\sigma_Y^2}, \\ &= \frac{Cov[X, Y] \cdot \sigma_X}{\sigma_X \sigma_Y \cdot \sigma_Y}, \\ &= \frac{\rho_{XY} \cdot \sigma_X}{\sigma_Y}, \end{aligned} \quad (5.6)$$

$$Z = E(X) - \frac{\rho_{XY} \cdot \sigma_X}{\sigma_Y} \cdot E(Y). \quad (5.7)$$

The minimum value of  $d$  can be obtained by substituting  $a$  and  $Z$  into (5.4);

$$\begin{aligned}
d &= E(X^2) - 2aE(XY) - 2(E(X) - aE(Y))^2 + a^2E(Y^2) + (E(X) - aE(Y))^2, \\
&= E(X^2) - 2aE(XY) + a^2E(Y^2) - (E(X) - aE(Y))^2, \\
&= E(X^2) - E(X)^2 + a^2(E(Y^2) - E(Y)^2) - 2a(E(XY) - E(X)E(Y)), \\
&= \sigma_X^2 + a^2\sigma_Y^2 - 2aCov[X, Y], \\
&= \sigma_X^2 + \left(\frac{Cov[X, Y] \cdot \sigma_X}{\sigma_X\sigma_Y}\right)^2 \cdot \sigma_Y^2 - 2\left(\frac{Cov[X, Y] \cdot \sigma_X}{\sigma_X\sigma_Y}\right) \cdot Cov[X, Y], \\
&= \sigma_X^2 - \rho_{XY}^2\sigma_X^2 = \sigma_X^2(1 - \rho_{XY}^2). \tag{5.8}
\end{aligned}$$

Equation (5.8) shows the relationship between minimum distortion ( $d$ ), and cross-correlation coefficient  $\rho_{XY}$ , where  $\frac{d}{\sigma_X^2} = 1$  for  $\rho_{XY} = 0$ , and  $\sigma_X^2(1 - \rho_{XY}^2) > d \geq 0$  is true for  $0 < \rho_{XY} \leq 1$ . Here  $\sigma_X^2(1 - \rho_{XY}^2) > d \geq 0$  satisfies equation (5.2), i.e.,  $R_{X/Y}$  should be set as 0 for  $d > \sigma_X^2(1 - \rho_{XY}^2)$  at a given  $\rho_{XY}$ .

### 5.2.1 Relationship Between $\rho_{XY}$ and $R_{X/Y}$

In Fig. 5.2, the theoretical behavior of the rate-distortion curve for conditional decoding (reconstruction of  $X$  with respect to  $Y$ ) is shown; variable correlation coefficients used in this graph are;  $\rho_{XY} = \rho = 0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9$ . As expected, for higher  $\rho_{XY}$  values, both the rate and the distortions are reduced.

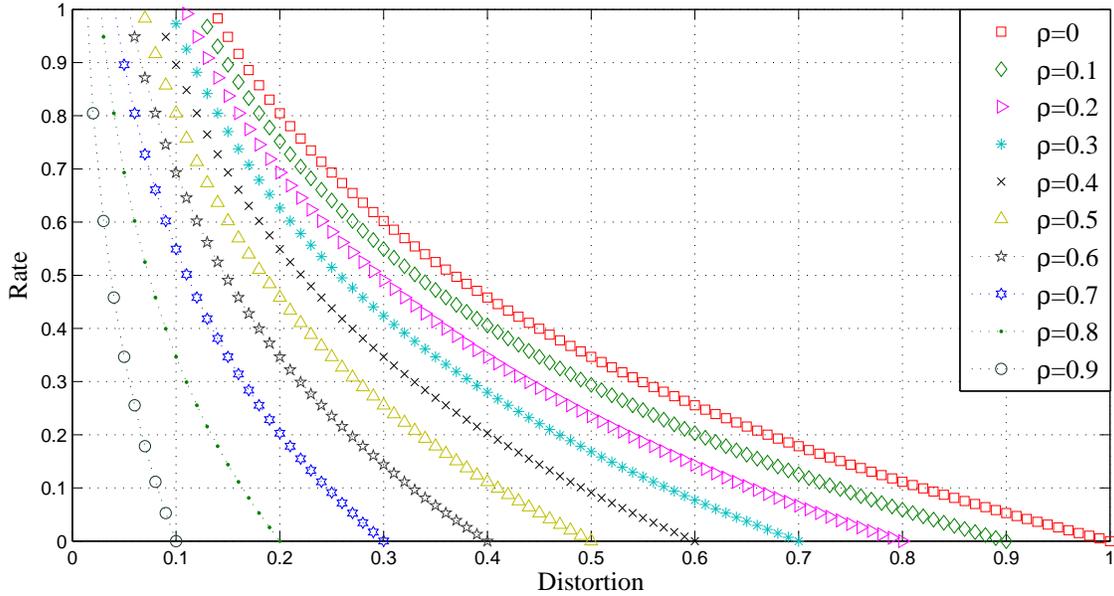


Figure 5.2: Theoretical Relationship of Rate-Distortion with Different Cross-correlation Values

### 5.3 Practical WZVC

Fig. 5.3 shows the block diagram of WZVC. Video sequence  $X \in \{X_1, X_2, \dots\}$  is fed into the WZ video encoder which consists of quantization, bit-plane extraction and Turbo encoder, where parity bit sequence  $s_0$  is generated to transmit to the decoder. At the receiver side, the corresponding side information sequence  $Y \in \{Y_1, Y_2, \dots\}$  is computed from key frames. Here, the key frames are a collection of frames obtained via intra-coding and previously reconstructed video frames.

The side information is known as the distorted systematic information of the original video frame, i.e.,  $Y = X + N$ , where  $N$  is the correlation noise between  $X$  and  $Y$ . The

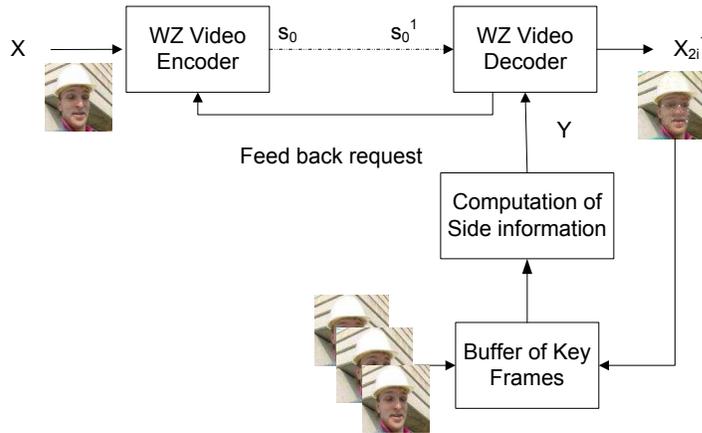


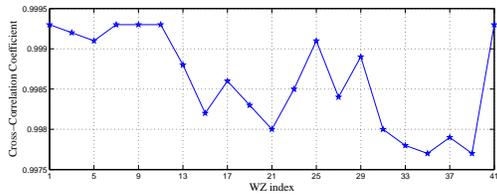
Figure 5.3: A Block diagram of WZVC

reconstruction of  $X$  is dependent upon  $Y$  and the received parity sequence  $s'_0$ . The transmission rate is measured from the number of parity bits requested by the decoder that is required to correct the error in  $Y$ .

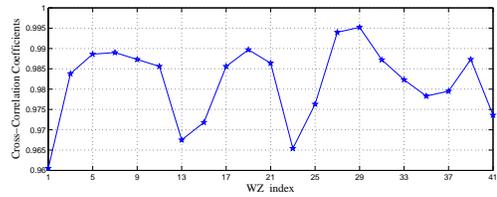
Looking back at Fig. 5.2, it is evident that the rate depends on the cross-correlation coefficient between  $X$  and  $Y$ ,  $\rho_{XY}$ . In practice,  $X$  is unknown to the decoder, thus, calculating  $\rho_{XY}$  is not possible. Therefore, it is important to discuss the computation of cross-correlation values based on the available key frames at the decoder, as described in Section 5.3.1

### 5.3.1 Observation of Cross-Correlations

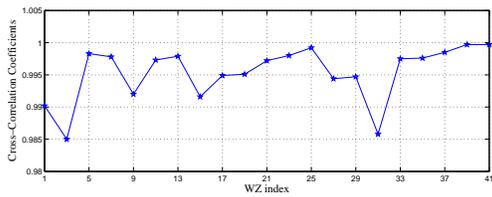
The cross-correlation values have been simulated with the knowledge of key frames. Fig. 5.4 shows the cross-correlation values for different video sequences: Foreman, Carphone, Hall,



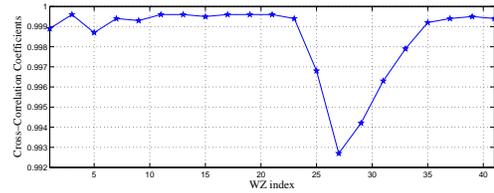
(a) Hall Video Sequence



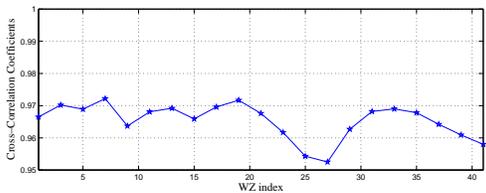
(b) Foreman Video Sequence



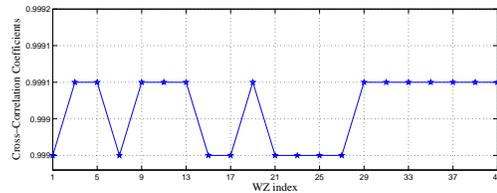
(c) Carphone Video Sequence



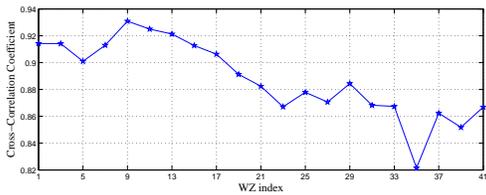
(d) Mother & Daughter Video Sequence



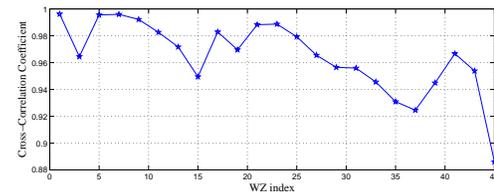
(e) Mobile Video Sequence



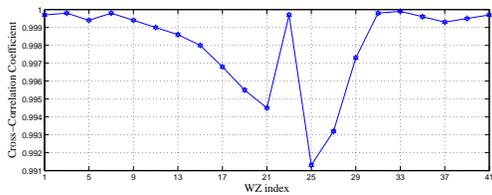
(f) Bridge Closer Video Sequence



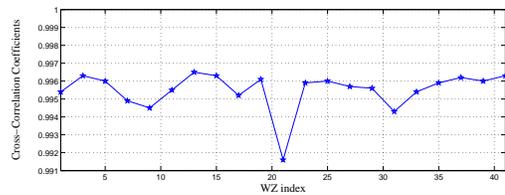
(g) Coast Guard



(h) Suzie Video Sequence



(i) Silent Video Sequence



(j) Highway Video Sequence

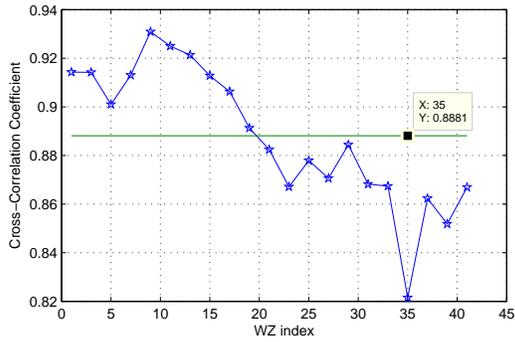
Figure 5.4: Cross-Correlations of Ten Different Video Sequences

Mobile, Mother & Daughter, Silent, Suzie, Bridge closer look, Cost guard, and Highway. The x and y axes of graphs in Fig. 5.4 show the index of WZ frames and the cross-correlation values, respectively. For instance, coordinate (5,0.9) ( $x = 5$  and  $y = 0.9$ ) is equivalent to  $\rho_{X_5 Y_5} = 0.9$ .

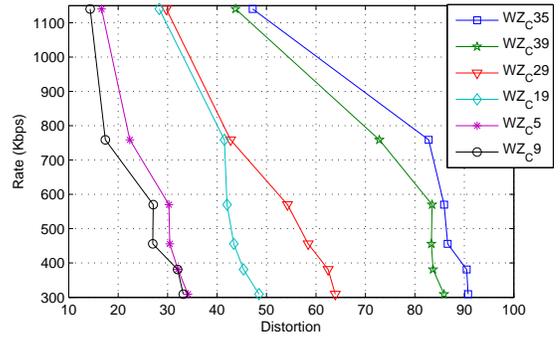
It is apparent that the cross-correlation values vary with different video sequences due to the type of motion in each video sequence. In our simulation, we have observed two categories of video samples; 1) slowly varying cross-correlation with few drops, e.g., Mother & Daughter, Silent, Highway, Mobile and Suzie, and 2) significant variation of cross-correlation values, or fast variation, e.g., Hall, Carphone, Coast Guard, Foreman and Bridge closer. The sudden drops in cross-correlation coefficients are observed in some of the video sequences such as Silent, Highway, and Mother & Daughter. It is due to a fast variation between the corresponding frames; e.g. higher motion in foreground or background of the given frame.

### 5.3.2 Observations of Rate-Distortion

Fig. 5.5, Fig. 5.6, and Fig. 5.7 show the simulation results of cross-correlation coefficient values for 41 WZ indices and rate-distortion behavior of selected frames of Coast Guard, Suzie, and Mother & Daughter, respectively. Fig. 5.5(a) depicts the cross-correlation pattern of Coast Guard video sequence, which exhibits a significant variations in the video

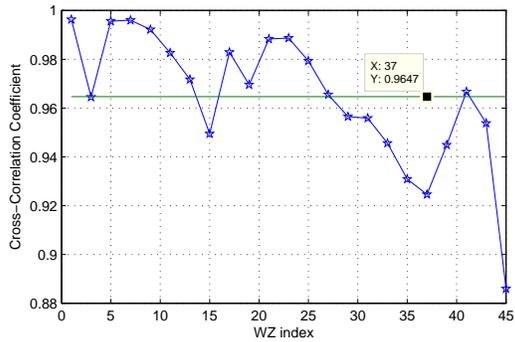


(a) Cross Correlation

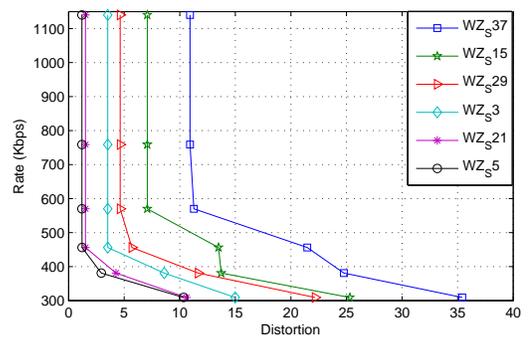


(b) Rate-distortion

Figure 5.5: Coast Guard Video Sequence

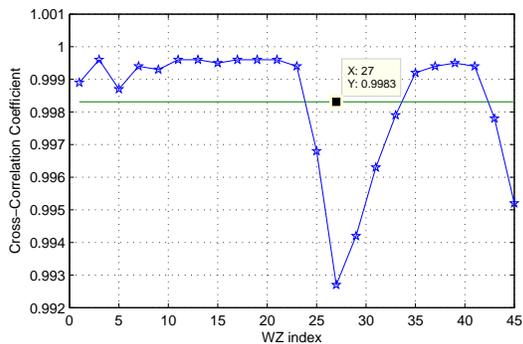


(a) Cross Correlation

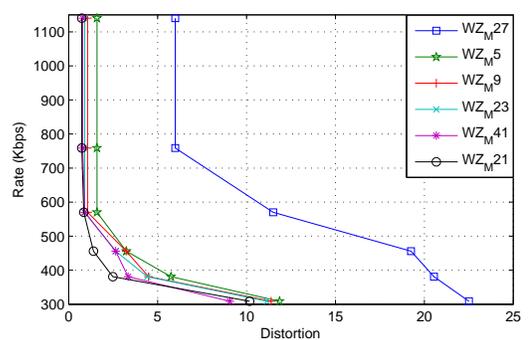


(b) Rate-distortion

Figure 5.6: Suzie Video Sequence



(a) Cross Correlation



(b) Rate-distortion

Figure 5.7: Mother and Daughter video Sequence

sequence. Rate-distortion of selected Coast Guard frames are shown in Fig. 5.5(b). Here, the ninth frame of Coast Guard sequence ( $WZ_C9$ ) shows lowest rate-distortion behavior, whereas  $WZ_C35$  is heavily distorted even at higher rate; frame  $WZ_C5$  exhibits similar behavior to  $WZ_C9$ . The other reconstructed frames  $WZ_C39$ ,  $WZ_C29$  and  $WZ_C19$  are bounded within the region covered by  $WZ_C9$  and  $WZ_C35$ . This behavior is validated in Fig. 5.5(a), which shows that the higher coefficient is at  $x = 9$  ( $\rho_{X_9Y_9}$ ) and the lower at  $x = 35$  ( $\rho_{X_{35}Y_{35}}$ ). Additionally, the cross-correlation values  $\rho_{X_{19}Y_{19}}$ ,  $\rho_{X_{29}Y_{29}}$  and  $\rho_{X_{19}Y_{19}}$  are between that of  $\rho_{X_9Y_9}$  and  $\rho_{X_{35}Y_{35}}$ .

The rate-distortion of Suzie video sequence is shown in Fig. 5.6(b). Here, the convex behavior is observed for frame  $WZ_S5$ ; however, higher rate-distortion is observed for  $WZ_S37$ . This is due to the higher cross-correlation value at  $x = 5$  and lower value at  $x = 37$  as shown in Fig. 5.6(a).

Fig. 5.7(a) shows the cross-correlation values of Mother & Daughter, which exhibits slow variation compared to Fig. 5.5(a) and Fig. 5.6(a). Therefore, it can be concluded that rate-distortion of the Mother & Daughter, in Fig. 5.7(b), shows a better performance than the other two sequences. However, due to the sudden decreases in cross-correlation at  $x = 27$ ,  $WZ_M27$  shows a higher rate-distortion.

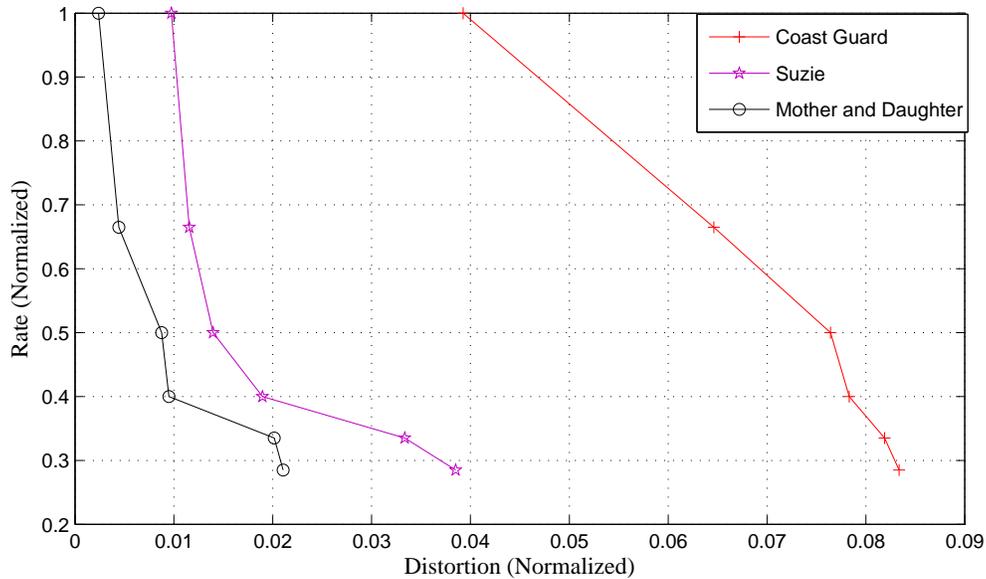


Figure 5.8: Average Rate-Distortion for Coast Guard, Suzie, and Mother & Daughter Video Sequences

Fig. 5.8 shows the average rate-distortion of the WZVC output of the video sequences Coast Guard, Suzie, and Mother & Daughter. It is apparent that Mother & Daughter and Suzie have better rate-distortion behavior than Coast Guard. In other words, the average cross-correlation of Suzie is higher, compared to Coast Guard, lower than the Mother & Daughter.

Fig. 5.9 compares the theoretical rate-distortion with practical rate-distortion behavior of Coast Guard, Suzie, and Mother & Daughter video sequences; here the practical result approaches the theoretical rate-distortion behavior.

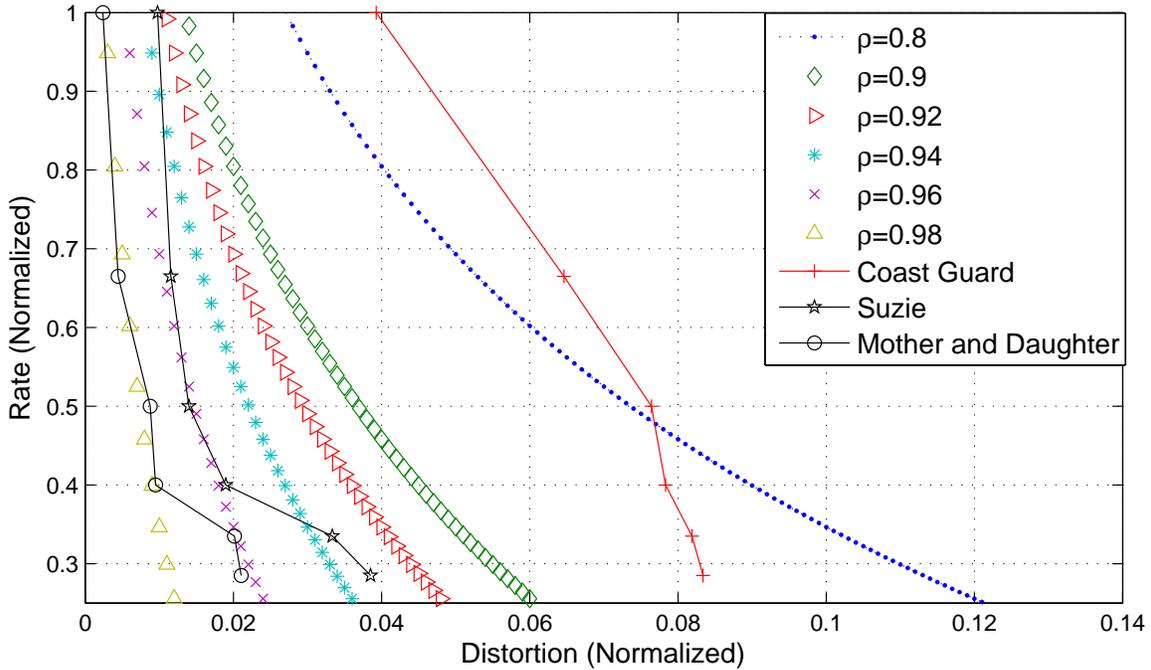


Figure 5.9: Theoretical Rate-Distortion Comparison with Coast Guard, Suzie, and Mother & Daughter Video Sequences

## 5.4 Cross-correlation Thresholds and Rate Control Algorithm

From Figures. 5.5(b), 5.6(b), 5.7(b) and 5.8, we have observed that the rate-distortion curves of the reconstructed video sequences vary with the cross-correlation of the key frames, and that these values approximate the theoretical curves in Fig. 5.9. Additionally, it can be claimed that; 1) *higher correlated key frames show lower rate-distortion*, and 2) *the higher rate-distortion relationships are observed at the lower correlated key frames*. Thus, the encoder rate can be pre-determined from the behavior of cross-correlation points at the decoder and additionally, the cross-correlation points vary with time. Therefore, we need to classify

the level difference of the cross-correlation points. Hence we propose an adaptive threshold level algorithm that separates the cross-correlation indices based on their importance. The following sections define the cross-correlation thresholds (CCTHs) and subsequently the encoder rate control algorithm.

## 5.5 Cross-correlation Thresholds (CCTHs)

This section proposes different cross-correlation thresholds (CCTHs) with respect to the cross-correlation values. Initially, mean (E) is chosen as the threshold. Then positive threshold (PTH) and negative threshold (NTH) are introduced in order to increase the number of cross-correlation regions. Finally, an adaptive threshold (ATH) is proposed based on the behavior of the video sequences.

### 5.5.1 Mean as The Threshold

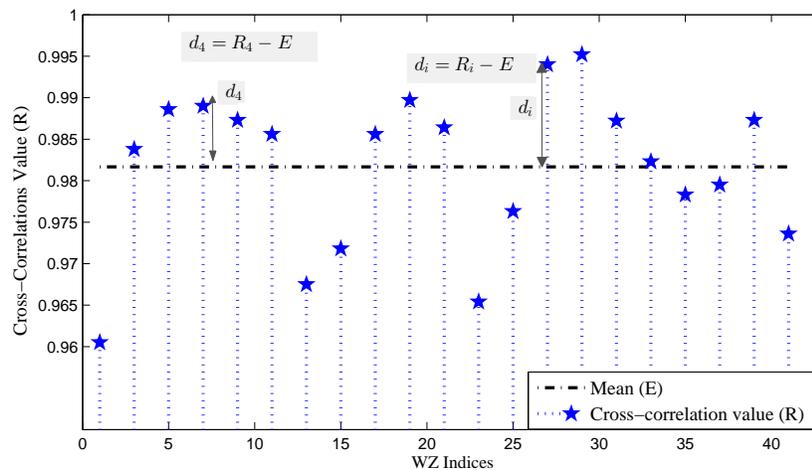


Figure 5.10: Cross-correlations of Foreman Sequences with Mean

Fig. 5.10 shows the cross-correlations ( $R$ ) of the Foreman sequence for 41 WZ indices. From the cross-correlation points,  $E$  is calculated as;  $E = 1/n \sum_{i=1}^n R_i$ , where  $R_i$  is the  $i^{th}$  cross-correlation value and  $n$  is the total number of WZ indices in the simulation. In Fig. 5.10,  $d_i$  is given as the distance from point  $R_i$  to threshold  $E$ , i.e.,  $d_i = R_i - E$ , thus  $d_i$  varies with different  $R_i$  values. If  $d_i \geq 0$ , then the cross-correlation region becomes a positive region, otherwise it is denoted as negative region ( $d_i < 0$ ). Since the length of  $d_i$  varies significantly across different cross-correlation points, another set of thresholds are introduced to minimize the distance between the threshold and cross-correlation points.

### 5.5.2 Positive and Negative Thresholds

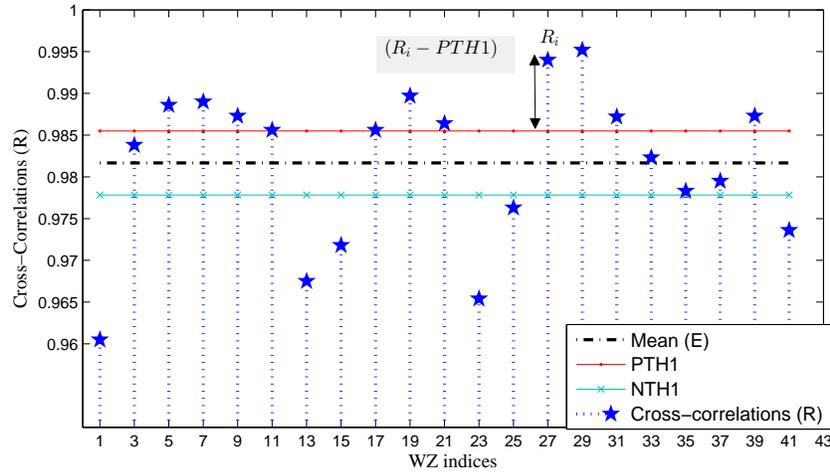


Figure 5.11: Cross-correlations of Foreman Sequences with Mean, PTH1 and NTH1

Fig. 5.11 shows the cross correlation points with three thresholds; mean  $E$ , positive threshold (PTH1) and negative threshold (NTH1). PTH1 is computed for  $d_i > 0$ ;  $PTH1 =$

$E + E(pd)$ , where  $E(pd) = 1/n' \sum_{j=1}^{n'} pd_j$  for  $j = 1, 2, \dots, n'$ .  $E(pd)$  is the mean of the positive differences  $pd_j$  and  $n'$  is the number of cross-correlation points when  $d_i > 0$ . Similarly, NTH1 is calculated from the negative differences  $nd(k)$  when  $d_i < 0$ ; i.e.,  $NTH1 = E + E(nd)$ , where  $E(nd)$  is the mean of the negative differences;  $E(nd) = 1/(n - n') \sum_{k=1}^{n-n'} nd_k$ .

### 5.5.3 Upper Limit of the Thresholds

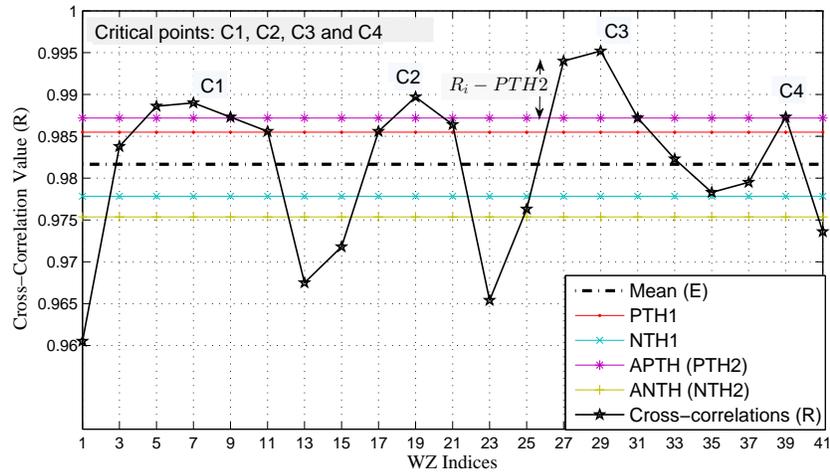


Figure 5.12: Cross-correlations of Foreman Sequences with Mean, PTH2, PTH1, NTH1, and NTH2

PTH1 is obtained from  $R_i - E$ ; similarly, PTH2 can be obtained from positive differences of  $R_i - PTH1$ , and NTH2 from the negative differences of  $R_i - NTH1$ . The set of CCTHs: E, PTH1, PTH2, NTH1 and NTH2 are graphed in Fig. 5.12. Furthermore, the computation of thresholds are continuous in nature, such as PTH3 & NTH3, PTH4 & NTH4, ... PTHn & NTHn, and therefore must be halted after an appropriate iteration. Since higher cross-correlation points become less significant, it is required to find the upper limit of CCTH;

i.e., the iterative calculation of positive threshold via analysis of positive envelop ( $d_i > 0$ ) of critical points must be determined. Here, the critical point is known as local maxima, which sets the upper limit of CCTH.

#### 5.5.4 Adaptive Threshold (ATH)

The definition of the local maxima critical points are a well-known property in calculus and linear algebra. For a given point to be a local maxima, the preceding tangent of the given point must be positive, and the following tangent of the given point must be negative. From Fig. 5.12, there is more than one critical point in the positive envelop.

Adaptive threshold is obtained from the combination of iterative computation of thresholds and maxima critical points ( $CR_i$ ), which are stored at the decoder buffer ( $CR_i$ ). Iteration is stopped if the threshold reaches the minimum of local maxima, i.e.,  $PTH \geq MIN(CR_i)$ ; thus PTH and NTH are denoted as upper and lower limits of CCTHs. Since these values vary with respect to cross-correlation behavior for different video sequences, they are called adaptive thresholds (ATH): adaptive positive threshold (APTH) and adaptive negative threshold (ANTH).

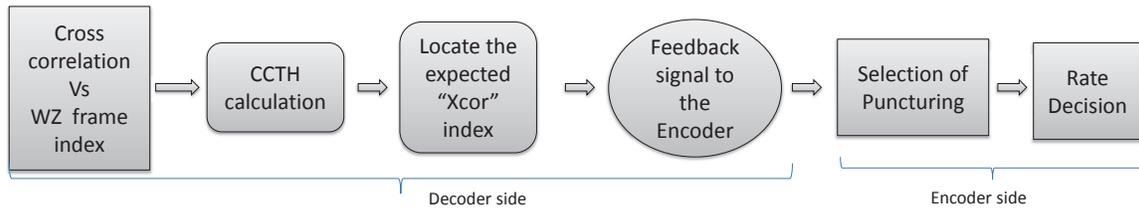


Figure 5.13: Block Diagram of Rate Control Process

## 5.6 Rate Control Algorithm

Fig. 5.13 shows the block diagram of the rate control process, where the process is divided into two; i) decoder-side process and ii) encoder-side process. The decoder-side process has four major operations; 1) buffering (tabulating) the cross correlation vs. WZ frame index, 2) calculating CCTH, 3) locating the cross-correlation point with respect to the WZ index, and 4) feedback signal generator to the encoder.

The encoder-side process determines the rate which is based on the received feedback pulses (P1, P2, ...), where Turbo encoder uses puncturing to determine the number of transmitted parity bits. The encoder has two operation blocks; 1) set the puncturing and 2) rate determination. Fig. 5.14 outlines how the algorithm works for sending feedback pulses, and Fig. 5.15 depicts the state diagram of rate determination with respect to the feedback pulses.

The decoder buffer contains the set of cross-correlation (Xcor) values with respect to the WZ frame index. Additionally, it also stores the calculated CCTHs. If the 'Xcor' of

expected WZ index is known, then the corresponding 'Xcor' is located from the decoder buffer. Thus, 'Xcor' is compared with CCTHs, and the decoder sends the feedback pulses to the encoder. In other words, the pulse generator synchronizes with the cross-correlation regions with respect to the CCTHs.

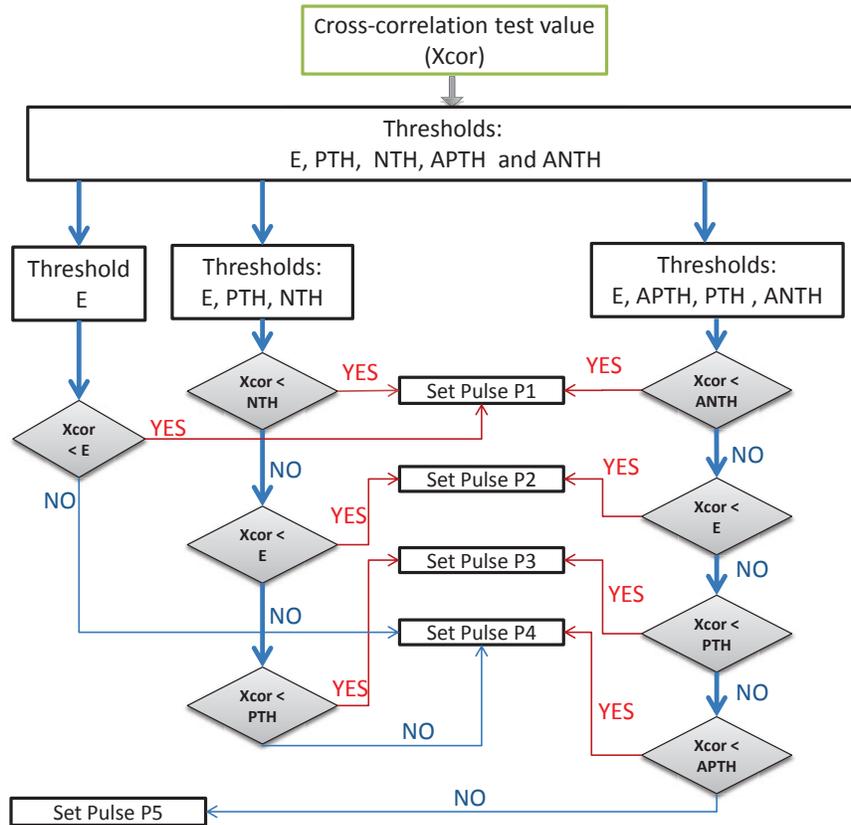


Figure 5.14: Flow Chart for Deciding Feedback Signal

Fig.5.14 shows a flowchart, which explains the algorithm of generating feedback signals. E, PTH, NTH, APTH and ANTH are the available CCTHs. The cross-correlation value 'Xcor' is compared with the known threshold values. Three cases are analyzed; i) CCTH as

E, ii) CCTHs as E, PTH and NTH, and iii) CCTHs as E, APTH, ANTH, and PTH.

### 5.6.1 Rate Control Based on Mean (E)

Table 5.1: Cross-correlation Coefficients and Rate Regions with Threshold E

Cross-correlations	Region	Required Rate
$X_{cor} < E$	Region 1	Pulse P1 (Higher rate required R1)
$X_{cor} \geq E$	Region 2	Pulse P4 (lower rate required R4)

The mean (E) divides the cross-correlation statistics into two regions. The expected WZ index 'Xcor' is compared with E; If  $X_{cor} < E$ , then signal pulse P1 is sent to the encoder (request rate is to be 'R1') otherwise P4 is sent (request rate is 'R4'), as shown in Table 5.1.

### 5.6.2 Rate Control Based on PTH and NTH

Table 5.2: Cross-correlation Coefficients and Rate Regions with Thresholds: PTH, NTH and E

Cross-correlations	Region	Required Rate
$X_{cor} \leq NTH$	Region 1	Pulse P1 (Higher rate required R1)
$NTH < X_{cor} < E$	Region 2	Pulse P2 (Medium higher rate required R2)
$E \leq X_{cor} < PTH$	Region 3	Pulse P3 (Medium lower rate required R3)
$X_{cor} \geq PTH$	Region 4	Pulse P4 (lower rate required R4)

There are four regions with respect to the thresholds NTH, E and PTH. The rate control algorithm in terms of the given 'Xcor' is described in Table 5.2, where the regions and corresponding request rates are shown. Three simulation cases are considered; 1) Both the PTH1 and NTH1 ("PTH1NTH1"), and 2) "PTH2NTH2", where PTH1 and NTH1 are

replaced by PTH2 and NTH2, respectively, which increase the space of region 2 and region 3 with respect to the PTH1NTH1, 3) region 1 and region 2 are unchanged, but only region 3 is altered, and this step is denoted as "PTH1NTH2".

### 5.6.3 Rate Control Based on Adaptive Thresholds APTH and ANTH

Table 5.3: Cross-correlation Coefficients and Rate Regions with Thresholds: APTH, ANTH, PTH and E

Cross-correlations	Region	Required Rate
$X_{cor} \leq ANTH$	Region 1	Pulse P1 (Higher rate required R1)
$ANTH < X_{cor} < E$	Region 2	Pulse P2 (Medium higher rate required R2)
$E \leq X_{cor} < PTH$	Region 3	Pulse P3 (Medium lower rate required R3)
$PTH \leq X_{cor} < APTH$	Region 4	Pulse P4 (Lower rate required R4)
$X_{cor} \geq APTH$	Region 5	Pulse P5 (Least punctured rate R5)

APTH and ANTH define four regions, whose sizes depend on the positive envelop curve. Here, the positive region is separated further using PTH to reduce distance from the cross correlation points to APTH and PTH. ANTH is marked as only one threshold in  $d_i < 0$  to give higher bit rate for lower cross correlated WZ index points. Table 5.3 shows the pulse generator with respect to the 'Xcor' position.

Fig. 5.15 shows the state diagram of the rate decision at the encoder. The required rates R1, R2, R3, and R4 are determined based on the received pulse signals P1, P2, P3, and P4, respectively.

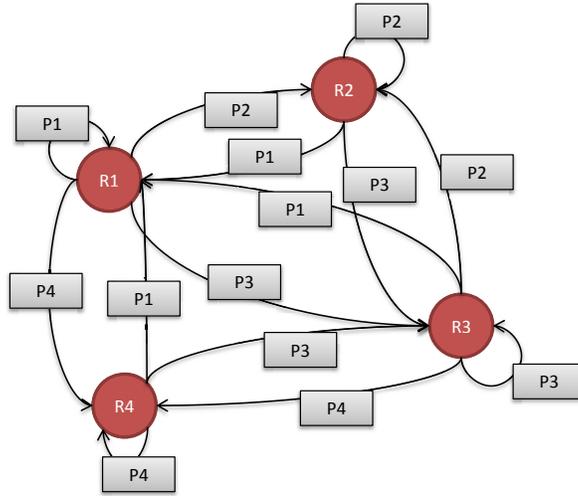
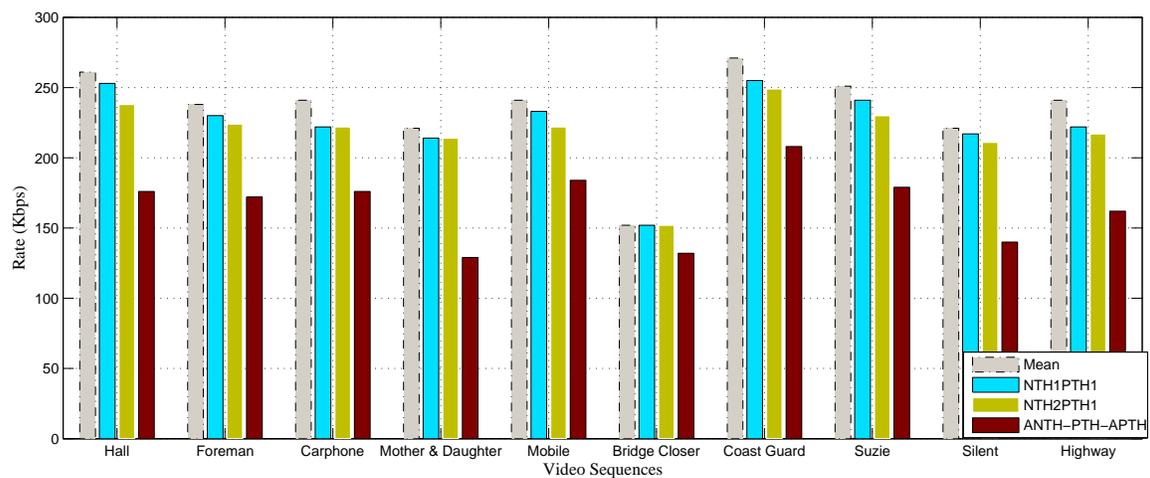


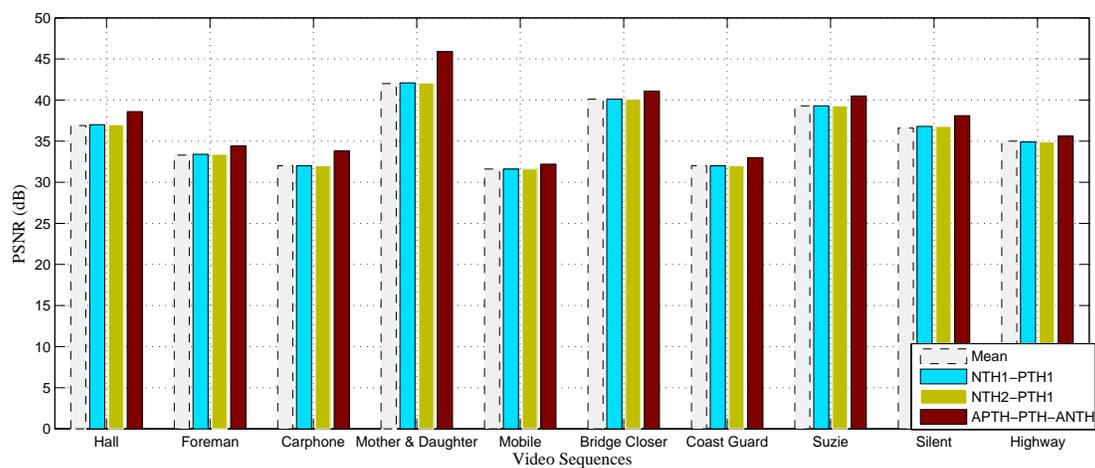
Figure 5.15: State Diagram for Deciding the Encoder Rate

## 5.7 Performance Analysis

In this section, we present the simulation results that illustrate the average PSNR and the average rate for ten different video sequences using our proposed rate control algorithm. We also show the optimum rate-distortion points with respect to the rate-distortion curves. The simulation is performed for five different cases of thresholds: E, NTH1TH1, NTH2PTH2, NTH2PTH1 and ANTH-PTH-APTH.

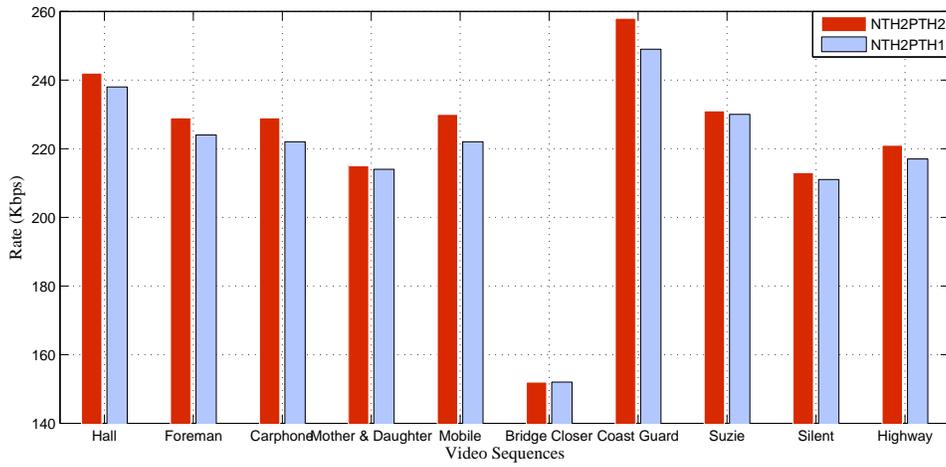


(a) Rate Comparison

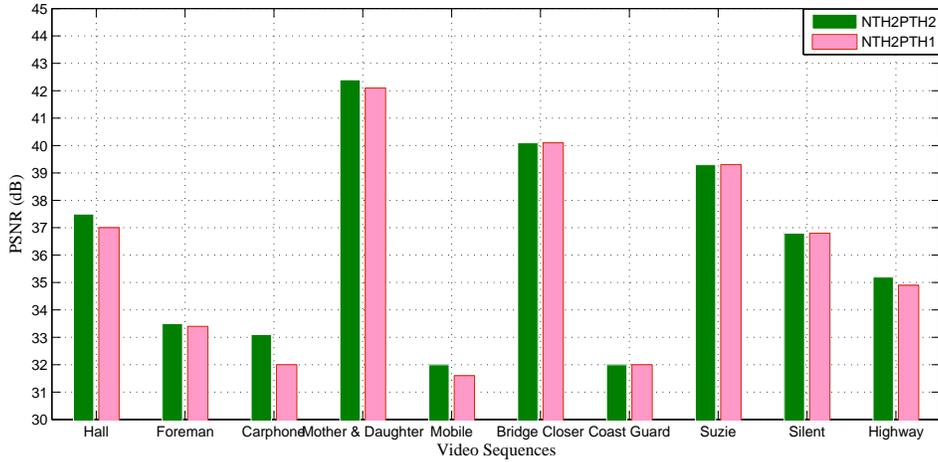


(b) PSNR Comparison

Figure 5.16: Rate and PSNR Comparison of Four CCTHs: Mean, NTH1PTH1, NTH2PTH2, and ANTH-PTH-APTH



(a) Rate Comparison



(b) PSNR Comparison

Figure 5.17: Rate and PSNR Comparison of CCTHs: NTH2PTH1 and NTH2PTH2

Fig. 5.16 shows the comparison of rate (Fig. 5.16(a)) and PSNR (Fig. 5.16(b)) with the proposed CCTH algorithms, where four simulation cases are compared to each other. In total, ten video sequences are considered in these simulations. The threshold E requests more bit rate in all the sequences since E separates the cross-correlation regions into two: either higher or lower rate. In NTH1PTH1 and NTH2PTH1, the NTH2PTH1 performs better at reducing the rate since NTH2 reduces the number of lower cross correlation points, which in turn request more bits.

Table 5.4: Observation of Rate Variation and PSNR Variation for ANTH-PTH-APTH Compared to 50% Compressed WZVC

Video Sequences	ANTH-PTH-APTH		Compression 50%	
	Variation of Rate (%)	Variation of PSNR (%)	Rate (Kbps)	PSNR (dB)
Hall	-7.4	3.2	190	37.4
Foreman	-9.5	9.2	190	31.5
Carphone	-7.4	6.3	190	31.8
Mother & Daughter	-32.1	13.3	190	40.5
Mobile	-3.2	2.2	190	31.5
Bridge Closer	-30.5	2.5	190	40.1
Coast Guard	9.5	8.2	190	30.5
Suzie	-5.8	5.2	190	38.5
Silent	-26.3	3.5	190	36.8
Highway	-14.7	3.2	190	34.5

Fig. 5.17(a) shows that the NTH2PTH2 shows higher rate requirement than NTH2PTH1; NTH2PTH2 reduces the number of points with higher cross correlation value, which request lower bits. In addition, Fig. 5.17(b) shows that NTH2PTH2 exhibits better PSNR com-

pared to NTH2PTH1. Therefore upper limit of PTH is the deciding factor in rate-distortion. Finally, the ANTH-PTH-APTH method shows a better performance than all other cases mentioned in Fig. 5.16, and it is due to the adaptive threshold. Adaptive threshold reduces the possibilities of lower cross correlation points, which request higher number of bits (reduce the rate), and reduces the higher cross correlation points, which request lower bits (increase PSNR). Furthermore, ANTH-PTH-APTH increases the number regions of cross-correlation values from four to five as shown in Table 5.3. Therefore, an improvement in PSNR and rate reduction can be seen for ANTH-PTH-APTH method in Fig. 5.16(b) and Fig. 5.16(a), respectively.

Table 5.4 shows the average rate and PSNR variations of the adaptive threshold technique compared to the 50% compressed WZVC; here it is compared for ten video sequences. We observed that the proposed method performs higher PSNR with lower rate in all the video sequences except Coast Guard. In Coast Guard, the proposed method achieved a minimum average rate as of 208 Kbps with  $PSNR = 33$  dB. Now, it is interesting to compare the PSNR performance of Coast Guard between the proposed (ANTH-PTH-APTH) method and constant rate (208 Kbps).

In Fig. 5.18, PSNR performance of proposed method is comparatively better than the

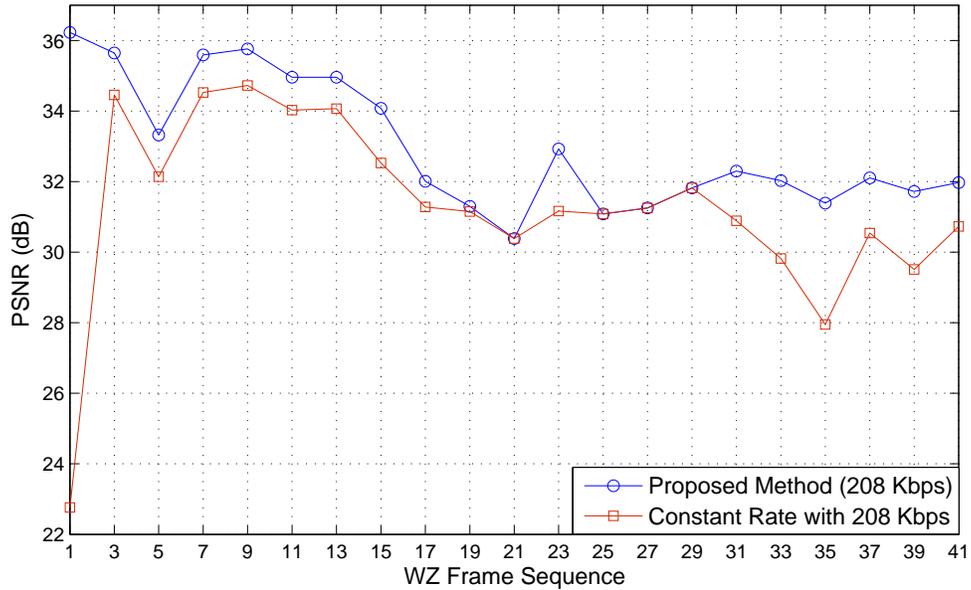


Figure 5.18: PSNR Variations vs Reconstructed WZ frames for Coast Guard Video

fixed rate of 208 Kbps. However, PSNR improvement is not that much higher since the cross-correlation behavior of Coast Guard exhibits fast variation as shown previously in Fig. 5.5(a). Fig. 5.19 shows PSNR performance of the proposed algorithm compared to the PSNR of the constant rate (172 Kbps) for Foreman video sequence. Again, PSNR of the proposed method performs better than PSNR in constant rate (172 Kbps); here, 172 Kbps is calculated as an average rate for 41 WZ frames based on the proposed adaptive threshold technique.

In Mother & Daughter, average rate of proposed method is 112 Kbps. Fig. 5.20 shows PSNR vs. WZ video sequence for Mother & Daughter video sequence; here, the proposed

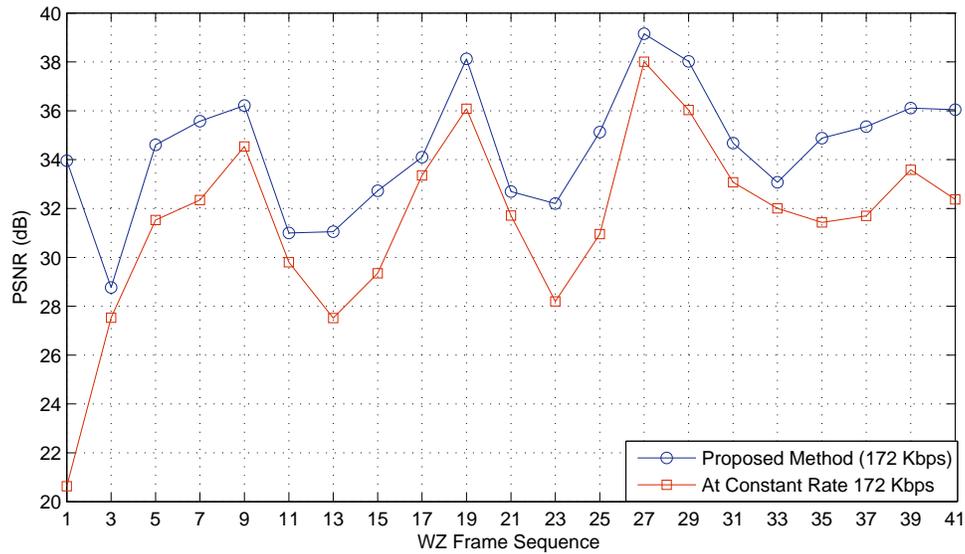


Figure 5.19: PSNR Variations vs Reconstructed WZ frames for Foreman Video

method (112 Kbps) is compared with constant rate (112 Kbps). The proposed method shows an improvement in PSNR by at least 5 dB. It also shows that PSNR-rate relationship of Mother & Daughter is much better than Foreman video sequence. From the rate-PSNR relationship of the ten video sequences, we have obtained possible minimum rate and best PSNR from the proposed algorithm. In order to compare the theoretical and practical rate-distortion relationship with our result (best rate-PSNR), we normalize our best rate-PSNR value as optimum rate-distortion points.

Fig. 5.21 shows the optimum rate-distortion points of ten video sequences with respect to the theoretical rate-distortion curves. It can be seen that the optimum points are closer to the lower rate-distortion curves. Fig. 5.22 shows optimum rate-distortion points with respect

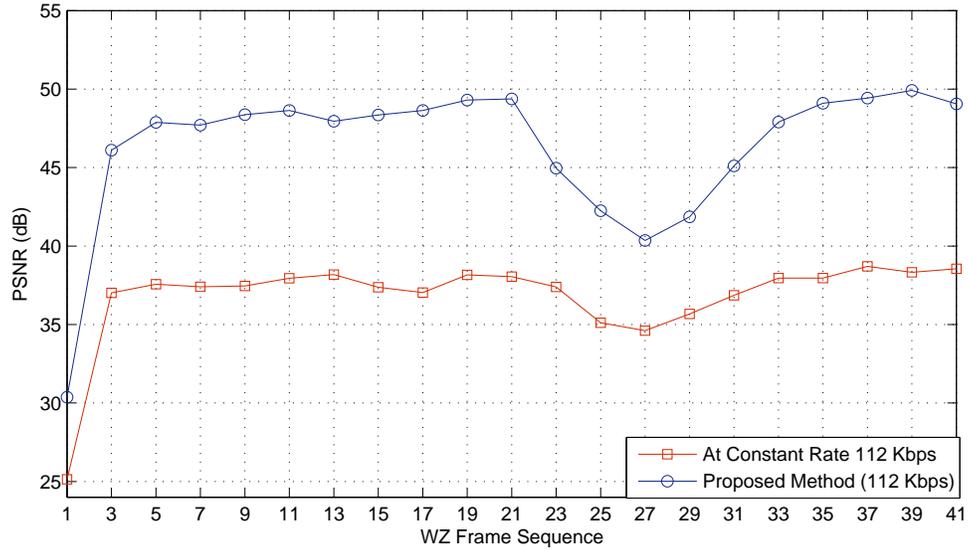


Figure 5.20: PSNR Variations vs Reconstructed WZ frames for Mother & Daughter Video to the practical and theoretical rate-distortion curves; here, practical rate-distortion curves and rate-distortion points are graphed for Coast Guard, Suzie, and Mother & Daughter video sequences. The optimum point is determined as the best value of rate and distortion for the given video sequences. For instance, for a given rate, any optimum point gives lower distortion, as well as a lower rate for a given distortion, in each video sequence.

## 5.8 Summary

Encoder rate/transmission rate in WZVC depends on the correlation between the video frames. In this chapter, we have investigated the cross-correlation coefficients ( $\rho_{XY}$ ) of video frames with respect to the conditional WZ rate  $R(X/Y)$ . Additionally, the rate-distortion relationship of the reconstructed video is examined and compared to the cross-correlation

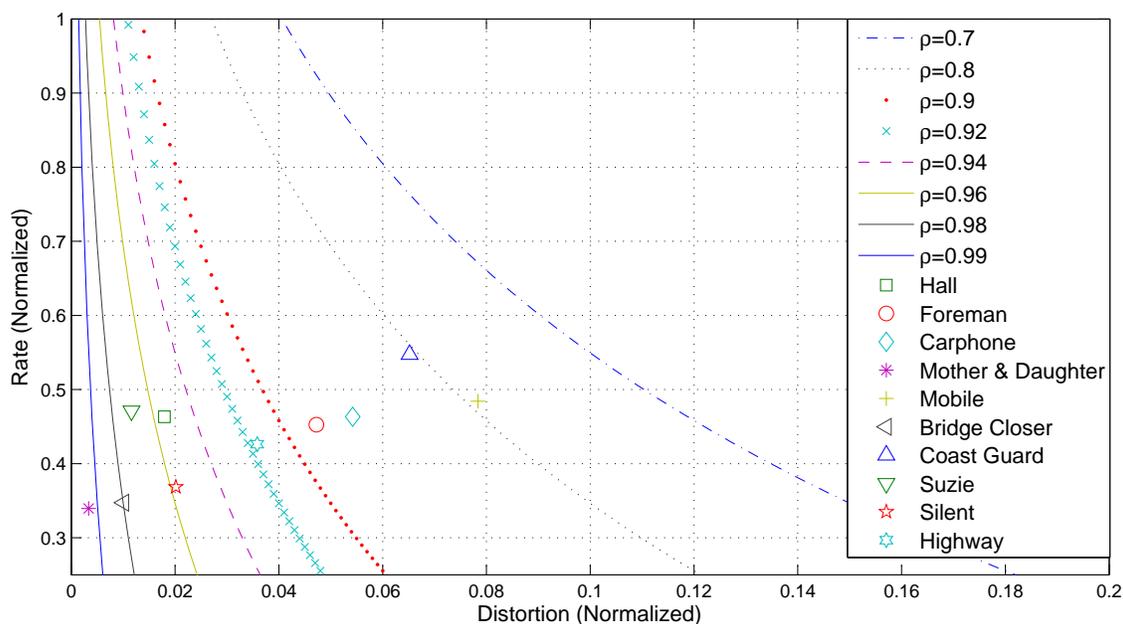


Figure 5.21: Best Rate-Distortion (RD) Points with Theoretical Rate-Distortion Curves

coefficient values of the key frames. We found that the rate-distortion relationship is directly related to the cross-correlation coefficients of the key frames. However, it is proven that the theoretical conditional rate (*rate-distortion relationship*) is a function of theoretical cross-correlation coefficient for Gaussian random variables, and it is justified in our practical simulation model by using the available key frames. We then introduced a cross-correlation threshold (CCTH) technique to control the encoder rate, and defined different levels of thresholds to divide possible regions for cross-correlation coefficients. Furthermore, adaptive threshold (ATH) technique is a potential threshold level in our research which will vary depending on the behavior of the video sequences. A significant reduction of encoder rate was obtained with higher PSNR by our proposed method. Simulation results (rate-PSNR values) have shown better performance than the 50% compressed WZVC, which is also

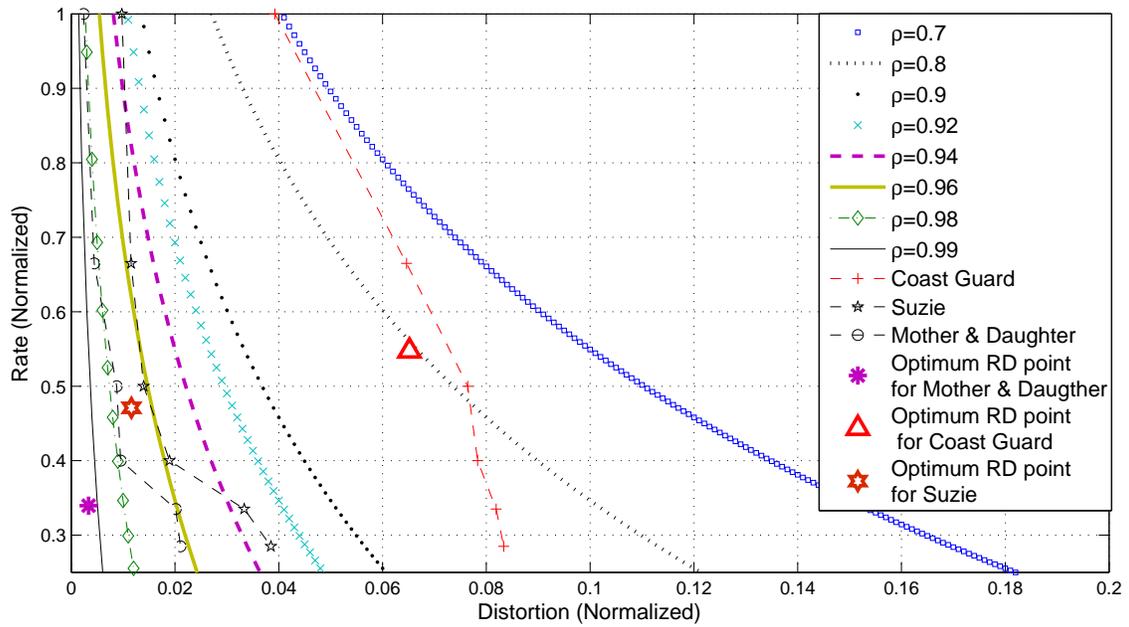


Figure 5.22: Best RD Points with Practical and Theoretical Rate-Distortion curves

compared with the theoretical and practical rate-distortion relationship.

# Chapter 6

## Bit Based Correlation Noise Estimation at the Decoder

### 6.1 Introduction

In WZVC,  $X$  and  $Y$  are correlated video frames;  $Y$  is side information, which is a noisy version of WZ frame  $X$ . The correlation noise between  $X$  and  $Y$ ,  $N$ , can be shown as  $N = Y - X$ ;  $X$  and  $Y$  are available at the encoder and decoder side, respectively. Typically, the probability distribution of  $N$  is approximated to Gaussian distribution [45]. From error-control coding (ECC), the encoder needs to transmit the parity bit sequence of an input video frame  $X$ , and decoder reconstructs  $X$  ( $X'$ ) with these received parity bits and side information  $Y$  [58].

Estimating the noise  $N$  due to the residual difference between side information and input video frame is one of the areas of greatest focus in WZVC research. A few previous

researchers have used a Laplacian distribution as a noise distribution as shown [5];

$$f(X - Y) = \frac{\beta}{2} \exp(\beta(-|X - Y|)) \quad (6.1)$$

where, the Laplacian distribution parameter  $\beta$  is given by  $\beta^2 = \frac{2}{\sigma^2}$  and  $\sigma^2$  is the variance of the correlation noise. The  $\beta$  has been computed over the entire video sequence off line at the encoder before the Wyner-Ziv coding procedure starts [4]. This computed  $\beta$  is then kept constant throughout the decoding of all the Wyner-Ziv frames [4]. However, note that the value of  $\beta$  is essential to convert the side information into soft-value information that is required for Turbo decoding. The soft value of side information can be computed with respect to the bit-metric value of each bit of side information. The bit-metric value depends on the noise variance ( $\sigma^2$ ), therefore, it must be estimated for each bit of side information. A few researchers have proposed variance estimation in terms of frame or pixel; here, estimation is done either online or off-line [16], [77]. However, iterative decoding uses only pixel-based variance in each bit of the side information. Therefore, it is worthwhile to investigate variance in terms of each bit from the side information.

This chapter proposes a new dynamic algorithm to estimate the *bit-based variance* in the side information. The bit-based variance can be used to compute the soft-value of side

information, which is used as the received systematic value at the decoder end.

This chapter is organized as follows: Section 6.2 proposes a technique for correlation noise estimation between the WZ frame and the side information. Section 6.3 elaborates the results and discussion of our research. Finally, conclusion is presented in Section 6.4.

## 6.2 Correlation Noise Variance Estimation

Log likelihood ratio(LLR) of a transmitted bit  $u_l$  of WZ frame, given the corresponding received  $r_l$  of side information, is shown as  $L(u_l|r_l)$  in Chapter 3. From (3.4), LLR is:  $L(u_l|r_l) = L_c r_l + L_a(u_l)$ , where  $L_c = 4(E_s/N_0) = \frac{2E_s}{\sigma_{pic}^2}$  and  $\sigma_{pic}^2$  is variance of correlation noise.

At the decoder, side information is generated using the existing key frames, as shown in [17]. Pixel values of side information  $Y_{2i}$  are obtained by interpolating the pixels of key frames. These pixel values will be converted into a bit stream  $r$  with respect to the bit plane extractor at the receiver. The soft-value,  $L_c \times r$  is computed and input into the Turbo decoder as the systematic information, where  $L_c = \beta = \frac{2}{\sigma^2}$ . This formula shows that the soft-value of the systematic information of Turbo decoder depends on the variance  $\sigma^2$  of correlation noise.

First, we have analyzed the actual distribution of the residual difference between Wyner-Ziv frame and side information; i.e., the luminance difference between corresponding pixels of Wyner-Ziv frame and side information, and then compared it to the theoretical distri-

bution. We have taken the quarter common intermediate format (QCIF) video sequence of Foreman and Carphone. In practice, it is impossible to calculate the variance using Wyner-Ziv and side information since the Wyner-Ziv frame is not known to the receiver. From our simulation result, the correlation between key frames has an effect on correlation noise. Therefore, the statistical relationship between key frames affects the amount of correlation noise between Wyner-Ziv and side information. The weighted mean squared error (WMSE) between the two key frames is;  $WMSE(x, y) = (\frac{1}{2})^2 \frac{\sum_{(x,y) \in SI} (X_B(x,y) - X_F(x,y))^2}{L}$  where,  $L$  stands for frame size, and  $X_B$  and  $X_F$  denote backward and forward key frames, respectively.  $WMSE(x, y)$  is the variance  $\sigma^2$  of the correlation noise, which in practice will vary from pixel to pixel; i.e.,  $\sigma^2$  depends on the difference between pixel values of key frames.

Fig. 6.1 shows two sample bit-planes of key frames. From the observation of correlation noise and the two key frames, it is clear that the noise variance is influenced by the lower bit planes, i.e., due to the least significant bits (LSB). The most significant bits (MSB) do not vary between the key frames in most cases. If it is changed then entire key frames will be altered from the other one.

Fig. 6.2 tabulates the bit-plane of the side information. Here, noise variance is divided among each bit-planes:  $VBP_i$ , where  $i = 0, 1, 2, \dots, 7$ . The correlation noise in side information

affects each bit-plane based on their importance. Since the decoder takes bit by bit from side information, we need to incorporate the noise with each bit. In other word, bit-plane characteristics can be used to convert the distribution of pixel variance into bit-variance. In the bit pattern of each pixel, MSB affects the variance of each pixel much more significantly than the LSB. Correlation noise variance of  $i^{th}$  bit-plane can be shown as  $\sigma_i^2 = (n - i) \times \frac{\sigma_c^2}{n}$ , where  $i$  ( $= 0, 1, 2...7$ ) is bit-plane number and  $n$  is number of bits used for quantization.  $L_c$  is the bit-metric value, and is dependent on the bit-based variance;  $L_c = \beta = \frac{2}{\sigma_i^2}$ .

### 6.3 Results and Discussion

Fig. 6.3(a) and 6.3(b) show the input and reconstructed Foreman frames, respectively. The PSNR value of reconstructed Foreman frame is 29.53 dB. Fig. 6.4(a) and 6.4(b) show the input frame and reconstructed frame of Carphone with  $PSNR = 31.56$  dB, respectively. The correlation between the key frames of Carphone sequence is higher than that of Foreman video sequence as shown in Fig. 6.5 and 6.6. It is due to a small region in the background of the Carphone sequence that appears to be influenced by motion; i.e., only the car window is in a high level of motion. On the other hand, the key frames of Foreman display the high level of motion in the face, as shown in Fig. 6.5, and the background is highly correlated. Subsequently, simulation result shows that there is higher PSNR in Carphone than in Foreman.

Fig. 6.7, and Fig. 6.8 show the probability distribution of correlation noise for Foreman and Carphone video sequences, respectively, where theoretical Laplacian model is compared with the actual distribution of the residual error between Wyner-Ziv and side information frames. We observe that the actual noise distribution closely approximates to Laplacian distribution from Fig. 6.7, and Fig. 6.8. It shows that the probability of practical value of correlation noise is higher for Foreman compared to Carphone.

Fig. 6.9 shows the PSNR performance of the received 87 Foreman video frames. The proposed bit-based noise variance estimation method shows better PSNR compared to the other three curves with the following fixed variances values;  $L_c = 2$ ,  $L_c = 3$ , and  $L_c = 4$ . The lower PSNR points at the fixed variance methods are improved significantly by the proposed bit-based noise estimation method. For instance, 1<sup>st</sup> and 77<sup>th</sup> frames show 4.4 dB and 5.1 dB improvement in PSNR respectively. If the correlation noise is higher, then the reconstructed PSNR value decreases. Our proposed method performs better for the video frames with higher correlation noise. This is because our method protects the most significant bits of each pixels from the distortion.

## 6.4 Summary

In this Chapter, we investigated correlation noise of side information at the receiver, and compared it for Foreman and Carphone video frames. The reconstructed Carphone frame appears to be better quality than the reconstructed Forman frame since its correlation of key frames in the video sequence is higher. Hence, correlation noise is influenced by correlation between key frames.

A novel algorithm is developed to generate soft-values of the side information in this chapter, known as dynamic correlation noise estimation. This Chapter contributes to WZVC research by proposing "bit-based variance" estimation, which supports to compute the soft-value of systematic information. The proposed dynamic bit-based estimation enhances the output video quality especially at higher correlation noise as shown in Fig. 6.9.

Pixels	1	2	3	-	-	-	-	n-1	n
Bit level	176	92	132	-	-	-	-	131	65
Bit Plane 7	1	0	1	-	-	-	-	1	0
Bit Plane 6	0	1	0	-	-	-	-	0	1
Bit Plane 5	1	0	0	-	-	-	-	0	0
Bit Plane 4	1	1	0	-	-	-	-	0	0
Bit Plane 3	0	1	0	-	-	-	-	0	0
Bit Plane 2	0	0	0	-	-	-	-	0	0
Bit Plane 1	0	1	0	-	-	-	-	1	0
Bit Plane 0	0	0	0	-	-	-	-	1	1

(a) Key Frame One

Pixels	1	2	3	-	-	-	-	n-1	n
Bit level	192	96	128	-	-	-	-	129	64
Bit Plane 7	1	0	1	-	-	-	-	1	0
Bit Plane 6	1	1	0	-	-	-	-	0	1
Bit Plane 5	0	1	0	-	-	-	-	0	0
Bit Plane 4	0	0	0	-	-	-	-	0	0
Bit Plane 3	0	0	0	-	-	-	-	0	0
Bit Plane 2	0	0	0	-	-	-	-	0	0
Bit Plane 1	0	0	0	-	-	-	-	0	0
Bit Plane 0	0	0	0	-	-	-	-	1	0

(b) Key Frame Two

Figure 6.1: Bit-planes of Two Key Frames

Pixels	1	-	-	n	Noise Variance
Bit level	$P_1$	-	$P_j$	$P_n$	
Bit Plane 7	1	-	-	0	$VBP_7$
Bit Plane 6	0	-	-	1	$VBP_6$
Bit Plane 5	1	-	-	0	$VBP_5$
Bit Plane 4	1	-	-	0	$VBP_4$
Bit Plane 3	0	-	-	0	$VBP_3$
Bit Plane 2	0	-	-	0	$VBP_2$
Bit Plane 1	0	-	-	0	$VBP_1$
Bit Plane 0	0	-	-	1	$VBP_0$

Figure 6.2: Bit-plane of Side Information



Figure 6.3: (a) Input Wyner-Ziv Frame , (b) Reconstructed Frame of Foreman



Figure 6.4: (a) Input Wyner-Ziv Frame , (b) Reconstructed Frame of Carphone



Figure 6.5: (a) Previous Frame of Wyner-Ziv Frame (Key Frame 1) , (b) Next Frame of Wyner-Ziv Frame (Key Frame 2)



Figure 6.6: (a) Previous Frame of Wyner-Ziv Frame (Key Frame 1), (b) Next Frame of Wyner-Ziv Frame (Key Frame 2)

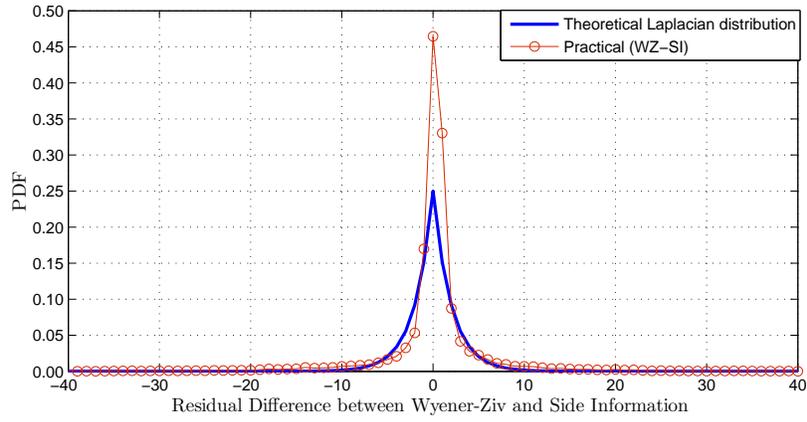


Figure 6.7: Probability Distribution of Correlation Noise for Foreman Video Sequence

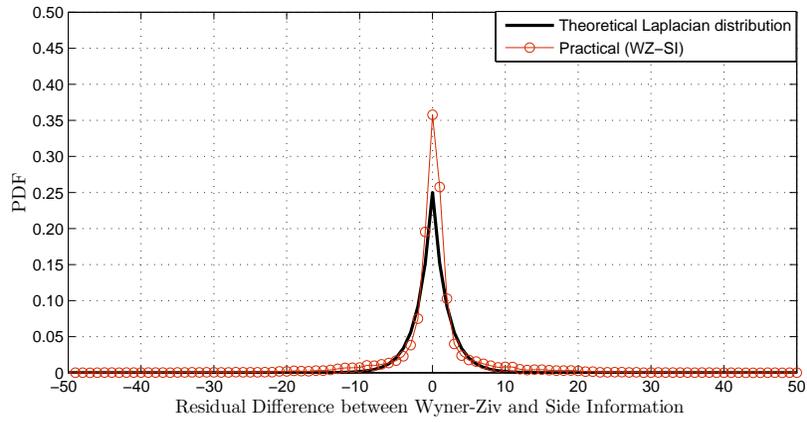


Figure 6.8: Probability Distribution of Correlation Noise for Carphone Video Sequence

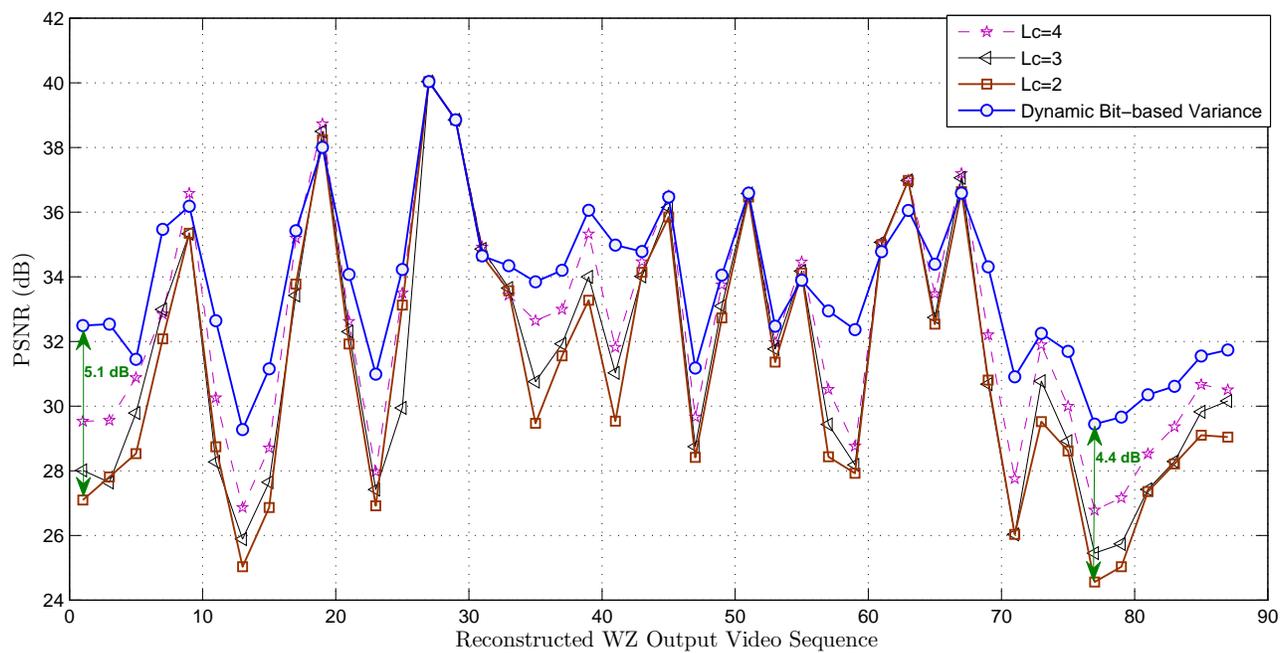


Figure 6.9: PSNR Performance of Foreman Video Sequence with Dynamic Bit-based Variance Estimation Compared to Fixed Variance Methods

# Chapter 7

## Conclusions and Future Work

### 7.1 Conclusions

In this thesis, we began by considering a practical Wyner-Ziv Video Coding (WZVC) with the effects of wireless fading channel and additive white Gaussian noise channel (AWGN). At the decoder, channel-metric values were defined with different channel conditions. The bit-metric value was computed based on the variance of correlation noise between the WZ frame and the side information. We analyzed the quality of the reconstructed (received) video in depth with respect to *channel* signal to noise ratio (SNR). Our results, which relate the channel SNR to peak signal to noise ratio (PSNR) will be useful for practical implementations of WZVC. Initial results have shown that the reconstructed WZ video quality in a wireless channel is markedly inferior in quality compared to noiseless or Gaussian noise channel. Therefore, improvement of video output quality was an essential goal in WZVC in a wireless environment.

Fundamental motivation behind the development of WZVC is to obtain a less complicated encoder, however, it comes at the expense of higher decoder complexity. Based on these fundamentals, we have investigated WZVC over a diverse array of multiple transmit-receive of wireless system, which has shown to improve the reconstructed video frames compared to that of a single wireless channel model. PSNR values of configuration of two-transmitting and four-receiving antennas approached to the PSNR value of noiseless channel at a lower value of channel SNRs. Since multiple input multiple output WZVC (MIMO-WZVC) performs better at the lower values of channel SNRs, it is postulated as a potential approach to overcome the wireless channel impairment in a WZVC. Additionally, MIMO-WZVC is suitable to use in wireless video sensor networks with low power sensors.

We examined the rate penalty in WZVC in a wireless channel environment, and derived expressions for the rate penalty of WZVC in AWGN ( $\epsilon_{AWGN}$ ) and fading channels ( $\epsilon_{Fading}$ ). Doing so, we have shown that the rate penalty increases with the channel hostility, which includes random channel noise and a fading coefficient. In addition, we have proposed WZVC with receiver diversity (WZVC-RD) which reduces the impact of rate penalty. Simulation results have shown that the rate penalty is reduced with the proposed WZVC-RD<sub>n</sub> for the same channel SNR, where  $n > 1$ ; i.e., WZVC-RD<sub>n</sub> shows improved PSNR at lower bit rates.

Next, we investigated the cross-correlation behavior at the decoder to control the transmitting rate at the encoder. The reconstruction of the WZ frame depends on the amount of distortion in the side information. We studied residual information and its relationship with the cross-correlation coefficient, and proposed a cross-correlation threshold (CCTH) technique at the decoder in order to control the transmitting parity bits. Additionally, we introduced an adaptive threshold (ATH) algorithm at the decoder, and used this to compare the different level of thresholds that is required to control the rate for different video sequences. By utilizing this technique, we obtained significant reduction in average rate with an improved PSNR compared to the 50% rate reduction method.

Finally, we have elaborated of developing a novel algorithm to dynamically estimate the correlation noise at the receiver which will generate soft-values of the side information. In this section, our contribution to WZVC research is by proposing a method for soft-value estimation of the side information for a continual varying video sequence as well as introducing the concept of *bit-based noise variance*.

## 7.2 Future Work

As an extension of this thesis, we propose the following possible directions for future research.

- Implementing WZVC in narrow-band wireless communications by transmitting key frames from parallel cameras; here, WZVC could be merged with H.264. Side information will be influenced by correlation noise and channel noise when it is transmitted via the wireless channel, thus the rate penalty and reconstruction will be a challenging future work.
- Rate control in terms of cross-correlation threshold performs better for slowly varying video sequences than for fast varying video sequences. Therefore it puts forth a challenge to investigate the group of picture (GOP) at the decoder with the available frames to improve the rate in fast varying video sequences in WZVC.

# Bibliography

- [1] T. Wiegand, G. Sullivan, G. Bjontegaard, and A. Luthra, “Overview of the H.264/AVC video coding standard,” *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 13, no. 7, pp. 560–576, 2003.
- [2] D. Slepian and J. Wolf, “Noiseless coding of correlated information sources,” *Information Theory, IEEE Transactions on*, vol. 19, no. 4, pp. 471–480, 1973.
- [3] A. Wyner and J. Ziv, “The rate-distortion function for source coding with side information at the decoder,” *Information Theory, IEEE Transactions on*, vol. 22, no. 1, pp. 1–10, 1976.
- [4] A. Aaron, R. Zhang, and B. Girod, “Wyner-Ziv coding of motion video,” in *Signals, Systems and Computers, 2002. Conference Record of the Thirty-Sixth Asilomar Conference on*, vol. 1, pp. 240–244, IEEE, 2002.
- [5] A. Aaron, S. Rane, R. Zhang, and B. Girod, “Wyner-Ziv coding for video: Applications

- to compression and error resilience,” in *Data Compression Conference, 2003. Proceedings. DCC 2003*, pp. 93–102, IEEE, 2003.
- [6] S. Pradhan and K. Ramchandran, “Distributed source coding using syndromes (DISCUS): Design and construction,” *Information Theory, IEEE Transactions on*, vol. 49, no. 3, pp. 626–643, 2003.
- [7] S. Pradhan and K. Ramchandran, “Enhancing analog image transmission systems using digital side information: A new wavelet-based image coding paradigm,” in *Data Compression Conference, 2001. Proceedings. DCC 2001.*, pp. 63–72, IEEE, 2001.
- [8] R. Puri, A. Majumdar, P. Ishwar, and K. Ramchandran, “Distributed video coding in wireless sensor networks,” *Signal Processing Magazine, IEEE*, vol. 23, no. 4, pp. 94–106, 2006.
- [9] R. Puri, A. Majumdar, and K. Ramchandran, “PRISM: a video coding paradigm with motion estimation at the decoder,” *Image Processing, IEEE Transactions on*, vol. 16, no. 10, pp. 2436–2448, 2007.
- [10] C.-C. Chiu, S.-Y. Chien, C.-h. Lee, V. S. Somayazulu, and Y.-K. Chen, “Hybrid distributed video coding with frame level coding mode selection,” in *Image Processing (ICIP), 2012 19th IEEE International Conference on*, pp. 1561–1564, IEEE, 2012.
- [11] J. Park, H. J. Shim, and B. Jeon, “Mobile to mobile video communication using dvc

- framework with channel information transmission,” in *Broadband Multimedia Systems and Broadcasting (BMSB), 2011 IEEE International Symposium on*, pp. 1–5, IEEE, 2011.
- [12] C. Li, J. Zou, H. Xiong, and C. W. Chen, “Joint coding/routing optimization for distributed video sources in wireless visual sensor networks,” *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 21, no. 2, pp. 141–155, 2011.
- [13] A. Aaron, S. Rane, and B. Girod, “Wyner-Ziv video coding with hash-based motion compensation at the receiver,” in *Image Processing, 2004. ICIP’04. 2004 International Conference on*, vol. 5, pp. 3097–3100, IEEE, 2004.
- [14] A. Aaron, D. Varodayan, and B. Girod, “Wyner-Ziv residual coding of video,” in *Proceedings Picture Coding Symposium*, Citeseer, 2006.
- [15] J. Ascenso, C. Brites, and F. Pereira, “Improving frame interpolation with spatial motion smoothing for pixel domain distributed video coding,” in *5th EURASIP Conference on Speech and Image Processing, Multimedia Communications and Services*, Smolenice, Slovak Republic, 2005.
- [16] C. Brites, J. Ascenso, and F. Pereira, “Studying temporal correlation noise modeling for pixel based Wyner-Ziv video coding,” in *Image Processing, 2006 IEEE International Conference on*, pp. 273–276, IEEE, 2006.

- [17] A. Aaron, E. Setton, and B. Girod, “Towards practical Wyner-Ziv coding of video,” in *Image Processing, 2003. ICIP 2003. Proceedings. 2003 International Conference on*, vol. 3, pp. III-869, IEEE, 2003.
- [18] C. Brites, J. Ascenso, and F. Pereira, “Improving transform domain Wyner-Ziv video coding performance,” in *Acoustics, Speech and Signal Processing, 2006. ICASSP 2006 Proceedings. 2006 IEEE International Conference on*, vol. 2, pp. II-II, IEEE, 2006.
- [19] M. Morbée, J. Prades-Nebot, A. Pizurica, and W. Philips, “Rate allocation algorithm for pixel-domain distributed video coding without feedback channel,” in *Acoustics, Speech and Signal Processing, 2007. ICASSP 2007. IEEE International Conference on*, vol. 1, pp. I-521, IEEE, 2007.
- [20] T. Sheng, G. Hua, H. Guo, J. Zhou, and C. Chen, “Rate allocation for transform domain Wyner-Ziv video coding without feedback,” in *Proceedings of the 16th ACM international conference on Multimedia*, pp. 701-704, ACM, 2008.
- [21] C. Brites and F. Pereira, “Encoder rate control for transform domain Wyner-Ziv video coding,” in *Image Processing, 2007. ICIP 2007. IEEE International Conference on*, vol. 2, pp. II-5, IEEE, 2007.
- [22] S. Alamouti, “A simple transmit diversity technique for wireless communications,” *Selected Areas in Communications, IEEE Journal on*, vol. 16, no. 8, pp. 1451-1458, 1998.

- [23] S. Alamouti and V. Tarokh, “Transmitter diversity technique for wireless communications,” Feb. 6 2001. US Patent 6,185,258.
- [24] V. Tarokh, H. Jafarkhani, and A. Calderbank, “Space-time block coding for wireless communications: performance results,” *Selected Areas in Communications, IEEE Journal on*, vol. 17, no. 3, pp. 451–460, 1999.
- [25] P. Parsons, “Defining cable television: Structuration and public policy,” *Journal of Communication*, vol. 39, no. 2, pp. 10–26, 1989.
- [26] D. Marpe, T. Wiegand, and G. Sullivan, “The H.264/MPEG4 advanced video coding standard and its applications,” *Communications Magazine, IEEE*, vol. 44, no. 8, pp. 134–143, 2006.
- [27] J. Ostermann, J. Bormans, P. List, D. Marpe, M. Narroschke, F. Pereira, T. Stockhammer, and T. Wedi, “Video coding with H.264/AVC: tools, performance, and complexity,” *Circuits and Systems Magazine, IEEE*, vol. 4, no. 1, pp. 7–28, 2004.
- [28] R. Collins, A. Lipton, T. Kanade, H. Fujiyoshi, D. Duggins, Y. Tsin, D. Tolliver, N. Enomoto, O. Hasegawa, P. Burt, *et al.*, *A system for video surveillance and monitoring*, vol. 102. Carnegie Mellon University, the Robotics Institute, 2000.
- [29] I. Akyildiz, T. Melodia, and K. Chowdury, “Wireless multimedia sensor networks: A survey,” *Wireless Communications, IEEE*, vol. 14, no. 6, pp. 32–39, 2007.

- [30] A. Aaron, S. Rane, E. Setton, B. Girod, *et al.*, “Transform-domain Wyner-Ziv codec for video,” in *Proc. SPIE Visual Communications and Image Processing*, pp. 520–528, 2004.
- [31] D. Rowitch and L. Milstein, “On the performance of hybrid FEC/ARQ systems using rate compatible punctured turbo (RCPT) codes,” *Communications, IEEE Transactions on*, vol. 48, no. 6, pp. 948–959, 2000.
- [32] L. Hanzo, T. Liew, B. Yeap, and J. Wiley, *Turbo coding, turbo equalisation and space-time coding*. Wiley Online Library, 2002.
- [33] C. Berrou, A. Glavieux, and P. Thitimajshima, “Near shannon limit error-correcting coding and decoding: Turbo-codes. 1,” in *Communications, 1993. ICC 93. Geneva. Technical Program, Conference Record, IEEE International Conference on*, vol. 2, pp. 1064–1070, IEEE, 1993.
- [34] K. Gracie and M. Hamon, “Turbo and turbo-like codes: Principles and applications in telecommunications,” *Proceedings of the IEEE*, vol. 95, no. 6, pp. 1228–1254, 2007.
- [35] C. Berrou and A. Glavieux, “Near optimum error correcting coding and decoding: Turbo-codes,” *Communications, IEEE Transactions on*, vol. 44, no. 10, pp. 1261–1271, 1996.
- [36] S. Benedetto and G. Montorsi, “Unveiling turbo codes: Some results on parallel con-

- catenated coding schemes,” *Information Theory, IEEE Transactions on*, vol. 42, no. 2, pp. 409–428, 1996.
- [37] C. Schlegel and L. Perez, *Trellis and turbo coding*, vol. 10. Wiley-IEEE Press, 2004.
- [38] A. Viterbi and J. Omura, *Principles of digital communication and coding*, vol. 286. McGraw-Hill New York, 1979.
- [39] M. Pinson and S. Wolf, “Comparing subjective video quality testing methodologies,” in *SPIE Video Communications and Image Processing Conference*, vol. 5150, pp. 573–582, 2003.
- [40] A. Webster, C. Jones, M. Pinson, S. Voran, and S. Wolf, “An objective video quality assessment system based on human perception,” in *SPIE Human Vision, Visual Processing, and Digital Display IV*, vol. 1913, pp. 15–26, 1993.
- [41] M. Rabbani and P. Jones, “Digital image compression techniques,” SPIE-International Society for Optical Engineering, 1991.
- [42] Z. Wang, H. Sheikh, and A. Bovik, “Objective video quality assessment,” *The handbook of video databases: design and applications*, pp. 1041–1078, 2003.
- [43] C. Brites and F. Pereira, “Correlation noise modeling for efficient pixel and transform domain Wyner-Ziv video coding,” *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 18, no. 9, pp. 1177–1190, 2008.

- [44] J. Slowack, S. Mys, J. Skorupa, P. Lambert, R. Van de Walle, and C. Grecos, “Accounting for quantization noise in online correlation noise estimation for distributed video coding,” in *Picture Coding Symposium, 2009. PCS 2009*, pp. 1–4, IEEE, 2009.
- [45] J. Ascenso, C. Brites, and F. Pereira, “Motion compensated refinement for low complexity pixel based distributed video coding,” in *Advanced Video and Signal Based Surveillance, 2005. AVSS 2005. IEEE Conference on*, pp. 593–598, IEEE, 2005.
- [46] F. Abrishamkar and E. Biglieri, “An overview of wireless communications,” in *Military Communications Conference, 1994. MILCOM’94. Conference Record, 1994 IEEE*, pp. 900–905, IEEE, 1994.
- [47] J. G. Proakis and D. G. Manolakis, *Digital communications*, vol. 3. McGraw-hill New York, 1995.
- [48] W. Stark and L. Milstein, “Wireless communications system architecture and performance,” *RF technologies for low power wireless communications*, pp. 9–38, 2001.
- [49]
- [50] P. Hoehner, “A statistical discrete-time model for the WSSUS multipath channel,” *Vehicle Technology, IEEE Transactions on*, vol. 41, no. 4, pp. 461–468, 1992.
- [51] R. Hoyt, “Probability functions for the modulus and angle of the normal complex variate,” *Bell Syst. Tech. J.*, vol. 26, pp. 318–359, 1947.

- [52] S. Rice, *Statistical properties of a sine wave plus random noise*. Bell Telephone Laboratories, 1948.
- [53] M. Nakagami, “The m-distribution-A general formula of intensity distribution of rapid fading,” *Statistical Method of Radio Propagation*, 1960.
- [54] P. Bello, “Characterization of randomly time-variant linear channels,” *Communications Systems, IEEE transactions on*, vol. 11, no. 4, pp. 360–393, 1963.
- [55] E. Biglieri, J. Proakis, and S. Shamai, “Fading channels: Information-theoretic and communications aspects,” *Information Theory, IEEE Transactions on*, vol. 44, no. 6, pp. 2619–2692, 1998.
- [56] M. Patzold and F. Laue, “Statistical properties of jakes’ fading channel simulator,” in *Vehicular Technology Conference, 1998. VTC 98. 48th IEEE*, vol. 2, pp. 712–718, IEEE, 1998.
- [57] D. Borah and B. Hart, “Frequency-selective fading channel estimation with a polynomial time-varying channel model,” *Communications, IEEE Transactions on*, vol. 47, no. 6, pp. 862–873, 1999.
- [58] J. Garcia-Frias, “Compression of correlated binary sources using turbo codes,” *Communications Letters, IEEE*, vol. 5, no. 10, pp. 417–419, 2001.

- [59] J. Garcia-Frias and Y. Zhao, "Compression of binary memoryless sources using punctured turbo codes," *Communications Letters, IEEE*, vol. 6, no. 9, pp. 394–396, 2002.
- [60] Y. Zhao and J. Garcia-Frias, "Data compression of correlated non-binary sources using punctured turbo codes," in *Data Compression Conference, 2002. Proceedings. DCC 2002*, pp. 242–251, IEEE, 2002.
- [61] J. Bajcsy and P. Mitran, "Coding for the Slepian-Wolf problem with turbo codes," in *Global Telecommunications Conference, 2001. GLOBECOM'01. IEEE*, vol. 2, pp. 1400–1404, IEEE, 2001.
- [62] P. Mitran and J. Bajcsy, "Coding for the Wyner-Ziv problem with turbo-like codes," in *Information Theory, 2002. Proceedings. 2002 IEEE International Symposium on*, p. 91, IEEE, 2002.
- [63] A. Liveris, Z. Xiong, and C. Georghiades, "Compression of binary sources with side information at the decoder using LDPC codes," *Communications Letters, IEEE*, vol. 6, no. 10, pp. 440–442, 2002.
- [64] A. Liveris, Z. Xiong, and C. Georghiades, "Joint source-channel coding of binary sources with side information at the decoder using IRA codes," in *Multimedia Signal Processing, 2002 IEEE Workshop on*, pp. 53–56, IEEE, 2002.
- [65] Q. Xu, V. Stankovic, A. Liveris, and Z. Xiong, "Distributed joint source-channel coding

- of video,” in *Image Processing, 2005. ICIP 2005. IEEE International Conference on*, vol. 2, pp. II-674, IEEE, 2005.
- [66] A. Majumdar and K. Ramchandran, “PRISM: an error-resilient video coding paradigm for wireless networks,” in *Broadband Networks, 2004. BroadNets 2004. Proceedings. First International Conference on*, pp. 478–485, IEEE, 2004.
- [67] S. Rane, P. Baccichet, and B. Girod, “Systematic lossy error protection of video signals,” *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 18, no. 10, pp. 1347–1360, 2008.
- [68] C. Tonoli, P. Migliorati, and R. Leonardi, “Error resilience in current distributed video coding architectures,” *Journal on Image and Video Processing*, vol. 2009, p. 8, 2009.
- [69] B. Vucetic and J. Yuan, *Turbo Codes Principles and Applications, 2000*. Kluwer Academic Publishers, Norwell, Massachusetts.
- [70] J. Cavers, “An analysis of pilot symbol assisted modulation for Rayleigh fading channels [mobile radio],” *Vehicular Technology, IEEE Transactions on*, vol. 40, no. 4, pp. 686–693, 1991.
- [71] A. Wyner, “The rate-distortion function for source coding with side information at the decoder\3-II: General sources,” *Information and Control*, vol. 38, no. 1, pp. 60–80, 1978.

- [72] K. Thambu, X. Fernando, and L. Guan, “Dynamic algorithm for correlation noise estimation in distributed video coding,” in *Proceedings of SPIE*, vol. 7543, p. 754310, 2010.
- [73] T. Kuganeswaran, X. Fernando, and L. Guan, “A pixel based dynamic algorithm for correlation noise estimation in distributed video coding,” in *Sarnoff Symposium*, 2009.
- [74] K. Thambu, X. Fernando, and L. Guan, “Transmit rate reduction of Wyner-Ziv video coding (WZVC) using multiple uncorrelated wireless receivers,” in *Communications (QBSC), 2012 26th Biennial Symposium on*, pp. 174–177, IEEE, 2012.
- [75] R. G. Gallager, *Information theory and reliable communication*. Wiley, 1968.
- [76] R. Yates and D. Goodman, *Probability and stochastic processes*. Wiley, 1999.
- [77] F. Pereira, “Distributed video coding: Basics, main solutions and trends,” in *Multimedia and Expo, 2009. ICME 2009. IEEE International Conference on*, pp. 1592–1595, IEEE, 2009.