INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI

A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor MI 48106-1346 USA
313/761-4700 800/521-0600

Robust Video Communication Over Wireless Packet Networks

by

Gubbi R. Rajugopal, B. E., M. Tech.

A thesis submitted to the
Faculty of Graduate Studies and Research
in partial fulfillment of
the requirements for the degree of

Doctor of Philosophy

Ottawa-Carleton Institute for Electrical Engineering
Faculty of Engineering
Department of Systems and Computer Engineering
Carleton University

April 1997

© Copyright

1997, Gubbi R. Rajugopal

National Library
of Canada

Acquisitions and
Bibliographic Services

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque nationale
du Canada

Acquisitions et
services bibliographiques

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

Our file Notre référence

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-22175-X

The undersigned hereby recommend to
the Faculty of Graduate Studies and Research
acceptance of the thesis,

Robust Video Communication Over Wireless Packet Networks

submitted by

Gubbi R. Rajugopal, B. E., M. Tech.

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

Chair, Department of Systems and Computer Engineering

Professor Samy A. Mahmoud, Thesis Co-Supervisor

Professor Roshdy H. M. Hafez, Thesis Co-Supervisor

Professor Y. Wu, External Examiner

Carleton University

April, 1997

Abstract

This thesis is an investigation of a high performance universal video communication system for mobile (indoor or outdoor) environment with an ATM backbone network. This work is in line with the upcoming MPEG-4 proposals which is being planned for future systems. Channel and Bandwidth constraints in mobile environment limit the achievable data rate. However, the end users demand good quality video and would prefer gracefully degrading video quality with decrease in available Bandwidth. The two main issues addressed are video compression and channel adaptation of the system.

A good source coder is required to provide high compression gain and gracefully degrading video quality with decrease in the available Bandwidth. Also it is known that variable bit rate (VBR) source coders outperform constant bit rate source coders under network transport conditions. Hence investigation of combination of wavelet transform, multi-resolution motion estimation and zero tree coder is a good choice for VBR video compression at low bit rates. Robust design of zero tree coder is proposed to make the source coder error resistant under hostile environment.

We suggest two stages of channel coding, with unequal error protection at the inner stage to better protect the data of higher importance. Varying channel conditions demand dynamically changing error protection for better use of bandwidth. A source level rate controller is proposed to enable adaptation to the channel conditions and to provide VBR video. The rate controller needs to efficiently analyze the channel conditions and adjust the system parameters for best performance. It is suggested to sense the channel at the receiver and send it as feedback. This information is further processed and used by the rate controller to dynamically adapt the system. Outer channel coding is suggested to be as per the standard air interface specifications such as in CODIT, DECT etc.

Another important system requirement fulfilled is the requirement of low overhead when the source coder robustness or the channel adaptation are not required. This will be the case at low BER (or CLR) environments. The Robust design of source coder and dynamic channel adaptation of the system without much overhead enables its application in a variety of channel conditions like narrowband, broadband, indoor and outdoor.

Acknowledgements

First and the foremost, I thank God for the love and affection showered on me and, for holding my hand and leading my way all through. I beg Him for His mercy for the rest of my life.

I would like to express my gratitude to my supervisors, Prof. Roshdy Hafez and Prof. Samy Mahmoud, for their guidance and encouragement throughout this research work. Their constructive criticism and practical views have improved the quality of this research work.

I am grateful to Prof. David Cole and Andre Mainguy, Communications Research Centre, for providing me with useful documents and also for the useful discussions I had with them. Also, I take this opportunity to thank Naren, Janet and the rest of the department staff whose assistance helped me throughout this thesis work.

The financial assistance provided by the Canadian Government through the *Canadian Commonwealth Scholarship and Fellowship Program* and the Telecommunication Research Institute of Ontario is gratefully acknowledged. I am grateful to the Ministry of Human Resources and Development, Government of India, and the Commonwealth Scholarship committee for recommending me to the Commonwealth Scholarship.

I wish to thank my wife, Prabha, and my little darling son, Madhu, for their unconditional love and great understanding extended to me during this preoccupation. I express my deepest gratitude to my friends Umesha, Rudra and Seena for their love and constant encouragement in all my endeavors.

Last but not the least, I thank all my friends and colleagues in the *Wireless Communications and Simulations Lab* who have extended their help whenever I was in need.

*karmanyevādhikāraṣṭe mā phaleṣu kadācana
mā karmaphalaheturbhūr mā te saṅgo' stvakarmaṇi*

**You have a right to perform your prescribed duty, but
you are not entitled to the fruits of action. Never
consider yourself the cause of the results of your
activities, and never be attached to not doing your duty.**

- Bhagavad-Gītā, Chapter 2

Table of Contents

Abstract	iii
Acknowledgements	iv
Table of Contents	vi
List of Figures	xii
List of Tables	xviii
List of Abbreviations	xix
List of Mathematical Symbols and Notation	xxiii
Chapter 1: Introduction.....	1
1.1 Motivation and Objectives.....	1
1.2 Problem Description and Proposed Solution.....	2
1.3 Thesis Contributions.....	3
1.4 Thesis Organization	4
Chapter 2: Review of Video Compression	6
2.1 Introduction.....	6
2.2 Source Coding Theory	7
2.2.1 Lossless and Lossy Coders.....	7
2.2.2 Distortion.....	8
2.3 Image Transforms	11
2.3.1 Discrete Cosine Transform.....	11
2.3.2 Wavelet Transform.....	12

2.4	Motion Estimation	17
2.4.1	Block Based and Pixel Based Motion Estimation	17
2.4.2	Multi-Resolution Motion Estimation	20
2.5	Quantization	22
2.5.1	Background	22
2.5.2	Zero Tree Quantization	22
2.5.2.1	Significance Map Encoding	23
2.5.2.2	Successive Approximation Quantization	26
2.5.2.3	Reconstruction	27
2.5.2.4	Symbol Level Encoding	28
2.6	Existing Standards	28
2.6.1	H.261 and H.263 Standards	28
2.6.2	MPEG Standard	29
2.6.3	Comments on Video Coding Standards	32
2.7	Future Trends	33
2.8	Summary	34
Chapter 3:	Review of Mobile Communications	36
3.1	Introduction	36
3.2	Channel Characterization	36
3.3	Techniques to Improve Performance at the Receiver	39
3.3.1	Diversity	39
3.3.2	Channel Coding	41
3.3.3	Comments	42

3.4	Channel Modeling.....	42
3.4.1	Network Channel.....	44
3.4.1.1	Narrowband Wireless Model	44
3.4.1.2	Broadband Wireless Model.....	45
3.5	Standards and Systems.....	47
3.5.1	Present Systems	47
3.5.1.1	Cellular Systems	48
3.5.1.2	Cordless Systems	50
3.5.1.3	Wireless Data Systems.....	51
3.5.2	Future Systems	51
3.6	Summary	52
Chapter 4:	Problem Definition.....	53
4.1	Introduction.....	53
4.2	Problem Definition.....	53
4.3	Overview of the Proposed System.....	56
4.3.1	Video Coder.....	57
4.3.1.1	Robust Quantization.....	57
4.3.1.2	Stream Classification	59
4.3.2	Channel Protection	60
4.3.3	Rate Controller	61
4.3.4	Test Sequences	62
4.4	Summary	62

Chapter 5:	Video Coder	64
5.1	Introduction.....	64
5.2	Video Coder Configurations.....	64
5.2.1	Wavelet Transform.....	65
5.2.2	Multiresolution Motion Estimation	65
5.2.3	Zero Tree Coder	68
5.3	Performance Evaluation.....	68
5.3.1	Experimental Setup	68
5.3.2	Simulation Results.....	69
5.4	Summary	72
Chapter 6:	Robust Quantization.....	77
6.1	Introduction.....	77
6.2	Robust Quantizer Designs.....	77
6.2.1	Sequential Partitioning	78
6.2.2	Orthogonal Partitioning.....	80
6.3	Performance Evaluation.....	82
6.3.1	Experimental Setup	82
6.3.2	Simulation Results.....	85
6.3.2.1	Performance in Narrowband Channel.....	85
	MRME Based Video Coder	85
	3D-WT Based Video Coder	90
6.3.2.2	Performance in Broadband Channel	90
	MRME Based Video Coder	90
	3D-WT Based Video Coder	101

6.3.2.3	Comparison with H.263 Based Coder.....	101
6.4	Summary	101
Chapter 7:	Adaptive Dynamic Rate Control	109
7.1	Introduction.....	109
7.2	Dynamic Rate Controller	109
7.2.1	Channel Measurement and Classification	110
7.2.2	Dynamic Bit Budget Allocation	111
7.2.2.1	Information Spread	111
7.2.2.2	Bit Budget Variations.....	111
7.2.3	Dynamic Scheduling of UEP.....	112
7.3	Performance Evaluation.....	116
7.3.1	Experimental Setup	116
7.3.2	Simulation Results.....	118
7.3.2.1	Adaptation Window length	118
7.3.2.2	Rate control.....	121
	Narrowband Channel	121
	Broadband Channel	121
	Performance at 16 Kbps and 32 Kbps ...	130
	Color Video Transmission	130
7.4	Summary	130
Chapter 8:	Conclusions.....	142
8.1	Conclusions.....	142
8.2	Suggestions for Future Research	144

References.....	146
Appendix - A: Arithmetic Entropy Coding	159
A.1 Principles of Arithmetic Coding	159
A.2 Implementation	160
Appendix - B: Rate Compatible Punctured Convolutional Codes.....	162
B.1 Punctured Codes Derived from Half Rate Codes	162
B.2 Selection of Constraint Length	163
Appendix - C: ATM Protocol.....	165
Appendix - D: Computational Complexity of Video coder	169
D.1 Wavelet Transform	169
D.2 Multiresolution Motion Estimation	170
D.3 Zero Tree Quantization	172

List of Figures

Figure 2-1:	Rate distortion function for a Gaussian source.....	10
Figure 2-2:	Illustration of one stage of one dimensional DWT.....	14
Figure 2-3:	A four stage one dimensional wavelet decomposition	14
Figure 2-4:	2D wavelet filtering using 1D filters	15
Figure 2-5:	A three level wavelet decomposition of picture frame	16
Figure 2-6:	Motion search illustration for 1/4th pixel accuracy	19
Figure 2-7:	Illustration of multiresolution motion estimation	21
Figure 2-8:	Parent-child dependency and scanning order of coefficients	24
Figure 2-9:	Zero tree construction (dominant pass) for a given threshold T.....	25
Figure 2-10:	Encoder - decoder structure for H.261/H.263 standard	30
Figure 2-11:	Simplified block diagram of MPEG encoder.....	31
Figure 2-12:	Relation between different frame types of MPEG standard	32
Figure 2-13:	Comparison of image compression techniques.....	35
Figure 3-1:	Different types of channels encountered in radio communication	43
Figure 3-2:	Mapping of user bit rates to chip rates in CODIT environment	46
Figure 3-3:	Generic transmission scheme in CODIT environment	46
Figure 4-1:	Concept of personal communication	55
Figure 4-2:	System block diagram for robust video communication	58

Figure 5-1:	Structure of video encoder.....	66
Figure 5-2:	Comparative performance of MRME schemes at 85 Kbps.	70
Figure 5-3:	Comparative performance of MRME schemes at 128 Kbps.	70
Figure 5-4:	Comparative performance of MRME schemes at 160 Kbps.	71
Figure 5-5:	Comparative performance of MRME schemes at 200 Kbps.	71
Figure 5-6:	Comparative performance of MRME algorithms on at 300 Kbps.....	73
Figure 5-7:	Performance curves for 3D wavelet transform at low bit rates	73
Figure 5-8:	Comparative performance of 3D-WT and MRME1 schemes at 64 Kbps.....	74
Figure 5-9:	Bit budget allocation for 3D-WT and MRME1 schemes at 64 Kbps.	74
Figure 5-10:	Comparative performance of MRME schemes at 85 Kbps for <i>carphone</i> sequence.....	75
Figure 6-1:	Robust zero tree quantizer design.....	79
Figure 6-2:	Experimental setup to compare different zero tree coder designs	83
Figure 6-3:	Performance comparison of the zero tree coders in MRME video configuration at 32 Kbps under narrowband channel conditions.....	86
Figure 6-4:	Performance comparison of the zero tree coders in MRME video configuration at 64 Kbps under narrowband channel conditions.....	87
Figure 6-5:	Performance comparison of the zero tree coders in MRME video configuration at 128 Kbps under narrowband channel conditions	88
Figure 6-6:	Performance comparison of the zero tree coders in MRME video configuration at 256 Kbps under narrowband channel conditions	89
Figure 6-7:	Frame #15 of <i>miss america</i> sequence: Comparison between the	

	original and the reconstructed output of different zero tree coder schemes (MRME configuration) under narrowband channel conditions with 10^{-2} CLR at 64 Kbps.....	91
Figure 6-8:	Performance comparison of the zero tree coders in 3D-WT video configuration at 32 Kbps under narrowband channel conditions	92
Figure 6-9:	Performance comparison of the zero tree coders in 3D-WT video configuration at 64 Kbps under narrowband channel conditions	93
Figure 6-10:	Performance comparison of the zero tree coders in 3D-WT video configuration at 128 Kbps under narrowband channel conditions	94
Figure 6-11:	Performance comparison of the zero tree coders in 3D-WT video configuration at 256 Kbps under narrowband channel conditions	95
Figure 6-12:	Frame #73 of <i>miss america</i> sequence: Comparison between the original and the reconstructed output of different zero tree coder schemes (3D-WT configuration) under narrowband channel conditions with 10^{-2} CLR at 64 Kbps.....	96
Figure 6-13:	Performance comparison of the zero tree coders in MRME video configuration at 32 Kbps under broadband channel conditions	97
Figure 6-14:	Performance comparison of the zero tree coders in MRME video configuration at 64 Kbps under broadband channel conditions	98
Figure 6-15:	Performance comparison of the zero tree coders in MRME video configuration at 128 Kbps under broadband channel conditions	99
Figure 6-16:	Performance comparison of the zero tree coders in MRME video configuration at 256 Kbps under broadband channel conditions	100
Figure 6-17:	Performance comparison of the zero tree coders in 3D-WT video configuration at 32 Kbps under broadband channel conditions	102
Figure 6-18:	Performance comparison of the zero tree coders in 3D-WT video	

	configuration at 64 Kbps under broadband channel conditions	103
Figure 6-19:	Performance comparison of the zero tree coders in 3D-WT video configuration at 128 Kbps under broadband channel conditions	104
Figure 6-20:	Performance comparison of the zero tree coders in 3D-WT video configuration at 256 Kbps under broadband channel conditions	105
Figure 6-21:	Frame #74 of <i>miss america</i> sequence: Comparison between the original and the reconstructed output of different zero tree coder schemes (3D-WT configuration) under broadband channel conditions with 5×10^{-3} BER at 64 Kbps.	106
Figure 6-22:	Comparing performance of RZCO with that of H.26P/M	107
Figure 7-1:	Flow chart for dynamic rate control.....	113
Figure 7-2:	Experimental setup to evaluate the adaptive rate controlled video communication system.	117
Figure 7-3:	Comparison of performance of different window lengths in dynamic rate control at 64 Kbps under narrowband channel conditions of 10^{-2} CLR.....	119
Figure 7-4:	Comparison of performance of different window lengths in dynamic rate control at 64 Kbps under broadband channel conditions of 5×10^{-3} BER.....	120
Figure 7-5:	Performance comparison of the RZCO scheme with and without dynamic rate control for MRME configuration at 64 Kbps under narrowband channel conditions	122
Figure 7-6:	Performance comparison of the RZCO scheme with and without dynamic rate control for MRME configuration at 128 Kbps under narrowband channel conditions	123
Figure 7-7:	Performance comparison of the RZCO scheme with and without	

	dynamic rate control for 3D-WT configuration at 64 Kbps under narrowband channel conditions	124
Figure 7-8:	Performance comparison of the RZCO scheme with and without dynamic rate control for 3D-WT configuration at 128 Kbps under narrowband channel conditions	125
Figure 7-9:	Performance comparison of the RZCO scheme with and without dynamic rate control for MRME configuration at 64 Kbps under broadband channel conditions	126
Figure 7-10:	Performance comparison of the RZCO scheme with and without dynamic rate control for MRME configuration at 128 Kbps under broadband channel conditions	127
Figure 7-11:	Performance comparison of the RZCO scheme with and without dynamic rate control for 3D-WT configuration at 64 Kbps under broadband channel conditions	128
Figure 7-12:	Performance comparison of the RZCO scheme with and without dynamic rate control for 3D-WT configuration at 128 Kbps under broadband channel conditions	129
Figure 7-13:	Visual quality comparisons of RZOC and DRC-RZOC in MRME configuration at 64 Kbps under 10^{-2} CLR (narrowband) and 5×10^{-3} BER (broadband) channel conditions	131
Figure 7-14:	Visual quality comparisons of RZOC and DRC-RZOC in 3D-WT configuration at 64 Kbps under 10^{-2} CLR (narrowband) and 5×10^{-3} BER (broadband) channel conditions	132
Figure 7-15:	Performance of the DRC-RZCO scheme in MRME configuration at 16 Kbps and 32 Kbps under narrowband channel conditions	133
Figure 7-16:	Performance of the DRC-RZCO scheme in 3D-WT configuration at 16 Kbps and 32 Kbps under narrowband channel conditions	134

Figure 7-17:	Performance of the DRC-RZCO scheme in MRME configuration at 16 Kbps and 32 Kbps under broadband channel conditions.....	135
Figure 7-18:	Performance of the DRC-RZCO scheme in 3D-WT configuration at 16 Kbps and 32 Kbps under broadband channel conditions.....	136
Figure 7-19:	Visual quality performance of DRC-RZOC in MRME and 3D-WT configurations (left and right columns respectively) at 16 Kbps and 32 Kbps under narrowband channel condition of 10^{-2} CLR.....	137
Figure 7-20:	Visual quality performance of DRC-RZOC in MRME and 3D-WT configurations (left and right columns respectively) at 16 Kbps and 32 Kbps under broadband channel condition of 5×10^{-3} BER.....	138
Figure 7-21:	Performance of DRC-RZCO for color video transmission at 64 Kbps under narrowband channel conditions.	139
Figure 7-22:	Performance of DRC-RZCO for color video transmission at 64 Kbps under broadband channel conditions.	140
Figure A-1:	Arithmetic encoding process for the message eaii!.	161
Figure B-1:	Procedure to obtain $(n-1)/n$ rate convolution code from half rate code..	164
Figure C-1:	ATM cell structure: AAL-2 format.....	166
Figure C-2:	Header of ATM cell at (a) User-Network interface (UNI) and (b) Network Node interface (NNI) and (c) Lookup Table at the switching unit.....	166
Figure C-3:	Virtual path and virtual channel connections.....	167

List of Tables

Table 6-1:	UEP schedule for different channel conditions	84
Table A-1:	Fixed model example for alphabet set {a, e, i, o, u, !}	161

List of Abbreviations

1D-WT	One Dimensional Wavelet Transform
2D-WT	Two Dimensional Wavelet Transform
3D-WT	Three Dimensional Wavelet Transform
AAL	ATM Adaptation Layer
ADPCM	Adaptive Differential Pulse Code Modulation
AMPS	Advanced Mobile Phone Service
ARQ	Automatic Repeat reQuest
ATDMA	Advanced Time Division Multiple Access
ATM	Asynchronous Transfer Mode
BER	Bit Error Rate
BZC	Basic Zero tree Coder
CDMA	Code Division Multiple Access
CELP	Code Excited Linear Prediction
CLP	Cell Loss Priority
CLR	Cell Loss Rate
CODIT	COde DIvision Testbed
CRC	Cyclic Redundancy Check
dB	deciBel
DCT	Discrete Cosine Transform
DECT	Digital European Cordless Telecommunications

DM	Delta Modulation
DPCM	Differential Pulse Code Modulation
DRC	Dynamic Rate Control
DS-CDMA	Direct Sequence Code Division Multiple Access
DSP	Digital Signal Processor
DWT	Discrete Wavelet Transform
EEP	Equal Error Protection
FDD	Frequency Division Duplex
FEC	Forward Error Correction
FIR	Finite Impulse Response
FLOP	FLoating point OPerations
FM	Frequency Modulation
FSK	Frequency Shift Keying
GOF	Group Of Frames
GSM	Global System for Mobile Communications
HEC	Header Error Correction
HVS	Human Visual System
IFC	Intra-Frame Coding
ISO	International Standard Organisation
KL	Karhunen-Loeve
MAD	Minimum Absolute Distance
MAVT	Mobile Audio Visual Terminal
MCC	Maximum Cross Correlation
MMSE	Minimum Mean Square Error

MPEG	Moving Picture Encoding Group
MRME	Multi-Resolution Motion Estimation
MSE	Mean Square Error
NMSE	Normalised Mean Square Error
PBX	Private Branch eXchange
PCM	Pulse Coded Modulation
PCS	Personal Communication System
PDC	Personal Digital Cellular
PN	Pseudo-random Noise
PSNR	Peak Signal to Noise Ratio
QCIF	Quarter Common Interface Format
QOS	Quality Of Service
RACE	Research and technology development in Advanced Communications technologies in Europe
RCPC	Rate Compatible Punctured Convolutional codes
REL P	Residual Excited Linear Prediction
RF	Radio Frequency
RZCO	Robust Zero tree Coder based on Orthogonal partitioning
RZCS	Robust Zero tree Coder based on Sequential partitioning
SAQ	Successive Approximation Quantization
SNR	Signal to Noise Ratio
TACS	Total Access Communication System
TDD	Time Division Duplex
TDMA	Time Division Multiple Access
UEP	Unequal Error Protection

VBR	Variable Bit Rate
VQ	Vector Quantizer
VSELP	Vector Sum Excited Linear Prediction
W-CDMA	Wideband Code Division Multiple Access
WACS	Wireless Access Communication systems
ZTQ	Zero Tree Quantization

List of Mathematical Symbols and Notation

$\delta()$	Delta function
$\Delta(x, y)$	Correction to motion vector
ψ	Mother wavelet
σ_x^2	Variance of the random variable x
τ_k	Time delay at k th multipath component
τ_{NF}	Instantaneous non-fade duration
$\bar{\tau}_{NF}$	Mean non-fade duration
$\theta_k(t)$	Phase difference at k th multipath component
a	Dilation constant
a_k	Amplitude of k th multipath component
b	Translation constant
$c(n_x, n_y)$	Discrete cosine transform matrix
D	Distortion
E	Expectation
$f(x)$	Input signal
$g(k)$	Low pass filter coefficient
$g^{-1}(k)$	Inverse low pass filter coefficient
H	Entropy
$h(k)$	High pass filter coefficient

$h^{-1}(k)$	Inverse high pass filter coefficient
$h(t, \tau)$	Channel impulse response
I_k	Information content in k th symbol
k_i	Number of information bits in a block
L	Number of stages of decomposition in wavelet transform
m	Dilation index
n	Translation index
n_c	Number of bits in a channel coded block
$N(\tau)$	Number of multipath components
N_d	Size of discrete cosine transform matrix
N_x	Size of image or video frame along X axis
N_y	Size of image or video frame along Y axis
N_s	Number of symbols
p_k	Probability of occurrence of k th symbol
R_c	Channel coder rate
R_D	Rate distortion function
t	time
T_i	Threshold at i th iteration
$V(x, y)$	Motion vector
x_g	Gaussian random variable
x_{ij}	Two dimensional input signal
x_r	Reconstructed form of the Gaussian random variable
y_{ij}	Reconstructed two dimensional signal
w	Wavelet coefficient

Chapter 1

Introduction

1.1 Motivation and Objectives

This research finds motivation in the worldwide quest to achieve the universal portable and personal communications. In order to realize this vision a high performance, robust communication system must be designed to complement current radio systems into a unified architecture. There are several issues to be addressed and hence there is scope for enormous research work.

The goal of personal communication becomes even more complicated when features like user mobility and system mobility are desired. This brings out the worst of channels into picture. Adding to this, for universal communication it is required to adhere to a universal network protocol conditions. Thus the robust communication under wireless packet environment becomes the core part of any reliable communication system. Further, video is becoming the most convincing and most sought after means of communication. This is especially true when it accompanies speech signals to communicate the information. This is leading to a colorful thought of wireless multimedia. However it is well known that the achievable user data rates are limited under wireless environment. Hence this research concentrates on low bit rate, robust video communication over wireless packet networks as a step towards making the wireless multimedia a reality. The main class of applications kept in view in this thesis are the video conference, video phones and slow motion video surveillance.

Two key objectives of this research work are reliable video communication and a

minimum of acceptable video quality. The other objectives include catering to user requirements, network and channel-characteristics and quality of service.

1.2 Problem Description and Proposed Solution

The important considerations in this work are:

- Robust video communication over wireless packet networks.
- A minimum of acceptable video quality.
- Variable bit rate communication depending on the channel conditions and the quality of service restrictions. This necessitates the system to be capable of learning and adapting to the channel conditions through channel sensing and be able to jointly optimize the source and channel coders.
- low overhead in ruggedizing the source coder and channel adaptation.
- Smooth bit rate variations from frame to frame to avoid the requirement of bit rate smoothing at the network level.
- Mean bit rate, peak bit rate and duration of peak bit rates adaptable to the network connection agreements.

The channel experienced by the received signal could be indoor or outdoor, narrowband or broadband depending on the user and terminal location and speed. This channel can vary drastically from time to time depending on the mobility of the user and changes in surrounding environment. Thus the proposed system employs an error resilient video coder and channel adaptation of the system parameters to combat the changes in channel conditions.

The video coder is based on a combination of wavelet transform, multiresolution motion estimation and zero tree quantization in order to achieve good compression gain, smooth bit rate variations and more importantly, a gracefully degrading video quality with decrease in available bit budget. Robust zero tree quantizer designs, along with two stages of error protection, are proposed to combat the channel errors. The channel adaptation of the system is achieved by making use of the feedback information on channel conditions from the receiver and thereby dynamically varying the source rate and the channel protection for robust communication.

1.3 Thesis Contributions

The contributions of this research are highlighted below:

- Studying and designing of robust video communication system for wireless packet networks.
- Devising a scheme for adaptive rate control which learns the channel characteristics and adapts the system to the current channel conditions so as to make the system robust against the channel errors and hence improve the system performance.
- Deriving a robust, error resilient zero tree coder design based on orthogonal partitioning to combat uncorrected channel errors.
- Showing that the sequential or symbol level partitioning of zero tree coder output does not perform reliably under noisy channel conditions.
- Presenting the idea of sensing the channel at the source decoder level and its feedback to the source encoder for channel adaptation. This includes the simple way suggested to measure channel in terms of the data available at the source level.
- Suggesting and evaluating two stage channel coding and interleaving, to combat burst errors in mobile channel and cell losses in ATM channel, without incurring extra delay.
- Presenting the idea of data classification and dynamically changing UEP schedule, as per the channel variations, to improve the bandwidth efficiency.
- Presenting the idea of spreading the information over several frames in order to make the output bit rate vary smoothly and hence improve the robustness of the system.
- Establishing that the MRME1 scheme is the best candidate for low bit rate applications. This is a new result contradicting those reported previously in [151].
- Presenting a new MRME scheme, MRME5, for motion estimation in subband coding.
- Providing a complete system evaluation for both 3D-WT and MRME configurations of the video coder under narrowband and broadband channel conditions.
- Demonstrating the use of zero tree coder to provide a simple bit rate control for smooth bit rate variations, apart from being used as a good compression scheme.

1.4 Thesis Organization

The two main issues addressed in this thesis are video compression and communication over mobile channel. Various aspects of video compression, the presently available video coders and the existing video coding standards are reviewed in chapter 2. The mobile channel characterization, ways and means to improve signal reception quality, mobile channel models and the existing mobile standards are reviewed in chapter 3.

A detail description of the problem chosen and an overview of the proposed solution are discussed in chapter 4. The low bit rate video compression and robust video communication makes the core part of the problem. The proposed solution involves error resilient designs of video coder and channel adaptation of the system in order to provide improved error protection at severe channel conditions and to provide better reconstruction quality when the channel conditions are favorable. This also involves classifying the compressed video into streams of different importance and providing a dynamically changing unequal error protection to achieve a minimum of acceptable video quality at the receiver.

The video coder employs two dimensional wavelet transform for spatial redundancy removal. The temporal redundancy can be exploited either through motion estimation-compensation techniques or a third dimension wavelet transform in the temporal domain. The video coder configurations using multiresolution motion estimation (MRME) and three dimensional wavelet transform (3D-WT) are discussed in chapter 5. The simulations are carried out to evaluate and compare the performance at different bit rates. The results show that the MRME1 scheme is best suited for low bit rate applications as it requires the lowest overhead. It is also shown that the 3D-WT configuration performs better than the motion estimation configuration but comes with the price of increased delay.

The error resilient zero tree quantizer designs are proposed in chapter 6. These new designs are based on sequential (quantizer symbol level) and orthogonal (wavelet coefficient level) partitioning. The simulations are performed and the subjective and objective performances are compared with the basic zero tree quantizer under narrowband and broadband channel conditions. Comparisons with H.263 based coder is also provided. The results show that the new design based on the orthogonal partitioning performs

superior compared to the rest of the designs and the performance is consistent under noisy channel conditions.

Varying channel conditions necessitate the channel adaptation of the system in order to improve the bandwidth efficiency of the system. This is discussed in detail in chapter 7. The channel measurement at the receiver involves indexing the ATM cells and measuring the cell loss rate at the receiver. This measured information is fed back to transmitter where it is used to dynamically allocate the bit budget to each of the compressed video streams and also to decide the degree of channel protection to be provided based on the channel conditions, available bit budget and the importance of the stream. The flow chart for the proposed dynamic rate control scheme is also provided. The simulations are performed and the results are compared with non-rate-controlled situation at different bit rates and channel conditions. It is shown that the dynamic rate controlling improves the system performance and its robustness against the channel variations.

In chapter 8 the thesis is concluded and some suggestions for future research are presented.

Arithmetic coding, rate punctured convolutional coder and the ATM protocol are briefly introduced in Appendix A, Appendix B and Appendix C respectively. In Appendix D the computational complexity analysis of the individual components of video coder is presented. The purpose of this appendix is to give a rough estimation of the video coder complexity. It is intended to show that this coder, though more complex than the DCT schemes, is realizable, unlike the fractals and the model based coders, using the present day DSP technology.

Chapter 2

Review of Video Compression

2.1 Introduction

The demand for video compression [115], [104], [76], [117], [6], [8] is rapidly increasing because of revolution in the fields of personal workstations and multimedia, and the raising needs of visual communication. Popular video compression schemes combine Discrete Cosine Transform (DCT) based image coder with motion estimation to achieve high compression. International Standard Organization (ISO) has standardized such schemes, via MPEG, H.261 and H.263 standards for public usage [8], [41], [15], [143]. Recently multiresolution schemes like Discrete Wavelet Transform (DWT) have emerged as efficient schemes for image and video compression and these have been found to achieve better picture quality, especially at low bit rates. The images transformed using multiresolution transforms possess hierarchically related correlation properties across different resolutions. Thus there are a whole new set of motion estimation and quantization techniques designed specifically to exploit these correlation properties. This chapter gives a review of different schemes of video compression and existing standards.

The chapter is organized as follows. The basic theory of source coding, source entropy and distortion are reviewed in section 2.2. The transforms and their applications are dealt with in section 2.3. Two important transforms, DCT and DWT, and their properties are discussed in details.

Motion estimation is essential in exploiting the temporal correlation. The block based and multiresolution approaches for motion estimation are reviewed in section 2.4. The

quantization procedure is reviewed in section 2.5 with the emphasis on zero tree quantization. Existing standards like H.261, H.263 and MPEG are reviewed in section 2.6. All these standards are DCT based and hence come with the disadvantage of *blocky noise* at low bit rates. Future trends in the video compression are discussed in section 2.7 where we brief the ideas of fractal coding and model based image coding. Finally, section 2.8 summarizes the discussion on video compression.

2.2 Source Coding Theory

Signals like speech, image and video have high amount of correlation leading to redundancy in their representation. *Source coding* [76] reduces the redundancy in such signals and represents them using smaller amount of data bits. This is also termed as *source compression*. Hence the source coding theory deals with the compression of data generated by an *information* source that sends out a sequence of symbols chosen from a finite set. The source coding can be of two types; *lossless* and *lossy*.

2.2.1 Lossless and Lossy Coders

The achievable compression gain in lossless coder is always limited by the actual average information generated by the source which is also termed as *entropy* of the source. Considering a source which generates a discrete set of independent symbols r_k , with probabilities p_k , $k = 1, 2 \dots N_s$, the information content in r_k is defined as,

$$I_k = -\log_2(p_k) \quad \text{where } 0 \leq p_k \leq 1 \quad \text{and} \quad \sum_{k=1}^L p_k = 1$$

The entropy of the source is defined as,

$$H = -\sum_{k=1}^{N_s} p_k \log_2(p_k) \quad \text{bits per symbol} \quad (2.1)$$

Note that both I_k and H are non-negative. The examples of lossless coders are Huffman coders [65], Arithmetic coders [146] (see Appendix-A also) etc., which are broadly classified as entropy coders.

The lossy coders are those which compromise on loss of information to achieve higher

compression gain. Most of the present day lossy coders exploit the human audio-visual system (HVS) in coding only a selected part of the information. For example, HVS is least sensitive to the loss of information at high frequency region. This is exploited by the coders by allocating most of the available bit budget to the information at low frequency region. Some coders go to the extent of not coding complete or good part of high frequency region. The examples of lossy coders [76] are Pulse Coded Modulation (PCM), differential PCM (DPCM), Delta modulation (DM) and Vector Quantizers (VQ).

The lossy coders are widely used because of their high compression gain. However, the lossless coders are used in concatenation with lossy coders to exploit the correlation in the quantized symbol output of the latter and hence to improve the compression gain. thus the lossless coders like huffman coders and arithmetic coders are still part of many popular video compression schemes.

2.2.2 Distortion

Distortion [76], [104], [73] is defined as the deviation of the reconstructed signal, reconstructed using digitized and encoded data, from the original analog signal. Hence the distortion is essentially a cost function involved in compact representation of the original signal. Rate distortion function relates the number of bits required to represent the data with admissible level of distortion. In other words, the rate distortion function gives the minimum average rate R_D (in bits per sample) required to represent a signal while allowing a fixed distortion D in its reconstructed value.

Let x_g be a Gaussian random variable of variance σ_x^2 and x_r be its reconstructed value. If the distortion is measured as mean square error (MSE) i.e.,

$$D = E[(x_g - x_r)^2] \quad (2.2)$$

then the rate distortion function of x_g is defined as,

$$R_D = \begin{cases} \left(\frac{1}{2}\right) \log_2(\sigma_x^2/D) & D \leq \sigma_x^2 \\ 0 & D > \sigma_x^2 \end{cases} \quad (2.3)$$

where the maximum value of D is σ_x^2 . This relation between R_D and D is depicted in Fig. 2-1. Using this relation it is possible to determine the output bit rate required to achieve a given distortion or alternatively the minimum attainable distortion for a given bit rate.

The distortion measure is an important criterion in evaluating a source coding scheme. The choice of distortion function is affected by many factors like its computational complexity, analytical tractability and its adaptability to human audio-visual system characteristics. The last requirement arise from the fact that the final judge of any coding system is the human system. The available distortion measures in the literature [76] are the Maximum Cross Correlation (MCC), Mean Square Error (MSE), Minimum MSE (MMSE), Normalized MSE (NMSE), Minimum Absolute Distance (MAD), Signal to Noise Ratio (SNR), Segmental SNR and Peak SNR (PSNR). Out of these, SNR has been the traditionally popular distortion measure in signal processing as it is mathematically tractable and serves as a useful measure for most of the applications. The computationally simplest distortion measure available is MAD. The PSNR is widely used in speech, image and video compression. Considering a two dimensional signal x_{ij} of size $N_x \times N_y$ (like image or a video frame) with its reconstructed form as y_{ij} , the definitions for these popular distortion measures are as follows;

$$MAD = \frac{1}{N_x \times N_y} \sum_{i=0}^{(N_x-1)} \sum_{j=0}^{(N_y-1)} |x_{ij} - y_{ij}|$$

$$MSE = \frac{1}{N_x \times N_y} \sum_{i=0}^{(N_x-1)} \sum_{j=0}^{(N_y-1)} (x_{ij} - y_{ij})^2$$

$$NMSE = \frac{\sum_{i=0}^{(N_x-1)} \sum_{j=0}^{(N_y-1)} (x_{ij} - y_{ij})^2}{\sum_{i=0}^{(N_x-1)} \sum_{j=0}^{(N_y-1)} (x_{ij})^2}$$

$$SNR = 10 \log_{10} \left(\frac{1}{NMSE} \right)$$

Figure 2-1: Rate distortion function for a Gaussian source

$$PSNR = 10 \log_{10} \left(\frac{Peak\ Signal\ Value^2}{MSE} \right)$$

It is known that none of the above measures correlate well with the subjective ratings, especially at low bit rates. Hence there are several new distortion measures suggested in the recent literature [92], [77], [126] for better adaptation of HVS properties.

2.3 Image Transforms

As one dimensional signals can be represented by a set of *basis functions*, two dimensional signals like images can be represented by a set of *basis images* [73], [104], [115], [150], [29], [31]. These basis images can be generated using unitary matrices. The term *image transform* refers to a class of such unitary matrices used to represent images. The advantage in using transforms before quantization is that these unitary transforms have a tendency of packing a large portion of the signal into a small set of transform coefficients without altering either the total energy content or the entropy of the signal. Another interesting advantage is that when the input signal is correlated, which is usually the case, the output transform coefficients tends to be uncorrelated. For the example of an image, this results in stronger diagonal coefficients and relatively very weak off-diagonal coefficients. Hence in the transform domain the quantizer can make a choice to encode only those coefficients above a certain threshold and send the rest as a sequence of zeroes, thus leading to compression. The compression gain achieved increases with the threshold but at the cost of increase in information loss. Some of the important transforms in image and video coding are DCT and DWT which are reviewed below.

2.3.1 Discrete Cosine Transform

Two dimensional DCT matrix for size $N_d \times N_d$ is given by,

$$c(n_x, n_y) = \begin{cases} \frac{1}{\sqrt{N_d}} & n_x = 0, 0 \leq n_y \leq N_d - 1 \\ \sqrt{\frac{2}{N_d}} \cos\left(\frac{\pi(2n_y + 1)n_x}{2N_d}\right) & 0 \leq n_x \leq N_d - 1, 0 \leq n_y \leq N_d - 1 \end{cases} \quad (2.4)$$

The DCT has some interesting properties as applied to picture compression. They are; first, it is real and orthogonal. This means that the inverse transform matrix is the transpose of the forward transform matrix. Second, it is a fast transform which needs $O(N \log_2 N)$ computations. Faster algorithms which are optimized specifically for a set of commonly used transform parameters are also available in the literature [104]. Third, it has excellent energy compaction for highly correlated signal and it performs very close to Karhunen-Loeve (KL) transform, which provides the best compaction, under certain circumstances. Unfortunately use of DCT comes with a price. At high compression gain, use of DCT results in blocky noise and it fails to provide a acceptable picture quality. This has set the picture coding community searching for new transforms and new ways of coding.

2.3.2 Wavelet Transform

Wavelets [150], [29], [35], [122], [99] are functions generated by dilations and translations of one single function ψ , called as *mother wavelet*. i.e.,

$$\psi^{a,b}(t) = |a|^{-1/2} \psi\left(\frac{t-b}{a}\right) \quad (2.5)$$

The mother wavelet has to satisfy two important conditions. They are; providing some oscillations and decaying in both directions of time to zero. This means that the mother wavelet has to have *zero average value* and *finite energy*. Using wavelet transform it is possible to represent any arbitrary function f as a superposition of scaled wavelet functions. Thus wavelets are used as basis functions to represent any function f .

The discrete wavelet functions are defined as,

$$\psi_{m,n}(t) = \psi^{a_0^m, nb_0 a_0^m}(t) = a_0^{-m/2} \psi(a_0^{-m/2} t - nb_0) \quad (2.6)$$

where m, n are non-negative integers and a_0 and b_0 are constants with $a_0 > 1$ and $b_0 > 0$. $\psi_{m,n}$ constitute the orthonormal basis so that the wavelet coefficients corresponding to a function f are given by,

$$c_{m,n}(f) = \langle \psi_{m,n}, f \rangle = \int \psi_{m,n}(x) f(x) dx \quad (2.7)$$

The realization of discrete wavelet transform is shown in Fig. 2-2, where $g(k)$ and $h(k)$ are

decomposition filters defined by discrete wavelet functions and $g^{-1}(k)$ and $h^{-1}(k)$ are the reconstruction filters. In practice these filters are realized as Finite Impulse Response (FIR) filters. The multiresolution wavelet transform will involve several stages of those shown in Fig. 2-2 connected in a cascade fashion. The subband split multiresolution pyramid, called as subband structure, is shown in Fig. 2-3. The output at the last stage of the cascade will be the residual or 'dc' portion of the signal which is basically a blurred (or coarse) version of the original input.

The spatial transform is formed using two dimensional wavelet filtering. Studies of 2D subband analysis/synthesis using both separable and non-separable wavelet filter banks are reported in the literature [35], [100], [131], [1], [7], [39], [45], [132]. The computationally efficient approach to splitting and merging subband images results from using separable filter banks. With separable wavelet filters 2D filtering can be implemented using a set of 1D filters [99] as illustrated in Fig. 2-4. Let the input picture frame, taken as LL_0 , is of size $N_x \times N_y$. First, all the rows are low pass filtered and decimated by two to form an intermediate frame IL . Additionally, the same input rows are high pass filtered and decimated by two to form another intermediate frame IH . Both IL and IH are of size $(N_x/2) \times N_y$. Next, both IL and IH are separately filtered column-wise and decimated to form LL_1 , LH_1 , HL_1 and HH_1 subbands of size $(N_x/2) \times (N_y/2)$. The LL band is called as dc subband, LH band is called as horizontal subband, HL band is called as vertical subband and HH band is called as diagonal subband. The three subbands LH , HL and HH are together called as *directional subbands*. It can be noted that the resolution of LL_1 and other directional subbands at this stage are half that of LL_0 . This process of band splitting is continued till the required lowest resolution is achieved. The reconstruction procedure will be in the reverse order with decimations replaced by interpolations and each of the filters replaced by an appropriate inverse filter. A three level decomposition of picture frame, also called as hierarchical subband structure, is shown in Fig. 2-5.

For three dimensional wavelet transform [105], [51], [89] it is required to perform an additional one dimensional wavelet filtering among a group of frames which have separately undergone spatial 2D wavelet transform.

Higher number of levels of band splitting allows more error resilience as the errors could

$g(k)$ and $h(k)$ are low-pass and high-pass filters.

$g^{-1}(k)$ and $h^{-1}(k)$ are respective inverse filters.

$\downarrow 2$ and $\uparrow 2$ are decimation and interpolation, respectively, by 2.

y_i 's are the wavelet coefficients and x_0 is the original input signal.

Figure 2-2: Illustration of one stage of one dimensional DWT

y_i are wavelet coefficients and x_4 is the residual signal at the output of LPF4 and decimator pair.

Figure 2-3: A four stage one dimensional wavelet decomposition

a. Decomposition of i th resolution LL band into four $(i+1)$ th resolution bands b. Reconstruction of i th resolution LL band from four $(i+1)$ th resolution bands

$g(k)$ and $h(k)$ are 1D lowpass and high pass filters respectively and

$g^{-1}(k)$ and $h^{-1}(k)$ are their respective inverse filters.

$\downarrow$ and $\uparrow$ are decimation and interpolation by 2 respectively and

Col and Row correspond to column and row of pixels in the frame

IL and IH are intermediate frames.

Figure 2-4: 2D wavelet filtering using 1D filters

Figure 2-5: A three level wavelet decomposition of picture frame

be more evenly spread. The number of levels of band splitting is limited only by the original picture frame size and the filter length used for LP and HP filters. An interesting point to note is that, as each of the subbands, starting from highest resolution level, is eliminated (assumed to be zeros) from reconstruction procedure, the quality of the reconstructed picture degrades only gradually [99], [131], [1], [7], [12], [16], [52], [134]. The same is true with the precision with which the coefficients are represented. Thus by not encoding certain high frequency subbands or by encoding coefficients with lower precision, one can trade reconstructed picture quality to the output bit rate. Another important feature is that the directional subbands at each level of resolution are highly correlated as they represent the same information at different resolution and direction. In other words they are different frequency bands of an highly spatially correlated picture frame. Using proper quantization schemes this can be exploited to improve compression gain. Finally, these coders are free of any annoying blocky noise at the receiver which is a common problem in DCT based video coders. These make the wavelet transform based video coders superior to those based on DCT. Many design techniques for wavelet filters are reported in [35], [100], [131], [1], [7], [39], [132] and fast algorithms are discussed in [123].

2.4 Motion Estimation

2.4.1 Block Based and Pixel Based Motion Estimation

The motion estimation techniques can be broadly divided into two groups; block based and pixel based techniques. A block based scheme [151], [86], [64], [80] divides the present picture frame into a number of small blocks and, for each of these blocks, a search is conducted for a closest matching block within a *search window* in the previous frame. The relative coordinates of the matching block, called as motion vector, is sent to the receiver. Then the difference block is computed as the difference between the block in the present frame and the matching block in the previous frame. all the difference blocks, comprising of difference frames, are quantized and sent to the receiver. At the receiver the matching block is located using the coordinates received and it is simply added to the difference block received to reconstruct the picture block. One important assumption made in any block based motion estimation technique is that all pixels within a block

belong to the same object and thus the nature of their motion activity is same. This need not always be a valid assumption. However, since the block based techniques provide a simple and computation efficient structure for implementation they are adopted in various image and video compression schemes. In contrast to the block based schemes, the pixel based schemes estimate the motion at pixel level. These schemes generally perform better than the block based schemes at the cost of increased complexity and output bit rate. Each of these techniques can be employed to work either in picture domain or in parameter domain, like in wavelet domain or in DCT domain.

In any motion estimation scheme [115], [104], [151], the matching criterion and the search window size needs to be selected. A good matching criterion is a must in finding a proper match to a block/pixel in the present frame from the previous frame. The frequently used matching criteria are the maximum cross correlation (MCC), the minimum mean square error (MSE) and the minimum absolute distance (MAD). Out of these, MAD is the least computational expensive and hence is the most widely used criteria. The search window size decides amount of the motion activity that can detected and the number of bits needed to encode a motion vector. The objects moving at higher speed are tracked better with bigger search area and is essential in a fast motion movie or television applications. For applications such as video-phones and video-teleconferencing, a small search area will be sufficient as the motion is slow and easily predictable. However, if the search window size is too small it adversely affects the performance of the motion estimation scheme and on the other hand if the search area is large the overhead in terms of search time and number of bits per motion vector increases.

The performance of a motion estimation scheme can be improved by using finer search grid than the actual pixel grid in the picture frame [48], [151]. This needs the search frame to be interpolated by an integer value. Then the accuracy of the search algorithm is said to be the reciprocal of the interpolation factor. This is illustrated in Fig. 2-6. However the cost of search increases with the interpolation factor.

Fast implementations of motion estimation involve hierarchical methods [98], [63] where motion estimation is carried for a lower resolution of the image and are refined for the actual resolution.

Figure 2-6: Motion search illustration for 1/4th pixel accuracy

2.4.2 Multi-Resolution Motion Estimation

The motion activity at different resolutions of wavelet frame is correlated as the information at those resolutions is highly correlated. This is exploited in Multi-Resolution Motion Estimation (MRME) schemes [151]. In this scheme, the motion vectors for the motion blocks in the lowest resolution subband are estimated first and then those for higher resolution subbands are obtained by refining the scaled versions of these vectors. This process is diagrammatically illustrated in Fig. 2-7.

There are four algorithms of MRME tested in [151]. They are,

MRME1: Motion vectors are estimated for LL subband of lowest resolution. Appropriately scaled versions of these are used as motion vectors at all other subbands.

MRME2: Motion vectors for all the four bands in lowest resolution are estimated. For the rest of the resolutions, the scaled version of the motion vectors from same directional band in the lowest resolution are used as motion vectors without any further refinement.

MRME3: Motion vectors are estimated for LL subband of lowest resolution. For all the directional bands in all resolutions, the scaled versions of these are used as bias motion vectors and refining process is carried out by searching for matching block in a smaller search window around the pixel pointed to by the obtained bias.

MRME4: Motion vectors are estimated individually for all the four subbands in lowest resolution. For the rest of the resolutions, the scaled version of the motion vectors from same directional subband in the lowest resolution are used as bias for refining.

The scaling factor at i th resolution will be 2^{L-i} , where L is the total number of resolutions with zeroth being the highest resolution. The search burden and the bit budget requirement in MRME1 is proportional to the number of motion vectors in LL band. The same is four times higher for MRME2 and $(3L+1)$ times higher for MRME3 and MRME4. The search burden and the bit rate requirements in MRME3 and MRME4 can be reduced by choosing smaller search windows at higher resolutions at the cost of reduced performance.

- Position of original block $V(x, y)$ is the bias motion vector
■ Position of matching block $\Delta(x, y)$ is the correction factor

Figure 2-7: Illustration of multiresolution motion estimation

2.5 Quantization

2.5.1 Background

Quantization [76] is the most important part of any lossy compression algorithm, as this is the stage where the actual data compression happens. If the quantization is carried out in a coarse manner, a high compression gain can be achieved, but with poor reconstruction quality. On the other hand, a fine quantization can provide good reconstruction quality at the cost of reduced compression gain. The quantizers are classified into *scalar quantizers* and *Vector Quantizers* (VQ). The scalar quantizer [76] quantizes one pixel or sample at a time with or without any memory of the past samples and hence the optimization criteria is sample based. The vector quantizer [43] encodes a group of signal samples, termed as a vector, together with a vector based optimization criteria. In many applications quantizer of variable bit rate output is preferred. In the case of VQ, variable bit rate needs the storage of a set of codebooks, each corresponding to a particular bit rate or compression gain. This limits the application to only a fixed set of bit rates. Additionally, the computational requirement of VQ is very high because of the codebook search procedure. On the other hand, the scalar quantizer can provide fine variations in bit rate with low complexity. The following section reviews zero tree quantization which is a scalar quantization scheme built to exploit the correlation in multiresolution structure of the wavelet pyramid.

2.5.2 Zero Tree Quantization

The Zero Tree Quantizer (ZTQ) [129] is a successive approximation type embedded quantizer which compactly quantizes the multiresolution wavelet pyramid. In this scheme the positions of significant coefficients, collectively known as *significance map*, are efficiently represented through an exponentially growing trees. This allows the prediction of non-significant coefficients without much overhead in terms of bit bandwidth or complexity. Once the position of significant coefficients are known, Successive Approximation Quantization (SAQ) is performed iteratively leading to embedded coding. Very recently improved versions [127] of this coder and a 3D extension [26] are also reported.

2.5.2.1 Significance Map Encoding

Past studies [115], [104], [76] have shown that the costs involved in coding the significance map is considerably higher than the cost involved in coding the magnitudes of coefficients, especially at low bit rates. Hence it follows that a good improvement in encoding the significance map translates into a corresponding gain in compression. To improve the compression gain of significance maps of wavelet coefficients, a data structure called as *zero tree* is defined.

A coefficient w is said to be insignificant (or zero) with respect to a given threshold, T , if magnitude of w is smaller than T . The zero tree concept is based on the hypothesis [150], [29], [89], [129] that, if a wavelet coefficient at a lower resolution band is insignificant with respect to a threshold T , then all the wavelet coefficients at the same orientation and spatial location at higher resolutions are likely to be insignificant with respect to T . More specifically, in a hierarchical subband system, with the exception of the highest frequency subbands (or highest resolution subbands), every coefficient at a given resolution can be related to a set of coefficients at the next higher resolution in the same direction. This coefficient at a lower resolution is called as *parent*, and all the related coefficients at the next higher resolution are called as *children*. For a given parent the set of all the related coefficients at all higher resolutions in the same direction are called as *descendants*. Similarly for a given child, the set of all related coefficients at all lower resolutions in the same direction are called as *ancestors*. These relationships of child, parent, descendants and ancestors are shown in Fig. 2-8. It can be seen that each parent has four children in the next higher resolution, all corresponding to the same spatial location and direction.

The zero tree construction for a given threshold T is shown in Fig. 2-9. A coefficient w is said to be an element of a *zero tree*, for threshold T , if w itself and all its descendants are insignificant with respect to T . On the other hand, if w is insignificant with significant descendants then it is coded as *isolated zero*. The coefficient of origination of zero tree is coded as *zero tree root* and all its descendants are left uncoded (for threshold T only) as they can be easily predicted at the decoder. Hence, any coefficient at a given threshold T is encoded as a zero tree root or an isolated zero or a positive/negative significant coefficient. These four symbols are collectively called as zero tree symbols, each of which needs two

- Wavelet coefficient
- Parent-child dependencies of subbands
- Scanning order of the subband coefficients

Figure 2-8: Parent-child dependency and scanning order of coefficients

Figure 2-9: zero tree construction (dominant pass) for a given threshold T

bits for representation. When encoding the highest resolution coefficients that are not part of an already existing zero tree, the zero tree root symbol need not be used as there is no higher resolution for a zero tree to continue.

The zero tree symbol formation needs the wavelet coefficients to be scanned in an order such as to always scan the parent coefficient before any child coefficient. Additionally, the lower frequency subband at a resolution is scanned before any higher frequency subband at the same resolution so as to give priority to the low frequency part of the picture. This scanning order is illustrated in Fig. 2-8.

2.5.2.2 Successive Approximation Quantization

The zero tree quantizer algorithm allows a successive approximation quantization (SAQ), with a sequence of thresholds $T_0, T_1, \dots, T_{N-1}$ applied in an iterative fashion to determine the significance of the wavelet coefficients. The threshold at any iteration i , T_i , is determined as one half of the previous threshold T_{i-1} . The initial threshold (T_0) is chosen to be greater than one half the magnitude of the largest coefficient. i.e., $|w|_{max} < 2T_0$. The initial threshold T_0 is communicated to the receiver at the beginning of the symbol stream.

Each iteration in zero tree quantization and reconstruction involves two separate passes, a dominant pass and a subordinate pass. Each of these passes use a separate list of wavelet coefficients termed as dominant list and subordinate list respectively. The dominant list contains the coordinates of those coefficients which are yet to be found significant. Hence in the beginning, the dominant list contains the coordinates of all the wavelet coefficients scanned in the order shown in Fig. 2-8. This is such that the subbands are ordered and within each subband, the set of coefficients are ordered sequentially. Thus all the coefficients in a given subband appear on the initial dominant list prior to coefficients in the next subband. The subordinate list contains the magnitudes of those coefficients that are already found significant. Hence at the start of the first iteration the subordinate list will be empty.

During the dominant pass, the coefficients with coordinates in the dominant list, i.e., those that have not been found significant, are compared to the threshold T_i to determine their significance and the zero tree symbols are formed as described in section 2.5.2.1.

Whenever a coefficient is found significant, the dominant list is updated by removing the coordinates of the coefficient from the list and the subordinate list is updated by inserting the magnitude of the coefficient into the list. While updating the dominant list it is important not to disturb the relative order of the rest of the coefficients. In addition, the original input list of the wavelet coefficients is updated by resetting the value of the newly found significant coefficient to zero so as not to prevent the occurrence of a zero tree in the future dominant pass with smaller thresholds.

During the subordinate pass, all the coefficients in the subordinate list are scanned and they are refined to an additional bit of precision. In other words, the effective quantization step, which defines the uncertainty interval for the true magnitude of the coefficient, is halved. For each coefficient in the subordinate list, this refinement is encoded as a binary *ONE* to indicate that the true value falls in the upper half of the old uncertainty interval and a binary *ZERO* to indicate the lower half. Note that prior to this refinement the width of the uncertainty interval for all the coefficients found significant is equal to the current threshold value.

It can be observed that any wavelet coefficient in the original list is found significant only once and gets encoded in a successive approximation fashion, one bit at each of the succeeding iterations. Thus at each iteration a set of zero tree symbols corresponding to the dominant pass and a set of binary symbols corresponding to the subordinate pass are sent to the receiver. The encoder stops whenever the target condition is met, meaning that the coding need not always stop at the end of a dominant or a subordinate pass.

2.5.2.3 Reconstruction

The dominant and subordinate pass reconstructions are performed as and when the symbols are available at the decoder. During the dominant pass, if a symbol corresponding to a dominant list entry is received as significant, its coordinates are removed from the dominant list and its reconstruction value is updated to the subordinate list. The reconstruction value needs to be at the centre of the uncertainty interval to satisfy the minimum mean square error (MMSE) criterion. Hence it is set to one and half times the current threshold. In the subsequent subordinate passes, a binary symbol is received for this entry in the subordinate list and is used to further halve the uncertainty interval of its

reconstructed magnitude. At the end of the iteration the current threshold is halved to form the threshold for the next iteration. Thus if the encoder stops in between an iteration, the decoder also stops at that same point in the same iteration resulting in an effective usage of all symbols sent from the encoder.

2.5.2.4 Symbol Level Encoding

Note that the particular encoder alphabet at any given time contains either 2, 3 or 4 symbols depending on whether the encoding is for a subordinate pass, or a dominant pass with no zero tree root symbol, or a dominant pass with zero tree root symbol. This gives the advantage in adapting a symbol level entropy coder which improves compression. If the entropy coder is adaptive, a short memory is sufficient to quick learning and constant tracking of changing symbol probabilities.

It is advised [129] that, during the subordinate pass the subordinate list be reordered as per the reconstructed magnitudes (which are available at both encoder and decoder) so as to give priority to higher magnitude coefficients. If encoder stops in between a subordinate pass, this reordering slightly improves the performance for a given bit budget.

2.6 Existing Standards

Compatibility among different applications and manufacturers is very desirable and is often essential. This lead to the need for standardization among different available video coding schemes. The standards are also crucial to the growth of visual services in communication and computer systems. Without standards, the communication between terminals and systems become extremely cumbersome and costly affecting the economies of scale in the manufacture of systems, user devices and microelectronics chips. Efforts by International Organization for Standards (ISO) have resulted in video coding standards which are widely known as H.261, H.263 and MPEGs.

2.6.1 H.261 and H.263 Standards

The H.261 standard [102], [69], [56] specifies scheme for video compression at bit rates of $px64$ Kbps, where p can range from 1 to 30, meant for low bit rate channels like PSTN. H.263 [70] is developed in close relation to H.261 for bit rates less than 64 Kbps. Both

H.261 and H.263 share the same basic codec structure based on DCT and motion estimation. Improvements in H.263 design enables the coder to provide significantly better subjective quality with less than half the bit rate compared to H.261.

The encoder and decoder structures for both standards are shown in Fig. 2-10. In both the schemes only the first picture frame is completely intra-frame coded. A block based motion compensation is performed on the subsequent frames. The input picture frame is divided into blocks of size 16x16 called as *Macroblocks* for motion estimation. While H.261 allows only integer pel accuracy for motion estimation/compensation, H.263 allows half pel accuracy too. The DCT is employed to code the difference frame in both the standards. The DCT compressed stream is encoded using variable length Huffman code to further improve the compression gain. As these coders are designed for real time communications, they use the closest previous frame for prediction and offer minimum encoding delay. The complexity of the encoder and the decoder are balanced as it is necessary in real time applications.

The H.263 provides four extra options of unrestricted motion vector mode (UMV mode), advanced prediction mode (AP mode), PB-frames mode (PB mode) and syntax based arithmetic coding mode (SAC mode) to improve the performance of the coder. In the UMV mode, the motion vectors can reference pixels outside the coded picture. The pixels outside the picture are approximated by the closest edge pixel. This avoids the restriction on the motion search window to be completely within the picture borders and provides accurate motion estimation at the borders. The AP mode operation allows overlapped motion compensation and the possibility of four motion vectors per macroblock to improve the performance of the coder. The PB-frames mode provides combined coding of a forward predicted (P) frame followed by a bidirectionally predicted (B) frame to achieve higher compression. Finally, SAC mode replaces Huffman coding with arithmetic entropy coding to improve the compression gain.

2.6.2 MPEG Standard

The MPEG standard [41], [67], [68] is designed to address the needs of video compression at about 1.0Mbps with good video quality and meant for applications involving the digital storage of compressed video. Two conflicting requirements in the design of MPEG

Figure 2-10: Encoder - decoder structure for H.261/H.263 standard

standard are the video quality and the random access feature. For a given bit rate, good quality needs high compression ratio not achievable with intraframe coding alone while the random access requirement needs pure intraframe coding. Thus the compression algorithm was chosen to be a combination of intraframe and interframe modes with interframe mode being coded with both predictive and interpolative techniques.

Similar to H.261/H.263, the MPEG coder constitutes of block based motion estimation and DCT as shown in Fig. 2-11. The motion estimation, with the macroblock size of

Figure 2-11: Simplified block diagram of MPEG encoder

16x16, is used with both predictive and interpolative coding. The motion estimation can be based on forward prediction, backward prediction or bidirectional prediction. The bidirectional prediction uses macroblocks from both past and future frames. Hence the motion information consists of one vector for forward or backward predicted macroblocks and two vectors for bidirectionally predicted macroblocks.

The compressed frames are classified as intraframes (*I*), predicted frames (*P*) and

Figure 2-12: Relation between different frame types of MPEG standard

interpolated or bidirectionally predicted frames (*B*) as shown in Fig. 2-12. The intraframes provide access points to random access as they do not depend on any past or future frames. However only a moderate compression gain is achieved for intraframes. The predicted frames are the forward predicted frames predicted using the past intraframes. The interpolated frames will have macroblocks which can be of type intra, forward-predicted, backward-predicted or average of past and future frames. Hence these frames provide the highest compression gain, but require both a past and a future reference frame for prediction. Both intraframe and predicted frame are used as reference frames for the interpolated frames. The interpolated frames are never used as reference frames. The organization of these frames is flexible in MPEG and depends on the specifications of a particular application.

The quantization of the motion vectors and the DCT coefficients are performed using a combination of fixed length and variable length codes to achieve maximum efficiency and good visual quality.

2.6.3 Comments on Video Coding Standards

The performance comparisons for H.261 and H.263 are detailed in [49]. It is shown that the H.263 coder, with all the four options, provide better picture quality, about 3 dB better,

than H.261 coder. Main drawback of these coders is the unacceptable blocky noise at low bit rates. This results because of the use of block based DCT for spatial redundancy reduction. The performance worsens further under noisy channel conditions like in mobile environment [145]. Recently an extension of H.263 to mobile environment, called as H.26P/M, is evaluated in [66] for DECT channel. In this coder, unequal error protection (UEP) is found to perform better than equal error protection (EEP) under noisy channel conditions. However it is observed that the PSNR degrades continuously as the errors propagate between subsequent frames. In [112], [113] a derivative of H.261 coder is tested on DECT channel as part of European project RACE termed as the Mobile Audio-Visual Terminal (MAVT). Again it is concluded that the use of UEP improves the performance. In [83] and [84] a two rate source coder derived from H.261 is tested for mobile channel. It is shown that changing the source rate, and channel rate accordingly, depending on the channel conditions result in improved performance.

Two major drawbacks of H.261, H.263 and MPEG compression schemes are: first, at low bit rates the subjective quality becomes poor as DCT based algorithms suffer from blocking effects and second, the quality varies drastically as the bit rate is varied.

2.7 Future Trends

It is an accepted fact that the present day video coding standards are unable to cope with the growing needs of better quality at low bit rates and demand for robust communication of video signal. Hence the researchers are looking at video coders using wavelet transform, fractal coding and model based coding [73], [57], [72], [91]. The research in adapting coders to human visual system (HVS) [28] is also underway.

The fractal coding exploits the self similarity in the picture while the model based coding attempts to model different connected elements in the picture, like human body, tree etc. These schemes claim high compression gain with good picture quality compared to other schemes. However, presently these schemes are highly computational intensive and hence are not affordable. The wavelet transform offers good picture quality and at the same time is realizable using the present day microelectronics technology. The wavelet transform combined with HVS adaptable quantization scheme is thought to become a third generation video coder [57], [91], [28]. A comparison of three generations of coders is

given in Fig. 2-13.

To cope with the birth of new applications at low bit rates, ISO is planning MPEG-4 standard [139], [124], [139] which is to be released in 1998. This standard promises to provide new ways for compression, communication, access and manipulation of audio-visual information at a wide range of bit rates and hence catering to a wide variety of applications. Currently details of services and features are being worked out.

2.8 Summary

In this chapter the important aspects of video compression and the existing schemes were reviewed. In the first section the fundamental concepts of source coding, entropy and distortion as applied to image/video compression were reviewed. This was followed by a review of existing transforms, motion estimation schemes and quantization. In these sections we emphasized DCT, DWT, block based motion estimation, MRME scheme and the zero tree quantization. It was observed that the DCT based schemes suffer from blocky noise at low bit rates. The DWT based schemes are becoming popular for their good subjective quality performance. The MRME scheme and the zero tree quantizer make use of the multiresolution structure of DWT and efficiently exploit the correlation among the subbands. Later, existing standards for video compression were reviewed. All these standards are DCT based and hence provide unacceptable subjective performance at low bit rates. Finally, a glimpse of future trends was provided in which it was observed that the DWT combined with an HVS based quantization can provide good performance.

Keeping these observations in mind, we decided to investigate the use of wavelet transform as the spatial transform in this work. In order to exploit the subband structure and inter-subband correlation, we chose the multi-resolution motion estimation. It is possible to use different MRME schemes as per the available bit budget and assist in achieving the variable compression gain. For the same reasons we have chosen to examine the zero tree coding as the quantization scheme. As the zero tree coding is capable of providing smooth bit rate variations, any bit rate smoothing filters at the multiplexer can be avoided resulting in reduced cost at the network level.

Figure 2-13: Comparison of image compression techniques

Chapter 3

Review of Mobile Communications

3.1 Introduction

The ultimate objective of communications is to enable anyone to communicate with anyone else from anywhere. Mobile radio communication is essential to achieve this goal. Early mobile communication systems used Morse-coded on-off keying scheme to communicate messages among ships and from ships to shore. Developments from there on include mobile telephone, paging, navigation control, status reporting and so on. the demand for mobile radio communication services are increasing everyday. As the available frequency spectrum is limited, efficient spectrum utilization is a key factor in coping with such increasing demand. Other important factors to be kept in mind while providing a service are security, reliability, delay involved and quality of service.

This chapter gives a brief review of some of the important topics in mobile communication. The mobile channel behavior and it's characterization is reviewed in section 3.2. Techniques like diversity and channel coding which are essential to improve the signal quality are reviewed in section 3.3. It is important to model the channel for system design and analysis. This is discussed in section 3.4. A brief review on the available standards and systems for mobile communication is provided in section 3.5. Finally, a summary of the discussion is given in section 3.6.

3.2 Channel Characterization

There are three important characteristics of mobile communication channel. They are *multipath fading*, *shadowing* and *path loss*. The multipath fading is further divided into

envelope fading (non selective), *Doppler spread* (time selective) and *time-delay spread* (frequency selective). The multipath fading is also termed as *small scale variations* as it results in fast fluctuations in the envelope of the received signal. Similarly shadowing is called *large scale variations* because of its effects on fluctuations in the envelope of the received signal.

When a pulse is transmitted from the transmitter, it is received as a train of overlapped pulses at the receiver due to multipath propagation. These paths are random in direction and are accompanied by randomly varying channel conditions. The range of temporal width over which these multipath signals are spread is termed as time-delay spread (T_m). At the receiver, multipath propagation results in fluctuations in the envelope of the overall received signal when measured over a distance of few hundreds of wavelengths [74], [111], [40], [32], [138]. This is termed as envelope fading which is shown to be *Raleigh* distributed when the number of multipath signals involved are sufficiently large and random. If there are fixed scatterers or signal reflectors in the medium they may cause some of the received paths to have fixed and low attenuation. In this case the envelope of the received signal will have significant non-zero mean and hence can not be modeled as having zero mean. Hence the envelope is better modeled using *Ricean* distribution instead of *Raleigh* distribution. In both the cases the rate at which the envelope fluctuates is proportional to the vehicle speed and the carrier frequency.

The multipath propagation coupled with motion causes fast phase changes in each of the received paths resulting in random FM noise on the received signal. This effect of multipath propagation is termed as *Doppler spread* and is found to impose a limit on the performance of digital phase modulated transmission systems. Hence the Doppler spread (B_d) is a measure of the spectral width of a received carrier when a single sinusoidal carrier is transmitted through the multipath fading channel. The *Coherence time* (Δt_c) is defined as the time interval over which the channel exhibits correlation at the received signal envelope. The coherence time is approximately equal to the reciprocal of the Doppler spread of the channel.

The *Coherence bandwidth* (Δf_c) is the required frequency spacing between two carriers and is approximately equal to the reciprocal of the time-delay spread. When the symbol

duration (T_s) is small compared to T_m , the multipath components with larger delay will spill over to the next symbol causing waveform distortions in the received signal. This effect of time-delay spread is regarded as *frequency selective fading* or *intersymbol interference* and imposes an upper limit on the performance of high-speed digital transmission systems. The channel under this condition is considered as *narrowband channel*. On the other hand if T_s is large compared to T_m , then it is considered as *broadband channel*. Under broadband conditions, the multipath components can be used as a kind of time diversity which is discussed in the next section.

Fast multipath fading can be removed by averaging over distances of few tens of wavelengths. However, this can not alleviate the *nonselective shadowing* effects caused by the terrain features of the land mobile radio propagation. The shadowing imposes a slowly changing average upon the fading statistics. Usually log-normal distribution is used to model the shadowing effect. Average value of this log-normal shadowing is determined by the path loss which varies with the distance between the base station and the mobile.

Taking all the above mentioned effects into account, the complex radio channel impulse response can be modeled for each point in three dimensional space as,

$$h(t, \tau) = \sum_{k=0}^{N(\tau)-1} a_k(\tau) \delta(\tau - \tau_k(t)) e^{j\theta_k(t)} \quad (3.1)$$

where t and τ are the time instant and delay associated. $N(\tau)$ are the number of multipath components. $\{a_k(t)\}$, $\{\tau_k(t)\}$, $\{\theta_k(t)\}$ are the random time-varying amplitude, arrival time and phase sequences, respectively. δ is the delta function.

When the fading is assumed to be slow relative to the symbol period, the time variance of the impulse response can be ignored for the symbol duration and hence the new approximated channel impulse response is,

$$h(\tau) = \sum_{k=0}^{N(\tau)-1} a_k \delta(\tau - \tau_k) e^{j\theta_k} \quad (2.2)$$

Under this assumption, the channel varies between the limits of narrowband and broadband conditions.

The indoor radio channel is non-stationary in time and frequency domain [42] because of movements of people and objects. The Doppler shift in the received signal is negligible as only the low speed motion is experienced. Path loss in indoor environment is severe within a confined room and are between 65 to 85 dB [133]. The impulse response profiles are correlated for close points in space but can be assumed to experience independent channel parameters [60]. Another observation [140] for indoor environment is that the signal coverage can be substantially increased by locating the transmitter in a large room in the middle of the building. This is desirable for one-cell-per-building signal coverage. In the case of multi-cell configuration other factors like interference should also be considered.

The outdoor radio channel is usually modeled using Ricean distribution because of presence of line-of-sight path. However, in places like side streets in urban areas this direct path will be absent and hence the signal strength varies wildly as the receiver is moved, even by a small distance, away from or towards the direct path [42], [61]. In rural areas the signal tends to be more stationary with shadowing being the major problem.

3.3 Techniques to Improve Performance at the Receiver

The popularly used techniques for combating the severe fading in mobile channel are classified into *equalization*, *diversity* and *channel coding*. The equalization is aimed at fighting the inter-symbol interference effect of multipath propagation when the available signal bandwidth is comparable to the coherence bandwidth of the channel. The diversity and channel coding techniques are discussed in the following sections.

3.3.1 Diversity

The basic idea behind diversity techniques [61] is that; if a message is received over a number of paths or branches which are subjected to different fading statistics, then it is possible to choose various branches and combine them properly in order to get better performance than just using any single branch. The diversity used in any system can be *space diversity*, *polarization diversity*, *angle diversity*, *frequency diversity*, *time diversity* and *code diversity*. There are many studies reported in the literature using diversity to combat fading channel effects [38], [50].

The space diversity employs multiple antennas in the space positioned with the

separations on the order of tens of wavelengths. This distance separation between the antennas is required in order to capture paths with uncorrelated fading. However presence of local scatterers in mobile environment reduces this distance to the order of a quarter wavelength.

The polarization diversity makes use of the horizontal and vertical polarizations of the transmitted wave to improve performance. The angle diversity needs multiple antenna beams that point to slightly different directions. This involves the use of multiple antenna feeds along with a single reflector-type antenna. The frequency diversity is realized by transmitting the same message over multiple channels which are sufficiently separated in frequency to obtain uncorrelated components at the receiver. This technique needs only a single antenna with an high price in spectrum efficiency. An elegant way of frequency diversity is realized in the *spread spectrum* systems, using either *direct-sequence* or *frequency-hopping*, where the information is spread over a wide spectrum to combat frequency selective fading.

The time diversity means to repeat the message several times with each repetitions made with sufficient time gap so as to obtain uncorrelated fading conditions. The repetitions could be either intentional or unintentional. The scatterers in the media cause multipath wave propagation resulting in unintentional signal repetition. This can be used for time diversity if the available signal bandwidth is very large compared to the channel coherence bandwidth, as in the case of spread spectrum systems. The intentional message repetition is best realized using *Automatic Repeat reQuest* (ARQ) [94]. In ARQ, the receiver sends out a request for re-transmission when the received data is found to be corrupt which will be obliged by the transmitter within certain time. There are many schemes using ARQ for mobile communication that are reported in the literature [9], [71], [83].

Another way of diversity is code diversity which can be used in spread spectrum systems. In this diversity scheme, message is repeated over multiple orthogonal codes which are basically Pseudo-random Noise (PN) sequences. The receiver uses as many correlators as the number of codes and then combines the obtained information in order to improve the performance. The price is in terms of reduced number of code-sets and hence reduced number of users supported by the system.

3.3.2 Channel Coding

The channel coding [18], [93], [116] is one of the effective ways of combating the mobile channel losses [135], [154], [25], [95], especially when combined with interleaving. The interleaving is used to randomize the burst errors so that the decoder can correct them with ease. The channel coders are classified as *block codes* and *convolutional codes* with the decoding schemes being either *soft-decision* or *hard-decision* based.

The block codes involve encoding each block of k_i information bits into corresponding block of n_c bits, where $n_c > k_i$. Hence the code rate R_c , defined as (k_i / n_c) , will be the measure of redundancy introduced by the encoder. The coders like Hamming codes, Golay codes, Bose-Chaudhary-Hacquenghem (BCH) codes are popular in this group. These coders have certain nice mathematical structures which ease their performance analysis. Unfortunately these coders can not support variable length channel coding which is essential in most of the emerging applications like VBR video, VBR speech, etc.

The second type of channel coders are called as convolutional codes which may be viewed as a linear finite state shift register with an output sequence consisting of a selected set of linear combinations of the input sequence. The redundancy rate is the number of output bits from the shift register for each of the input bit. The reciprocal of the redundancy rate is the code rate R_c . The performance analysis of these coders is difficult and hence usually their upper bounds are computed. Another disadvantage of these coders is that the decoding process involving Viterbi algorithm for optimum decoding is computational intensive. The main advantage of convolution coders is that they support variable length coding. With the invent of punctured convolutional codes [23], [58], [149] (see Appendix-B) these codes have become even more useful in providing unequal error protection with just one 1/2 rate or 1/3 rate basic encoder. Further, the punctured convolutional codes can be used to dynamically change the coding rate of any particular information sequence. This makes the convolutional coders very important in the channel coder selection.

The block/convolution coded bits are modulated into an elementary signal waveform before transmission. At the receiver, the demodulated sequence may or may not be quantized before decoding. If the demodulator is allowed to make firm decisions on whether each received bit is a 0 or 1, then the decoding technique is hard-decision based.

On the other hand, if the demodulator lets the unquantized (analog) or finely quantized (discrete) samples to be accessed by the decoder, then the decoder can make use of the additional information contained in these samples to better recover the information sequence. This is termed as the soft decision decoding. It is observed in [116] (section 5.2.6 and section 5.3.8) that the soft decision decoding performs roughly about 2 dB (of channel signal to noise ratio) better than the hard decision decoding.

3.3.3 Comments

All the techniques discussed in this section can be used in both the narrowband and broadband conditions with the following exceptions. The equalization is a necessity for the narrowband channel conditions, which is dogged by the intersymbol interference. Hence the diversity due to multipath is not available for the narrowband systems. On the other hand, in broadband channels the equalization becomes unnecessary while the multipath provides diversity. The other techniques like channel coding, antenna diversity, ARQ are widely used in both the channel conditions to improve the system performance.

3.4 Channel Modeling

The channel can be modeled at many levels in any system under consideration [136], [85]. Some of these are; the propagation channel, the radio channel, the modulation channel and the digital channel. The overall picture describing the relation between these different channels is depicted in Fig. 3-1.

The propagation channel is the physical medium and hence it consists of everything that influences propagation between two antennas. This channel is assumed to be linear, reciprocal and time variant, due to the motion of the mobile. The radio channel constitutes of transmitter and receiver antenna and the propagation channel. Hence the reciprocity of the radio channel depends on the characteristics of the antennas used. The modulation channel extends from output of transmitter modem to the input of receiver modem and hence is of particular interest to the designers of modulation schemes and trellis coding schemes. Assuming linear radio channel the linearity of the modulation channel is determined by the characteristics of the front ends of the transmitter and receiver. The issues involved in achieving linearity of the modulation channel are cost and the power efficiency. Further, the modulation channel is non-reciprocal as the amplifiers and other

Figure 3-1: Different types of channels encountered in radio communication

front-end components are non-reciprocal. This brings in the need for having separate front-ends for transmission and reception operations at the same station. Hence in a cellular radio environment the modulation channel from base to mobile could be different from that in the reverse direction. The digital channel which includes all the system components from transmitter modem input to the receiver modem output is of great interest to the source code and channel code designers. This channel is nonlinear as the output can take only binary values.

One could define another type of channel which has become the most commonly experienced channel in the present day applications. It is called as the network channel as experienced at the application level of any network. Any network involves certain protocol and transfer of data over less erroneous, but load sensitive wired medium. Hence the network channel can have more data corruptions than the digital channel and data erasures because of corruptions at protocol header. The popular protocols used are X.25, Frame relay and ATM (see Appendix-C).

Out of the above described channel types, the network channel is the one that is used in this thesis. Hence we further review this channel type and its modeling in the next section.

3.4.1 Network Channel

The channel experienced due to indoor and outdoor propagation are very different because of totally different multipath and shadowing effects and path losses. In this section we review two separate channel models that are typical for indoor and outdoor scenario. The examples of indoor and outdoor environments considered are DECT and CODIT respectively.

3.4.1.1 Narrowband Wireless Model

In indoor scenario, the received signal envelope experiences very deep fades only occasionally in which the outage rate rises up very rapidly [74], [22], [38], [59], [60], [119]. In addition, for a given average SNR a small increase in outage probability result in a jump in bit error rate (BER) [128]. Hence it is reasonable to model the channel by two states: fade and non-fade (or OFF and ON) states with respect to the received SNR going below certain threshold value [128]. The distribution of ON and OFF states are affected by

the techniques used, like diversity and channel coding, to improve received SNR.

In [96] it is shown that for a general situation where memory-less thresholding is used, the non-fade duration with respect to certain threshold can be modeled as exponentially distributed. Thus the probability density function of the non-fade duration is expressed as,

$$f_{\tau_{NF}}(\tau_{NF}) = \frac{1}{\bar{\tau}_{NF}} e^{-\frac{\tau_{NF}}{\bar{\tau}_{NF}}} \quad (2.3)$$

where $\bar{\tau}_{NF}$ is the mean non-fade duration. Similarly it is shown [101] that the exponential distribution can be used to characterize the fade duration also.

The network with certain protocol support, especially one with packet transmission, the channel can be assumed to be constant over a period of one packet. Thus the ON-OFF periods get quantized in terms of integral number of packets [101].

3.4.1.2 Broadband Wireless Model

The broadband wireless systems mostly use Code division Multiple Access (CDMA) to achieve the required broadband conditions. Here the low bandwidth baseband signal is spread over a large RF bandwidth before transmission. At the receiver, the received signal is de-spread and combined to recover the baseband signal. One of the broadband environments providing medium data rates required by video applications and aiming at VBR services is the COde DIvision Testbed (CODIT).

The CODIT [14], [2], [3], [5], [17], [21], [30] is a proposed standard emerging out of the European RACE project. This environment is based on Direct Sequence (DS-CDMA) CDMA and is classified as a third generation mobile and personal communication system environment [14], [108]. This environment claims some attractions such as potential for high spectrum efficiency, soft capacity, soft handover etc. The key features [14], [137], [4] are;

- Multiple RF channel bandwidths at chip rates of 1M, 5M and 20Mchips/sec (see Fig. 3-2)
- Wide variety of applications supported including VBR speech and VBR video

Figure 3-2: Mapping of user bit rates to chip rates in CODIT environment

Figure 3-3: Generic transmission scheme in CODIT environment.

- Support to mixed-cell architecture
- User bit rates up to 144Kbps
- Separate, reliable control channels for each user data channel (see Fig. 3-3)
- QPSK modulation with coherent detection for down link and offset QPSK with coherent detection for uplink
- Open loop power control for shadowing and path loss, and closed loop power control to track fading at low Doppler frequencies and Doppler cancel/reduction of 244Hz ($f_{D,max}$), corresponding to vehicle speed of 120Kmph, to couple of Hz.
- Four path RAKE receiver with maximum-ratio combining at the mobile
- Channel estimation to assist RAKE demodulation
- Two branch antenna diversity to further improve the performance

The performance of the testbed is evaluated in [14] and it is observed that an SNR of 6.2 dB is required to achieve a bit error rate of 10^{-3} with rate 1/2 convolutional coding and interleaving. It is also shown that this figure compares favorably with the 7 dB SNR that is claimed for the system in [47].

At the digital channel, the CODIT environment with the above mentioned features can be safely assumed to provide uniformly distributed random errors. However at the network channel, the accrued packet loss rate depends on the type of protocol used, the packet header size and the type of error correction used for the packet header.

3.5 Standards and Systems

3.5.1 Present Systems

The present day mobile communication systems [108], [110] can be broadly divided into three groups. They are; cellular systems, cordless telephone systems and wireless data systems. The cellular systems are meant for voice and low rate data within areas covered by the base station. The cordless systems provide a wireless counterpart to the standard telephone with the handset operating within 50 to 100m of the user's base station. However, the development of digital cordless systems has led to higher data rate and more sophisticated applications like wireless private branch exchanges (PBX). The wireless

data systems are developed to support the asynchronous, packet-switched communication. Three main parts of the globe are involved in the development of wireless systems and standards. They are; North America, Europe and Japan. In this section we discuss different systems developed at these parts with emphasis on the digital systems.

3.5.1.1 Cellular Systems

The cellular systems are divided into analog and digital cellular systems. The analog cellular systems are classified as *first generation cellular systems* and used frequency modulation (FM) for speech transmission and frequency shift keying (FSK) for signalling. The system parameters like number of channels, frequency spectrum and FM deviation etc. varied from system to system. Some examples of analog systems are *Advanced Mobile Phone Service* (AMPS) in United States, *Total Access Communication System* (TACS), *Nordic Mobile Telephone* (NMT), *C-450* and *Radiocom 2000* etc. in Europe, and *NTT* and *JTACS/NTACS* in Japan.

The digital cellular systems evolved due to the development of low-rate digital speech coders and advances in integrated circuits technology. These include *Global System for Mobile Communications* (GSM) in Europe, IS-54 and IS-95 in USA and *Personal Digital Cellular* (PDC) in Japan which are classified as *second generation cellular systems*.

The IS-54 is based on *Time Division Multiple Access* (TDMA) and supports three full rate or six half rate user slots on a single frequency channel of 30 KHz bandwidth. This system supports three times the number of users supported by AMPS using the same spectrum. This uses Vector Sum Excited linear Prediction (VSELP) speech coder at a bit rate of 7.95 Kbps (13 Kbps with error protection) at full rate. The IS-54 provides an adaptive equalizer to combat intersymbol interference, but due to the low channel rates supported the equalizer will not be used in most of the situations. This system also provides many facilities like point-to-point short messaging, broadcast messaging, group addressing, private user groups, hierarchical cell structures, and slotted paging channels.

The IS-95 is more recent compared to IS-54 and is based on *Code Division Multiple Access* (CDMA). The basic user bit rate of 9.6 Kbps is spread to a channel chip rate of 1.2288 Mc/sec using various techniques. On the forward link (base to mobile)

coherent detection is employed with the aid of a base-specific pilot code for channel estimation, where as on the reverse link (mobile to base) the base station is forced to use non-coherent detection. A combination of open-loop and closed-loop power control is used for tight control of the mobile's transmit power. The RAKE receivers are used on both links to exploit the diversity effect of multipath propagation. The IS-95 offers a number of advantages including increased capacity, flexibility to accommodate different rates, variable bit rate feature, power control, reduced fade margin and forward error correction (FEC) which contribute in reducing the power requirements at the mobile.

Both IS-54 and IS-95 are built to co-exist with and operate as AMPS systems when required. The IS-54 systems are already deployed and are being used mainly in USA and parts of Canada, while IS-95 systems are still entering the market.

The GSM is based on TDMA with eight user slots per frequency channel and is expected to provide better speech quality, low cost implementation, increased spectral quality, pan-European roaming, and facility for data transmission. This system supports both full-rate (22.8 Kbps per slot, eight time slots) and half-rate (11.4 Kbps per slot, sixteen time slots) with Residual Exited Linear Prediction (RELP) speech coder at 13 Kbps (full rate) and 6.5 Kbps (half rate). The synchronous and asynchronous data services at 9.6 Kbps, 4.8 Kbps and 2.4 Kbps are defined for both full-rate and half-rate operations. The GSM systems started being deployed in 1993 and ever since it's popularity is increasing. There is an estimated two million plus subscribers in a total of 26 European countries and another 26 non-European countries. Recently emerging cellular markets in India and China have also accepted GSM for deployment.

The PDC system is similar to IS-54 with TDMA, three time slots/frequency and VSELP speech coder for full rate operation. The main differences are being 25 KHz channel bandwidth, 6.7 Kbps rate of VSELP (11.2 Kbps with FEC), antenna space diversity and feature of handoff, which facilitates the use of small cells for efficient frequency usage. The half rate speech coder is planned and is based on Code Exited Linear Prediction (CELP) at the bit rate of 5.6 Kbps. This system supports the data transmission at 2.4 Kbps (fax) and 4.8 Kbps (modem) rates. There is an estimated 250,000 subscribers and is gaining popularity because of high quality, high security, a longer handset battery life, etc.

3.5.1.2 Cordless Systems

The cordless systems are meant to be for high quality speech in a quasi-static (very low motion) environment with low complexity, low power equipments. In United States, *Wireless Access Communication systems* (WACS) and *Personal Access Communication Systems* (PACS) were developed for this purpose. The CT2 (CT for *Cordless Telephones*) and *Digital European Cordless Telecommunications* (DECT) were developed in Europe and *Personal Handyphone System* (PHS) was developed in Japan. All these systems used TDMA combined with either *Time Division Duplex* (TDD) or *Frequency Division Duplex* (FDD).

The salient features of CT2 are the use of FDMA with TDD and the use of high quality, 32 Kbps *Adaptive Differential Pulse Code Modulation* (ADPCM) speech coder. The FDMA channels are separated by 100 KHz spacing in the range of 864-868 MHz. The transmission power allowed is 10 mw with two level power control to prevent the desensitization of base station receivers. The CT2 allows data transmission up to 2.4 Kbps through the speech codec and higher rates by accessing the 32 Kbps bearer channel directly. The antenna diversity is allowed both at the base station and the mobile. A Canadian enhancement of CT2, called as CT2+, provides many more mobility management functions like location registration, location updating, paging, etc.

The DECT is much more flexible than CT2 and provides cost effective communication services to high user densities, like in picocells. Functionally the DECT is closer to a cellular system than to a cordless telephone by providing functionalities such as, multiple bearer channels for speech and data, handover, location registration, paging, etc. The DECT uses TDMA-TDD with 12 time slots/carrier in each direction and a high signalling rate at 1152 Kbps. One important feature of DECT is that it can allocate multiple time slots to a single call to enhance the data rate in either directions. This enables the DECT system to cope with the traffic asymmetries arising out of applications such as database access. Due to the high signalling rate, either equalization or antenna diversity is a typical requirement in more dispersive microcells in applications like Telepoint.

The PHS system uses TDMA and TDD, like in DECT, with four bearer channels at a RF bit rate of 384 Kbps. This bit rate is chosen as trade-off between maximizing the

multiplexing number and minimizing the effects of frequency selective fading, particularly in the outdoor environment. A fully dynamic channel assignment is employed by measuring the field strength and selecting a channel on which the field strength is below a prescribed level. Like DECT and CT2, PHS uses 32 Kbps ADPCM speech coder. The error detection is achieved using a cyclic redundancy check (CRC), though there is no error correction. This system can support fax at 4.2 to 7.8 Kbps and full-duplex modem transmission at 2.4 to 9.6 Kbps through the speech codec.

The WACS systems focus on low speed portable applications and small-cell systems. The WACS air interface is similar to DECT with two exceptions. They are; use of FDD instead of TDD and greater optimization of link budget and frequency reuse. This system also uses 32Kbps ADPCM speech coder with CRC for error detection. A two branch polarization diversity is used at both base and mobile with the feedback giving an advantage equivalent to that of a four branch reception diversity. The PACS is designed as a combination of WACS and PHS to create a better industry standard. This is intended as a low-tier air interface for upcoming Personal Communications Services (PCS).

3.5.1.3 Wireless Data Systems

Two important divisions in wireless data systems are *wide area data service* and *Wireless Local Area Networks* (WLAN). The former uses licensed spectrum to provide services to customers while the latter is mostly privately owned and operated to provide a high-rate data communication over a small area, as in the case of in-building communications. Some of the wide area data service providers are Advanced Radio Data Information Systems (ARDIS) and RAM Mobile Data (RMD) using dedicated networks, and Cellular Digital Packet Data (CDPD) using the idle time on existing AMPS channels. These can provide packet data services at various bit rates up to 19.6 Kbps. The examples of WLANs are Freeport, WaveLAN, ALTAIR, HIPERLAN etc. These systems provide services at bit rates higher than one Mbps.

3.5.2 Future Systems

Continuing advances in microelectronics technology and communication techniques have resulted in increasingly sophisticated features and services. Further work in identifying the

services, spectrum allocation and system design are being carried out around the globe. In Europe, systems based on Advanced TDMA (ATDMA) and CODIT (see section 3.4.1.2) are being evolved. These systems are capable of providing data rates above 64 Kbps, which is the basic need of video applications. Parallely efforts to combine the major standards in certain fashion for PCS is also taken up. It has been recognized that PCS standards fall into two categories: *high-tier* supporting macrocells and high-speed mobility; and *low-tier* optimized for low power and low complexity. These two tiers correspond to the existing digital cellular and digital cordless systems respectively. It has been identified that the PACS and DECT form the low-tier standards while the GSM, IS-64 and IS-95 form the high-tier standards. Another two standards based on hybrid TDMA/CDMA approach and Wideband CDMA (W-CDMA) are also being planned for high-tier.

3.6 Summary

In this chapter the different aspects of mobile communication were reviewed. The channel characterization was discussed in section 3.2, where we observed that the signal experiences multi-path fading, shadowing and path loss in the mobile channel. In order to combat these effects, the techniques like diversity and channel coding are available. These techniques were reviewed in section 3.3. It is possible to visualize the concept of channel at many levels in a system. They are; propagation level, radio level, digital level and network level. A brief review of these channel types was presented in section 3.4. In this work our focus is restricted to network level channel models as this is the channel that is experienced at the application layer. Hence we further elaborated on two different channels; the narrowband channel using ON-OFF model and the broadband channel using the CODIT environment. These two channels are chosen in order to make the system evaluation complete in terms of indoor/outdoor and TDMA/CDMA environments. In the section 3.5 we reviewed the presently available standards and noted that, except for the DECT, none of the systems can support bit rates around 64 Kbps which is required for the video applications. The DECT can support these bit rates, but is limited to the indoor environment. On the other hand the CODIT which is being developed can support bit rates more than 64 Kbps and aims at providing VBR video services.

Chapter 4

Problem Definition

4.1 Introduction

This chapter presents the problem investigated in this thesis, the reasons for its selection and the proposed solution. The problem is defined in section 4.2 where it is noted that the robust video communication under mobile channel conditions is still an unsolved problem. The network scenario is chosen to be a combination of mobile and ATM channels, as this channel combination makes the core part of the universal communication channel from the view point of channel impairments. The overview of the proposed system design to achieve robust video communication is presented in section 4.3. The system design emphasizes the low bit rate communication as the mobile channel poses restrictions on the achievable data rates. The system consists of wavelet transform based video coder, stream/data classifier, two-stage channel coder and an adaptive rate controller. Two video coder configurations based on motion estimation and 3D wavelet transform are investigated. The rate controller is designed to play the role of adapting the system to the channel conditions and hence introduce 'robustness' against the channel. This is done through learning the channel variations using feedback from the receiver. Finally, a summary of the discussion is provided in section 4.4.

4.2 Problem Definition

This research is based on the guidelines of MPEG-4 which is being planned to cater for a variety of future audio-video communication applications over a wide range of bit rates and communication conditions. The important considerations in this work that are derived

from the guidelines of MPEG4 are as follows.

- Universal communication channel.
- Robust video communication.
- A minimum of acceptable video quality.
- Variable bit rate (VBR) communication depending on the channel conditions and the quality of service restrictions. This necessitates the system to be capable of learning and adapting to the channel conditions through channel sensing and be able to jointly optimize the source and channel coders.
- Low overhead in source coding and channel adaptation.
- Smooth bit rate variations from frame to frame to avoid the requirement of bit rate smoothing at the network level.
- Mean bit rate, peak bit rate and duration of peak bit rates adaptable to the network connection agreements.

It is envisaged that in the future, it would be possible to communicate from any where to any where else at any time through any form of data using personal communication systems (PCS). Various concepts that are being promoted are; terminal mobility using wireless access, personal mobility based on personal numbers, and service portability through management of user service profiles. The various options available in terms of the location and the device/equipments used are shown in Fig. 4-1. The different types of networks/channels that can get involved are also shown. As it can be seen, this involves communication over mobile networks, satellite networks and fixed networks. The worst of these channels is known to be the mobile channel. It is the worst part of any communication scenario due to the burst errors and large variations in the channel conditions. The use of network protocol can only worsen the situation as few errors in the protocol header can result in the loss of the entire payload accompanying the header. In this work we investigate the mode of video communications over mobile networks with certain protocol restrictions. The standard ATM [20], [11] is chosen, as it is the widely accepted protocol.

Given the above illustrated scenario of PCS and the future systems with mobility to users, the channel experienced by the received signal keeps changing with time and positions of

Figure 4-1: Concept of personal communication

the transmitter and the receiver. For a given transmission bandwidth, if the received signal experience a broadband channel when the user is outdoor, the same can experience a narrowband channel when the user moves into a building due to a large increase in the multipath. For the system to be robust to these changing channel conditions, it is essential to dynamically adapt the system to the channel. This is achieved by jointly optimizing the source and channel coding to the observed channel conditions.

The problem of robust video communication and channel adaptation remains largely unsolved despite some of the lucid efforts reported in [66], [113], [84], [145] and [88] (see section 2.6.3). All these schemes are DCT based and derivative of H.261/H.263. Some image communication schemes using wavelet transform over mobile channels are reported in [75], [97] and [135]. In [28] an attempt is made to communicate 3D-wavelet compressed video over wireless channels. This work suggests error concealment technique at the decoder and does not attempt any kind of channel adaptation. However the reported performance is encouraging, though it comes with the price of large inherent delay involved in the 3D-wavelet schemes. It can also be seen that none of the above reported efforts, except [84], have attempted the combined optimization of video compression and communication which seems to be the solution for the robust communication of video signals. In [84], a two layered coding approach with the local decoder operating with only the essential (first) layer is suggested. This needs the second layer to be strictly non-essential. The increase in number of layers and the increase in information content of the second layer (and subsequent layers) leads to serious performance degradation even under no error condition.

4.3 Overview of the Proposed System

Keeping the above observations in mind, this thesis proposes a robust video communication system comprised of error resilient source coder and a channel adaptable rate controller. The rate controller makes use of the feedback from the receiver regarding the channel status and adapts the system by suitably changing the source rate and the degree of channel protection. Further, the source coder is designed to efficiently slice out any uncorrected errors that can sneak in after the channel decoder.

The block diagram of the proposed system to cater for the above described environment is

shown in Fig. 4-2. It is proposed to use a wavelet based video coder in combination with the multiresolution motion estimation and the zero tree quantization to achieve the required high compression gain.

A two stage channel coder is employed to combat the effects of mobile and ATM channels on the transmitted video signal. The channel coding at the outer stage provides equal error protection (EEP) while that at the inner stage provides unequal error protection (UEP). The stream classification is carried out to mark the importance of each compressed output stream and thus assist in scheduling the UEP.

The variable bit rate required for high performance under network transport conditions is achieved by the use of adaptive rate controller. The rate controller changes the output video bit rate and UEP schedule depending on the quality of service requirement and the channel conditions.

The system design emphasizes the low bit rate communication due to the rigid restrictions posed by the mobile channel on the achievable data rates. However the same design can be employed at higher bit rates by choosing the appropriate values for the system parameters.

4.3.1 Video Coder

The proposed video coder structure makes use of the two dimensional wavelet transform (2D-WT) to exploit the spatial redundancy in the input picture frames. In this work we are investigating both the temporal wavelet transform (3D-WT) and the multiresolution motion estimation (MRME) techniques to exploit the temporal redundancy in the input image sequence. The design details, the criteria considered and the performance evaluation of the schemes are discussed in chapter 5.

4.3.1.1 Robust Quantization

We are investigating the zero tree technique (section 2.5.2) for quantizing the wavelet coefficients. The main disadvantage of this coder is its intolerance to errors. Hence any uncorrected error result in invalid symbol input at the zero tree decoder making the rest of the received data useless. Considering this disadvantage of the zero tree scheme under channel errors, our proposed designs concentrate on slicing the erroneous portion of the received data and using the rest effectively. In order to be able to slice out the erroneous

Figure 4-2: System block diagram for robust video communication

portion of the received data, the zero tree coder output is partitioned and each partition is separately channel coded and transmitted. The zero tree coder output partitioning is categorized into two types; *sequential partitioning* and *orthogonal partitioning*. The schemes based on both the types of partitioning are discussed and their performances are evaluated in chapter 6.

4.3.1.2 Stream Classification

The efficient utilization of bandwidth takes high priority in low bit rate applications. For this purpose it is useful to classify the compressed data and provide different degrees of protection, or UEP, to the data streams of different importance. In this work UEP is provided to the compressed stream at the inner channel coding stage. Hence the data classification is carried out right after the compression.

In the video configuration using MRME, the LL band of Y component in the IFC frame is of highest importance as the information in this band is later spread to other bands of the same frame and the next frames. Hence this band is given the highest protection when the RZCO scheme is used. Otherwise in general, the Y component of IFC frame is given high protection. The motion vectors take less number of bits than the corresponding residual frame, but are of more importance than the latter as they usually represent most of the energy in that frame. Hence the motion vectors are given better protection than Y component of residual frames. The same is applicable to the chroma components.

In the video coder using 3D-WT, the first frame is *dc* frame and hence has more information. However, it is not advantageous to provide higher protection to a larger part of the data which could lower the overall performance. Thus the protection to Y component of first frame is increased only when it is feasible.

Summarizing the classification for both MRME based and 3D-WT based video coders the order of importance is as follows:

- LL band of Y component of first frame
- LL band of Y component of subsequent frames
- Motion vectors (only for residual frames in MRME configuration)

- Non-LL bands of Y component of first frame
- Non-LL bands of Y component of subsequent frames
- LL bands of UV components of first frame
- LL bands of UV components of subsequent frames
- Non-LL bands of UV components of first frame
- Non-LL bands of UV components of subsequent frames

For monochrome video transmission, the chroma components are simply ignored.

4.3.2 Channel Protection

Error protection is an inevitable part of any communication system. In the scenario described in section 4.2, the video signal traverses two channels; a mobile channel and an ATM channel. The mobile channel introduces fading and shadowing, and results in bursts of errors in the transmitted bit stream. The ATM channel induces mainly the cell losses in the transmitted cell stream. Hence in this work we propose two stages of error protection, inner and outer stages, one before and one after the ATM cell construction stage.

The burst error conditions of mobile channel necessitates the interleaving in order to randomize the burst errors. Then the channel coder is used to correct these random errors. EEP is preferred at this stage for the following reasons.

- In order to provide UEP at the outer stage, the complete stream of any particular class needs to be fit into an integer number of cells or the last cell in the stream needs to be padded with dummy bits. This results in wastage of bandwidth and hence is not desirable.
- After the inner stage interleaving it is difficult to maintain the separation between the streams of different importance for UEP at the outer stage. Further, the variable length nature of the streams and fixed length nature of interleaver complicates the issue.
- The classification of data needs to be known at the mobile modem, i.e. after ATM cell construction, for UEP. This calls for interaction between the source coder and the outer stage coder, which is considered to be part of the modem. This again is not desirable as it customizes the modem design for this particular application.

Hence in this work the outer stage of channel coding is taken to be EEP and is considered to be part of the modem along with the outer stage interleaving. The span of interleaving is assumed to be about one frame size for MRME configuration and about four frame size for 3D-WT configuration. The sizes are chosen in order to maximize the performance keeping the delay to be within the tolerable limits.

Any uncorrected mobile channel errors in the header of the ATM cell results in rejection of the entire cell. Additionally the traffic conditions in the ATM channel offer certain cell loss rates. These lost cells appear as burst data erasures at the receiver. One way to recover the data in the lost cells is by providing error protection and interleaving to the payload portion of the ATM cell. As there needs to be an interleaver at the outer stage to help combat the mobile channel, we propose to use another interleaver of smaller size to help combat the ATM cell losses. The interleaver parameters need to be carefully chosen so as not to incur additional delay in the reconstructed video. The source encoder and variable precision quantization result in variable length of bit streams for each of the frame to be transmitted. In order to provide channel protection to these streams, we need the inner stage channel coder to fulfil two important criteria. First, it should be a variable length coder and the second, it should be able to provide different rates of coding for different classes of input data. Hence we chose to investigate the rate punctured convolutional coder (see Appendix-B) at inner stage as these coders are of variable length in nature and by puncturing process it is possible to easily control the coding rate using the same basic coder-decoder set. The use of these coders to dynamically change the coding rate of any stream is discussed further in section 7.2.3. The interleaver at the inner stage is adjusted to be half the average number of information bits in the span of outer stage interleaver so as to avoid any additional delay.

4.3.3 Rate Controller

The rate controller is used to dynamically allocate the bit budget and decide the channel coder rate for each of the compressed bit stream depending on the channel conditions. For this purpose the channel conditions are measured at the receiver and sent as feedback to the transmitter. This feedback information is further classified into number of channel states. The bit budget and channel coding rates are accordingly varied in order to

efficiently make use of the available bandwidth and better combat the channel variations. This is discussed in detail in chapter 7.

It should be noted that in this work the dynamic rate controlling is proposed at the application level or the user level instead of being carried out at the network level. This makes the system more robust and versatile as only the application knows the importance of any particular compressed stream. In [19], [27], [46], [55], [62], [79], [81], [87], [103], [114], [120], [141], [144], [148] and [152] attempts are made to control the video rate at the network level. As the network can not have any knowledge about the compression scheme or the importance of a particular stream, these methods can not be used in achieving the robust video communication. Besides, the basic aim of most of these reported works is controlling the traffic and not reliable communication.

4.3.4 Test Sequences

All the experiments performed in this work are carried out on all the five sequences; *miss america*, *susie*, *claire*, *salesman* and *carphone*. Due to space limitations and repeated performance behavior of the system, only one sequence per experiment is chosen to be presented. However, a different sequence is chosen at each experiment to illustrate the sequence independency of the system.

All the test sequences are at 10 frames/sec. The standard frame rates used in the literature are 5, 7.5 and 10 frames/sec. Some of the works by the European community also use 12.5 frames/sec. In this thesis we have used 10 frames/sec as it seems to be the best trade-off between the amount of *jerkiness* in the video and the quality achievable at around 64 Kbps.

4.4 Summary

In this chapter we defined the chosen problem for the thesis and looked at the system design aspects to solve this problem. In section 4.2 we defined the problem where we discussed the scenario chosen, previous work reported and the reasons for the selection of this problem. It was noted that the problem of robust video communication is still unsolved and this thesis makes an effort to arrive at an elegant solution to it. The important factors emphasized are robustness, low overhead and smooth bit rate variations. The

system design aspects to achieve this were briefed in section 4.3. The video coder comprises of the wavelet transform, MRME/3D-WT and the zero tree quantizer. It was noted that robust quantizer designs are required in order to improve the error resilience of the video coder. We propose two stages of error protection to combat burst errors in the mobile channel and cell losses in the ATM channel. The classification of compressed streams based on their importance and unequal error protection for better bandwidth utilization are proposed. The design details and the performance evaluations are discussed in the following chapters.

Chapter 5

Video Coder

5.1 Introduction

This chapter discusses the performance of the video coder configurations that are used in the system. The video coder configurations using multiresolution motion estimation (MRME) schemes and 3D wavelet transform (3D-WT) are discussed in section 5.2. Their performances at low bit rates are evaluated in section 5.3. It is shown that at low bit rates the MRME schemes perform differently than as reported in [151]. This is due to the fact that at low bit rates the available bit budget for motion vectors is limited and hence the scheme requiring a low bit budget with less disturbance to the subband correlation becomes the winner. The performance comparison of MRME and 3D-WT schemes is also provided. Note that all these performance measurements are without any channel considerations and hence involve only the video coder and decoder in the experiment.

5.2 Video Coder Configurations

Three important criteria to be satisfied in the video coder design are; VBR output, acceptable quality and gracefully degrading performance with increase in the compression gain. The first criterion is posed by the network transport conditions and varying mobile channel conditions, and the second and third criteria are as demanded by the end user.

The proposed video coder structure makes use of the two dimensional wavelet transform (2D-WT) to exploit the spatial redundancy in the input picture frames. In this work we are investigating both the temporal wavelet transform and the motion estimation-compensation techniques to exploit the temporal redundancy in the input image sequence.

The video coder configurations using these schemes are shown in Fig. 5-1. In both the configurations the zero tree coder is employed for quantization.

5.2.1 Wavelet Transform

In section 2.3.1 it was observed that the DCT based schemes suffer from blocky noise at low bit rates and the DWT based schemes are becoming popular for their good visual quality performance. Hence in this work we chose to examine the use of wavelet transform based video coder. The two dimensional wavelet transform (section 2.3.2) is realized using one dimensional wavelet filter banks and is used to exploit the spatial redundancy in the video frame. An extra one dimensional wavelet filtering is carried in temporal domain for 3D-WT.

A six tap Daubechies filter is used at 2D-WT and a two tap Haar filter is used at the temporal filtering of 3D-WT. These filters are perfect reconstruction filters [35], [90] and are short in length leading to lower complexity implementation and minimum abnormalities at the frame borders. These filters are orthonormal filters meaning they are their inverses too. Thus we need only one set of filters for both decomposition and reconstruction. They are also reported to provide good energy compaction and found to offer comparable performance in image compression applications when compared with other subband filters with much longer taps [147], [90]. It is always possible to slightly improve the performance using higher tap filters from other subband filter families [147], [1], [100] with techniques to reduce distortions at the frame border. However higher number of taps significantly increase the complexity of the video coder [29], [150].

5.2.2 Multiresolution Motion Estimation

In the multiresolution motion estimation (section 2.4.2) schemes, the motion vectors are estimated for the lowest resolution subband and then those for higher resolution subbands are obtained by refining the scaled versions of these vectors. As only the refining process is required at all the higher resolutions, the searching burden is significantly reduced when compared to that of a full frame motion search scheme. In this process, the MRME schemes efficiently exploit the subband structure and the correlation in motion information at different levels of resolution. At the same time, when the restrictions demand different MRME schemes can be used to trade the bit budget requirement with the

WT - Wavelet Transform

SW - SWitch, IFC - Intra-Frame Coding

MRME - Multi-Resolution Motion Estimation

MRMC - Multi-Resolution Motion Compensation

ZC - Zero tree Coder, IZC - Inverse ZC

AEC - Arithmetic Entropy Code, FM - Frame Memory

(a) Video coder using motion estimation

(b) Video coder using 3D wavelet transform

Figure 5-1: Structure of video encoder

performance. For example, the first scheme MRME1 requires encoding of the motion vectors only in the LL band. MRME2 requires motion vectors in four subbands and hence needs four times more bit budget compared to MRME1. In MRME3 and MRME4 it could be as much as number of subbands, in the pyramid, times the bit budget for LL band. Again, by adjusting the size of search window at higher resolutions in MRME3 and MRME4 different bit budget requirements can be met. Thus, in this work we chose to investigate the MRME schemes for motion estimation.

Another important factor to be considered in motion estimation is the fractional pel accuracy motion estimation which is presently used in standards like H.263. The fractional pel accuracy motion estimation provides finer sub-pixel motion estimation after interpolating the pixels in the reference frame by an appropriate factor. However, it is to be noted that both the search overhead and the motion vector overhead increase significantly with the accuracy. On the other hand, if the search window is not increased with the accuracy, then the amount of motion that is being catered by the motion estimation is reduced by an equivalent amount. Additionally, with MRME schemes, there are other problems in using fractional pel accuracy. As the subband structure provides doubled resolution at every stage, the accuracy or factor of interpolation increase from bottom of the pyramid to the top by a factor of two at each stage. This results in exponentially increasing computational complexity along the resolutions in the pyramid. Apart from this, due to large interpolation factor at the top of the pyramid, the interpolation itself tends to be loose resulting in lower performance compared to integer pel accuracy. Hence in this thesis it is decided to use only the integer pel accuracy motion estimation in all the experiments.

Along with four schemes suggested in [151] we propose a fifth scheme to be examined. In this scheme, MRME5, the motion vectors are estimated for the LL band at the lowest resolution. For all the directional bands in the lowest resolution, the LL subband motion vectors are used as bias for further refining. For the rest of the resolutions, the scaled version of the motion vectors from same directional subband in the previous resolution are used as bias for refining. This scheme is similar to MRME4 except that the bias vectors are derived from the motion vectors in the previous resolution instead of the LL band. For the same reason the scaling factor for motion vectors in MRME5 is always two.

In order to cope with the low bandwidth availability, it is proposed to use the motion vectors computed for luminance (Y) component of the picture frames at the corresponding chroma (UV) components also.

5.2.3 Zero Tree Coder

In this chapter we are using the zero tree technique discussed in section 2.5.2 for quantizing the wavelet coefficients. This coding is conceptually similar to the combination of successive approximation quantization (SAQ) and bit plane coding. One major advantage of this technique is that the encoding can be stopped at any precision at any coefficient and the receiver will still be able to decode the frame for that precision up to that coefficient, with all the remaining coefficients being reconstructed to the previous lower precision. Hence the bit rate variations from frame to frame can be as smooth as required providing a 'true VBR' output. Other advantages of this coder are; it is fairly general leading to easy implementation and provides gracefully degrading quality as the bit rate is decreased.

5.3 Performance Evaluation

5.3.1 Experimental Setup

In this section the evaluation of different MRME schemes and 3D-WT scheme at low bit rates is provided. Additionally the new MRME scheme, MRME5, suggested in section 5.2.2 is compared with the other MRME schemes that are reviewed in section 2.4.2. In these experiments there is no channel consideration as these are only the evaluation of compression schemes for their basic performance. In other words a clean channel is assumed in the experiment. Thus the experimental setup (not shown) comprises of video coder and decoder simply connected back to back.

The input video frames are divided into blocks of eight called *group of frames (GOF)*. In MRME configuration the first frame in each GOF is intraframe coded (IFC). In the successive frames, only the difference frame after removing the estimated motion activity is encoded. The IFC enables the coder and decoder to be refreshed at the beginning of each GOF and hence limits the error propagation. At the same time it should be noted that the IFC requires higher bit budget. In the case of 3D-WT configuration, the 1D temporal

wavelet filtering is performed over GOF. The size of GOF in both the configurations is a compromise among performance, error propagation and the delay involved. Lower size of GOF increases the number of intraframes and hence reduces the error propagation. At the same time it results in lower achievable video quality due to higher bit budget requirement of intraframe coding. In reverse, higher GOF size increases the error propagation, but also increases the achievable video quality due to reduced number of intraframe coding. Hence as a trade-off, a GOF size of eight frames is chosen in this work.

In all the following illustrations, except for Fig. 5-10, QCIF resolution (176x144 pixels) monochrome *miss america* video is used at a frame rate of 10 frames/sec. A three level wavelet decomposition is performed on each of the input frames. In all the MRME schemes 2x2 size motion blocks are used at the lowest resolution. In the 3D-WT scheme a three level one dimensional wavelet decomposition is carried out in the temporal domain. The compressed bit stream is decoded at the decoder and the performance is evaluated.

5.3.2 Simulation Results

Fig. 5-2 through Fig. 5-6 shows the performances obtained at five different output bit rates. It can be seen that at every eighth frame, all the MRME schemes perform the same because of an intraframe coding inserted to make the schemes more practical.

The bit rate of 85 Kbps, Fig. 5-2, is chosen as it is the minimum bit rate required for the schemes MRME3, MRME4 and MRME5 to operate with reasonable reconstruction quality. At this bit rate, it is seen that the scheme MRME1 performs 2 dB better than the rest of the schemes. The scheme MRME2 performs 0.5-1dB better than MRME3 and the scheme MRME4 performs worst compared to other schemes reported in [151]. The reasons for these results to be in contradiction to those reported in [151] are; first, these evaluations are carried out at low bit rates where every few hundred bits contribute significantly to the reconstruction quality which is not the case at a bit rate of few Mbps. The MRME1 needs to code only a few motion vectors and hence needs lesser bit budget for motion vectors. This makes available more bits for difference frame coding and thus results in better performance. Secondly, in MRME1 the motion vectors for all the non-LL bands are appropriately scaled versions of the motion vectors at the LL band. This keeps the inter-resolution subband correlation in the difference frame mostly undisturbed. Hence

Figure 5-2: Comparative performance of MRME schemes at 85 Kbps.

Figure 5-3: Comparative performance of MRME schemes at 128 Kbps.

Figure 5-4: Comparative performance of MRME schemes at 160 Kbps.

Figure 5-5: Comparative performance of MRME schemes at 200 Kbps.

the zero tree quantizer made available with more bit budget provides better quality. In the case of other schemes the correction factors for the motion vectors at non-LL bands disturb the inter-subband correlation. This makes the zero tree quantizer to need more bits to code the difference frame for the same quality. Hence for a given low bit rate the reconstruction quality offered by MRME1 is higher. The same argument holds good for the MRME2 performing better than the MRME3. Relatively, the inter-resolution subband correlation is retained slightly better in the difference frame in MRME3 than that in MRME4. This is due to the fact that all the higher resolution motion vectors in MRME3 had the same bias motion vectors from the LL band. However note that at low bit rates there is only a very slight difference in performance of these two schemes.

The new scheme MRME5 performs roughly same as MRME3 and MRME4 due to similar bit budget and complexity requirements. The difference in bias arise mainly because of difference in bias vectors at higher resolutions.

The performance curves for 3D-WT scheme at different bit rates is given in Fig. 5-7. It can be seen that the performance of 3D-WT is fairly flat. The bit budget allocation is prefixed and is in proportion to the estimated energy content in each frame. The comparison of 3D-WT and MRME1 schemes at the average output bit rate of 64 Kbps is illustrated in Fig. 5-8. The bit budget allocation in this comparison are shown in Fig. 5-9. It should be noted that, though the 3D-WT scheme perform superior to the MRME schemes it comes with the price of large delay and large variations in the output bit rate from frame to frame.

The results of experiments performed on the other four sequences *claire*, *susie*, *salesman* and *carphone* exhibited similar performance behavior. The results for *carphone* sequence at 85Kbps are shown in Fig. 5-10.

5.4 Summary

In this chapter we discussed the proposed video coder configurations and their performance evaluation. The video coder configurations make use of two dimensional wavelet transform for spatial redundancy removal and either MRME or 1D-WT for temporal redundancy removal. The zero tree quantization is chosen as it efficiently exploits the pyramid structure of the wavelet subbands and provides gracefully degrading video quality with decrease in the output bit rate.

Figure 5-6: Comparative performance of MRME algorithms on at 300 Kbps.

Figure 5-7: Performance curves for 3D wavelet transform at low bit rates

Figure 5-8: Comparative performance of 3D-WT and MRME1 schemes at 64 Kbps.

Figure 5-9: Bit budget allocation for 3D-WT and MRME1 schemes at 64 Kbps.

Figure 5-10: Comparative performance of MRME schemes at 85 Kbps for *carphone* sequence

At low bit rates, the MRME1 scheme outperforms the rest of the three MRME schemes by about 2 dB peak signal to noise ratio (PSNR). The MRME2 performs better than MRME3 and MRME4, but is inferior to MRME1. The new scheme MRME5 performs roughly same as MRME3 and MRME5. As the bit rate is increased, the difference in performance of these schemes reduce as the exact motion compensation gains significance. It is seen that all the five schemes perform equally well at around 300 Kbps. After this mark, the MRME3 scheme is expected to perform better than the other schemes as shown in [151]. The scheme MRME5 is also expected to perform at least as good as MRME3 at high bit rates because of improvement in bias vectors at each higher resolution. As this work concentrates on video communication below 300 Kbps, we have chosen to use MRME1 scheme in the remaining experiments.

The 3D wavelet transform is shown to perform superior to the motion estimation schemes. It can also be observed that the performance variation is smooth in the case of 3D-WT which is a desirable feature. However the main disadvantage of the 3D-WT scheme is the large delay involved, which is not tolerable in any two-way communication applications. Another disadvantage of 3D-WT is that the energy content varies drastically from frame to frame causing large variations in the output bit rate. However this can easily be circumvented by making use of the large inherent delay in the scheme and buffering the output bit stream before transmission to smooth the output bit rate.

Chapter 6

Robust Quantization

6.1 Introduction

This chapter presents a discussion on the proposed robust quantization schemes and their performance evaluations. The new robust designs of the zero tree quantizer which are based on symbol level and coefficient level partitioning are discussed in section 6.2. Section 6.3 deals with the performance evaluation of these new designs of zero tree coder under both the video coder configurations that were discussed in section 5.2. The evaluations are carried out under different channel conditions and compared with the basic zero tree coder. It is shown that the basic scheme proposed in [129] and the schemes based on the sequential partitioning perform adversely and are unpredictable under noisy channel conditions. On the other hand the scheme based on orthogonal partitioning performs well and is reliable under noisy channel conditions. The performance of the scheme based on orthogonal partitioning is compared with H.261 based coder and is shown that the former has better robustness against the channel errors. Finally, a summary of the discussion is provided in section 6.4.

6.2 Robust Quantizer Designs

In this thesis we are investigating the zero tree technique (section 2.5.2) for quantizing the wavelet coefficients. We term the scheme proposed in [129] as the *basic zero tree coder* (BZC) from here on. The main disadvantage of BZC is its intolerance to errors. Hence any uncorrected error results in invalid symbol input at the zero tree decoder making the rest of the received data useless. Thus in the presence of uncorrected errors the basic zero tree

quantizer design often results in significant loss of data as illustrated in Fig. 6-1 (a). This makes the quality of received picture highly varying and unpredictable.

Our proposed designs concentrate on slicing the erroneous portion of the received data and using the rest effectively. In order to be able to slice out the erroneous portion of the received data, the zero tree coder output is partitioned and each partition is separately channel coded and transmitted. One important criteria to satisfy while partitioning is the independence of all the positions achieved. This means that when the erroneous part is sliced out, the next error free part should always improve the picture quality. The number of partitions that can be achieved is mainly limited by the overhead required to synchronize the states of encoder and decoder. Two types of partitioning are considered; *sequential partitioning* and *orthogonal partitioning*. The schemes based on these types of partitioning are discussed below.

6.2.1 Sequential Partitioning

The sequential partitioning is achieved by dividing the symbol output of the zero tree encoder into many parts as shown in Fig. 6-1 (b). In order to simplify the synchronization process, the partitions are always made at the end of an iteration and at each partition the starting threshold is communicated to the receiver. In this work we considered two schemes based on the sequential partitioning. In the first scheme, RZCS1, the encoder starts encoding the input frame at the beginning of the first partition. At the subsequent partitions the difference of the original input frame and the locally reconstructed frame is encoded. The reconstruction is carried out in an accumulative fashion. In the second scheme, RZCS2, the significance map consisting of position and sign of all the coefficients already found significant, is communicated to the receiver at the beginning of each partition. Hence the coding and decoding are continuous as in the original scheme.

The coding schemes RZCS1 and RZCS2 are similar to the layered coding scheme suggested in [84] except that the local decoder gets all the layers of information for decoding. This avoids the degradation in the picture quality due to increase in the number of layers or their information content. However the disadvantage is that, each partition contains the information about the coefficients found significant in previous partitions along with those found significant in the current partition. This results in inter-dependence

(a) Channel error effect on BZC

DP - Symbols from dominant pass

SP - Symbols from subordinate pass

(b) Sequential partitioning of zero tree coder output

(c) Orthogonal partitioning of wavelet frame into sets of arbitrary shapes.

Figure 6-1: Robust zero tree quantizer design

of partitions. This means that if a partition is lost due to uncorrected errors, the next partition arriving at the decoder need not necessarily reduce the mean square error in the reconstructed frame. This causes the picture quality to be varying rather unpredictably and hence making the schemes unreliable under noisy channel conditions. Another disadvantage of these schemes is the large overhead involved at the beginning of each partition, i.e. in restarting the coder in RZCS1 and encoding the initial significance map in RZCS2. This causes a significant degradation in the base performance of the schemes. The overhead becomes unaffordable at bit rates of 32 Kbps and lower, and hence making these schemes inapplicable.

6.2.2 Orthogonal Partitioning

The orthogonal partitioning is carried out by dividing the wavelet coefficients, instead of zero tree symbol stream, into orthogonal sets. In order to guarantee the independence of each set, it is necessary that the entire coefficient tree starting from the LL band belongs to one set. Thus only the LL band coefficients need to be partitioned into several sets. Each set is then independently zero tree coded and transmitted. The orthogonal partitioning also minimizes the overhead requirement as there is no history between the encoded sets. This process of orthogonal partitioning is illustrated in Fig. 6-1 (c) for arbitrary shapes of sets.

The main disadvantage of orthogonal partitioning is that when a set is lost due to uncorrected errors, all the coefficient trees in that set will be either lost or truncated in magnitude. This appears as black (or gray) square patches in the reconstructed picture and severely degrade the visual quality. To improve the quality, the following measures are suggested in this work.

- The LL band is encoded independently with higher channel protection and communicated to the receiver. This ensures the proper reception of most of the low frequency energy.
- The partitioning is performed on the rest of the frame with the coefficient trees originating from the non-LL bands of the lowest resolution instead of the LL band. This means that the partitions need to be carried out separately at three directional subbands at the lowest resolution.

- The trees chosen for each set must be separated by maximum possible Euclidean distance so that if a set is lost, the trees in that set will be interpolated by the surrounding trees from other sets during the inverse wavelet transformation. Further, the trees that are close by are allocated to different sets with maximum possible difference in their indices. This is done in order to tolerate the loss of two or more consecutive sets due to large burst errors. This tree allocation problem is similar to the frequency allocation problem [36], [37], [82] in cellular scenario. Assuming one frequency per cell and a fixed number of available frequencies, the cells using the same frequency are required to be positioned at a maximum possible distance in order to reduce the cochannel interference. Further, the frequencies allocated to the neighboring cells need to be farther separated in the spectrum in order to minimize the adjacent channel interference. In this work the best possible solutions for various values of number of partitions are pre-computed, hand picked and stored as tables to be used at the time of quantization.
- The trees at the same spatial location in different directional subbands are allocated to different sets separated by a maximum possible number of sets. This avoids the loss of all three trees at the same spatial location when any one set is lost. The lost tree at one of the subband will be interpolated by the other two trees during the inter-subband interpolation at inverse transform stage. This measure in the tree allocation can be achieved by first formulating a tree allocation table as described above and then by simply offsetting the table entries by a third of total number of partitions at each directional subband. This way trees at the same spatial location in each direction is placed in farthest possible sets.

In the coder based on orthogonal partitioning, RZCO, a zero-tree-like coding is used for the LL band. The zero trees and their construction are absent from the dominant pass as we need to code only the LL band coefficients. Hence the *isolated zero* and the *zero tree root* symbols are merged into a *zero* symbol leading to a 3 symbol encoding at the dominant pass. The subordinate pass is used as before to improve the quantization precision by a factor of two.

6.3 Performance Evaluation

6.3.1 Experimental Setup

The experimental set up for comparing the different zero tree coders is shown in Fig. 6-2. The four schemes considered are the basic zero tree coder (BZC), the zero tree coders based on sequential partitioning (RZCS1 and RZCS2) and the zero tree coder based on orthogonal partitioning (RZCO). As the rate controller is absent from the setup, the source rate is varied within the GOF in a predetermined fashion. The channel coding rates are also prefixed for all the compressed streams. In this work it is assumed that the outer channel coder stage is within the mobile modem. Hence the bit rates, bit error rates (BER) and cell loss rates (CLR) used are at the output of ATM cell construction stage.

In all experiments, the QCIF resolution (176x144 pixels) monochrome *miss america* video sequence is used at 10 frames/sec. The performance evaluation of the schemes are carried out at the output bit rates of 32 Kbps, 64 Kbps, 128 Kbps and 256 Kbps. Both the video coder configurations (see section 5.2) using MRME1 scheme and 3D-WT scheme are considered. The size of GOF is chosen as eight video frames as in the previous experiment (see section 5.3.1). A three level wavelet decomposition is performed on each of the input frames. In the case of motion estimation, the MRME1 scheme at one pixel accuracy is used. The motion blocks at the lowest resolution are chosen to be of 2x2 size. In the 3D-WT scheme a three level one dimensional wavelet decomposition is carried out in the temporal domain. The number of partitions in each of the proposed zero tree coder designs vary at each frame in proportion to the bit budget allocated for that particular video frame.

The evaluations are carried out for both the narrowband and the broadband channel conditions which are discussed in section 3.4. The UEP schedule, in terms of the convolutional coder rates, for each of the channel conditions considered is shown in Table 6-1. The UEP schedule is kept the same for all the schemes. It can be seen that the channel protection is increased at higher channel error conditions and lowered at lower channel error conditions. In the channel model the seed of the random number generator is changed at each experiment to bring in the flavor of statistical behavior of the channel.

Figure 6-2: Experimental setup to compare different zero tree coder designs

UEP schedule	10^{-2} CLR (narrowband) and 5×10^{-3} BER (broadband)	10^{-3} CLR (narrowband) and 10^{-3} BER (broadband)
Motion vectors	1/2	2/3
LL band (first frame in GOF)	1/2	2/3
LL band (subsequent frames in GOF)	2/3	3/4
Zero tree quantized symbol stream	2/3	3/4

Table 6-1: UEP schedule for different channel conditions

6.3.2 Simulation Results

6.3.2.1 Performance in Narrowband Channel

The narrowband channel is modeled using ON-OFF states as discussed in section 3.4.1.1. The mean of the OFF period is set at three cells while that of the ON period is adjusted to the required CLR. The cell loss rates of 10^{-2} and 10^{-3} are considered in this experiment.

MRME Based Video Coder

The obtained results for the MRME based video coder are shown in Fig. 6-3 through Fig. 6-6. At CLR of 10^{-2} , RZCO outperforms the rest of the schemes which is clearly seen at deep fade conditions in the plots shown. The schemes RZCS1 and RZCS2 are dogged by their large overhead requirements and non-independence of the partitions. These factors affect their performance, making them no better than BZC most of the time. At CLR of 10^{-3} there are less uncorrected errors seen at the zero tree decoder and hence all the four schemes perform equally well. It is to be noted that the performance of RZCO is very close to that of BZC because of the lower overhead requirement. The schemes RZCS1 and RZCS2 tend to incur exponentially increasing overhead as the number of partitions are increased. This is clearly visible at high bit rates as shown in Fig. 6-5(b) and Fig. 6-6(b).

At many frames it is observed that there is little or no partitioning carried out in the schemes RZCS1 and RZCS2, due to lack of bit budget available for the overhead. However in RZCO the partitioning is carried out for all frames and larger number of partitions are achieved at higher bit rates. This makes the overall overhead in RZCO larger than that in RZCS1 or RZCS2. This further results in the lower peak performance by RZCO but with the advantage of increased robustness. This is visible at all GOF at 32 Kbps and some GOF at other bit rates.

It is noted that occasionally the scheme RZCO perform inferior to BZC at certain GOF, mostly for the entire period of those specific GOF. It is observed that at these GOF the intraframe LL band is corrupted degrading the performance at subsequent frames. However the probability of this happening is quite low as the intraframe LL band is always provided with better channel protection. As the bit rate is increased, the scheme RZCO

(a) 10^{-2} CLR (b) 10^{-3} CLR

Figure 6-3: Performance comparison of the zero tree coders in MRME video configuration at 32 Kbps under narrowband channel conditions

(a) 10^{-2} CLR (b) 10^{-3} CLR

Figure 6-4: Performance comparison of the zero tree coders in MRME video configuration at 64 Kbps under narrowband channel conditions

(a) 10^{-2} CLR (b) 10^{-3} CLR

Figure 6-5: Performance comparison of the zero tree coders in MRME video configuration at 128 Kbps under narrowband channel conditions

(a) 10^{-2} CLR (b) 10^{-3} CLR

Figure 6-6: Performance comparison of the zero tree coders in MRME video configuration at 256 Kbps under narrowband channel conditions

performs better due to increase in number of partitions which in turn increases the error tolerance.

The visual quality comparison of different schemes is given in Fig. 6-7 for 64 Kbps. Only the best among RZCS1 and RZCS2 is shown. It can be seen that the picture is muffled in the case of BZC and RZCS2 due to the loss in difference frames, i.e. frame 12-16 in Fig. 6-4 (a). This can be clearly seen at the left side eye lid and the space between lips. The scheme RZCS2 performs worse than BZC with added distortion at the bottom-center of the picture. The picture quality in the scheme RZCO is better than the rest of the schemes.

3D-WT Based Video Coder

The results for 3D-WT based video coder are shown in Fig. 6-8 through Fig. 6-11. It can be seen that the scheme RZCO performs consistent compared to the other three schemes at 10^{-2} CLR and performs close to BZC at 10^{-3} CLR. This again proves that the parameters in the scheme RZCO can be adjusted for reliability under worst channel conditions and can be changed to require low overhead under low error conditions. The visual quality comparison is provided in Fig. 6-12. It is noted that the errors in the early part of the stream has completely degraded the performance of BZC while the error resilience is seen effective in the scheme RZCO. The performance of RZCS1 is only slightly better than that of BZC.

6.3.2.2 Performance in Broadband Channel

The broadband channel is simulated for the bit error rates of 5×10^{-3} and 10^{-3} using the CODIT environment (see section 3.4.1.2). In this channel any ATM cell with more than one residual error in the header is considered to be corrupted and is thrown away. This contributed to a CLR of 0.76×10^{-2} at BER of 10^{-3} and a CLR of 1.72×10^{-2} at BER of 5×10^{-3} .

MRME Based Video Coder

The simulation results for the MRME1 based video coder are shown in Fig. 6-13 through Fig. 6-16. As in the narrowband channel conditions, the scheme RZCO proves its error

(a) Original

(b) BZC

(c) RZCS2

(d) RZCO

Figure 6-7: Frame #15 of *miss america* sequence: Comparison between the original and the reconstructed output of different zero tree coder schemes (MRME configuration) under narrowband channel conditions with 10^{-2} CLR at 64 Kbps.

(a) 10^{-2} CLR (b) 10^{-3} CLR

Figure 6-8: Performance comparison of the zero tree coders in 3D-WT video configuration at 32 Kbps under narrowband channel conditions

(a) 10^{-2} CLR (b) 10^{-3} CLR

Figure 6-9: Performance comparison of the zero tree coders in 3D-WT video configuration at 64 Kbps under narrowband channel conditions

(a) 10^{-2} CLR (b) 10^{-3} CLR

Figure 6-10: Performance comparison of the zero tree coders in 3D-WT video configuration at 128 Kbps under narrowband channel conditions

(a) 10^{-2} CLR (b) 10^{-3} CLR

Figure 6-11: Performance comparison of the zero tree coders in 3D-WT video configuration at 256 Kbps under narrowband channel conditions

(a) Original

(b) BZC

(c) RZCS1

(d) RZCO

Figure 6-12: Frame #73 of *miss america* sequence: Comparison between the original and the reconstructed output of different zero tree coder schemes (3D-WT configuration) under narrowband channel conditions with 10^{-2} CLR at 64 Kbps.

(a) 5×10^{-3} BER (b) 10^{-3} BER

Figure 6-13: Performance comparison of the zero tree coders in MRME video configuration at 32 Kbps under broadband channel conditions

(a) 5×10^{-3} BER (b) 10^{-3} BER

Figure 6-14: Performance comparison of the zero tree coders in MRME video configuration at 64 Kbps under broadband channel conditions

(a) 5×10^{-3} BER (b) 10^{-3} BER

Figure 6-15: Performance comparison of the zero tree coders in MRME video configuration at 128 Kbps under broadband channel conditions

(a) 5×10^{-3} BER (b) 10^{-3} BER

Figure 6-16: Performance comparison of the zero tree coders in MRME video configuration at 256 Kbps under broadband channel conditions

resilience in the broadband channel through its robust performance at 5×10^{-3} BER. At 10^{-3} BER the scheme requires low overhead and performs close to BZC.

3D-WT Based Video Coder

The results for 3D-WT based video coder are shown in Fig. 6-17 through Fig. 6-20. It can be seen that at BER of 5×10^{-3} the scheme RZCO performs better than the rest of the schemes. The low overhead requirement of RZCO is proved again with its performance close to that of BZC at 10^{-3} BER. The visual quality comparison of the schemes is given in Fig. 6-21 where it can be seen that the scheme RZCO performs superior.

6.3.2.3 Comparison with H.263 Based Coder

In this section we compare RZCO scheme with the proposed H.26P/M standard [66] (see section 2.6.3) which is a mobile adaptation of the H.263 standard. In order to make the comparisons fair, the frame rate is set at 5 frames/sec and the bit rate is set at 32 Kbps. The experiment is conducted under the narrowband channel environment at 10^{-2} CLR, as the DECT channel is used in [66]. The comparison of RZCO in both the MRME and the 3D-WT configuration is illustrated in Fig. 6-22. From the figure it is clear that the scheme RZCO is able to combat the channel errors more effectively than H.26P/M coder.

The RZCO coder performs superior to H.26P/M coder due to two main reasons. First, because of orthogonal partitioning, the reconstructed video tends to have better quality even after losing bunch of consecutive cells. Second, the error propagation is reduced to the current GOF. In the case of H.26P/M coder, the error immunity at the source coder is absent. Additionally, the insertion of an explicit intraframe requires huge bit budget in H.263. To alleviate this problem in H.263, intra-motion-block coding is used instead. However in the worst behaved channels like mobile channel, the intra-motion-block coding results in extended error propagation as seen in Fig. 6-22.

6.4 Summary

In this chapter we presented new error resilient designs for the zero tree coder. These designs were based on the sequential (symbol level) and the orthogonal (coefficient level)

(a) 5×10^{-3} BER (b) 10^{-3} BER

Figure 6-17: Performance comparison of the zero tree coders in 3D-WT video configuration at 32 Kbps under broadband channel conditions

(a) 5×10^{-3} BER (b) 10^{-3} BER

Figure 6-18: Performance comparison of the zero tree coders in 3D-WT video configuration at 64 Kbps under broadband channel conditions

(a) 5×10^{-3} BER (b) 10^{-3} BER

Figure 6-19: Performance comparison of the zero tree coders in 3D-WT video configuration at 128 Kbps under broadband channel conditions

(a) 5×10^{-3} BER (b) 10^{-3} BER

Figure 6-20: Performance comparison of the zero tree coders in 3D-WT video configuration at 256 Kbps under broadband channel conditions

(a) Original

(b) BZC

(c) RZCS1

(d) RZCO

Figure 6-21: Frame #74 of *miss america* sequence: Comparison between the original and the reconstructed output of different zero tree coder schemes (3D-WT configuration) under broadband channel conditions with 5×10^{-3} BER at 64 Kbps.

Figure 6-22: Comparing performance of RZCO with that of H.26P/M

partitioning. The schemes were used in both the video coder configurations employing MRME1 and 1D-WT for temporal redundancy removal. The evaluation of the schemes under narrowband and broadband channel conditions were carried out.

From the results it is clear that the schemes employing the sequential partitioning does not perform well as the independence of partitions is not guaranteed. Apart from this, these schemes require large overhead leading to inferior base performance at a given bit rate. The orthogonal partitioning has the inherent property of orthogonality among the partitions and hence is reliable under noisy channel conditions. This fact is proven in the performance at 10^{-2} CLR and 5×10^{-3} BER. Additionally, the individual partitions do not possess any history which enables them to be coded with little or no overhead. In other words, for the same overhead, or given bit budget, the number of partitions achieved in RZCO is more than that in RZCS1 and RZCS2 which results in increase in error resilience. This is clear from the view graphs for high bit rates. The visual or subjective quality performance of RZCO is also shown to be significantly better than the other schemes.

Further, the RZCO scheme is compared with the H.26P/M coder in section 6.3.2.3 and is shown that the RZCO is superior to H.26P/M coder in combating against the channel errors.

Chapter 7

Adaptive Dynamic Rate Control

7.1 Introduction

It is clear from the discussions in the previous chapter that by adjusting the parameters in the RZCO scheme, it is possible to achieve robust video communications under high error conditions and at the same time, it is possible to reduce the overhead and provide improved picture quality at low error conditions. Given this scheme, there is a need for an intelligent decision box which can adjust the parameters depending on the channel conditions. This is achieved using an adaptive rate controller working on the principle of feedback on channel condition. The adaptive rate controller scheme proposed in this thesis is discussed in section 7.2 and the performance evaluation is detailed in section 7.3. The rate controlling scheme is tested for the video coder configurations discussed in chapter 5. The evaluation is carried out for various channel conditions at different bit rates. The summary of the discussion is provided in section 7.4.

7.2 Dynamic Rate Controller

The dynamic rate controller (DRC) is used to dynamically allocate the bit budget and determine the channel coder rate for the compressed bit stream depending on the channel conditions. The channel condition is measured at the receiver and sent as feedback to the transmitter. This feedback information is further classified into number of channel states. The bit budget and channel coding rates are accordingly varied in order to efficiently make use of the available bandwidth and better combat the varying channel conditions.

The bit budget allocation discussed in this section is meant for combined requirements of

the source and channel coding of a stream. The source coder rate is automatically limited by the channel coder rate for a given overall bit budget.

The important factors in rate controlling are:

- Network connection agreements in terms of average bit rate, peak bit rate, maximum time duration over which the peak bit rate can be sustained.
- Feedback on channel condition from the receiver.
- Quality of service (QOS) requirements in terms of PSNR of the reconstructed video.

Since communication robustness is considered to be more important in this work, the order of priority for these factors is in the same order as they are listed above.

7.2.1 Channel Measurement and Classification

Channel errors result in the loss of ATM cells at the receiver. Hence the count of number of cells lost within a certain time period provides a measure of channel condition. Thus, we choose to measure the channel conditions in terms of the number of cells lost at the receiver within the inner interleaver span. This is monitored by indexing the ATM cells before transmission and detecting the holes in the cell sequence at the receiver. Only the cell loss rate measured over a period equal to, or longer than, the interleaver span makes sense as there is a fixed delay involved in the interleaver before the ATM cell construction. In this work it is decided to sense the channel, and feedback, at every one half the picture frame duration.

It should be noted that the cell indexing results in extra overhead in each cell and hence reduces the effective payload of the cell. Additionally, any error in the index causes the entire cell to be in error, hence increasing the overall cell loss error. The measured channel condition is compared with different error thresholds to mark the channel into different states. With this, the channel can be seen as hopping from state to state, with each state being associated with a range of error conditions.

7.2.2 Dynamic Bit Budget Allocation

7.2.2.1 Information Spread

In the case of MRME scheme, loss of cells in intraframe (or reference frame) result in poor quality throughout that particular GOF. The proposed solution in this thesis for such problems is to spread the information over many frames than trying to code the same in one frame. To achieve this, when the conditions are not favorable, the rate controller allows the intraframe coding (IFC) to be done coarsely. The coarse quantization of an intraframe is compensated by allowing more number of bits during the coding of subsequent interframes. The spread of information over many frames adds to the robustness of the system and has the advantage of smoothing the bit rate variations among the picture frames. The smooth bit rate variation at the source eliminates the requirement of a bit rate smoothing filter at the network level. However the price of this spreading is a lower quality achieved due to lower bit budget. Thus a trade-off is required between system robustness and bit rate variations. A positive note to this problem is that, in the previous experiment it was observed that the difference in performance of IFC compared to the peak is only about 2-3 dB of PSNR which is tolerable. The subjective performance varies noticeably but was observed to be very much in the tolerable limits.

When the 3D-WT configuration is used, the above information spread is not possible. This is due to the fact that the energy content of the first frame, or *dc* frame, is high and it has to be communicated to the receiver. However the 3D-WT schemes possess the inherent delay of one complete GOF. This can be exploited by buffering the compressed and coded bit stream of the entire GOF before transmission in order to smooth the output bit rate.

7.2.2.2 Bit Budget Variations

In order to obey the network connection agreements, the instantaneous bit rate for any picture frame is derived from the average over a moving time window called as "*adaptation window*". A moving window is chosen because a fixed window can not provide smooth variations under varying channel conditions unless the window size is very long. Another disadvantage of fixed window is that, if there is a severe channel condition towards the tail end of the window, it will not be possible to provide the required

high protection due to the lack of bit budget. When a moving window is used, these problems are easily circumvented as the available bit budget at any frame is derived from the average over one complete window.

The adaptation window size has to be an integral multiple of GOF in order to take care of the variations within the GOF. At the beginning of each GOF, the bit budget distribution for the entire GOF is computed, taking into account the currently available channel condition. This bit budget schedule is stored and used throughout the current GOF except when there is change in the channel condition. If the channel changes to a lower error condition, then only the channel coding rate is lowered which automatically increases the source coder rate for the same bit budget. On the other hand, if the channel changes to a higher error condition, the bit budget for the frame is recomputed under the new time window. This is done so as to maintain the QOS requirements. The flow chart for the dynamic rate control is illustrated in Fig. 7-1. Within GOF, the bit budget allocation is performed as per the amount of information-spread specified in the MRME configuration and as per the expected energy distribution in the 3D-WT configuration. However these have to be in accordance with the currently available bit budget for the entire GOF and the sustainable peak bit rate. Hence the following makes the important factors in deciding the bit budget for any compressed stream.

- Current channel condition as received from the receiver
- Available bit budget for the present frame
- Energy content in the stream (in 3D-WT configuration)
- Information spread (or bit rate variations) specified
- Quality of service (QOS) requirements

7.2.3 Dynamic Scheduling of UEP

The following are the important factors in deciding the degree of error protection to be provided to any compressed stream.

- Current channel condition
- Previous channel condition
- Importance of the stream as specified by the stream classification

Figure 7-1: Flow chart for dynamic rate control

Check_channel() function in the above flow chart

Figure 7-1: (continued...) Flow chart for dynamic rate control

WIN_BB is the bit budget available in the current adaptation window

GOF_BB is the bit budget available in the current GOF

FRM_BB is the bit budget available in the current frame

WIN_SZ is the size of the window in terms of number of GOF

AVE_GOF_BB is the average bit budget available for a GOF

$\text{AVE_GOF_BB} = (\text{ave. bit rate} / \text{frame rate}) * \text{number of frames in GOF}$

Figure 7-1: (*continued...*) Flow chart for dynamic rate control

- Available bit budget for the stream
- Quality of service (QOS) requirements

The UEP schedule is initially decided for all the streams in the current frame and is altered as and when the channel changes. If the channel worsens then the channel coding rate is improved to provide better protection. On the other hand if the channel improves, the degree of protection is lowered only by one level of coding rate. This is required in order to keep the system robust against consecutive short spurs of deep channel fading. The channel coding rate scheduling is illustrated in *Check_channel()* function in Fig. 7-1. This kind of varying error protection as per the channel conditions improve the bandwidth efficiency of the system keeping the robustness intact.

7.3 Performance Evaluation

7.3.1 Experimental Setup

The experimental setup to evaluate the dynamic rate controlled video communication over noisy channel is shown in Fig. 7-2. The RZCO scheme is considered in this experiment as it was shown, in chapter 6, as the only reliable design of zero tree quantization under noisy channel conditions.

The main assumption in this experiment is that the channel varies slowly so that the measured channel conditions can be used as it is without any further processing. This assumption is necessary due to the reasons of the restriction on the channel measurement (see section 7.2.1) and the absence of long term correlation in the measured channel. For this purpose the parameters of the narrowband model are set to make the mean values of the ON and OFF times longer. This however affects the performance of RZCO as it increases the probability of loss of LL band and the probability of loss of multiple successive orthogonal partitions. It also increases the effort and efficiency required at the dynamic rate control. In this experiment the mean value of OFF period is increased to 12.5 msec (roughly two cell period at 64 Kbps) as it is sufficient to cater for typical indoor channel without incurring very long burst errors. However, the price of increasing the burst lengths is the increased failure rates at the channel decoder. This causes increased degradation in the performance of the non-rate-controlled system.

Figure 7-2: Experimental setup to evaluate the adaptive rate controlled video communication system.

Another important assumption in the experiment is the perfect feedback channel. The performance degradation due to noisy feedback channel is not addressed in this work.

The ATM cell indices, used for the CLR measurement, are provided with single error correction capability. However in the broadband conditions the cell loss errors increase compared to those in the previous experiments in chapter 6 due to the possibility of multiple errors in the cell index bits.

The sustainable peak bit rate is set at 75% more than the average bit rate with its sustainable duration not to exceed one picture frame period. The peak bit rate duration is chosen so as to make it possible to further smoothen the bit rate, if required, by using a buffer of one picture frame size.

7.3.2 Simulation Results

7.3.2.1 Adaptation Window length

It is essential to optimize the length of the adaptation window in the dynamic rate control for efficient rate control and best performance. Hence the window lengths of 3, 5, 7, 10 and 15 GOF at 10^{-2} CLR under narrowband channel condition and 5×10^{-3} BER under broadband channel condition are tested. The monochrome video sequence *susie* is used in the experiment at the bit rate of 64 Kbps. The performances are plotted in Fig. 7-3 and Fig. 7-4. At the beginning of the transmission the channel model is executed for one measurement duration and the feedback is made available to the encoder in order to avoid hassles involved in the initial conditions.

From the results it can be seen that the window lengths of 7, 10 and 15 perform roughly same due to the saturation in the available bit budget at worst channel conditions. The window lengths of 3 and 5 lack bit budget at certain picture frames associated with heavy cell losses. The advantage of larger window length is that, if high bit budget is consumed at any frame, there need not be drastic reduction in bit budget for the succeeding picture frames in order to compensate for the current frame's high bit budget. This is due to the fact that the compensation is allowed to extend at least as far as the length of the adaptation window. On the other hand this can be disadvantageous if the high bit budget is requested often and the bit budget to be compensated grows beyond limit so as to keep the

Figure 7-3: Comparison of performance of different window lengths in dynamic rate control at 64 Kbps under narrowband channel conditions of 10^{-2} CLR.

Figure 7-4: Comparison of performance of different window lengths in dynamic rate control at 64 Kbps under broadband channel conditions of 5×10^{-3} BER.

history forever. Considering these factors it is decided to keep the adaptation window length to 7 GOF in the future experiments.

7.3.2.2 Rate control

In this section the rate control algorithm illustrated in Fig. 7-1 is evaluated. The monochrome video sequence *claire* is used in the experiment at the bit rates of 64 Kbps and 128Kbps. The scheme is evaluated for both the narrowband and broadband channel conditions. The performance of the rate controlled scheme, DRC-RZCO, is compared with that of the RZCO scheme whose parameters are optimized at each channel condition and bit rate.

Narrowband Channel

The cell loss rates of 10^{-2} and 10^{-3} are considered under the narrowband conditions. The performance curves for MRME configuration are plotted in Fig. 7-5 and Fig. 7-6, and those for 3D-WT configuration are shown in Fig. 7-7 and Fig. 7-8. It can be seen that the rate controlled performance at least matches the performance of the parameter optimized RZCO. When the channel offers less errors, it is seen that the DRC-RZCO provides better picture quality by providing more bit budget to the source coder. At the same time when the channel is severe, DRC-RZCO dynamically improves the channel protection and performs better than the RZCO.

Broadband Channel

The bit error rates of 5×10^{-3} and 10^{-3} are experimented in the broadband environment. The performances of the schemes under MRME configuration are illustrated in Fig. 7-9 and Fig. 7-10, and those for 3D-WT configuration are shown in Fig. 7-11 and Fig. 7-12. As in the narrowband channel, it is seen that the DRC-RZCO adapts to the channel and provides better quality in the low error condition and, improves channel protection and retains the robustness at high error conditions.

In both the narrowband and broadband conditions, occasionally the DRC-RZCO provides lower PSNR than the RZCO. It is observed that these are mainly due to the channel changing abruptly from high error to low error condition. In this situation the rate

(a) 10^{-2} CLR (b) 10^{-3} CLR

Figure 7-5: Performance comparison of the RZCO scheme with and without dynamic rate control for MRME configuration at 64 Kbps under narrowband channel conditions

(a) 10^{-2} CLR (b) 10^{-3} CLR

Figure 7-6: Performance comparison of the RZCO scheme with and without dynamic rate control for MRME configuration at 128 Kbps under narrowband channel conditions

(a) 10^{-2} CLR (b) 10^{-3} CLR

Figure 7-7: Performance comparison of the RZCO scheme with and without dynamic rate control for 3D-WT configuration at 64 Kbps under narrowband channel conditions

(a) 10^{-2} CLR (b) 10^{-3} CLR

Figure 7-8: Performance comparison of the RZCO scheme with and without dynamic rate control for 3D-WT configuration at 128 Kbps under narrowband channel conditions

(a) 5×10^{-3} BER (b) 10^{-3} BER

Figure 7-9: Performance comparison of the RZCO scheme with and without dynamic rate control for MRME configuration at 64 Kbps under broadband channel conditions

(a) 5×10^{-3} BER (b) 10^{-3} BER

Figure 7-10: Performance comparison of the RZCO scheme with and without dynamic rate control for MRME configuration at 128 Kbps under broadband channel conditions

(a) 5×10^{-3} BER (b) 10^{-3} BER

Figure 7-11: Performance comparison of the RZCO scheme with and without dynamic rate control for 3D-WT configuration at 64 Kbps under broadband channel conditions

(a) 5×10^{-3} BER (b) 10^{-3} BER

Figure 7-12: Performance comparison of the RZCO scheme with and without dynamic rate control for 3D-WT configuration at 128 Kbps under broadband channel conditions

controller takes some time to reduce the channel protection and increase the source coder rate. However this low PSNR is never drastically lower than the average PSNR indicating that the degradation is not severe.

The visual quality comparisons at 64 Kbps transmission are shown in Fig. 7-13 and Fig. 7-14. It is seen that the visual quality performance of the dynamic rate controlled RZCO is better than the plain RZCO scheme.

Performance at 16 Kbps and 32 Kbps

The performance of DRC-RZOC at lower bit rates of 16 Kbps and 32 Kbps are plotted in Fig. 7-15 through Fig. 7-18. The picture frame rate is lowered to 5 frames/sec so as to achieve a reasonable visual quality. The typical visual quality obtained is shown in Fig. 7-19 and Fig. 7-20. It can be seen that the rate control is effective and hence the system is able to withstand the channel variations.

Color Video Transmission

The results of color video transmission at 64 Kbps are plotted in Fig. 7-21 and Fig. 7-22. The color video sequence *carphone* is used in the experiment. The motion vectors computed for the luminance component are also used at the chroma components so as to reduce the overhead. The zero tree coding of the chroma components is provided with roughly 10% of the bit budget that is allocated for the luminance component. The bit budget and the channel protection for the chroma components are dynamically decided by the rate controller along with those for the other components of the video. However, due to the lack of bit budget, the chroma components are encoded using the basic zero tree coder without any partitioning.

The PSNR, shown in Fig. 7-21 and Fig. 7-22, is the average over Y, U and V components of the video sequence. It is observed that the variation in PSNR is increased due to the added loss of chroma components when the channel is severe.

7.4 Summary

This chapter discussed the adaptive rate control scheme to combat the severe channel variations. The RZCO scheme parameters and the channel protection were varied depending on the feedback information on the channel to achieve the robust video

(a) Original (frame 67 of *claire*)

(b) RZOC and DRC-RZOC at 10^{-2} CLR (narrowband)

(c) RZOC and DRC-RZOC at 5×10^{-3} BER (broadband)

Figure 7-13: Visual quality comparisons of RZOC and DRC-RZOC in MRME configuration at 64 Kbps under 10^{-2} CLR (narrowband) and 5×10^{-3} BER (broadband) channel conditions.

(a) Original (frame 108 of *claire*)

(b) RZOC and DRC-RZOC at 10^{-2} CLR (narrowband)

(c) RZOC and DRC-RZOC at 5×10^{-3} BER (broadband)

Figure 7-14: Visual quality comparisons of RZOC and DRC-RZOC in 3D-WT configuration at 64 Kbps under 10^{-2} CLR (narrowband) and 5×10^{-3} BER (broadband) channel conditions.

(a) 16 Kbps

(b) 32 Kbps

Figure 7-15: Performance of the DRC-RZCO scheme in MRME configuration at 16 Kbps and 32 Kbps under narrowband channel conditions

(a) 16 Kbps

(b) 32 Kbps

Figure 7-16: Performance of the DRC-RZCO scheme in 3D-WT configuration at 16 Kbps and 32 Kbps under narrowband channel conditions

(a) 16 Kbps

(b) 32 Kbps

Figure 7-17: Performance of the DRC-RZCO scheme in MRME configuration at 16 Kbps and 32 Kbps under broadband channel conditions

(a) 16 Kbps

(b) 32 Kbps

Figure 7-18: Performance of the DRC-RZCO scheme in 3D-WT configuration at 16 Kbps and 32 Kbps under broadband channel conditions

(a) Original (frame 28 of *claire*) (d) Original (frame 38 of *claire*)

(b) 16 Kbps (MRME)

(e) 16 Kbps (3D-WT)

(c) 32 Kbps (MRME)

(f) 32Kbps (3D-WT)

Figure 7-19: Visual quality performance of DRC-RZOC in MRME and 3D-WT configurations (left and right columns respectively) at 16 Kbps and 32 Kbps under narrowband channel condition of 10^{-2} CLR

Figure 7-20: Visual quality performance of DRC-RZOC in MRME and 3D-WT configurations (left and right columns respectively) at 16 Kbps and 32 Kbps under broadband channel condition of 5×10^{-3} BER

MRME configuration at 64 Kbps

3DWT configuration at 64 Kbps

Figure 7-21: Performance of DRC-RZCO for color video transmission at 64 Kbps under narrowband channel conditions.

MRME configuration at 64 Kbps

3DWT configuration at 64 Kbps

Figure 7-22: Performance of DRC-RZCO for color video transmission at 64 Kbps under broadband channel conditions.

communication. The channel measurement and classification, dynamic bit budget allocation and dynamic UEP scheduling were discussed in section 7.2. The algorithm for dynamic rate controlling, with flow chart illustration in Figure 7-1, was also discussed. In this scheme the bit budget is recomputed whenever the channel conditions degrade so as to make more bit budget available for higher channel protection. Additionally when the conditions demand the channel protection is improved at the cost of reducing the source rate in order to make the system robust against the channel errors. However, when the channel conditions improve the channel protection is decreased cautiously, one step at a time, to take care of short spurs of non-fading duration between adjacent fading durations.

The performance evaluation of adaptive rate control scheme was discussed in section 7.3. The results show that the adaptive rate control is very effective in improving the robustness of the system. Different sets of experiments were carried out in order to establish the reliability of the scheme under different video coder configurations, bit rates and channel conditions. Both the narrowband and broadband environments were taken in to consideration in the evaluation process. The very low bit rates of 16 Kbps and 32 Kbps were also experimented and shown that the achieved performance is acceptable. Finally, the color video transmission at 64 Kbps which includes the dynamic bit budget allocation and channel protection for chroma components was also discussed.

Chapter 8

Conclusions

8.1 Conclusions

This research is motivated by the worldwide quest to achieve the universal portable and personal communications, and the idea of wireless multimedia. An important issue in this task is to achieve a robust video communication over wireless packet networks. In this thesis, the robustness is realized by means of error resilient quantization and channel adaptation of the system. Adapting the system to channel conditions requires that the channel to be sensed at the receiver and the same to be used at the transmitter to dynamically alter the system parameters so as to combat the channel variations. At the transmitter this is achieved by classifying the compressed video data into smaller streams of different importance and dynamically changing the bit budget and UEP schedules as per the channel variations.

The combination of wavelet transform, MRME schemes and zero tree quantizer performs well in variable bit rate video coding and provides gracefully degrading video quality with a decrease in output bit rate. At low bit rates, the MRME1 scheme outperforms the rest of the MRME schemes by about 2 dB PSNR as shown in chapter 5. With the increase in the output bit rate the difference in the performance of these schemes is reduced, as the exact motion compensation gains significance.

Further in chapter 5, it is observed that the three dimensional wavelet transform perform better than the motion estimation based configuration of the video coder. However it is noted that the price is a considerable increase in delay at the encoder and decoder.

The basic zero tree quantizer proposed in [129] can not tolerate channel errors and hence can not be reliably used under noisy channel conditions. For this reason, new error resilient designs of the zero tree coder based on sequential and orthogonal partitioning were presented in chapter 6 (see section 6.2). The simulation results show that the schemes based on the sequential partitioning are not reliable due to the dependence among the partitions achieved and the large overhead. On the other hand, the scheme based on the orthogonal partitioning (RZCO) has a superiorly consistent performance and is shown to be consistent under various channel conditions. In section 6.3.2.3 the RZCO scheme is compared with the H.26P coder and is shown that the former performs better under noisy channel conditions.

For the system to be robust against changing channel scenario and provide good performance, it is required to vary the source rate and channel protection from time to time as per the variations in the channel conditions. This necessitated the channel sensing and measurement at the receiver and the feedback of the measured information. In chapter 7, experimental results using this information to alter the parameters of RZCO scheme are presented. The bit budget and the channel coder rate are varied to provide better protection when the channel is severe and, to improve the source rate and provide better picture quality when the channel offers less errors. The dynamic rate control scheme devised to achieve this is shown to be very effective in improving the performance compared to the non-rate-controlled RZCO scheme.

In the rest of chapter 7, the simulation results for low bit rates of 16 Kbps and 32 Kbps, and also the color video transmission at 64 Kbps are presented. The results at 16 Kbps and 32 Kbps further prove that the dynamic rate controlling is effective even when the available bit bandwidth is low. Similar conclusion can be drawn from the results of color video transmission carried out at 64 Kbps.

Finally, in Appendix D we presented the computational complexity analysis for the video coder. The purpose of this is to show that the video coder is realizable using present day micro-electronics technology unlike most of the third generation video coders like fractals and model based coders.

8.2 Suggestions for Future Research

- The main goal of this thesis was to arrive at a reliable video communication system for mobile channel. In order to focus our attention towards this, we decided to use the available channel models which happens to be simple in nature. There is a lack of good channel models for wireless packet environment. Hence our first suggestion is for **more research in modeling the mobile channel at the network level, especially for the packet mode of transmission**. Further, we like to suggest that the work in this thesis can be extended by simulating the channel to bit level involving modulation, diversity etc. and validate the results presented in this work. The channel simulation should take into account the user and terminal mobility from indoor to outdoor and vice versa so as to confirm the adaptability of the system to change in the type of channel experienced.
- The main assumption in chapter 7 is the channel being slow or stationary at least for the inner interleaver span period. In general, the channel could vary faster than this and hence there is a need to extend this work to such cases. Our suggestion in this regard is to **investigate the use techniques like ARQ along with the proposed DRC scheme to combat fast channel variations**. The ARQ scheme can be intelligent and multi-priority based so as to exploit the advantage of partitioning by providing high priority to the cells at the beginning of the partition. The ARQ scheme can also make use of the stream classification information to decide the priority with which the retransmission is allowed.
- Another important assumption in this work is the perfect feedback channel. **Further study is required to analyze the effects of noisy feedback channel on the overall system performance**.
- Another important issue that is not addressed in this work is the **real time implementation of the system**. This would require some implementation tricks in optimizing the schemes for speed. One needs to investigate the possible compromise between speed and quality of the implemented system over a given platform.
- **Adaptation of the proposed DRC scheme with other quantization schemes with their own stream classification strategies** would be an interesting work. Any VBR quantizer including the existing ones based on DCT can be researched

to improve the quality and make the schemes robust against the channel errors.

- Other interesting research directions would be;
 - to evaluate the system performance for multiple access schemes other than TDMA and CDMA.
 - efforts towards improving the system performance using better techniques within the video coder. Some examples are the use of longer tap wavelet filters, fractional pel accuracy motion estimation at higher bit rates, orthogonal partitioning on improved zero tree coders, etc.

References

- [1] E. Adelson and E. Simoncelli, "Orthogonal Pyramid Transform for Image Coding," *Proceedings of SPIE, Visual Communication and Image Processing*, Vol. 845, pp. 50-58, Oct 1987.
- [2] P. G. Andermo and G. Brismark, "CODIT, a Testbed Project Evaluating DS-CDMA for UMTS/FPLMTS," *Proceedings of IEEE Vehicular Technology Conference*, pp 21-25, 1994.
- [3] P. G. Andermo and L. M. Ewerbring, "A CDMA-Based Radio Access Design for UMTS," *IEEE Personal Communications*, pp. 48-53, Feb 1995.
- [4] P. G. Andermo, "Overview of CODIT Project," *Proceedings of RACE Mobile Telecommunications Summit*, pp. 33-42, 1995.
- [5] T. Andersson, "Tuning Macro Diversity Performance in a DS-CDMA System," *Proceedings of IEEE Vehicular Technology Conference*, pp 41-45, 1994.
- [6] J. S. Angiolillo, H. E. Blanchard and E. W. Israelski, "Video Telephony," *AT&T Technical Journal*, pp. 7-20, May-Jun 1993.
- [7] M. Antonini, M. Barlaud, P. Mathieu and I. Daubechies, "Image Coding Using Wavelet Transform," *IEEE Transactions on Image Processing*, Vol. 1, No. 2, pp. 205-220, Apr 1992.
- [8] R. Aravind, G. L. Cash, D. L. Duttweiler, H. M. Hang, B. G. Haskell and A. Puri, "Image and Video Coding Standards," *AT&T Technical Journal*, Vol. 72, pp. 67-89, Jan/Feb 1993.
- [9] F. Argenti and G. Benelli, "An ARQ Protocol for Mobile Radio systems," *Proceedings of ICC*, pp. 1323-1326, 1992.
- [10] F. Argenti, G. Benelli and A. Mecocci, "Source Coding and Transmission of HDTV Images Compressed With the Wavelet Transform," *IEEE Journal on Selected Areas of Communications*, Vol. 11, No. 1, pp. 46-58, Jan 1993.

- [11] H. Armbruster and G. Arndt, "Broadband Communication and its Realization with Broadband ISDN," *IEEE Communications Magazine*, Vol. 25, No. 11, pp. 8-19, Nov 1987.
- [12] A. Averbuch, D. Lazar and M. Israeli, "Image Compression Using Wavelet Transform and Multiresolution Decomposition," *IEEE Transactions on Image Processing*, Vol. 5, No. 1, pp. 4-15, Jan 1996.
- [13] J. J. Bae and T. Suda, "Survey of Traffic Control Schemes and Protocols in ATM Networks," *Proceedings of the IEEE*, Vol. 79, No. 2, pp. 170-189, Feb 1991.
- [14] A. Baier, U. C. Fiebig, W. Granzow, W. Koch, P. Teder and J. Thielecke, "Design Study for a CDMA-Based Third-Generation Mobile Radio System," *IEEE Journal on Selected Areas of Communications*, Vol. 12, No. 4, pp. 733-743, May 1994.
- [15] D. Ballantyne, "Digital Video Compression," *DSP Applications*, Vol. 2, pp. 26-35, May 1993.
- [16] M. Barlaud, T. Gaidon, M. Antonini and P. Mathieu, "Pyramidal Lattice Vector Quantization For Multiscale Image Coding," *IEEE Transactions on Image Processing*, Vol. 3, No. 4, pp. 367-381, Jul 1994.
- [17] E. Berruto, T. Brannlund, J. Gustafsson, D. Maiwald, W. Schott and P. Whitworth, "Radio Protocol Architecture of the CODIT UMTS System," *Proceedings of Zurich Seminar on Digital Communication*, pp. 417-427, 1994.
- [18] R. E. Blahut, *Theory and Practice of Error Control Codes*, Addison-Wesley Publishing Company, 1983.
- [19] J. C. Bolot and T. Turletti, "A Rate Control Mechanism for Packet Video in the Internet," *Proceedings of INFOCOM*, pp. 1216-1223, 1994.
- [20] J. Y. L. Boudec, "The Asynchronous Transfer Mode: A Tutorial," *Computer Networks and ISDN Systems*, No. 24, pp. 279-309, 1992.
- [21] G. Brismark, M. Ewerbring, T. Sandin and P. Teder, "A Coherent Detection Scheme for the Uplink Channel in a CDMA System," *Proceedings of IEEE Vehicular Technology Conference*, pp 729-732, 1994.
- [22] R. J. C. Bultitude, "Measurement, Characterization and Modelling of Indoor 800/900 MHz Radio Channels for Digital Communications," *IEEE Communications Magazine*, Vol. 25, No. 6, pp. 5-12, Jun 1987.

- [23] J. B. Cain, G. C. Clark, Jr. and J. M. Geist, "Punctured Convolutional Codes of Rate $(n-1)/n$ and Simplified Maximum Likelihood decoding," *IEEE Transactions on Information Theory*, Vol. 25, No. 1, pp. 97-100, Jan 1979.
- [24] J. R. Casas and L. Torres, "Coding of Details in Very Low Bit-Rate Video Systems," *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 4, No. 3, pp. 317-327, Jun 1994.
- [25] L. F. Chang and N. R. Sollenberger, "Comparison of Two Interleaving techniques for CDMA Radio Communications Systems," *Proceedings of IEEE Vehicular Technology Conference*, pp. 57-62, 1991.
- [26] Y. Chen and W. A. Pearlman, "Three-Dimensional Subband Coding of Video Using the Zero-Tree Method," *Proceedings of SPIE: Visual Communications and Image Processing*, Vol. 2727, pp. 1302-1312, Mar 1996.
- [27] S. Chong, S. Q. Li and J. Ghosh, "Predictive Dynamic Bandwidth Allocation for Efficient Transport of Real-time VBR Video Over ATM," *IEEE Journal on Selected Areas of Communications*, Vol. 13, No. 1, pp. 12-23, Jan 1995.
- [28] C. H. Chou and C. W. Chen, "A Perceptually Optimized 3-D Subband Codec for Video Communication Over Wireless Channels," *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 6, No. 2, pp. 143-156, Apr 1996.
- [29] C. K. Chui, *An Introduction to Wavelets*, Academic Press, Inc. 1992.
- [30] R. D. Cideciyan, E. Eleftheriou and M. Rupf, "Data Traffic Channels for the CODIT Testbed," *Proceedings of RACE Mobile Telecommunications Workshop*, pp. 231-235, 1994.
- [31] R. J. Clarke, *Transform Coding of Images*, London: Academic, 1985.
- [32] D. C. Cox, "Universal Digital Portable Radio Communications," *Proceedings of the IEEE*, Vol. 75, No. 4, pp. 436-477, Apr 1987.
- [33] R. M. R. Dagnino and A. L. Garcia, "Modelling VBR Video Sequences Containing Scene Transitions: A Case Study," *Technical Report*, Univ. of Toronto, May 1992.
- [34] R. M. R. Dagnino and A. L. Garcia, "Broadband Traffic Characterization," *Proceedings of Canadian Conference on Electrical and Computer Engineering (CCECE)*, Sep 1992.

- [35] I. Daubechies, "Orthonormal bases of compactly supported wavelets," *Communications on Pure and Applied Mathematics*, Vol. 41, pp. 909-996, Nov 1988.
- [36] M. C. Delfour and G. A. De Couvreur, "Interference-Free Assignment Grids-Part I: Basic Theory," *IEEE Transactions on Electromagnetic Compatibility*, Vol. 31, No. 3, pp. 280-292, Aug 1989.
- [37] M. C. Delfour and G. A. De Couvreur, "Interference-Free Assignment Grids-Part II: Uniform and Nonuniform Strategies," *IEEE Transactions on Electromagnetic Compatibility*, Vol. 31, No. 3, pp. 293-305, Aug 1989.
- [38] C. L. B. Despins, D. D. Falconer and S. A. Mahmoud, "Compound Strategies of Coding, Equalization, and Space Diversity for Wide-Band TDMA Indoor Wireless Channels," *IEEE Transactions on Vehicular Technology*, Vol. 41, No. 4, pp. 369-379, Nov 1992.
- [39] R. A. DeVore, B. Jawerth and B. J. Lucier, "Image Compression Through Wavelet Transform Coding," *IEEE Transactions on Information Theory*, Vol. 38, No. 2, pp. 719-746, Mar 1992.
- [40] K. Feher, ed., *Advanced Digital Communications: Systems and Signal Processing*, John Wiley, New York, 1974.
- [41] D. L. Gall, "MPEG: A Video Compression Standard for Multimedia Applications," *Communications of The ACM*, Vol. 34, No. 4, pp. 47-58, Apr 1991.
- [42] J. Gardiner and B. West, ed., *Personal Communication Systems and Technologies*, Artech House, Boston - London, 1995.
- [43] A. Gersho and R. M. Gray, *Vector Quantization and Signal Compression*, Boston, Mass., Kluwer Academic Publishers, 1992.
- [44] M. Ghanbari, "An Adopted H.261 Two-Layer Video Codec for ATM Networks," *IEEE Transactions on Communications*, Vol. 40, No. 9, pp. 1481-1490, Sep 1992.
- [45] H. Gharavi, "Subband Coding Algorithms for Video Applications: Videophone to HDTV-Conferencing," *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 1, No. 2, pp. 174-183, Jun 1991.
- [46] H. Gharavi and M. H. Partovi, "Video Coding and Distribution Over ATM for Multipoint Teleconferencing," *Proceedings of GLOBECOM*, pp. 1-7, 1993.

- [47] K. L. Gilhousen, I. M. Jacobs, R. Padovani, A. J. Viterbi, L. A. Weaver, Jr. and C. E. Wheatley III, "On the Capacity of Cellular CDMA Syatem," *IEEE Transactions on Vehicular Technology*, Vol. 40, No. 2, pp. 303-312, May 1991.
- [48] B. Girod, "Motion Compensating Prediction with Fractional-Pel Accuracy," *IEEE Transactions on Communications*, Vol. 41, No. 4, pp. 604-612, Apr 1993.
- [49] B. Girod, E. Steinbach and N. Farber, "Comparison of the H.263 and H.261 Video Compression Standards," *Standards and Common Interfaces for Video Information Systems*, Editor K. R. Rao, SPIE Optical Engineering Press, Vol CR60, pp. 233-251, 1995.
- [50] B. Glance and L. J. Greenstein, "Frequency-Selective Fading Effects in Digital Mobile Radio with Diversity Combining," *IEEE Transactions on Communications*, Vol. 31, No. 9, pp. 1085-1094, Sep 1983.
- [51] K. H. Goh, J. J. Soraghan and T. S. Durrani, "New 3-D Wavelet Transform Coding Algorithm for Image Sequences," *Electronics Letters*, Vol. 29, No. 4, pp. 401-402, 18th Feb 1993.
- [52] K. H. Goh, J. J. Soraghan and T. S. Durrani, "Wavelet Transform Based Adaptive Bit-plane Run-length Coding for Still Images," *Electronics Letters*, Vol. 30, No. 5, pp. 395-396, 3rd Mar 1994.
- [53] S. J. Golestani, "A Framing Strategy for Congestion Management," *IEEE Journal on Selected areas in Communications*, Vol. 9, No. 7, pp. 1064-1077, Sep 1991.
- [54] D. J. Goodman, "Trends in Cellular and Cordless Communications," *IEEE Communications Magazine*, pp. 31-40, Jun 1991.
- [55] A. Guha and D. J. Reininger, "Multichannel Joint Rate Control of VBR MPEG Encoded Video for DBS Applications," *IEEE Transactions on Consumer Electronics*, Vol. 40, No. 3, pp. 616-623, Aug 1994.
- [56] H.261 Software Codec, Portable Video Research Group, Stanford (Internet:ftp://havefun.stanford.edu/pub/p64/).
- [57] A. Habibi, "Future Trends in Image Coding," *Proceedings of SPIE; Applications of Digital Signal Processing XV*, Vol. 1771, pp. 406-412, 1992.
- [58] J. Hagenauer, "Rate Compatible Punctured Convolutional Codes (RCPC Codes) and Their Applications," *IEEE Transactions on Communications*, Vol. 36, No. 4, pp. 389-400, Apr 1988.

- [59] H. Hashemi, "Impulse Response Modeling of Indoor Radio Propagation Channels," *IEEE Journal on Selected Areas in Communications*, Vol. 11, No. 7, pp. 967-978, Sep 1993.
- [60] H. Hashemi, "The Indoor Radio Propagation Channel," *Proceedings of the IEEE*, Vol. 81, No. 7, pp. 943-968, Jul 1993.
- [61] G. C Hess, *Land-Mobile Radio System Engineering*, Artech House, Boston - London, 1993.
- [62] D. P. Heyman, A. Tatabai and T. V. Lakshman, "Statistical Analysis and Simulation Study of Video Teleconference Traffic in ATM Networks," *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 2, No. 1, pp. 49-59, Mar 1992.
- [63] D. Houlding and J. Vaisey, "Pyramid Decompositions and Hierarchical Motion Compensation," *Proceedings of SPIE - Conference on Digital Video Compression: Algorithms and Technologies*, Vol. 2419, , pp. 201--209, Feb. 7--10 1995.
- [64] Y. Huang and X. Zhuang, "An Adaptively Refined Block Matching Algorithm for Motion Compensated Video Coding," *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 5, No. 1, pp. 56-59, Feb 1995.
- [65] D. A. Huffman, "A Method for the Construction of Minimum-Redundancy Codes," *Proceedings of the I. R. E.*, pp 1098-1101, Sept. 1952.
- [66] K. Illgner and D. Lappe, "Mobile Multimedia Communications in a Universal Telecommunications Network," *Proceedings of SPIE Conference on Visual Communication and Image Processing*, 1995.
- [67] ISO/IEC 11172-2, "Coding of Moving Pictures and Associated Audio - for Digital Storage Media at up to about 1.5 Mbps," *Recommendation MPEG*, Nov 1992.
- [68] ISO/IEC 13818-2, "Generic Coding of Moving Pictures and Associated Audio," *Recommendation MPEG-2*, Nov 1994.
- [69] ITU, Study Group 15, "Video Codec for Audiovisual Services at p x 64 kbits/s," *Recommendation H.261*, Geneva 1990.
- [70] ITU, Study Group 15, "Video Coding for Narrow Telecommunication Channels at < 64 kbits/s," *Draft Recommendation H.263*, Jun 1995.
- [71] V. Iyengar, A. Jalali, P. Memelstein and J. Michaelides, "Comparison of ARQ Protocols for Asynchronous Data Transmission Over Rayleigh Fading Channels," *Proceedings of GLOBE-COM*, pp. 1703-1707, 1993.

- [72] A. Jacquin, "Fractal Image Coding: A Review," *Proceedings of the IEEE*, Vol. 81, No. 5, pp 796-818, Jun 1992.
- [73] A. K. Jain, *Fundamentals of Digital Image Processing*, Eaglewood Cliffs, N.J., Prentice Hall, 1989.
- [74] W. C. Jakes, Jr., ed., *Microwave Mobile Communications*, John Wiley, New York, 1974.
- [75] E. S. Jang and N. M. Nasrabadi, "Subband Coding With Multistage VQ for Wireless Image Communication," *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 5, No. 3, pp. 247-253, Jun 1995.
- [76] N. S. Jayant and P. Noll, *Digital Coding of Waveforms: Principles and Applications to Speech and Video*, Englewood Cliffs, N.J., Prentice-Hall, 1984.
- [77] N. S. Jayant, J. Johnston and R. Safranek, "Signal Compression Based on Models of Human Perception," *Proceedings of the IEEE*, Vol. 81, No. 10, pp. 1385-1422, Oct 1993.
- [78] S. H. Jung, S. K. Mitra and D. Mukherjee, "Subband DCT: Definition, Analysis, and Applications," *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 6, No. 3, pp. 273-286, Jun 1996.
- [79] G. Karlsson and M. Vetterli, "Packet Video and Its Integration in to the Network Architecture," *IEEE Journal on Selected Areas in Communications*, Vol. 7, No. 5, pp. 739-751, Jun 1989.
- [80] J. Katto, J. Ohki, S. Nogaki and M. Ohta, "A Wavelet Codec With Overlapped Motion Compensation for Very Low Bit-Rate Environment," *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 4, No. 3, pp. 328-338, Jun 1994.
- [81] M. Kawashima, C. T. Chen, F. C. Jeng and S. Singhal, "Adaptation of the MPEG Video-Coding Algorithm to Network applications," *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 3, No. 4, pp. 261-269, Aug 1993.
- [82] M. G. Kazantzakis, P.P. Demestichas and M. E. Anagnostou, "Optimum Frequency Reuse in Mobile Telephony Systems," *International Journal on Communication Systems*, Vol. 8, pp. 185-190, 1995.
- [83] M. Khansari, A. Jalali, E. Dubois and P. Mermelstein, "Robust Low Bit-Rate Video Transmission Over Wireless Access Systems," *Proceedings of International Conference on Communications*, pp. 571-575, 1994.

- [84] M. Khansari, A. Jalali, E. Dubois and P. Mermelstein, "Low Bit-Rate Video Transmission Over Fading Channels for Wireless Microcellular Systems," *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 6, No. 1, pp. 1-11, Feb 1996.
- [85] W. C. Y. Lee, *Mobile Communication Engineering*, McGraw-Hill, New York, 1982.
- [86] X. Lee and A. L. Garcia, "A Fast Feature Algorithm of Multi-Resolution Motion Estimation," *Proceedings of GLOBE-COM*, pp. 320-324, 1992.
- [87] H. Lee, "Joint Wavelet-based Coding and Packetization for Video Transport over Packet-Switched Networks," *Proceedings of SPIE; Visual Communications and Image Processing*, Vol. 2727, pp. 591-602, Mar 1996.
- [88] J. M. Lervik and T. A. Ramstad, "Robust Image Communication Using Subband Coding and Multilevel Modulation," *Proceedings of SPIE; Visual Communications and Image Processing*, Vol. 2727, pp. 524-535, Mar 1996.
- [89] A. S. Lewis and G. Knowles, "Video Compression Using 3D Wavelet Transforms," *Electronics Letters*, Vol. 26, No. 6, pp. 396-398, 15th Mar 1990.
- [90] A. S. Lewis and G. Knowles, "Image Compression Using 2-D Wavelet Transforms," *IEEE Transactions on Image Processing*, Vol. 1, No. 2, pp. 244-250, Apr 1992.
- [91] H. Li, A. Lundmark and R. Forchheimer, "Image Sequence Coding at Very Low Bitrates: A Review," *IEEE Transactions on Image Processing*, Vol. 3, No. 5, pp. 589-609, Sep 1994.
- [92] J. O. Limb, "Distortion Criteria of the Human Viewer," *IEEE Transactions on Systems, Man and Cybernetics*, pp 778-793, Dec 1979.
- [93] S. Lin and D. J. Costello, *Error Control Coding; Fundamentals and Applications*, Englewood Cliffs, N.J., Prentice-Hall, 1983.
- [94] S. Lin, D. J. Costello, Jr. and M. J. Miller, "Automatic-Repeat-Request Error-Control Schemes," *IEEE Communications Magazine*, Vol. 22, No. 12, pp. 5-17, Dec 1984.
- [95] F. Ling and D. D. Falconer, "Combined Orthogonal/Convolutional Coding for a Digital Cellular CDMA System," *Proceedings of IEEE Vehicular Technology Conference*, pp. 63-66, 1991.
- [96] J. P. Linnartz, *Narrowband Land-Mobile Radio Networks*, Artech House, 1993.

- [97] Y. J. Liu and Y. Q. Zhang, "Wavelet-Coded Image Transmission Over Land Mobile Radio Channels," *Proceedings of GLOBE-COM*, pp. 235-239, 1992.
- [98] B. Liu and A. Zaccarin, "New Fast Algorithms for the Estimation of Block Motion Vectors," *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 3, pp. 148-157, Apr. 1993.
- [99] S. G. Mallat, "A Theory for Multiresolution Signal Decomposition: The Wavelet Representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 11, No. 7, pp. 674-693, Jul 1989.
- [100] S. G. Mallat, "Multifrequency Channel Decompositions of Images and Wavelet Models," *IEEE Transactions on Acoustics, Speech and Signal Processing*, Vol. 37, No. 12, pp. 2091-2110, Dec 1989.
- [101] N. Movahhedinia, "Bandwidth Management in Indoor Broadband Wireless Networks (IBWN)," *Ph.D. Thesis Dissertation*, Carleton University, OCIEE-96-01, Jan 1996.
- [102] H. G. Musmann and J. Klie, "TV Transmission Using a 64 kbit/s Transmission Rate," *Proceedings of International Conference on Communication*, pp. 23.3.1-23.3.5, 1979.
- [103] T. Naveen, F. Bosveld, J. W. Woods and R. L. Lagendjik, "Rate Constrained Multiresolution Transmission of Video," *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 5, No. 3, pp. 193-206, Jun 1995.
- [104] A. N. Netravali and B. G. Haskell, *Digital Pictures; Representation, Compression, and Standards*, Second Edition, Plenum Press, New York and London, 1995.
- [105] K. N. Ngan and W. L. Chooi, "Very Low Bit Rate Video Coding Using 3D Subband Approach," *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 4, No. 3, pp. 309-316, Jun 1994.
- [106] Nikolaidis and R. O. Onvural, "A Bibliography on Performance Issues in ATM Networks," *Computer Communication Review, ACM SIGCOMM*, Vol. 22, No. 5, pp. 8-23, Oct 1992.
- [107] M. Nomura, T. Fujii and N. Ohta, "Basic Characteristics of Variable Rate Video Coding in ATM Environment," *IEEE Journal on Selected Areas in Communications*, Vol. 7, No. 5, pp. 752-760, Jun 1989.

- [108] J. E. Padgett, C. G. Gunther and T. Hattori, "Overview of Wireless Personal Communications," *IEEE Communications Magazine*, pp. 28-41, Jan 1995.
- [109] S. Panchanathan, E. Chan and X. Wang, "Fast Multiresolution Motion Estimation Scheme for a Wavelet Transform Video Coder," *Proceedings of Visual Communications and Image Processing '94*, SPIE publications, Vol. 2308, pp. 671-681, 1994.
- [110] R. Pandya, "Emerging Mobile and Personal Communication Systems," *IEEE Communications Magazine*, pp. 44-52, Jun 1995.
- [111] J. D. Parsons and J. G. Gardiner, *Mobile Communication Systems*, Blackie, London and Halstead Press, New York, 1989.
- [112] R. M. Pelz, "Transmission of Error Protected Video Over DECT," *RACE Mobile Telecommunications Workshop*, Amsterdam, Holland, Vol. 2, pp. 526-529, May 1994.
- [113] R. M. Pelz, "An Unequal Error Protected px8 kbit/s Video Transmission for DECT," *Proceedings of IEEE Vehicular Technology Conference*, pp. 1020-1024, 1994.
- [114] M. P. Pickering and J. F. Arnold, "A Perceptually Efficient VBR Rate Control Algorithm," *IEEE Transactions on Image Processing*, Vol. 3, No. 5, pp. 527-532, Sep 1994.
- [115] W. K. Pratt, *Digital Image Processing*, Second Edition. New York, Wiley, 1991.
- [116] J. G. Proakis, *Digital Communications*, Second Edition, McGraw-Hill, Inc. 1989.
- [117] M. D. Prycker and A. D. Gelman, "Video on Demand," *IEEE Communications Magazine*, pp. 67-80, May 1994.
- [118] G. R. Rajugopal, S. A. Mahmoud and R. H. M. Hafez, "Multiresolution Motion Estimation at Low Bit Rates," Accepted to be presented at TENCON'96, Perth, Australia, Nov 1996.
- [119] T. S. Rappaport, S. Y. Seidel and K. Takamizawa, "Statistical Channel Impulse Response Models for Factory and Open Plan Building Radio Communication System Design," *IEEE Transactions on Communications*, Vol. 39, No. 5, pp. 794-807, May 1991.
- [120] A. R. Reibman and A. W. Berger, "On VBR Video Teleconferencing Over ATM Networks," *Proceedings of GLOBE-COM*, pp. 314-319, 1992.

- [121] D. Reininger, D. Raychaudhuri, B. Melamed, B. Sengupta and J. Hill, "Statistical Multiplexing of VBR MPEG Compressed Video on ATM networks," *Proceedings of INFOCOM*, pp. 919-926, 1993.
- [122] O. Rioul and M. Vetterli, "Wavelets and Signal Processing," *IEEE Signal Processing Magazine*, pp. 14-38, Oct 1991.
- [123] O. Rioul and P. Duhamel, "Fast Algorithms For Discrete and Continuous Wavelet Transforms," *IEEE Transactions on Information Theory*, Vol. 38, No. 2, pp. 569-586, Mar 1992.
- [124] K. Rob and P. Fernando, "MPEG-4: A New Way Towards Multimedia Mobile Communications," *Proceedings of RACE Mobile Telecommunications Summit*, pp. 528-534, 1995.
- [125] J. W. Roberts, "Variable-Bit-Rate Traffic Control in B-ISDN," *IEEE Communications Magazine*, pp. 50-56, Sep 1991.
- [126] J. A. Saghri, P. S. Cheatham and A. Habibi, "Image Quality Measure Based on a Human Visual Model," *Journal of Optical Engineering*, pp 813-818, July 1989.
- [127] A. Said and W. A. Pearlman, "A New, Fast, and Efficient Image Codec Based on Set Partitioning in Hierarchical Trees," *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 6, No. 3, pp. 243-250, Jun 1996.
- [128] A. A. M. Saleh and L. J. Cimini, Jr., "Indoor Radio Communications Using Time-Division Multiple Access with Cyclical Slow Frequency Hopping and Coding," *IEEE Journal on Selected Areas in Communications*, Vol. 7, No. 1, Jan 1989.
- [129] J. M. Shapiro, "Embedded Image Coding Using Zerotrees of Wavelet Coefficients," *IEEE Transactions on Signal Processing*, Vol. 41, No. 12, pp. 3445-3462, Dec 1992.
- [130] E. A. B. D. Silva and M. Ghangari, "On the Performance of Linear Phase Wavelet Transforms in Low Bit-Rate Image Coding," *IEEE Transactions on Image Processing*, Vol. 5, No. 5, pp. 689-706, May 1996.
- [131] E. P. Simoncelli, W. T. Freeman, E. H. Adelson and D. J. Heeger, "Shiftable Multiscale Transforms," *IEEE Transactions on Information Theory*, Vol. 38, No. 2, pp. 587-607, Mar 1992.
- [132] M. J. T. Smith, S. L. Eddins, "Analysis/Synthesis Techniques for Subband Image Coding," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, Vol. 38, No. 8, pp. 1446-1456, Aug 1990.

- [133] P. F. M. Smulders and A. G. Waggmans, "Wide-band Measurements of MM-Wave Indoor Radio Channels," *Third International IEEE Symposium on PIMRC, Boston*, pp. 10.5.1-10.5.5, Oct 1992.
- [134] P. Sriram and M. W. Marcellin, "Image Coding Using Wavelet Transforms and Entropy-Constrained Trellis-Coded Quantization," *IEEE Transactions on Image Processing*, Vol. 4, No. 6, pp. 725-728, Jun 1995.
- [135] R. Stedman, H. Gharavi, L. Hanzo and R. Steele, "Transmission of Subband-Coded Images via Mobile Channels," *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 3, No. 1, pp. 15-26, Feb 1993.
- [136] R. Steele, *Mobile Radio Communications*, Pentech Press, London 1992.
- [137] T. Stefansson and M. Allkins, "Real-Time Testbed for Assessing a CDMA-Based System," *IEEE Personal Communications*, pp. 75-80, Oct 1995.
- [138] S. Stein, "Fading Channel Issues in System Engineering," *IEEE Journal on Selected Areas in Communications*, Vol. 5, No. 2, pp. 68-89, Feb 1987.
- [139] R. Talluri, "MPEG-4 Status and Directions," *Standards and Common Interfaces for Video Information Systems*, Editor K. R. Rao, SPIE Optical Engineering Press, Vol CR60, pp. 252-262, 1995.
- [140] A. M. D. Turkmani and A. F. Toledo, "Radio Transmission at 1800 MHz in to and Within Multistorey Buildings," *IEE Proceedings*, Part I, Vol. 138, No. 6, pp. 577-584, 1991.
- [141] W. Verbiest and L. Pinnoo, "A Variable Bit Rate Video Codec for Asynchronous Transfer Mode Networks," *IEEE Journal on Selected Areas in Communications*, Vol. 7, No. 5, pp. 761-770, Jun 1989.
- [142] A. J. Viterbi, "Convolutional Codes and Their Performance in Communication Systems," *IEEE Transactions on Communication Technology*, Vol. 19, pp. 751-772, Oct 1971.
- [143] G. K. Wallace, "The JPEG Still Picture Compression Standard," *Communications of the ACM*, Vol. 34, No. 4, pp. 31-45, Apr 1991.
- [144] L. Wang, "Bit Rate Control for Hybrid DPCM/DCT Video Codec," *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 4, No. 5, pp. 509-517, Oct 1994.

- [145] X. Wang, Y. Liu and S. D. Blostein, "Video Image Transmission via Mobile Satellite Channels," *Proceedings of International Conference on Communications*, pp. 352-356, 1995.
- [146] I. H. Witten, R. M. Neal and J. G. Cleary, "Arithmetic Coding for Data Compression," *Communications of the ACM*, Vol. 30, No. 6, pp. 520-540, Jun 1987.
- [147] J. Woods and T. Naveen, "A Filter Based Bit Allocation Scheme for Subband Compression of HDTV," *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 1, No. 3, pp. 436-440, July 1992.
- [148] Z. D. Wu, "Analysis of Rate Control Mechanisms for Transport Layer Protocols in High speed Networks," *Proceedings of International Conference on Communications*, pp. 1319-1325, May 1994.
- [149] Y. Yasuda, K. Kashiki and Y. Hirata, "High-Rate Punctured Convolutional Codes for Soft Decision Viterbi Decoding," *IEEE Transactions on Communications*, Vol. 32, No. 3, pp. 315-319, Mar 1984.
- [150] R. K. Young, *Wavelet Theory And Its Applications*, Kluwer Academic Publishers, 1993.
- [151] S. Zafar, Y. Q. Zhang and B. Jabbari, "Multiscale Video Representation Using Multiresolution Motion Compensation and Wavelet Decomposition," *IEEE Journal on Selected Areas in Communications*, Vol. 11, No. 1, pp. 24-35, Jan 1993.
- [152] Y. Q. Zhang, W. W. Wu, K. S. Kim, R. L. Pickholtz and J. Ramasastry, "Variable Bit-rate video Transmission in the Broadband ISDN Environment," *Proceedings of the IEEE*, Vol. 79, No. 2, pp. 214-221, Feb 1991.
- [153] Y. Q. Zhang and S. Zafar, "Motion-Compensated Wavelet Transform Coding for Color Video Compression," *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 2, No. 3, pp. 285-296, Sep 1992.
- [154] Q. F. Zhu, Y. Wang and L. Shaw, "Coding and Cell-Loss Recovery in DCT-based Packet Video," *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 3, No. 4, pp. 248-258, Aug 1993.

Appendix - A

Arithmetic Entropy Coding

A.1 Principles of Arithmetic Coding

In arithmetic coding [146], a symbol is represented as an interval of real numbers between 0 and 1. The interval size is chosen to be proportional to the probability of occurrence of the corresponding symbol. As the number of symbols in the set (or size of dictionary) increases, the interval size reduces requiring higher precision representation and hence needs more bits for its representation.

An entire message consisting of several symbols can also be represented as an interval in the same region between 0 and 1. As the number of symbols in the message increase, the interval size will reduce requiring more bits for its representation. Successive symbols in the message reduce the interval size in accordance with the symbol probabilities generated. The more likely symbols reduce the range to a lesser extent than those which occur rarely and hence add fewer bits to the message.

There are two kinds of probability models suggested for the arithmetic coder. They are fixed probability model and adaptive probability model. In the case of a fixed probability model, the symbol probabilities are prefixed and so will be the width of the intervals allotted to them. Whereas in the case of an adaptive probability model, the symbol probability is computed dynamically using histogram counters. Hence the interval allotted to each symbol keeps changing each time a symbol is encoded or decoded. The fixed model is chosen to describe the idea of arithmetic coding as it is easy to understand.

Before anything is transmitted, the range for the message is the entire interval $[0, 1)$ (i.e., 0

$\leq x < 1$). As each symbol is processed, the range is narrowed to the portion which is allocated to the message. For example, suppose the symbol set is $\{a, e, i, o, u, !\}$ with the probabilities as shown in Table A-1. The arithmetic encoding of the message *eaui!* using these probabilities is illustrated in Fig. A-1. Initially both the encoder and the decoder will start from the range $[0, 1)$. After getting the first symbol, *e*, the encoder narrows the range to $[0.2, 0.5)$ and allocates the intervals to individual symbols in proportionate to the initial allocation. The second symbol, *a*, will narrow this range to its first one fifth and splits this new range $[0.2, 0.26)$ to allocate the intervals for all the symbols in the set. This process of splitting the range and then narrowing of the same by an incoming symbol continues till the end of the message. After the symbol *!*, the range will be $[0.2336, 0.23354)$. The decoder needs to know any one point in this range and end of message to decode the entire message.

A.2 Implementation

In the description given above, the arithmetic encoder needs to wait till all the symbols of the message arrive before beginning any transmission. In practice, an incremental transmission and reception scheme is preferred. Also in practice, a fixed point (or integer) arithmetic coder is preferred over an infinite precision arithmetic coder. To achieve this the lower and upper ends of range are defined to be some N bit integers, which initially will be 0 and $2^N - 1$ respectively. As and when the symbols arrive the precision of these bounds will be increased by scaling them up. The scaling is done by shifting the ZEROS into the lower order bits of lower bound and shifting ONES into those of higher bound. As the code range narrows, the top bits of the bounds become same. These bits can be transmitted immediately as they are not going to be refined anymore. This results in an incremental transmission with fixed point arithmetic implementation. The care should be taken to check underflow or overflow as they affect any fixed point arithmetic implementations.

Symbol	Probability	Range
a	0.2	[0, 0.2)
e	0.3	[0.2, 0.5)
i	0.1	[0.5, 0.6)
o	0.2	[0.6, 0.8)
u	0.1	[0.8, 0.9)
!	0.1	[0.9, 1.0)

Table A-1: Fixed model example for alphabet set {a, e, i, o, u, !}

Figure A-1: Arithmetic encoding process for the message eaii!.

Appendix - B

Rate Compatible Punctured Convolutional Codes

It is well known that the convolutional coders are the best bet for variable length channel coding. Building a separate coder for each of the rates required in a multi-rate coding environment is a tedious and costly solution. Also it is known that the Viterbi decoding, which needs 2^{n-1} -ary comparisons, becomes highly complicated as the coding rate is increased. Studies [23], [58], [149] have shown that, using rate compatible punctured convolution coders (RCPC) it is possible to provide an efficient control over the output redundancy rate and also provide a cost effective implementation with only slight performance degradation. The rate punctured convolutional coders are comprised of a basic 1/2 rate or 1/3 rate convolutional coder with the desired rate of output puncturing before transmission. Hence these coders reduce the computations to binary comparisons and result in $(2^{n-1}-1)/(n-1)$ less binary comparisons per decoded bit. Using these coders it is possible to change the puncture rate of each of the bit stream encoded. However, the puncture rate used at the encoder needs to be known at the decoder for proper decoding. At the receiver the punctured bits, punctured at encoder, are stuffed as erasures into the stream and decoded as if the stream were an half rate coded.

B.1 Punctured Codes Derived from Half Rate Codes

The procedure of obtaining $(n-1)/n$ rate convolution codes from half rate convolution code

[23], [58], [149] is shown in Fig. B-1. The original half rate coded stream consists of sequentially arranged blocks of 2 bits, one from each of the generators. Specific m bits among every l such blocks ($2l$ bits) of original coded sequence are periodically deleted. The deletion will be according to the puncture map, common to both encoder and decoder, which indicates the deleting bit positions. When l is chosen as $(n-1)$ with m as $(n-2)$, an output code rate of $(n-1)/n$ is achieved. At the receiver, these deleted bits are inserted as erasures (according to the puncture map) and decoded treating the sequence as an half rate coded one. In this case the Viterbi's upper bound [142] of the bit error probability (P_B) is given by,

$$P_B \leq \frac{1}{l} \sum_{k=d}^{\infty} C_k P_k$$

Where d is the minimum free distance of the punctured code, C_k is the total number of error bits produced by all incorrect paths with distance k which diverge from the correct path at each of l blocks of the original code, and P_k is the probability that one such incorrect path is selected in the Viterbi decoding process. The above upper bound is shown to be very tight for large SNR [142] and hence is suitable for theoretical evaluation of the BER performance of these codes.

The distance property of the punctured codes largely differ depending on the deleted bit positions. Thus the puncture map should be decided in order to obtain the best BER performance under given conditions. The best puncture maps for n varying from 2 to 14 for different constraint lengths are given in [149].

B.2 Selection of Constraint Length

The factors governing the selection of constraint length are; the error correcting capability, the length of encode/decode shift register and the length of buffered symbol sequence used in current symbol decision. Note that the last two factors in turn decide the length of the history involved with the coder. In this work we have chosen a constraint of length of 6 to cater for bit error rates as low as 5×10^{-3} with low history (less than 75 symbols) required in coding short streams like motion vectors.

Figure B-1: Procedure to obtain $(n-1)/n$ rate convolution code from half rate code

Appendix - C

ATM Protocol

The Asynchronous Transfer Mode (ATM) [20], [106] is widely accepted as the mode of operation for future communication systems, particularly the broadband ISDN systems. At the ATM layer of the protocol, the information is carried in fixed size packets called *cells*. ATM application layer (AAL) format of ATM cells contain seven parts as shown in Fig. C-1. The header contains the routing information which is also called as *label*. The second part of the cell is 8 bit wide serial number of the cell. B/E is a 2 bit begin or end tag and is followed by 6 unused reserved bits. The effective payload is 44 bytes long followed by 6 bits of length indicator (LI) which is used to encode the size of the useful payload in the cell leaving out any padded bits. A 10 bit cyclic redundancy code (CRC) is added at the end to enable some error detection at the receiver. At the ATM layer, only the 5 byte header will be used for switching and the rest of the cell is treated as information field.

In an ATM cell, the header is short and of fixed size. Hence the label is not an explicit address, as an explicit address is usually lengthy and of variable size. Whenever a switching unit reads an incoming cell with routing label, it consults the *switch routing table* to find out the next logical address and/or the port to which this cell is to be sent. The label contains a virtual path identifier (VPI) and virtual channel identifier(VCI) as the present logical address. The complete description of the header is given in Fig. C-2 and the usage of label as the logical address is shown in Fig. C-3. The header also contains a HEC (Header Error Correction) byte. The HEC byte is initially used for packet synchronization and afterwards it is used to correct a single error in the header.

In contrast to the existing packet networks where some kind of window flow control is

Figure C-1: ATM cell structure: AAL-2 format

Figure C-2: Header of ATM cell at (a) User-Network interface (UNI) and (b) Network Node interface (NNI) and (c) Lookup Table at the switching unit

Figure C-3: Virtual path and virtual channel connections

used, the ATM uses *rate control* scheme between the user and the network access. Each traffic connection is associated with a kind of traffic contract defining the mean bit rate, peak bit rate and the duration for which this peak bit rate is sustained. The network continuously monitors all the connections for possible contract violations. If a user violates the contract by sending more packets, the extra packets will be either discarded or marked as low priority depending on the network congestion status. The cell loss priority (CLP) bit in the header is used for this purpose. In the case of overload at the network node, those packets with CLP bit set to low priority will be discarded. There are many source policing and congestion control schemes suggested in the literature [13], [125], [53] to reduce the risk of losing data packets.

Appendix - D

Computational Complexity of Video coder

This Appendix gives the computational complexity analysis of the individual components of video coder, i.e. wavelet transform, multiresolution motion estimation and zero tree quantization. The basic aim of this appendix is to give a rough estimation of the video coder complexity. It is intended to show that this coder, though more complex than the DCT schemes, is realizable unlike the fractals and the model based coders where the complexity is huge and can not be achievable using the present day DSP technology. While considering the above calculations it should be noted that the memory fetch and other program overhead are not taken into account. For video applications these overhead are known to be higher than the algorithm complexity itself. It should also be noted that this analysis does not take in to account the parallelism of the algorithm or the special instruction level capabilities available in DSP architectures.

D.1 Wavelet Transform

We will start from finding the computational complexity of 1D wavelet transform of a N point signal for L stages of decompositions, and then extend it to 2D transform of a $N_x \times N_y$ picture frame. Note that the wavelet filter is realized as a finite impulse response filter (FIR) of length K . Then the computation for one output sample will be K multiplications and K additions both in floating point (FP) domain. Taking each pair of a FP multiplication and FP addition as a floating point operation (FLOP) we need K floating point operations

per each output sample of the filter. Also, at each stage of decomposition both low pass and high pass filtering of the same input signal are needed. Thus the complexity of an L stage wavelet transform on an N point signal would be,

$$\begin{aligned}
 C_{WT,N} &\approx 2\left(N + \frac{N}{2} + \frac{N}{4} + \dots L \text{ terms}\right)K \text{ FLOPs} \\
 C_{WT,N} &\approx 2NK\left(1 + \frac{1}{2} + \dots + \frac{1}{2^{L-1}}\right) \text{ FLOPs} \\
 C_{WT,N} &\approx 4NK(1 - 2^{-L}) \text{ FLOPs} \tag{D.1}
 \end{aligned}$$

Extending this to a 2D wavelet transform of N_x rows and N_y columns frame, if L is the number of stages, then we need 4^L number of individual (low pass or high pass) filtering to be performed. But at each stage the input length to these filters will reduce by a factor of two. Thus we get,

$$\begin{aligned}
 C_{WT,N_x N_y} &\approx K\left(2N_x N_y + \frac{2N_x N_y}{4} + \frac{2N_x N_y}{16} + \dots L \text{ terms}\right) \\
 C_{WT,N_x N_y} &\approx 2N_x N_y K\left(1 + \frac{1}{4} + \frac{1}{16} + \dots + \frac{1}{4^{L-1}}\right) \\
 C_{WT,N_x N_y} &\approx \frac{8N_x N_y K}{3}(1 - 4^{-L}) \tag{D.2}
 \end{aligned}$$

Thus for a QCIF frame of 176x144 size, using (D.2), it takes 0.27MFLOPs for a 4 stage 2D wavelet transform with a 4 tap wavelet filter. If a third dimension transform has to be performed on a bunch of 8 frames using a 2-tap filter, an extra 1.7MFLOPs would be required. The equation (D.2) is valid for the reconstruction process also.

D.2 Multiresolution Motion Estimation

The computational complexity [115], [104], [151] of any block based motion estimation scheme is mainly governed by the search window size and searching method employed. For a search window of size of $(n \times n)$, the search complexity of a full search scheme is $O(n^2)$ per pixel. Hence for a frame size of $(N_x \times N_y)$, the computational complexity will be

$O(N_x N_y n^2)$. There are many suboptimal search techniques suggested in the literature to reduce this computational burden at the cost of reduced performance.

Here we first present the computational complexity calculations for MRME1 algorithm (see section 2.4.2) with full-pixel accuracy and then extend it to other algorithms and sub-pixel accuracies. The computational burden of MRME1 algorithm is much lesser than that for full frame motion vector schemes as we need to compute the motion vectors for only the LL band. Let the motion block size is $B_x B_y$, the search window size is $S_x S_y$ and frame size is $N_x N_y$. Each of the block wise comparison operation in a search would involve MSE (or MAD) calculation and finding minimum MSE location. Hence this involves $B_x B_y$ subtractions, $B_x B_y$ square (multiplication) operations and $B_x B_y$ additions for each search vector location. In total, each search vector check needs $2B_x B_y$ FLOPs and hence for each motion vector decision the complexity is,

$$C_{MV} \approx 2B_x B_y S_x S_y \text{ FLOPs} \quad (D.3)$$

The computational complexity for a frequency band of size $F_x F_y$ would be,

$$C_{ME1} = C_{MV} \left(\frac{F_x F_y}{S_x S_y} \right) = 2B_x B_y S_x S_y \left(\frac{F_x F_y}{B_x B_y} \right) \text{ FLOPs}$$

$$C_{ME1}(S_x, S_y, F_x, F_y) = 2F_x F_y S_x S_y \text{ FLOPs} \quad (D.4)$$

For MRME1 algorithm we need to find motion vectors only for the LL band. Thus for a frame size of 176x144, LL band size of 22x18 (after 3 stage wavelet decomposition), motion block size of 2x2 and search window size of 16x16, the computational complexity for motion estimation in the LL band is 0.16 MFLOPs. In MRME1 we also need to find the residual bands for non-LL bands, which needs one subtraction per pixel, meaning an extra 0.02 MFLOPs. For MRME2 algorithm the motion search complexity is four times that of MRME1. thus in the above example, MRME2 needs 0.64MFLOPs.

In MRME3 through MRME5, we need to take care of increasing motion block size at each resolution. Keeping the search window size same, for a L stage decomposition we get,

$$\begin{aligned}
C_{ME2} &= C_{ME1}(S_x, S_y, F_x, F_y) + 3 \sum_{r=1}^L C_{ME1}(S_x, S_y, 2^{r-1}F_x, 2^{r-1}F_y) \\
C_{ME2} &= C_{ME1}(S_x, S_y, F_x, F_y) \left(1 + 3 \sum_{r=0}^{L-1} 2^{2r} \right) \\
C_{ME2} &= C_{ME1}(S_x, S_y, F_x, F_y) \left(1 + 3 \left(\frac{1-4^L}{1-4} \right) \right) \\
C_{ME2} &\approx 4^L C_{ME1}(S_x, S_y, F_x, F_y) \\
C_{ME2} &= 2^{2L+1} F_x F_y S_x S_y \tag{D.5}
\end{aligned}$$

Thus the MRME3 through MRME5 needs 64 times more computational power than MRME1, requiring 10.25MFLOPs in the above example. The motion compensation in full pixel accuracy needs only additions and thus requires about 0.026 MFLOPs. For the sub-pixel accuracies there will be additional burden of interpolation by the same factor as accuracy.

D.3 Zero Tree Quantization

The zero tree coding requires significance check and zero tree symbol formation in the dominant pass and one bit SAQ in the subordinate pass of each iteration. This means that the ZTQ needs only additions, subtractions and comparisons and does not involve any multiplications and vector search etc. Thus the computational requirement of this scheme is much smaller when compared to that of VQ. In finding the upper bound for calculations we make use of the fact that each symbol at n th resolution gets compared at most n times during significance map encoding. Thus for a frame size of $N_x \times N_y$ with L level decomposition, the computational complexity at each iteration of ZTQ will be,

$$C_{ZTQ} \leq 3N_x N_y + \frac{N_x N_y}{2^{2L}} + 3 \left(\frac{N_x N_y}{2^{2L}} \right) \sum_{r=1}^L 2^{2(r-1)} r + 3N_x N_y \tag{D.6}$$

The first term in (D.6) represents one significance check and two comparisons for sign

determination at each coefficient during dominant pass. The second and third terms are for zero tree symbol comparisons, of LL and non-LL bands respectively, during encoding of the significance map. The last term is for two comparisons and one addition/subtraction during one bit quantization at subordinate pass. Simplifying (D.6) we get,

$$C_{ZTQ} \leq N_x N_y \left(6 + \frac{1}{2^{2L}} \left(1 + 3 \sum_{r=1}^L 2^{2(r-1)r} \right) \right) \quad (D.7)$$

Thus, for ZTQ on a three level decomposition of 176x144 size frame the maximum computational requirement is 0.22 MFLOPs.