DIGITAL VIDEO TRANSMISSION OVER WIRELESS NETWORKS

by

Saviour Zammit B.Elec.Eng. (Hons), M.Sc.

Submitted for the degree of

DOCTOR OF PHILOSOPHY

at

The University of Aston in Birmingham

December, 1995.

DIGITAL VIDEO TRANSMISSION OVER WIRELESS NETWORKS

by

Saviour Zammit B.Elec.Eng. (Hons), M.Sc.

Submitted for the degree of
DOCTOR OF PHILOSOPHY
at the University of Aston in Birmingham
1995.

SUMMARY

The advent of the Integrated Services Digital Network (ISDN) led to the standardisation of the first video codecs for interpersonal video communications, followed closely by the development of standards for the compression, storage and distribution of digital video in the PC environment, mainly targeted at CD-ROM storage. At the same time the second-generation digital wireless networks, and the third-generation networks being developed, have enough bandwidth to support digital video services. The radio propagation medium is a difficult environment in which to deploy low bit error rate, real time services such as video. The video coding standards designed for ISDN and storage applications, were targeted at low bit error rate levels, orders of magnitude lower than the typical bit error rates experienced on wireless networks.

This thesis is concerned with the transmission of digital, compressed video over wireless networks. It investigates the behaviour of motion compensated, hybrid interframe DPCM/DCT video coding algorithms, which form the basis of current coding algorithms, in the presence of high bit error rates commonly found on digital wireless networks. A group of video codecs, based on the ITU-T H.261 standard, are developed which are robust to the burst errors experienced on radio channels.

The radio link is simulated at low level, to generate typical error files that closely model real world situations, in a Rayleigh fading environment perturbed by co-channel interference, and on frequency selective channels which introduce inter symbol interference. Typical anti-multipath techniques, such as antenna diversity, are deployed to mitigate the effects of the channel. Link layer error control techniques are also investigated.

Using the robust codecs developed in this thesis, it is possible to transmit good quality video on burst-error radio channels at average bit error rates in the region of 10^{-3}. A comparison of interleaved forward error correction codes and retransmission techniques has been undertaken. It is shown that higher bandwidth efficiency can be obtained by using error detection and retransmission in an adaptive error control configuration. Good results are also obtained using layered codecs which use unequal error protection and joint source-channel coding techniques to deliver digital video on a variety of multipath fading channels, in a robust and efficient manner.

KEY WORDS

WIRELESS TRANSPORT, LAYERED CODING, ERROR CONTROL, REAL-TIME ARQ

This thesis is dedicated to my wife Patricia,

my children Joseph and Aidan

and my Parents

Agnes and Joseph,

for all their love, support and sacrifices.

One picture is worth more than a thousand words.

Anonymous

Chinese Proverb

ACKNOWLEDGEMENTS

First of all, I wish to thank my supervisor Dr. Geoff Carpenter, for his unrelenting support, guidance and encouragement during the research period and the thesis write-up.

I would also like to thank my wife Patricia and my in-laws Joseph and Lelina Rizzo, for their constant support, and for looking after my family whilst I was away.

I would also like to thank the University of Malta for granting me study leave and for sponsoring my studies at Aston University, and my colleagues who gave me their support when it was most needed.

Finally, I would like to acknowledge the help of the Committee of Vice-Chancellors and Principals of the Universities of the UK for part-sponsoring my tuition fees at Aston through an ORS award.

Contents

Title 1

Abstract 2

Dedication 3

Acknowledgements 4

Contents 5

List of Figures 14

List of Tables 17

1. Introduction 18

1.1 The emergence of digital video services 18

1.2 The emergence of Wireless Networks 19

1.3 Wireless video services 20

1.4 Digital Video transmission over Wireless Networks 20

1.5 Aims 21

1.6 Method 21

1.6.1 Literature Reviews 21

1.6.2 Modelling and Simulation 21

1.6.3 Video Codec 23

1.6.4 Performance Assessment 23

1.7 Thesis Outline 24

2. Background 26

2.1 Introduction 26

2.2 Digital Video Compression 26

2.2.1 Intraframe coding techniques 27

2.2.1.1 Predictive Coding 28

2.2.1.2 Transform Coding 28

2.2.2 Interframe Image Sequence Coding Techniques 30

2.2.2.1 Interframe DPCM 30

2.2.2.2 Conditional Replenishment 31

2.2.2.3 Motion Compensation 31

2.2.2.4 Foreground/Background Prediction 32

2.2.2.5 Frame Interpolation Coding 32

2.2.2.6 Hybrid Interframe Coding 33
2.2.2.7 Other Image Sequence Coding Techniques 34
2.2.3 Objective and Subjective Fidelity Measures 34
2.3 Wireless Networks 35
2.3.1 The Physical Medium 35
2.3.2 The Physical Layer 36
2.3.2.1 Wireless Network Topologies 36
2.3.2.2 The Cellular Concept 37
2.3.2.3 Digital Wireless Networks 37
2.3.2.4 Spectrally Efficient Modulation 38
2.3.2.5 Radio Propagation 38
2.3.2.6 Bit Error Mechanisms on Multipath fading channels 39
2.3.2.7 Countermeasures 41
2.3.2.8 System Limitations 41
2.3.3 Medium Access Control 42
2.3.4 Data Link Layer - Error Control 43
2.3.5 Wireless Network Services 43
2.3.6 Current Wireless Network Offerings 43
2.3.7 Evolution of Wireless Networks 46
2.3.7.1 Personal Communications Networks and Systems 46
2.3.7.2 UMTS 46
2.3.7.3 FPLMTS 47
2.4 Error Control Techniques 47
2.4.1 Error Control Fundamentals 47
2.4.2 The phases of Error Control 48
2.4.3 Forward Error Correction using Block Codes 48
2.4.4 ARQ 49
2.4.5 Hybrid ARQ/FEC 49
3. State of the Art Literature Review 50
3.1 Introduction 50
3.2 Video Coding Standards 50
3.2.1 H.120 50
3.2.2 H.261 50
3.2.3 MPEG-1 51
3.2.4 MPEG-2 51
3.2.5 H.263 and MPEG-4 52

3.2.6 Video codecs for real-time interactive mobile services. 52
3.3 Wireless networks in support of Video Transmission 53
3.3.1 Second-Generation Digital Mobile Networks 53
3.3.2 Second-Generation Digital Cordless Networks 55
3.3.3 Wireless Local Area Networks 56
3.3.4 Integrated Services Wireless Transport Systems 56
3.3.5 Third-Generation Wireless Networks 57
3.4 Error Control For Video Transmission 57
3.4.1 Intraframe Error Control 57
3.4.1.1 Error control for DPCM coded images 57
3.4.1.2 Error control for DCT-coded images 58
3.4.1.3 The effect of channel errors on Variable length Codes 59
3.4.1.4 Joint Source-Channel techniques 59
3.4.2 Interframe Error Control 60
3.4.2.1 Conditional replenishment and channel errors 60
3.4.2.2 The effect of channel errors on Interframe DPCM 60
3.4.2.3 Motion Compensated Prediction and Channel Errors 60
3.4.2.4 Interframe Error control Techniques 60
3.4.2.5 Error control and H.261 61
3.5 Packet Video 62
3.5.1 Error control techniques for packet video 62
3.5.2 Cell Loss Compensation: Non-Layered Techniques 63
3.5.3 Cell Loss Compensation: Layered Coding 66
3.5.4 Subjective Testing 68
3.6 Developments in Digital Terrestrial TV 68
3.7 Video Transmission Over Wireless Networks 69
3.7.1 Unequal Error Protection 69
3.7.2 Error Control by Retransmission 71
3.7.3 Intramode codecs 72
3.7.4 Other Techniques 73
3.8 Conclusions 73
4. A Hybrid Interframe DPCM\DCT Video Codec 75
4.1 Introduction 75
4.2 CCITT Visual Telephone System Overview 76
4.3 H.261 Video Codec Overview 77
4.4 Video Source Format 78

4.5 Video Frame Subdivisions 79
4.6 The Interactive H.261 Codec 80
4.7 The IHC coder 80
4.7.1 The Source Coder 80
4.7.1.1 Intramode Coding 80
4.7.1.2 Intermode Coding 81
4.7.1.3 Motion Estimation 82
4.7.1.4 The Discrete Cosine Transform 83
4.7.1.5 Quantization 85
4.7.1.6 Macroblock Coding Mode Selection 86
4.7.2 Video Multiplex Coder 87
4.7.3 Rate Control 89
4.7.3.1 Rate Control Strategies 90
4.7.3.1.1 Intracoded Frame Algorithm 92
4.7.3.1.2 Intercoded Frame Algorithm 92
4.7.4 Transmission Coder 93
4.8 The IHC decoder 94
4.8.1 Transmission decoder 94
4.8.2 Source Decoder 95
4.9 IHC functionality 95
4.9.1 The IHC Coder 96
4.9.1.1 Coder Parameters 96
4.9.1.2 Variable Bit Rate Coding 96
4.9.1.3 Fixed Bit Rate Coding 97
4.9.1.4 Layered Coding 97
4.9.1.5 Video frame/sequence manipulation 97
4.9.2 The IHC Decoder 97
4.10 Validation tests 98
4.10.1 Test Sequences 98
4.10.2 Results 98
4.11 Conclusions 105

5. Coded Video Transmission on High Bit Error Rate Channels 107

5.1 Introduction 107
5.2 Error resilience of hybrid Interframe DPCM/DCT codecs 108
5.2.1 Error resilience of the source coding algorithm 108
5.2.2 Error resilience of the video multiplex coder 110

5.2.3 H.261 error control mechanisms 111
5.3 The design of robust H.261 codecs 112
5.4 Implementation of a robust H.261 decoder 113
5.4.1 Error Detection and Correction 113
5.4.2 Error Containment and Concealment 114
5.4.3 Error Recovery 114
5.5 Testing the Robust Decoder 114
5.5.1 Performance measures 116
5.5.2 Results 117
5.5.2.1 QM48 stream 118
5.5.2.1.1 Subjective Assessment 118
5.5.2.1.2 Objective Results 120
5.5.2.1.3 Discussion 120
5.5.2.2 MA96 stream 121
5.5.2.2.1 Subjective Assessment 121
5.5.2.2.2 Objective Assessment 123
5.5.2.2.3 Discussion 123
5.5.2.3 SM384 stream 124
5.5.2.3.1 Subjective Assessment 124
5.5.2.3.2 Objective Assessment 124
5.5.2.3.3 Discussion 126
5.5.2.4 Discussion of results 126
5.6 Improvements 127
5.6.1 Protecting the Picture and GOB headers 127
5.6.2 Fixed field length DC intraframes 128
5.7 Error recovery using periodic frame Intracoding 130
5.8 Error recovery using forced updating. 131
5.9 Error recovery using the Fast Update Request facility. 135
5.10 Transcoding 135
5.10.1 Transcoding for video transmission. 136
5.10.2 Transcoding delay 136
5.10.3 Transcoding techniques 137
5.10.4 FEC 137
5.10.5 Structured packing 138
5.10.6 QCIF re-synchronisation 139
5.10.7 Fixed sized Macroblocks. 139

5.11 Summary of Conclusions 139

6. Modelling and Simulating the Digital Radio System 143

6.1 Introduction 143
6.2 The Radio Propagation channel 144
6.2.1 Fading Multipath Channel Models 145
6.3 Modulation 147
6.4 The digital radio system model 148
6.4.1 Transmitter Model 148
6.4.2 The Channel Model 149
6.4.3 The Receiver Model 151
6.5 System Simulation 151
6.5.1 Simulating the transmitter 151
6.5.1.1 Pseudo random number generation 152
6.5.1.2 Digital Encoding 152
6.5.1.3 Pulse Shaping filters 153
6.5.2 The Receiver 153
6.5.2.1 Carrier Recovery 153
6.5.2.2 Symbol Timing Recovery 154
6.5.2.3 Receiver filters 154
6.5.2.4 Statistics collection 154
6.5.3 Simulating radio channels 154
6.5.3.1 Simulating an AWGN channel 154
6.5.3.2 Simulating Rayleigh fading channels 155
6.5.3.2.1 Validation Results 156
6.5.3.3 Rayleigh Fading Channel with Co-Channel Interference 159
6.5.3.4 Simulating wideband channels 159
6.5.3.5 Two-ray channel model 160
6.6 Simulation Results 160
6.6.1 Simulation results for the AWGN channel 160
6.6.2 Rayleigh fading channel 161
6.6.3 Simulation results for the Rayleigh fading channel with CCI 162
6.6.4 Frequency-selective channel 165
6.7 Conclusions 165

7. Video Transmission Over a Fading Radio Link 167

7.1 Introduction 167

7.2 System Description 168
7.2.1 Wireless Network Topology. 169
7.2.2 Network Architecture 169
7.2.3 The Physical Layer 170
7.2.4 The Multiple Access Control Sub-layer 170
7.2.5 The Error Control Layer 171
7.2.6 Transcoder 172
7.3 Digital Radio Modem Performance in Narrowband, Rayleigh Fading channels 172
7.4 Video Transmission over Rayleigh fading channels 176
7.5 Unprotected H.261 streams on the Rayleigh fading Channel 177
7.6 Decoding H.261 streams using FEC and error detection 178
7.7 Decoding H.261 with Forced Updating and fixed length DC frames 178
7.8 The effect of increasing the mobile speed 182
7.9 The effect of increasing the modulation level 185
7.10 Interleaved FEC 186
7.11 Summary of Conclusions 187

8. Error Control by Retransmission for Video Transmission over Wireless Networks 191

8.1 Introduction 191
8.2 Basic ARQ techniques 192
8.2.1 Stop-And-Wait ARQ 193
8.2.2 Go-Back-N ARQ 194
8.2.3 Selective Retransmission ARQ 194
8.2.4 Hybrid ARQ 194
8.3 ARQ techniques for data transmission over fading radio channels. 195
8.4 Feedback techniques for video transmission. 196
8.4.1 ARQ for Video Transmission over Mobile Radio links 197
8.4.2 Multiple Access Control considerations 198
8.4.3 Forward and reverse channels 198
8.5 The BTRL ARQ System 198
8.6 Retransmission Buffer Control 199
8.6.1 Flow Control Techniques 200
8.6.1.1 Flow Control by discarding Frames 201
8.6.1.2 Flow Control by discarding Coefficients 202
8.6.2 Bandwidth-on-Demand Techniques 204
8.6.2.1 Variable Rate MAC 204

8.6.2.2 Variable Rate Modulation 205
8.6.2.3 Shared Retransmission Channel 205
8.6.2.4 Up link packet pre-emption 206
8.6.2.5 Down link packet pre-emption 206
8.7 SAW ARQ for video transmission 206
8.7.1 System Description 206
8.7.2 The SAW protocol 207
8.7.3 Flow Control 208
8.7.4 Simulation 208
8.7.5 Results - Flow control on a Rayleigh fading channel 209
8.7.6 Results - Flow control on a Rayleigh fading channel with CCI 211
8.8 SR ARQ for video transmission 212
8.8.1 Selective Retransmission ARQ System - Model and Description 213
8.8.2 SR ARQ Algorithm 213
8.8.3 System and Channel Model 214
8.8.4 Results 214
8.9 The Trace-Back Decoder 216
8.10 Summary of conclusions 216

9. Layered Coding 220

9.1 Introduction 220
9.2 Layered Coding 221
9.3 Adapting Layered Coding Techniques For Wireless Networks. 224
9.4 A Two-Layer Video Transport System For Wireless Networks 225
9.4.1 System Description. 226
9.4.2 Two-Layer Transcoder 227
9.4.3 Two-Layer Decoder. 228
9.4.4 Analysis 228
9.5 Layered Video Transport Over A Third-Generation Wireless Access Network. 229
9.5.1 A Third-Generation Wireless Access Network 229
9.5.2 A Three-Layer Video Codec 230
9.5.3 Error Control Sub-System 232
9.5.4 MAC Requirements 233
9.5.5 Third generation MAC 233
9.5.6 Simulation Results 233
9.6 Conclusions 235

10. Conclusions and Further Work 238

10.1 Introduction 238

10.2 Simulation Framework 239

10.2.1 Video source coding and transcoding 240

10.2.2 Real-time video display 240

10.2.3 Digital radio modem simulation 240

10.2.4 Radio channel modelling 240

10.2.5 Countermeasures 241

10.2.6 Channel coding 241

10.2.7 Statistical analysis 241

10.2.8 Burst error generation 241

10.2.9 Subjective and objective video quality measures 241

10.3 Hybrid Interframe DPCM / DCT in high BER 242

10.4 Hybrid Interframe DPCM / DCT on Rayleigh fading channels 245

10.5 Error Control by Retransmission for Video Transmission 247

10.6 Layered Coding 250

10.7 Further Work 251

Publications 254

References 255

Appendix One - Report on Current Wireless Telecommunications Networks 268

Appendix Two - IHC Functionality 276

Appendix Three - Error Trapping Conditions 281

Appendix Four - Abbreviations and Acronyms 283

List of Figures

Figure 1-1 Generic Digital Communication System Model 22
Figure 1-2 OSI Model of a Wireless Access Network 22
Figure 1-3 Simplified Simulation System 23

Figure 2-1 The interaction of the Research Fields 26
Figure 2-2 Image Source Coding Operations 27
Figure 2-3 Block diagram of a predictive codec 28
Figure 2-4 Hybrid Intraframe/DPCM coder and decoder 33
Figure 2-5 Hybrid DPCM/Intraframe coder and decoder 34
Figure 2-6 Channel Model 47

Figure 3-1 Leaky DPCM predictor 58
Figure 3-2 ATM based video transmission system 63
Figure 3-3 Two-layer video coding for ATM networks 66

Figure 4-1 Visual Telephone System 77
Figure 4-2 Video codec block diagram 78
Figure 4-3 CIF and QCIF Frame Decomposition 79
Figure 4-4 Source Coder 81
Figure 4-5 Block Matching Motion Estimation 83
Figure 4-6 AAN DCT Algorithm 84
Figure 4-7 Macroblock Type Decision Tree 86
Figure 4-8 Video Multiplex stream hierarchy 88
Figure 4-9 Block Zig Zag Scan 89
Figure 4-10 Rate Control Strategy 91
Figure 4-11 Source Decoder 95

Figure 5-1 Error Control Simulation System 115
Figure 5-2 Error Simulation Menu 115
Figure 5-3 Error statistics for the colour QM48 sequence 119
Figure 5-4 Error statistics for the colour MA96 sequence 122
Figure 5-5 Error statistics for the SM384 sequence 125
Figure 5-6 Fixed length DC intramode frame 129

Figure 5-7 Error statistics for the QM48 sequence with Forced Updating....132
Figure 5-8 Error statistics for the MA96 sequence with Forced Updating....133
Figure 5-9 Error statistics for the SM384 sequence with Forced Updating....134
Figure 5-10 Structured Packing....138

Figure 6-1 Generic QAM Communications System Model....149
Figure 6-2 Rayleigh fading Channel Model....150
Figure 6-3 Rayleigh fading channel with CCI....150
Figure 6-4 Two-ray frequency-selective model....151
Figure 6-5 Rayleigh Fading Envelope....156
Figure 6-6 Simulated Rayleigh Fading Envelope Statistics....158
Figure 6-7 RF Spectrum of Rayleigh Fading Envelope....159
Figure 6-8 Tapped delay line Wideband channel model....159
Figure 6-9 Symbol Error Rate versus Eb/ No for the Static AWGN Channel....163
Figure 6-10 Symbol Error Rate versus Average Eb/No in Rayleigh Fading....163
Figure 6-11 Symbol Error Rate versus average SIR in Rayleigh fading....164
Figure 6-12 Bit Error versus rms delay spread in the two-ray frequency-selective channel....164

Figure 7-1 Video Transmission over a Wireless Access Network - System model....168
Figure 7-2 Wireless Access Network Topology....169
Figure 7-3 Wireless Communications Network Architecture....170
Figure 7-4 MAC Frame and Packet Structure....170
Figure 7-5 Error Detection Packet Format....171
Figure 7-6 FEC Packet Format....171
Figure 7-7 MDPSK BER on the Rayleigh Fading Channel....173
Figure 7-8 BER per block at 4 km/hr....173
Figure 7-9 BER per Block at 64 km/hr....174
Figure 7-10 Consecutive Block Errors....174
Figure 7-11 Errored Block Gap Distribution at 64 km/hr....175
Figure 7-12 Errored Block Gap Distribution at 4 km/hr....175
Figure 7-13 Forced Updating - PSNR Penalty....176
Figure 7-14 QM32 Sequence in Rayleigh Fading - QDPSK at 4 km/hr....179
Figure 7-15 QM32 Sequence in Rayleigh Fading - QDPSK at 64 km/hr....181
Figure 7-16 QM32 Sequence in Rayleigh Fading - 16-DPSK at 4 km/hr....183
Figure 7-17 QM32 Sequence in Rayleigh Fading - 16-DPSK at 64 km/hr....184
Figure 7-18 Interleave Buffer....187

Figure 8-1 Block diagram of a basic ARQ error control system 193
Figure 8-2 Generic Hybrid ARQ/FEC Technique 195
Figure 8-3 DECT MAC TDD frame structure 197
Figure 8-4 The BTRL ARQ system for H.261 transmission over DECT 199
Figure 8-5 A transcoder based ARQ system 200
Figure 8-6 Simple Buffer System 200
Figure 8-7 MSE with Frame Discard 202
Figure 8-8 MSE with Coefficient Discard 203
Figure 8-9 MSE with Continuous Coefficient Discarding 204
Figure 8-10 Retransmission Bandwidth Allocation 205
Figure 8-11 Flow control / SAW ARQ system model 209
Figure 8-12 Buffer Occupancy in SAW ARQ with flow control - Rayleigh Fading channel 210
Figure 8-13 PSNR results for SAW ARQ with Flow Control in Rayleigh fading 211
Figure 8-14 Buffer Occupancy with SAW ARQ and Flow Control - CCI channel 211
Figure 8-15 PSNR results for SAW ARQ with flow control on a CCI channel 212
Figure 8-16 SR ARQ System model 213
Figure 8-17 Channel Model for the SR ARQ system 213
Figure 8-18 Probability of Packet loss P_e for SR ARQ (with diversity) 215

Figure 9-1 Two Independent Layer Codec System 222
Figure 9-2 Generic Two-Layer Video Coding System 223
Figure 9-3 An H.261 Based Two-Layer Video Coder 223
Figure 9-4. A Hybrid DPCM/DCT Two-Layer Video Coder 224
Figure 9-5 Down-Link Model For Two-Layer Video Transport 226
Figure 9-6 Generating The Guaranteed And Enhancement Macroblocks 227
Figure 9-7 Third-Generation Wireless Access Network 230
Figure 9-8 A Three-Layer Video Codec 231
Figure 9-9 Block Diagram Of The Hybrid ARQ-FEC Error Control Scheme 232
Figure 9-10 Receive Buffer Delay in 0.1 Delay Spread 234
Figure 9-11 Layered Coder Bit Rate Division 235
Figure 9-12 Layered Coder - PSNR for Base and Enhancement Layers 235

Figure A.1 Wireless Network Taxonomy 268

List of Tables

Table 2-1 A Taxonomy of Current Wireless Networks....44
Table 3-1 Second Generation Digital Cellular Systems....54
Table 3-2 Second Generation Cordless Systems....55

Table 5-1 Five point Subjective Usefulness Scale....117
Table 5-2 QM48 Stream Results....118
Table 5-3 MA96 stream results....121
Table 5-4 SM384 stream results....124

Table 6-1 Gray coded dibits in π/4 DQPSK....152
Table 6-2 Theoretical and Simulated mean and variance for the Rayleigh fading Process....156
Table 6-3 Ideal and measured mean an variance of the Gaussian Components....157

Table 7-1 Block Error Burst Means....174
Table 7-2 Error Stream Summary....176
Table 7-3 Summary of results - QM32_11 corrupted by Stream A....177
Table 7-4 Summary of Results - QM32_11 corrupted by Stream B....182
Table 7-5 Summary of Results - QM32_11 corrupted by Stream C....185
Table 7-6 Summary of Results - QM32_11 corrupted by Stream D....185
Table 7-7 Summary of Results - QM32_23 sequence....186

Chapter 1
Introduction

This thesis is concerned with the transmission of digitized video over wireless networks. The impetus to study digital video transport arises from the emergence of video services as an important, wide area network application. Furthermore, digital wireless networks are being developed, standardized and installed which can support data rates above 32 kbit/s. Thus it seems natural to expect that wide area, video services will also be accessed over wireless networks. Hitherto, wireless access research has concentrated on access to data and voice services. The aim of this thesis is to extend the research to wireless-accessed video services.

Current video coding techniques generate high data rates, require very low bit error rates and are delay sensitive. On the other hand, wireless networks operate in a very difficult radio propagation environment and it is difficult to meet the objectives of low bit error rate and low delay simultaneously, efficiently and economically, over the scarce radio spectrum resource.

This thesis reports on the results obtained whilst investigating video transport over wide area networks accessed, at one or both ends, via wireless local area networks. In this thesis the term *Wireless Access Networks* describes such hybrid wide area networks, comprising wireless local area networks interconnected by fixed, wide area networks.

1.1 The emergence of digital video services

Real time, interactive video communications, in the form of videophony and videoconferencing, is expected to take off in this decade. Traditionally, video transmission services to the general public have been on a broadcast-only basis, mainly because of the lack of adequate throughput in the local loop. The Integrated Services Digital Network (ISDN) is set to change all this with digital access at the user's premises, at transmission rates which allow acceptable videophone and videoconference applications. The CCITT has ratified a video codec standard, recommendation H.261 [CCITT-H.261], for operation over narrowband ISDN at n x 64 kbit/s (where n = 1 .. 30).

The development of Broadband ISDN (B-ISDN) using the (optical) Fiber To The Home (FTTH) concept, will accelerate the penetration of video services and applications. B-ISDN will be cell based and can support both fixed and variable rate traffic. This has led to the current interest in packet video coding and transport.

Interactive video communications is expected to penetrate the LAN environment as well. ISDN compatible Local Area Network (LAN) terminals have already been manufactured. Real-time internet video and audio communications have already been standardized [RFC1889], though a degree

of compatibility with ISDN will have to be maintained. The same is true for multimedia communications over Metropolitan Area Networks (MAN). Eventually, interactive video services will be extended to the cordless and mobile users as well.

The launch of broadcast digital video services is imminent. These include digital terrestrial and satellite broadcast TV and High-Definition TV (HDTV), as well as digital video transmission over cable.

Video forms an integral part of the interactive multimedia information concept which is set to revolutionize communications and computing. Multimedia is the synergistic aggregate of video, audio, text and computer data. Until now these services have been available individually, but the possibility of combining the four together opens up a number of new applications such as Distance Learning, Tele-Working and Tele-Shopping.

Video services can be classified as Interactive Services comprising Conversational, Messaging and Retrieval services or Distributive Services comprising Broadcast and Video On Demand [Stallings90]. Armbruster and Arndt [Armbruster90] identify the following video services for commercial and private applications:-

Person-to-Person video communication:-

- Video telephony (Conversation, domestic scenes, objects, graphic material, procedures)
- Broadband Message Handling (video mail)
- Videoconference

Access to Video Information

- Broadband videotext
- Video on Demand

Broadcast of programs and Data

- Common TV
- Pay TV
- HDTV
- CableText

1.2 The emergence of Wireless Networks

Mobile and cordless communications systems have enjoyed unparalleled growth in the last decade. These systems are attractive because they allow user mobility, easy system setup and reconfiguration, and reduce installation and modification costs.

First-Generation analogue cellular, mobile and cordless systems have been designed for voice services; support for data services was added later, but there is not enough bandwidth to support video services. At least one Second-Generation system, the Digital European Cordless Telecommunications

(DECT) system, has been designed to support voice and data services. Recently, [MacDonald92] has reported on initial trials to transmit H.261 coded video over a DE(Stedman *et al.* [Stedman92] have also developed a joint source/channel coding scheme motion compensated, sub-band codec, which can be used for low bit rate video services o wireless system.

Third-Generation wireless networks being developed will support integrated servic video [Goodman91a]. Recent wireless LAN standardization efforts by the IEEE P802.11 cc also considering support for isochronous services (e.g. voice and video) [Black92]. Performance European Radio LAN (HIPERLAN) standard being developed by th Telecommunication Standards Institute (ETSI) will support terminal rates of up to 20 Mbit/ WARC 92 set aside spectrum for the FPLMTS (Future Public Land Mobile Telecom System) being developed by the International Telecommunication Union (ITU), which is provide a wide range of global services, including bearer services equivalent to ISDN B& [Gardiner90]. It is thus apparent that the wireless communications systems being dev support data rates that are compatible with good quality video services.

1.3 Wireless video services

Wireless video makes novel applications possible. Mobile Audio Visual Termin; [RACE2072] allow videophone and videoconference services to be accessed on the mov(facilitating interactive access to multimedia databases. Video transmitters worn by pol: services and disaster relief services will make it possible to gather video intelligence qu operational logistics. Mobile video transceivers, in conjunction with Virtual Reality (VF enhance VR applications in medicine, remote maintenance and entertainment. M communications could be used to pilot vehicles remotely and can serve as safety and bac for autonomous vehicles on factory floors and grounds.

1.4 Digital Video transmission over Wireless Networks

Voice, data and video have different requirements in a telecommunication syste voice can tolerate Bit Error Rates (BER) up to 10^{-3} and packet cell loss rates up to about service is sensitive to delay. Data services must have very low BER and can tolerate no (they can tolerate higher transmission delays. Both Constant Bit Rate (CBR) and Varia (VBR) video coding techniques make stringent demands on the transmission sys transmission rates can be substantially higher than voice. Current video codec standards low BER similar to data services [Yamamoto89] and exhibit the same level of delay s voice, since interactive video services will be accompanied by voice.

1.5 Aims

The main aims of the research were:-

- To identify current trends in video coding .
- To identify existing video coding standards and those in development.
- To familiarize with the techniques used in current and proposed coding standarc
- To identify current trends in mobile communications.
- To identify current and proposed wireless access standards capable of suppo services.
- To study the transport characteristics of identified video capable wireless netwo
- To develop a simulation and modeling framework to assist in the study of d transport over Wireless Networks (WN).
- To analyze the resilience of the selected video standards to high bit error rates.
- To study the resilience of unprotected coded video to wireless network impairm
- To study error control techniques to mitigate the effects of the radio channel i on the coded video.

1.6 Method

Two main tools were used in the pursuance of the above goals: literature reviews an and simulation.

1.6.1 Literature Reviews

Literature reviews were conducted in three main areas: Video coding, Wireless Ne Error Control. Initially an extensive literature survey was conducted to identify current trer coding, existing video coding standards and those in development.

A literature review was undertaken to identify the state-of-the-art in wireless netw identify future trends. This led to the identification of wireless networks, both current an capable of supporting video services. The channel characteristics for these networks an techniques were then studied.

The third extensive literature research concerned error control techniques. Three m areas were identified: general error control techniques, general error control techniques channels and previous work in error control techniques for digital video transport ov networks.

1.6.2 Modelling and Simulation

All the relevant results in this thesis were obtained by simulation. Two telecom

systems models served as a simulation framework: The generic telecommunication sys [Sklar83] shown in Figure 1-1 and the International Standards Organization's (ISO) Ope Interconnection (OSI), seven layer model [Tanenbaum89].

The OSI seven layer model was used to conceptualize and model the end-to-er access networks. The layering helped to focus the attention on relevant detail but also prove the classification and development of error control techniques. A model of a wireless acce comprising two wireless networks interconnected over ISDN is shown in Figure 1-2

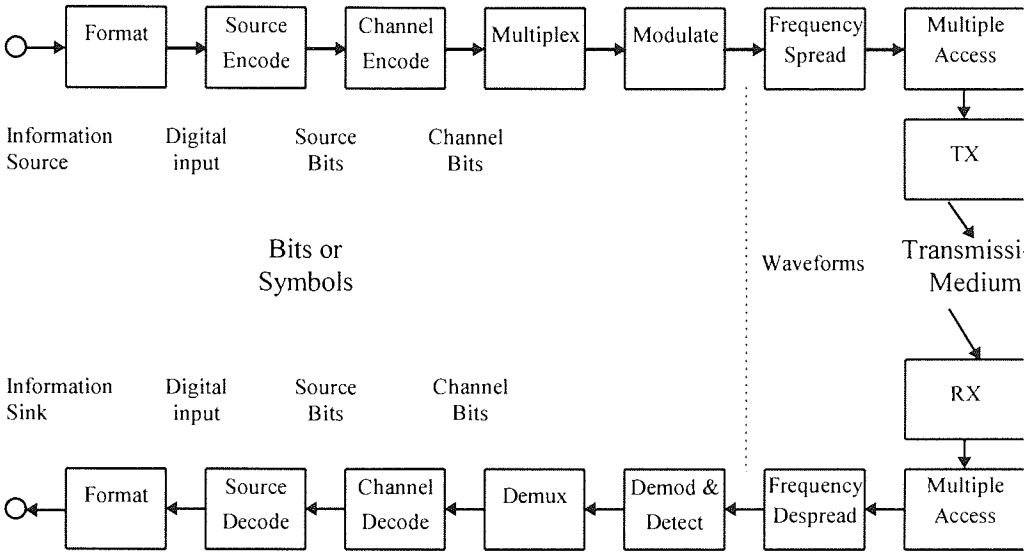


Figure 1-1 Generic Digital Communication System Model

The generic communication system model is used to organize the low level simul all subsystems in Figure 1-1, however, are needed simultaneously. A simplified block diagr in Figure 1-3, suffices to explain the overall simulation rationale.

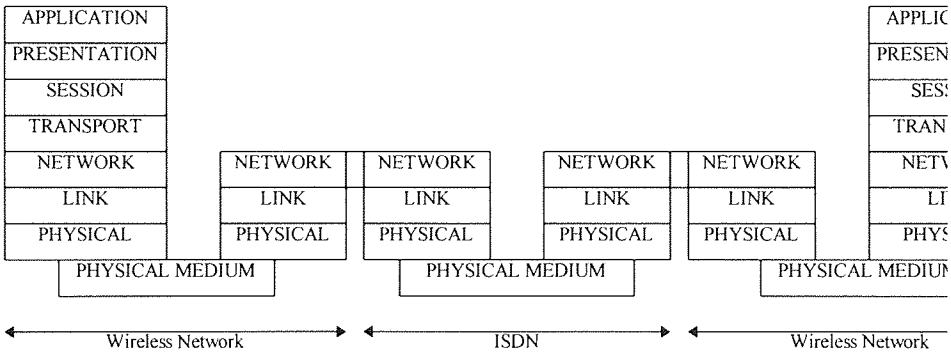


Figure 1-2 OSI Model of a Wireless Access Network

The simulations generally proceed in three phases. In the first instance a channel err developed by analyzing the bit error data generated by simulating a particular modulation scl

one of the radio channel models. Alternatively, the actual error data can be stored in an used directly.

Secondly, a digital video stream is generated by applying the video coding a standard image sequence. The bit stream is packetized if necessary and any forward coding applied to the video data blocks.

In the final phase the video stream is corrupted by modulo-two addition with th generated by the error model or with the actual error vectors obtained by simulation. access operations, and diversity techniques are simulated at this stage (e.g. by using r channel model for macro diversity techniques). The corrupted video stream is then deco in a file for later real-time replay.

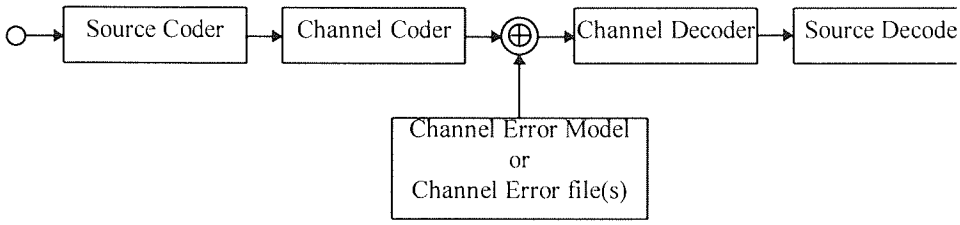


Figure 1-3 Simplified Simulation System

In certain cases, when feedback techniques are simulated, it is necessary to perfo and three simultaneously, since the forward video coding algorithm can be influence control data fed back by the decoder. In general, the generation of channel error v undertaken whilst being independent of the rest of the simulation, thereby improving throughput.

1.6.3 Video Codec

Rather than develop new video codecs it was decided to investigate a stanc codec. The H.261 codec was chosen as the only available interactive (conversational standard. The Motion Picture Expert Group (MPEG-1) standard [MPEG1] developed t digital coded video storage and retrieval, draws very heavily on the H.261 standard and stream compatible, the major H.261 operational blocks are contained as a subset of the M tool-set. The MPEG-2 standard, now reaching its final stages of development, and like basis of satellite and terrestrial digital TV and HDTV transmission, still retains the mai blocks. Thus the results derived for H.261 may be pertinent to a large number of curr video codecs.

1.6.4 Performance Assessment

Two performance measures are used in this study:- coding efficiency and

Coding efficiency is measured using transmission bit rates for fixed bit rate transmission anc and peak bit rates for variable rate transmission. Video quality can be measured using both and subjective measures. A number of objective measures are developed to measure sp temporal error propagation, and these are used to compare and report results. Formal subject tests were not conducted in this project. Instead, informal subjective results are repo summarized using a five grade "usefulness" measure.

In order to evaluate the results subjectively, the corrupted video sequences had to be v real time. A real-time display system was developed for SUN workstations and IBM compa capable of displaying colour video sequences. The length of the displayed sequence is limit size of the computer memory.

1.7 Thesis Outline

Chapter one covers the motivations behind the thesis, its aims and the methods used t these goals. Chapter two introduces background material which is necessary for the developm arguments in the thesis. This material covers video coding, wireless networks, and erro techniques.

Chapter three identifies wireless networks that can support video services and revie areas which are directly related to video transmission over wireless networks:- video coding s error control techniques for video, packet video techniques, developments in digital t television broad-casting and previous work in coded video transmission over wireless network

The hybrid motion-compensated, interframe DPCM/DCT (Differential Pulse Code M / Discrete Cosine Transform) video codec studied in this thesis is described in chapter four. A codec was implemented in accordance with the CCITT H.261 standard which is representati class of codecs. The Interactive H.261 Codec (IHC) implementation is discussed and the a used to implement the various modules are described. Validation results are presented for t using a number of standard image sequences.

Chapter five analyses the error resilience of H.261 to high, random Bit Error Rates (B main deficiencies of the codec are identified and several techniques are proposed to imp resilience to high BER. Performance results are presented comparing the proposed technic regards to effectiveness and efficiency.

Chapter six describes the theory and simulations used to generate radio channel erro Four radio channel models are used in the simulations: an Additive White Gaussian Noise model, a Rayleigh fading channel, a Rayleigh fading channel with cochannel interference, and ray, Rayleigh fading, frequency-selective channel. These channel models are used to gene files by simulating the transmission of digital modulated signals over these channels. The erro used in subsequent chapters to investigate the behaviour of coded video on the respective chan

The errors encountered on fading radio channels are bursty in nature. The behavio
codecs on burst error channels is investigated in chapter seven using the error files develope
six. The performance of a number of schemes used to mitigate the effects of multipath fadi
errors is measured and compared.

Chapter eight investigates the use of ARQ techniques to protect video transmission
networks. Traditionally these techniques have not been favoured for real time services su
and video. A justification of the use of ARQ techniques for video transmission over wirele
is presented and the requirements imposed by these techniques are investigated. Th
associated with video transmission over a wireless network are identified and a number of p
made to solve these problems. Two ARQ systems are then proposed and investigated by sim

The use of layered video codecs for robust transmission over wireless networks is i
in chapter nine. Similar codecs have been proposed for packet video transport over Variat
(VBR), Asynchronous Transfer Mode (ATM) networks. Two layered coding systems are pr
based on a layered video transcoder and the other on an H.261 compatible layered
transcoder based system is used to study Fixed Bit Rate (FBR) video transmission over a
The video codec is used to investigate video transmission over a third-generation wire
network. The techniques developed in the previous three chapters are used to guarantee
channel.

Finally, the results are summarized in chapter ten. A discussion of the results leads t
of recommendations and conclusions. Some problems identified in this thesis are dis
solutions of which form the basis of further work.

Chapter 2
Background

2.1 Introduction

The subject areas which are of direct relevance to this thesis are digital video codin networks and error control techniques. These three areas overlap as shown in Figure 2.1. T introduces topics which lie mainly outside the regions of overlap in Figure 2.1 The review p the next chapter concentrates on the regions of overlap. The main contributions in this thesi centre of Figure 2.1, where the three fields overlap.

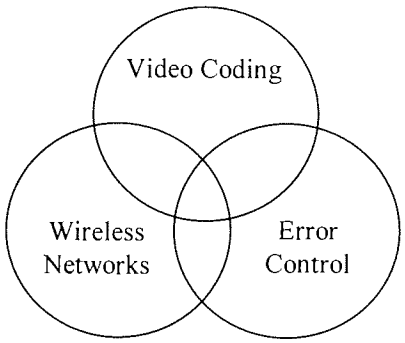


Figure 2-1 The interaction of the Research Fields

The next section reviews digital video coding techniques. This is followed by an o wireless networks. Finally, error control coding techniques are introduced and categorised ir compatible with the work in the thesis.

2.2 Digital Video Compression

Digitisation of video signals using Pulse Code Modulation (PCM) results in very hig for example TV pictures coded to ITU-R (ITU Radio) recommendation Rec. 601 [CCIR601 216 Mbit/s. The target bit rates for the majority of commercial transmission and storage ap however, are much lower than this, for example videophone and videoconference codecs ar to operate over basic rate and primary rate ISDN with bit rates of between 64 kbit/s and 2 I typically below 384 kbit/s. Thus video compression techniques must be used to meet these rates.

Digital image coding is concerned with the compression of digital images a sequences. It is a mature subject and a large number of coding techniques have been pub reviewed in the literature [Pratt78, Netravali80, Jain81, Musmann85, Gonzales87, N Frochiemer89, Lim90].

There are three ways in which techniques developed to code still images can be u image sequences. Firstly, an image coding technique can be used repeatedly to code separately in intramode. This does not capitalise on the large frame-to-frame correlations, h transmission errors are contained within the frame in which they occur and hence are not Secondly, two-dimensional coding techniques can be extended to three dimensions, with third dimension. The third possibility is to combine interframe DPCM with an Intrafr Encoding (ISE) technique to build Hybrid DPCM/ISE codecs.

2.2.1 Intraframe coding techniques

The majority of intraframe image source coding techniques can be broken down int operations [Gonzalez87, Lim90] namely Mapping, Quantization, Statistical Coding, ar Control as shown in Figure 2-2.

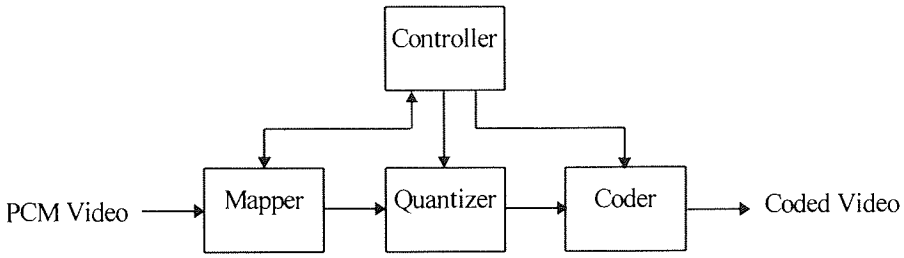


Figure 2-2 Image Source Coding Operations

The mapper maps digital pixel vectors into digital or analogue vectors whic amenable to compression. The mapper itself does not compress the image, indeed sometim the storage requirements as in pyramid coding [Burt81]. Mapper operations include transform and hybrid coding functions [Pratt78].

The quantizer quantizes the analogue or finely quantized mapper output. Th designed so that quantization results in fewer levels and fewer non-zero quantized Quantization achieves compression by discarding information and is the lossy stage in the chain. It reduces the subjective redundancy in the image and is usually designed to reduce of the quantization distortion by capitalising on properties of the human vision system [F quantizer input vector components can be quantized individually, using scalar quantiz vectors using vector quantization [Gray84].

The coder assigns codewords to the quantizer output. Both fixed-length and va codes can be used. Variable length codes can remove the statistical redundancy in the qua and hence achieve additional lossless compression.

The controller is used in adaptive schemes to optimise compression in a rate\dis by changing the source coding operational parameters in accordance with the local sub-im Image statistics are non-stationary over the image and hence adaptive compression can a

results than a non-adaptive scheme.

The main image coding techniques such as Differential PCM (DPCM), the Discr Transform (DCT) [Ahmed74], Vector Quantization (VQ) [Nasarbadi89], Pyramid Coding Sub-band Coding [Woods86], Wavelet coding, Fractal Coding [Barnsely88] and others, l reviewed in [Zammit91]. Two intraframe techniques are reviewed below, namely DPCM anc because both are of crucial importance to interframe image sequence coding and a hybrid DPCM\DCT scheme forms the basis of most standards available today [CCITT-H.261, MPE(

2.2.1.1 Predictive Coding

Intraframe predictive coding was amongst the first techniques to be studied compression because of its low complexity [Pratt78, Netravali80, Jain81]. A generic predic [Netravali88] is shown in Figure 2-3. In predictive coding the differential or error signal is fi by subtracting the predicted value of the current input sample from the value of the sample differential signal is then quantized, coded and transmitted. The received signal is decoded to the local predicted value to reconstruct the transmitted sample. The predictor is a fi previously transmitted samples, and it is the same in the transmitter and receiver. If the quantizes the differential signal to two levels then the predictive coder is also called a Delta (DM), otherwise it is referred to as Differential PCM (DPCM).

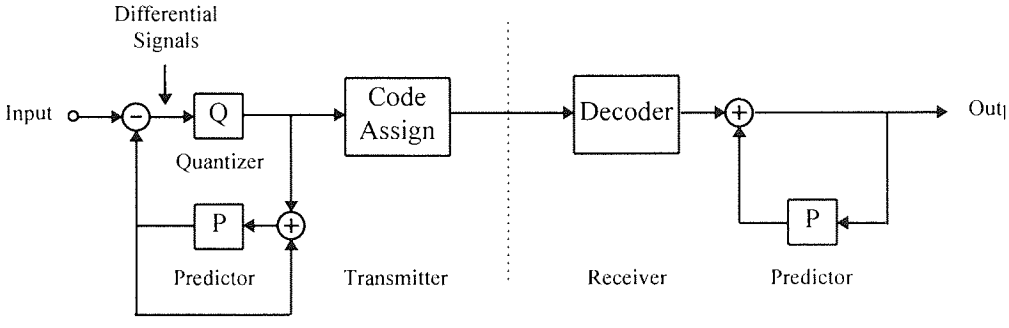


Figure 2-3 Block diagram of a predictive codec

In intraframe DPCM coding of image sequences, the predictor uses previous pixel the same line or previous transmitted lines, but in the current frame or field. The predictor ca or adaptive. Adaptivity can be switched or continuous, with or without the transmissi adaptation control signal (explicit or implicit signalling) [Pratt78, Netravali80, Jain81].

DPCM strategies have low complexity and are, therefore, attractive for application: coding rates in the region of 2-3 bpp (bits per pixel), e.g. for high quality digital TV transmis:

2.2.1.2 Transform Coding

In transform coding [Pratt78, Netravali88] linear, unitary, often separable, trans applied to subimage blocks, resulting in a less correlated, better compacted representation

better to code than the original image.

The image is first divided into LxL pixel subimage blocks **B** with n bits per pixel. Ea is then independently transformed into an LxL block of transform coefficients **C**. The coeffic quantized to produce the quantized coefficients matrix **q**=Q(**C**). Q(•) is the quantization functi quantizes each transform coefficient $c_{i,j}$ into the quantized coefficients $q_{i,j}$ with $l_{i,j}$, levels. ' entropy H_T of the quantized coefficients, assuming independent $q_{i,j}$, is given by:-

$$H_T = \sum_{i=1}^{L}\sum_{j=1}^{L}\sum_{k=1}^{\ell_{i,j}} p\left(q_{i,j}^{k}\right)\log_2\left[\frac{1}{p\left(q_{i,j}^{k}\right)}\right]$$

where $p(q_{i,j}^{k})$ is the probability of the k th level of the quantized coefficient $q_{i,j}$, i,j =1. compression is achieved, provided H_T < (LxLxn), by coding the matrix **q** coefficients indivi produce the coded matrix **m** = ζ(**q**) where ζ(•) represents the coding operation. If $q_{i,j}$ independent then H_T is not the minimum total entropy, and better compression can be ach conditional coding of the coefficients. Thus the transformation must de-correlate $q_{i,j}$ but qua Q(•) must deliver H_T substantially below (LxLxn) to ensure compression.

Energy compaction means that most of the energy is concentrated in a few t coefficients. Here, many low-energy transform coefficients can be coarsely quantized a discarded without incurring a heavy quantization noise penalty in the spatial frequency doma the transformation conserves energy, by Parseval's Theorem [Jeruchim92], inverse transf delivers the high signal to noise ratio to the reconstructed pixel block in the pixel domain.

The human vision system is less sensitive to noise in the high spatial frequencies desirable to compact the energy towards the low spatial frequency end of the transform, such high signal to noise ratio coincides with high perceptual image quality.

Let **T** be a linear orthonormal (or unitary) LxL matrix. Then [Netravali88]:-

$$\mathbf{T'T} = \mathbf{T^tT^*} = |\delta_{i,j}|$$

where **T'** is the conjugate transpose of **T**, $\mathbf{T^t}$ is the transpose of **T**, **T*** is the conjugate of **T**, Kronecker delta which is zero except when i=j, and i,j = 1,2..L. Then the forward separable t of the LxL matrix **B** can be written as [Netravali88]:-

$$\mathbf{C} = \mathbf{TBT^t}$$

and the inverse transform may be written as [Netravali88]:-

$$\mathbf{B} = \mathbf{T'CT^*}$$

The problem in transform coding is finding **T** which de-correlates the coefficients and optimal energy compaction. The Karhunen-Loeve Transform (KLT) [Pratt78, Netravali88] shown to have optimal energy compaction and de-correlation. Its computation, however, computation intensive. The Discrete Cosine Transform (DCT) [Ahmed74] performs very clc KLT, has a fast algorithm and is currently the most popular transform. Transforms such as the

Walsh-Hadamard Transform (DWHT) have square basis functions and are very easy to compute [Netravali88]. They do not perform as well as the DCT, however. The N-point DCT transformation matrix $\mathbf{T} = t_{mi}$ is given by [Netravali88]:-

$$t_{mi} = \sqrt{\frac{2-\delta_{m-1}}{N}} \cos\left\{\frac{\pi}{N}\left(i-\frac{1}{2}\right)(m-1)\right\} \qquad \text{(2-5)}$$

$$i, m = 1..N$$

The subjective performance of transform coders depends on the size and shape of the subimage blocks, the transformation used, the selection of the coefficients to quantize, the quantization process and code assignment [Pratt78].

Transform coders perform well at low bit rates and give good results at 16:1 compression ratios. The degradations introduced are mainly visible as blocking artefacts; one of the main criticisms of block orientated techniques. Post processing of the image at the block boundaries can reduce the blocking effects. The Lapped Orthogonal Transform (LOT) [Cassereau89] has been proposed to combat this defect by using overlapping blocks, but without a computational penalty vis-à-vis non overlapped TC.

Transform coding has been successfully applied to still, moving, monochrome, colour and multi-spectral images. The Discrete Cosine Transform forms the heart of the ITU-T compression algorithm for a nx64kbit/s video codec for use on ISDN, the ISO JPEG (Joint Picture Expert Group) still image standard, the MPEG-1 and MPEG-2 motion picture storage algorithms, and the CCIR 34 Mbit/s digital TV distribution codec standard [Stenger89].

2.2.2 Interframe Image Sequence Coding Techniques

Generally, image sequences hardly change from frame to frame. There is thus more scope for compressing video sequences using interframe techniques which compress groups of frames. Interframe compression ratios can typically be an order of magnitude higher than intraframe compression [Forchiemer89]. The amount of frame-to-to frame redundancy depends on the dynamicity of the video source [Netravali88]. In videophone-type head-and-shoulders sequences, there is limited motion due to the slow-changing nature of the scene and the absence of camera movement and zoom. Videoconference-type sequences consist of "waist-up" displays, with multiple subjects, greater motion and scene cuts. Broadcast-type sequences, with camera pan and zoom, and frequent scene changes, are the most dynamic and non-stationary. Statistics indicate that broadcast type sequences change by 15% on average between frames, conference sequences material by 11% and videophone sequences by 8% [Chin89].

2.2.2.1 Interframe DPCM

In interframe DPCM the predictor function P in Figure 2-3 is based on samples within both the previous and current frame or fields. The simplest interframe DPCM coder uses the corresponding

pixel in the previous frame as the predictor. Since an image seldom changes by more than 8%-15% on average [Mounts69, Chin89], this previous frame pixel predictor is a good prediction strategy. In regions where there is substantial motion the previous frame pixel prediction is no longer accurate, and an intraframe prediction will be more appropriate. Its simplicity makes pure interframe DPCM very attractive and it has been widely studied and used. In combination with the DCT it forms the heart of many proposed and standardised image sequence coding schemes.

2.2.2.2 Conditional Replenishment

Conditional replenishment [Mounts69] was developed by Mounts for video coding. Rather than transmitting information about all pixels in a frame, each pixel is compared to the previous frame pixel and if it has changed sufficiently it is coded, transmitted and used to update or "replenish" the displayed-frame store. In this scheme the positions of the changed pixels must be transmitted as overhead information. Conditional replenishment generates variable bit rates and introduces the need for buffers and buffer control strategies to prevent buffer overflow, if fixed bit rate transmission is required. Statistics show that only around 10% of all pixels change by more than 2% between frames at a frame rate of 30 Hz [Mounts69, Netravali88, Chin89]. This represents a 10:1 compression (ignoring side information) at a Peak Signal to Noise Ratio (PSNR - defined in section 2.2.3) of 34dB.

Candy, Haskell and Limb [Candy71] refined conditional replenishment by combing it with interframe DPCM and by using cluster-oriented change significance testing. The compression rate is increased because the pixels that have to be replenished are transmitted using interframe DPCM. Also clustering reduces the address overhead by having one cluster address for a block of pixels. The technique was further improved by intraframe differential encoding of the cluster address to further reduce the address overhead.

2.2.2.3 Motion Compensation

Rocca [Rocca69] proposed the use of movement compensation to utilise the frame-to-frame correlation to compress digitised television signals. Motion Compensated Prediction (MCP) reduces interframe error by estimating the motion of objects between two coded frames, and uses the estimated motion vectors to compute the prediction error between re-aligned pixels. MCP allows considerable coding gains over non-MC predictions. Gains of 1.5 to 2 are usually quoted [Netravali88]. The motion vectors describing the motion have to be transmitted together with the prediction errors to enable the receiver to reconstruct the frame.

Musmann and Pirsch [Musmann85] classified motion estimation algorithms into pixel recursive and block matching algorithms. The former use recursive techniques to estimate individual pixel displacement. Block matching techniques use spatial correlation on rectangular MxN pixel blocks. Motion estimation is also possible in the transform domain since translational and rotational motion will effect the phase of the transform coefficients, but not the magnitude [Netravali88].

Current standards use block matching techniques because these can be implemented in VLSI, thereby accomodating the large computational burden. Block matching searches for the two-dimensional vector $\boldsymbol{D}$ which minimises some measure of prediction error $PE(\boldsymbol{D})$ given by [Netravali88]:-

$$PE(D) = \sum_{b \in B, \hat{b} \in S} N(b(z,t) - \hat{b}(z - D, t - \tau)) \qquad \textbf{(2-6)}$$

using a distance metric $N(\bullet)$, where $b(z,t)$ is the monochrome intensity at position z in frame t, $z=\{x,y\}$, x is the horizontal and y the vertical dimension, b($\bullet$) is an element of the current MxN block B and $\hat{b}$ is an element of the (M±Y/2)x(N±X/2) search area S, in the previous frame (t-τ). In practice, the distance measure used to compute $N(\bullet)$ is either the Mean Absolute Distance (MAD) or the Mean Square Distance (MSD).

The computation of the motion vector $\boldsymbol{D}$ is very intensive if a full search of the search area is performed. A number of algorithms [Gharavi90, Ghanbari90a] have been proposed to reduce the number of search steps, compromising between accuracy and computational complexity. Usually square search blocks and areas are employed with 8x8 to 16x16 pixels.

2.2.2.4 Foreground/Background Prediction

Translation, rotation and deformation of objects in an image cause occlusion and exposure of the background. Brofferio [Brofferio89] proposed a three-state scene model which classifies pixels into an object, background and new scene pixels, and described an algorithm to build and maintain a background reference store. The updated background can be used to improve the prediction in uncovered background regions [Hepper90]. Two predictions are performed, one using the background information and the other using the previous frame. A selection circuit selects the minimum residual error and signals which predictor has been used to the decoder. Recently, Lettera and Masera [Lettera89] used foreground/background techniques to improve image segmentation into moving and stationary regions. Thoma and Bierling [Thoma89] have proposed the use of background information to improve the quality of motion compensated frame interpolation.

2.2.2.5 Frame Interpolation Coding

In fixed bit rate video transmission the encoder skips frames in periods of high image activity to lower the generated bit rate. The receiver then displays the last received frame n times (for n skipped frames) but this leads to loss of temporal resolution, evident as jerkiness in areas of motion. Alternatively, temporal frame interpolation [Musmann85, Thoma89, Caffario90] can be used at the receiver to improve the quality of the images. In temporal linear interpolation the missing frames are reconstructed by averaging pixels from a past frame and a future frame. This leads to loss of spatial resolution, however, evident as image blurring and a n-frame delay.

Motion Compensated Frame Interpolation (MCFI) [Thoma89, Caffario90] combines frame inter-polation with MCP to reduce temporal and spatial loss of resolution. The skipped frame blocks $b(z,t)$ (using the same notation in 2.2.2.3) are estimated from a past frame block $b(z+D',t+\tau)$ and a future frame block $b(z+D'',t-\tau)$ where D' and D'' are the motion vectors to the past and future frames from the current (skipped) frame. Then if only the vectors are transmitted, the receiver interpolates the skipped frame blocks using:-

$$b(z,t) = \frac{1}{2}\left(b\left(z+D',t+\tau\right)+b\left(z+D'',t-\tau\right)\right) \qquad (2\text{-}7)$$

Good quality MCFI depends on accurate motion estimation and proper selection of predictors; *e.g.* it is better to predict an uncovered (relative to the past frame) area in the skipped frame from the future frame, and it is better to predict areas that will be uncovered in the future frame, from the past frame.

Studies have shown that motion compensated frame interpolation can improve the PSNR of the reconstructed frames by 5 dB over field repetition and 2 dB over temporal linear interpolation [Thoma89, Caffario90]. MCFI is used in the MPEG standard in the bi-directionally predicted B frame [MPEG1].

2.2.2.6 Hybrid Interframe Coding

Hybrid coding combines pure interframe DPCM with Intraframe Source Encoders (ISE) described in section 2.2.1. There are two possible ways of combining the two techniques; Hybrid Intraframe Coding/DPCM as depicted in Figure 2-4 and Hybrid Interframe DPCM/Intraframe coding as depicted in Figure 2-5. The two schemes are not identical because the intraframe coder is not linear in general. MCP is easier to incorporate in Hybrid DPCM/Intraframe coders, but Hybrid Intraframe coding/DPCM is simpler because it does not require an Intraframe Decoder at the encoder end.

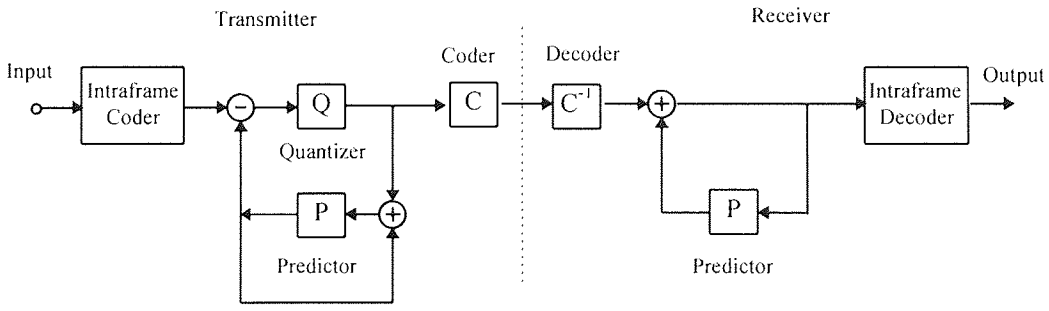


Figure 2-4 Hybrid Intraframe/DPCM coder and decoder

Hybrid coding is very popular because it allows the use of the relatively simple DPCM technique to capitalise on the frame-to-frame redundancy and to allow straightforward incorporation of conditional replenishment, background prediction, motion compensated prediction and frame

interpolation. Hybrid DPCM/Intraframe codecs have been implemented by combing DPCM with SBC [Gharavi89], VQ [Tubaro90], DCT [CCITT-H.261] and many other intraframe coders. The Hybrid DPCM/DCT codec is of special importance, as mentioned above.

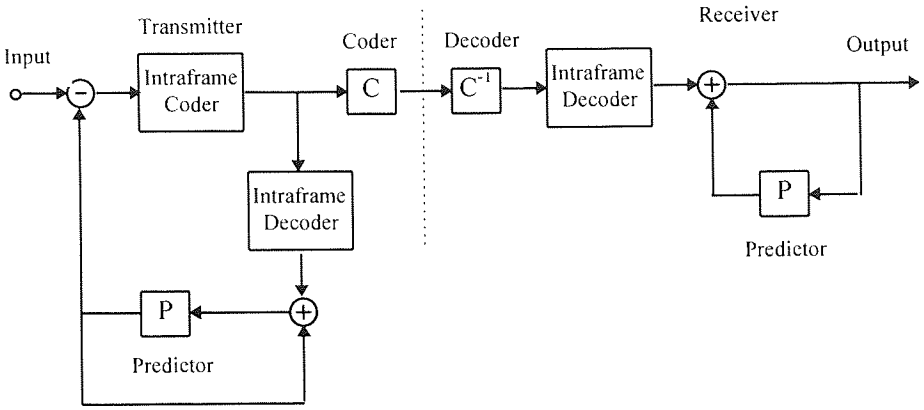


Figure 2-5 Hybrid DPCM/Intraframe coder and decoder

2.2.2.7 Other Image Sequence Coding Techniques

Numerous other image sequence coding techniques have been proposed in the literature based on DPCM, Vector Quantization, Sub-band Coding, Wavelets, Transform Coding, Fractal Transforms, Neural Networks, Model Based Coding, Object Based Coding, Pyramid coding and Region Based Coding etc. Only the areas relevant to the work in this thesis have been reviewed here. A taxonomy for image and image sequence coding has been developed in [Zammit91] and used to present a broader review of image coding and image sequence coding techniques based on the references reviewed above and in [Zammit91].

2.2.3 Objective and Subjective Fidelity Measures

Objective fidelity measures are required to assess and compare the performance of image communication systems in terms of the quality of the output image. Two often used measures in image coding are the Mean Square Error (MSE) and the Peak Signal-to-Noise Ratio (PSNR). For a monochrome NxM pixel image the MSE is given by [Pratt78]:-

$$MSE = \frac{1}{NM}\sum_{i=1}^{N}\sum_{j=1}^{M}\left(U_{i,j} - \hat{U}_{i,j}\right)^2 \qquad \textbf{(2-8)}$$

where $U_{i,j}$ is the pixel in the PCM frame and $\hat{U}_{i,j}$ is the pixel in the reconstructed frame with frame size of NxM pixels. For a monochrome image, the PSNR is given by [Pratt78]:-

$$PSNR = 10\log_{10}\left(\frac{A^2}{MSE}\right) \qquad \textbf{(2-9)}$$

where A is the peak-to-peak pixel amplitude, usually 255 for byte sized pixels.

These measures are easy to compute, but equally graded images often differ markedly when

assessed subjectively. A number of alternative measures, based on the characteristics of the Human Visual System (HVS), have been suggested which are claimed to perform consistently with subjective tests [Voran92].

Ultimately, quality measurements must depend on subjective evaluations either directly or indirectly. Two possible subjective evaluation methods are the rating-scale method and the comparison method [Netravali88]. In the rating-scale method, subjects assess image properties, such as quality and imperfections, and assign a value from a rating scale. In comparison methods, noise is added to one image until it is deemed to be of comparable quality to a reference image. The set of reference images have to be rating-scaled, such that the tested images can then be assigned a quality measure.

Subjective evaluations are expensive and time consuming. There is a requirement for an image quality measure which is easy to compute, can be applied objectively to images, and grades images consistently on a perceptual scale.

Recently, Voran and Wolf [Voran92] have developed an objective video quality assessment system that emulates human perception. They corrupted a library of test scenes using impairment generators, collected subjective and objective quality ratings and used joint statistical analysis to match the objective results to the subjective evaluations. Three objective measures were defined; one spatial measurement based on the Sobel edge detector and two temporal measurements based on the difference image. The objective assessment sought is then obtained by weighting and adding the three measures together. The correlation between the predicted scores and the true scores is very good for sequences outside the training set.

In the literature, the MSE and the PSNR are invariably used to report picture quality. In standardisation work these measures are used as an initial guide to determine the likely candidates, but the schemes are also tested subjectively [Yasuda89]. *In this thesis the mean square error and the PSNR are amongst the objective measures used to report the results. Informal subjective evaluations were also carried out and are reported together with the objective results.*

2.3 Wireless Networks

Wireless communications is attractive because it allows user mobility, ease of use and reconfiguration, and decreased installation and modification costs. A diverse range of applications can be supported, a number of them new and made possible only because of the mobility and increased flexibility. This section presents a broad overview of the technical aspects of Wireless Networks, concentrating on the Physical Medium, the Physical Layer, the MAC Layer and the Data Link Layer in OSI terms.

2.3.1 The Physical Medium

Wireless Networks (WN) use electromagnetic radiation, either radio or infrared, to achieve wireless communications. Currently, infrared systems are limited largely to short-range indoor Wireless

Networks (WLANs) [Barry91, Kahn94, Chu87] implemented either as Line Of Sight (LOS) or non-directive, diffused systems. In these applications infrared systems enjoy a number of advantages over radio based systems: a virtually unlimited and unregulated bandwidth, high data rates (of up to around 100 Mbit/s currently [Barry91]) and better confinement (and hence better security).

Radio-based systems do not have the drastic range limitations of infrared systems and are suitable for short, medium and long range wireless applications. Radio signals are difficult to confine, leading to security issues and some environmental and health concerns. *Radio-based systems predominate in virtually all wireless network applications, however, and only radio-based wireless networks are considered in this thesis.*

A major problem in implementing radio-based WNs is the acute shortage of radio spectrum. Radio spectrum allocations are made at a national, continental and international level. For example, the Federal Communications Commission (FCC) and the Department of Trade and Industry (DTI) regulate and licence the use of radio spectrum in the USA and the UK respectively. The European Radio-communications Committee (ERC) regulates the use of radio spectrum for pan-European applications. The highest world authority, however, is the International Telecommunications Union (ITU). The ITU organises the World Administrative Radio Conference (WARC) where its 170 nation members meet to agree on the allocation of radio spectrum on a global scale and decisions taken here affect frequency assignments at all lower levels. The last WARC was held in 1992 and a number of important allocations have been made for new and emerging wireless network systems [Adams92, Young92].

2.3.2 The Physical Layer

Issues which are relevant to the physical layer concern the wireless network topology, the efficient use of the radio bandwidth, the properties of the Radio Propagation Channel (RPC), the effects of the RPC on digital modulation, the limits the RPC imposes on the wireless networks and the countermeasures used to alleviate these limitations.

2.3.2.1 Wireless Network Topologies

A radio based wireless network comprises mobile stations and base stations inter-communicating using radio links. The base stations are stationary and allow the mobile stations to access fixed communications networks. The organisation of the network nodes and the communications links determines the network topology. It is possible to contemplate a one-layer system in which the mobile stations communicate directly with each other and the base stations. The most common topology, however, is one in which the base stations assume a central role in the system, with the mobile stations transmitting directly to the base stations which relay the signals to the receiving mobile stations. The base stations can communicate with each other over fixed networks and thus allow mobile stations outside radio reach of each other to communicate. *This two-level, star*

hierarchy is used in the majority of wireless networks in current or contemplated use, and is the only topology considered in this thesis.

2.3.2.2 The Cellular Concept

As already mentioned, radio is a very scarce resource which must be used as efficiently as possible. Cellular systems adopt the frequency re-use approach to use spectrum as efficiently as possible. In this technique the service area is split into cells each served by a base station. The base stations are connected to each other through switching and call processing equipment. Cells are grouped into K-cell clusters, which are organised in a regular pattern which tessellates the whole service area. The whole system frequency bandwidth is re-used within each cluster. Each cell is allocated 1/K th of the overall system bandwidth. The K frequency allocations are assigned so that no adjacent cell uses the same frequencies so as to minimise Cochannel Interference (CCI). The cochannel Signal-to-Interference ratio SIR is defined by:-

$$\text{SIR} = \frac{\text{Average Power of wanted Signal}}{\text{Average Power of Interfering Signals in the same bandwidth}} \qquad \textbf{(2-10)}$$

and it is usually expressed in dB.

The reuse distance D is defined as the minimum distance between cell sites using the same frequency. For hexagonal cell shapes of radius R_C and K-cell clusters the following relationship can be shown to hold [Parsons89]:-

$$D / R_C = \sqrt{3K} . \qquad \textbf{(2-11)}$$

The smaller the value of K the greater the frequency re-use, but the higher the CCI. A design compromise exists, but cellular networks are interference limited.

When a mobile station moves outside the service area of the current base station it is handed over to the closest base station whilst the call is still in progress. Thus mobile stations can roam the whole service area with a low probability of dropped calls. Some cordless systems do not have this capability.

The cell sizes vary from system to system. Macrocellular systems use large cell diameters sometimes above 30 km. Microcellular systems use smaller cell diameters from several hundred metres to several kilometres. Indoor systems may use picocells with a diameter of several tens of metres. *The work in this thesis concentrates mainly on microcellular and picocellular systems which are associated with higher data rate systems.*

2.3.2.3 Digital Wireless Networks

First-Generation analogue mobile cellular systems (refer to Appendix 1) use analogue frequency modulation to transmit voice. Second-Generation mobile networks use digital techniques and use the available bandwidth more efficiently. Digitized voice can be compressed such that after modulation, the bandwidth occupied is substantially less than with the analogue voice signal. As an

example, the 3 kHz baseband voice signal occupies 30 kHz after frequency modulation in the ETACS system [Parsons89]. The Electronics Industry Association's Interim Standard 54 (IS-54) in the USA [Goodman91b] uses digital techniques to transmit three voice channels in a bandwidth of 25 kHz; six voices channels when half-rate codecs are introduced. The digital Global System for Mobile (GSM) developed for Pan-European service [Goodman91b] achieves the same order of capacity increase with full and half-rate voice codecs.

2.3.2.4 Spectrally Efficient Modulation

The second-generation digital mobile systems in operation use modulation schemes with a modulation efficiency below 2 bit/s/Hz *e.g.* 1.35 bit/s/Hz for GSM and 1.62 for IS-54 [Goodman91b]. The need for ever higher data rates, and the restrictions on the available bandwidth means that there is a need for more spectrally efficient modulation techniques. A number of spectrally efficient modulation techniques have been studied recently, such as M-ary PSK techniques (M>4), 16, 32 and 64 QAM [Webb92b], and 16 and 32 Star QAM [Webb91, Nix91]. Higher spectral efficiency comes at the cost of a lower immunity to thermal noise and system interference.

2.3.2.5 Radio Propagation

In a wireless environment the radio waves experience attenuation, shadow fading and multipath fading [Jakes74, Hashemi93]. Radio wave attenuation in a wireless environment follows an inverse n power law, where n can be lower than for free space ($n \approx 1.6$) in line-of-sight propagation, and greater than for free space (n between 3 and 6) for obstructed paths.

Multipath fading arises because the radio signal from the transmitter arrives at the receiving antenna via multiple paths due to reflection, scattering and diffraction. This multipath propagation environment leads to two phenomena, namely fading in narrowband systems and pulse dispersion in wideband systems.

A narrowband system can be described as one where the propagation delays associated with each path are extremely small compared to the inverse of the signal bandwidth [Acampora87b]. In this case the arriving multiple wavefronts add vectorially at the receiver with random phase, leading to a Rayleigh distributed envelope where there is no line of sight between transmitting and receiving antennae [Proakis89]. If a line-of-sight exists, then the distribution is Ricean. The vectorial addition can lead to very deep fades, typically exceeding 30 dB in both indoor and outdoor environments. If a mobile moves within this propagation environment it will experience time varying fading at a rate directly related to the transceiver speed and the carrier frequency. The time varying envelope statistics can also be characterised by Doppler fading in the frequency domain [Proakis89].

Large obstacles, such as buildings in the mobile environment and walls and floors in an indoor environment, obstruct the direct path between the transmitter and the receiver and cause shadow fading. Shadow fading causes variations in the local mean of the Rayleigh distribution described above, as the

mobile moves in the shadowed environment. The variations are gradual within a large area and shadow fading is also called slow fading, to distinguish it from Rayleigh fading which exhibits many complete cycles of variation within the local area and is, hence, also called fast fading. Shadow fading has been found to have a lognormal distribution in the mobile environment and in some indoor radio environments [Hashemi93], though the variance is markedly less in the latter case. Shadow fading is spectrally flat.

A wideband system is one in which the delay variations between the arriving rays cannot be ignored and lead to signal distortion. This delay spread is experienced as frequency selective fading in the frequency domain. The wideband properties of the channel can be modelled by its impulse response h(t) which can be measured directly [Devasirvatham87, Saleh87]. The excess power delay profile is given by $|h(t)|^2$ [Devasirvatham87]. The second moment of the excess power delay profile, also called the delay spread, is a very important parameter in wideband digital systems because the BER experienced by wideband digital modulation techniques is found to be proportional to the rms delay spread [Chuang87].

The excess delay spread depends on the environment, the separation between the transmitter and the receiver and the type of antenna system used [Acampora87b]. Thus in a mobile environment delay spreads above 5 μs are normal. In large open plan office buildings or factories, with a central antenna, the delay spread may be around 250 ns [Rappaport91a]. In a typical office room it reduces to 50 ns. With a distributed antenna system, the delay spread in most buildings is around 25ns [Acampora87b, Rappaport91a].

2.3.2.6 Bit Error Mechanisms on Multipath fading channels

In the absence of fading, the Bit Error Rate (BER) decreases rapidly as the received Signal-to-Noise Ratio (SNR) increases. When a digital signal is transmitted over a multipath fading channel, however, the BER characteristics decrease much more gradually [Proakis89]. The error characteristics on multipath channels depend not only on the received average SNR but also on the bandwidth of the signal, the delay spread of the propagation environment, the mobile speed and the quantity and distribution of interferers. The delay spread and the signalling rate combine to determine whether the system is narrowband or wideband, and this determines the characteristics of the error generating processes.

Error Mechanisms on narrowband fading channels [Yoshida88]

On a narrowband channel two separate mechanisms generate errors. The first is the Rayleigh fading envelope which causes periodical SNR deterioration, leading to bursts of error. The performance of digital modulation schemes is substantially worse on a Rayleigh fading channel than on the Additive White Gaussian Noise (AWGN) channel [Proakis89].

The average BER due to Rayleigh fading is substantially independent of mobile speed (ignoring the irreducible error floor described below). The higher-order burst error statistics, however, are intimately related to the mobile speed [Jalal91]. At low speeds the error bursts and error gaps are long, but become shorter as the mobile speed increases. The average BER caused by Rayleigh fading can be decreased by increasing the SNR e.g. by increasing the transmitted power.

The second mechanism generates errors due to the rapid phase changes which occur at the bottom of a deep fade, where the In-phase and Quadrature components are both small. Then, provided the bit period is sufficiently long, the resulting phase change may become large enough to cause errors. The rate of change of the envelope increases with speed and, hence, the errors induced by this phenomenon increase with vehicle speed. As the baud rate increases, the symbol period decreases and there is less time for phase change accumulation between symbols. Thus the error rate decreases with increased transmission rate. The BER caused by 'fast' Rayleigh fading does not decrease with increased SNR unlike that caused by 'slow' Rayleigh fading. Thus these errors constitute an irreducible error floor as the SNR tends to infinity. The errors are strongly correlated to deep fades, so that antenna diversity can be used to alleviate their intensity. The irreducible errors occur in bursts which are correlated to the fading envelope.

Error mechanisms on wideband fading channels [Yoshida88]

At high transmission rates, the irreducible error floor due to fast Rayleigh fading disappears. As the system bandwidth approaches and exceeds the coherence bandwidth [Proakis89] of the channel, however, Intersymbol Interference (ISI) sets in. ISI arises because of the pulse dispersion introduced by the channel as described above. This ISI may cause eye closure independent of SNR and hence represents another irreducible error floor [Yoshida88].

The optimal receiver for ISI channels is the RAKE receiver [Proakis89]. A less complex option is to use antenna diversity [Jakes74, Proakis89, Parsons89] to alleviate the BER due to ISI since uncorrelated paths are less likely to suffer ISI simultaneously. *This thesis investigates relatively simple counter measures to combat ISI namely antenna diversity, coding, frequency hopping and variable rate techniques.*

Interference

Wireless links are afflicted by internal and external interference. Wireless networks generate internal or self interference in the form of Cochannel Interference (CCI) and Adjacent Channel Interference (ACI). CCI is due to the re-use of frequencies in nearby cells which, under normal fading conditions, are received sufficiently strongly to interfere with the legitimate signals. ACI is due to non-perfect filtering in the receiver which admits a certain amount of energy from adjacent frequency channels. CCI is of particular importance in cellular systems. The CCI causes burst errors as the main signal fades with respect to the interfering fading signals.

Wireless links are also affected by external noise sources which are frequently impulsive in nature such as atmospheric noise and car ignition noise [Parsons89]. These systems are also prone to jamming either intentionally or accidentally, although some systems, spread spectrum systems in particular, are more immune to this kind of noise than others.

2.3.2.7 Countermeasures

Various countermeasures have been devised to combat the effects of fading. Antenna diversity [Jakes74, Proakis89, Parsons89] is very effective in combating Rayleigh fading and CCI. For example, Proakis shows that second order antenna diversity can improve the BER for QDPSK by more than an order of magnitude at an average SNR/bit of 20 dB [Proakis89].

The optimal receiver in a wideband, frequency selective environment is the RAKE receiver [Proakis89]. An often employed solution is adaptive equalisation which continuously sounds the channel impulse response and takes measures to equalise the delay and phase of the detected echoes. Spread spectrum is another technique which can be used to avoid ISI when combined with RAKE decoders [Proakis89]. Recently, Orthogonal Frequency Division Modulation (OFDM) has also been proposed to combat ISI. A Frequency hopping system avoids interference by jumping to a new frequency when the link deteriorates below some minimum threshold. The hopping distance is determined by the coherence bandwidth of the channel [Saleh91]. Variable rate transmission techniques have also been proposed [Acampora87a ,Zhang90] in which the transmission rate is chosen dynamically, depending on some measure of the current link performance.

2.3.2.8 System Limitations

A wireless network is designed to meet certain targets such as user and traffic density, terminal access rate, reliability, bit error rate and quality of service [Acampora87b]. The latter is usually expressed as the probability of encountering system degradations.

Maximum Range

The maximum range is reached when the terminal fails to meet the system performance objectives (*e.g.* BER > 10^{-4} for less than 10^{-4} of the time). The maximum range is limited by the available transmit power, propagation attenuation, the receiver sensitivity, shadow fading, Rayleigh fading, thermal and impulsive noise, and system interference. The range can be improved by using countermeasures such as diversity [Nix91, Nix92], coding and frequency hopping [Saleh91].

Maximum Transmission Rate

The Shannon channel capacity equation limits the transmission rate on AWGN channels. Pulse spreading causes ISI which leads to an error floor substantially above the Shannon limit which cannot be improved just by increasing the transmit power. The maximum rate R_{max} without equalization and diversity is given by [Saleh91]:-

$$R_{max} \approx (\sigma_{RMS})/10 \quad (2\text{-}12)$$

where σ_{RMS} is the root mean square delay spread. Thus in a building with 250 ns delay spread, the maximum transmission rate without equalization and diversity is approximately 400 kbit/s. If the spread is less, 50 ns for a typical office, then rates up to 2 Mbit/s are achievable. The transmission rate can be improved by diversity [Nix91, Nix92], coding and frequency hopping [Saleh91], variable-rate systems [Acampora87a, Zhang90] and equalization [Webb91].

2.3.3 Medium Access Control

In wireless networks the Medium Access Control (MAC) layer regulates the access to what is essentially a broadcast channel. It determines which user can access the channel, at what time and for how long. In star networks the bandwidth access is controlled by a central base station. Here, the down-link (or forward channel; base-station to mobile) access represents no problem since the base-station multiplexes all the forward channels on the one broadcast channel transmitted to all listening stations, which select their own channels as directed by the base-station. The base-station multiplexes the signals using Time Division Multiplexing (TDM), Frequency Division Multiplexing (FDM) or Code Division Multiplexing (CDM).

Access on the up-link (or reverse channel; mobile to base-station) is more problematic because the base station has to co-ordinate the transmissions from the mobiles so that they do not interfere with each other's transmissions. It is here that the MAC protocols assume their importance.

Existing multiple access protocols may be categorised as random access or scheduled access, with a spectrum of variants in between [Li87]. The ALOHA protocol is the original random access protocol originally incepted for a radio-based computer network by Abramson *et al.* at the University of Hawaii in the 1970's [Tanenbaum89]. Random access protocols are used on packet radio networks but scheduled access protocols are more favoured by voice oriented networks due to the real time nature of voice.

There are two basic types of scheduled access MAC protocols; static channel allocation and dynamic channel allocation. Static allocation MAC protocols assign resources for the duration of the call, whereas dynamic channel allocation protocols assign resource on the fly, when the user demands service during the duration of the call.

There are three basic fixed channel allocation MAC protocols: FDMA, TDMA and CDMA. In Frequency Division Multiple Access (FDMA), the available bandwidth is divided into frequency channels, and a user is assigned one of the channels following a successful call establishment phase. In Time Division Multiple Access (TDMA) the whole system bandwidth is organised into timeslots, and each user is assigned the whole bandwidth for the duration of the allotted timeslot and. the timeslot identifies the user. In Code Division Multiple Access (CDMA) the whole system bandwidth may be used by a number of users simultaneously. The system issues an unique code to each user, which is

then used to code the users transmission by multiplication with the bit stream. The base station can separate each user's transmission by correlating the received signals from all users with the code of the particular user.

In Dynamic Channel Allocation (DCA) a signalling channel is available either per user or it may be contention based. When the user wishes to transmit packets to the base station it signals its intention to the BS and waits for a reply. The BS then grants the use of a channel for a particular duration and signals the information to the MS which proceeds to transmit its data as scheduled. An example of contentionless DCA is the Packet Reservation Multiple Access (PRMA) protocol [Goodman89].

The MAC has a very important role to play in wireless networks as it determines the throughput available to each user and the access delay characteristics which are of paramount importance for real time services.

2.3.4 Data Link Layer-Error Control

It has been shown that wireless channels suffer from severe error bursts. Countermeasures are necessary to reduce the frequency and severity of these bursts, however they cannot be eliminated altogether at modem level. One of the main functions of the data link layer is that of error control to recover from these error bursts. Error control techniques are introduced in section 2.4 below.

2.3.5 Wireless Network Services

The first-generation wireless networks were designed for voice communications. Starting from the second generation digital networks data services have become an integral part of the service offering. The number of data services on offer is on the increase and includes Fax transmission, computer data transmission, text, messaging services *etc*. Third-generation networks will support integrated multimedia applications, integrating video and audio services with the current voice/data services

2.3.6 Current Wireless Network Offerings

Numerous wireless networks are on offer currently. Table 2-1 attempts to categorise these wireless networks coherently according to a taxonomy developed in appendix one. The last column in Table 2-1 includes particular wireless networks overviewed in appendix one which contains a comprehensive list of references. The first level in the taxonomy distinguishes between Wide Area Wireless Networks (WAWNs) and Local Area Wireless Networks (LAWNs). WAWNs have global coverage whereas LAWN coverage is limited to a relatively small local area. LAWNs can still access global networks, but the global coverage must be planned and executed by the end user, whereas in WAWNs global coverage is the responsibility of the service provider. Thus LAWNs can be said to provide *Access* to global networks.

WAWNs can be further split into Terrestrial and Satellite based networks. Satellite networks can be broadly classified into Land, Maritime and Aeronautical networks although the distinction between the Land and Maritime networks is not very clear. Thus the INMARSAT offerings can be used by both Land and Maritime users. Another possible classification is into GEO, MEO and LEO systems depending on the height of the **E**arth **O**rbit, whether **G**eostationary, **M**edium or **L**ow.

Terrestrial based WAWNs currently dominate the mobile scene. These can be further classified according to the main service provision which can be Voice, Data or Messages. Note that Voice oriented networks still support data and fax transmission using appropriate radio data modems. Data oriented networks are targeted at computer and information gathering services and do not offer voice communications. Message-oriented services include Paging systems, which allow text based messages to be sent or received on pagers.

Wide Area Wireless	Terrestrial	Voice-Oriented	Analogue Cellular	AMPS, TACS, ETACS, NMT
			Digital Cellular	GSM900, DCS1800, IS-54, IS-95
			Public Mobile Radio (PMR)	
		Data-Oriented	Mobitex, Cellular Data	
		Message-Oriented (Paging)	POCSAG, ERMES	
	Satellite	Land	Euteltracs, PRODAT	
		Maritime	INMARSAT A, B, C	
		Aeronautical		
Local Area Wireless	Cordless	Analogue	CT1	
		Digital	CT2, CT3, DECT	
	Wireless LANS	Standards-Based	ETSI-based	CAT 1, 2, 3
			IEEE 802.12	
		Proprietary	Altair, Wavelan etc.	
	Radio Local Loop	PACS, WACS		

Table 2-1 A Taxonomy of Current Wireless Networks

Voice-oriented WAWNs are further classified into Analogue and Digital mobile networks. Analogue mobile networks belong to the first generation of wireless networks whereas the Digital mobile networks belong to the second generation. Second Generation digital networks include Global System for Mobile (GSM), Digital Communication System at 1800 MHz (DCS1800), IS-54 and IS-95. DCS1800 is based on the GSM system but operates at 1800 MHz instead of 900 MHz (GSM is also referred to as GSM900). IS-54 is a digital extension of AMPS (also referred to as D-AMPS). IS-95 uses a slice of the AMPS bandwidth to provide a digital CDMA based service. GSM and DCS1800 are incompatible with earlier systems. Dual mode sets exist for IS-54 and IS-95 which also allow access to the analogue AMPS service.

Data-oriented networks allow packet-based data communications. A number of networks exist such as Mobitex and ARDIS. Un-utilised bandwidth in mobile cellular systems is also being used to offer cellular data services.

Paging networks initially offered only one way messages to the mobile pagers. The main paging standards are POCSAG and ERMES, being developed for the European market. A new generation of pagers offer two way messaging, and even access to electronic mail and the Internet.

Local Area Wireless Networks can be subdivided into Cordless, Wireless LAN and Radio Local Loop systems. The first Cordless Telephone (CT) systems were analogue and are now referred to as CT-1. The second-generation CT-2 systems is digital and embraces the Universal Mobile Telephone (UMT) concept which envisages the same handset to be used in the home, office and outdoors in the vicinity of base stations. The first CT2 systems deployed in the UK were incompatible and allowed outgoing calls only in the outdoor environment. The Common Air Interface (CAI) standard made the systems compatible, and later improvements allowed both incoming and outgoing calls, and handover between contiguous cells. The Digital European Cordless Telephone (DECT) system is a second-generation digital CT system designed by the European Telecommunication Standards Institute (ETSI). DECT also embraces the UMT concept. CT-3 refers to a variant of DECT offered as a commercial wireless PBX.

Wireless Local Area Networks (WLANs) allow mobile computers to access computer LANs but may also function independently of the fixed LAN. Proprietary WLAN systems abound like Altair, WaveLan *etc.* A new generation of WLANs has been developed for the 2 GHz band, initially in the USA, but now also in Europe and the rest of the World which allows unlicensed, but regulated, use (at least in the USA)

Two standards bodies are working to regulate WLANs. ETSI in Europe is standardising three categories of WLANs. CAT-1 LANs are data-oriented and are intended for low user density, low data rate usage. CAT-2 systems allow a higher user density, a higher access rate and integrated voice-data services. DECT is a typical CAT-2 system. CAT-3 systems have the highest user density and allow data terminal rates up to 10 Mbit/s. The High Performance LAN (HIPERLAN) is being standardised by ETSI in this category.

In the USA the IEEE 802.11 committee has been developing standards for WLANs with a minimum data rate of 1 Mbit/s [Links94]. The committee has developed one MAC protocol to support all physical layers. The MAC - Carrier Sense Multiple Access - Collision Avoidance (CSMA/CA), can also support real-time, contention free services. In addition, a number of physical layers have been defined.

Radio Local Loop systems use wireless links to replace the copper local loop thereby allowing wireless operators to compete with the public switched telephone network providers. PACS and WACS developed in the USA target such applications and aim to deliver ISDN services as well as normal

voice services.

2.3.7 Evolution of Wireless Networks

Wireless communications systems have experienced very rapid growth in the last two decades. The first-generation analogue mobile and cordless networks have been joined by the second-generation digital networks, and the planning of third generation wireless networks is well underway which will see the merger of mobile and cordless systems [Gardiner90, Steele94]. The ethos of third generation systems is encapsulated in the Personal Communications Systems (PCS) concept [Steele92] which will be delivered by Personal Communications Networks (PCNs). The third generation wireless networks are being designed jointly by the European ETSI body through its UMTS (Universal Mobile Telecommunications System) initiative [Chia92] and by the ITU-R through its FPLMTS (Future Public Land Mobile Telecommunications System) initiative [Callendar94].

2.3.7.1 Personal Communications Networks and Systems

The Personal Communications Systems concept will allow users to carry a personal communicator (PC) wherever they may be. This calls for lightweight handsets that are easy to use, and which have a wide area coverage, mobility and high user density both indoors and outdoors. The idea is that the PCs will be as ubiquitous as the telephone. Given the scarce spectrum allocations it would seem that to achieve the target penetration, especially in large conurbanations, smaller cell sizes would have to be used [Steele89].

The practical realisation of the PCN has been approached by the DCS1800 system [Potter92]. It is based around the GSM standard but operates at 1.8GHz, to provide a hand portable service. 150 MHz of spectrum has been made available in the 1.8 GHz band. However it will may not be as pervasive as the concept of PCN would have.

The PCN merges the 'cellular' and 'cordless' concepts. In fact a PCN could be envisaged, made up of combined GSM, DECT and ERMES (paging) base stations which can capitalise on the reach of ERMES, the mobility of GSM and the economy of DECT.

2.3.7.2 UMTS

The RACE UMTS project is a component of the RACE IBCN initiative [Gardiner90]. The RACE mobile is intended to supersede the pan-European GSM. It is planned to support voice and data services up to 200 kbit/s close to the 2 Ghz band. The 60 GHz band is also being considered for a system which will allow 2 Mbit/s data rates over limited areas. It will thus integrate existing type of services, cellular, cordless and paging and provide data transmission up to 2 Mbit/s. This will make possible indoor wireless LANs and integrated voice/data networks.

The RACE mobile initiative may be based on a three layer hierarchy of macrocells, microcells and picocells similar to that proposed by Steele [Steele89]. In this scheme the microcells would carry

the burden of traffic within 100m to over 3 Km distance from the BS at error rates below 10^{-4}. The picocells will be deployed in office and other regions of high traffic concentrations e.g. at telepoints. The macro cells (>3 Km) would be deployed to support microcells and for use in low density regions, such as in rural areas. The concept sees BSs being deployed within domestic and commercial buildings as well as along roads, typically on lampposts, and connected by optical fibre networks. Experiments have shown [Steele89] that Basic Rate (BRA) ISDN could be deployed via microcells, whereas picocells could support primary rates.

2.3.7.3 FPLMTS

FPLMTS [Callendar94] shares much the same ideals with the UMTS and PCS concept and ETSI and ITU-T are collaborating on this project. Whilst ETSI's UMTS is a pan-European project, ITU-R FPLMTS has global implications. Thus FPLMTS is targeting both terrestrial and satellite base wireless systems. The target bit rate is similar to UMTS, with ubiquitous BRA ISDN rate access, peaking up to 2 Mbit/s for some wideband applications.

2.4 Error Control Techniques

Error control is a mature field, with an established theoretical foundation rooted in the original work of Shannon [Shannon64] in the 1940s. The theoretical aspects in this thesis are based on [Lin83] and [Michelson85] whereas the simulation aspects are based on [Jeruchim92], with algorithms from [Capellini85] for Forward Error Correction (FEC) simulations, and [Tanenbaum89] and [Halsall92] for ARQ simulations.

2.4.1 Error Control Fundamentals

The telecommunications system channel model is shown in Figure 2-6. The channel encoder is responsible for adding error control fields to the source bit stream. The modulator transforms the bit codewords into real-valued symbols, which are transmitted over the channel where noise and interference corrupt the symbols. The demodulator reverses the modulation operation to deliver the channel coded bit stream. The decoder uses the potentially corrupted channel coded data to recover an uncorrupted copy of the transmitted source data.

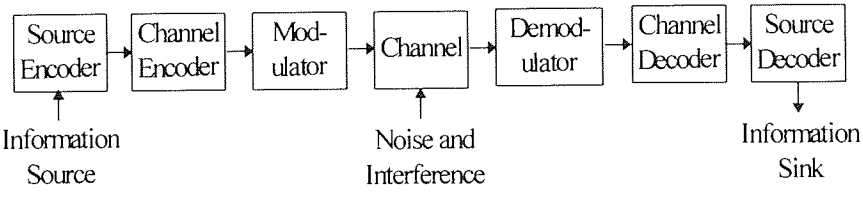


Figure 2-6 Channel Model

In this thesis the following nomenclature is used to describe system wide applications. End-to-End error control will refer to error control operations which span the whole network, which may

comprise many links. Link error control will refer to the case were the error control is active across one link in the network. Link error control operations may be nested inside end-to-end error control systems, often transparently. Link error control operates at or below the network layer in the OSI model in Figure 1-1. End-to-End error control must necessarily operate at or above the transport layer.

2.4.2 The phases of Error Control

The following phases can be identified in error control:-

- **Error Prevention** Error prevention is an important design exercise. This may be achieved by selecting robust source coding techniques, by resorting to low Bit Error Rate (BER) modulation techniques, and by deploying transmission impairment countermeasures, such as diversity and equalisation.
- **Error Detection** Error detection is of basic importance to a number of error control techniques. A large number of very efficient error detection codes are available to the system designer, and the probability of undetected errors can be made as low as required, often with a relatively small system overhead.
- **Error Recovery** Having detected an error, the next step is to recover from it. This is accomplished by Error Concealment and Error Correction.
- **Error Concealment** When the transacted data is audio/visual in nature it is sometimes preferable to conceal detected errors until these are removed by error correction. Sometimes the real time nature of the audio/visual material rules out complete error control and error concealment terminates the error recovery phase.
- **Error Correction** Two main techniques are available for error correction namely Forward Error Correction and Automatic Repeat Request. In forward error correction enough redundancy is coded into the transmitted data such that corrupted messages can be recovered completely, with a given probability. When a feedback channel is available Automatic Repeat Request (ARQ) techniques can be used instead. The receiver can now, upon detecting errors, request the transmitter to repeat the corrupted data, leading to highly reliable systems.

2.4.3 Forward Error Correction using Block Codes

Two broad classes of FEC techniques exist; Block Codes and Convolutional Codes. *Only block codes have been investigated in this thesis.*

An (n,k,t) block code collects k information symbols and encodes them into n-symbol codewords, with each block coded independent of any other code block. In systematic block codes the k information symbols are stored unchanged in the block code. The additional $(n-k)$ parity symbols introduce enough redundancy in the block code such that at least any t symbols in error can be corrected. This is possible because each codeword differs from any other in at least d positions, where

d is the minimum distance of the code. Then, provided $t = \lfloor (d-1)/2 \rfloor$ (where $\lfloor x \rfloor$ represents the integer part of x), the original transmitted codeword can be determined.

Binary block codes operate with binary symbols. The Bose-Chaudhuri-Hocquenghem (BCH) codes [Lin83] are a powerful class of multiple-error detection and correction codes. These block codes are linear, cyclic codes with a complete algebraic formulation, which allows long and efficient codes to be encoded and decoded in hardware or software with a reasonable complexity. The Reed-Solomon codes [Lin83], are a generalisation of the binary BCH codes to the non-binary case. The RS codes can correct random symbol errors, but since the symbols consist of bit fields, they are particularly suited for correcting bit error bursts.

2.4.4 ARQ

When a feedback path is available between the transmitter and the receiver, ARQ techniques can achieve very high reliability with low overhead when the BER is low, or on bursty channels [Lin84]. Three basic ARQ techniques have been developed namely Stop and Wait (SAW) ARQ, and two continuous ARQ techniques; Go-Back-N (GBN) and Selective Repeat (SR) [Lin83]. In SAW ARQ the transmitter continues retransmitting the same packet until it receives a valid acknowledgement of delivery. In SR ARQ the transmitter transmits data packets continuously and retransmits packets which are in error only. In GBN ARQ the transmitter transmits continuously, but upon detection of packet error, retransmits the packet and all other packets that succeeded the original packet transmission.

2.4.5 Hybrid ARQ/FEC

When the BER is above 10^{-3} on random error channels, the throughput efficiency of the basic ARQ schemes deteriorates rapidly because the probability of packet error is high for practical packet sizes [Lin84]. If a FEC code is used to correct many of the low weight error patterns that occur frequently, in conjunction with an ARQ scheme for uncorrected, but detected, error patterns, the throughput efficiency will remain high in the region of interest i.e. average BER below 10^{-2}.

If the technique is used on burst error channels, such as those encountered on radio based wireless networks, the throughput efficiency will be higher than on a random error channel for the same BER because the error bursts corrupt less packets. Thus the efficiency increases with the burstiness of the channel.

Chapter 3
Video Transmission Over Wireless Networks - A Review

3.1 Introduction

This chapter reviews a number of topics related to the issue of video transmission over wireless networks; video coding standards, video compatible wireless networks, and error control techniques for digital video transmission.

The first section reviews the current state in video coding standards and identifies trends which lead to the selection of an appropriate algorithm for further study. The next section identifies current and planned wireless networks which can support video services. This is followed by an investigation of coded video transmission in the presence of channel errors. Recent advances in error control for packet video transmission, and the development of robust digital video delivery techniques for digital terrestrial television broadcasting are reviewed. Finally, a review of previous work on error control techniques for digital video transmission over wireless networks is presented.

3.2 Video Coding Standards

Currently, there are a four digital video standards for consumer based video services, although others exist for specialist applications [Stenger89]. These are H.120 [CCITT-H.120], H.261 [CCITT-H.261], MPEG-1 [MPEG1] and MPEG-2 [MPEG2]. Two more standards are being defined; H.263 and MPEG-4. H.263 [Sherif94] is targeted at very low bit rate transmission of video over the PSTN whereas MPEG-4 [Reader95] is targeted at very low bit rate, interactive, mobile, multimedia services.

3.2.1 H.120

The CCITT recommendation H.120 [CCITT-H.120] video codec is targeted at videoconference applications at primary rates, that is at 1.544 and 2 Mbit/s. The standard uses 16 level intraframe DPCM to code moving area pixels. Variable length codes are used to code the DPCM levels. Horizontal adaptive sub-sampling is used to reduce the data rate when necessary. This standard has been superseded by the H.261.

3.2.2 H.261

The CCITT recommendation H.261 [CCITT-H.261] codec was ratified in 1990 and is targeted at videophones and videoconferencing codecs operating over fixed bit rate ISDN links at n x 64 kbit/s, with n in the range 1 to 30. The source coder operates on colour CIF and Q-CIF format images (refer to

section 4.4). A block diagram of the coder is shown in Figure 4.2. It consists of a source coder, a video multiplex coder, a transmission coder, a smoothing buffer and a coding controller.

The source coder uses the DCT to code intraframes and hybrid DPCM/DCT to code interframes. Motion compensation prediction is optional. A loop filter can be used to improve compression. The motion estimation operates on the luminance macroblock (16x16 pixels) with integer motion vectors in the range +/-15. The DCT operates on 8x8 pixel blocks. There are 32 uniform quantizers with dead zone, one for the DC component (non-adaptive) in intraframe mode and the other 31 for the rest of the AC components. The coding control maintains a fixed bit rate output and prevents transmission buffer overflow by controlling certain parameters such as frame skipping and quantizer selection. The data and side information are coded using variable length Huffman-like codes. The transmission coder uses a BCH (511,493,2) FEC to protect the data. An H.261 codec emulation is described in chapter 4.

3.2.3 MPEG-1

The MPEG-1 [MPEG1] video codec standard was developed by a joint ISO/IEC committee for multimedia applications and was ratified in 1991. MPEG-1 is optimized for the storage and playback of video using digital media, such as CD-ROMs, at bit rates in the region of 0.9 to 1.5 Mbit/s. The MPEG committee strove to maintain compatibility with the H261 standard, and although MPEG-1 is not a superset of H.261, there is much commonality between the two [Zammit96]. The MPEG-1 block diagram is identical to that in H.261 except that there is no transmission coder in MPEG-1.

The source coder in MPEG-1 is very similar to H.261. It is not quite backward compatible, because it does not include the prediction loop filter of H.261. In addition, the motion estimation in MPEG uses 1/2 pixel accuracy over a larger range, reflecting the more dynamic material targeted by this standard.

MPEG-1 requires periodic intraframes for fast forward and reverse, and random access. These would render the scheme less efficient than H.261 which can use interframe prediction continuously. The 1/2- pixel accuracy motion estimation and a highly efficient bi-directional mode using motion compensated prediction frame interpolation, however, make MPEG-1 slightly more efficient than H.261. In addition, special care has been taken to code intraframes more efficiently than in H.261, by DPCM coding of the DC transform coefficients and lower overhead in the macroblock header. The video multiplex coder in MPEG-1 shares many of the variable-length code tables with H.261, but there are significant differences in the stream syntax which render the two incompatible [Zammit96].

3.2.4 MPEG-2

The MPEG-2 standard [MPEG2] was ratified in 1994. It is a generic standard and is intended for applications of higher quality than MPEG-1, including applications in the fields of broadcasting, consumer electronic products and telecommunications. The MPEG -2 standard is very similar to

MPEG-1 (and hence H.261) being based on motion compensated prediction and the DCT. Like MPEG-1, it supports intraframes, interframes and bi-directionally predicted pictures. MPEG-2 has more features, however. It can code frames or fields in interlaced mode, it offers a choice of zigzag or slanted scanning of the transform coefficients, it supports linear and non-linear quantization tables and multiple resolutions for the DC coefficient in intraframes. MPEG-2 supports scalability which allows a stream to be viewed at various resolutions.

3.2.5 H.263 and MPEG-4

Current standardization activities have targeted low and very low bit rate video transmission. The ITU-T is finalizing the first phase of the H.263 recommendation for video codecs operating at p x 8 kbit/s and suitable for modem use over PSTN [Pereira93, Mickos94]. The first indications are that the H.263 codec is an evolution of the H.261 and MPEG-1 standards, and uses motion-compensated prediction hybrid DPCM/DCT techniques. It uses 1/2 - pixel motion estimation, optionally using 8x8 blocks instead of 16x16, a combined predicted/interpolated two frame mode, and more efficient variable length codes. First results indicate that it achieves 50% to 60% more compression than H.261. The H.263 standard has not been ratified yet, however it is reasonable to expect that the results obtained in this thesis should also be applicable to this closely related standard.

The ISO has started developing the MPEG-4 video codec standard for mobile multimedia services at very low bit rates [Reader95]. The first results are expected by the end of 1996. The very low bit rates targeted by MPEG-4 suggestthat it is time to look at other coding options, such as model based, object based or region based schemes. The results being published by the H.263 proponents have, however, breathed new life into the well known hybrid scheme, even at very low bit rates.

3.2.6 Video codecs for real-time interactive mobile services.

From the above, it is obvious that hybrid DPCM/DCT with motion compensation must be a strong contender for further study, since it is used by both the H.261 and MPEG standards (and the proposed H.263 interim standard). Furthermore, H.261 is targeted at real-time, interactive services. The MPEG codecs usually employ bi-directional prediction which introduces substantial coding delay which may interfere with real time video conversations. An MPEG coder should disable bi-directional prediction, or at least limit its use to single frames (as in H.263), if it is to be employed for real time services.

H.261 is selected for further study in this thesis but, because of the strong similarity between the two algorithms, results obtained for H.261 should also apply to the MPEG codecs (and H.263). H.261 is optimized for n x 64 kbit/s (n=1..30). *Since 64 kbit/s can be considered as a high bit rate on most currently available digital wireless networks, lower bit rates are also studied.*

At least two other projects have studied the problem of video transport over wireless networks, largely in parallel with the work presented in this thesis. The Mobile Audio Visual Terminal (MAVT)

project was under-taken as part of the RACE Mobile initiative [Pereira94, Roser93]. The MAVT team opted to develop a new codec and targeted it at p x 8 kbit/s operation, whereas the approach adopted in this thesis is an investigation of the use of H.261-compatible codecs.

The Portable Video Radio Group (PVRG) at Stanford University has looked at the problem of implementing a portable multimedia terminal to access good quality video material from video-on-demand servers [Meng94]. The video codecs developed by the team are optimized for retrieval, interactive services unlike the conversational, interactive services considered in this thesis. The techniques developed in this thesis can be easily extended to cover retrieval services, especially since much video material is being coded and stored on CD-ROM in MPEG-1 format

3.3 Wireless networks in support of Video Transmission

In the previous chapter (and Appendix 1) a number of second generation, digital wireless networks were introduced. *In this thesis the accent is on terrestrial based wireless networks capable of supporting real time services*, which excludes satellite mobile systems and data and message-oriented networks. The salient specifications of second generation digital mobile and cordless wireless networks, with the potential to support video services, are tabulated in Table 3.1 and Table 3.2.

The main criterion in deciding whether a digital wireless network can potentially support video services, is the raw bit rate supported. The target video bit rates in this thesis are 32 kbit/s and above, which sets the minimum bit rate for the support of video services. The next important network characteristic is the MAC method employed in the network. A network designed to support voice services already has an MAC capable of supporting real-time services. The question that remains is whether the bit error rates and packet error rates targeted for voice services are adequate for video. At face value, the answer is in the negative. The video codec standards available are targeted at very low bit error rates, usually below 10^{-7}, orders of magnitude lower than the target bit rates of second generation voice-oriented networks which have a typical BER of 10^{-3}.

Integrated voice/data wireless networks support data services which have a BER requirement which is much more stringent than that for voice services, and may be compatible with video transport requirements. If the data transport mechanism is supported using the voice MAC, but implemented using error correction techniques, then the data transport mechanism can potentially support video services. One would then have to look at the error control overhead to determine the residual data rate and at the delay introduced by the error control mechanism. If the data transport mechanism is implemented using data- oriented packet-switched techniques, then it is unlikely that the transport mechanism can support video transport without modification.

3.3.1 Second-Generation Digital Mobile Networks

The four main second generation, digital mobile networks considered are GSM, DCS1800, IS-54 and IS-95, as shown in Table 3-1 [Goodman91, Kucar91, Potter92, Katz94, Padgett95, Cox95]. The

first three use TDMA/FDMA multiple access and FDD duplexing, whereas IS-95 is a spread spectrum system and uses FDD/CDMA. A GSM and DCS1800 [Potter92] channel supports 8 time slots each with a raw bit rate of 22.8 kbit/s. These channels can be aggregated to form bearer data channels with data rates up to 182.4 kbit/s uncoded. These channels are circuit switched and can thus potentially support video transport even though, when coded, the rate drops to 91.2 kbit/s.

IS-54 supports 3+3 channels with a raw bearer bit rate of 13 kbit/s available per channel excluding headers [Padgett95]. Thus if the channels were to be aggregated, this would give a total channel capacity of 39 kbit/s. This is not enough to support 32 kbit/s video, voice and the error correction overhead.

The system bandwidth used by IS-95 is 1.228 Mbit/s (chip rate), however each CDMA channel has a raw bit rate of 16 kbit/s in the forward direction and 24 kbit/s in the reverse direction. Thus, unless IS-95 can aggregate channels together, it cannot support video services. Note that TDMA systems can use the same receiver to receive any combination of time multiplexed channels. In the case of CDMA each received channel needs a separate receiver.

SYSTEM	GSM	DCS1800	IS-54	IS-95
Region	Europe/World	Europe/World	US	US
Multiple Access	TDMA/FDMA	TDMA/FDMA	TDMA/FDMA	FDMA/CDMA
Duplexing	FDD	FDD	FDD	FDD
Frequency Band (MHz)	935-960	1710-1785	869-894	869-894
	890-915	1805-1880	824-849	824-849
Carrier Spacing (kHz)	200	200	30	1250
Bearer channels/carrier	8	8	3	
Channel bit rate kbit/s	270.833	270.833	48.6	1228.8 (Chip Rate)
Modulation	GMSK	GMSK	$\pi/4$ QDPSK	BPSK/QPSK
Speech Coding (kbit/s)	13	13	8	1-8
Frame Duration (ms)	4.615	4.615	40	20
Error Control	Equalizer	Equalizer	Equalizer	Rate 1/2 fwd.
	Rate -1/2	Rate -1/2	Rate 5/8	Rate 1/3 rev.
	Convolutional	Convolutional	Convolutional	Convolutional

Table 3-1 Second Generation Digital Cellular Systems

Of the four systems reviewed, only GSM and DCS1800 appear to have the throughput required to support videophone services at or above 32 kbit/s. Furthermore, these systems are targeted at macrocellular operation so that at 270 kbit/s the delay spread is still high enough to cause ISI. Therefore equalizers have to be used. In addition to this, the systems operate in a strong CCI environment and a powerful half rate convolutional code is necessary to render the voice channel useful.

3.3.2 Second-Generation Digital Cordless Networks

The second generation digital cordless systems of interest are CT2, DECT, PHS, and PACS as shown in Table 3-2 [Goodman91, Kucar91, Katz94, Padgett95, Cox95]. The four systems use TDMA/FDMA multiple access. PACS uses FDD duplex whereas the other three use TDD. The systems are designed to transport ADPCM coded voice at 32 kbit/s without forward error correction or equalization. CT/2 can only support 32 kbit/s full-duplex and thus it cannot support the 32 kbit/s video, voice and additional error correction overhead.

DECT supports twelve 32 kbit/s full duplex channels using 24 paired timeslots. The time slots can be aggregated together to support 384 kbit/s full duplex. Asymmetrical aggregates are also supported so that 768 kbit/s is possible with half duplex. DECT was designed to support integrated services including video as an ISDN application [Owen90]. The DECT signal is rather wideband and the average ISI-induced BER is on the order of 10^{-3} for a 40 ns delay spread [Schultes92]. With antenna diversity DECT tolerates around 18 ns delay spread [Schultes92], but it cannot cope with much higher delay spreads, *e.g.* the 300 ns delay spread found in outdoor environments for which DECT is targeted.

The PHS and PACS systems are quite similar. Both use TDMA/FDMA multiple access and have a channel bit rate of 384 kbit/s. PACS uses FDD and $\pi/4$ QPSK whereas PHS uses TDD and $\pi/4$ QDPSK. PACS can support 256 kbit/s full duplex whereas PHS can reach 128 kbit/s full duplex, by aggregating timeslots. These rates are well capable of supporting videophone services. The 300 kHz channel bandwidth render the systems non-dispersive in many environments and only CRC is used for error control for voice. Whether this is enough for video services needs to be established.

SYSTEM	CT2	DECT	PHS	PACS
Region	Europe	Europe	Japan	U.S.
Multiple Access	TDMA/FDMA	TDMA/FDMA	TDMA/FDMA	TDMA/FDMA
Duplexing	TDD	TDD	TDD	FDD
Frequency Band (MHz)	864-868	1880-1900	1895-1918	1850-1910/ 1930-1990
Carrier Spacing (kHz)	100	1728	300	300/300
Bearer channels/carrier	1	12	4	8/pair
Channel bit rate kbit/s	72	1152	384	384
Modulation	GMSK	GMSK	$\pi/4$ QDPSK	$\pi/4$ QPSK
Speech Coding (kbit/s)	32 kbit/s	32 kbit/s	32 kbit/s	32 kbit/s
Frame Duration (ms)	2	10	5	2.5
Error Control	None	CRC	CRC	CRC

Table 3-2 Second Generation Cordless Systems

Three of the four second generation cordless systems studied can support video services. PACS and PHS appear to require the least complex error control due to the substantially lower bandwidth occupied than DECT. DECT however can support higher data rates but requires antenna diversity even in low delay spread environments.

3.3.3 Wireless Local Area Networks

Currently, the wireless local area network offerings are mainly from the US and operate in the unlicensed ISM bands. The permitted transmit power for non-CDMA use in this band is very limited, so that most of the offerings use CDMA. Wireless LANs certainly have the bandwidth to support wireless video services. Most of the systems offered, however, are targeted at ethernet access and seem to implement predominantly data-oriented multiple access strategies.

Currently, Wireless LANs are being standardized by ETSI in Europe and IEEE 802.11 in the US. The target bit rates are well capable of supporting video services and, in fact, both standards bodies have identified video transport as an important WLAN service and both sets of standards will support isochronous traffic. Literature about these standards appeared too late to influence the work in this thesis.

3.3.4 Integrated Services Wireless Transport Systems

A number of researchers have studied the impact of integrated services on third-generation wireless transport systems. Goodman [Goodman90] proposed a single network architecture unifying Broadband ISDN networks with third-generation mobile networks. The central proposal is the Cellular Packet Switch (CPS), a cellular mobile network implemented by interconnecting base stations over an optical fiber metropolitan area network. The CPS supports integrated services using Packet Reservation Multiple Access (PRMA) [Goodman89]

Mermelstein *et. al.* [Mermelstein93] have studied the impact of Integrated Services on TDMA and CDMA multiple access schemes. The required performance for the various services was stated in terms of the maximum BER and the sensitivity to delay. Low resolution video (64, 128 kbit/sec) is listed as delay sensitive which can tolerate a BER of up to 10^{-5}. They conclude that in TDMA systems, time slots must be allocated dynamically to achieve statistical multiplexing gain.

Raychaudhuri and Wilson [Raychaudhuri93, Raychaudhuri94] have studied multimedia transport in next generation Personal Communication Networks (PCN). They propose a Multiservice Dynamic Reservation (MDR) TDMA based multiple access protocol for wireless access to ATM-based B-ISDN networks, capable of supporting both constant bit rate and variable bit rate traffic. The MDR TDMA frame is divided into request slots accessible in contention mode and message slots which comprise fixed allocation slots for CBR traffic and dynamic allocation slots for VBR traffic and packet switched data.

3.3.5 Third-Generation Wireless Networks

Third-Generation mobile and personal communications systems such as the global FPLMTS and the European UMTS are being jointly standardized by ITU-T and ETSI. Both of these systems recognize video as an integral part of the multimedia traffic which they are being designed to support. The RACE Mobile program is focusing on the Universal Mobile Telecommunications System (UMTS). UMTS targets ISDN rates (144 kbit/s) with the possibility of higher rate services up to 2 Mbit/s. FPLMTS has similar targets. Thus both systems will support video services.

A number of projects within the UMTS initiative are tackling issues which are of interest. Besides MAVT described above, the ATDMA [Dunlop95] and the CODIT [Baier94, Grillo95] projects are developing third generation MAC schemes based on PRMA and CDMA.

3.4 Error Control For Video Transmission

This section investigates the effects of channel errors on Image Sequence Coding (ISC) and reviews error control techniques developed to combat their effects. The inherent robustness of ISC techniques is investigated as well as means of improving this robustness. Although FEC can be deployed to combat channel errors, it is important to study the behaviour of ISC techniques on high BER channels and on burst error channels when the residual BER is still high, even after FEC. Intraframe techniques are investigated first, followed by interframe coding techniques and hybrid DPCM/DCT algorithms.

3.4.1 Intraframe Error Control

Intraframe coding techniques code a frame independently of the previous frames. Thus intracoded frames do not suffer from temporal error propagation from previous frames. Channel errors cause errors in the current frame, and the high compression ratios used leave the system vulnerable to spatial error propagation.

3.4.1.1 Error control for DPCM coded images

In DPCM codecs a channel error corrupts the current pixel and becomes embedded in the predictor from where it corrupts subsequent pixels. If a two-dimensional predictor is used the errors can potentially propagate to the rest of the frame, in a pattern depending on the prediction equation [Netravali88]. Some predictors may not be stable in the presence of channel errors. With one dimensional, unity-gain predictors, errors propagate to the end of the current line. It is usual for the first pixel on each line to be coded non-differentially, so that the first pixel terminates the error condition. Aguello *et al.* [Aguello71] proposed inserting periodic PCM updates in DPCM coded streams, to halt error propagation in 4 bit DPCM ISC. Good results were reported at a BER of 10^{-3} with periodic updates every 26 pixel block. Steele *et al.* [Steele79] extended the idea to allow error detection and

correction of the corrupted blocks in DPCM coded speech. Ngan and Steele [Ngan82] then extended the idea to block DPCM coded video with error propagation limited to square MxM pixel blocks.

Connor [Connor73] studied techniques for reducing the visibility of transmission errors in digitally encoded video signals. Three basic techniques were introduced. The first was the selection of a coding strategy which was inherently robust. Secondly, the errors were concealed by using line averaging and replacement. Thirdly, a leaky predictor was used to recover from DPCM errors as shown in Figure 3-1.

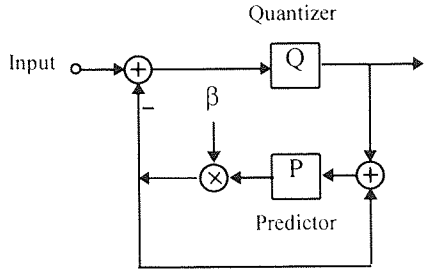


Figure 3-1 Leaky DPCM predictor

These three techniques, robust coding, error concealment and error recovery, are basic error control strategies employed in most error control techniques to be described.

3.4.1.2 Error control for DCT-coded images

The Discrete Cosine Transform is used extensively in image coding applications. The transform itself is very robust to errors [Netravali88]. Channel errors are usually diffused over the whole transform block and may remain invisible, especially when these affect high spatial frequency coefficients, since the human vision system tends to mask such errors [Netravali88]. Large errors or errors which affect low spatial frequency coefficients are usually quite evident as significantly altered pixel blocks. This assumes that only a few transform coefficients are affected, *i.e.* code synchronization is preserved.

Brewster and Dodgson [Brewster91] studied Zonal and Threshold coding of the DCT coefficients. Zonal coding uses a fixed bit assignment table to code the transform coefficients and was found to be robust to errors. In threshold coding all coefficients above a certain threshold have to be coded. The particular coefficient pattern is not predeterminable and has to be coded as side information. Thus both coefficients and side information may become corrupt. If the latter is corrupted then synchronization is lost in the current block and in subsequent blocks unless loss of synchronization can be detected and re-synchronization codewords are inserted in the stream.

Mitchell and Tabatai [Mitchell81] proposed an error concealment strategy for an adaptive Zonal DCT coder with classification side information. Their technique detects errors by searching for cross DCT block discontinuities. Concealment is then achieved by locating the transform coefficient causing the discontinuity and subtracting its basis function from the block.

3.4.1.3 *The effect of channel errors on Variable length Codes*

Variable length coding of quantized levels introduces a potential for loss of synchronization. Variable length codes lose synchronization in the presence of channels, and re-synchronization periods are difficult to predict. Maxted and Robinson [Maxted85] modeled variable length codes using a Markov model and used this to determine the synchronization properties of the code. The codeword probabilities are required to obtain meaningful results. Ferguson and Rabinowitz [Ferguson84] proposed a technique for designing self-synchronizing optimal Huffman codes. The code is guaranteed to synchronize by the inclusion of a synchronizing codeword. When this code appears in an un-synchronized situation the stream will re-synchronize starting from the next codeword. Yin and Yu [Yin93] replaced the H.261 transform coefficient VLC table with one using a non-optimal synchronizable variable length code (similar to an S code [Pratt78]). They reported very good results at BER of 10^{-4} on a random error channel.

Some systems avoid VLC synchronization problems by using VQ followed by fixed-length codewords instead of SQ followed by VLC codewords, since VQ is very resilient to errors and the fixed-length code cannot lose synchronization. Zager and Gersho [Zager90] improved the inherent resilience of VQ by using Pseudo-Gray Encoding of code vectors, such that codewords close in Hamming space represent codevectors close in Euclidean space. Then single bit errors do not result in large reconstruction errors. Farvardin [Farvardin90] also studied the optimal assignment of codewords to code vectors, and introduced hierarchical bit-assignment such that some of the codeword bits are more significant than the rest and allow unequal error protection.

3.4.1.4 *Joint Source-Channel techniques*

Shannon [Shannon64] demonstrated that the source and channel coding functions are fundamentally separable. According to [Farvardin87] Viterbi and Omura clearly indicate that this assumption is justifiable only in the limit of arbitrarily complex encoders and decoders. Thus, Farvardin and Vaishampayan [Farvardin87] studied the interrelationship between the source and channel coders for practical systems with limited complexity, and proposed an algorithm for the joint optimization of the source and channel encoder-decoder pair. This indicates that joint source-channel coding may be superiour to separate source-channel optimization, under the constraint of limited hardware complexity.

Modestino and Daut [Modestino79] applied joint source-channel coding to 2D-DPCM coded images and obtained the best results by reducing the source coding rate and using the extra bits for channel coding. Modestino, Daut and Vickers [Modestino81] later studied joint source-channel coding of DCT images by applying channel coding unequally to the transform coefficients. Unequal error protection as demonstrated by Modestino *et al.* [Modestino81] has now become an important technique for combatting channel errors in digital video transmission [Viasey92, Stedman92].

3.4.2 Interframe Error Control

In intraframe coding, errors are confined to one frame. In interframe coding, errors can propagate temporally between frames and may lead to permanently corrupted areas in the sequence.

3.4.2.1 Conditional replenishment and channel errors

Conditional replenishment (CR) [Mounts69] serves to highlight important aspects of interframe error propagation. Conditional replenishment generates two types of data - replenishment data and data addresses. Both can become corrupted, leading to permanently corrupted pixels in the sequence until the corrupted pixels are replenished by correct data. When conditional replenishment is applied using pixel clusters or blocks [Candy71], address errors lead to even more visible errors as a whole block may be deposited in an erroneous section of the frame. Thus CR causes temporal error propagation and the cluster addressing causes spatial error propagation. Differential encoding of the pixel block address leads to even larger areas being corrupted. Thus it is important to protect address information and to limit the extent of differential encoding of this sensitive data.

3.4.2.2 The effect of channel errors on Interframe DPCM

In interframe DPCM the difference between the current and previous frame pixel is coded and transmitted. If channel errors corrupt the differential data it will become embedded in the reference frame and will propagate temporally to other frames until it is refreshed in *non-differential* mode. When used with conditional replenishment and block addressing, the errors may propagate spatially in future frames.

3.4.2.3 Motion Compensated Prediction and Channel Errors

Motion compensated prediction introduces two error propagation mechanisms which cause temporal and spatial error propagation. Firstly of all, channel errors may corrupt motion vectors, leading to corruption in the current frame which then propagates temporally since prediction is being used. Secondly, the corrupted blocks may be subsequently referenced by otherwise uncorrupted blocks leading to spatial and temporal error propagation. MCP-induced errors are not as visible as block address errors, since the former consists of prediction data whereas the latter may consist of intraframe coded replenishment data.

3.4.2.4 Interframe Error control Techniques

The intraframe error control techniques introduced above can be adapted to counter temporal error propagation in interframe image sequence coding schemes. Thus leaky prediction can be used to limit temporal error propagation [Ghanbari93]. Periodically refreshing frames in intramode, similar to the periodic PCM coded samples in DPCM, can be used to clear the reference memory of any previous errors [Horst93]. Concealment can also be used to limit the visibility of errored blocks until the

corrupted data is force updated.

To summarize, channel errors may corrupt large areas within one frame due to spatial error propagation and these errors can then propagate temporally to other frames in the image sequence. A reasonable error control strategy requires:-

- Selection and Implementation of an inherently robust source coding scheme with spatial re-synchronization mechanisms and spatial error containment designed in.
- The ability to detect errors and correct them if possible
- The ability to conceal errors
- The ability to recover from temporal error propagation e.g. using forced updating or leaky prediction.

3.4.2.5 Error control and H.261

The H.261 codec is a complex assembly of conditional replenishment, motion compensation prediction, DCT, and variable length coding, and hence inherits the error characteristics expounded above. The H.261 design incorporates a number of the error control measures introduced above. It has 12 re-synchronization points per 352x288 pixel CIF (Common Intermediate Frame [CCITT-H261]) frame (3 in a 176x144 Quarter-CIF frame), which allow recovery from VLC synchronization loss and hence limit spatial error propagation. Error detection and correction is catered for by a BCH (511,493,2) code.

Temporal error propagation is not catered for explicitly in H.261. There is no leaky prediction but a forced updating mechanism introduced to clear accumulated arithmetic errors in DCT computation, can be used to achieve periodic forced updating. Another possibility is to use the external fast update request signal, included to facilitate multipoint, multiparty videoconferences, but which can be used by a decoder to request the encoder to code the next frame in intramode, following a detected channel error.

H.261 has been optimized for ISDN circuits, however, where BERs are typically below 10^{-7}. On wireless networks the target error rates *after error correction* are sometimes as high as 10^{-3}. It is thus not clear whether the H.261 error control mechanisms can cope with such high error rates. *This thesis investigates the behaviour of hybrid DPCM/DCT algorithms, based on H.261, on high bit error rate wireless networks and investigates how to render the decoded sequences useful and acceptable.*

Recently, numerous investigations have been conducted on the behaviour of image sequence coding algorithms under high BER conditions in two fields: Packet Video transmission and Digital Terrestrial Television Broadcasting. Contributions from these two fields are reviewed in the next two sections.

3.5 Packet Video

Packet Video [Verbiest88, Karlsson89, Pearson90] refers to the transmission of digitized video information over packet switched telecommunication networks such as computer LANs and MANs. Recently, the Asynchronous Transfer Mode (ATM) has been chosen as the multiplexing, transmission and switching technology for B-ISDN which will deliver integrated multimedia services to the home over optical fiber at 155 Mbit/s.

This technology is attractive because it can integrate disparate multimedia services on the one network and supports variable bit rate services. It can capitalize on the variable bit rate nature of the traffic to realize a statistical multiplexing gain, leading to more efficient use of the telecommunications infrastructure. This is achieved by multiplexing VBR users based on their average bit rate requirements, such that their total peak bit rate is substantially greater than the channel capacity. This leads to a significant probability of congestion, however, during which the multiplexing node discards excess cells.

The variable bit rate nature of the cell based ATM transport is attractive to video transmission because video codecs inherently generate highly variable bit streams when coding video with fixed video quality. Hence the variable bit rate transport mechanism can be exploited to provide constant quality video and at the same time allow more users to be multiplexed for a given channel capacity. The complexity of the video codec is reduced because there is no need for the smoothing buffer and the complex bit rate control logic needed on fixed bit rate networks. In addition, the removal of the smoothing buffer reduces the video coding delay substantially, usually by around 100 ms.

There are two main problems in transporting VBR video over ATM; packet delay jitter and the significant cell loss probability. Delay jitter is not a major problem and can be handled by introducing a small buffer at the receiver to absorb the jitter. The size of this buffer is much smaller than the smoothing buffer found in FBR codecs. Cell loss is a more difficult problem to solve. The loss of an ATM cell means a loss of 48 data octects (384 bits) and this is catastrophic for H.261 and MPEG coders designed to operate at a BER below 10^{-7}.

Although the issue of whether to opt for VBR video transport is still unresolved, a large number of highly robust video codecs have been proposed in the packet video literature. The next section attempts to classify the *error control techniques* found in these proposals, prior to an in-depth review of the major contributions.

3.5.1 Error control techniques for packet video

Nomura *et al.* [Nomura91] defined two classes of cell loss: random cell loss and congestion cell loss. The first loss mechanism is due to random errors corrupting cell headers. The probability of this is very small because the random BER component is below 10^{-9} and the headers are protected by an FEC code. Congestion related cell loss is due to cell discard by congested access nodes and cells

discarded at the decoder because of excessive time delay [Kishimoto91]. The ATM video transmission system is modeled as in Figure 3-2 and error control techniques for the two classes of cell loss are categorized by the level at which they operate in this model *i.e.* THL, ANL, PL and CL in Figure 3-2.

Congestion related error control techniques are specified at the Access Node Level and at the Codec Level. At the Access Node Level, three error control techniques can be considered (a) bandwidth re-allocation so as to avoid cell loss (b) encoder bit rate control by feedback from the access node to the encoder preventing it from generating too much data and (c) selective packet discarding which, in conjunction with a layered coder, discards only cells of a certain priority. At Codec Level error control is achieved by using a layered coder. The first two techniques can be classified as preventive methods whereas the third and fourth describe layered coding.

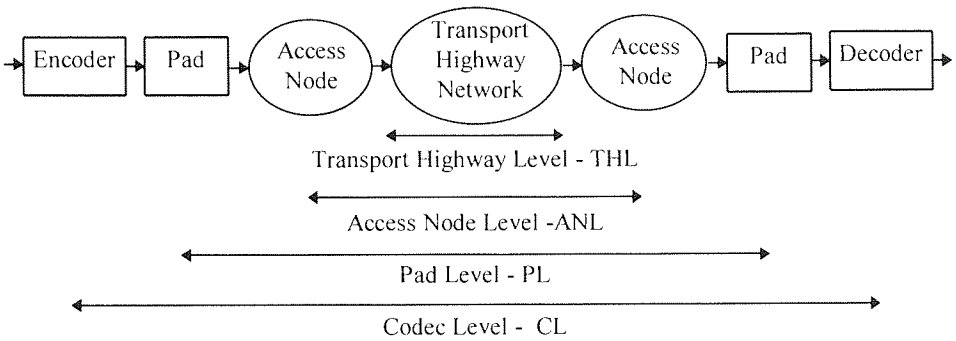


Figure 3-2 ATM based video transmission system

Karlsson and Vetterli [Karlsson89] recognize two approaches to error recovery:- (a) Error Control Coding, which can recovery completely from errors and (b) Error Concealment, which relies on visual redundancy. They identify two codec requirements for successful error concealment: limited error propagation and error location detection.

Yasuda *et al.* identify three classes of protection/recovery methods (a) Source Coding such as the Layered coding method, (b) Channel Coding such as re-transmission or FEC and (c) Joint Source-Channel coding such as the re-transmission of priority packets.

Morrison and Beaumont [Morrison91] identify five cell loss compensation techniques. The first two are preventive: Zero cell loss through low network loading and Zero cell loss through source characterization and policing functions. The other three techniques are (a) Interleaved FEC (b) Error Concealment and (c) Two-Layer Coding.

The identification of two-layer coding (or joint source-channel coding) as an error control class implicitly defines another class of non-layered techniques. The packet video coding and error-control techniques reviewed below are thus divided into non-layered and layered techniques.

3.5.2 Cell Loss Compensation: Non-Layered Techniques

Forward Error Correction techniques have been proposed for coded video transmission in ATM networks. Kishimoto and Irie [Kishimoto91] proposed a video coding algorithm for HDTV

transmission on an ATM network. The proposed algorithm uses sub-band coding to create four sub-sampled bands. The low frequency band is coded using a DCT which adaptively codes data from the motion compensated, intra-field or inter-field predicted frames. The horizontal low-pass and vertically high-pass filtered band is coded using horizontal DPCM. The other two bands are coarsely quantized. All four bands are variable length coded. A cell loss compensation method based on random error correction codes and interleave structures is then applied to the multiplexed signal. The continuous video stream is divided into x byte packets to which a header is added, comprising a destination address and a packet sequence number. Then the FEC parity bits are computed and added to the packet to generate 40 byte cells. Twenty such cells are byte interleaved and transmitted. The receiver can detect cell losses because the ATM network numbers each cell. Cell losses are treated as erasures therefore and after de-interleaving the FEC code is used to recover the erasures. Using an RS (40,36,2) code with 90% efficiency, the post-decoding cell loss probability is less than 10^{-5} for a cell loss probability of 10^{-3} before compensation. The interleaver introduces a delay of around 40 ms at 384 kbit/s, which increases to around 240 ms at 64 kbit/s.

Harasaki and Yano [Harasaki94] also considered interleaved forward error correction for variable bit rate video coding. They pointed out that the interleave delay in a VBR environment is not constant because the interleave may have to wait for a cell to complete the interleave. Thus, cells may be unduly delayed and would have to be discarded. A time-out period was thus introduced to bound the interleave period which decreases the efficiency slightly by forcing interleave transmissions with empty cells.

Error control by retransmission is not a favored approach in ATM transmission mainly because the round-trip delay is large on wide area B-ISDN networks. Therefore the re-transmitted data would cause excessive delay on the link. More important still, cell loss is usually caused by congestion on ATM networks. Then, attempts to retransmit data with small delay, is bound to increase the congestion and lead to more lost cells. However feedback can be used to demand refresh, that is if cell loss is detected the decoder will notify the encoder which can transmit the next frame in intramode. Again this is bound to increase congestion since intraframes generate more than three times the data of interframes.

Wada [Wada87] proposed an elegant cell loss recovery method based on selective demand refresh for variable bit rate video which side-steps these problems. When the decoder detects an error it tries to conceal the error, transmits the identity of the lost cell to the coder, and continues decoding. When the coder receives the lost cell ID it performs a calculation based on the lost cell frame and the last frame transmitted, to determine how many corrupted blocks have resulted at the decoder. It then codes these blocks using intraframe coding. This scheme works even in the presence of a long propagation delay. It has a higher efficiency than demand refresh, no additional information is needed

for error recovery, the method is applicable to conventional encoding algorithms and it is suitable for multipoint transmission.

Zhang et al. [Zhang93] proposed a cell loss compensation technique based on the immediate feedback of congestion and cell loss information from the local access node to an ATM network. The encoder then takes steps to refresh the decoder over a number of frames to minimize network traffic overload. This scheme capitalizes on the property of ATM, that once a cell is accepted at the local access node the probability of cell loss is very small.

Horst and Hoeksema [Horst93] studied six, one layer error control techniques for H.261 transmission over ATM. They found that Implicit Refresh, using the forced updating of H.261 macroblocks to clear DCT computation errors, to be insufficient since the refresh period is long - a macroblock is refreshed every 132 times it is transmitted. Temporal error propagation recovery by transmitting Periodic Intraframes was found to result in excessive overhead. Demand refresh using the fast update request mechanism in the H.261 standard was found to have a fast response time with very low overhead, provided the mean time between cell loss was greater than 10 seconds. Attempts to conceal errors at the decoder were found to have limited success. The use of FEC codes with interleaving was also investigated, and revealed that a large interleave buffer was required leading to substantial excess delay. The best way to transmit one-layer H.261 was found to decrease congestion by allowing a slightly smaller number of video users to access the link simultaneously.

Hamano *et al.* [Hamano93] proposed a robust cell-loss compensation scheme for an MPEG-2 like video coding scheme over ATM networks. The error control strategy involved structured packing to limit the spatial error propagation caused by cell loss, concealment at the decoder by macroblock replacement and temporal error recovery using a leaky interframe predictor.

Structured packing is used frequently in packet video to recover from cell loss. The continuous video stream is divided into cells and if the data therein contains a synchronization entry point, *e.g.* a picture, group-of-blocks or macroblock header in H.261 terms, then the bit number of the first bit of the VLC code is included in a pointer field in the structured packet header. The absolute address of the synchronization point within the frame and the frame number are also included in this header. If any data following the synchronization point is coded in differential format, than a absolute reference must be inserted in the structured packet header.

Finally, the transmission of intraframe coded sequences has been discussed in the literature. Chia *et al.* [Chia93] studied the transmission of Motion JPEG intraframes over an ATM network. The scheme is made more robust to cell loss by the inclusion of 'marker codes' in the JPEG specification, which allow re-synchronization of the Huffman variable length code. Error concealment was then studied to hide corrupted areas until the reception of the following frame. Intraframe schemes are very effective because errors are limited to one frame and are usually invisible. However the bit rate overhead penalty is usually above 300%.

3.5.3 Cell Loss Compensation: Layered Coding

Layered coding [Verbiest88, Ghanbari89, Kishimoto89, Kishino89] is a coding technique which is considered to be very attractive for packet transmission. The technique is based on unequal error protection and splits the image to be coded into two or more layers with unequal importance. The first layer conveys the important structural and synchronization data, whereas the other layers convey the detail of the image and serve to improve quality. For the technique to work a priority scheme must be implemented on the packet network, in which several levels of priority are assigned to the packets. A number of techniques have been used to generate two-layer schemes. The majority reviewed here are based on some hybrid of DPCM with DCT including modified H.261 and MPEG codec.

M. Ghanbari [Ghanbari89] proposed a two layer coding scheme which is shown in Figure 3-3. The input video is split into two streams, one called the guaranteed stream, the other the enhancement stream. It is essential for the guaranteed layer data to arrive at the destination, and hence it is assigned the highest priority. Data from the enhancement layer is less important and can be assigned a lower priority. The network treats higher priority packets preferentially. If traffic congestion leads to queuing buffer overflows, the low priority packets will be discarded first.

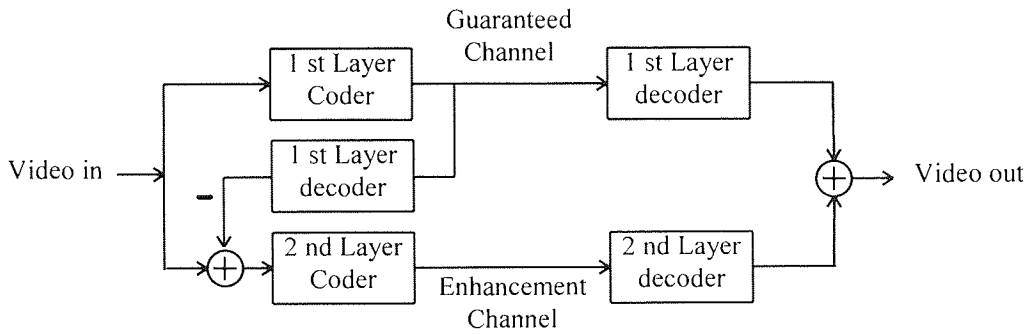


Figure 3-3 Two-layer video coding for ATM networks

The scheme proposed by Ghanbari uses an H.261 compatible coder in the first layer and an intraframe DPCM encoder in the second. Since the second layer cells may be lost, then the non-layered techniques identified above have to be used to compensate for the lost cells. In this scheme Ghanbari uses intraframe DPCM to code the difference between the original and the reconstructed frame using 8x8 blocks. The first pixel in each block is PCM coded so that each block is a self contained unit and can be transmitted independently. Then the error due to lost packets is limited to one block and does not cause loss of synchronization. Thus the block structure limits spatial error propagation and intraframe coding limits the temporal error propagation in the vulnerable second layer.

In [Ghanbari92] Ghanbari replaced the second layer DPCM coder with a more efficient DCT coder The DCT coding is not performed explicitly, but implicitly by requantizing the difference between the DCT output and the coarsely quantized output of the H.261 coder in the first layer, using a finer quantizer to generate the second layer data. In [Ghanbari93] Ghanbari and Seferidis generated the second layer by using an interframe DPCM coder with leaky prediction to limit temporal error

propagation. In a second proposal, they used a second H.261 coder to code the second layer data and found that the implicit macroblock refresh of the second layer H.261 coder was enough to recover from errors.

Morrison and Beaumont [Morrison91] have also proposed a two-layer coder. They use an H.261 coder in the base layer with the smoothing buffer disabled so that the coder generates data with fixed quality. The second layer is formed as in [Ghanbari92] by requantization. Cell loss compensation in the second layer is achieved by structured packing.

Kishino *et al.* [Kishino89] proposed a layered coding scheme which is robust to cell loss. A hybrid DPCM\DCT scheme was used to code the input video. The layering was executed in the transform domain by splitting the transform coefficients into a most significant part (MSP) composed of the DC and lower spatial frequency components, and the least significant part (LSP) which was composed of the remainder of the coefficients. MSPs take priority over LSPs when congestion occurs. The method has been shown to be effective for 20% LSP lost cells.

Nomura *et al.* [Nomura91] implemented a two layer coder using a DCT followed by interframe DPCM coding of the transform coefficients with a leaky predictor. The quantized DPCM coefficients are then split to generate the low and high priority streams. In this scheme the leaky prediction protects both streams against temporal error propagation. The two streams are protected against spatial error propagation using structured packing.

Tubaro [Tubaro91] proposed a two layer scheme based on a H.261 coder in the base layer. The second layer is divided into 8x8 pixel blocks which are coded in intraframe DPCM mode to prevent temporal error propagation. The first pixel in each DPCM coded block is PCM coded to prevent spatial error propagation. The block elements are transmitted using an element and block interlace pattern which attempts to break up the effects of cell loss over the whole frame. Error concealment of the lost blocks can then be undertaken using interpolation. Alternatively, the second layer can be coded using an 8x8 DCT using the same interlace technique on the transformed coefficients.

A number of proposals have been made to adapt MPEG encoders to ATM transmission. In [Raychaudhuri94] Raychaudhuri describes the MPEG++ ATM transport scheme which divides an MPEG stream into two layers and then uses extensive structured packing to protect both the high priority and the low priority streams.

Sub-band coding schemes have been found to perform well in VBR applications. Karlsson and Vetterli [Karlsson88] have proposed a 3D SBC scheme for packet video. Two frames are coded simultaneously by filtering the image in the three dimensions to produce 11 down-sampled sub-bands. The lowest band retains a large degree of variance and is DPCM coded whereas the other bands are PCM coded with coarse quantization. The scheme is well suited for layered coding in which case the lowest band has the highest priority and the PCM-coded second layer is inherently resilient to cell loss.

3.5.4 Subjective Testing

The performance of two layer coding schemes has been tested subjectively. Yasuda, Ohta and Kishino [Yasuda89] performed subjective tests on four image sequences subjected to random packet loss (equivalent to packet loss on one layer codecs) and to selective packet loss (equivalent to packet loss on layered codecs). The results show that the random loss schemes perform acceptably up to about 0.1% packet loss, whereas a selective packet loss of 10% was just perceptible. Another test showed that a variable bit rate coder can achieve the same visual quality as a fixed bit rate codec, using one half the bit rate on average.

Seferidis, Ghanbari and Pearson [Seferidis93] have conducted subjective tests based on the CCIR 5-grade impairment scale to study the effects of cell loss rate on subjective video quality. The results show that cell loss rates between 10^{-3} and 10^{-2} return good Mean Opinion Scores (MOS) close to 4. The authors show that the MOS increases as the corrupted sequence duration is decreased at the same cell loss rate. They refer to this phenomenon as the *forgiveness factor*.

Raychaudhuri [Raychaudhuri94] subjected the MPEG++ layered video transport to cell loss rates between 10^{-5} and 10^{-1} and reports that acceptable video quality is achievable at a cell loss rate of 10^{-2}. These results show that two-layer codecs can deliver good quality video when the cell loss rate in the enhancement stream is 1% and even,possibly, as high as 10%

3.6 Developments in Digital Terrestrial TV

Digital Terrestrial Television Broadcasting (DTTB) has recently progressed rapidly and various prototypes have been tested. It has been shown that a High Definition TV (HDTV) signal can be digitally transmitted in the existing TV bands. DTTB transmission must take place in "taboo" channels which suffer from strong co-channel interference from adjacent analog TV transmitters, and must coexist without causing interference.

These DTTB systems are of some relevance to the work in this thesis because they constitute digital video transmission over a radio channel. However it must be pointed out that DTTB is quite different from the interactive video services of interest in this thesis. First of all DTTB are targeted at mainly non-mobile users. Thus it is assumed that the TV sets can be bulky and make use of the mains supply. Thus there is not such much concern about minimizing complexity and power consumption. Secondly, The broadcast nature of the service allows a substantial channel coding delay due to the non interactive nature of the service.

DTTB systems are being standardized in EUROPE, JAPAN and the USA. The systems appear to be converging in one respect,; the use of some MPEG derivative for the video coding part either a derivative of MPEG-1 such as MPEG++ [Siracusa93, Hulyalkar93] or MPEG-2 [Anastassiou94]. However several other systems have been proposed [Ramchandran93, Argenti93]. The common denominator is some form of layered coding or multiresolution approach, which allows scalability of

service and graceful degradation such that users far from the transmitter still receive a useful image albeit at a lower resolution.

In addition to layered coding, the proposed algorithms attempt to protect the streams using techniques such as synchronization point insertion [Argenti93], concealment, and leaky prediction [Ramchandran93]. MPEG++ transport [Siracusa93, Raychaudhuri94] has already been described above. It makes extensive use of structured packing to protect both high priority and normal priority signals, and then conceals lost blocks by some form of reconstruction.

Europe and the USA have taken different routes with respect to the radio transport approach. Europe is trying to implement DTTB using Coded Orthogonal Frequency Division Multiplexing (COFDM), whereas the USA has concentrated on Pulse Amplitude Modulation (PAM) in Vestigial Side Band (VSB) and Quadrature Amplitude Modulation (QAM) techniques, although recently the preferred solution seems to be 8-PAM-VSB. COFDM has already been successfully applied to Digital Audio Broadcasting in Europe. It has the significant advantage of being inherently robust to frequency selective fading and research is being undertaken to extend the technique to DTTB [Helard91, Anastassiou94].

Besides the modulation techniques, the error control mechanisms employed are also of interest. Here the unconstrained power budget at both ends of the broadcast chain means that a plethora of error control schemes have been proposed.

Hulyalkar et al. [Hulyalkar93] propose 32-QAM transmission using interleaved RS codes, concatenated with a convolutional coder. The receiver uses a combined equalizer carrier recovery unit and a soft decision convolutional decoder. The CNR threshold for the Low Priority signal is 16.1 dB and 11.1 dB for the High Priority signal. Ramchandran and Ortega [Ramchandran93] propose various multi-resolution, embedded QAM schemes with Trellis Coded Modulation (TCM) of the fine (Low Priority) channel. Helard and Floch [Helard91] proposed a COFDM system in combination with rate 2/3 8-PSK TCM.

This brief overview of DTTB research indicates that the techniques adopted in this area may not be compatible with a low-power, hand-held, interactive, mobile videophones.

3.7 Video Transmission Over Wireless Networks.

This section reviews proposed systems for interactive video transmission over wireless networks. The material reviewed here is directly related to the work in this thesis, and a number of the papers, especially those on unequal error protection and ARQ (as well as layered coding techniques reviewed above), inspired the main original contributions in this thesis.

3.7.1 Unequal Error Protection

A number of authors proposed systems employing unequal error protection. The technique operates by partitioning the video stream into two or more bit field classes, and using bit significance

testing to order them by vulnerability to transmission errors. Then the more vulnerable fields are given better protection by using FEC codes of varying rate.

Vaisey, Yuen and Cavers [Vaisey92] proposed a 64 kbit/s video transmission system based on a modified QCIF (Quarter CIF 176x144 format- refer to section 4.4) H.261 codec. The target frame rate was approximately 4 fps (frames per second). Having identified catastrophic variable code failure as a major problem, Vaisey et al. proposed to code as much as possible using fixed length codes and to provide a high level of error protection for the remainder using unequal error protection.

The H.261 algorithm was modified to reduce the sensitivity to bit errors, by eliminating the use of variable length codes for the transform coefficients. The DCT coefficients are quantized using non-uniform quantizers and then coded with 12, 30, 50 or 70 bits, depending on a classification of the DCT coefficients in the block. The block class is transmitted as side information. Temporal error propagation is reduced by coding every tenth frame in intramode.

The source code bits are classified into three classes. Class A contains the more sensitive bits, such as headers and DCT coefficient block size. Class B contains bits which cause severe degradation of quality but not loss of synchronization such as motion vectors and the DC coefficient. Class C consists of the DCT coefficients except the DC coefficient. The three classes are protected using Rate Compatible Punctured Convolutional (RCPC) codes at three different rates: 1/4 for class A, 1/2 for class B and 3/4 for class C. A convolutional interleaver was proposed to randomize the error. The delay associated with the proposed interleaver was estimated to be in the order of one second. The coder was not tested with simulated errors.

Stedman, Steele, Gharavi and Hanzo [Stedman92] proposed a 22 KBd mobile video telephone scheme and simulated its performance on a Rayleigh fading channel. The image coder is based on the sub-band decomposition of the motion compensated frame difference (MCP Hybrid DPCM/SBC). The seven sub-bands are quantized with the low frequency bands receiving more bits than the less significant high frequency bands. The quantized bands are then run-length coded using variable length codewords. An end-of-line code is inserted after each coded line in a sub-band. This allows the decoder to re-synchronize if errors occur, but also to correct one bit in error by tentative bit flipping.

The variable length stream is passed through a smoothing buffer which outputs video data at a fixed bit rate. Buffer fullness information is fed back to the quantizer controller to adjust the coder rate to maintain a fixed bit rate.

The video data is split into two streams, with the more important information grouped in one stream. These two streams are then unequally protected using BCH codes and transmitted over a flat Rayleigh fading channel using a bandwidth efficient 16-QAM modem. The 16-QAM modem supports two channels with unequal BER characteristics and a joint source-channel coding exercise was undertaken to assign the more important stream to the channel with the least residual BER following interleaved forward error correction.

The coder achieves an average PSNR of 38 dB at 55 kbit/s for the QCIF Miss America sequence at 10 fps. After BCH coding and addition of headers the bit rate rose to 88 kbit/s. This requires a baud rate of 22 KBd. When the SBC/BCH/QAM signal was transmitted over a Rayleigh fading channel at a carrier frequency of 1.8 GHz, using automatic gain control and second order diversity, at a pedestrian speed of 4 mph, the decoded image sequence remained error free for an average channel Signal-to-Noise Ratio (SNR) in excess of 16-18 dB.

Hanzo, Stedman, Steele and Cheung [Hanzo94] recently described a mobile speech, video and data transceiver scheme and studied its performance on an integrated multimedia wireless network using Integrated PRMA as the multiple access protocol. The video codec they proposed uses DCT for intraframes and hybrid motion compensated DPCM/DCT for interframes. The DCT coefficients are quantized using a fixed number of bits per coefficient, depending on the classification of the block according to its energy. The coded information is partitioned into two classes; A) motion vectors, and block classifiers and B) transform coefficients. Class A data was coded using a BCH(63,30,6) code and class B data using a BCH(63,51,2) code. A one frame interleaver was used. The QCIF Miss America sequence coded at 10 fps generated 25.1 kbit/s, and 39.06 kbit/s when coded (with 56% redundancy) at a PSNR of 39 dB. When transmitted over a 16 QAM Rayleigh fading link, the image quality deteriorated at an average SNR below 35 dB. With a 64 QAM modem, deterioration occurred below 40 dB.

Recently, Pelz [Pelz94] proposed an unequal error protected px8 kbit/s video transmission system for DECT. The video codec uses the DCT to code intraframes and motion compensated DPCM with structure coding to code the interframes. The coded information per frame is partitioned into seven variable length coded classes given be 1) contours, 2) motion vectors, 3) block positions, 4) DCT coefficients, 5)quantizer level, 6) quadtree, 7) and structure indices. Each class is preceded by a fixed length class size indicator.

Pelz performed a bit sensitivity analysis on the fourteen classes (including the class size indicators) and partitioned the fourteen classes into three error protection levels. Each level was coded using a RCPC code at a different rate. Redundancy rates vary from 50% to 200%.

Streit and Hanzo [Streit94] used unequal error protection in conjunction with an Intraframe fractal coder and two source sensitivity matched shortened binary BCH codes to code the video at a redundancy rate close to 100%. The two streams were transmitted over a 16-QAM radio link similar to that in [Stedman92]. The simulated system baud rate was 1037 KBd and the 39 KBd codec attained a PSNR of 30dB with the SNR above 15dB, with the channel simulated as flat Rayleigh fading.

3.7.2 Error Control by Retransmission

Error control by retransmission has not been favoured for real time services, because of the non-deterministic delay associated with the techniques in this class, and the excessive delay experienced on some long round-trip delay circuits. Error control by retransmission or Automatic

Repeat Query (ARQ) can, however, achieve very reliable communications, very efficiently and adaptively on full duplex channels.

Recently MacDonald [MacDonald92] described an experimental system designed at BT Laboratories for transmitting compressed video over DECT. The system uses selective-repeat ARQ which, although more complex than SAW and GOB-N ARQ, has the highest efficiency. H.261 compressed video was transmitted over a DECT demonstrator system at 32 kbit/s and 128 kbit/s. When a Reed-Solomon (63,59) code was used to correct errors in FEC mode with an interleaver, the picture quality was much degraded. The quality improved when the RS code was used to detect and correct errors, with retransmission of detected corrupted blocks, although it was still corrupted by errors which were incorrectly decoded by the RS code. When the FEC code was used for detection only, the picture quality was very good with no visible degradations. MacDonald concluded that the best solution was to use error detection and retransmission without an interleaver. A number of unresolved issues were identified such as the flow control of a remote H.261 codec, the rate mismatch between DECT and non-DECT codecs, delay problems and treatment of audio.

Scorse [Scorse93] described a system for transmitting JPEG compressed images over a fading HF link using a Selective reject ARQ technique. The receiver only acknowledges correctly received data, and the transmitter retransmits any un-ackowledged blocks following the end of the current pass. Thus the image is built up progressively at the receiver.

3.7.3 Intramode codecs

Hybrid codecs suffer from prolonged temporal error propagation. Some proposals avoid this by encoding frames in intramode only. In so doing, the encoded bit rate is between three to six times higher than when interframe coding. The advantage of the scheme lies in the simplicity of the codec and may be preferred in situations where very low power consumption is a priority.

Meng et al. [Meng94] proposed a low power video codec for wireless access to video-on-demand services. Their design centers on compression efficiency and error recovery. They reject variable rate entropy encoding which is highly susceptible to transmission errors and requires complex error correction which increases power consumption and adds overheads of about 40% to 60% to the overall bit rate. They propose, instead, a hybrid Sub-band Coder with Pyramid Lattice Vector Quantizer (SBC/PLVQ). The SBC decomposes the image into frequency bands which are then quantized at different bit rates using fixed length PLVQ. Each VQ codeword index is built up using a product code consisting of pattern, shape and sign information, thereby reducing the susceptibility to single bit errors. However the technique is entirely intraframe and achieves little improvement over JPEG in quality, although it is more robust to transmission errors than JPEG.

Belzer, Liao and Villasenor [Belzer94] are developing a robust, adaptive video codec for mobile wireless networks. The codec is based on sub-band decomposition and owes much of its robustness to the fact that it is an intraframe only technique. The codec uses variable length Huffman

codes to code the quantized sub-band coefficients. These are then protected by low-overhead Reed-Solomon block codes. End-of-frame and End-of-Sub-band symbols help maintain synchronization under channel error conditions. Good results were reported when transmitting the coded stream over a spread spectrum link with CDMA multiple access, operating at 915 MHz.

3.7.4 Other Techniques

Hemami and Meng [Hemami93] proposed a one-layer video coding technique for video transmission over a radio link. The video coder codes every sixth frame in intramode using a DCT. The intervening frames are coded using bi-directional interframe, motion compensated prediction, quantized using a vector quantizer with fixed length codewords. A structured packing strategy is adopted such that a packet loss results in the loss of a known quantity of information. Then lost intraframe blocks are linearly interpolated from the adjacent block edges. If interframe VQ data is lost it is simply ignored, since it does not contribute much to image degradation. If only a motion vector is lost, then the vector is estimated by averaging the motion vectors of the surrounding blocks, and the estimate used with the prediction data.

Roser *et al.* [Roser93] proposed the use of MPEG-1 for mobile use in the RACE MAVT project. A subjective evaluation has been carried out based on three aspects (i) Visibility of degradations, (ii) acceptability of degradations and (iii) recoverability from degradation.

3.8 Conclusions

This chapter attempted to identify a video coding algorithm on which to base the study of digital video transport over wireless networks. The main current standards (and at least one other in preparation) are based on the hybrid DPCM/DCT algorithm with motion compensation. Given the investment in the three standards - H.261, MPEG-1 and MPEG-2, and the fact that the market has not reached maturity, it is reasonable to expect that hybrid DPCM/DCT with motion compensation will dominate the market for some time. It is therefore reasonable to select this algorithm for further investigation.

This chapter also reviewed current and planned digital wireless networks to identify trends in digital modulation, error control techniques, and MAC techniques on which to base the wireless networks simulated in this thesis.

Of the current second generation digital mobile networks only GSM and DCS1800 can support digital video services at around 32 kbit/s, *provided they can aggregate time slots*. Both use FDMA/TDMA with FDD multiple access. Three of the second-generation cordless systems investigated can potentially support video services. These systems use TDMA/FDMA multiple access. Two use TDD duplex and the other FDD. It is thus apparent that TDMA/FDMA is worthy of further investigation. However a number of third generation proposals have concentrated on dynamic channel allocated MAC protocols like PRMA and MDR TDMA. DECT also supports dynamic channel

allocation. *TDMA, and TDMA with some form of dynamic channel allocation will be studied in this thesis.*

Furthermore, these cordless systems operate in an environment which does not require complex error control to support voice. *It is to be seen whether this low complexity approach can be applied to video.* Low complexity countermeasures identified in the previous chapter, such as antenna diversity, frequency hopping, and variable rate systems can improve the resilience of these channels without increasing the complexity of the system.

The bandwidth efficiency of the systems using $\pi/4$ QDPSK and $\pi/4$ QPSK is slightly higher than that of the systems using GMSK. Also, third generation systems like UMTS are also targeting 4-QAM for higher data rates. Other authors have studied higher level modulation schemes such as 16-QAM for video transport. *It is thus reasonable to study modulation schemes with four to sixteen modulation levels.*

Four classes of error control techniques for video were reviewed; generic techniques, packet video oriented techniques, techniques for DTTB and error control techniques developed specifically for video transmission over wireless networks.

Three important phases in error control for ISC have been identified namely; selection of a robust coding technique, error concealment and error recovery. A number of basic protection techniques have been identified and these appear in various guises in most of the recently developed techniques.

Layered coding was identified as a very robust technique against cell loss. Subjective tests show that it can tolerate between 1% to 10% packet loss in the enhancement layer without excessive deterioration in quality. Layered coding is thus a good candidate for video transmission over wireless networks. One important problem needs to be solved though, that of protecting the "guaranteed" layer on the undiscriminating radio channel.

Another interesting proposal is the use of ARQ on radio links, not only because the technique has not been previously considered suitable for real-time services, but also because it promises to deliver a highly reliable transport mechanism efficiently by adapting to the conditions on the radio link. *This technique together with layered coding will be investigated further in this thesis.*

Chapter 4
A Hybrid Interframe DPCM/DCT Video Codec

Three video coding standards have emerged recently: the CCITT H.261 standard, and the ISO MPEG-1 and MPEG-2 standards. The three codecs are based on a hybrid of interframe prediction and intraframe coding using the Discrete Cosine Transform (DCT). Although many other coding algorithms exist or are being researched, as amply documented in the previous chapters, the hybrid interframe DPCM\DCT algorithm is in wide-spread use and will continue to do so for some years. The H.261 video codec was selected for further study in this thesis. It is currently the only standard targeted at real time videophone and videoconference applications.

This chapter describes the software implementation and testing of a H.261 compatible codec: the Interactive H.261 codec (IHC). This codec forms the basis of all video coding algorithms described in the thesis. It implements the video coding part of the simulation infrastructure developed to study video transmission over wireless networks. The link to the wireless network part of the simulation infrastructure is provided through the error control module in the IHC, which can import error files generated by the low-level wireless radio link simulations.

All software development on the codec is original work. The fast DCT algorithm uses the one dimensional DCT butterfly developed by Arai *et al.* [Arai88]. The motion estimation algorithm is a fast, full search algorithm with a *short circuit* abort to speed it up. An original rate control algorithm was developed and is shown to perform well on a range of video material. A fast frame display system has also been developed to view and assess the decoded images in real time.

Validation results are presented which show that the IHC is H.261 compliant and can inter-work with other H.261 compliant codecs. Comparative tests show that the performance of the IHC is on a par with other H.261 codecs. The rate control algorithm is shown to be effective and efficient over a range of sequences and coding rates, even when coding highly dynamic sequences above 1 Mbit/s.

4.1 Introduction

The CCITT (International Telegraph and Telephone Consultative Committee - now ITU-T) Study Group XV developed the H.261 recommendation, Video Codec for Audiovisual Services at p x 64 kbit/s (p = 1 .. 30) , in recognition of the need for ubiquitous video services provision using the Integrated Services Digital Network (ISDN) [CCITT-H.261]. The H.261 recommendation was ratified in December 1990.

A joint ISO (International Standards Organization), IEC (International Electromechanical Committee) committee promulgated the MPEG-1 (Motion Picture Experts Group) video compression standard for multimedia applications [Gall91]. MPEG-1 was optimized for CD-ROM storage applications at bit rates in the region of 0.9 to 1.5 Mbit/s.

In a second phase, the ISO/IEC joint committee is currently finalizing the MPEG-2 standard which is targeted at coding of higher resolution video signals at bit rates up to 10 Mbit/s. MPEG-2 will be used for digital terrestrial and satellite television broadcasting. The standard was also found to be suitable for HDTV (High Definition TV) coding at bit rates between 20 and 40 Mbit/s [Anastassiou94].

The H.261 standard is the only one of the above three standards to address bi-directional, conversational, videophone and videoconference services, although MPEG-1 can also be used in this mode if precautions are taken to reduce the coding delay. A recent survey of PC-based videoconferencing products showed that H.261 compliant offerings accounted for two thirds of the surveyed sample [Gold94]. The main public domain videoconference systems used on the Internet support H.261 video coding. It is mainly for the above reasons that H.261 has been chosen as the basic video coding algorithm in this thesis.

The MPEG committee strove to maintain compatibility with the H.261 standard, and although the MPEG-1 standard is not a superset of H.261, there is much commonality between the two and a common hardware implementation is quite plausible [LeGall91]. Recently it has been shown that it is possible to transcode between MPEG-1 and H.261 in software alone [Zammit96a], such that an H.261 video stream can be viewed on MPEG hardware and vice versa. This is possible because the main video coding subsystems, namely motion estimation, motion compensated prediction, and the discrete cosine transform are common to both standards. The main modules used in H.261 and explained below are to be found in MPEG-1. Thus the results obtained for H.261 are relevant to MPEG-1. Indeed, certain error control techniques may be more easily incorporated into MPEG-1 than H.261.

An overview of the overall audiovisual teleconferencing system recommendations developed by the CCITT (now ITU-TS) is presented next. This is followed by a more detailed description of the H.261 video codec standard. The H.261 codec was simulated in software using the C programming language in a UNIX environment. The H.261 standard allows implementation innovation in a number of areas such as motion estimation, the forward and reverse discrete cosine transforms, and the buffer control algorithm. Details of how these functions were implemented are presented, followed by a description of the main operating modes of the Interactive H.261 Coder (IHC) and Decoder (IHD) programs. Validation and test results for the IHC are then presented using standard image sequences.

4.2 CCITT Visual Telephone System Overview

CCITT recommendation H.320 [CCITT-H.320], covers the technical requirements for visual telephone services at channel rates up to 1920 Kbit/s including the terminal equipment, the network multipoint control units and the other system operation entities as shown in Figure 4-2 [CCITT-H.320].

Only the H.261 video codec is studied in this thesis. H.261 is either studied in isolation or it is assumed, unless otherwise stated, that the necessary system control and indication signals, as well as the call establishment functions, function adequately for the duration of the video call.

4.3 H.261 Video Codec Overview

The H.261 recommendation describes video coding and decoding for fixed bit rate operation over the Integrated Services Digital Network (ISDN) at p x 64 Kbit/s, where p varies from 1 to 30. The recommendation covers both point-to-point and multipoint video conference services. The recommendation leaves much scope for innovative implementation of the various subsystems, especially in the encoder, such as the buffer control algorithm, motion estimation, and the forward and reverse discrete cosine transforms. An outline block diagram of the video codec is shown in Figure 4-2 below [CCITT-H.261].

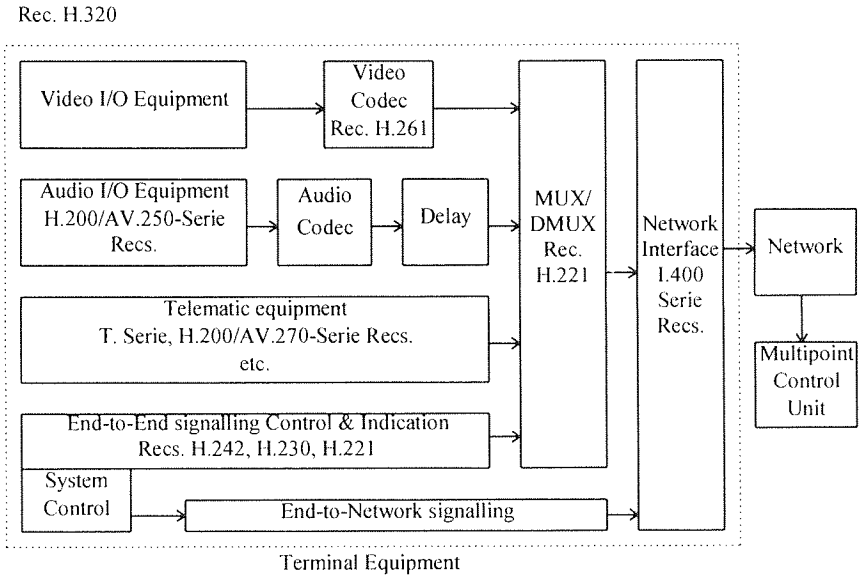


Figure 4-1 Visual Telephone System

The video coder operates on digital video signals in one of two formats defined by the standard, known as the Common Intermediate Format (CIF) and the Quarter-CIF (QCIF) format.

The source coder uses a hybrid algorithm to reduce the spatial and temporal redundancies in the image sequence in a lossy manner. Temporal redundancies are removed by previous frame prediction and conditional replenishment. Spatial redundancies are removed by applying a two dimensional Discrete Cosine Transform (DCT) on eight by eight, pixel or prediction error blocks. The transform coefficients are then quantized using uniform quantizers with dead zone.

The decoder also accepts one motion vector per 16x16 pixel block, so that the encoder can optionally operate in a Motion Compensated Prediction (MCP) mode. An optional filter can be used in MCP mode to improve compression.

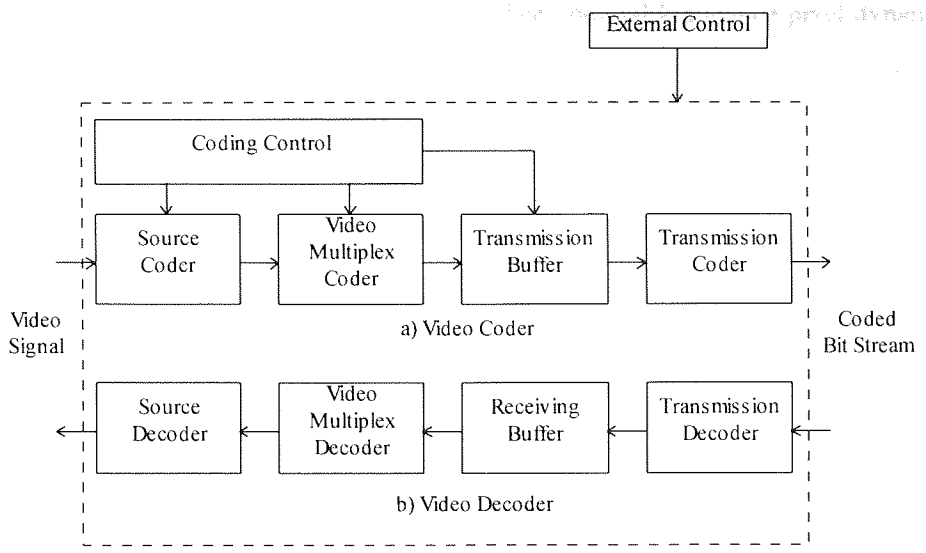


Figure 4-2 Video codec block diagram

The video multiplex coder uses fixed length and variable length codes to code headers, coding parameters and video data. It then combines the various fields according to a precise video stream syntax. The video multiplex coder achieves lossless compression through the use of Huffman-like Variable Length Codes (VLC) on many of the coded parameters, including the quantized transform coefficients.

The output bit rate of the Video Multiplex Coder (VMC) is highly variable. Since H.261 is intended for operation over fixed bit rate ISDN links, the transmission buffer and the coding control are required to smooth the variable VMC output. The transmission buffer occupancy is used by the coding control sub-system to vary various coding parameters, to adjust the output bitstream to meet the requirements of a Hypothetical Reference Decoder (HRD), in order to prevent buffer overflows and underflows.

The transmission coder adds a BCH(511,493,2) Forward Error Correction (FEC) code to the bit stream. An error correction frame structure is required to allow the decoder to synchronize. The standard specifies the re-lock conditions and maximum re-lock time following loss of error-correction framing.

4.4 Video Source Format

The Common Intermediate Format (CIF) is a compromise video format ensuring compatibility between NTSC (525 lines, 30 fps) and PAL (625 lines, 25 fps) video formats. The CIF format is based on the CCIR Rec. 601 for digitizing television signals [CCIR601].

A CIF frame consists of 360 Y (Luminance) pixels and 180 C_R and 180 C_B (C_R, C_B colour differences) pixels per line [CCIR601]. There are 288 Y lines and 144 C_R and C_B lines per frame. The frames are non-interlaced and occur at a rate of 29.97 frames per second (fps). The luminance and

colour difference samples are represented using 8 bits. The nominal luminance pixel dynamic range is from 16 (Black) to 235 (White). The nominal colour difference ranges from 16 to 240 at the peaks with the zero value shifted to 128. The algorithm must however handle input values from 1 to 254. The transformations between the Y, C_B, C_R and the Red, Green, Blue (R, G, B) display colour spaces are given by [CCIR601]:-

$$\begin{bmatrix} Y \\ C_B \\ C_R \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.1687 & -0.3313 & 0.5 \\ 0.5 & -0.4187 & -0.0813 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (4\text{-}1)$$

and,

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1.402 \\ 1 & -0.34414 & -0.71414 \\ 1 & 1.772 & 0 \end{bmatrix} \begin{bmatrix} Y \\ C_B \\ C_R \end{bmatrix} \quad (4\text{-}2)$$

4.5 Video Frame Subdivisions

The coding algorithm subdivides CIF and QCIF frames into Group of Blocks (GOB), Macro Blocks (MB) and Blocks (BLK) as shown in Figure 4-3 below. A CIF frame consists of twelve GOBs whereas a QCIF frames consists of only three GOBs (GOBs 1,3, & 5 in Figure 4-3). The group of blocks is implemented as an error control mechanism. If an error occurs during transmission of a GOB, the decoder can discard the errored block and continue decoding the next GOB, such that errors are contained spatially.

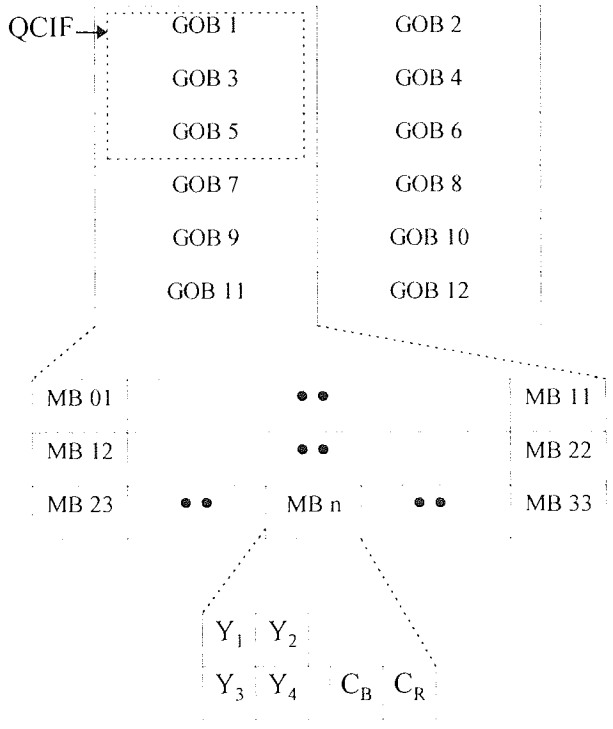


Figure 4-3 CIF and QCIF Frame Decomposition

A macroblock consists of six, 8x8 pixel blocks. Four of the pixel blocks are luminance blocks and the other two are the corresponding colour difference blocks. Together these comprise the colour information for a 16x16 pixel block. The macroblock is the basic unit for motion estimation and compensation. Motion estimation is carried out using the four luminance blocks in 16x16 block format. The chrominance blocks motion vectors are calculated from the Luminance motion vectors by halving the components and rounding down to the nearest integer.

The discrete cosine transform operates on eight-by-eight pixel blocks. This block size was chosen to allow cost effective hardware implementation, given that larger block sizes do not lead to a significant improvement in compression [Netravali88].

4.6 The Interactive H.261 Codec

The research project depended on the availability of a readily modifiable H.261 codec. This aim was reached by developing the Interactive H.261 Codec (IHC) software package. The IHC was coded in C and executes on UNIX based platforms such as SUN workstations and on IBM compatible PCs under DOS. The software was developed to be as flexible, re-usable and upgradeable as possible. Some coding speed was sacrificed to allow better debugging and provide more statistical information. However the object of re-usability was reached however, and the software can be readily modified and expanded. Key implementation details of the IHC coder are presented in section 4.7 followed by the IHC decoder description in section 4.8.

4.7 The IHC coder

The IHC coder block diagram is identical to the coder in Figure 4-2. It consists of four main blocks: the source coder, the variable length coder, the transmission coder and the rate control function which comprises the coding control and the transmission buffer.

4.7.1 The Source Coder

The source coder is shown in Figure 4-4. It reduces the spatial and temporal redundancies in the image sequence using lossy techniques. It has two modes of operation, an intraframe coding mode or intramode, and an interframe mode or intermode. In intramode a macroblock is coded using a DCT without reference to other frames and thus reduces spatial redundancies only. In intermode, a hybrid interframe prediction \ DCT algorithm is used to remove both temporal and spatial redundancies.

4.7.1.1 Intramode Coding

In intramode a macroblock is coded without reference to the previous frame. This mode must be used to code all frame macroblocks in the first transmitted frame, when a scene change occurs and when requested for multipoint operation (fast update request). In addition, macroblocks can be coded in intramode either due to forced updating, to prevent the accumulation of DCT arithmetic errors, or when

indicated by the macroblock coding mode decision function (e.g. as in the case of uncovered background).

In intramode, spatial redundancies are removed by applying a two dimensional discrete cosine transform on 8x8 pixel blocks in the current frame. The DCT concentrates the energy into a few low spatial frequency coefficients, with much less correlation than in the original subimage block, and which can be coded independently, close to their actual entropy (section 2.2.1.2).

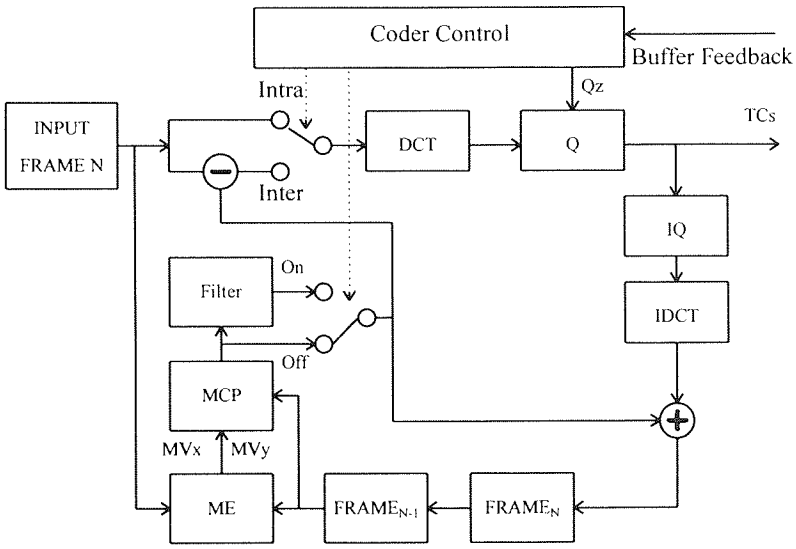


Figure 4-4 Source Coder

The coefficients are then quantized using uniform quantizers with dead zone except for the DC coefficient which is coded by an independent uniform quantizer without a dead zone. Quantization introduces unrecoverable quantization noise, thus rendering the source coding lossy. The human vision system is less sensitive to quantization noise in the high frequency coefficients than in the low frequency ones [Netravali88]. Thus high frequency coefficients can be coarsely quantized and even discarded without much loss in subjective image quality.

The quantized coefficients so obtained are then used to reconstruct a local copy of the transmitted image. At the end of the current frame this information passes into the previous frame store where it can be used for interframe prediction.

4.7.1.2 Intermode Coding

In intermode coding, temporal redundancies are removed by predicting current frame macroblocks from the previous frame. Only the prediction error is transmitted to allow the decoder to reconstruct the next frame. Since there is often very little difference between two consecutive frames, the prediction error is zero or close to zero over most of the frame. Whole regions of the image where no change occurs need not be transmitted. Additional compression is achieved by DCT coding the prediction error when this must be transmitted. The encoder and decoder must use the same reference

frame data for prediction, so that the encoder must decode the quantized DCT coefficients to reconstruct a local copy of the transmitted frame.

The decoder can use one motion vector to access the prediction macroblock offset by up to ±15 pixels in the horizontal and vertical directions. This will reduce the prediction error where the macroblock represents image objects in motion, either due to object motion or camera panning, hence the name Motion Compensated Prediction (MCP). The use of motion compensation by the encoder is optional. It can improve compression by a factor of two or more, however, such that its use is desirable in most applications, despite the large computational burden.

A two-dimensional low-pass filter applied to the MCP macroblocks has been found to greatly improve the coding performance. The use of the filter is optional and its selection is determined at macroblock level.

4.7.1.3 Motion Estimation

Block based motion estimation matches the current, 16x16 pixel luminance macroblock in the current frame F_N with 16x16 pixel luminance blocks in the previous frame $F_{(N-1)}$ over a 46x46 pixel search area. The position of the best match relative to the current luminance macroblock position, yields the required horizontal and vertical motion vector components as shown in Figure 4-5.

Various matching measures have been proposed, such as the cross correlation function, the mean square error and the mean absolute error. Irrespective of the distance measure used, motion estimation is very computation intensive. One macroblock distance measure requires three mathematical operations per pixel, or 768 operations. A full search over the search area requires nearly three quarters of a million mathematical operations per macroblock.

Hardware implementation is usually accomplished using parallel processing techniques (e.g. using VLSI systolic arrays). For software simulation, various non-optimal motion estimation algorithms have been proposed [Ghanbari90a]. The performance with these sub-optimal techniques, however, can be markedly less than with a full search algorithm.

A full search block matching algorithm with a Mean Absolute Distance (MAD) measure is used in the IHC implementation. The MAD measure at vector offset *{x,y} D(x,y)* is given by:

$$D(x,y) = \frac{1}{256}\sum_{i=0}^{15}\sum_{j=0}^{15}\left|P_{i,j}^{n} - P_{i+y,j+x}^{n-1}\right| \qquad \text{(4-3)}$$

Where $P_{i,j}^{n}$ is the pixel in frame n, in the i th row and j th column, and x and y are motion vector components ranging from -15 to 15.

The algorithm employs a *short circuit* speed up feature which is an improvement on the short circuit feature used in the p64 coder, a public domain H.261 codec developed at Stanford [Hung94]. In the modified algorithm $D(0,0)$ is computed first. If $D(0,0)<256$, motion estimation on the current block is immediately aborted since then the block would be coded without motion vectors anyway (refer to

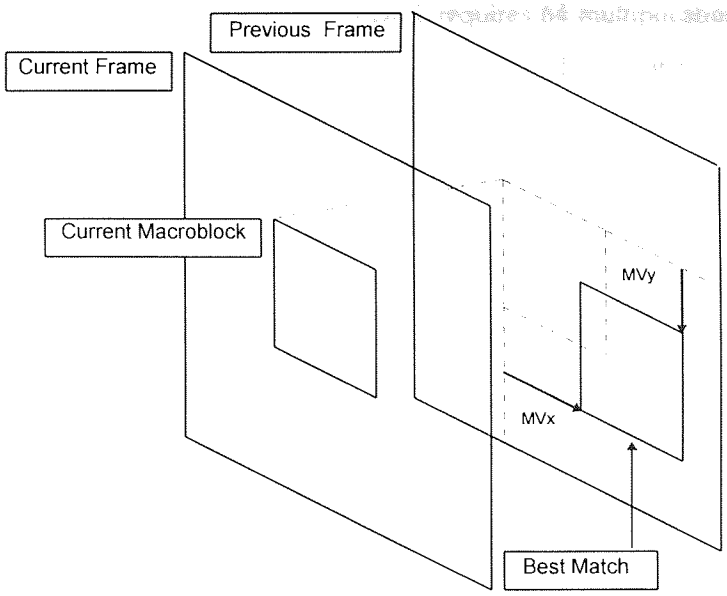


Figure 4-5 Block Matching Motion Estimation

section 4.7.1.6 below). If $D(0,0) > 256$, motion estimation continues, but a MAD computation is aborted if the partial sum exceeds $D(0,0)/x$ where x is 2 if $D(0,0) < 768$ and 1.1 otherwise. This again makes use of the fact that motion compensation is not used if the best match distortion exceeds $D(0,0)/x$ in the ranges described (refer to section 4.7.1.6 below).

4.7.1.4 The Discrete Cosine Transform

The transmitted image pixel blocks in intramode and the prediction blocks in intermode, are processed by a separable, 2-dimensional discrete cosine transform. The standard specifies the Inverse DCT (IDCT), the accuracy to which it should be computed, and the forced intramode updating rate to clear accumulated arithmetic errors in predicted macroblocks, but leaves the implementation details to the designer. The forward DC is given by [CCITT-H.261]:-

$$F(v,u) = \frac{1}{4}C(u)C(v)\sum_{x=0}^{7}\sum_{y=0}^{7} f(y,x)\cos((2x+1)u\pi/16)\cos((2y+1)v\pi/16) \qquad (4\text{-}4)$$

and the inverse DCT by [CCITT-H.261]:-

$$f(y,x) = \frac{1}{4}\sum_{v=0}^{7}\sum_{u=0}^{7} C(u)C(v)F(v,u)\cos((2x+1)u\pi/16)\cos((2y+1)v\pi/16) \qquad (4\text{-}5)$$

where y,x are the coordinates in the pixel domain, v,u are the coordinates in the transform domain, $F(v,u)$ are the DCT coefficients, $f(x,y)$ are the pixels and

$$C(x) = \sqrt{1/2} \qquad x=0 \qquad (4\text{-}6)$$

$$C(x) = 1 \qquad x\neq 0 \qquad (4\text{-}7)$$

The complexity of the DCT can be simplified in the same way as the discrete Fourier transform. Separability means that the computation of the transform can proceed first on the columns and then on the rows (or vice-versa) reducing the computation to sixteen DCT operations on eight

component vectors. Full computation of an 8 point DCT requires 64 multiplications and 63 additions, leading to 1024 multiplications and 1008 additions for an 8 by 8 DCT computation.

Many fast DCT algorithms have been described in the literature. A recent survey can be found in [Hung94]. The fast DCT algorithm proposed by Arai, Agui and Nakajima (AAN) [Arai88] has been used to implement the forward and reverse DCTs. The AAN butterfly shown in Figure 4-6 is for the inverse, eight point DCT. The forward DCT is computed using the same butterfly but with the pixel values input from the right and the transform values output from the left.

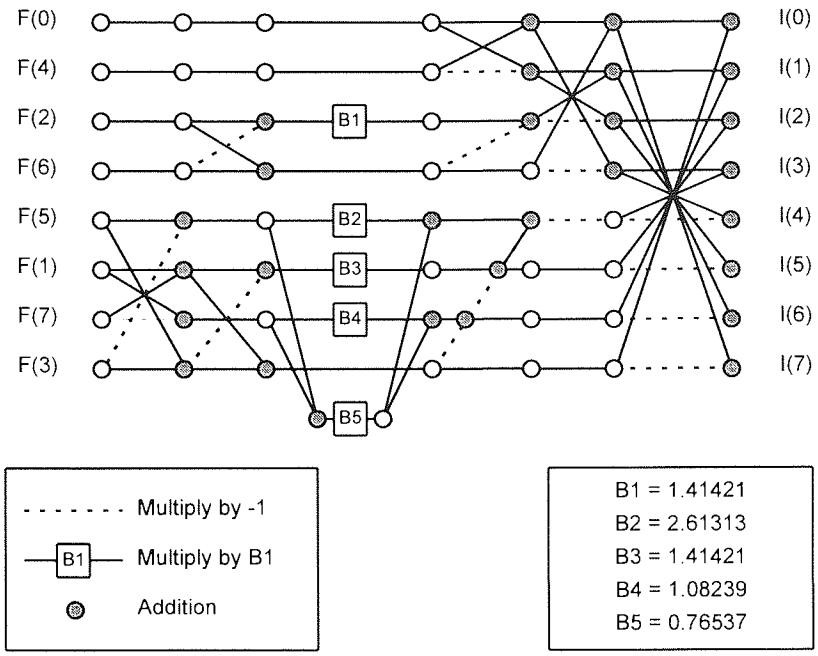


Figure 4-6 AAN DCT Algorithm

The AAN algorithm is a scaled DCT algorithm; the DCT coefficients require post scaling and the IDCT coefficients pre-scaling to arrive at the correct pixel values. The frequency components *F(x)* are transformed to the intermediate components *I(x)*. The time domain components *f(x)* are derived by multiply-ing *I(x)* by appropriate scaling factors *S(x)*. When computing a 2D DCT the scaling multiplications can be applied following the one dimensional row and column DCTs. The scaling factors can then be combined into an 8x8 array *S(v,u)* given by

$$S(v,u) = C(v)C(u)\cos(u\pi / 16)\cos(v\pi / 16) \qquad (4\text{-}8)$$

The pixel value *f(x,y)* is then obtained from *I(x,y)* by multiplication with the scaling factor *S(x,y)*.

The eight point DCT requires only five multiplications and 29 additions, allowing an 8x8 DCT computation with 80 multiplications and 464 additions. The post-scaling and pre-scaling operations require a further 64 multiplications yielding a total of 144 multiplications.

In the H.261 specification the DCT is succeeded by quantization and the IDCT is preceded by inverse quantization. These operations involve a multiplication of each DCT coefficient with a value

from a (inverse) quantization matrix. Thus the scaling multiplications can be combined with the (inverse) quantization operation saving 64 multiplications.

When computing the IDCT there is even more scope for complexity reduction since many of the input transform coefficients are zero and many arithmetic operations can be avoided. During inverse quantization and pre-scaling, the zero coefficients can be determined and actions taken to selectively eliminate unnecessary operations [Hung94].

The H.261 recommendation specifies the DCT accuracy in Annex A. The AAN DCT and IDCT implementations in IHC and IHD meet all the requirements set down. H.261 also specifies that each macroblock should be refreshed after 132 transmissions in predicted mode. Since there are 396 macroblocks, this criterion can be met by force updating three macroblocks every transmitted frame i.e. coding them in intramode. The algorithm implemented force updates three consecutive macroblocks per frame, in macro-block and GOB order.

4.7.1.5 Quantization

The DCT coefficients are quantized to 256 levels. Uniform quantizers with dead zone are used such that many of the least significant DCT coefficients are quantized to zero. Quantization accounts for the lossy nature of the source coding algorithm.

In intercoded blocks the transform coefficients are quantized using one quantizer with step size 2Q, where Q is the quantizer number. In intracoded blocks the DC coefficient $C(0,0)$ is quantized using a quantization step size of 8 independent of the Q number, whereas the AC coefficients are quantized as in intercoded blocks. The quantizer number may be specified at the GOB level (GQUANT) or at the macroblock level (MQAUNT). MQUANT overrides GQUANT until the next GQUANT.

H.261 specifies 32 different inverse quantizers. The forward and inverse quantizers for the intramode DC coefficient are defined by [CCITT-H.261]:-

$$C(u,v) = F(u,v)/8 \qquad \textbf{(4-9)}$$

and

$$F(u,v) = 8.C(u,v) \qquad \textbf{(4-10)}$$

where $C(u,v)$ is the quantized coefficient and $F(u,v)$ is the DCT coefficient. The remaining 31 quantizers are defined according to whether the quantizer Q is odd or even. If Q is odd

$$C(u,v) = F(u,v)/2Q \qquad \textbf{(4-11)}$$

and

$$F(u,v) = (2C(u,v) \pm 1)Q \qquad \textbf{(4-12)}$$

and if Q is even

$$C(u,v) = (F(u,v) \pm 1)/2Q \qquad \textbf{(4-13)}$$

and

$$F(u,v) = (2C(u,v) \pm 1)Q \mp 1 \qquad \textbf{(4-14)}$$

In equation (4-13) the positive sign in ± corresponds to positive $F(u,v)$ and the negative sign to negative $F(u,v)$. In equations (4-12)and (4-14) the positive sign in ± corresponds to positive $C(u,v)$ and the negative sign to negative $C(u,v)$, with the correspondences reversing for the ∓ sign.

4.7.1.6 Macroblock Coding Mode Selection

A macroblock can be coded in one of 10 ways ([CCITT-H.261] table 2). An optimal but computationally expensive method for selecting the macroblock type is to exhaustively apply each type and then select the one generating the least number of bits for the same coding accuracy. A simpler method described in [MPEG1] is to resolve the selection into a series of sequential decisions. The decision tree employed in the IHC is portrayed in Figure 4-7.

The first decision is whether to code the entire frame in intramode or not. Usually the first frame is intracoded but other instances require intracoded frames such as scene cuts, and fast update requests for multipoint operation. In addition, one of the proposed error control schemes makes use of periodic intra-coded frames as described below. If a frame is intracoded, all blocks are coded in INTRA mode.

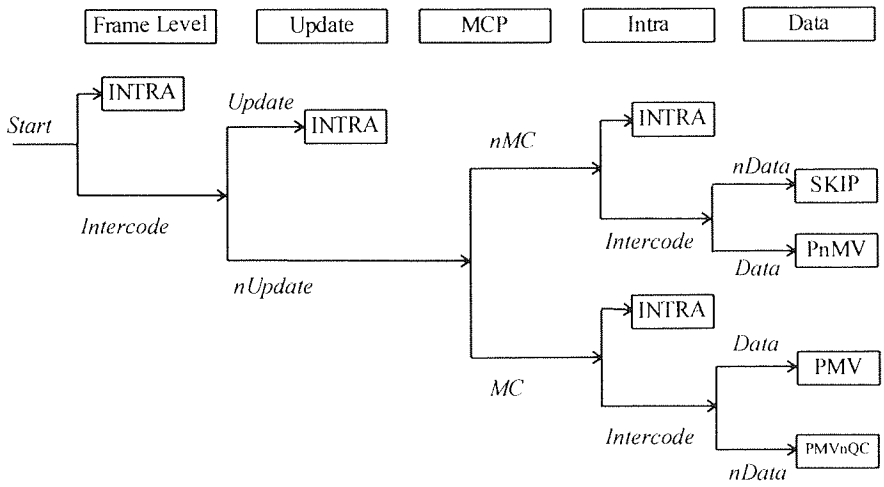


Figure 4-7 Macroblock Type Decision Tree

According to H.261 macroblocks must be intracoded after 132 intercoded transmissions. Rather than keep a separate count on the individual macroblocks, three macroblocks every frame can be intracoded. Since there are 396 macroblocks in a CIF frame the objective of forced updating every 132 transmission is thus met with reduced complexity. A more sophisticated update pattern generator has been incorporated in IHC as part of an error control technique by cyclic refresh. The update pattern generator can force update an arbitrary number of macroblocks per frame, in a variety of update patterns. Thus if the decision is to update the current macroblock it is intracoded, in which case all the six component blocks are coded.

The next decision is whether to use motion compensated prediction (MCP) or not. During motion estimation the sum of absolute differences at the zero offset position Z, and at the best match

position M are computed and stored. Then if Z/256<1, MCP is not used (nMC). For Z/256 between 1 and 3, MCP is used (MC) if M/256>0.5*Z/256. Above 3, MCP is used if M/256>(1/1.1)*(Z/256). This decision characteristic was used during the development of the H.261 standard, and attempts to avoid the objectionable drag-along effect of the background by moving objects as reported in the MPEG-1 standard [MPEG-1].

A method, used during the development of the MPEG-1 standard, was used to determine whether to use MCP or not. The variances *Varc* and *Vard* of the current macroblock and the pixel difference block are first computed using [MPEG1]:-

$$Varc = \frac{1}{256}\sum_{i=0}^{15}\sum_{j=0}^{15}\left(P_{i,j}^{n}\right)^2 \qquad (4\text{-}15)$$

$$Vard = \frac{1}{256}\sum_{i=0}^{15}\sum_{j=0}^{15}\left(P_{i,j}^{n} - P_{(MVy+i),(MVx+j)}^{n-1}\right)^2 \qquad (4\text{-}16)$$

where P^n is the pixel in the current frame n and P^{n-1} is the motion compensated pixel in the previous frame with offset motion vectors *MVx* and *Mvy*. If *Vard*<64 the macroblock is intracoded, otherwise it is intercoded if *Varc>Vard.*

Following quantization a block is skipped if all the coefficients quantize to zero. If all the blocks in a macroblock are zero then the macroblock is SKIPped if the motion vector is also zero. If the motion vector is non zero, the macroblock is coded as predicted with motion vectors but without quantized coefficients (PMVnQC). If at least one coefficient in any one block is non-zero then the macroblock is coded as predicted with motion vectors (PMV) or predicted with no motion vectors (PnMV), depending on whether the motion vector is zero or not.

The use of a prediction loop filter is optional. Results show, however, that if the filter is not used the motion compensation prediction does not result in much greater compression than using predictive coding alone. It is thus recommended to use the filter with motion compensated prediction. In the IHC the filter is selected for motion compensated blocks with transform coefficients to transmit.

H.261 allows the selection of one quantizer per macroblock. This feature is not implemented in the fixed bit rate codec. The rate control algorithm calculates the required quantizer with GOB granularity. Thus four macroblock types with a MQUANT field are never coded in this mode. In variable bit rate mode the quantizer can be set at the macroblock level.

4.7.2 Video Multiplex Coder

The source coding algorithm generates the quantized coefficients and side information which consists of motion vectors, coded macroblock addresses, the coded blocks in a macroblock, intra or predicted mode flags, the filter flag, and the quantizer number. The Video Multiplex Coder (VMC) uses Variable Length Codes (VLC) to code the quantized coefficients and the side information fields and then collates this information in the appropriate sequence, together with the various headers, to generate the Video Multiplex Stream (VMS).

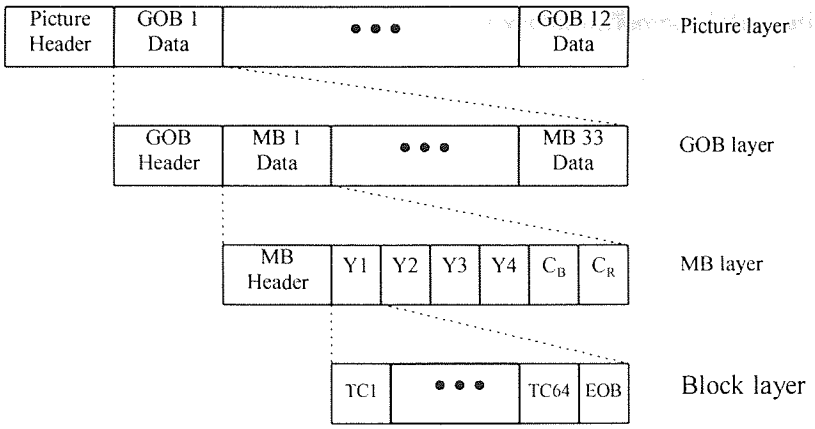


Figure 4-8 Video Multiplex stream hierarchy

The H.261 VMS hierarchy is shown in Figure 4-8. At the picture level the VMS consists of a picture header and twelve GOBs (three for QCIF). All the GOB headers must be present in a picture even if the GOBs do not contain coded macroblocks. The picture header and the GOB header fields have fixed lengths and contain fixed length sub-fields. At the GOB level, each GOB consists of up to 33 macroblocks. Any macroblock may be skipped. If it is present it consists of up to four luminance blocks and two colour difference blocks. Any block may be skipped if all the quantized coefficients are zero. The MB header is variable length and consists of variable and fixed length sub-fields. At the block level there are up to sixty four variable length coded coefficients. The blocks are separated by a 2-bit end-of-block (EOB) field.

The picture header consists of a unique 20-bit Picture Start Code (PSC), a five-bit Temporal Reference field (TR), a six-bit Picture Type (PT) field followed by an extra insertion information bit (PEI). If the PEI bit is set to one than an eight-bit PSPARE field plus one more PEI bit may follow. No use for the extra insertion information has been defined yet. The decoders must discard PSPARE in the H.261 specification. The PTYPE field has two spare bits, one bit specifying a QCIF or CIF format, a document camera indicator, a split screen indicator and a freeze mode release bit which allows a decoder to exit from Freeze mode.

The GOB header field consists of a sixteen-bit GOB Start Code (GBSC), a four-bit GOB Number (GN) (of which 0 is used for the PSC and 13,14, and 15 are reserved), a five-bit GOB quantizer (GQUANT), and a GOB extra information insertion bit (GEI) with the same function as PEI described above, but for the GOB header.

Every macroblock header starts with a variable length Macroblock Address (MBA) and Macroblock Type (MTYPE) fields. The MBA field is the difference between the current macroblock address and the last transmitted macroblock. The first MBA in a GOB is an absolute address with range 1 to 33. The MTYPE field indicates the macroblock coding mode. The coding mode then indicates which of the following four sub-fields are present.

MQUANT (five bits) is present if the quantizer value is to be changed for the current macroblock. As described above the IHC codec does not use this subfield in fixed bit rate mode.

The MVD subfield contains the X and Y motion vector difference information in two variable length codewords. MVD represents absolute motion vectors for macroblocks 1,12 and 23; and if the immediately preceding macroblock has been skipped or coded without motion vectors.

The Coded Block Pattern (CBP) is a variable length field indicating the coded blocks in the block data field. All data blocks are present in INTRA coded macroblocks so CBP is not transmitted.

The coded Transform Coefficients (TC) block is present whenever there are TCs to transmit. The video multiplex coder encodes the block quantized coefficients using variable length codewords which indicate {RUN,LEVEL} events. The coefficients in the blocks are scanned in the zig zag manner indicated in Figure **4-9**. The RUN indicates the number of zero coefficients preceding the current coefficient in the scan. The Level is its quantized value. A two bit EOB terminates the block data.

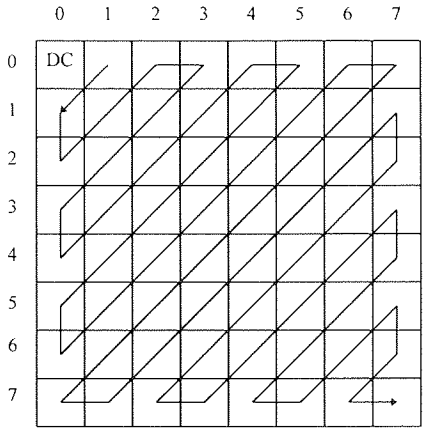


Figure 4-9 Block Zig Zag Scan

4.7.3 Rate Control

The output of the video multiplex coder is highly variable due to the nature of the source coding algorithm and the extensive use of variable length codes. H.261 is, however, targeted for fixed bit rate applications over ISDN. Three main mechanisms are included in the standard to smooth the output bit rate; the transmission buffer, the coding control function and the Hypothetical Reference Decoder (HRD). The generated video data is buffered before transmission. The state of the transmission buffer can be monitored by the coding control function, which must then vary key source coding parameters to avoid buffer underflow, that would reflect bad link utilization, and overflow, which would lead to lost data. The exact coding control strategies have not been defined in H.261, although a number of possible control parameters are indicated, such as pre-processing, quantizer control, block significance testing, and temporal sub-sampling by discarding complete pictures. H.261 specifies the maximum bit rates for CIF and QCIF frames, and the HRD model. The coding control strategies must adhere to the HRD model to ensure that the receive buffer in the decoder does not overflow leading to lost data. H.261-compliant codecs must adhere to this model.

The HRD model (Appendix B [CCITT-H.261]) stipulates that the encoder and HRD operate synchronously with the same clock rate and CIF frame rate. The HRD receiving buffer size is set at

(B+256Kbits) where B = $4R_{max}/29.97$ and R_{max} is the maximum link speed. Starting with an empty buffer the HRD examines the receiver buffer at CIF sampling instances and removes the data of complete pictures instantaneously. After each removal the buffer occupancy must be less than B after ignoring the bits pertaining to the transmission coder. This requirement is met if

$$d_{N+1} > b_N + \int_{t_N}^{t_{N+1}} R(t)dt - B \qquad \textbf{(4-17)}$$

where d_{N+1} are the number of bits used to code frame N+1, b_N is the buffer occupancy just after the removal of the N th picture at time t_N and R(t) is the video bit rate at time t.

4.7.3.1 Rate Control Strategies

The coding control adapted in the IHC follows a fixed frame rate strategy, that is the coding control attempts to code a constant number of frames per second with constant interframe separation. Rate variations are then countered by varying the quantization step size at GOB boundaries. Occasional frame skipping does occur, at scene changes for example, but although the coding control algorithm can deal with this, frame skipping is not exercised as a rate control technique.

Figure 4-10 illustrates the coding control strategy. The frames to code are $f_0,f_1,..,f_r,..,f_{N-1}$. The nominal frame period is T_F and the nominal video frame rate $R_F=1/T_F$. The first frame f_0 is allocated n times the bit rate of subsequent frames since it must be coded in intramode. Frames f_1 to f_{n-1} are not coded. The ideal target bit allocation for the r th. frame B_{Tr} is given by

$$\begin{aligned} B_{T0} &= nR_vT_F && r=0 \\ B_{T1} &= .. = B_{T(n-1)} = 0 && 0<r<n \\ B_{Tr} &= R_vT_F && r\geq n \end{aligned} \qquad \textbf{(4-18)}$$

where R_v is the maximum video transmission bit rate and is equal to the maximum link rate R_L less the transmission encoder overhead. It is assumed that every frame takes the same time to code - T_{DC}, and that after this coding delay the generated data for the frame is deposited into the transmit buffer ready for transmission.

The actual data generated by the r th frame is designated by b_r and this can be less than, equal to or greater than, the target allocation B_{Tr}. Ideally the two are equal but this may not achieved in practice. If $b_r>B_{Tr}$ the extra bits are taken from the following coded frame. The transmission time the previous frame encroaches into the transmission period allocated for frame r, is designed T_{Er}. Thus frame f_0 encroaches into frame f_3 by T_{E3} in Figure 4-10. Frame f_5 on the other hand, does not utilize all the transmission time allotted to it.

Occasionally, a frame encroaches heavily into the next frame, e.g. frame f_6 encroaches heavily into frame f_7 in Figure 4-10. The remaining transmit time for frame f_7 is T_x and it is equal to (T_F-T_{E7}). If T_x is less than an arbitrary minimum threshold T_{min}, than the frame is skipped. The coder than codes the next closest frame in the sequence instead, that is f_8 instead of f_7 in Figure 4-10. After a skipped

frame, coding continues at the same frame rate and with the same interframe separation. A T_{min} value of $0.75xR_vT_F$ was found to be a good compromise, allowing enough encroachment to prevent frequent frame skipping in normal operation, whilst ensuring that a frame is coded with enough bits to ensure good picture quality. The un-utilised time for the r th frame is designated T_{Ur}. This un-utilised time can be used by macroblock stuffing, by transmitting fill bits (refer to 4.7.4) or it can be assigned to the next frame. This last option is preferable because it improves the sequence quality. A buffering delay T_{DB} must be inserted in the stream to ensure that the data is available for transmission, otherwise the transmit encoder would request the video frame data for transmission before it has been encoded, which is physically unrealizable. The buffer delay must be equal to $max(T_{MIN},max(T_U))$ where $max(x)$ denotes the maximum value of x. Now the transmit time for the r th frame T_R is equal to (T_F-T_{Ur}) with un-utilised time T_{Ur}. T_R must be greater than T_{MIN} since T_{MIN} is the minimum coded frame transmit duration. Then $(T_F-T_{Ur})\geq T_{MIN}$ and $(T_F-T_{MIN})\geq T_{Ur}$. Finally since T_{MIN} has been chosen equal to $0.75T_F$, $T_{Ur} \leq 0.25T_F \leq 0.75T_F=T_{MIN}$. Thus $T_{DB}=T_{MIN}$.

The overall encoder delay T_{DE} is given by $T_{DC}+T_{DB}$. The maximum encoder buffer size must be equal to the maximum allowed actual bit rate for one frame: $max(b_r)$, plus the buffer occupancy during one frame time T_F: T_FR_V, i.e. Maximum Buffer Size = $max(b_r)+T_FR_V$.

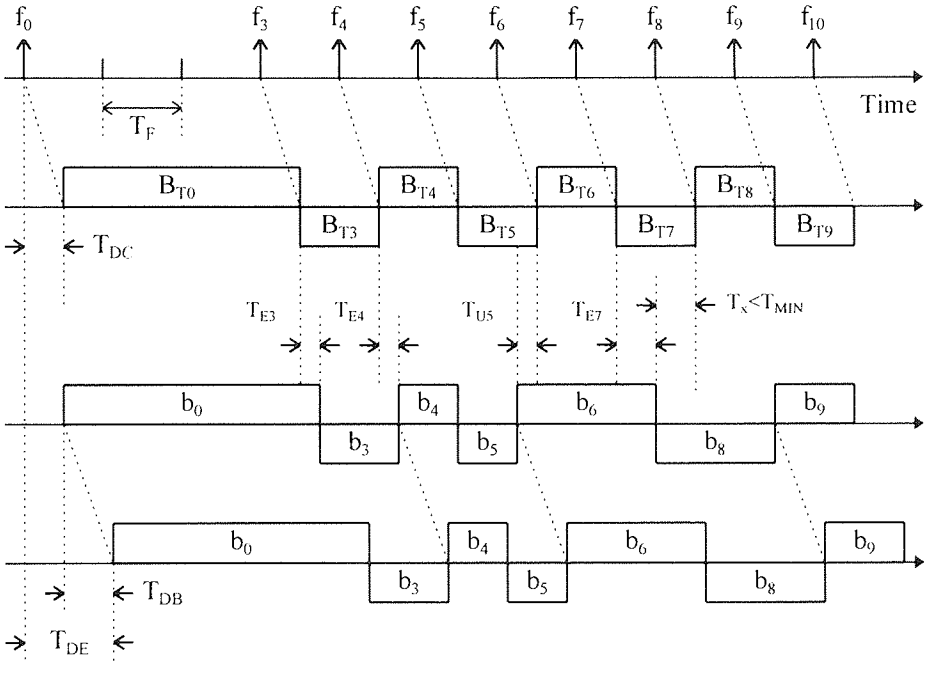


Figure 4-10 Rate Control Strategy

The algorithm describe can be shown to meet the HRD requirements as follows. First consider a receiver buffer from which the data of a video frame is removed as soon as the last bit of that frame enters the buffer. Then it is obvious that the buffer occupancy immediately after the frame removal is zero. If the receive buffer is inspected at CIF intervals, than the buffer occupancy immediately after frame removal is at most $(T_{CIF}R_V-1)$ provided each video frame is coded with at least $T_{CIF}R_V$ bits. Thus

the buffer occupancy is always less than B (which is equal to $4T_{CIF}R_{max}$+256 kbit) thereby satisfying the HRD requirement provided $T_{min} \geq .R_V T_{CIF}$.

The coding algorithm requires accurate control over the bit rate generated per frame. This is achieved by varying the quantization step size at GOB boundaries except for the first intracoded and the first intercoded frames, for which only one quantizer value is calculated for the whole frame. The algorithms used to meet the preset targets are described next.

4.7.3.1.1 Intracoded Frame Algorithm

The first frame in a sequence must be intracoded. Usually intracoded frames generate between three to five times more data than intercoded frames. This means that at low frame rates the first intracoded frame transmission time approaches one second.

The bit rate for a minimal intracoded CIF frame, that is one containing just one DC coefficient per block, was calculated to be 26084 bits. The minimal QCIF frame requires 6545 bits. Setting Q to 31 generates more data than the minimal calculated value since not all AC coefficients are quantized to zero even at this setting. These values must be taken into consideration when the target bit rate for the first intracoded frame, B_{T0}, is computed.

The 'brute force' algorithm to meet the target bit rate with one quantizer would be to intracode the frame 31 times incrementing Q from 1 to 31, and then selecting the Q value that generates the rate closest to the target. The brute force algorithm can be refined by noting that the Rate versus Q curves are monotonic so that starting with Q = 1 the search can stop as soon as the generated rate drops below the target, and the Q value is determined from the rates about the target value. The algorithm used in the IHC is a further refinement which makes use of a logarithmic step search to find the Q value in six steps. The algorithm used to meet the target T_0 is as follows

Step 1	Pre-compute and store the transform coefficients for the whole frame.
Step 2	Set Q to 16, Q_s to 8, and R_{min} to infinity.
Step 3	Code the transform coefficients and calculate the size of the Code R_i. If $R_i(Q)<R_{min}$ store Q as Q_{min} and $R_i(Q)$ as R_{min}.
Step 4	If $R_i(Q)=T_0$ or $Q_S=1$, then the required $Q=Q_{min}$. Go to Step 6.
Step 5	If $R_i(Q)>$Target let $Q=Q+Q_S$, and go to Step 3, else set $Q=Q-Q_S$ and go to Step 3.
Step 6	Code the frame with Q_{min} and Exit.

4.7.3.1.2 Intercoded Frame Algorithm

The first intercoded frame is also coded with one Q value for the whole frame, which is found using the same algorithm described above, except that during the first pre-computation step, MCP

hybrid interframe prediction is executed prior to transform coding. Subsequent interframes are coded with variable Q per frame computed for GOB segments, which incurs less computational delay, emulates real world rate control algorithms more realistically, but which is more robust to errors since the GOB header and hence the GQUANT field, can be protected from error as described below. The algorithm requires two inputs, the target bit rate for the whole frame T_T^n, and the array R_i^{n-1} for i=1..N_G (N_G=12 for CIF frames and 3 for QCIF frames). R_i^{n-1} is the bit rate generated whilst coding the *i* th GOB in the previous intercoded frame n-1, with Q=16. Then the algorithm proceeds as follows to code frame *n*.

Step 0 Set i=1

Step 1 Pre-compute the *i* th GOB up to the DCT and store the coefficients.

Step 2 Set Q to 16, Q_S to 8, and R_{min} to infinity.
Compute T_i^n, the target bit rate for the *i* th GOB from

$$T_i^n = \frac{T_T^n R_i^{n-1}}{\sum_{r=i}^{N_G} R_r^{n-1}} \qquad \textbf{(4-19)}$$

Step 3 ode the transform coefficients and calculate the size of the code $R_i(Q)$
If $R_i(Q)<R_{min}$ store Q as Q_{min} and $R_i(Q)$ as R_{min}.
If Q=16 store $R_i(Q)$ as R_i^{n-1}.

Step 4 If $R_i(Q)$ is equal to T_i^n or Q_S=1, then the required $Q=Q_{min}$. Goto Step 6

Step 5 If $R_i(Q)> T_i^n$ let $Q=Q+Q_S$ and goto Step 3, else set $Q=Q-Q_S$ and go to Step 3.

Step 6 Code the *i* th GOB using Q_{min}.
If $i=N_G$ then Stop else let i=i+1 and go to Step 1.

In the above $R_{min} = R_i(Q_{min})$ is the closest rate to the target I th GOB bit rate T_i^n, Q_{min} is the value of Q to achieve R_{min}. Q_S is the quantization step size which is initially set to eight, and halved following each iteration.

The performance of the algorithm is very good as evidenced from the validation tests and the results shown in Figures 4.12 to 4.16. The algorithm is very stable and copes well with both QCIF and CIF frames at frame rates from 48 kbit/s to above 1 Mbit/s. The maximum algorithmic delay introduced by the algorithm is equal to the duration of one GOB, assuming that the encoder can code the GOB within that period.

4.7.4 Transmission Coder

The transmission coder removes data from the transmission buffer in 492 bit packets, adds one fill bit indicator and then codes the resulting 493 bit information word using a BCH (511,493,2) forward error correction code. The transmission coder clock is provided by the ISDN interface, and the

coder must provide video data on every valid clock cycle. Thus if the transmission buffer does not contain at least 492 bits, the transmission coder sets the fill indicator bit to '0' and the other 492 bits to '1' before BCH coding. The receiver must then discard the fill packets upon reception.

The inclusion of the BCH code necessitates error control framing to allow the decoder to distinguish the parity bits from the data bits. This is achieved by appending one framing bit S_i to the 511 bit codeword yielding a 512 bit multiframe packet. An eight packet multiframe is used with the framing pattern $S_1S_2S_3S_4S_5S_6S_7S_8 = (00011011)$.

The BCH FEC code is not implemented implicitly in the IHC coder. Thus video streams generated by the IHC coders are raw video streams without FEC parity bits, fill bits or framing bits. The BCH is simulated within the IHC decoders as described below.

4.8 The IHC decoder

The IHC decoder block diagram is shown in Figure 4-2 It consists of a transmission decoder which decodes the BCH code and deposits the decoded video data in the receive buffer. The VLC decoder decodes the VLC codewords and passes the quantized transform coefficients and side information to the source decoder. The source decoder operates on this data to reconstruct the transmitted frames.

4.8.1 Transmission decoder

The IHC transmission decoder *simulates* generic BCH (n,k,t) FEC block codes. The simulation uses an error source to corrupt n bit codewords. Then, depending on the decoding mode and the FEC code distance properties, the errors may be deemed to be correctable, or else detectable but not correctable. Bounded distance decoding is assumed [Lin83, Michelson85] and the probability of incorrect decoding is usually assumed to be negligible with respect to the post decoding error probability.

The BCH(511,493,2) defined in the H.261is decoded using a bounded distance decoder [Lin83], which can correct up to two errors. When more than two errors occur the error pattern may be detectable, or else it may lead to an undetectable error pattern. The transmission coder and decoder simulations are described in chapters five and seven.

The IHC can use two error sources; a internal random error source or an external error file. The random error source is used to test the robustness of the decoder and the effectiveness of proposed error control techniques on random error channels with high bit error rates. The external error file option provides the essential link between the video coding and the radio link simulations in this thesis. The bit error vector files generated by the radio link simulations can be read in by the IHC decoders, optionally operated upon by the BCH simulation function, and then used to corrupt the video data. The internal FEC simulation can be disabled such that the effects of other FEC and/or ARQ error control techniques can be pre-computed and input via the error file option.

The IHC decoder can use error detection events in a number of ways. It can ignore errors and continue decoding, it can discard the errored block and jump to the next GOB, or it can try to conceal the error.

4.8.2 Source Decoder

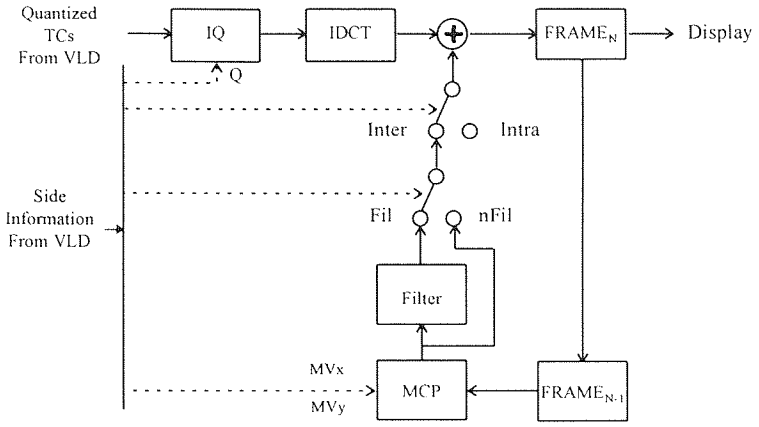


Figure 4-11 Source Decoder

The source decoder is shown in Figure 4-11. It. uses the inverse quantization equations described in section 4.7.1.5 to reconstruct the transform coefficients. The AAN IDCT described in section 4.7.1.4 is used to reconstruct the sub-image blocks or the prediction error blocks. In intermode, the motion compensation block can accept one motion vector per macroblock, to predict the current macroblocks from the reference frame. The absence of the motion estimation function simplifies the source decoder drastically, and the IHC decoder typically decodes H.261 streams at ten times the frame rate that the IHC coder encodes them.

4.9 IHC functionality

A comprehensive description of the IHC functions is presented in appendix 2. An overview of the major functions and the rationale behind their inclusion is presented in this section.

The main design objectives driving the IHC development were flexibility and re-usablity. The complexity of the included functions, however, added another necessity; that of ease of use. Thus an interactive, text based, menu driven user interface was included in the programs.

Four modules were developed; a CIF coder, a CIF decoder, a QCIF coder and a QCIF decoder. Separate CIF and QCIF codecs have been implemented to achieve some increase in speed by optimizing the code. The CIF codecs can handle normal 12 GOB CIF frames or the 10 GOB frames generated by some NTSC orientated codecs. The codecs can display the sequences in 24 bit RGB colour or in colour component mode which displays the selected frames as monochrome Y,U,V components. The QCIF decoder can decode QCIF streams at 10 to 15 fps on fast SUN and Pentium workstations. The CIF decoder is four times slower and a real-time display program has been coded to view decoded sequences in real time.

4.9.1 The IHC Coder

The IHC software coder has a text based, interactive front end which provides control over all file manipulation, video file coding and display. A simple to use, text based, two layer menu system was developed. The sub-menus allow:-

- Selection of all coder parameter options.
- Control of the three main coding modes:-
  - The Variable Bit Rate Coder
  - The Fixed Bit Rate Coder and
  - The Layered Coder.
- Video frame and coded stream file manipulation.
- Debugging mode selection.
- Statistics collection control.

4.9.1.1 Coder Parameters

The following coder parameters can be set from the coder parameters sub-menu:-

Motion estimation can be switched on or off to evaluate the additional compression achievable through motion compensated prediction.

For VBR modes it is possible to select either a fixed quantizer, or a fixed target MSE per macro-block. If the latter (experimental) option is selected, an automatic quantizer mode becomes operational, which searches for the quantizer that returns a macroblock MSE closest to the target. In FBR mode it is possible to set the target bits per frame and the frame skipping distance.

A macroblock forced updating pattern generator has been implemented as part of a cyclic refresh error protection scheme. The type of update pattern and the number of macroblocks to force update per frame can be chosen for one frame or the whole sequence.

During zig-zag scanning the number of non-zero coefficients per block is calculated. If this number is lower than a preset threshold, then the coefficients are forced to zero and the block is not transmitted. Setting the threshold to one or two results in additional compression with little loss in SNR. The streams used in the thesis are, however, coded with the threshold set to zero, to disable this feature.

4.9.1.2 Variable Bit Rate Coding

The basic VBR mode codes a sequence by intracoding the first frame and intercoding all subsequent frames. Alternatively the coder can code the whole sequence using intracoded frames only. It is also possible to code individual frames in intramode or intermode. The locally decoded sequence can be displayed during coding. Statistical and .debugging data files can be generated and stored to hard disk.

The rate-distortion R(D) curves for intracoded and intercoded frames can also be produced. The curves are generated by repeatedly coding the frames, whilst varying Q from 1 to 32 and measuring the bit rate and the mean square error.

4.9.1.3 Fixed Bit Rate Coding

In normal FBR mode the first frame is intracoded. The rate control algorithm assigns it n times the bit rate per frame set in the parameters menu (n is set to 2 unless otherwise specified). The coder then skips (n-1) frames and codes all subsequent frames in intermode. The rate control algorithm calculates one quantizer value for the first intramode and first intermode frames (frame level rate control). Subsequent frames are coded in intermode with GOB level rate control, meaning that one quantizer is calculated per GOB. It is also possible to code individual frames in intramode or intermode using frame level rate control, and in intermode using GOB level rate control.

4.9.1.4 Layered Coding

A layered codec has been implemented as a specific error control technique based on the H.261 codec. Detailed description of the codec and its modes of operation are deferred to chapter 9.

4.9.1.5 Video frame/sequence manipulation

The IHC coder utilizes four frame buffers which can be viewed individually. Two displays are available: a 24 bit RGB colour display and a monochrome display of the YUV colour components. A video frame can be loaded from hard disk to the reference frame buffer where it can be viewed or coded. A sequence stored on hard-disk or CD-ROM can be viewed in 24 bit colour, or as YUV components.

4.9.2 The IHC Decoder

The decoder follows the same design philosophy as the coder. It is menu driven and has been coded to ease development, reusability and upgradeability, often at the expense of speed. The colour H.261 decoder can decode colour or monochrome H.261 sequences using any number of GOBs per frame. The decoder decodes an H.261 stream file frame-by-frame or continuously, starting from any specified byte aligned position in the stream. The output can be sent to the screen or to a file. Statistical data generation is optional and can be stored to file if necessary. Several levels of debug and check levels can be selected by the user.

A stream decode mode allows the decoder to locate all picture and GOB headers in a sequence. A VLC decode mode decodes the variable length codes in the stream and stores the header locations and macroblock size information to a statistics collection file. The decoder can store all the intermediate operations it performs on the data stream so that it is possible to select a macroblock and

trace the decoding operations on a convenient coefficient matrix display (option only available in X-Windows version).

An error simulation mode allows the user to introduce an error in any number of macroblocks in a frame, by setting all macroblock pixel values to, say, zero. The temporal and spatial diffusion of the error can be investigated by subsequent frame-by-frame decoding. The error generation menu can also simulate a range of BCH block codes with bounded distance decoding. A programmable random error generator can be used to corrupt the streams. Alternatively an error file can be read into the decoder and used to corrupt the video stream on a bit by bit basis. A more comprehensive description of this important function is deferred to chapter five.

4.10 Validation tests

Initially the IHC codes were validated by comparing the PSNR and rate performance of the codecs with published results using well known test sequences. In particular the colour Salesman sequence at 10 frames per second was coded and the results compared with those of Morrison and Beaumont [Morrison91]. The results shown in Figure 4.13a agree very closely with the graph in [Morrison91-Fig. 7].

Subsequently, the codecs were validated by decoding streams generated by public domain software codecs as these became available. IHC streams have been successfully decoded using the Stanford Portable Video Research Group's H.261 codec [Hung93] and the Inria Videoconferencing System [Turletti92] H.261 codec, and *vice versa*.

4.10.1 Test Sequences

Three standard test sequences were used in the tests; the Miss America (MA) sequence, the Salesman (SM) sequence and the Table Tennis (TT) sequence . The Miss America sequence is a typical head and shoulders, videophone-type sequence with very limited motion initially, until the subject sways sideways in front of the fixed camera.

The salesman sequence depicts a salesman seated at his desk describing a component he holds in his hand. This sequence is more detailed and dynamic than the previous one but the salesman occupies less frame space than Miss America's head and shoulders. The SM sequence is typical of a person to person(s) videoconference session.

The table tennis sequence features a short clip of a table tennis match. The action is fast, there is fast camera zooming and panning, and a number of scene cuts. This sequence is typical of a video-on-demand retrieval from CD-ROM or a video broadcast. This sequence was used during the MPEG-1 standard development phase and generates bit rates in the region of 1 Mbit/s.

4.10.2 Results

The PSNR and bit rate curves for five sequences coded with the IHC codec are plotted in

Figures 4.12 to 4.16. Also plotted are the PSNR and bit rate results for the same sequences coded using the well known public domain p64 software codec produced by the Stanford Portable Video Research Group [Hung93].

The first two sets of comparative results relate to two QCIF sequences generated by low pass filtering the MA and SM CIF sequences and subsampling them in the horizontal and vertical direction to QCIF format. The 2D low pass filters used were of the separable type, with 1D kernel (1/4,1/2,1/4). These two QCIF sequences were then coded in VBR and FBR mode.

The QCIF Miss America (QMA) colour sequence was coded at 12.5 fps with the quantizer parameter Q set to 8 using both the IHC and p64 coders. The bit rate per frame and the luminance PSNR are shown in Figures 4.12 (a) and 4.12(b) respectively. The graphs show that the IHC coder generates a lower bit rate per frame but achieves a marginally better luminance PSNR consistently for all frames.

The PSNR and bit rate for the QMA sequence coded at a fixed bit rate of 48 kbit/s, using the rate control algorithms described in section 4.7.3, are plotted in Figures 4.12(c) and 4.12(d). The bit rate generated is very constant and the luminance PSNR is also relatively constant reflecting the low dynamicity of the QMA sequence. In comparison, the p64 coder produces a rather unstable bit rate per frame and a more variable luminance PSNR which is consistently lower than that generated by the IHC.

The graphs in Figure 4.13 were obtained for the colour QSM sequence at 10 fps. The VBR results again indicate that the IHC performs better, generating a higher PSNR at a lower bit rate per frame. The 132nd frame generated by the p64 coder shows an instantaneous high peak, revealing the force updating strategy adopted by the p64 coder, which is to code every 132nd frame in intramode. In contrast, the IHC coder updates three macroblocks every frame to distribute the forced updating burden over the whole sequence and hence avoiding the five times increase in bit rate every 132nd frame.

The IHC rate control algorithm is clearly more stable than the p64 algorithm. The p64 coder produces larger PSNR variations than the IHC coder, which produces quite constant PSNR results. Despite the extreme bit rate peaks for the p64 coder, the IHC PSNR is superiour for most of the frames in the sequence.

The next three sets of results were obtained using three different CIF sequences. The 12.5 fps colour MA sequence was coded in VBR mode with Q set to 8. The results shown in Figures 4.14 (a) and 4.14(b) show that the CIF sequence generated around four times the bit rate of the QCIF sequence. When viewed at the same display size the QMA frames are visibly inferior to the MA frames despite the PSNR being higher in the former. When viewed at the same resolution (i.e. with the QCIF frames at one half the height and width at the same viewing distance) the subjective quality of the two sequences is similar.

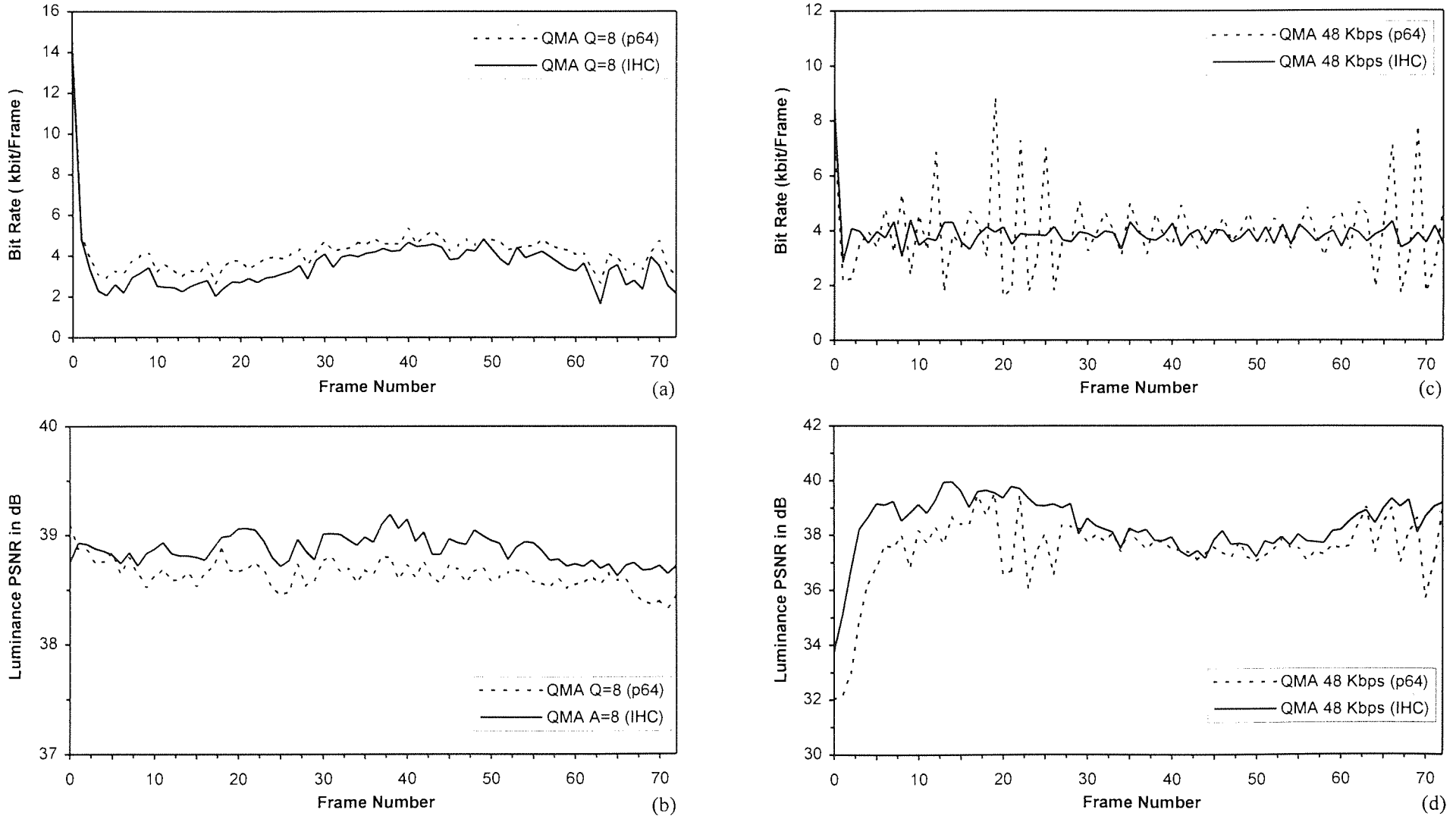


Figure 4.12 Comparative curves for the QCIF Miss America (QMA) sequence coded at 12.5 fps with the IHC and the p64 coders. (a) Variable bit rate with Q=8, (b) Luminance PSNR for case (a), (c) Fixed bit rate at 48 kbit/s, (d) Luminance PSNR for case (c).

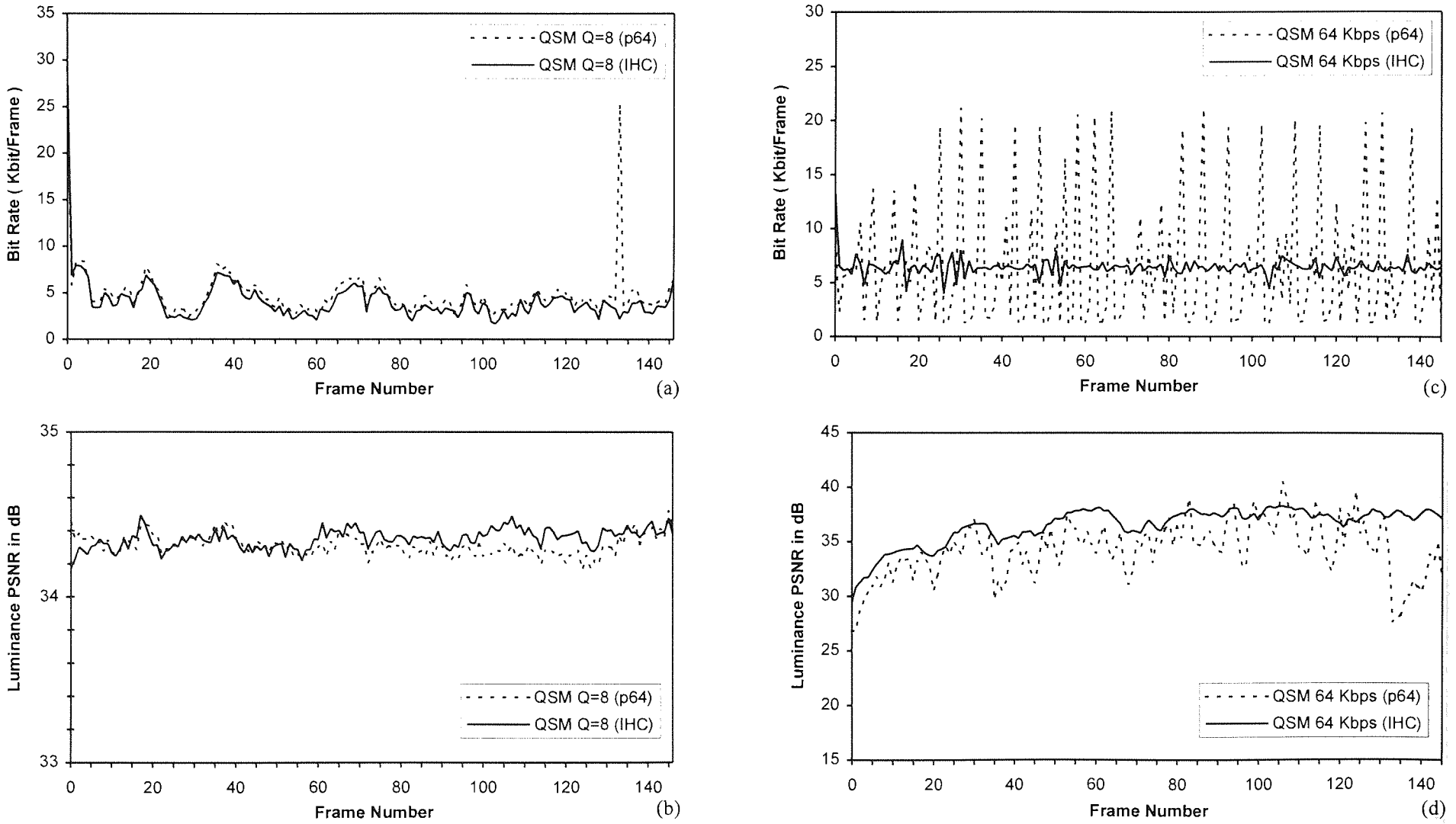


Figure 4.13 Compartive curves for the QCIF Salesman (QSM) sequence at 10fps coded with the IHC and p64 coders. (a) Variable bit rate with Q=8, (b) Luminance PSNR for case (a), (c) Fixed bit rate at 64 Kbps, (d) Luminance PSNR for case (c).

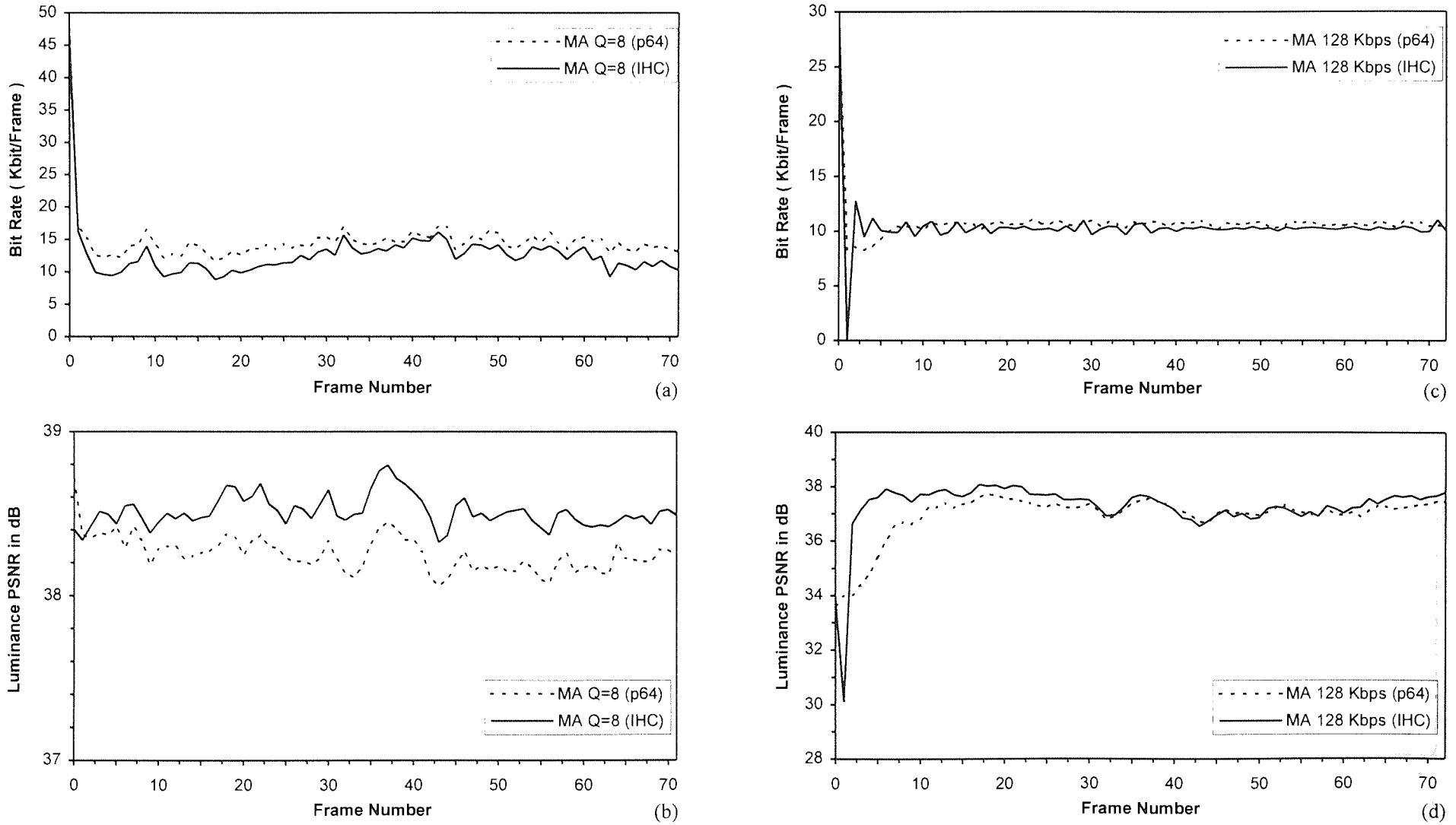


Figure 4.14 Comparative curves for the CIF Miss America (MA) sequence at 12.5 fps coded with the IHC and p64 coders. (a) Variable bit rate with Q=8, (b) Luminance PSNR for case (a), (c) Fixed bit rate at 128 Kbps (d) Luminance PSNR for case (c).

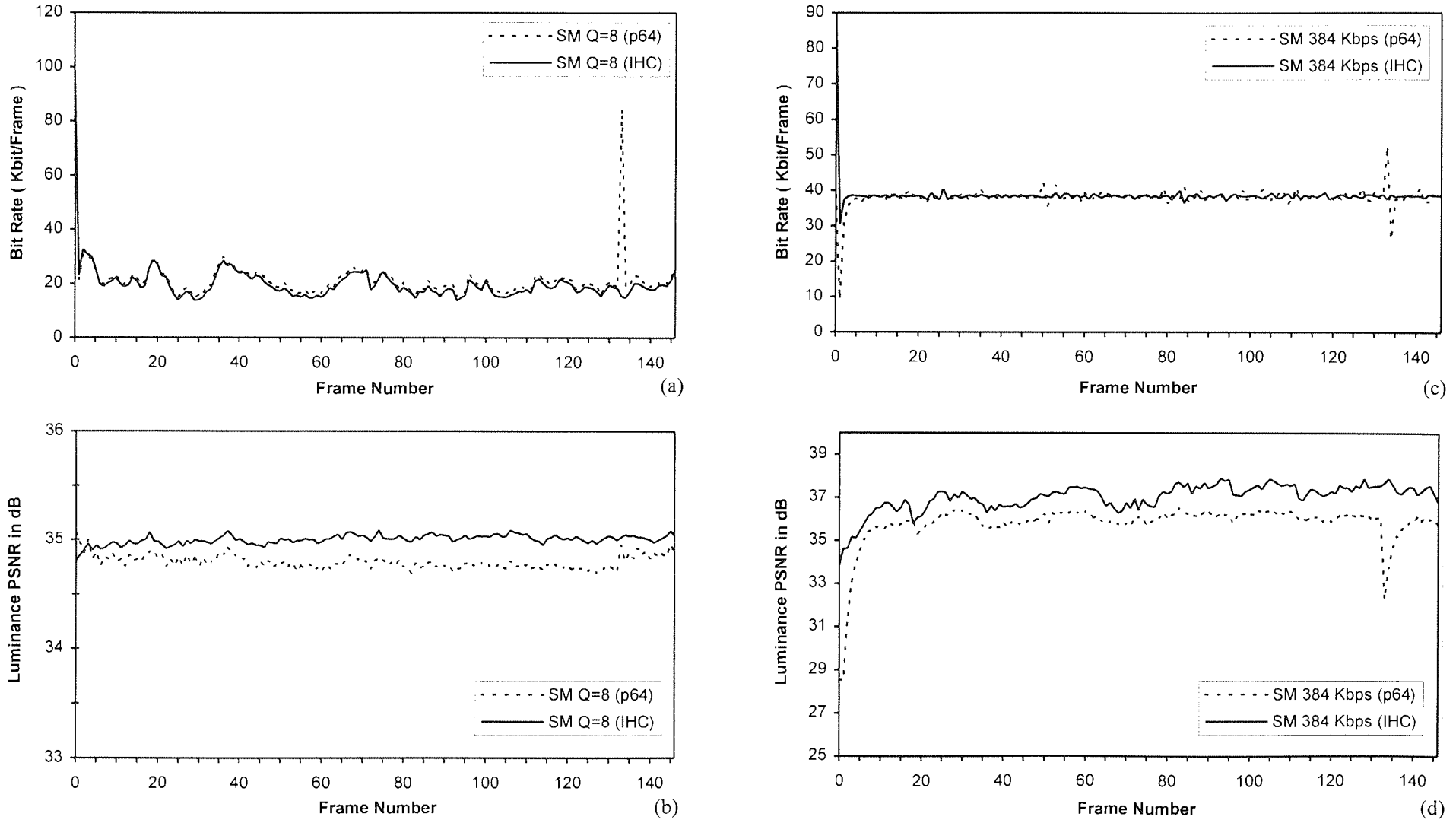


Figure 4.15 Comparative curves for the CIF Salesman (SM) sequence at 10 fps coded with the IHC and p64 coders. (a) Variable bit rate with Q=8, (b) Luminance PSNR for case (a), (c) Fixed bit rate at 384 Kbps, (d) Luminance PSNR for case (c).

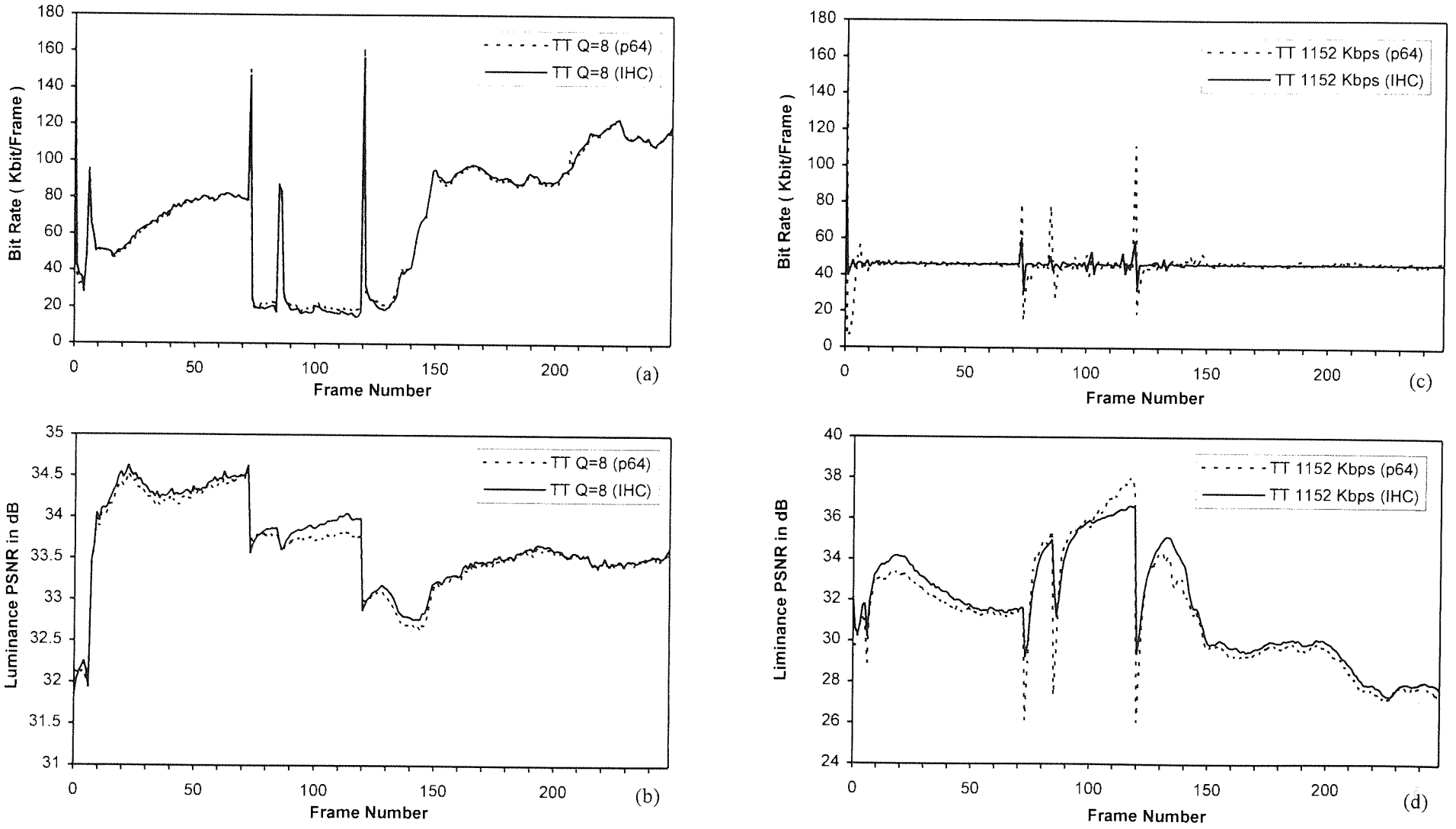


Figure 4.16 Comparative curves for the CIF Table Tennis (TT) sequence at 25 fps coded with the IHC and p64 coders. (a) Variable bit rate with Q=8, (b) Luminance PSNR for case (a), (c) Fixed bit rate at 1152 Kbps, (d) Luminance PSNR for case (c).

The marginally superior performance of the IHC algorithm experienced when coding the QCIF sequence was again evident when coding the CIF MA sequence as shown in Figure 4.14. The FBR performance of the p64 coder, however, improved dramatically to close the performance gap between the two coders, although the IHC performance remained slightly superior when the MA sequence was coded at 128 Kbit/s.

The trend continued when the 10 fps, colour SM sequence was coded in VBR mode as shown in Figure 4.15. However the performance of the IHC coder in FBR mode at 384 Kbit/s was again distinctly better than the p64 coder. The forced updating strategy adopted in the p64 coder causes an increase in bit rate at the 132nd frame which is less pronounced than in VBR mode. This leads to a reduction in PSNR which lasts for around 5 frames (i.e. 500ms at 10 fps).

The last sequence tested was the ten second long, 25fps, CIF colour Table Tennis sequence. The results for the sequence coded in VBR mode are presented in Figures 5.16(a) and 5.16(b). These results reveal the dynamic nature of the sequence which is characterized by large rate and quality variations. Four of the bit rate peaks evident in Figure 4.16(a) correspond to the four scene changes in the sequence whereas the peak at the 84th frame is due to a one pixel frame displacement to the left present in the original 84th frame, which led to the jump in bit rate for this frame.

When the sequence was coded in FBR mode with a target bit rate of 1152 kbit/s, the IHC generated a more regulated output than the p64 with less bit rate variability, although bit rate variations do occur at scene changes as shown in figure 4.16(c). The PSNR curves show that the IHC rate control algorithm delivers better quality over most of the sequence expect for the central scene where the p64 coder performed marginally better.

4.11 Conclusions

This chapter discussed the implementation of an H.261 compliant video codec capable of compressing video so that it can be carried effectively and efficiently over telecommunication networks. The video codec represents the first part of the simulation infrastructure developed in this thesis to study video transmission over wireless networks. The software implementation allows a large degree of control over all aspects of the coding algorithm, and allows detailed measurements at all points in the compression/decompression chain.

Software implementation meant that key, computationally intensive operations had to be coded in an optimized way. The fast DCT/IDCT algorithms described by Arai *et al.* [Arai88] have been used in the IHC codecs. A modified version of the full search algorithm used in the PVRG p64 codec, was used to implement the motion estimation. The modifications allow a motion estimation to be aborted if it becomes apparent that the macroblock will not be coded using motion vectors.

A rate control algorithm was developed to produce fixed rate output with a constant number of bits per coded frame. The rate control algorithm is of the one pass type and does not introduce additional delays.

A number of standard test sequences have been coded and decoded successfully. The video streams have been validated against published results and have been successfully decoded using other H.261 compliant video codecs.

A study comparing the performance of the IHC codec with that of a public domain software coder - p64, produced by the Stanford Portable Video Research Group, shows that the IHC coder performs consistently better over a range of sequences, coded at different frame rates, for both CIF and QCIF sequences in both VBR and FBR modes. The rate control algorithm achieved the aim of delivering FBR with good quality. The good performance of the algorithm when coding the TT sequence was unexpected, as the fixed frame nature of the rate control algorithm was not expected to perform well with highly dynamic material, although it is not possible to generalize to other sequences from this one sequence. However the good IHC performance in this validation phase meant that the sequences generated were representative of the class of streams generated by H.261 compliant codecs.

The H.261 video codec was used as described above in a number of the experiments that follow. It also forms the basis of modified video codecs which were developed to combat the deleterious effects of the wireless access channel. These modifications and enhancements will be described in subsequent sections. A number of the enhancements were designed to be H.261 compatible, meaning that an unmodified H.261 compatible decoder would be able to decode the enhanced streams. In other cases the modifications require modified decoders to be decoded successfully.

Chapter 5
Coded Video Transmission on High Bit Error Rate Channels

The performance of an H.261 video codec on a binary-symmetric, random-error channel with high BER is investigated in this chapter. The main aim of this investigation is to identify weaknesses in the H.261 standard, and to develop robust decoders which can survive high bit error rates in the region of 10^{-4} to 10^{-3}, which are common on wireless networks. The un-coded wireless network channel is not well modelled as a binary-symmetric, random-error channel; however, by developing robust codecs for the random error channel it is hoped to improve the resilience of the codecs to burst errors which are typical of wireless networks.

The investigation of the robustness of the H.261 codec to high bit error rates is an original contribution, although other studies of hybrid DPCM/DCT schemes have been reported in the literature as are documented in chapter three. The techniques proposed to improve the resilience of the H.261 standard are original contributions. The techniques of cyclic refresh, forced updating and demand refresh have been proposed for error control on packet video networks and have been adapted to H.261 in high BER environments. Transcoding is widely used in audio wireless networks. The technique has now been proposed for use with video wireless networks. The techniques proposed in the transcoding section are original contributions with adaptations of previous work indicated in the relevant sections.

5.1 Introduction

The H.261 standard addresses video transport over the ISDN network where the expected bit error rates (BER) are substantially lower than what is considered adequate for voice. In fact, on voice-oriented wireless networks, post decoding bit error rates in the region of 10^{-3} and packet error rates in the region of 10^{-2} are considered sufficient for acceptable voice communications. In ISDN networks the un-coded BER is often better than 10^{-7}.

The high compression ratios achieved by the H.261 codec means that there is very little redundancy left in an H.261 coded stream. This leaves the stream very vulnerable to bit errors. The motion-compensated, hybrid interframe DPCM/DCT nature of the source coding algorithm means that errors propagate not only spatially in the current frame, but also temporally into subsequent frames. The variable length codes generated by the entropy encoding schemes employed in the standard, accentuate the error sensitivity of the streams since one bit error causes loss of synchronisation, corrupting large areas of a frame and propagating into subsequent frames. The standard does incorporate error control techniques, but these are based on lower error rates than anticipated on

wireless networks. This chapter investigates the resilience to high bit error rates of the hybrid interframe DPCM/ DCT codecs developed in the previous chapter.

The chapter is organised as follows. An analysis of the inherent error resilience of the hybrid interframe DPCM/DCT is presented presented first. The design of robust H.261 codecs is discussed followed by a description of a robust H.261 codec for high BER applications. The resilience of this robust design is tested by corrupting four test sequences using random errors. The performance of the decoder is reported using four criteria, the bit error rate per frame, the mean square pixel error per frame, the number of macroblocks in error, and the histogram of Error Pipe Lengths (as defined below). These criteria are meant to assess both the spatial and temporal error propagation in the sequences. An informal subjective assessment of the corrupted sequences is reported together with the objective reports. A subjective quality rating is defined which attempts to summarise the 'usefulness' of the corrupted sequences.

Two solutions are then presented to overcome shortcomings identified in the robust decoders. The first relates to the protection of video multiplex stream headers and the second concerns the improvement of the resilience to errors of intracoded frames.

The use of forced updating to limit temporal error propagation is investigated next. The same sequences studied before are re-coded with forced updating and corrupted using the same random error patterns. The decoded results are compared with the decoded sequences without forced updating.

Initially, only H.261-compatible measures are studied. Subsequently, error control measures which generate non H.261-compatible streams are proposed. These can be incorporated in a non-obtrusive method using low-complexity, low-delay transcoders. Four transcoder based techniques are then proposed and discussed.

5.2 *Error resilience of hybrid Interframe DPCM/DCT codecs*

Video coding standards such as H.261 comprise three stages which impact on the error resilience of the codecs, namely, the source coding algorithm, the variable length coder and the transmission coder (the MPEG-1 standard does not include a specific transmission coder). The properties of these three stages are analysed in the following sections.

5.2.1 Error resilience of the source coding algorithm.

The inherent error sensitivity of the hybrid DPCM/DCT source coding algorithm depends on the four major mechanisms employed; namely, interframe prediction, motion compensation, conditional replenishment and the two-dimensional DCT. The DCT contributes to the spatial propagation of errors, DPCM and conditional replenishment cause temporal error propagation, whilst motion compensation and can cause both spatial and temporal error propagation.

The DCT compacts the energy in a few lower frequency components and the inverse DCT re-distributes the energy over the entire block again. Thus errors which alter one DCT coefficient will, after inversion, affect the whole block. Errors in the DC and low spatial frequency components are much more visible to the human vision system than are errors in high frequency components. In particular, DC coefficient errors cause highly visible and objectionable block contrast errors. On the other hand, the block nature of the two dimensional DCT means that coefficient errors are contained within the 8x8 pixel block and do not propagate outside this limited area.

The DCT in H.261 and MPEG codecs employs a threshold coefficient selection scheme and the algorithm codes all non zero coefficients following quantization. A coefficient addressing scheme is required to identify the coefficients. In H.261 this is achieved by scanning the coefficients in a zigzag pattern and forming {run,level} events. The zig zag scan transforms the two dimensional coefficient array into a one dimensional stream, with the run field determining the position of the next coefficient relative to the previous non-zero coefficient. Thus errors can corrupt the run address data leading to a miss-registration of all subsequent coefficients. This also leads to highly visible and objectionable block patterns, with the subjective degradation decreasing with errors affecting higher frequency coefficients.

In the MPEG algorithm the DC coefficients are DPCM coded with respect to the DC coefficient of the adjacent block in macroblock order. Thus if an error occurs in one DC coefficient this is propagated to other blocks. This feature is absent in H.261.

The H.261 algorithm employs previous frame prediction to capitalise on interframe redundancies. The previous frame is subtracted from the current frame to generate the frame difference image. The 2D DCT is then applied to 8x8 pixel subimage blocks from the frame difference image and any block with non-zero quantized coefficients is coded and transmitted. At the decoder, the received data is added to the previous frame buffer to generate the next displayed frame which is also the reference frame for the following frame. Errors that corrupt the difference data thus become embedded in the reference frame, from where they affect all subsequent blocks which refer to this corrupted data.

In H.261, transmitted blocks must be coded in intramode after 132 transmissions in intermode to prevent the accumulation of arithmetic truncation errors arising from the non-exact implementation of the DCT and IDCT arithmetic operations. Thus for continuously updated blocks an error will remain in the reference frame for up to 132 transmitted frames. This translates to around 9 seconds at a transmitted frame rate of 15 frames per second (fps). As worded, however, the H.261 standard does not guarantee that blocks are refreshed every 132 frames: only actively transmitted blocks need to be refreshed. Thus if a background block is corrupted and it is not transmitted again, it will remain corrupted for the duration of the whole sequence. An easier and more frequent implementation of the refresh rule is to intracode three macroblocks every transmitted frame irrespective of whether these have reached the re-transmit threshold count. This scheme guarantees the refreshing of the decoder

reference frame every 132 frames and thus limits the lifetime of corrupted blocks. This technique is referred to as *Implicit Refreshing*.

The classical technique to combat errors in DPCM coding schemes is to employ a fractional prediction coefficient α such that the current value x_n is given by $x_n = \alpha x_{n-1} + e_n$, where e_n is the transmitted prediction error. Then the contribution of a corrupted value x_{n-1} is reduced by α^r following r subsequent predictions. Thus since $\alpha<1$ the deleterious effects of the error diminish in time. This technique is not used in H.261 which uses unity α.

Without motion compensated prediction, corrupted predicted blocks propagate temporally but are confined spatially. Motion-compensated prediction causes the errors to spread spatially outside the original blocks by allowing reference to the corrupted blocks in the reconstruction of otherwise uncorrupted blocks. The region of influence of the motion estimation algorithm extends ±15 pixels horizontally and vertically around the corrupted macroblock in H.261.

The motion vectors themselves represent an additional error risk. If the motion vectors become corrupted then the reference point for a 16x16 pixel macroblock is altered leading to visible degradation. In H.261 the motion vectors are differentially encoded with respect to motion vectors of the previous, horizontally adjacent macroblock, except for macroblocks on the edge of the GOB and where the horizontally adjacent macroblock is not coded with motion vectors. This means that a motion vector error can potentially propagate for the width of a GOB affecting up to eleven horizontally adjacent macroblocks.

The H.261 (and MPEG) source coding algorithm is based on the conditional replenishment concept of transmitting only those blocks which change significantly between frames. Thus the macroblock transmission pattern is not predictable and a macroblock address field is necessary in the macroblock header to identify the macroblock position. The absolute macroblock address increments from left to right, top to bottom in the current GOB. A differential macroblock address is transmitted, however, which represents the difference between the absolute addresses of the current and the previous transmitted macroblock. A macroblock address error thus causes the miss-registration of all subsequent blocks in the GOB leading to highly objectionable spatial image degradation.

The transmitted blocks within a macroblock are indicated by the Coded Block Pattern. If this field becomes corrupted, then some of the blocks will be decoded in error.

5.2.2 Error resilience of the video multiplex coder

The video multiplex coder codes the source coder image data and side information using variable length codes and collates it with other header information into an H.261 video multiplex bit stream. The use of variable length codes to code most of the side information fields and the {run.length} transform coefficient events using interleaved, completely loaded encoding trees means that if a bit error occurs synchronisation is lost for all the subsequent fields. It is also very difficult to

detect loss of synchronisation at the variable length decoder level, since the decoding trees are nearly complete, and other means must be relied upon to detect errors and then limit error propagation.

Not all errors affecting the variable length fields cause loss of synchronisation however; *e.g.* errors affecting the least significant bits of equal length codewords sometimes cause a decoding error without loss of synchronisation. The effect which the decoding error has depends on the particular field affected. Survivable bit errors occur quite frequently, but are mainly restricted to transform coefficients substitution errors.

An analysis of the bit error resilience of the video multiplex fields is presented in Appendix 3. The main results indicate that the most critical fields are the Picture, GOB and Macroblock header fields. In particular the Picture and GOB headers must be protected to ensure successful operation of the error control schemes supported within the H.261 specification (as described below in section 5.4). The macroblock header fields have no synchronisation facility whatsoever. If an error occurs then the nearest synchronisation point is at the GOB header.

In conclusion, the video multiplex coder output is very vulnerable to channel errors. The design may be valid in low BER conditions but is insufficient for high BER channels. It is not even particularly efficient when it comes to compression efficiency. The H.263 codec under development achieves twice the compression mainly through the use of better variable-length coding. The error sensitivity of the new multiplex coder may be even worse however, since channel error sensitivity should increase with information redundancy reduction.

5.2.3 H.261 error control mechanisms

The H.261 specification provides two error control mechanisms; the GOB frame segmentation, and the BCH(511,493,2) forward error correction code. The GOB frame segmentation prevents spatial error propagation outside the corrupted group of block by providing re-synchronisation points at the GOB headers. Thus errors may potentially affect one twelfth of the picture with CIF frame format and one third of the frame in QCIF format.

The use of the BCH (511,493,2) forward error correction code by the decoder is optional. This code can correct all one and two-bit random error patterns. Alternatively it can be used to detect random errors. When used with bounded distance decoding with an (n,k,t) block code, the probability of correct decoding P_{CD} is given by [Michelson83]:-

$$P_{CD} = \sum_{i=0}^{t} \binom{n}{i} p^i (1-p)^{(n-i)} \qquad \textbf{(5-1)}$$

where p is the channel bit error probability for the binary symmetric channel with random errors. The post-decoding block error probability P_b is given by [Michelson83]:-

$$P_b = (1 - P_{CD}) \qquad \textbf{(5-2)}$$

The post decoding block error probability for the BCH(511,493,2) code for a channel bit error rate of 10^{-3} is equal to 0.0152, thus more than one block per hundred will contain errors. At a channel bit error rate of 10^{-4} the block error probability reduces to 2.1×10^{-5} and the expected block error rate is a more acceptable one block every five minutes at 64 kbit/s. It is to be noted that current wireless networks usually support voice services at a BER of 10^{-3} and a packet error rate of 10^{-2}. It is thus probable that the BCH code does not offer adequate protection on these networks and additional protection may be needed.

The BCH code requires the start of the error control blocks to be identified. This is achieved by introducing error correction framing using an eight frame multiframe consisting of eight framing bits, each bit preceding a 511 bit codeword. The frame alignment pattern is $(S_1S_2S_3S_4S_5S_6S_7S_8)$ = (00011011). If frame alignment is lost the minimum decoder re-lock duration is 12288 transmitted bits and the maximum is 34000 bits. If the probability of framing loss increases because of the higher BER, then the framing mechanism poses a serious problem.

Data must be transmitted on every clock cycle according to the recommendation. When video data is unavailable the error framing packet is filled with all ones instead. A bit is included in the error frame to indicate the presence of video or fill data. A corrupted fill bit can lead to the loss of the whole packet. The fill bit forms part of the BCH protected information word, reducing the risk of this happening.

5.3 The design of robust H.261 codecs

A robust H.261 decoder has to be designed by making full use of the error control mechanisms in the H.261 specification. Error detection relies mainly on the detection capabilities of the BCH code in the transmission coder but codeword violations detected by the VLC decoder also permit a certain measure of error detection. Error concealment can be implemented by using the GOB synchronisation fields to confine errors to a GOB and then displaying the previous frame GOB video data. No error recovery mechanism is explicitly specified in the H.261 decoder specification, however two specified mechanisms can be exploited to achieve error recovery. The first is the forced updating mechanism and the second is the fast picture update mechanism included for multipoint operation.

The encoded video stream can be made more resilient to errors by judicious use of the specified encoding mechanisms. Examples include periodic intraframes or cyclic macroblock intracoding to improve error recovery, the use of the GSPARE and PSPARE fields for error control purposes, and the coding of intraframes with fixed bit length fields.

Video streams incorporating the above techniques can be decoded using standard H.261 decoders, so that the error control mechanisms can be generically referred to as H.261 compatible techniques. However the last two mechanisms require decoders which are aware of the extra error

control measures in the coded streams. Codecs equipped to exploit these extra features can be referred to as enhanced H.261 codecs.

Finally, it is possible to transcode H.261 streams into more robust streams using video algorithms designed for high BER operation. Transcoders are best used in situations where the overall telecommunications network can be partitioned into sub-networks with widely differing BER characteristics. Then the H.261 codecs would be used on the low BER segment, the robust codec would be used on the high BER segment, with the transcoder interfacing the two segments and facilitating interoperability between the two. This technique requires the design of a new codec and transcoder. The approach taken here is to use low complexity, low delay transcoders in conjunction with modified H.261 codecs which are stream incompatible with standard H.261 codecs.

5.4 Implementation of a robust H.261 decoder

The design and error resilience of a robust H.261 compatible codec are described in this section. The codec uses the BCH code to correct and detect errors. VLC codeword violations are also used to detect errors. The decoder has the option of using implicit error concealment inherent in the robust decoder or it may use explicit GOB level error concealment by discarding the current decoded GOB video data and displaying the previous frame data. The robust decoder recovers from errors using the implicit refresh mechanism.

5.4.1 Error Detection and Correction

The BCH (511,493,2) code is decoded using a simulated bounded distance decoder [Michelson83]. It corrects up to two codeword bits in error. The simulated decoder is based on syndrome computation and matching and is thus a faithful simulation of bounded distance decoding techniques such as the Meggit decoder or the Peterson algorithm [Michelson83, Lin83].

When more than two errors occur, the error pattern may or may not be detectable. A t bounded distance decoder tries to decode a codeword by introducing up to t bit corrections. If the codeword is not corrected, then it may have up to t more bits in error following decoding. The most probable cause of an error in a bounded distance decoder on a Binary Symmetric Channel (BSC) is the occurrence of a $t+1$ bit error pattern, and the most likely outcome, in the case of a decoding failure, is $2t+1$ decoded error bits, which is equal to the minimum distance of the code d. A transmission error is detected by the decoder when the error pattern syndrome is not equivalent to a decodable error pattern with t or less bits in error.

By assuming bounded distance decoding, some error correcting power of the BCH code is wasted. For example, maximum likelihood decoding can correct a further $(2^{n-k}-{}_nC_t)$ error patterns since there are 2^{n-k} possible syndromes but only ${}_nC_t$ t error patterns. Substituting n=511, k=493 and t=2

shows that a further 131839 error patterns can be corrected. However this only represents 0.589 % of all possible three error patterns so that the performance loss can be tolerated.

An examination of the video multiplex stream indicated that some errors in the fixed and variable length fields could be detected, so that field decoding failure could be also be used as an error indicator, in addition to the BCH code. The error trapping capabilities of all video multiplex stream fields are reported in appendix 3. It was found that most of the variable length codes employed are nearly complete. The codes appear to be modified Huffman codes with the all zero codeword removed and preceding zero runs reduced to prevent start code emulation. Thus their error detection capability is not very reliable and long false decoding runs are normal. Nevertheless, these code violations are used by the decoder to flag errors to augment the error detection using the BCH code.

5.4.2 Error Containment and Concealment

When an error is detected the decoder discards the corrupted codeword, stops decoding the current GOB and skips to the next GOB header. This is equivalent to *Implicit error concealment* because the decoder automatically displays the previous frame data in lieu of the skipped macroblocks.

The decoder may also take steps to try and conceal the error explicitly. Explicit GOB error concealment is optional in the robust decoder which was implemented. If the option is selected the decoder discards all the decoded data in the current GOB as soon as an error is detected, and displays the previous GOB data.

Error concealment is precipitated by both of the error detection mechanisms described above.

5.4.3 Error Recovery

Once an error has been detected the decoder must rely on the implicit forced updating feature defined in the standard to recover from the error. The decoder updates three macroblocks per frame in CIF mode and one macroblock per frame in QCIF mode. Thus a CIF picture is refreshed every 132 frames and a QCIF picture every 99 frames. The frame rates usually encountered at the bit rates of interest vary from 7.5 to 15 frames per second, so that a corrupted region will usually remain visible for some ten to twenty seconds.

5.5 Testing the Robust Decoder

The set-up used to test the robust decoder is shown in Figure 5-1 below. The video sequences are encoded and the results stored as H.261 streams. These are then read in by the IHC decoder and corrupted by X-OR addition with the output of a random bit error generator. The BER of the error generator is user programmable. The corrupted sequences are then decoded and the reconstructed frames are displayed and stored for further manipulation.

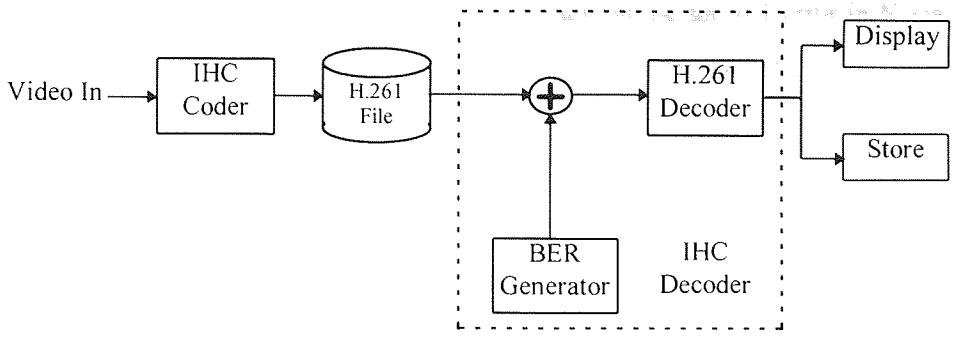


Figure 5-1 Error Control Simulation System

The error generation and control functions included in the IHC decoder are controlled via the "Simulate Errors Menu" shown in Figure 5-2 below. The first option can be used to test the temporal propagation of error macroblocks by setting the selected macroblock(s) to zero and tracing the diffusion of the error in subsequently decoded frames. The second option simulates the loss of one whole frame by skipping the selected frame.

SIMULATE ERRORS MENU

[1] Blank MB Row,Col	
[2] Skip Frame	1
[3] Error Mode	OFF
[4] Error Rate 1 in	1000
[5] Error File Name	error.file
[6] Error Packet Size	511
[7] Conceal	OFF
[8] Discard	ON
[9] BCH Mode	OFF
[a] Block Size N	511
[b] Word Size K	493
[c] Correct L	2
[d] Min. Dist. D.	5
[e] Protect Headers	OFF
[f] Fixed Len. DC Frame	OFF
[g] Transcoder Select	0
[h] Transcoder TrT	0
[i] Transcoder Zone	15

INPUT SELECTION

Figure 5-2 Error Simulation Menu

Options [3] to [6] control the error generation processes. Option [3] selects the error generator. Error generation can be OFF, RANDOM, or EXTERNAL. RANDOM turns on the internal random

error generator. The error rate of the internal error generator can be set to 1 error in N via option [4]. The EXTERNAL option reads the error vector from the file indicated via option [5]. The error file can be read in PS bits at a time, with the packet size PS selected using option [6].

Option [7] controls the error concealment function. When set, detected errors cause the IHC decoder to reject all the data pertaining to the current GOB being decoded and to display the previous GOB video data. If reset, implicit concealment is used as described above. When Option [8] is set to ON corrupted codewords are discarded, otherwise the IHC decoder tries to use the corrupted data.

Options [9] to [d] control the BCH decoder. Option [9] sets the operating mode of the BCH decoder. The BCH mode can be OFF, such that the BCH code is ignored; DETECT, which instructs the IHC decoder to use the BCH code to detect errors; or CORRECT, which instructs the decoder to correct up to T errors per N block codeword.

The codeword length N, the information word length K, the error correction capability L and the minimum distance of the code D (where known) of the BCH code are programmable via options [a]-[d]. The generic decoder will correct up to L errors and simultaneously detect up to D-L-1 errors. If (N,K,L) are set to (511,493,2) the simulation uses the exact bounded distance decoder for this code.

The remaining options [e] and [i] represent decoder enhancements and are described in later sections.

5.5.1 Performance measures

The resilience of the decoder was tested by corrupting the sequences, decoding them and measuring the reproduction quality using objective measures. The decoded sequences were also assessed subjectively and a short description is included for each decoded sequence. A five grade *usefulness* scale has been defined to assign a subjective usefulness quality measure to each sequence and hence summarise the subjective descriptions and allow comparisons.

Four objective parameters were chosen to characterise the performance of the robust decoders. The first parameter is the residual BER per frame, which is the number of bits in error per frame after bounded distance error correction using the BCH code. This parameter is used to relate the number of residual errors per frame to the error degradation in the sequence

The second parameter is the colour pixel Mean Square Error (MSE) which is the square of the colour pixel difference between a correctly decoded frame and the same, incorrectly decoded, frame. MSE is given by:-

$$MSE = \frac{1}{3N_C N_R} \sum_{i=1}^{N_c} \sum_{j=1}^{N_r} \sum_{k=0}^{2} \left(C_{i,j,k} - \hat{C}_{i,j,k} \right)^2 \qquad \text{(5-3)}$$

where N_c =352 and N_r =288 for CIF sequences and N_c= 176 and N_r= 144 for QCIF sequences. $C_{i,j,k}$ is the pixel in the kth colour plane, in the ith column and jth row and $\hat{C}_{i,j,k}$ is the corresponding pixel in the corrupted frame. There are three planes {R,G,B} corresponding to k={0,1,2}. The MSE gives an

indication of the quality of the frame. Ideally, the lower the average MSE for the sequence, the higher the usefulness of the sequence.

The third parameter is the number of corrupted macroblocks per frame, irrespective of whether the error is visible or not. This measure gives an indication of the spatial extent of the errors within a frame.

The fourth parameter centres around the concept of a "macroblock error pipe" (MEP) which begins existence on the occurrence of the first error within a macroblock and terminates when the macroblock is error free again. The measured parameter is the MEP length which is a direct measure of the temporal macroblock error propagation.

Since the objective and subjective assessments do not always agree [Netravali88], an informal subjective assessment of the corrupted sequences was also undertaken. A five point subjective usefulness scale was selected to grade the usefulness of the corrupted sequences as shown in Table 5-1. The grades range from five, for visibly un-corrupted reproduction, down to one where the errors propagate catastrophically and render the sequence unrecognisable. Grade four sequences contain a few, often persistent macroblock errors, but the sequence is perfectly recognisable and useful. Grade 3 is assigned to sequences which contain many corrupted macroblocks but which may still be defined as useful, although the errors are highly visible. Grade two is assigned to those sequences which although not useful for meaningful communications still retain some notion of image content. The qualifiers + and - are also used below to indicate slightly better and slightly worse subjective quality.

GRADE	MEANING
5	No errors visible
4	Useful reproduction with only a few visible errors.
3	Useful reproduction but with many visible errors
2	Un-useful - errors over most of the frame
1	Catastrophic sequence breakdown

Table 5-1 Five point Subjective Usefulness Scale

5.5.2 Results

Three FBR streams were generated to test the resilience of the robust decoder. Forced updating was enforced at the rate of three macroblocks every frame. The generated streams were:-

- QM48 QCIF colour Miss America sequence coded at 12.5 fps and 48 kbit/s.
- MA96 CIF colour Miss America sequence coded at 12.5 fps and 96 kbit/s.
- SM384 CIF colour Salesman sequence coded at 10 fps and 384 kbit/s.

Four runs were executed for each of the sequences generating two sets of results, one set conducted at a BER of 10^{-3} and the other at a BER of 10^{-4}. Each set consisted of a run without error correction and implicit concealment and another with error correction and GOB explicit concealment.

When the sequences were decoded at a BER of 10^{-4} with error correction applied, none of the reconstructed sequences experienced reconstruction errors. The sequences were then decoded at this bit rate with only error detection enabled. Therefore three results are presented per sequence. The sequences are typical videophone and videoconference sequences, at bit rates compatible with good quality video services using the H.261 codec. The bit rate per frame increases from QM48 to MA96 to SM384 to investigate the effect of increasing the bit rate per frame at the same bit error rate.

At a BER of 10^{-3} without error correction all the three tested sequences suffered rapid and catastrophic picture quality deterioration. The first frames of the sequences were also severely in error; that of the QM48 sequence suffered one completely corrupted GOB; the MA96 sequence suffered three GOB wipe-outs; whereas all the GOBs in the first frame of the SM384 sequence were corrupted.

5.5.2.1 QM48 stream

The pixel mean square error, the residual frame BER, the corrupted macroblocks per frame and the error pipe length histogram are plotted in Figure 5-3 (a) to (b). The sequences averages for the three runs A, B and C are tabulated in Table 5-2. The stream A results were obtained by decoding the colour QCIF Miss America sequence, coded at 48 kbit/s, using the robust decoder with implicit concealment and a BER of 10^{-3}. The B stream results were obtained using explicit concealment at the same BER. The third result C, was obtained using the BCH code to detect errors only, at a BER of 10^{-4}.

	QCIF Miss America	Residual BER	Average MSE	Average MBE	Average EPL	SUR
A	10^{-3} FEC	$1.12\text{x}10^{-4}$	21.8	33.8	33.3	5-
B	10^{-3} FEC+Conc	$1.12\text{x}10^{-4}$	31.5	45.6	40.0	5-
C	10^{-4}	$1.00\text{x}10^{-4}$	94.7	47.4	31.4	3

Table 5-2 QM48 Stream Results

5.5.2.1.1 Subjective Assessment

When the QM48 sequence was decoded with forward error correction and implicit error concealment there were no visible errors most of the time. The 65 frame sequence for which the results are plotted in Figure 5-3 and summarised in Table 5-2, did not suffer any visible errors apart from some slight increase in blockiness. The Subjective Usefulness Rating (SUR) for this sequence is thus 5-, indicating a sequence with barely any visible degradations.

When the same sequence was decoded using explicit concealment the results were the same. although there was an impression that this sequence suffered a slight loss of resolution with respect to stream A. Stream B was, however, also given a SUR score of 5-.

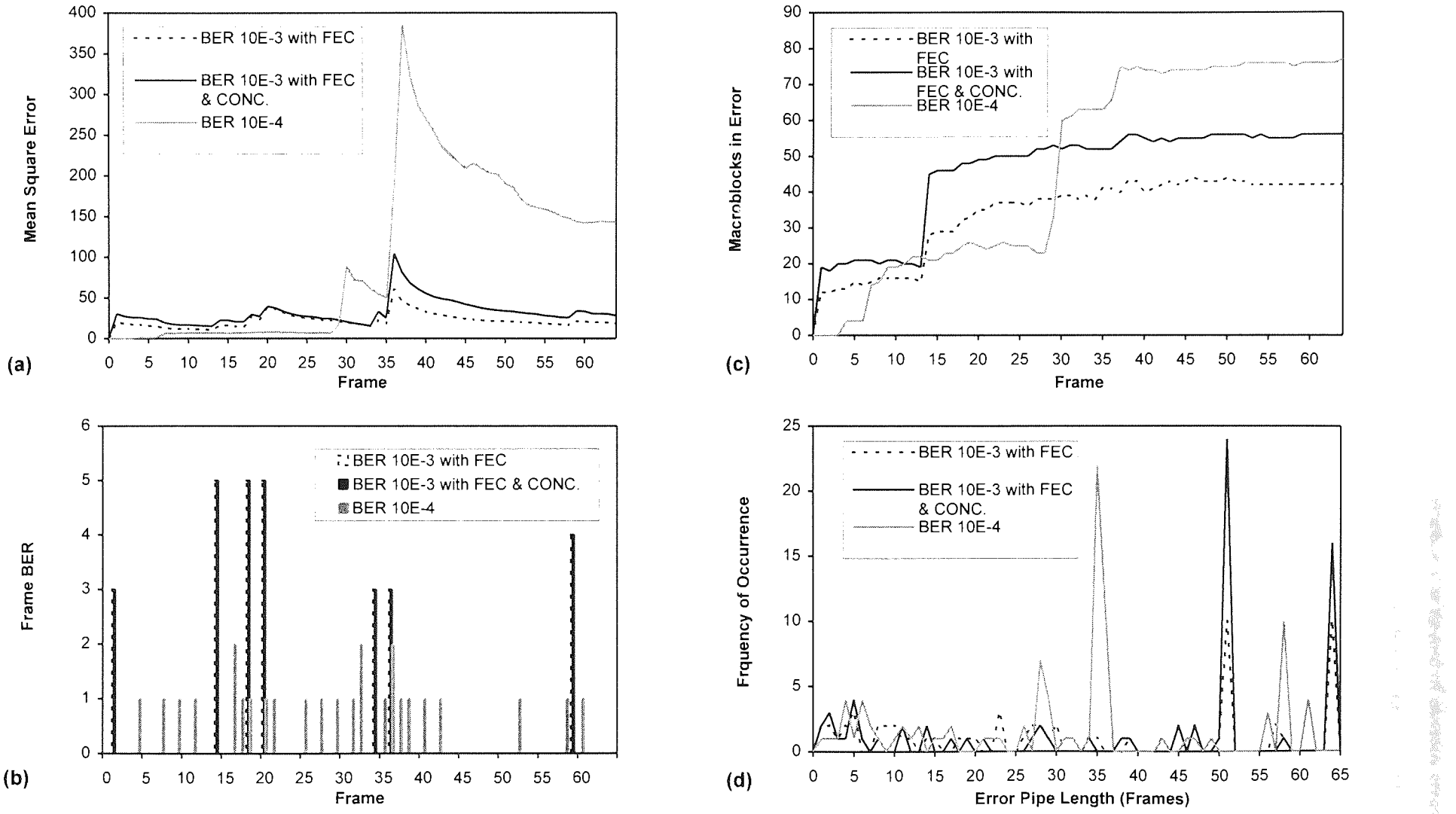


Figure 5.3 Error Statistics for the colour QM48 sequence coded at 48 kbit/s and corrupted with random errors at the rate of 10^{-3} and 10^{-4}, and decoded using the robust H.261 decoder; without the use of FEC, with the use of FEC and with the use of FEC and GOB Concealment. (a) The Frame Square Error, (b) The residual Frame BER, (c) the corrupted Macroblocks per Frame and (d) the Error Pipe Length histogram.

At a BER of 10^{-4} without error correction the sequence was devoid of obvious errors until the subject moved her head to the right, whereupon the picture quality deteriorated abruptly. However there was a perceptible error recovery as the areas in motion appeared to improve quickly after being corrupted.

5.5.2.1.2 Objective Results

The MSE of the decoded streams is shown in Figure 5-3 (a). The discontinuities in MSE occur at error events plotted in Figure 5-10 (b). Not all errors cause visible degradations. In fact streams A and B suffered no obvious errors. The large jump in stream C is due to a skipped frame, when an error caused the loss of a picture header. Curves A and B with implicit and explicit concealment have very similar MSE curves as expected. The MSE with explicit GOB concealment is worse than with implicit concealment.

Figure 5-3 (b) reveals that the BER at the output of the forward error correction module is bursty in nature. The errors with five bits correspond to one codeword error with three bits, having been incorrectly decoded into a close codeword at a Hamming distance of 5 bits from the correct codeword. The remaining four errors correspond to detected errors with three and four codeword bits in error. The un-corrected bit errors at a BER of 10^{-4} cause random codeword errors in the frames.

The Macroblocks in Error (MBE) curves in Figure 5-3 (c) indicate the strong temporal error propagation in the three decoded sequences. The number of contaminated macroblocks increases gradually with abrupt discontinuities at bit error events, until a large percentage of all the macroblocks are in error. Any reduction in the number of corrupted MBs is very gradual and short-lived. The errors spread quickly to most macroblocks even though there are very few corrupted codewords.

The Error Pipe Length (EPL) histogram reveals that there are a few short lived error pipes, but the sequences are characterised by long EPL's. This shows that the implicit updating used with the three sequences is not suitable for the curtailment of temporal error propagation.

5.5.2.1.3 Discussion

The BCH (511,493,2) FEC code in H.261 allows useful QCIF sequences even at a BER of 10^{-3}. Both implicit and explicit concealment prevented any obvious errors, and the strategy to discard corrupted codewords was successful. Explicit GOB concealment suffered from slightly more blockiness than with implicit concealment, also reflected in the MSE curves. The sequence can be decoded usefully at a BER of 10^{-4} without error correction but the picture quality is not very good.

Table 5-2 shows that A has the lowest average MSE, followed by stream B and C. Streams A and B have much lower average MSEs than stream C despite it having a slightly higher residual BER. This is due to the large error caused by a skipped frame in stream C. This is also reflected in the average MBE result but not in the average EPL result where stream C has the lowest average EPL. This is due to most MB errors occurring after frame 35 in stream C.

The usefulness rating shows good correlation with the average MSE result. However this cannot be extended to other sequences, and maybe not even to the same sequence with other error patterns.

The large discontinuities in the MSE for curve C led to very visible macroblock impairments. The errors coincided with the onset of the dynamic part of the sequence and subsequently there is a gradual decrease of MSE which corresponds to the reported improvement in picture quality above. This would seem to indicate that a more dynamic sequence may have better error recovery properties.

5.5.2.2 MA96 stream

The pixel mean square error, the residual frame BER, the number of corrupted macroblocks per frame and the error pipe length histogram are plotted in Figure 5-4 (a) to (b). The sequences averages for the three runs A, B and C are tabulated in Table 5-3. The stream A results were obtained by decoding the colour CIF Miss America sequence, coded at 96 kbit/s, using the robust decoder with implicit concealment and a BER of 10^{-3}. The B stream results were obtained using explicit concealment at the same BER. The third result C, was obtained using the BCH code to detect errors only, at a BER of 10^{-4}.

	CIF Miss America	Residual BER	Average MSE	Average MBE	Average EPL	SUR
A	10^{-3} FEC	$1.18\text{x}10^{-4}$	247	148	29.1	3
B	10^{-3} FEC+Conc	$1.18\text{x}10^{-4}$	260	167	30.2	3
C	10^{-4}	$1.00\text{x}10^{-4}$	112	238	32.1	3

Table 5-3 MA96 stream results

5.5.2.2.1 Subjective Assessment

The decoding of the MA96 sequence with FEC and implicit concealment at a BER of 10^{-3} suffered from a partial GOB wipe-out in the first frame. There was slow but persistent recovery. The image also suffered from very visible corrupted macroblocks with very coloured, chequered patterns, caused by incorrectly decoded colour blocks. However the sequence was useful though substantially corrupted and its quality can be summarised by a SUR score of 3.

When the sequence was decoded with explicit concealment, the first frame suffered a complete GOB wipe-out as the concealment algorithm discarded the macroblocks decoded correctly prior to the error. There was, however persistent recovery as in the previous case. Although the image was slightly more 'blocky' then stream A, there were no visible 'chequers'. The B stream SUR is 3.

Without error correction at a BER of 10^{-4} the decoded sequence suffered visible, annoying but non catastrophic errors. Some error recovery was perceptible and the sequence remained useable even during the subject's brisk motion of the head and shoulders. The C stream SUR was 3 again.

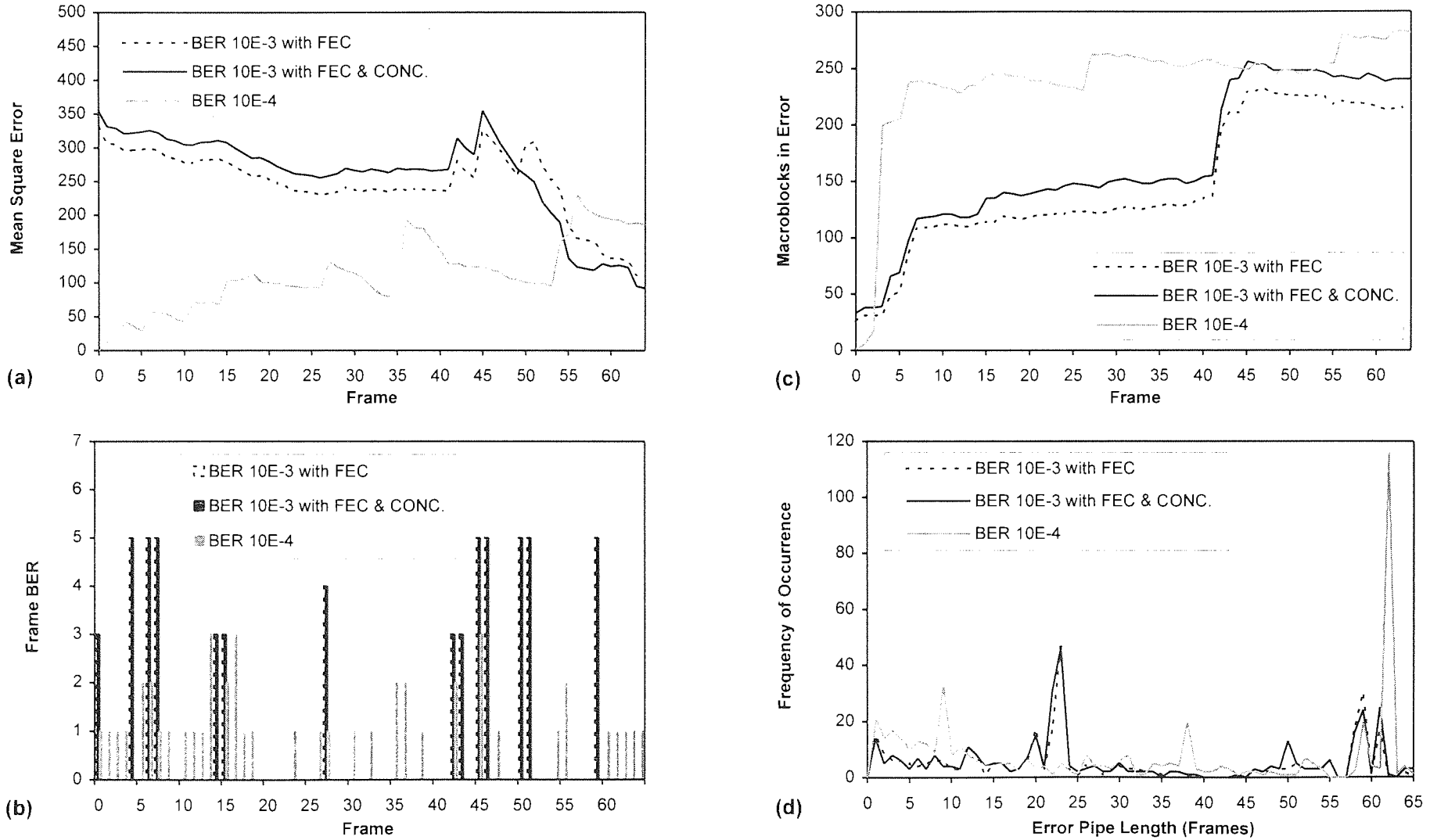


Figure 5.4 Error statistics for the colour MA96 sequence coded at 96 kbit/s and corrupted with random errors at the rate of 10^{-3} and 10^{-4}, and decoded using the robust H.261 decoder without the use of FEC, with the use of FEC and with FEC and GOB Concealment. (a) The Mean Square Error (b) The Residual BER per frame,(c) the Corrupted Macroblocks per frame and (d) The Error Pipe Lengths histogram.

5.5.2.2.2 Objective Assessment

The MSE curves for the MA96 sequence, when decoded at a BER of 10^{-3} with FEC and concealment, show in Figure 5-4 (a) that the first frame suffered from a high MSE, because of the corrupted GOB. The MSE curve B with explicit concealment is initially higher than for curve A because of the complete GOB wipe-out, and remains so for the first 50 frames, after which it dips below curve A as GOB concealment prevents high MSE 'chequers'. There is gradual MSE reduction as the sequence recovers form the GOB error. The improvement accelerates in the third, more dynamic part of the sequence. The MSE curve for the decoded stream C starts off with zero MSE (no GOB error in frame one) and then increases in jumps at each error event. The peaks at frames 3, 6, 27 and 56 are caused by skipped frames.

The residual frame errors in Figure 5-4 (b) follow the same pattern as that for the previous sequence QM48. There are approximately twice the number of residual frame errors, however, reflecting the doubling of the bit rate per frame from the QM48 sequence to the MA96 sequence.

The MBE curves in Figure 5-4 (c) again reveal the persistent spatial and temporal error propagation which very steadily spreads to corrupt most of the macroblocks. There is some gradual reduction in corrupted macroblocks between error events, but this is very gradual. The MBE curve B remains above curve A for the whole duration of the sequence. The discontinuity in these MBE curves at frame 42 is due to the error bursts at this point. The MBE for curve C discontinuities to 200 corrupted macroblocks out of 396 following the first skipped frame and reaches 75% corrupted MBs by the end of the sequence.

The EPL histogram curves in Figure 5-4 (d) show the presence of many long error pipes. The peaks on the three curves correspond to the distance of error events from the end of the sequence. Thus the first skipped frames in stream C after frame 3 lead to around 120 error pipes which survive till the end of the sequence 62 frames later.

5.5.2.2.3 Discussion

Despite the initial GOB wipe-out the image is graded as useful, because the corrupted area is small relative to the whole image. A one GOB error in a QCIF sequence is more serious since it leads to the loss of one third of the picture.

GOB Concealment removes the obvious, rather distracting errors, but results in a slightly poorer overall reproduction quality. This happens because GOB concealment discards correctly decoded MBs and hence increases the MSE.

Comparing the average MSE and the SUR scores in Table 5-3 it is clear that whereas streams A and B have the same MSE and SUR results, stream C has the same SUR with a lower MSE. This reveals the problem with using objective measures to predict subjective issues like 'usefulness'; in this case differing MSE results correspond to the same usefulness measure of 3. For this sequence the SUR

is well correlated with the residual BER, which is approximately the same for the three sequences at around 10^{-4}. The Average MBE is again shown to be a bad predictor for sequence 'usefulness'.

5.5.2.3 SM384 stream

The pixel mean square error, the residual frame BER, the number of corrupted macroblocks per frame and the error pipe length histogram are plotted in Figure 5-5 (a) to (b). The sequences averages for the three runs A, B and C are tabulated in Table 5-4. The stream A results were obtained by decoding the colour CIF Salesman sequence, coded at 384 kbit/s, using the robust decoder with implicit concealment and a BER of 10^{-3}. The B stream results were obtained using explicit concealment at the same BER. The third result C, was obtained using the BCH code to detect errors only, at a BER of 10^{-4}.

5.5.2.3.1 Subjective Assessment

The first frame in the decoded sequence A suffered three partial GOB wipe-outs. This error persisted for most of the sequence, but was practically erased over the whole sequence. The image sequence suffered from coloured 'chequers' and image break-up in regions of high dynamicity. The sequence SUR is 3- because despite the obvious, serious, impairments the image content was more than recognisable at all times.

	CIF Salesman Sequence	Residual BER	Average MSE	Average MBE	Average EPL	SUR
A	10^{-3} FEC	1.05×10^{-4}	589	309	50.47	3-
B	10^{-3} FEC+Conc	1.05×10^{-4}	937	338	53.75	3-
C	10^{-4}	1.03×10^{-4}	809	299	56.10	2+

Table 5-4 SM384 stream results

The subjective assessment of the decoded sequence B, was similar to that for the decoded sequence A. The first frame suffered three GOB wipe-outs which again recovered by the end of the sequence. However the coloured 'chequers' were all but absent for the whole sequence, although the image break-ups where slightly worse than for decoded sequence A. Thus, the sequence receives the same usefulness quality rating -3-.

The subjective quality without error correction at 10^{-4} was roughly equivalent to that with error correction at 10^{-3}, though slightly worse. Again the first frame contained three GOB wipe-outs, with other visible 'chequers'. The error recovery from the initial GOB errors was very slow but the sequence bordered on the useful. The usefulness rating is 2+; it is worse than the other two sequences but still useful.

5.5.2.3.2 Objective Assessment

The initial frame MSE in Figure 5-5 (a) is about three times higher than that in Figure 5-4 (a) as expected (three corrupted GOBs rather than one). The MSE curves show very little error recovery

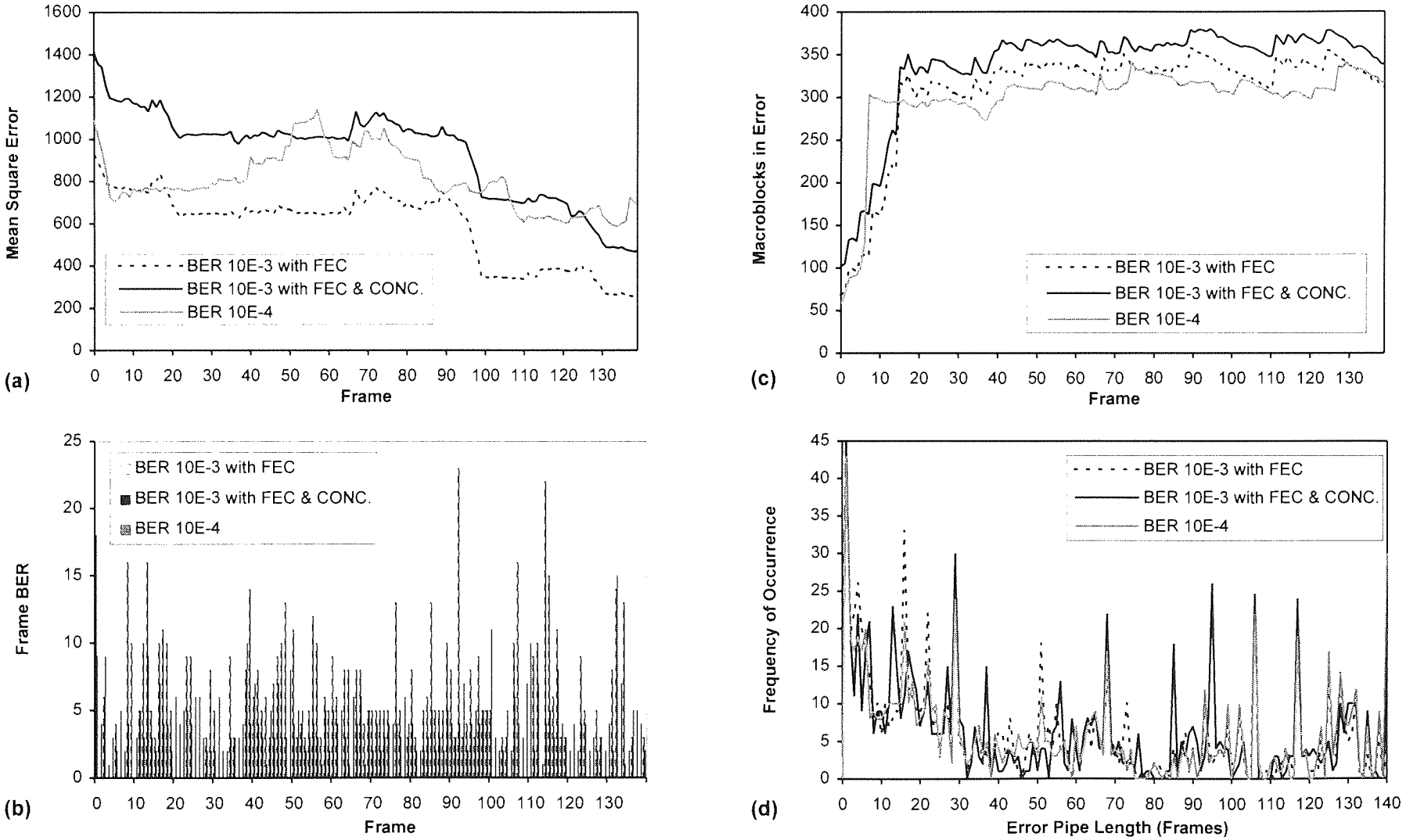


Figure 5.5 Error statistics for the colour SM384 sequence coded at 384 kbit/s and corrupted with random errors at the rate of 10^{-3} and 10^{-4}, and decoded using the robust H.261 decoder without the use of FEC, with the use of FEC and with FEC and Concealment. (a) The Frame Square Error (b) The Residual BER per frame,(c) the Corrupted Macroblocks per frame and (d) The Histogram of Error Pipe Lengths.

for the three decoded sequences, except for sequences A and B which show a dip in MSE after the recovery of the last corrupted GOB. The MSE for sequence A is again below that of the explicitly concealed sequence B.

The fivefold increase in bits per frame over the MA96 sequence is reflected in the increased bit rate per frame in Figure 5-5 (b). Most of the frames contain errors in this situation, often multiple codeword errors. The post-correction error events are still bursty with respect to the random errors at a BER of 10^{-4}. The corrupted frames, however, are much more randomly distributed than before.

More than 75% of the macroblocks are corrupted after the first twenty frames as evident in Figure 5-5 (c); yet the sequence is still useful, although it is highly impaired. The explicitly concealed sequence has the highest corrupted macroblocks most of the time.

The EPL distribution in Figure 5-5 9(d) confirms that temporal error propagation is unchecked as in the other two sequences.

5.5.2.3.3 Discussion

The averaged results in Table 5-4 again reveal that the residual BER for decoded sequences A and B is slightly above 10^{-4}. The decoded sequence with explicit GOB concealment again fare better in the subjective assessment than in the MSE assessment since the MSE for the B sequence is nearly twice that of the A sequence. Average MSE does not grade the sequences in the correct usefulness scale.

The higher usefulness rating of the corrected sequences, despite a higher residual BER than the un-corrected sequence C may be due to the fact that the errors effect less frames in the former case than in the latter, and point to the possibility that the major factor in determining image 'usefulness' is the number of errors per frame at these rates.

The Average MBE and EPL results show that the three decoded sequences suffer from acute spatial and temporal error propagation.

5.5.2.4 Discussion of results

The main result of the above tests is that the BCH forward error correcting code renders the sequences useful even at a BER of 10^{-3} , although some image degradation is still evident. The sequences do not suffer catastrophic failure and remain stable for tens of seconds.

The sequences are still useful at a BER of 10^{-4} with error detection and concealment, although on the random error channel the natural choice would be to use forward error correction.

The subjective results indicate a dual behaviour of explicit GOB error concealment. In some cases the technique is detrimental to picture quality because correctly decoded data is discarded. In other cases the picture quality improves if the discarded GOBs contain very visible impairments. From an MSE point of view the implicit concealment performs consistently better than explicit GOB concealment. It is clear, however, that error concealment can play a major role in error control

techniques, but it should be targeted at regions smaller than GOBs; at macroblock or even at block level.

The results indicate that the reconstructed sequence quality deteriorates and sequence usefulness decreases, with an increase in the bits per frame at the same video frame rate, since the average corrupted bits-per-frame are then higher at the same BER.

The results indicate that gross MSE difference is a good predictor of quality. Thus the average MSE increases from the QM48, through MA96 to SM384, and the usefulness rating decreases as expected. When the BER differences are small, however, and considering the same sequences, the MSE is not a good predictor of absolute usefulness order.

However the MSE is much higher than in the QM48 and MA96 sequences. This is due to the higher bit rate per frame, of five and ten times higher for the SM384 sequence than the MA96 and QM48 sequences respectively.

The tests conducted at a BER of 10^{-4} with forward error correction resulted in no residual BER. At a BER of 10^{-3} the sequences are still useful. Tests conducted at a BER of 10^{-2} using the BCH (511,493,2) forward error correction module revealed that the sequence is unusable. Further tests showed that rapid sequence deterioration sets in above a BER of $2x10^{-3}$.

5.6 Improvements

Two major impairments were evident during the testing of the robust decoders. The first was due to corruption of the Picture or GOB headers which cause the loss of one whole frame or one GOB of data. The GOB loss is more evident in QCIF sequences since there are only three GOBs in a frame. The second impairment is due to the high probability of error in the first frame. The first frame is intracoded and generates a bit rate which is 3 to 5 times higher than the intercoded frames. Thus there is more probability of an error occurring in this frame. When an error occurs it is highly visible since it cannot be concealed using previous frame data, and it tends to persist for a very long time such that the visibility of the impairments are severe. The resilience of the robust codecs to these two impairments can be improved as described below.

5.6.1 Protecting the Picture and GOB headers

The picture and GOB header fields can be protected using the PSPARE and GSPARE fields. The picture header field is terminated by a picture header extension indicator (PEI) bit. If this bit is high, a PSPARE field consisting of eight bits can be added followed by another PEI bit. Other PSPARE fields can be similarly attached to the end of the PH field. The last PSPARE field is terminated by a low PEI bit. An identical mechanism is used to insert GSPARE octet fields in the group of blocks header.

Currently H261 does not specify the use of the PSPARE and GSPARE and recommends that decoders should discard the two fields if the PEI and GEI bits are high. Thus if these spare fields are employed, current codecs would discard these fields and hence decode the video correctly. However it must be stressed that future amendments to the H261 standard involving the use of these two fields will cause streams using the proposed mechanisms to be incompatible with the new standard.

The proposed header protection mechanism takes the form of a BCH code applied to the headers with the parity bits occupying the PSPARE and GSPARE bits. Most of the header start code bits (0000 0000 0000 0001) and the GEI and PEI bits are known beforehand such that errors in these bits are erasures and hence the error correcting capability of the blockcode is enhanced. The main challenge is synchronisation of corrupted headers. This can be achieved by continuously performing a hamming distance measure on the known bit pattern and attempting decoding when the hamming distance in less than or equal to the error and erasure correcting capability of the code.

Three possible coded picture header (PH) structures employing shortened binary BCH (n-l,k-l,t) block codes are shown below.

C1 0000 0000 0000 0001 0000 NNNNN TTTTTT 1 SSPP PPPP 0
C2 0000 0000 0000 0001 0000 NNNNN TTTTTT 1 SSSS PPPP 1 PPPP PPPP 0
C3 0000 0000 0000 0001 0000 NNNNN TTTTTT 1 PPPP PPPP 1 PPPP PPPP 1 PPPP PPPP 0

The BCH codes employed are the (63,57,1) code shortened to (41,35,1), the (63,51,2) code shortened to (50,38,2) and the (63,39,4) code shortened to (59,35,4). The bits indicated as 1 and 0 have known values, N and T bits are indeterminate information bits, S are spare bits that can be set to known patterns and P are the parity bits.

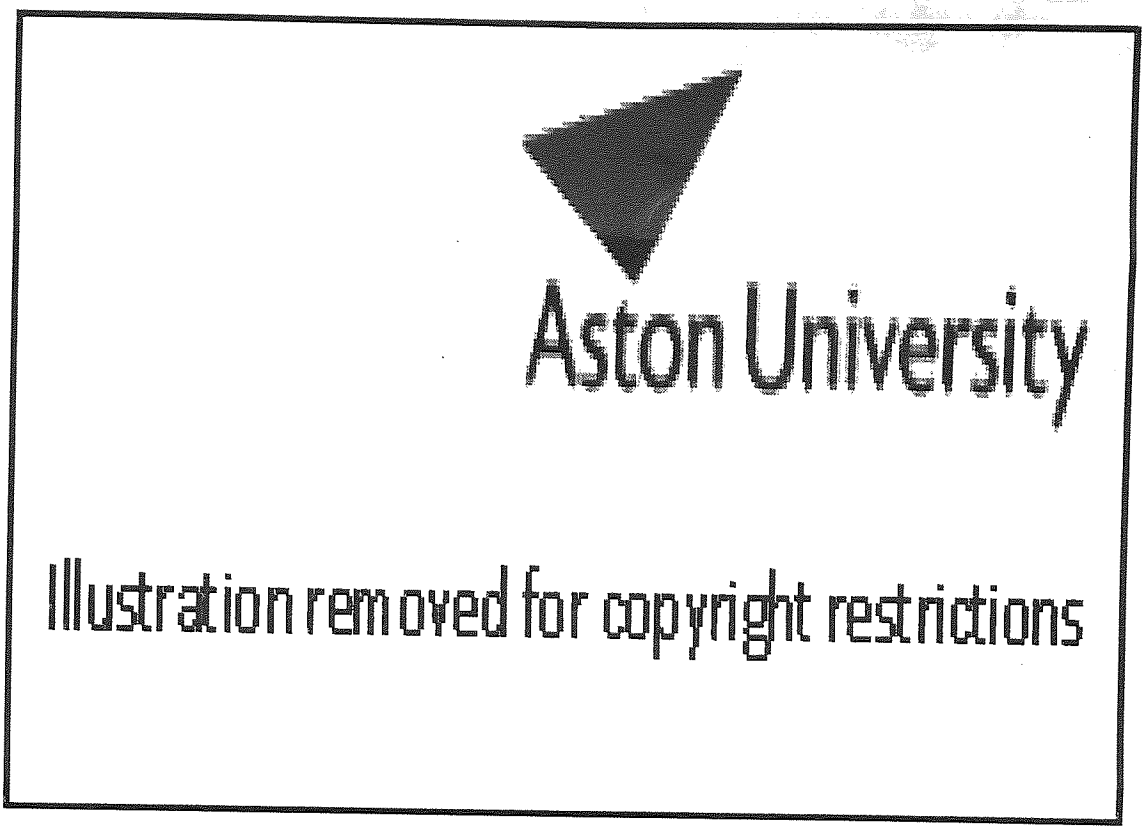
The use of these codes can be illustrated by describing the use of the C1 code. The C1 code is a shortened BCH (63,57,1) code. Thus it also has at least minimum distance 3 [Michelson83] and error correcting capability t=1. Consider the search pattern given by :-

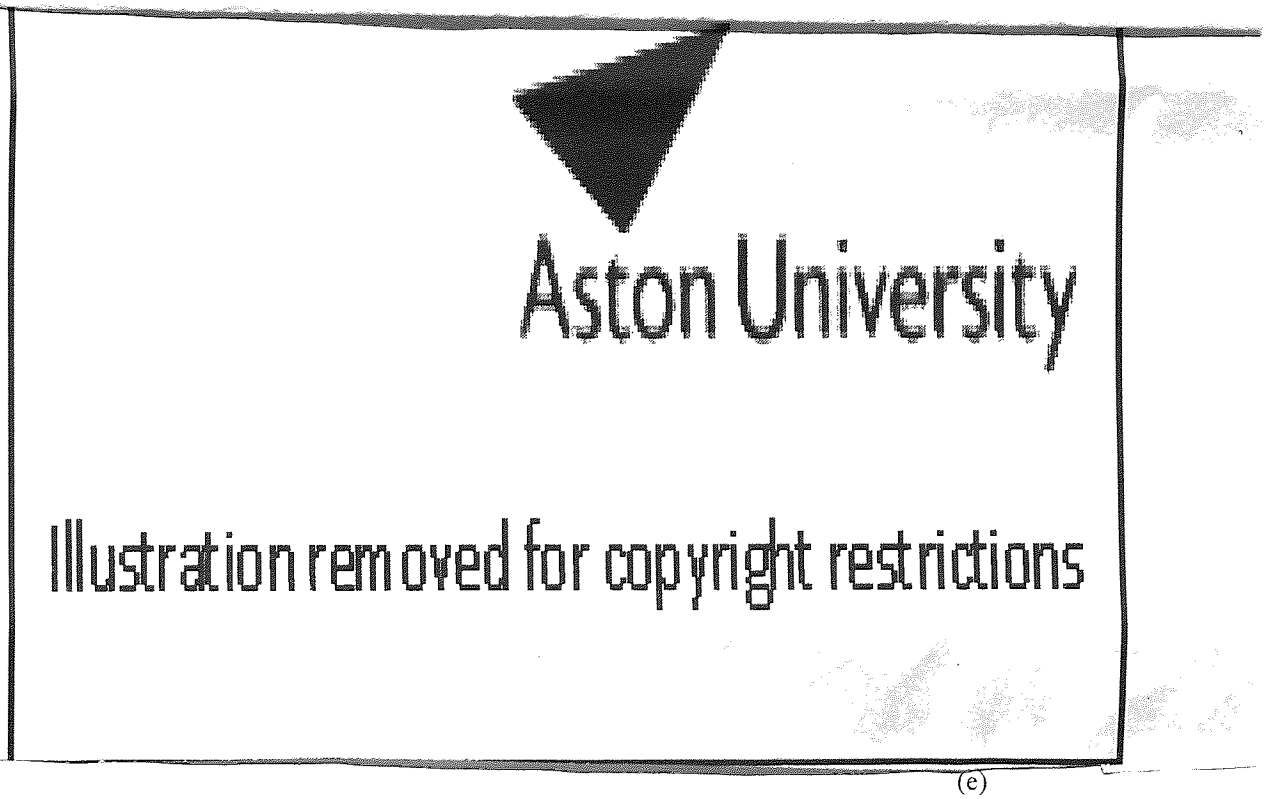
[0000 0000 0000 0001 0000 XXXXX XXXXXX 1 11XX XXXX 0]

Then if during synchronisation the hamming weight of the search pattern is zero the syndrome is computed. If it is equal to zero, then the PH is assumed to be correct, otherwise the error correction procedure is initiated to correct the error if possible. If the error cannot be corrected it is necessary to initiate an error recovery procedure. If during synchronisation the Hamming distance is one, the 'erroneous' bit is reversed and the PH accepted if the syndrome result is zero indicating one error in the known bit pattern. Otherwise error correction is initiated

5.6.2 Fixed field length DC intraframes.

The first frame in the sequence must be intracoded. Usually an intracoded frame generates two to five times the bit rate of intercoded frames. Thus the probability of corrupted data is substantially





(e)

Figure 5.6 (a) Second frame of the SM sequence VBR coded with Q=8. First frame coded in normal Intramode **(b)** Second frame of SM sequence VBR coded with Q=8. First frame coded with fixed length DC Intraframe **(c)** Same as case (a) but decoded with with a BER of 10^{-3}. **(d)** Same as case (b) but decoded using DC Intramode decoder with a BER of 10^{-3}. **(e)** Same as case (b) but decoded with a normal decoder with a BER of 10-3.

higher. Unfortunately error concealment in this particular case results in highly visible corrupted areas since there is no preceding GOB data with which to replace the missing data. Thus it may be better to switch off error concealment for the first frame. However the error visibility may still be substantial. The fixed field length DC intraframe is proposed to allow macroblock concealment in this first essential intracoded frame.

It is noted that if the encoder codes all blocks in intramode using only the DC coefficient, then all the VLC fields are in fact of fixed length: the picture headers and the GOB headers are always fixed length, the macroblock header is the same for all block '10001' indicating an MBA of +1 and an intracode MTYPE, and the DC coefficient blocks are coded with 8 coefficient bits and two EOB bits. It is thus trivial to design a decoder to decode such a frame without losing synchronisation whatever the error patterns encountered. Then errors affecting the luminance blocks are confined to 8x8 blocks whereas chrominance errors affect 16x16 pixel macroblocks.

Test results showed that the objective and subjective quality of DC intraframes is substantially worse than the same frames intracoded with Q=31, because even in this case not all AC coefficients quantize to zero. However the degradation is for a couple of frames only as the quality is recovered quickly in subsequent intercoded frames. It was found that the number of bits required by the DC intraframe scheme when the quality has recovered to normal levels is approximately the same as in the non DC intraframe sequence.

5.7 Error recovery using periodic frame Intracoding

Temporal error propagation is halted by periodic intracoded frames, although the intracoded data may itself be corrupted, thus merely leading to a change in displayed error patterns. This is a significant consideration: intracoded frames require three to five times the bit rate of intercoded frames and thus the chances are much higher of an error corrupting an intracoded frame. However it was found that by intracoding one frame out of ten, the situation was improved somewhat over the non periodically refreshed case. The bit rate fluctuations are very high, however, and very difficult to smooth out. There is also a marked loss in temporal resolution which is very noticeable.

The periodic frame intracoding is encountered in MPEG where the periodic intraframes are required to provided random access points in the video sequence, and facilitate fast forward and reverse searches. MPEG-I buffer control algorithms usually span group of pictures from one I frame to the next. The problem of increased rate has been solved be introducing highly efficient bi-directionally predicted frames which predict the current image using both past and future frames. However this introduces a substantial delay which is not tolerable in real time audio-visual communications. MPEG-1 pays more attention to intracoded frames than H.261 since I frames are much more common in the former. Thus better video multiplex coding and differential coding of the highly correlated DC coefficients led to a substantial improvement in compression of I frames. Thus periodic intracoding is

naturally included in MPEG-1 streams but would not be viable in H.261. MPEG-1 sequences thus have a much better error recovery time than H.261.

5.8 Error recovery using forced updating.

Rather than intracode a whole frame it may be possible to distribute the intracode burden over the entire sequence. It has already been indicated that the IHC forces updates at three macroblocks per frame by intracoding. Thus a frame is intracoded completely every 132 transmitted frames, which limits temporal error propagation and visibility. Thus it should be possible to improve error recovery by increasing the forced updating rate.

The force updating generator in the IHC is capable of force updating a programmable number of macroblocks per frame. Two update patterns have been defined. The first codes N macroblocks in GOB and MB order. The second pattern updates N macroblocks per GOB per frame in MB order. The second update pattern was found to be better suited to the IHC GOB level rate control algorithm since the same number of macroblocks are intracoded per GOB and hence the rate control algorithm can easily predict the forced updating burden from the previous coded frame. Hence the second forced updating pattern interferes less with the rate control algorithm which can be used without modifications.

With FBR operation, the extra force-updated blocks consume bits which would have otherwise gone to code MBs in intercode, and since intercoded blocks use bits much more efficiently, then there is the danger that the sequence quality degrades considerably if the number of force updated blocks is allowed to increase too much. Thus force updating leads to a decrease of PSNR with FBR coding. Force updating three macroblocks every GOB every frame, was found to be a good compromise between the error recovery rate and the decrease in quality which was marginal both subjectively and objectively. If the technique is used with VBR coding, the PSNR remains constant, but the bit rate increases monotonically with the number of intracoded blocks.

To test the error recovery potential of this technique the four sequences tested above were recoded, force updating three MBs every GOB. The sequences were then subjected to a BER of 10^{-3} and decoded using error correction but no concealment. The results with and without forced updating are compared in Figure 5-7 to Figure 5-9. The force updated MSE curves in Figure 5-7 (a) to Figure 5-9 (a) are seen to recover very quickly following an error spike, within 11 frames or less. The number of corrupted macroblocks also recover though the recovery is seldom complete.

The error recovery potential of this technique is documented in Figure 5-7 (d) to Figure 5-9.(d) where the error pipe histograms are compared. Very few error pipe lengths exceed 11 frames in length. This is as expected since the frame is refreshed every 11 frames. Note that neither header protection nor fixed length DC intrafield were used to generate these results, to study force updating in isolation.

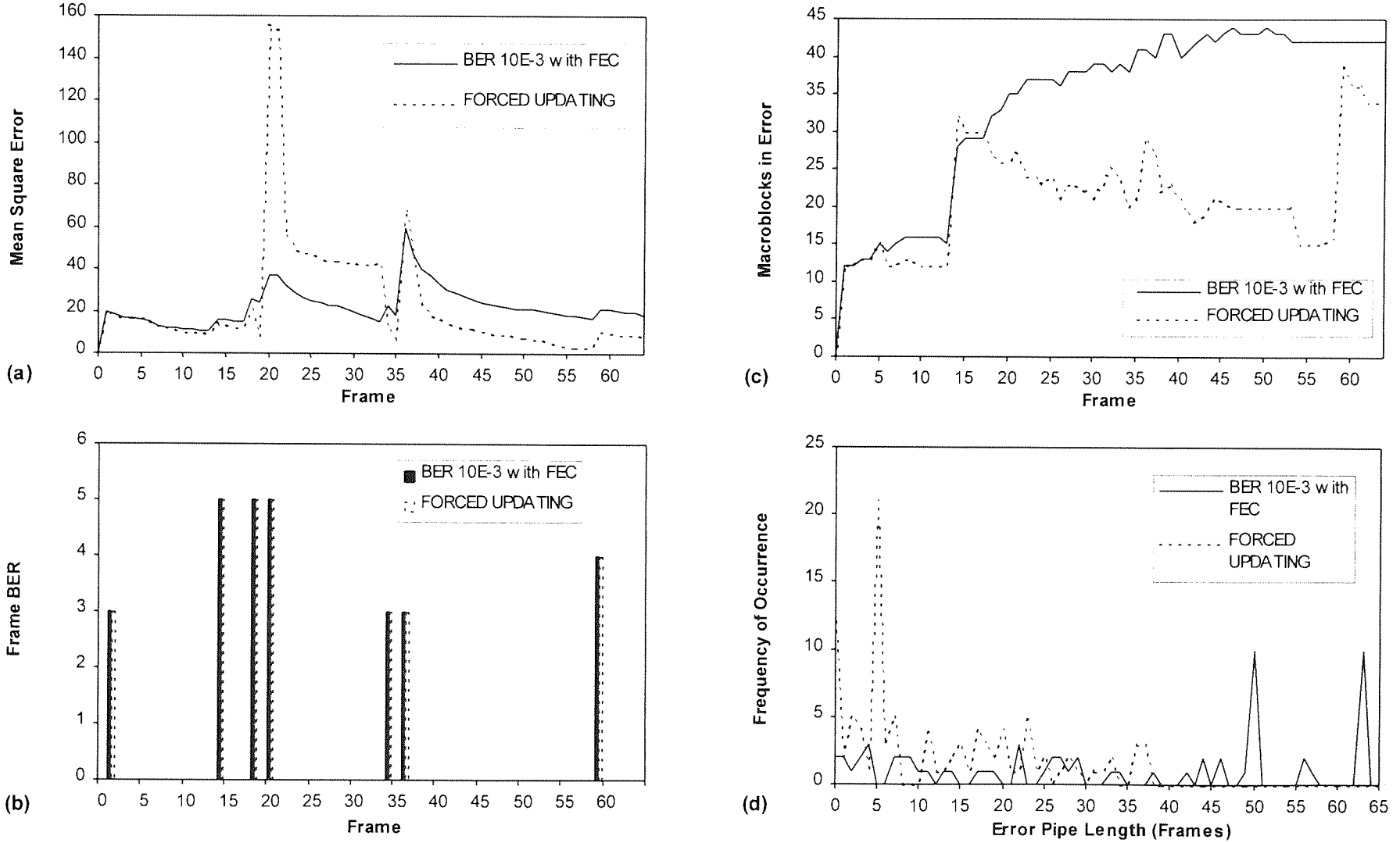


Figure 5.7 Forced Updating - QM48 The effect of Force Updating 3 MBs every GOB instead of 3 MBs per frame for the colour QM48 sequence coded at 48 kbit/s, corrupted by random noise at a BER of 10^{-3} and decoded by the robust H.261 decoder. (a) The Frame Square Error, (a) the residual BER per frame, (c) the errored Macroblocks per frame and (d) the Error Pipe Length Histogram.

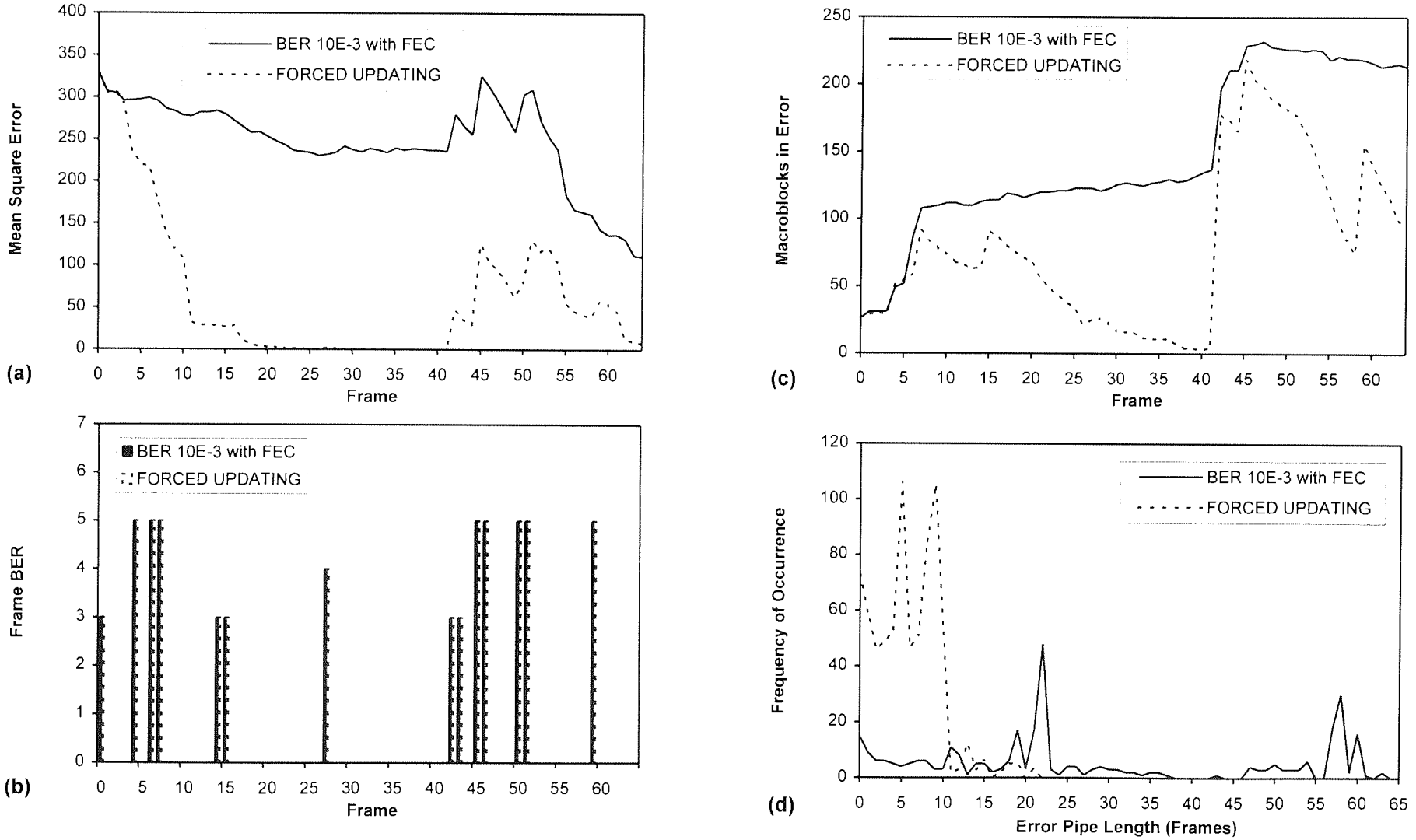


Figure 5.8 Forced Updating - MA96 The effect of Force Updating 3 MBs every GOB instead of 3 MBs per frame for the colour MA96 sequence coded at 96 kbit/s, corrupted by random noise at a BER of 10^{-3} and decoded by the robust H.261 decoder. (a) The Mean Square Error, (b) the residual BER per frame, (c) the corrupted Macroblocks per frame and (d) the Error Pipe Length histogram.

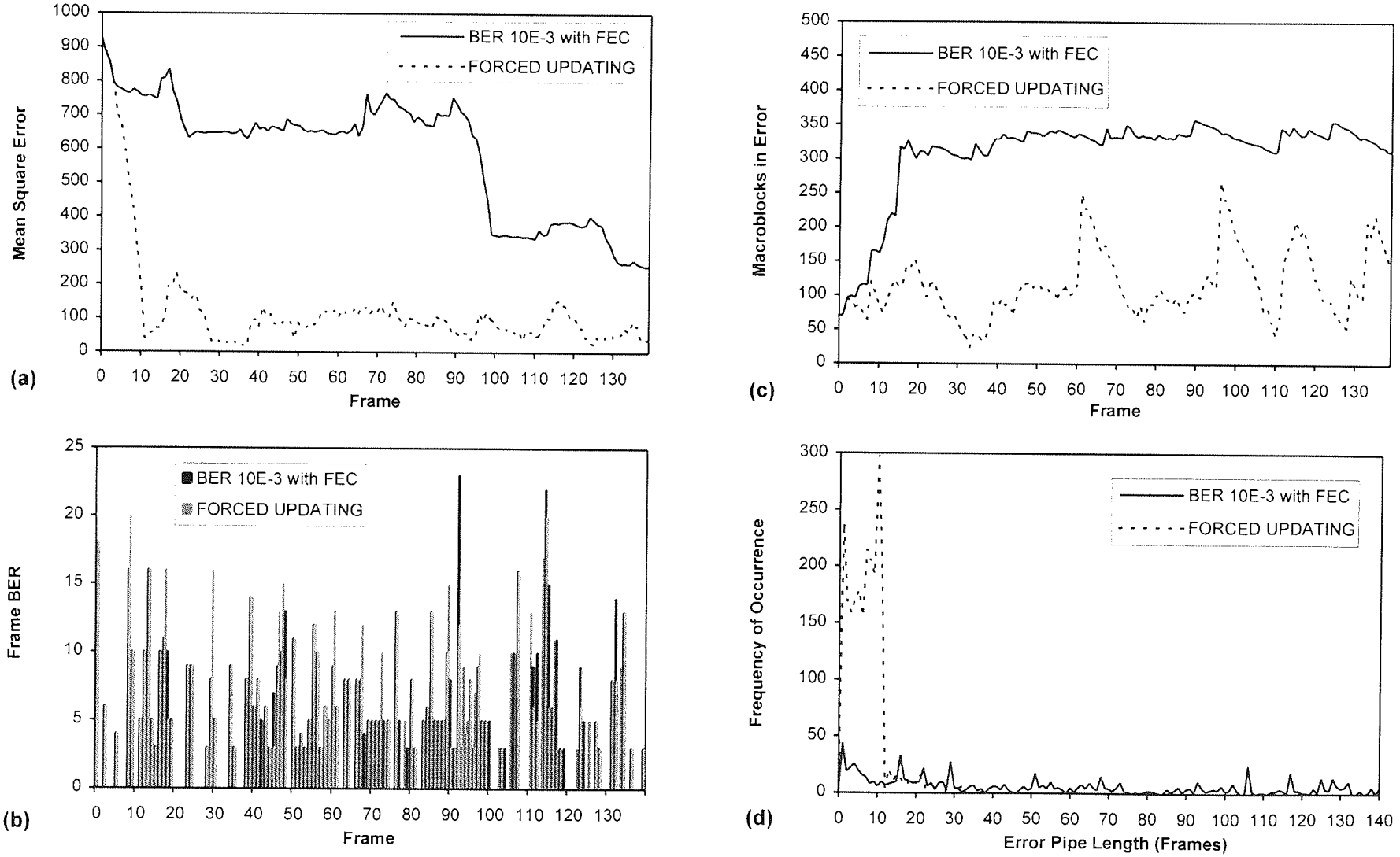


Figure 5.9 Forced Updating - SM384 The effect of Force Updating 3 MBs every GOB instead of 3 MBs per frame for the colour SM384 sequence coded at 384 kbit/s, corrupted by random noise at a BER of 10^{-3} and decoded by the robust H.261 decoder. (a) The Frame Square Error, (a) the residual BER per frame, (c) the errored Macroblocks per frame and (d) the Error Pipe Length Histogram.

5.9 Error recovery using the Fast Update Request facility.

H261 includes a fast update request mechanism whereby an additional party joining a conference session requests the transmission of the next frame in INTRA mode. This facility can also be used to recover from error as described below.

The BCH code is used to detect errors in the current frame. When an error is detected the previous frame display is frozen and an update request is issued via the mechanism originally intended for multipoint (multiparty) operation. When the encoder receives the update request, it codes the next frame in intramode and issues the picture release signal in the picture header. As soon as the receiver receives the picture release signal it resumes decoding and displaying the video data.

The technique can be used with unmodified H.261 codecs. A similar technique, but which does not freeze the picture when the fast update request is issued, has been studied in the context of packet video transmission over Asynchronous Transfer Mode (ATM) networks [Horst93], where packets can be lost due to congestion at statistical multiplexing nodes. One of the main conclusions in this study is that the fast update request technique or 'refresh on request' is an appropriate method to limit the cell loss effects temporally. The typical duration of error visibility was calculated to be of the order of 0.3 to 0.5 seconds which is around half that of forced updating with three MBs per GOB. Also, the overhead is substantially less using the fast update request provided the Mean Time Between Cell Loss (MTBCL) is of the order of 10 seconds.

At a bit error rate of 10^{-3} the probability of H.261 block error is of the order of 0.0152. At 64 kbit/s this represents two corrupted cells per second. At this MTBCL the fast update system would be intracoding frames continuously. However if the bit error rate is 10^{-4} the probability of block error is only 2.1×10^{-5} and the mean time between cell loss is of the order of 380 seconds at 64 kbit/s. At this BER MTBCL is more than ten seconds even at 2 Mbit/s.

It is thus clear that the effectiveness and efficiency of the fast update request technique is strongly sensitive to the BER between 10^{-3} and 10^{-4}. It is also expected that the performance of the technique should improve on burst error channels where the MTBCL is substantialy greater than on random error channels.

5.10 Transcoding

A transcoder is an apparatus which translates between two different source coding standards. The transcoded material can be audio or video. In the video field, transcoding is needed to translate between NSTC 525-line, 30 fps standard the PAL 625-line, 25 fps standard, for example. Digital wireless networks often employ audio compression algorithms and then transcode the audio to interface with the PCM oriented PSTN networks. Transcoding can also be used to interface H.261 codecs which are optimal over p x 64 kbit/s ISDN with more robust video coders optimized for wireless link operation.

5.10.1 Transcoding for video transmission.

A full-complexity transcoder decodes the input video stream down to the component frames, and then re-codes the frames to the required output standard. This arrangement makes no restrictions on the coded material. Indeed, transcoding is often performed to adjust for disparate frame sizes and frame rates. Although such an approach is the most flexible, it is can also be quite complex and expensive.

Transcoding creates two main problems in general. The first is the loss in quality due to the double coding and re-coding, and the second is the extra delay caused by the transcoder. However transcoders can be implemented which introduce no quality degradation [Zammit96], so that the transcoded video sequences have exactly the same subjective and objective quality, provided the two coding algorithm are similar.

5.10.2 Transcoding delay

The transcoder delay can be minimized by capitalizing on the similarity of the two algorithms and the fact that the input stream is already in FBR format, to eliminate the bit rate smoothing buffer.

The end-to-end delay on a video link used for two way communications must match that on the audio link, otherwise lip synchronization is lost, resulting in poor audiovisual reproduction. The maximum delay on a voice link at which it starts to interfere with communications intelligibility is of the order of 500ms. Loss of lip synchronization becomes noticeable above +180ms/-120ms [Karrlson]. Therefore this sets a maximum limit on the amount of delay on the video link.

The delay budget on a video link is shared amongst the coder, decoder, transcoder and the transmission functions. The coding and decoding delay in high compression video coding algorithms, especially those producing fixed bit rate output streams, are already high, and the transmission delay may be high as well, especially on links which include a satellite hop. It is therefore imperative to minimize transcoder delay.

Morrison [Morrison94] resolves the transcoder delay Δt_d into the following components :-

- **Computational Delay** (Δt_c):- Due to finite speed of execution of the coding algorithm in hardware and software and which can be reduced by faster hardware or software.
- **Algorithmic structure Delay** (Δt_a):- Due to the sequence of operation inherent in the algorithm and which cannot be reduced by faster hardware.
- **Buffering Delay** (Δt_b):- Due to the buffers needed to smooth the inherently variable bit rates generated by the coder to match the stream to the fixed bit rate streams usually available for communications.

Thus $\Delta t_d = \Delta t_c + \Delta t_a + \Delta t_b$. Morrison further describes how to reduce the effects of buffering delay in generic transcoders by translating the two transcoder buffers to the same location and merging them together into one unit, thereby eliminating one buffer.

The transcoding algorithm can be designed to minimize the transcoding delay Δt_d in two ways. First, the input stream is decoded immediately eliminating the need of an input buffer. Secondly, since the input stream is already in FBR configuration the rate control buffer can be eliminated. This minimizes the delay to less than one sixth of the normal encoder delay.

5.10.3 Transcoding techniques.

Transcoding permits an unlimited variety of error countermeasures to be introduced. However it is desirable to limit the complexity of the techniques to allow cost effective implementation since the transcoders would have to be fitted to all the links which support video services. (Initially only a shared transcoder pool could suffice until the video services become pervasive enough to justify one transcoder per interface link).

Thus a low complexity limit has been adopted which is implemented by considering only changes to the transmit and video multiplex coders, without touching the source coding algorithm. In this way the transcoding delay should also be kept low since it may be possible to transcode on the fly, without introducting algorithmic delays [Zammit96].

Furthermore, if the transcoder is designed carefully, the output stream will not require further rate control since the input stream would already be in FBR format. This eliminates the need for rate control and a rate control buffer and further simplifies the transcoder and decreases the delay. These ideas have been verified in an MPEG-to-H261 transcoder set [Zammit96] which is somewhat more complex than the techniques considered below.

5.10.4 FEC

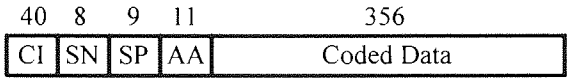
A low-complexity, low delay transcoder can be implemented by replacing the H.261 transmit coder's BCH (511,493) FEC code with a more robust FEC code. This approach has a number of advantages. Firstly, the H.261 error correction framing can be terminated at the transcoder. This practically eliminates the possibility of frequent re-synchronisation events. Secondly, the FEC block code size can be matched to the wireless network transport packet, thereby eliminating the need for separate error correction framing on the wireless network link. Thirdly, the rate of FEC code can be selected as necessary since by partitioning the overall link into segments it is possible to use a higher transmission rate on the high BER segment to allow for the application of low rate codes. In this case, the low complexity requirement can be easily met by just replacing the FEC code, leaving the VMC stream intact. Then, the transcoder delay is simply the time needed to decode the H.261 BCH code plus the time needed to re-code to the new FEC code specification, which is only a fraction of one frame period.

The effectiveness and efficiency of this approach can be easily demonstrated on the IHC simulation by protecting the MA 128 kbit/s and the SM 384 kbit/s sequences using low rate BCH

codes. On a Gaussian random error channel, this approach is optimal since an optimal system can be implemented by separating the source and channel coding and optimising both separately [Shannon48]. Thus it can serve as a benchmark against which the other techniques can be compared. When the channel is no longer memoryless, and when the coding delay is of importance, as in the case of real time communications, then Shannon's partition principle will not be applicable [Vembu95] and joint source channel coding techniques can do better within the required constraints.

5.10.5 Structured packing

Structured packing is a popular technique proposed for packet video transport over ATM [Hamano93]. Packet loss on ATM networks is mainly due to congestion and is of a bursty nature [Bito93]. A lost packet represents the loss of 384 video bits and leads to a loss of synchronisation if VBR video is used. The structured packing technique proposed for VBR MPEG transport includes three additional fields in the ATM transport packet as shown in Figure 5-10.



CI Cell Identification Header
SN Sequence Number
SP Start Pointer
AA Absolute Address

Figure 5-10 Structured Packing

The Sequence Number (SN) field is needed to detect cell loss. The start pointer (SP) points to the first bit of the first macroblock in the packet. If the block does not contain the start of a macroblock header this field is set to zero (say). The AA field contains the absolute address of the macroblock since the macroblock address in H.261 is differentially encoded with respect to the previous coded macroblock, which may have been lost. The motion vector for the first MB in the sequence must also be coded in absolute mode for the same reason.

The same technique can be applied to high BER channels. When the BER exceeds the FEC capability the errors cause the data in the rest of the block, and in the rest of the GOB to be lost due to the loss in synchronisation of the variable length codewords. Then it is advantageous to be able to recover synchronisation on the next un-corrupted or correctable packet since this would limit the error to those blocks within the corrupted packets. As it stands the technique does not allow advanced concealment techniques because it does not convey information about which macroblocks have been lost.

The overhead per transmit packet which encapsulates N_d data bits includes $\log_2 N_d$ bits for the SP start pointer field, the N_{SN} sequence number bits and the 10 bits to indicate the absolute MBA address, which can be 1 to 396. The absolute motion vector overhead cannot be estimated

deterministically since not all MBs are coded with motion vectors and some of those with motion vectors are coded in absolute mode irrespectively.

5.10.6 QCIF re-synchronisation

It was noted above, that loss of a GOB in QCIF sequences leads to very visible degradation since one GOB represents one third of the frame area. A more reasonable re-synchronisation point would be at the leftmost side of each of the 9 macroblock rows. This increases the frame size by 26 x 6 bits. In addition the macroblock addresses and the motion vectors of the first coded macroblock would have to be coded in absolute mode.

5.10.7 Fixed sized Macroblocks.

The idea here is to code macroblocks using a fixed number of bits per macroblock, such that the beginning of each macroblock could be identified, allowing the incorporation of macroblock level error concealment. However since macroblocks generate a variable number of bits, a better approach may be to align macroblock sizes to 16 bit quanta (say) such that the start of macroblocks could be more easy to identify whilst accomodating the variable nature of the coded macroblocks.

5.11 Summary of Conclusions

This chapter studied the resilience of H.261 video codecs to high bit error rates on a random error binary symmetric channel. Although the wireless network channels considered in this thesis are not well modelled as random error channel, the analysis of the behaviour of the H.261 codec on this channel revealed certain shortcomings in the standard and allowed certain modifications to be proposed and tested.

The error resilience of the H.261 codec was investigated by analysing the three main codec components separately. The motion compensated hybrid interframe DPCM/DCT source coding algorithm was considered first. The components of this complex algorithm were identified and treated separately. The DCT algorithm *per se* is shown to be relatively robust to channel errors, with the 8x8 block structure limiting the spatial propagation of coefficient errors. However the threshold coefficient selection schemes is less robust than a fixed zonal selection scheme.

The weakest part in the source coder is identified as the interframe DPCM coder, and the lack of an adequate temporal error propagation countermeasure in the standard. The conditional replenishment aspect of the algorithm requires the transmission of block address which introduces further weaknesses. The motion compensation scheme is also shown to cause both temporal and spatial error propagation.

The video multiplex coder was identified as a major cause for concern because of the liberal use of variable length codes in the standard. This is further compounded by the use of differential encoding of certain key parameters such as the macroblock address and the motion vectors. Although some damage limitation is in-built into the standard using absolute encoding in certain cases.

The transmit encoder is shown to be inadequate for bit error rates in the region of 10^{-3} which are typical of the bit error rates experienced on wireless networks. The BCH codeword framing mechanism is also cause of concern since loss of re-synchronization means the loss of the video data stream for hundreds of milliseconds.

The error control mechanism, besides the transmission encoder, consists of the picture and group of block headers which allow resynchronization should this be lost because of channel errors. The spatial error is adequate for CIF size images, but is inadequate for QCIF images.

The implementation of robust decoders was discussed, leading to a description of the IHC decoder which uses the full error control capabilities of the H.261 standard. It normally uses the FEC code to correct errors. Alternatively it can use the parity bits to detect errors in addition to the limited error detection afforded by VMC code violations. When an error is detected the corrupted block can be discarded and the decoder seeks the nearest re-synchronisation point. Concealment can be intrinsic or extrinsic. In the latter case the current GOB data is replaced by the previous GOB data.

The robust decoder was tested in high BER conditions. It has been shown that a properly designed H.261 decoder can survive around one bit in error per transmitted frame, without making use of the general BCH code or error concealment, but protecting the PH and GOBH header fields. The decoder traps field errors and skips to the next GOB when errors are detected. However the error detection is very unreliable since the detected error events are very few, due to the complete nature of the variable length codes. Occasionally highly visible errors occur but at this error rate these do not propagate catastrophically for the tested sequences and are kept in check by the forced updating of the unaltered algorithm. Although some errors are highly objectionable, the decoded image is still perfectly intelligible. When the error rate is increased to 10 bit errors per transmitted frame, the visible errors coalesce rapidly and the intelligibility is lost very quickly.

When the BCH code is used to detect errors, the error detection capability becomes highly reliable. The decoder still attempts to decode the PH and GOBH fields in corrupted blocks, though it will then discard the decoded macroblock data and skip to the next non-corrupted GOB.

When the BER is 10^{-4} the FEC code can correct most of the error patterns and the three sequences were decoded without errors. When the BER is 10^{-3} the FEC decoder commits a significant number of decoding errors and there is visible picture degradation. The sequence can be decoded usefully, however, and there is no catastrophic picture degradation.

It was argued that the robustness of H.261 can be improved in three ways:-

(a) Using H.261-compliant techniques which can be decoded by normal decoders.

(b) Using H.261-compatible techniques which require intelligent decoders to capitalise on the improved error robustness.

(c) Using non-H.261-compliant techniques, in which case a low-delay, low-complexity transcoder can be used to interface the robust codecs to normal H.261 codecs.

Five H.261 compliant or compatible techniques were proposed to improve the resilience of H.261.

(i) Header protection.

The header protection technique uses the PSPARE and GSPARE header fields which are ignored by normal H.261 decoders. It is proposed to use enough SPARE fields to implement a FEC code 'signature'. Then probable corrupted picture and GOB headers can be located by correlation and positive identification achieved by proper decoding of the FEC code 'signature'. This technique is H.261 compatible but not H.261 compliant since the use of the SPARE fields in prohibited in H.261.

(ii) Fixed length DC intraframes.

It was noted that corruption of the first intracoded frame led to very visible and persistent image degradation. The fixed length DC intraframe was proposed to implement a robust intraframe mode. The technique codes the first frame use the fixed length eight bit dc fields such that all the fields in the frame are fixed length and synchronisation cannot be lost, although dc coefficients may be corrupted. The error resilience is also improved because the frame format uses the lowest number of bits possible and hence stands a better chance of being delivered uncorrupted. Results confirm the robustness of the technique. The initial reduction in picture quality is generally recovered in the subsequent frame. The combined resilience of the first two frames is much improved with the proposed technique.

(iii) Cyclic Refresh using Periodic Intracoding.

Error recovery using periodic intracoding was shown to be inadequate for fixed bit rate video codecs because the intraframes consume three to five times the bit rate of interframes and lead to very visible temporal resolution reduction, especially at low frame rates. Also the intracoded frames are not very efficient when compared to MPEG-1 say.

(iv) Cyclic Refresh using Forced Updating.

Forced updating over and above the H.261 recommendation was found to provide very good error recovery capabilities. The good temporal error recovery is also experienced by normal H.261 decoders. Force updating three macroblocks per GOB per frame was found to be a good compromise between picture quality reduction and error recovery capabilities. At this rate, the picture usually recovers from errors within eleven frames.

(v) Error Recovery using Fast Update Request.

H.261 incorporates a fast update request signal which can be asserted by third parties wishing to join a multi-party conversation. This signal can be used by a decoder to force the far end encoder to code the next frame in intramode to recover from an error. It was found that this technique is very sensitive to BER in the region of 10^{-4} to 10^{-3} on a random error channel. It is anticipated, however, that the techniques would be more successful on a burst error channel at these error rates.

It was argued that the best way to access H.261 video services over a wireless network may be to use a robust non-H.261 compliant mobile video codec and use transcoders to interface the video stream with H.261 codecs on the fixed network. The two main issues involved are complexity and transcoding delay. Complexity should be low to reduce cost. The transcoding delay should be low to prevent excessive communications delays which are detrimental to speech intelligibility.

A number of low-complexity, low-delay techniques were proposed.

The optimal solution on an AWGN channel is to replace the H.261 BCH code with a more robust FEC code. However this solution may not be optimal if the channel is not memoryless and the system complexity and decoding delay are an important issue.

Structured packing is another low-delay, low-complexity option. The technique adds absolute pointers to video data packets to allow re-synchronisation at points other than the picture and GOB headers, and thus allows better concealment strategies to be adopted. The first macroblock header in each packet has to be identified by a pointer. The differentially coded fields in the macroblock, i.e. the macroblock address and the motion vector fields, must be augmented with absolute information.

For QCIF sequences the GOB re-synchronisation points do not offer adequate spatial error protection since the loss of one GOB signifies the loss of 33% of one frame. Therefore it is proposed to include a new synchronisation point at the beginning of macroblock rows.

Another proposal is to re-code the DCT coefficients using a number of fixed length block sizes to allow synchronisation at macroblock boundaries. This is the most complex proposal and may result in some inefficiencies.

Chapter 6
Modelling and Simulating the Digital Radio System

This chapter describes the development of modeling and simulation tools which are used to generate error vector files that realistically reproduce the error characteristics on digital radio telecommunication links. These error vector files are required to corrupt digital video streams, to evaluate the robustness of the proposed codecs to conditions encountered on digital radio networks.

The mobile radio propagation channel is introduced, and characterized as a time-varying, multipath fading channel. The parameters which determine the manner in which the radio channel influences the different digital modulation techniques are introduced, and these lead to the various classifications of the mobile radio channel, *e.g.* narrowband or wideband, fast fading or slowly fading *etc.*

A number of modulation techniques are identified, which are suitable for second and third generation, digital wireless networks. A Quadrature Amplitude Modulation (QAM) communications system model is developed to study the behavior of the modulation schemes over the various mobile radio channels considered.

A simulation infrastructure is developed, and the simulation techniques used to implement the various signal processing functions, such as filtering, and the Rayleigh fading channel simulation are described. The modeling and simulation techniques are validated by comparing the performance results obtained with published theoretical, simulation and experimental results.

6.1 Introduction

Bit error generators with accurate first and second-order statistics are crucial to the work in this thesis. The evaluation of error control techniques for real-time services such as video, requires knowledge of the second order statistics of the error generation process, especially for burst error channels encountered in digital radio systems. This requirement arises from the delay sensitivity of real time services, and hence the need to determine how much time delay is introduced by the error control technique.

Burst error models, like the Gilbert-Elliot Model (GEM) [Gilbert60] and its derivatives [Fritchman67], have been successfully employed to model and simulate burst error processes over Rayleigh fading channels [Jeruchim93]. These models use the fade and interfade statistics of the process they model to generate accurate error streams.

It is thus necessary to model and simulate the complete digital radio link, to generate burst error processes for direct use or from which the second order symbol and bit error statistics can be extracted to drive burst error models such as the GEM. In this chapter, the digital radio systems is first modeled mathematically and then simulated in discrete time. The simulations generate error vector files in which a '1' represents a bit error and a '0' a correct bit.

The simulation programs are coded in C and C++ and run on UNIX based platforms. The time domain models and simulations are used to generate bit error data for the listed, uncoded, multiphase digital modulation schemes over four different channels: the Additive White Gaussian Noise (AWGN) channel; the narrowband, non-frequency-selective, Rayleigh fading channel; the Rayleigh fading channel with co-channel interference; and the wideband, frequency-selective, two-ray, Rayleigh fading channel.

- DQPSK — differential quadrature phase shift keying
- π/4 DQPSK — π/4 shifted differential quadrature phase shift keying
- 8-DPSK — eight differential phase shift keying
- 16-DPSK — sixteen differential phase shift keying

The chapter is organized as follows. The mobile radio propagation channel is introduced first and the general behavior of communications systems on such channels is described. This is followed by a discussion of the factors influencing the choice of modulation schemes for further study. The communications system model used in the subsequent simulations is presented next, followed by detailed descriptions of its components. The last sections present a number of results. The simulations are validated by comparing the Symbol Error Rate (SER) and Bit Error Rate (BER) curves for the multiphase modulations on the AWGN, Rayleigh fading and two-ray frequency-selective channels, with published results.

6.2 The Radio Propagation channel

In mobile radio and wireless indoor channels, more than one propagation path exists between the transmitter and the receiver[Jakes74, Proakis89], due to reflection, scattering and diffraction caused by obstacles between and around the source and the destination. Such multipath propagation channels are characterized either by a number of discrete propagation paths or by a continuum of paths [Proakis89].

The attenuation exhibited by multipath channels in mobile and personal communications systems is time varying due to the movement of the mobile nodes and the varying nature of the immediate surroundings (e.g. movement of people and objects in the region of influence). These time variations cause amplitude variations in the received signals, called fading. If a direct path exists between the transmitter and the receiver the amplitude variations are found to have Ricean distribution, and Rayleigh if such a direct path is absent [Jakes74].

The multipath channel also suffers from phase variations. If the phase changes significantly between two transmitted symbols the channel is said to exhibit fast fading. It is said to be slow if the phase changes are insignificant over two symbol periods. Thus whether a multipath channel is fast or slow fading, depends on the mobile speed and the signaling rate.

If the delay variation between the various received paths is significant with respect to one symbol duration, the channel will introduce intersymbol interference. This is exhibited in the frequency domain as frequency-dependent amplitude and phase spectrum distortions and the phenomenon is frequently referred to as frequency-selective fading.

The performance of digital modulation techniques in the presence of multipath fading is significantly worse than on the non-fading additive white Gaussian noise channel. Fast phase variations and frequency-selective fading introduce error floors which cannot be eliminated by increasing the signal to noise ratio. Various diversity (*e.g.* frequency and antenna diversity) and equalization techniques are usually employed to combat fading.

In this thesis the emphasis is on the short term, fast fading channel. Shadowing and propagation loss are not considered in this thesis.

6.2.1 Fading Multipath Channel Models

The discrete multipath channel model is a simple, but often used, statistical channel model with the time-varying, lowpass, channel impulse response $c(\tau;t)$ given by [Proakis89]:-

$$c(\tau;t) = \sum_{n} \alpha_n(t) e^{-j\theta_n(t)} \delta[\tau - \tau_n(t)] \qquad \text{(6-1)}$$

Each of the n propagation path has associated with it, randomly time-varying attenuation $\alpha_n(t)$, phase delay $\theta_n(t)$ and propagation delay $\tau_n(t)$ factors which can be modeled as random processes. The determination of the statistical distributions of $\alpha_n(t)$, $\theta_n(t)$, $\tau_n(t)$ and the number of rays n in the model have been the subject of much research [Hashemi94]. A commonly used approach is to model $\alpha_n(t)$ as a Rayleigh distributed random variable, and $\theta_n(t)$ as a uniformly distributed random variable in the range 0 to 2π, as in [Chuang87] and, therefore, each impulse coefficient in $c(\tau;t)$ is a complex Gaussian random variable. Physically this can be justified by arguing that at any particular τ the impulse response consists of a large number of unresolved paths, and then invoking the central limit theorem so that the received path approaches a complex Gaussian distribution [Proakis89]. A number of other distributions have, however, been suggested in the literature. [Hashemi94] is a detailed review of indoor radio channel modeling and contains a comprehensive list of proposed distributions.

By characterizing $c(\tau;t)$ as a complex Gaussian process and assuming that it is wide sense stationary and that the attenuation and phase shift at a path τ_1 is un-correlated with the attenuation and phase shift at path τ_2 (i.e. un-correlated scattering), Proakis shows that the autocorrelation function is

given by:-

$$\phi_C(\tau_1,\tau_2;\Delta t)=\frac{1}{2}E\left[c^*(\tau_1;t)c(\tau_2;t+\Delta t)\right]=\phi_C(\tau_1;\Delta t)\delta(\tau_1-\tau_2) \qquad \textbf{(6-2)}$$

At $\Delta t = 0$ the autocorrelation function is simply a measure of the average power output of the channel as a function of τ, where $\phi_C(\tau_1;0)\equiv\phi_C(\tau)$ is referred to as *the multipath intensity profile* or *the power delay profile*. The multipath delay spread T_m is the range of τ over which $\phi_C(\tau)$ is essentially non-zero. The root mean-square (rms) delay spread σ_{rms} of the power delay profile *p(t)* is defined by [Chuang87], thus:-

$$\sigma_{rms}=\left[\frac{\int(t-D)^2\,p(t)dt}{\int p(t)dt}\right]^{1/2} \qquad \textbf{(6-3)}$$

where D, the average delay is given by:-

$$D=\frac{\int tp(t)dt}{\int p(t)dt} \qquad \textbf{(6-4)}$$

The performance of digital modulation techniques over multipath channels is strongly dependent on the rms delay spread [Chuang87] and is only weakly dependent on the shape of the power delay profile.

The multipath delay spread is experienced in the Frequency Domain as frequency selective fading. The reciprocal of the multipath delay spread is called the coherence bandwidth $(\Delta f)_C$. Two sinusoids separated by more than $(\Delta f)_C$ are treated independently by the channel. When the bandwidth of a passband signal is small relative to $(\Delta f)_C$ the signal is passed through the channel largely unchanged (probably attenuated) and the channel is said to be non frequency-selective. If the bandwidth is larger than the coherence bandwidth, the channel is referred to as frequency-selective and it will introduce gross distortions in the signal. In the time domain this distortion is called Inter-Symbol Interference (ISI).

The power delay profile is thus a very powerful characteristic of the channel and it can be measured directly by transmitting very narrow pulses and cross-correlating the received signal with a delayed version of itself [Parsons89].

If a bandpass signal *s(t)* is transmitted over the channel represented by equation (6-1) where

$$s(t)=\Re\left[u(t)e^{j2\pi f_c t}\right] \qquad \textbf{(6-5)}$$

an *u(t)* is the baseband waveform, the received signal *r(t)* is given by:-

$$r(t)=\Re\left(\left\{\sum_n \alpha_n(t)e^{-j2\pi f_c\tau_n(t)}u\left[t-\tau_n(t)\right]\right\}e^{j2\pi f_c t}\right) \qquad \textbf{(6-6)}$$

If an un-modulated carrier of frequency f_c is transmitted the received signal becomes:-

$$r(t) = \sum_n \alpha_n(t) e^{-j\theta_n(t)} \qquad (6\text{-}7)$$

where $\theta_n(t) = 2\pi f_c \tau_n(t)$. Proakis shows that if n is large, *r(t)* can be modeled as a complex-valued Gaussian random process by invoking the central limit theorem. The channel impulse response then becomes:-

$$C(\tau;t) = R(t) e^{-j\theta(t)} \delta(\tau) \qquad (6\text{-}8)$$

where *R(t)* is Rayleigh distributed and *θ(t)* is uniformly distributed. Then the channel is said to be a *Rayleigh fading channel.* If there are moving scatterers between the transmitter and the receiver, the impulse response cannot be modeled with zero mean and the distribution of *R(t)* has a Rican distribution and the channel is called a *Ricean fading Channel.* In this thesis only the Rayleigh fading channel is considered. This channel is often considered to produce the worst-case results. The statistics and radio frequency Doppler spectrum of the Rayleigh fading channel are summarized in section 6.5.3.2.

6.3 Modulation

Filtered minimum shift keying has been used in two second-generation digital wireless networks: in DECT as Gaussian filtered FSK (GFSK) and in GSM as GMSK. Its use in DECT is of particular interest since the transmission of H.261 coded video has already been demonstrated on a DECT test-bed [McDonald92]. Detailed bit error rate statistics obtained from a DECT test bed were available [BT] from which a reliable bit error rate model was built and incorporated into the simulation package.

π/4 shifted DQPSK has been extensively studied since its selection for the IS-54 standard and the Japanese JDC standard [Goodman91a]. It has also been studied in the context of third-generation PCS systems, for higher data rate microcellular and picocellular systems. Four QAM has also been proposed for the MBS system being developed by RACE [Zubrzycki94].

Filtered FSK and DQPSK achieve spectral efficiencies below 2 bit/s/Hz. Third-generation PCS systems, however, need higher spectral efficiencies if they are to meet the requirements of new wireless multimedia services [Stedman92]. Differential multiphase signaling is a spectrum efficient, constant envelope modulation technique. Quadrature Amplitude Modulation (QAM) performs optimally over the Additive White Gaussian noise (AWGN) channel, but there are problems with QAM signaling over multipath fading channels. First, there is the problem of carrier recovery which is non-trivial in fading channels. Secondly, there is also the probability of false locking [Webb91] due to non-linear constellation deformation as the receiver passes through a deep fade.

The first problem can be solved by using differential modulation. The second problem has been solved by employing rotation invariant modulation [Webb91]. Several researchers have studied 16 Star QAM or Differential Amplitude and Phase Shift Keying (DAPSK) [Webb91, Nix91, Adachi92]

in particular. DAPSK performs 4.1dB worse than coherent 16 QAM but only 1.9dB worse than differential encoded and detected 16 QAM on the Rayleigh fading channel. The less complex 16-DPSK performs only 1.5 dB worse than DAPSK in the presence of Rayleigh fading [Adachi92], and is simpler to implement and to analyze. It still retains all the advantages of DAPSK, namely differential detection, elimination of false locking, spectrum efficiency and adds the advantage of a constant envelope.

The modulations modeled and simulated in this chapter are DQPSK, π/4-DQPSK, 8-DPSK and 16-DPSK. GMSK is not studied directly though the available data allowed the design of a error generator which was incorporated into the simulation package.

6.4 The digital radio system model.

The modulation schemes of interest in this thesis can be modeled as quadrature amplitude modulation (QAM) systems. The digital communications system block diagram for generic QAM is shown in Figure 6-1. The generic model consists of three sub-models: the transmitter model, the channel model, and the receiver model.

6.4.1 Transmitter Model

The data source generates the bit stream $\mathbf{b_n}$. The serial to parallel converter accepts $\mathbf{\log_2(M)}$ bits from the bit stream to form the discrete level stream $\mathbf{A_k}$. The encoder and symbol mapper maps the $\mathbf{A_k}$ binary symbol into the complex symbol $\mathbf{a_m}$ chosen from an **M**-symbol alphabet. The mapping is usually done observing Gray encoding, such that if upon reception the symbol $\mathbf{a_m}$ is mistaken for one of its nearest neighbours in the constellation space (the most probable occurrence), only one bit is in error after inverse mapping at the receiver. In multiphase QAM the complex symbols are chosen from the alphabet $\mathbf{\exp(jn\pi/M)}$ where n={0..M-1} such that $a_m = e^{j\theta_m}$. If differential encoding is used, the output symbol $\mathbf{a_m}$ is given by :-

$$a_m = e^{j\theta_m}.e^{j\theta_{m-1}} = e^{j(\theta_m+\theta_{m-1})} \qquad \text{(6-9)}$$

The mapper then outputs the in-phase I channel and the quadrature Q channel impulses $\mathbf{I_k=Re(a_m)}$ and $\mathbf{Q_k=Im(a_m)}$. The impulses are filtered by the transmit filters whose impulse response defines the pulse shapes. The transmit filters are chosen to shape the spectral occupancy of the transmitted inphase and quadrature baseband signals **u(t)** and v(t). In addition the transmit filters, in cascade with the receiver matched filters, obey the Nyquist criterion for ISI free transmission [Proakis89]. Finally, the passband signals are modulated by multiplication with the cosine and sine functions at the carrier frequency ω_C, then added together and transmitted. The RF power amplification and the transmitting antenna are considered to be part of the channel.

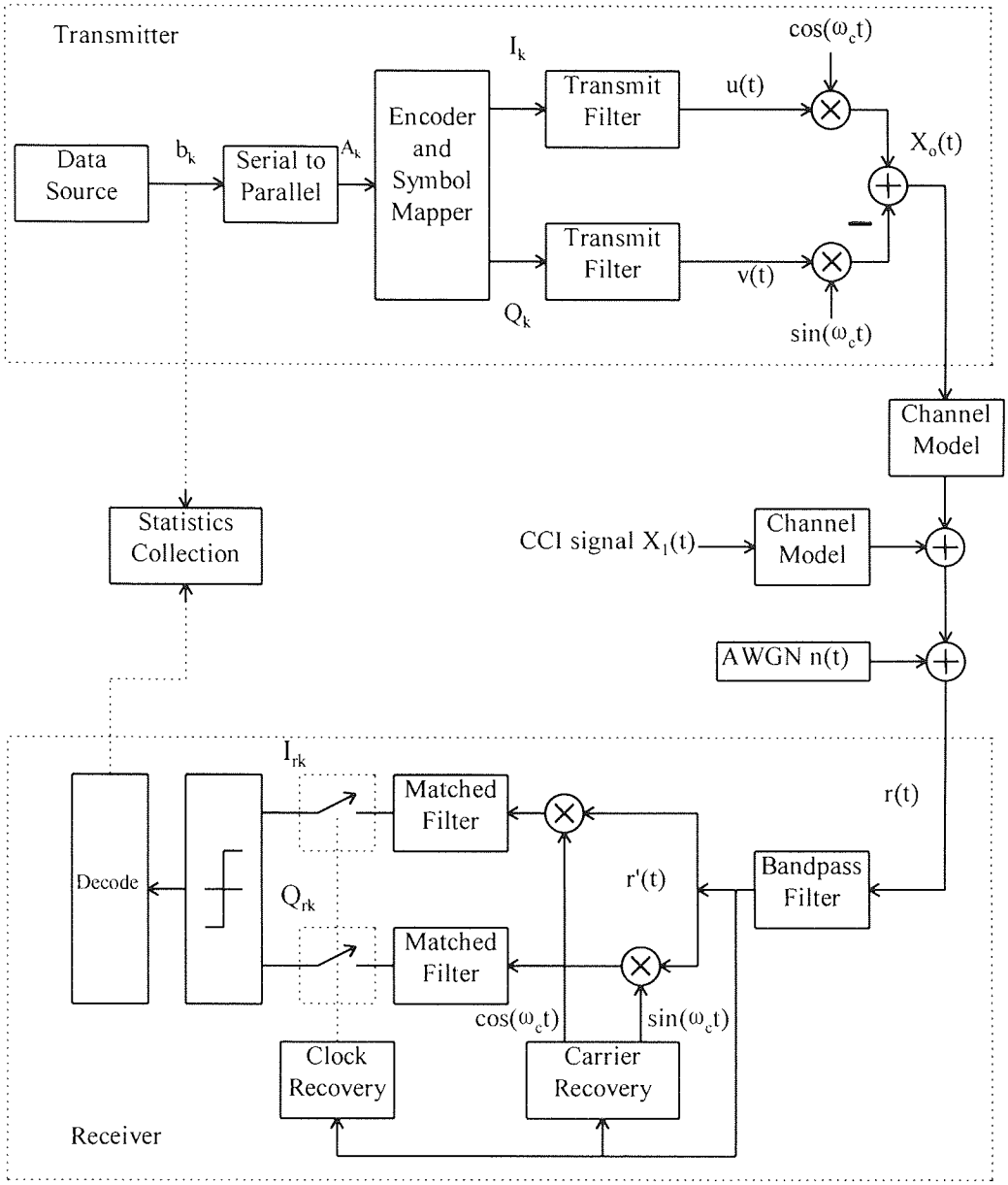


Figure 6-1 Generic QAM Communications System Model

6.4.2 The Channel Model

The mobile radio propagation channel characteristics, the transmitted symbol rate, the bandwidth occupancy and the mobile speed play an important role in the selection of an appropriate channel model. In cellular mobile systems, adjacent and co-channel signals may also interfere with the transmitted signal in addition to the ever present Additive White Gaussian Noise (AWGN), and are included in the overall channel model.

Three basic channel models are considered. The first is the AWGN channel which is perturbed by spectrally white noise with double sided spectral density $S_{NN}(f)$ equal to $N_o/2$. The noise fluctuations v_n have a Gaussian probability distribution given by [Proakis89] :-

$$p(v_n) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{(v_n^2/2\sigma^2)}. \qquad \textbf{(6-10)}$$

The second channel of interest is the Rayleigh fading channel shown in Figure 6-2 which multiplies the transmitted signal ***X(t)*** by a complex Gaussian process which, in polar form, becomes $R(t)e^{\theta(t)}$ with Rayleigh distributed ***R(t)*** and uniformly distributed $\theta(t)$. The Rayleigh fading channel is appropriate for narrowband systems over multipath fading channels usually found in mobile radio systems.

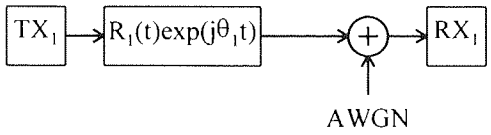


Figure 6-2 Rayleigh fading Channel Model

In cellular systems the main limitation is often due to co-channel interference (CCI), since it is desirable to re-use frequencies in cells as close as possible in order to utilize the available spectrum as efficiently as possible. The CCI channel model used is shown in Figure 6-3 below. The main transmission is received at the receiver RX_1 via the Rayleigh fading path R_1. The co-channel interferer TX_2 is received via an independent Rayleigh fading path R_2.

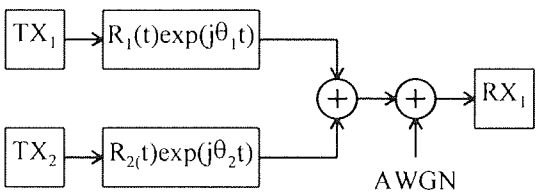


Figure 6-3 Rayleigh fading channel with CCI

The third channel of interest is the frequency-selective channel. Chuang [Chuang87] studied MPSK and MSK modulations on different discrete multipath channel models and concluded that the BER performance is more sensitive to rms delay spread than to the shape of the power delay profile. Fung *et al.* [Fung93] report, however, that this may not be valid for the Indoor Radio Channel.

In this section the discrete multipath, frequency selective channel is modeled using the two-ray, equal amplitude delay profile, with two independent Rayleigh fading rays separated by τ seconds, as shown in Figure 6-4. This channel model is the simplest example of a more generic wideband channel model usually implemented as a FIR filter with time varying, Rayleigh fading coefficients. For this channel the τ_{RMS} is equal to $\tau/2$.

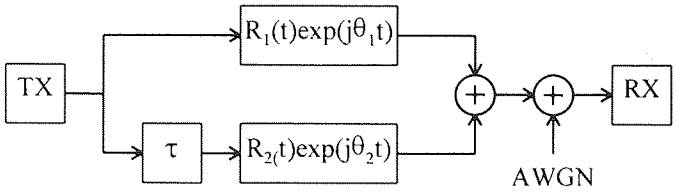


Figure 6-4 Two-ray frequency-selective model

6.4.3 The Receiver Model

The received signal ***r(t)*** is equal to the sum of the transmitted signal ***X(t)*** which is filtered by the channel model, the AWGN and the CCI signals. The bandpass filter selects the wanted signal and filters out-of-band noise. The bandpass filter output ***r'(t)*** is then multiplied by the in-phase and quadrature components $cos(\omega_c t)$ and $sin(\omega_c t)$ which are generated by the carrier recovery circuit, which recovers frequency and phase from the received signal. The matched filters are optimized to detect the signals, maximizing the signal-to-noise ratio (SNR), without introducing inter-symbol interference in the absence of channel nonlinearities. The matched filter outputs are then sampled at the appropriate time as determined by the symbol timing recovery circuit.

The quadrature samples are then operated upon by the inverse mapping function, which selects the constellation point which is closest to the sampled (I_{rk}, Q_{rk}) phasor. The decoder than outputs the bit stream associated with the selected constellation point.

6.5 System Simulation

The Monte Carlo simulation technique [Jeruchim92] was used throughout. The bit error rates of interest were quite high, in the region of 10^{-3} - 10^{-4}, so that the simulation runs to obtain reliable results with high confidence limits were not prohibitively long.

The simulation programs where coded in C++ and run on UNIX platforms. The programs allow both baseband and Intermediate Frequency (IF) simulations. Most of the results were obtained using baseband simulation to reduce the simulation time. The intermediate frequency simulation was mainly used to calibrate the baseband simulations by deriving the noise bandwidth of the receiver filters.

6.5.1 Simulating the transmitter

The transmitter block diagram is shown in Figure 6-1. It consists of a data source, a serial to parallel converter, a Gray encoder and mapper into the PSK constellations, pulse shaping filters and the cosinusoidal modulation functions (used only in the intermediate frequency simulation mode). The data source was generated using pseudo-random number generators.

6.5.1.1 *Pseudo Random Number Generation*

Both Uniform and Gaussian distributed pseudo-random number generators were used in the simulations. The uniform random number generators were used to generate the information sequences. Gaussian distributed random variables were used to simulate the additive white Gaussian noise samples.

The Wichmann-Hill algorithm was used to generate uniformly distributed random sequences with long periods and good distributions [Jeruchim92]. It has been shown that the sequence generated has a period of approximately $7x10^{12}$, which is excellent for the simulations in this thesis.

The Box-Muller algorithm [Jeruchim92] was used to generate two independent zero-mean Gaussian variables X and Y. The required results are obtained using two independent variables U_1 and U_2, uniformly distributed in the unit interval, where:-

$$X=[-2 \ln (U_1)]^{1/2} \cos (2\pi U_2) \qquad (6\text{-}11)$$

and

$$Y=[-2 \ln (U_1)]^{1/2} \sin (2\pi U_2) \qquad (6\text{-}12)$$

X and Y have unit variance. To obtain a variance σ^2, X and Y have to be multiplied by σ.

6.5.1.2 *Digital Encoding*

Three classes of encoders were used: absolute encoders for M-ary PSK, differential encoders for M-ary PSK and the differential encoder for $\pi/4$ DQPSK. Gray encoding was used with all encoding techniques. The absolute encoding function is described by:-

$$a_m^k = \exp(2\pi\Gamma(A_k)/M) \qquad (6\text{-}13)$$

where a_m^k is the kth, M-ary discrete symbol, A_k is the kth $\log_2(M)$ bit word, and $\Gamma()$ represents the Gray encoding function. Differential encoding can be described by:-

$$a_m^k = \exp(2\pi\Gamma((A_k + A_{k-1}) \bmod M)/M) \qquad (6\text{-}14)$$

where A_{k-1} is the previous word and x *mod M* returns the remainder of x divided by M.

DiBit MSB	DiBit LSB	$\Delta\phi_k$
0	0	$\pi/4$
0	1	$3\pi/4$
1	1	$-3\pi/4$
1	0	$-\pi/4$

Table 6-1 Gray coded dibits in $\pi/4$ DQPSK

The encoding function for $\pi/4$ DQPSK is given by [Chennakeshu93]:-

$$a_m^k = \exp(\phi_{k-1} + \Delta\phi_k) \qquad \text{(6-15)}$$

where ϕ_{k-1} is the absolute phase for symbol (k-1) and $\Delta\phi_k$ is the differential encoded phase, derived from the Gray coded dibits as illustrated in Table 6-1

6.5.1.3 Pulse Shaping filters

The symbols generated by the encoder are pulse shaped using square root raised cosine filters with excess bandwidth parameter α defined by [Chennakeshu93]:-

$$G(f) = \begin{cases} T, & 0 \le |f| \le (1-\alpha)/2T \\ T\sqrt{\frac{1}{2}\left(\left\{1-\sin\left[\pi\frac{T}{\alpha}\left(|f|-\frac{1}{2T}\right)\right]\right\}\right)} & \frac{(1-\alpha)}{2T} \le |f| \le \frac{(1+\alpha)}{2T} \\ 0, & |f| > (1+\alpha)/2T \end{cases} \qquad \text{(6-16)}$$

where *G(f)* is the filter response in the frequency domain. The baseband simulations were conducted in the time domain using the impulse response of the filter *g(t)* which can be shown to be [Chennakeshu93]:-

$$g(t) = \begin{cases} 1-\alpha+(4\alpha/\pi), & t = 0 \\ \frac{\alpha}{\sqrt{2}}\left[\left(1+\frac{2}{\pi}\right)\sin\left(\frac{\pi}{4\alpha}\right)+\left(1-\frac{2}{\pi}\right)\cos\left(\frac{\pi}{4\alpha}\right)\right] & t = \pm\frac{T}{4\alpha} \\ \left\{\sin\left[\pi(1-\alpha)\frac{t}{T}\right]+4\alpha\frac{t}{T}\cos\left[\pi(1+\alpha)\frac{t}{T}\right]\right\} \Big/ \pi\frac{t}{T}\left[1-\left(4\alpha\frac{t}{T}\right)^2\right] & t \ne 0, \pm\frac{T}{4\alpha} \end{cases} \qquad \text{(6-17)}$$

The filters were implemented as Finite Impulse Response (FIR) filters of user definable length. most of the reported results, however, were obtained using an FIR filter length of eight symbols with sixteen samples per symbol.

6.5.2 The Receiver

The receiver block diagram is shown in Figure 6-1. Both baseband and intermediate frequency (IF) simulations were coded. The IF simulations were used to calibrate the simulations and to derive the equivalent noise bandwidth of the square root raised cosine filters used in the receiver. Once calibrated, the baseband simulations were used to generate the required results at a much faster rate than allowed by the IF simulations. Implementation details relating to the receiver blocks are discussed next.

6.5.2.1 Carrier Recovery

The I-Q, differential detection techniques considered reduce the requirement of carrier phase recovery. The receiver oscillator, however, should oscillate at the transmitter frequency. In the

simulations conducted it is assumed that the transmitter and receiver oscillators operate at the same frequency, but are not phase synchronous.

6.5.2.2 Symbol Timing Recovery

Perfect symbol timing recovery was assumed in all simulations except in the case of the two-ray Rayleigh fading channel. In this case the symbol sampling time is set to t_x which is given by:-

$$t_x = \frac{R_1(t)\tau}{\left(R_1(t) + R_2(t)\right)} \qquad \textbf{(6-18)}$$

where $R_1(t)$ and $R_2(t)$ are the amplitudes of the two-Rayleigh fading processes at time t and τ is the time delay between the rays (refer to Figure 6-4).

6.5.2.3 Receiver filters

The receive filters are identical to the pulse shaping filter, with a square root raised cosine frequency spectrum given by equation (6-15) and time domain impulse response given by equation (6-16). The pulse and receive filters are thus matched and maximize SNR on the AWGN channel. The pulse and receive filters multiply in the frequency domain to generate a raised cosine filter response which obeys the Nyquist pulse shaping criterion and is hence ISI free in the absence of channel nonlinearities.

6.5.2.4 Statistics collection

The decoded receiver bit output is compared with the transmitted bit stream to generate the decoded error stream. This is achieved by X-ORing the two streams such that identical bits generate a '0' bit whereas dissimilar bits generate a '1' indicating a decoding error. The generated error streams can be read in directly by the IHC decoder and used to corrupt the video streams. Alternatively, the streams can be analyzed further to derive parameters needed for bit level models such as the GEM.

6.5.3 Simulating radio channels.

Four channels were simulated. The AWGN was mainly used to validate the simulation by comparison of the BEP results with published results. The AWGN generator developed is also used in the Rayleigh fading channel simulation. The simulation techniques developed for this channel are then used to simulate the Raleigh fading channel with one co-channel interferer, and the two-ray, frequency-selective fading channel.

6.5.3.1 Simulating an AWGN channel

The AWGN channel consists of a pseudo-random number generator with Gaussian distributed output, which corrupts an input stream by XOR addition. The Gaussian distributed random number generator is implemented using the Box Muller algorithm described above.

If white noise with two-sided spectral density $S_{NN}(f)=\eta/2$ is bandlimited to B, then the autocorrelation function is given by [Jeruchim92]

$$R_{NN}(\tau) = \eta B \frac{\sin 2\pi B\tau}{2\pi B\tau} \qquad \textbf{(6-19)}$$

Thus the ACF has zero crossings at $t=kT_s$ and $Ts=1/f_s$; *i.e.* the samples are uncorrelated in the time domain, and since the process is Gaussian they are also independent. Thus the bandlimited noise with variance σ^2 is simulated by generating one independent Gaussian Random Variable (RV) per sample time.

The two-sided spectral noise density $\eta/2$ is related to the variance and the sampling frequency through the equation:-

$$\sigma^2{}_{N_s} = \eta f_S / 2 \qquad \textbf{(6-20)}$$

6.5.3.2 Simulating Rayleigh fading channels.

The Jakes fading model [Jakes74] is a deterministic method for simulating time-correlated Rayleigh fading waveforms. This method was further refined by Dent *et al.* [Dent93] to produce more uncorrelated multiple fading waveforms. The method combines N_o randomly phased and appropriately weighted sinusoids to produce a complex Gaussian process X(t) with Power Spectral Density (PSD) S(v) and autocorrelation function $J_0(\omega_m\tau)$. The required result is obtained if

$$X(t) = \sqrt{\left(\frac{2}{N_o}\right)} \sum_{n=1}^{N_o} [\cos(\beta_n) + j\sin(\beta_n)]\cos(\omega_n t + \theta_n) \qquad \textbf{(6-21)}$$

where $\beta_n=\pi n/N_o$; $\omega_n=\omega_m\cos(2\pi(n-0.5)/N)$; $\omega_m=2\pi v/\lambda$ is the maximum Doppler frequency and $N_o=N/4$. Randomizing θ_n provides different waveform realizations. The mobile speed is v and the carrier wavelength λ. Multiple uncorrelated waveforms are realized by multiplying the summation terms by $A_j(n)$ the nth component of the Walsh-Hadamard (WH) codeword A_j giving

$$X(t) = \sqrt{\left(\frac{2}{N_o}\right)} \sum_{n=1}^{N_o} A_j(n)[\cos(\beta_n) + j\sin(\beta_n)]\cos(\omega_n t + \theta_n) \qquad \textbf{(6-22)}$$

The PSD of the electric field component $\mathbf{E_z}$ received by a moving vehicle, assuming an omnidirectional vertical monopole, is given by [Parsons89]

$$S_{E_z}(f) = \frac{1.5}{\pi . f_m}\left[1 - \left(\frac{f - f_c}{f_m}\right)^2\right]^{-1/2} \qquad \textbf{(6-23)}$$

where $f_m=v/\lambda$ and f_c is the carrier frequency.

6.5.3.2.1 Validation Results

The Rayleigh fading channel simulation was validated by generating 320,000 samples of the complex Gaussian process and comparing the Cumulative Distribution Function (CDF), the Autocorrelation Function (ACF), the Level Crossing Rate, the Fade Duration and the Normalized Power Spectral Density (PSD) with theoretical results from [Jakes74] and [Parsons89]. The simulated vehicle speed was 64 km/h, the carrier frequency was 2 GHz; and the simulated time was 100 seconds equivalent to 3200 samples per second. The first 0.3 seconds of the Rayleigh fading power envelope are shown in Figure 6-5.

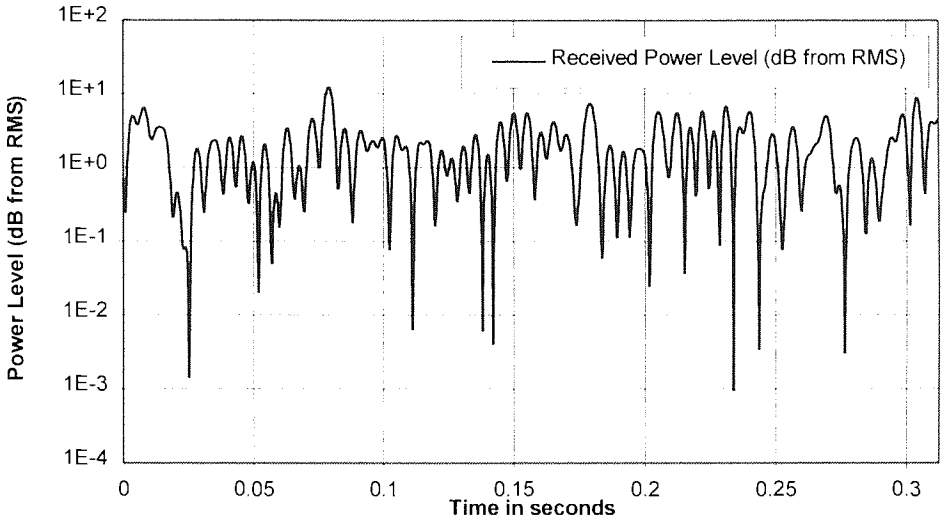


Figure 6-5 Rayleigh Fading Envelope

The mean and variance of the inphase and quadrature Gaussian processes are tabulated in Table 6-2. The mean μ_r and variance σ_r^2 of the Rayleigh envelope are given by [Proakis89]:-

$$\mu_r = \sqrt{\frac{\pi\sigma^2}{2}} \quad \text{and} \quad \sigma_r^2 = \left(2 - \frac{\pi}{2}\right)\sigma^2 \qquad \textbf{(6-24)}$$

where σ^2 is the variance of the zero mean Gaussian components. The theoretical and measured values are tabulated in Table 6-3. Both sets of values are close to the theoretical values.

	Theoretical value	Measured value
μ_r	0.886	0.893
σ_r^2	0.215	0.205

Table 6-2 Theoretical and Simulated mean and variance for the Rayleigh fading Process.

Component	Ideal Mean	Measured Mean	Ideal Variance	Measured Variance
In-Phase	0	0.000046	0.5	0.50225
Quadrature	0	0.000085	0.5	0.50039

Table 6-3 Ideal and measured mean an variance of the Gaussian Components

The cumulative distribution function of the Rayleigh envelope is given by [Proakis89]:-

$$F_R(r) = 1 - e^{-r^2/2\sigma^2} \qquad (6\text{-}25)$$

The autocorrelation function is given by equation (6.26) [Parsons89]:-

$$R_{xx}(\tau) = J_o(\omega_m \tau) \qquad (6\text{-}26)$$

the level crossing rate is given by [Parsons89]:-

$$N(R) = \sqrt{2\pi} f_m \rho \exp(-\rho^2) \qquad (6\text{-}27)$$

and the fade duration is given by [Parsons89]:-

$$\tau(\rho) = \frac{\exp(\rho^2) - 1}{\sqrt{2\pi} f_m \rho} \qquad (6\text{-}28)$$

The theoretical results given by equations (**6-25**) to (6-28) are plotted together with the simulation results in Figure 6-6 (a) - (d). The agreement between the two sets of results is excellent. The theoretical results were obtained by setting $\sigma^2 = 0.5$, $\omega_m = 2\pi f_m$, $f_m = v/\lambda_c = 17.77[m/s]/0.15[m] = 118.52$ *Hz*, $\rho = R/R_{rms}$ R is the specified level and R_{rms} is the rms amplitude of the fading envelope, in the above equations.

Finally, the RF spectrum of the simulated Rayleigh fading carrier is shown plotted in Figure 6-7 together with the theoretical RF Spectrum obtained by taking the square root of the PSD equation (6-23). The normalized frequency in the figure is given by $f_N = (f-f_c)/f_m$. The RF spectrum is also normalized to return unity at $f_N=0$. This result agrees well with the RF spectrum plotted in Figure 1.7-5 in [Jakes74].

The excellent agreement between the simulated and theoretical results validates the Rayleigh fading channel simulation, which is used extensively in this thesis.

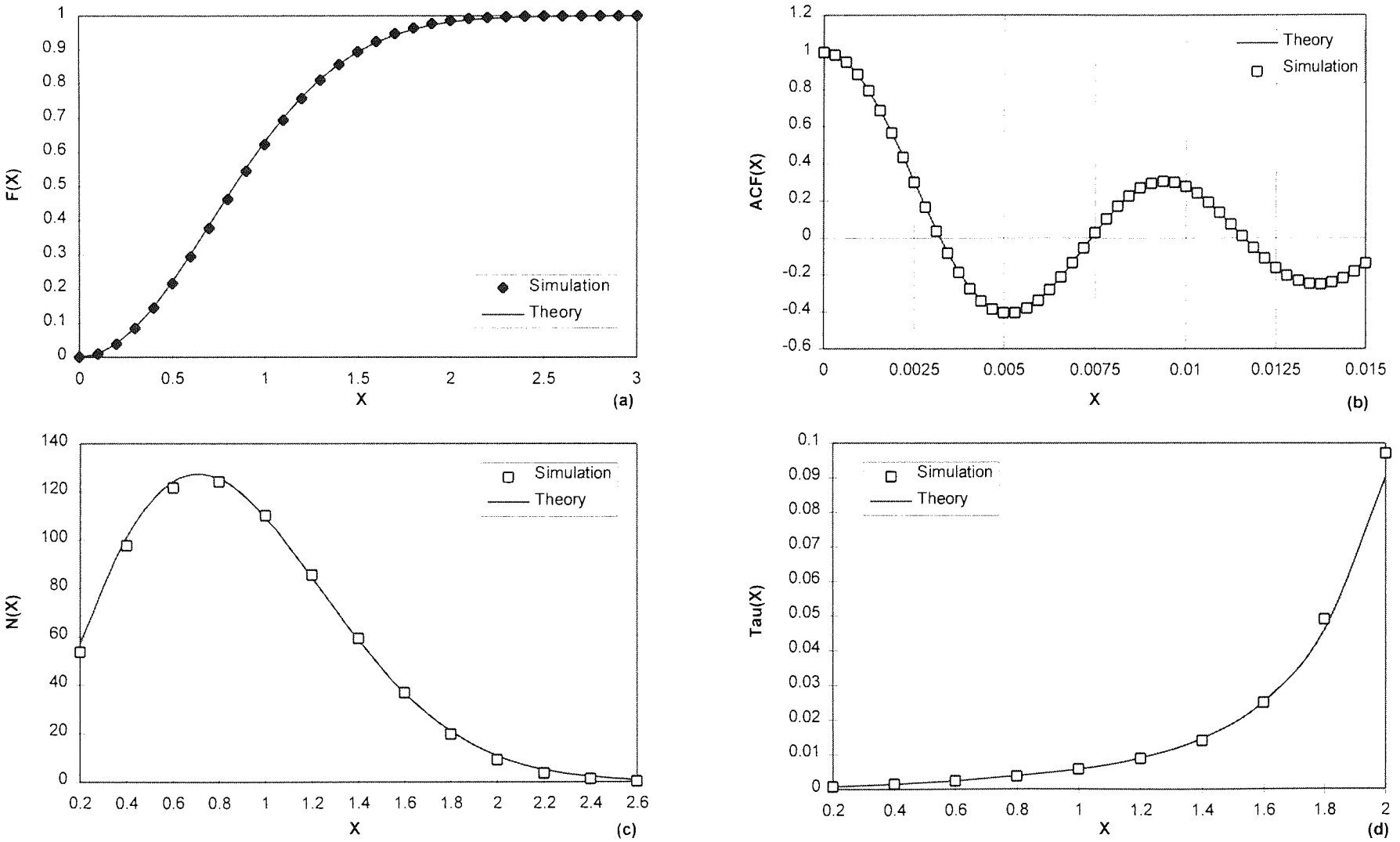


Figure 6.6 Comparison of the Simulated Rayleigh Fading Envelope Statistics with Theoretical Curves(a) Cumulative Distribution function (b) Autocorrelation Function (c) Level Crossing Rate (d) Interfade Duration.

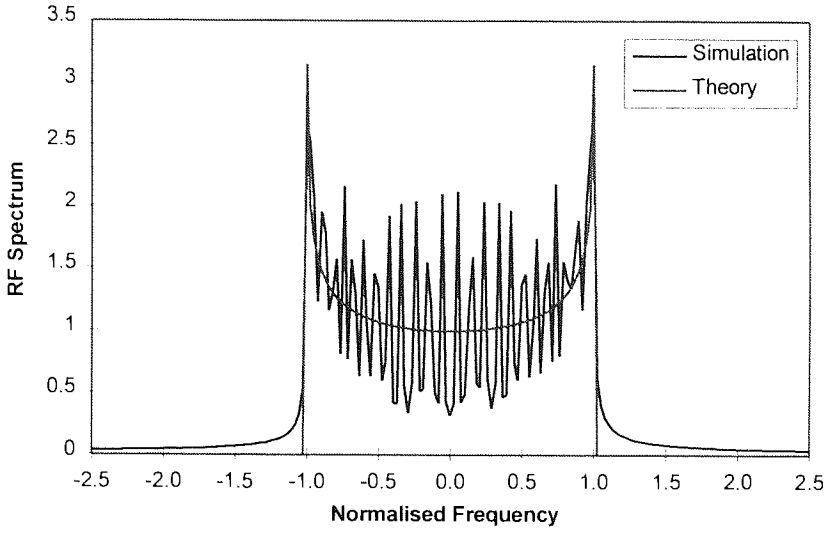


Figure 6-7 RF Spectrum of Rayleigh Fading Envelope

6.5.3.3 Rayleigh Fading Channel with Co-Channel Interference

The simulation of the Rayleigh fading channel with co-channel interference follows closely the theoretical description in section 6.4.2 and Figure 6-3. The Rayleigh fading channels and the AWGN are simulated using the techniques described above.

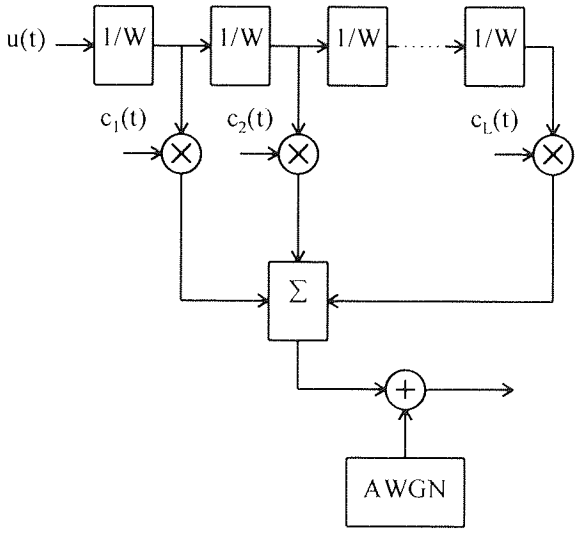


Figure 6-8 Tapped delay line Wideband channel model

6.5.3.4 Simulating wideband channels

The time variant wideband (frequency-selective) channel can be modeled as a truncated tapped delay line with tap spacing 1/W and time-variant tap weight coefficients { $c_n(t)$ }, where W is the bandwidth of the channel and $c_n(t)$ are zero mean complex-valued stationary Gaussian random processes [Proakis89]. The tapped delay line mode of the wideband channel is shown in Figure 6-8.

The resolution achieved in the multipath delay profile is 1/W and the maximum simulated delay spread is (L-1)/W.

Each of the L complex-valued stationary Gaussian random processes can be generated using the Rayleigh fading channel simulator described above. L independent fading rays are required. The AWGN is simulated as described above.

6.5.3.5 Two-ray channel model

The two-ray channel model (Figure 6-4) is often used to model wideband, frequency-selective channels. It is the simplest realization of the tapped delay line model, with just two taps. The model is mathematically tractable and is, therefore, frequently used in the literature. Furthermore, Chuang has shown [Chuang87] that the BER statistics of digital modulation techniques are proportional to the RMS delay spread, but are largely insensitive to the shape of multipath delay profile. Hence the two-ray channel model generates results which are as accurate as other models employing more rays, which better model the multipath delay profile.

Recently Fung *et al.* have shown [Fung93] that the two-ray channel underestimates the BER for the Indoor Radio channel, indicating that the two-ray channel model should be used with care, and that there are instances in which the multipath delay profile must be taken into account

The two-ray channel has been simulated by generating two complex-valued zero mean random Gaussian processes using the Jakes technique described above and introducing a delay τ between the two paths as shown in Figure 6-4. The received signal may also be corrupted by AWGN which is simulated as described above.

6.6 Simulation Results

The simulation methodology was validated by simulating the performance of QPSK, DQPSK, $\pi/4$ DQPSK, 8-PSK, 8-DPSK, 16-PSK and 16-DPSK on four channels. These were the AWGN channel, the Rayleigh fading channel; the Rayleigh fading channel with co-channel interference and the two-ray Rayleigh fading, frequency-selective channel. The average Symbol Error Probability (SEP) and the average Bit Error Probability (BEP) were then compared with theoretical and published results. Very close agreement was found in all the cases considered.

6.6.1 Simulation results for the AWGN channel

The probability of symbol and bit errors for the various modulation techniques were measured in AWGN. Both coherent and differential demodulation were considered. Perfect carrier recovery was assumed for coherent demodulation. For differential reception, the local oscillator was assumed to run at the correct frequency but without phase synchronization. Perfect symbol timing recovery was assumed for all cases. The simulation results are plotted together with the theoretical results in Figure 6.9. The theoretical results were obtained from equations (6-29) and (6-30).

The probability of symbol error for M-ary PSK (MPSK) $P_S(M)$ is approximated by [Proakis89]:-

$$P_S(M) \approx erfc\left(\sqrt{(Log_2 M)\gamma_b}\,\sin\left(\frac{\pi}{M}\right)\right) \qquad \textbf{(6-29)}$$

where $\gamma_b = E_b/N_o$, E_b is the Energy per bit and N_o is the spectral density of the AWGN. Exact theoretical results are also tabulated in [Lindsey73].

Lindsey and Simon [Lindsey73] derived an approximate equation for M level Differential PSK (M-DPSK) given by:-

$$P_S(M) \approx erfc(U) + \frac{U\exp(-U^2)}{4\sqrt{\pi}(1/8+\gamma_b k)} \qquad \textbf{(6-30)}$$

where $U = \sqrt{(2k\gamma_b)}\sin(\pi/2M)$, k = $\log_2$M, and $\gamma_b = E_b/N_o$ as above.

When Gray encoding is used, the bit error probability P_b can be approximated by [Proakis89]:-

$$P_b = \frac{1}{k}P_S(M) \qquad \textbf{(6-31)}$$

The coherent demodulators perform 2 to 2.5 dB better than the differential demodulators. QPSK performs almost 4 dB better than 8-PSK, which is about 4 dB better than 16-PSK. Although QPSK has a 8 dB advantage over 16-PSK, the latter carries twice the bit rate in the same bandwidth. The performance of $\pi/4$ DQPSK was found to be the same as DQPSK, so that the DQPSK curve in Figure 6.9 is also valid for $\pi/4$ DQPSK.

The close agreement between the simulations and the theoretical results confirm that the implementation of the different modulation and detection techniques are correct, and that the results obtained from the simulations can be used with confidence.

The errors on the AWGN channel are randomly distributed and are well modeled as a Poisson arrival process. Thus, given the probability of error, it is relatively straightforward to simulate the errors at bit level using a uniform random number generator. Alternatively, since interarrival times are negative exponentially distributed, a faster implementation uses a random number generator with negative exponential distribution to simulate the error-free gaps [Jeruchim92].

6.6.2 Rayleigh fading channel

The probability of symbol error versus average γ_b is shown if Figure 6.10. The simulated baud rate was 24 kbaud, the vehicle speed was 64 km/h and the carrier frequency 1.8 GHz. The transmitter pulse shaping filters and the receiver matched filters were square root raised cosine filters with α=0.35. The cascade of the two filters is ISI free on a linear channel. The sampling rate was sixteen samples per symbol period and the filters were eight symbols long.

As already mentioned, carrier recovery in the presence of Rayleigh fading can be quite difficult as the signal emerges from a deep fade. It has also been shown [Webb91] that some square

constellations e.g. 16-QAM, exhibit false locking due to non-linear constellation distortion, which results in very long error bursts. For this reason differential detection was selected for further study since it does not require perfect carrier phase recovery. Furthermore, the QAM constellations studied do not exhibit false lock positions.

Proakis [Proakis89]. derived an approximate equation for the probability of symbol error for M-ary DPSK (MDPSK), given by:-

$$P_S(M) \approx \frac{M-1}{\left(M \log_2 M\right)\left[\sin^2\left(\pi/M\right)\right]\bar{\gamma}_b} \qquad \textbf{(6-32)}$$

The equation was derived assuming a slowly fading channel such that it does not model the irreducible error floor introduced by large inter-symbol phase changes as the mobile passes through a deep fade. The probability of symbol error curves are linear when plotted on a logarithmic scale since equation (6-32) is of the form $P_S(M) = k/\bar{\gamma}_b$.

The experimental results agree well with the theoretical results predicted by equation (6-32) between 10 dB and 25 dB, before the irreducible error floor effects become significant. The curve for PI/4 DQPSK agrees very well with Figure 6 in [Chennakeshu93]. The irreducible error floors are very close at $f_D T_S = 4.5\text{x}10^{-3}$, where $f_D = v/\lambda_C$ is the Doppler frequency, v is the vehicle speed and λ_C is the carrier frequency.

A comparison of Figures 6.9 and 6.10 reveals that the performance in Rayleigh fading is appreciably worse than on an AWGN channel. The performance of DQPSK and π/4 DQPSK are again very close on the Rayleigh fading channel, and these two modulations perform better than 8-DPSK by about 5 dB. 8-DPSK performs 5 dB better than 16-DPSK.

During simulation, the transmitted and received bit streams are X-ORed together, such that an error generates a '1' and a correctly received bit generates a '0'. The resulting error stream is stored in an error file for later use. The close agreement of the first order statistics of the simulation results with theoretical and published simulation results, in combination with the validated Jakes Rayleigh fading channel model, guarantees the validity of the second order statistics, including the fade and the interfade duration statistics, of the generated error files.

6.6.3 Simulation results for the Rayleigh fading channel with CCI

The CCI channel model in Figure 6-3 was used to measure the performance of the MDPSK modulations on the Rayleigh fading channel with CCI. The simulated baud rate, vehicle speed and carrier frequency were the same as in section 6.6.2, for both Rayleigh fading channels. The additive white Gaussian noise generator was switched off to concentrate on the CCI.

The results from the simulation, for differential detection of the differential modulations, are shown in Figure 6.11 where the symbol error probability is plotted against the Signal to Interferer Ratio (SIR) given by:-

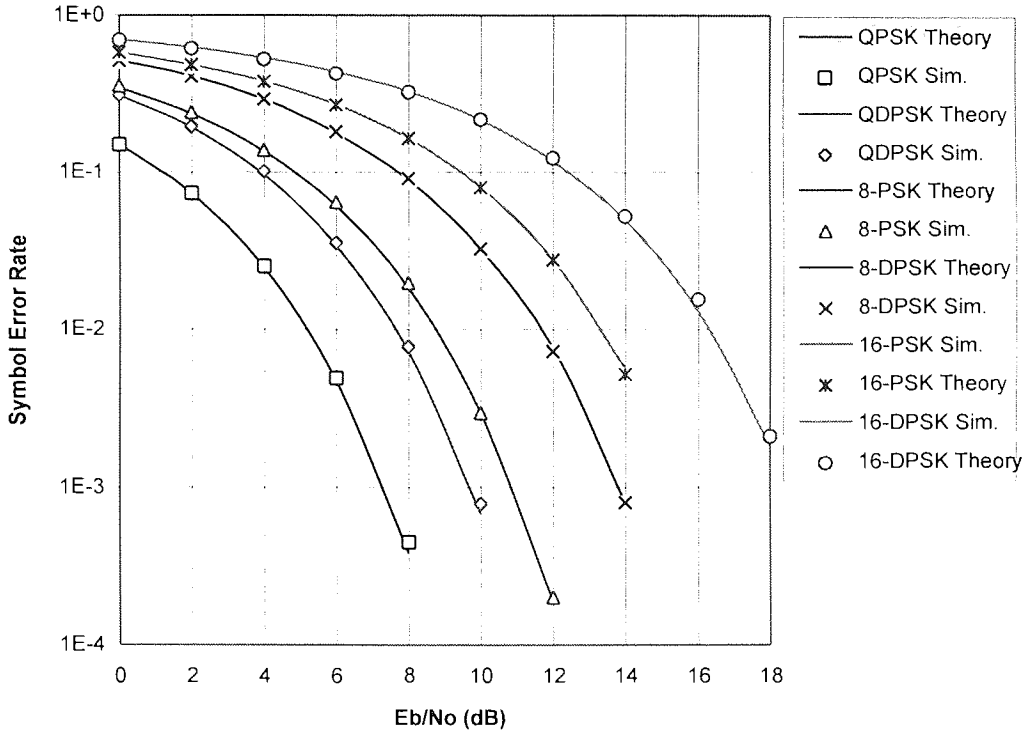


Figure 6.9 Symbol Error Rate versus E_b/N_o for the static AWGN Channel

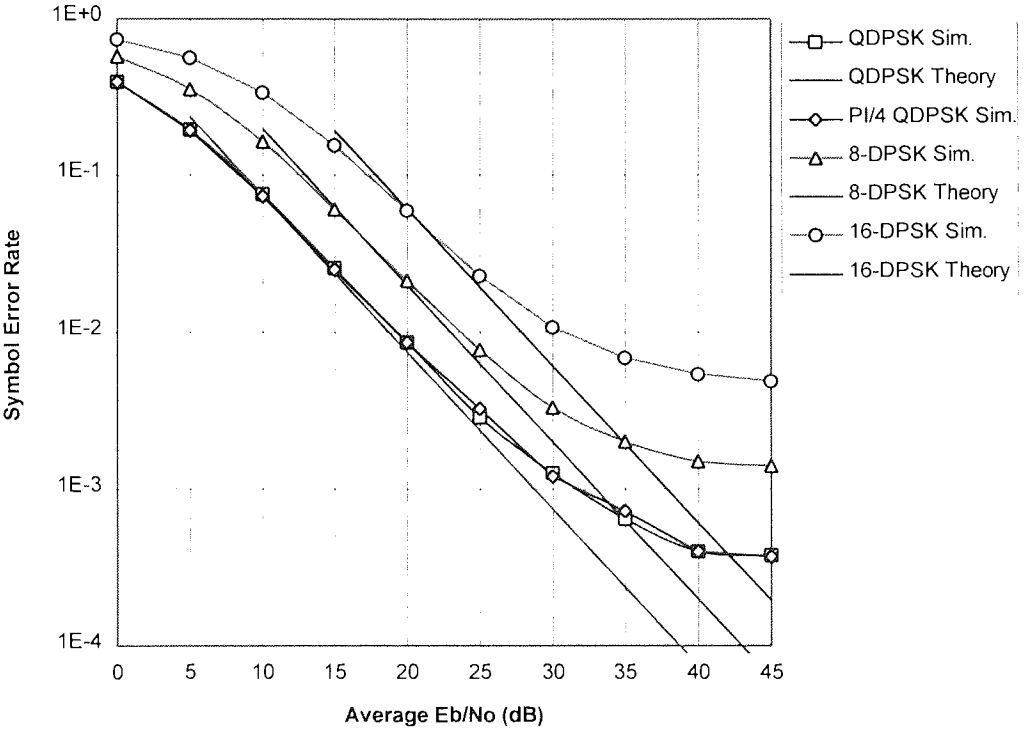


Figure 6.10 Symbol Error Rate versus Average E_b/N_o in Rayleigh Fading ($f_DT_S=4.5x10^{-3}$, C/I=∞)

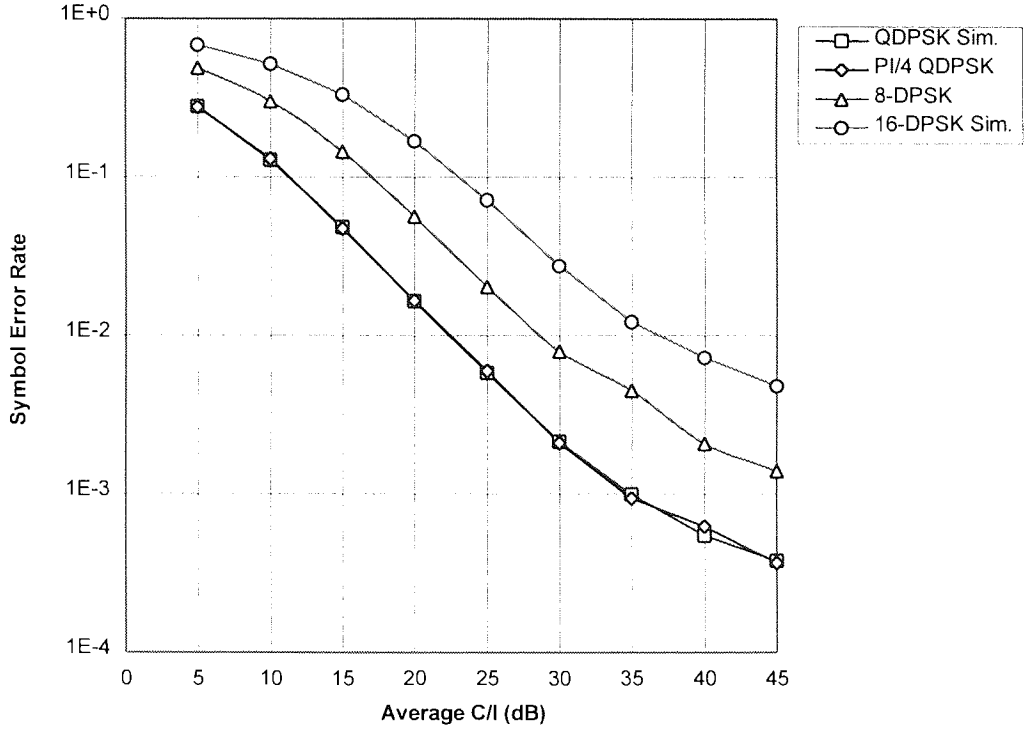


Figure 6.11 Symbol Error Rate versus Average C/I in Rayleigh Fading ($f_D T_S$=4.5x10-3, Eb/No=∞)

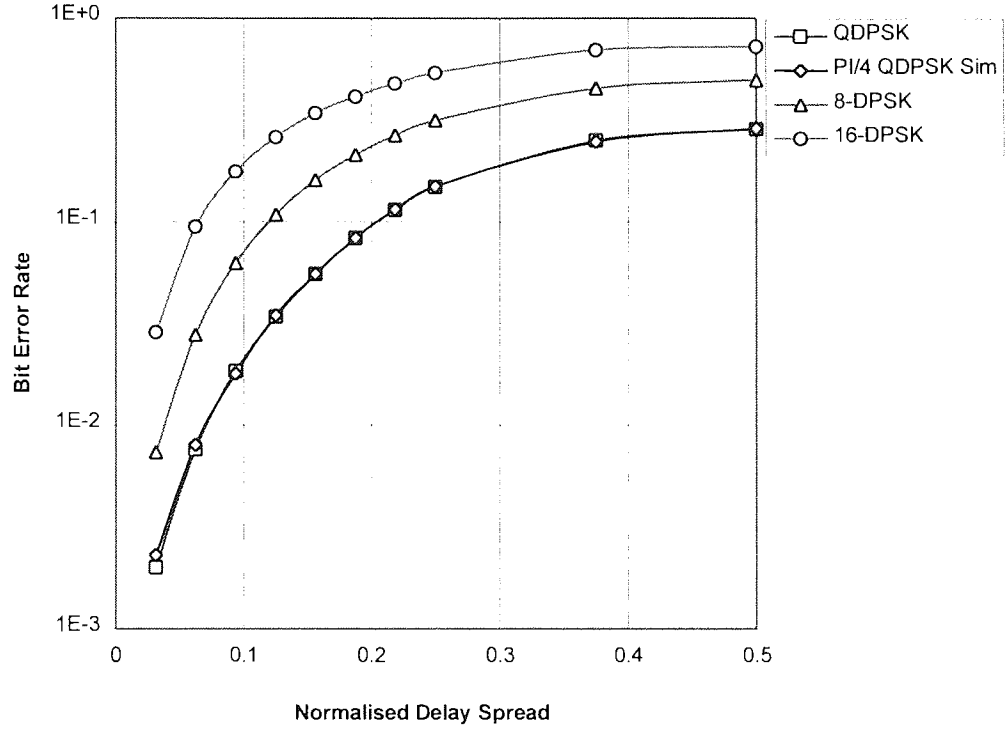


Figure 6.12 Bit Error Rate versus τ_{rms} in the two-ray frequency-selective channel

$$SIR = 10\log_{10}(P_S/P_I) \quad \textbf{(6-33)}$$

where P_S is the main signal power and P_I is the interfering signal power.

The π/4-DQPSK results agree very well with the results in Figure 10 [Chennakeshu93]. The irreducible error floors at $f_D T_S = 4.5 \times 10^{-3}$ are also in close agreement. Again, it is difficult to distinguish between the π/4-DQPSK and the DQPSK results, since the they are very close to each other

The main limitation on cellular mobile networks is due to CCI since the need of spectrum efficiency through frequency re-use, leads to reduced re-use distances and hence relatively high levels of CCI. For example, the DECT standard requires the BER to be less than or equal to 10^{-3} at a SIR of 10 dB [Schultes90].

The modulations studied here cannot approach this level of robustnes, and other countermeasures have to be studied to achieve the required performance. In the next chapter two countermeasures; interleaved FEC and diversity are used to reduce the BER due to CCI.

6.6.4 Frequency-selective channel

The two-ray, Rayleigh fading, frequency-selective channel has been widely used to study discrete multipath frequency-selective channels [Chennakeshu93, Fung93, Chuang91] and it has also been proposed as a an appropriate model for digital mobile modems research during the development of the TIA second generation digital radio system [Fung93]. This model has been used to simulate the error performance of the digital modulations of interest in a fading, multipath mobile radio environment. Chuang [Chuang87] studied MPSK and MSK modulations on different discrete multipath channel models and concluded that the BER performance is more sensitive to rms delay spread than to the shape of the power delay profile.

The channel model used is that shown in Figure 6-4. The delay τ was varied in steps of one symbol period T_S from $\tau = T_S$ to $\tau = 16.T_S$. The root mean square (RMS) delay spread for the two-ray, equal average power model with the two rays separated by τ, is equal to τ/2. Hence the RMS delay spread resolution is $T_S/2$. The baud rate, vehicle speed and the carrier frequency are the same as in the previous cases.

The results for differential reception of the differential modulations are plotted in Figure 6.12. The π/4-DQPSK results are in close agreement with the results plotted in Figure 6.11 [Chennakeshu93]. The π/4-DQPSK and DQPSK results are again superimposed. Figure 6.12 indicates that the selected modulations can tolerate very little delay spread and that anti-multipath measures have to be taken to render the modulations usable on this channel, unless the normalized delay spread is limited to below 0.01.

6.7 Conclusions

This chapter, described the simulation methodology used to generate error vector files with

good first-order and second-order statistics that match real-world situations as much as possible. The error vector files are essential for the study of video transmission over wireless networks. These simulation methodologies, in combination with the video coding simulations presented in chapter four, form the simulation framework developed in this thesis.

A group of digital modulation techniques have been selected for further study, based on their suitability for digital mobile environments. The modulations are M-ary differential PSK techniques with M=2,4 and 8 and $\pi/4$ shifted DQPSK. These modulations have good spectral efficiency, and do not require complex carrier phase recovery techniques.

The error performance of these modulations has been simulated on four different channels and the results validated with theoretical and other published results in the literature. The channel models chosen are typical models, used extensively in the literature to study digital modulation in a mobile radio environment. The channel models are the AWGN channel, the Rayleigh fading channel, the Rayleigh fading channel with CCI and the two-ray, Rayleigh fading, frequency-selective model.

The AWGN model is used to validate the simulations whereas the other three channel models are used to generate error vector files to be used in subsequent chapters. The Rayleigh fading channel is the basis of all three mobile channel models and hence the simulation of this channel was validated with particular care. The Jakes simulation methodology was employed to generate the complex Gaussian process which characterizes the Rayleigh fading channel and the CDF, ACF, Interfade Duration, Fade Duration and Electric Field Spectrum were measured and validated against the published theoretical results. The very good agreement between the two sets of results should guarantee the accuracy of the second-order statistics of the simulated error vectors used in this thesis.

The symbol error probability of the modulations as a function of E_b/N_o were obtained for the AWGN channel and compared with theoretical results. On the Rayleigh fading channel the symbol error probability was plotted against the average received E_b/N_o, whereas on the CCI channel the independent variable is the main signal power to interfering signal power ratio. On the frequency-selective channel, the symbol error probability was measured whilst varying the normalized RMS delay. In all the above cases, there was close agreement with theoretical and published results, further consolidating the validity of the simulation results.

The simulations produce error files where '0' bits signify correctly received bits and '1' bits signify transmission errors. These error files are used to corrupt digital video streams in the manner introduced in chapter 1. The results can also be analyzed to derive second order statistics and hence to abstract the appropriate parameters to drive Gilbert-Elliot Models [Gilbert60] or its derivatives [Fitchman67] to simulate typical burst error channels at bit level.

Chapter 7

Video Transmission Over Narrowband Rayleigh Fading Channels

This chapter studies the transmission of video over narrowband, Rayleigh fading channels. The transmission baud rate is selected such that the required bandwidth is lower than the coherence bandwidth of the channel. Then the channel does not exhibit frequency selective fading, ISI is absent and channel equalisation is not required.

A wireless access network is proposed and the radio link is modelled and simulated in software. The MAC layer uses TDM on the downlink, circuit switched TDMA on the uplink and frequency division duplex (FDD). The physical layer uses MDPSK modulation, a class of rotation invariant, low complexity modulation techniques. The architecture is star based, with a central base station servicing mobiles in a micro-cellular or pico-cellular environment.

The second-order error burst statistics of the radio link are obtained by simulation, and the error vectors generated are then used to corrupt H.261 coded video. A number of source coding, transcoding, channel coding and Rayleigh fading countermeasure techniques are proposed and studied, to achieve an effective and efficient digital video transmission over the wireless network.

7.1 Introduction

The classical narrowband mobile radio channel is a flat fading channel with a n th law attenuation loss function (n>2, typically 3 to 4), shadow fading with a Log-Normal distribution and a Rayleigh or Ricean distributed fading envelope [Hashemi93]. The performance of digital modulation techniques over this channel are much worse than over the AWGN channel as described in chapter six. The average E_b/N_0 ratios required to achieve a BER of 10^{-3}, typically associated with wireless voice transmission, are high, above 25 dB for the studied modulations. The next generation of microcellular and pico-cellular wireless networks, however, can deliver such high E_b/N_0 ratios [Stedman92]. In these systems the rms. delay spread is usually low, frequently well below 1 μs, and several hundred of kilo-symbols per second can be transmitted without the channel becoming dispersive.

In micro-cellular and pico-cellular systems both Ricean and Rayleigh multipath statistics have been observed [Hashemi93]. Ricean multipath fading statistics are usually observed where a direct line-of-sight path exists between the transmitter and the receiver. The performance of digital modulation techniques is generally worse on a Rayleigh fading channel than on the Ricean fading channel. Therefore, the Rayleigh fading channel, should produce worst case results.

A number of authors have studied video transmission over Rayleigh fading channels. Vaisey, Yuen and Cavers [Vaisey92] proposed a system for transmitting video over a mobile network at 64 kbit/s using a modified H.261 codec. Recently, Stedman *et. al.* [Stedman92] proposed a system for transmitting video over a narrowband channel using a sub-band based video codec and a 16-QAM radio link, with diversity, interleaving and BCH coding. In a later paper Hanzo *et. al.* [Hanzo94] studied video transmission in a wireless environment using a DCT based video codec using 16- and 64-QAM modems and a PRMA MAC system. The approach adopted in this chapter is to study the transmission of H.261 coded video over a wireless access network such as might be adopted in a micro-cellular or pico-cellular environment. The requirement to interface to H.230 standard videoconferencing systems is assumed, and this leads to the inclusion of a transcoder at the interface between the fixed and wireless networks.

The overall system description is presented in the next section. The basic burst error statistics of the MDPSK modems is then studied by simulation. The error vectors so obtained are used to investigate the transmission of coded video over the simulated link at 4 km/hr and 64 km/hr using QDPSK and 16-DPSK. Problems with these basic systems are identified and solutions based on diversity and interleaved FEC are proposed and studied. The issue of CCI is studied and an error control technique is proposed to allow efficient video transmission in CCI limited wireless networks. The conclusions are summarised in the final section.

7.2 System Description

The transmission system model used to study H.261 video transmission over digital radio channels is shown in Figure 7-1. The video codecs investigated are standard H.261 video codecs. The Modified H.261 codec shown on the mobile terminal side is basically a combined H.261 codec and transcoder. As already described, voice transcoders are widely employed to interface the robust, spectrum efficient, compressed voice codecs used on the radio links to the simpler, but less efficient, codecs used on fixed networks. The same concept is used here. The transcoders contemplated, however, have low complexity and encoding delay as described in chapter 5.

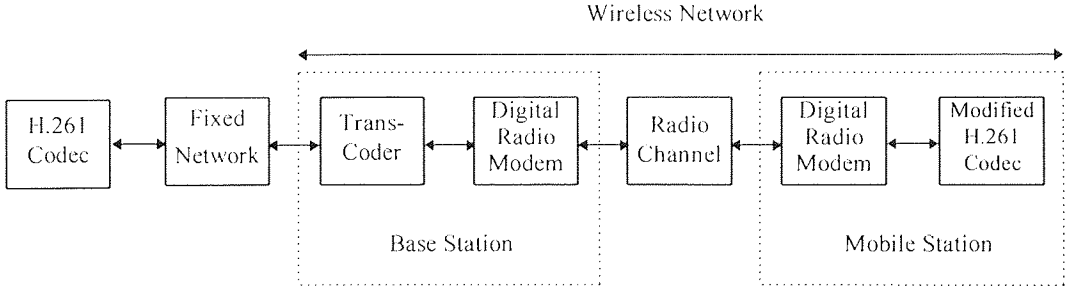


Figure 7-1 Video Transmission over a Wireless Access Network - System model

7.2.1 Wireless Network Topology.

The basic Wireless Access Network topology is shown in Figure 7-2 below. The wireless network consists of a number of base stations (BS) each of which provides cellular coverage to mobile stations (MS) in the vicinity of the base station. Regions of overlap exist between cells, where any one of the surrounding base stations could provide service, and the one offering the strongest signal is usually selected. A mobile may traverse from one cell to another, and the radio link must then be handed over to the new base station before the previous link deteriorates and becomes unusable.

The base stations are assumed to be internetworked using a fixed local area network. The local area network interfaces with the global, public and private Wide Area Networks (WAN) via WAN interface units. In this thesis the LAN and WAN fixed networks are grouped together and are referred to as the Fixed Network. This fixed network is assumed to have a bit error performance which is significantly better than on the Wireless network side, so that the fixed network BER can be ignored.

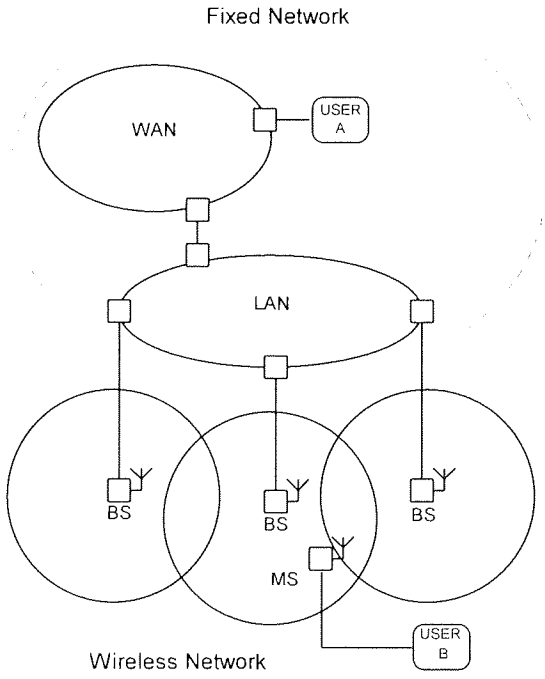


Figure 7-2 Wireless Access Network Topology

7.2.2 Network Architecture

A layered network architecture is assumed as shown in Figure 7-3. The emphasis in this thesis is on the physical and data link layers. The data link layer is resolved into the Multiple Access Control (MAC) and the Error Control sub-layers. The mobile video codecs use one communications stack, whereas the transcoder uses two stacks to interface the wireless and fixed networks.

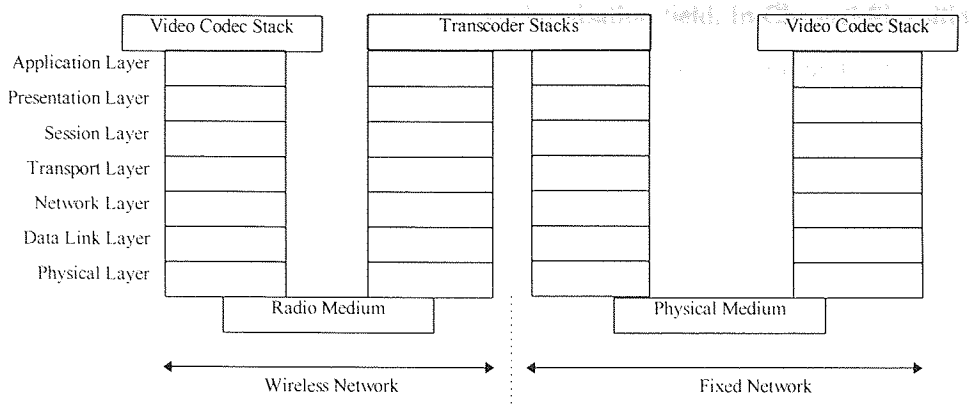


Figure 7-3 Wireless Communications Network Architecture

7.2.3 The Physical Layer

The digital radio modems are assumed to operate at a frequency of 1.9 GHz. The modulations studied are the class of rotation invariant, low complexity M-ary Differential PSK techniques introduced in chapter 6. The transmission baud rate is 192 kbaud and the radio channel is assumed to be Rayleigh fading, initially without either ACI or CCI. The multipath countermeasures of interest, *e.g.* diversity, belong to the physical layer.

7.2.4 The Multiple Access Control Sub-layer

The Multiple Access Control (MAC) adopted uses Time Division Multiplexing (TDM) on the down link (BS to MS) and circuit switched TDMA on the up link (MS to BS). The TDM structure is shown in Figure 7-4 below. Frequency Division Duplex (FDD) is used. There are 100 frames per second with each frame supporting a number of channels. A particular realisation consists of eight bearer channels per frame. Each channel consists of a 64 symbol header and a 160 symbol packet. A guard band of 16 symbols separates the channels. Each bearer channel supports a raw bit rate of 32 kbit/s. This configuration is similar to that used in DECT, PACS and PHS [Goodman91a].

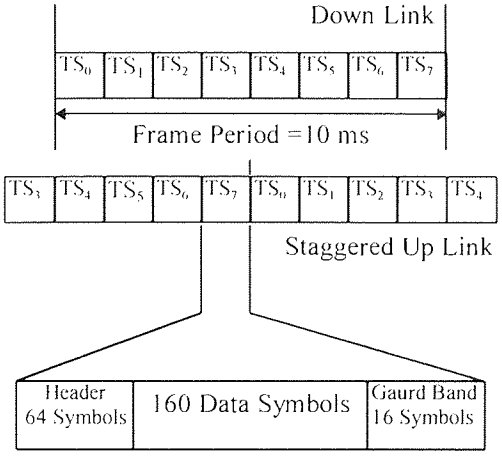


Figure 7-4 MAC Frame and Packet Structure

The header may include the following; the synchronisation field, In-Channel Signalling fields and a header cyclic redundancy check (CRC) field. The in-channel signalling field is needed to establish, service and terminate calls, to achieve handover and for techniques such as dynamic channel allocation, frequency hopping and Variable Rate Modulation control.

7.2.5 The Error Control Layer

The error control layer adds channel coding to the signalling and data fields prior to collation into MAC packets by the MAC layer. Thus the BCH codes considered in this chapter belong to the data link layer. Interleaving is performed here as well.

Two error control packet formats have been defined to utilise the 32 kbit/s MAC bearer channels described above. One packet structure, shown in Figure 7-5, is targeted at systems which use error detection rather than FEC, for error control. The error detection packet consists of 320 data bits and a 16 bit CRC with the remaining 56 bits used for the header. The CRC failure probability is approximately 2^{-16} which is significantly lower than the packet loss rate due to corruption, and can be neglected. If a more secure link is needed the CRC field can be augmented at the cost of header bits or guard band bits.

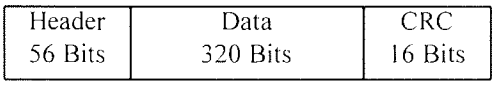


Figure 7-5 Error Detection Packet Format

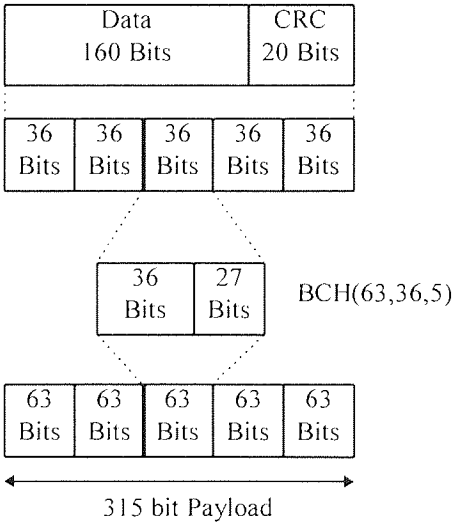


Figure 7-6 FEC Packet Format

The second packet structure shown in Figure 7-6 makes use of a low-rate BCH(63,36,5) forward error correction code to achieve a coded channel throughput of 16 kbit/s, which is 50%

efficient with respect to the error detection packet format described above. Each packet carries 160 information bits and a 20-bit CRC field which is used to detect FEC decoding failures. The 180-bit payload is divided into five 36-bit fields which are coded using the BCH(63,36,5) code to produce 315 bits which fit into the MAC bearer channel packets. This particular BCH code has been chosen, rather than a longer, more efficient code, because of its lower complexity. BCH(63,k,t) FEC codes have been proposed for other systems [Stedman92, Hanzo94]. Note that the five 63-bit codewords can be interleaved to randomise error bursts.

7.2.6 Transcoder

The transcoder performs the interfacing function between the fixed and wireless networks. In particular, it terminates the H.261 error control protocol on the fixed network side and introduces the more robust error control packets described above on the wireless network side. The transcoder must also match the rate of the two systems. In this example the H.261 video stream rate and the wireless transport are well matched. The H.261 video rate is assumed to be 32 kbit/s which can be accommodated using one error-detecting data link packet or two FEC, data link packets. The transcoder, however, must buffer the 492 information bits in the H.261 error framing packet and re-packetize these into the 320 or 160 data link packet formats. The transcoders considered here are low-complexity, low-delay transcoders as described in chapter five.

7.3 Digital Radio Modem Performance in Narrowband, Rayleigh Fading channels

The radio link described above has been simulated using the techniques described in chapter 6. Two mobile speeds are considered; 64 km/hr for vehicle bourne video codecs and a more pedestrian 4 km/hr for handheld video codecs. 512 symbols of the Rayleigh fading process are generated at 10 ms intervals, at the two speeds, for a simulated frequency of 1.9 GHz. The modem performance, using the four modulations of interest, at the two speeds and 192 Kbaud is shown in Figure 7-7. 1.5 million symbols have been simulated to obtain the results. The simulations use square root raised cosine filters with excess bandwidth α=0.35, implemented as FIR filters 16 symbols long, and with a sampling rate of eight samples per symbol. The results agree very well with theoretical results in [Proakis89]. The performance for QDPSK is nearly identical to π/4-QDPSK. There was no significant difference between the performance results at the two different speeds.

A comparison of Figure 7-7 with the performance charts in Figure 6.10 reveals the absence of an error floor in the current results. This is as expected, since the baud rate here is 192 Kbaud, substantially higher than the 24 Kbaud in chapter 6.

The QDPSK modulation achieves a BER of 10^{-3} at an average E_b/N_0 of about 26.5 dB. 8-DPSK requires 29 dB and 16-DPSK requires 32.5 dB to achieve the same BER. The error vector streams generated at this BER have been analysed. The BER per 256 symbol packets for the three

modulations at the two mobile speeds are shown plotted in figures 7.8 and 7.9. Both figures confirm the bursty nature of the error processes. The packet errors occur in nearly identical positions for all modulations and coincide with deep Rayleigh fades at both speeds. The number of errored packets increases from 50 per 3000 for the quartenary modulations to 72 for 8-DPSK and 106 for 16-DPSK at 4 km/hr. At 64 km/hr, the number of errored packets approximately doubles for all modulations with 116,175 and 217 errored packets for QDPSK, 8-DPSK and 16-DPSK respectively.

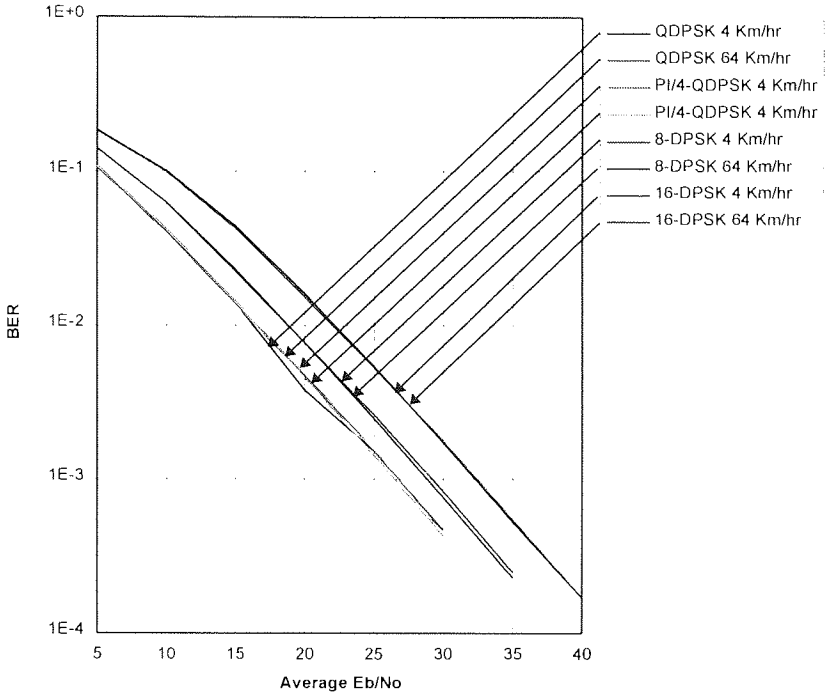


Figure 7-7 MDPSK BER on the Rayleigh Fading Channel

Figure 7-8 and 7-9 show that the mobile goes through more fades per second at the higher speed, such that there are more blocks in error in Figure 7-9 than in Figure 7-8. Since the number of errored packets increases at the higher speed and the BER remains the same, then the BER per packet should decrease at the higher speed as can be confirmed from Figure 7-8 and Figure 7-9.

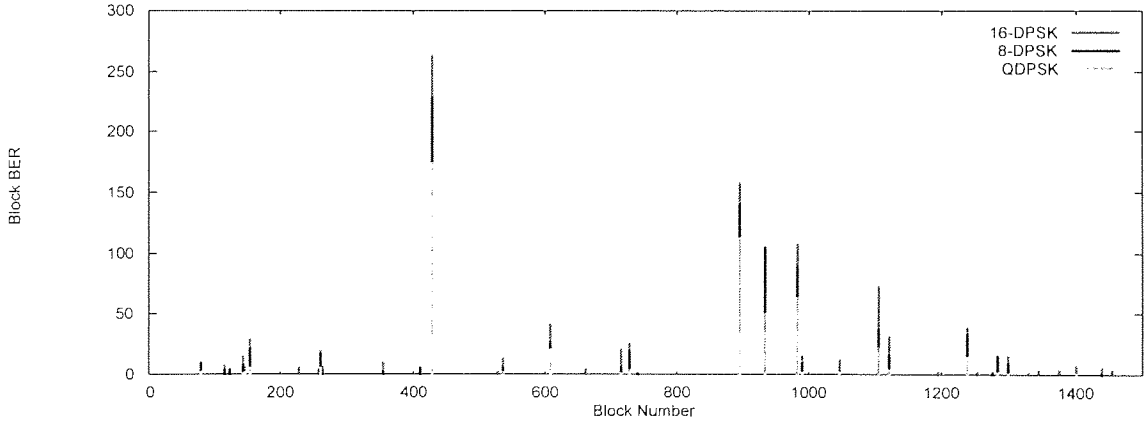


Figure 7-8 BER per block at 4 km/hr

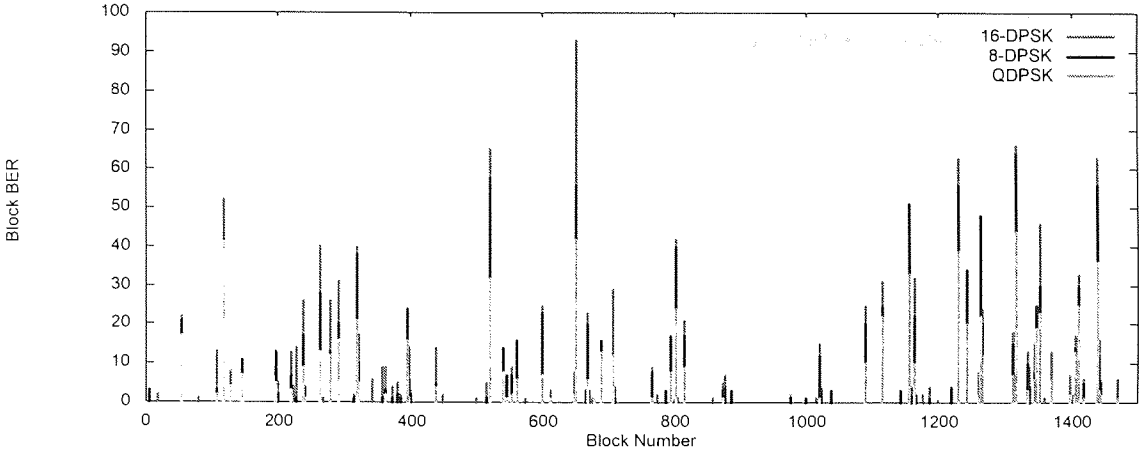


Figure 7-9 BER per Block at 64 km/hr

The block error statistics, in the form of the Consecutive Block Error distribution and the Errored Block Gap distribution have been derived and are plotted in Figure 7-10 to Figure 7-12 below. The mean Consecutive Block Error burst is tabulated in Table 7-1 below. The results show that the MAC frame structure breaks up the error bursts, such that errored blocks occur predominantly independently. This effect is more pronounced at higher speeds. At 4 km/hr the 16-DPSK modulated blocks have the highest burst mean at 1.395.

Velocity	16-DPSK	8-DPSK	PI/4-QDPSK	QDPSK
4 km/hr	1.395	1.263	1.064	1.04
64 km/hr	1.098	1.080	1.025	1.035

Table 7-1 Block Error Burst Means

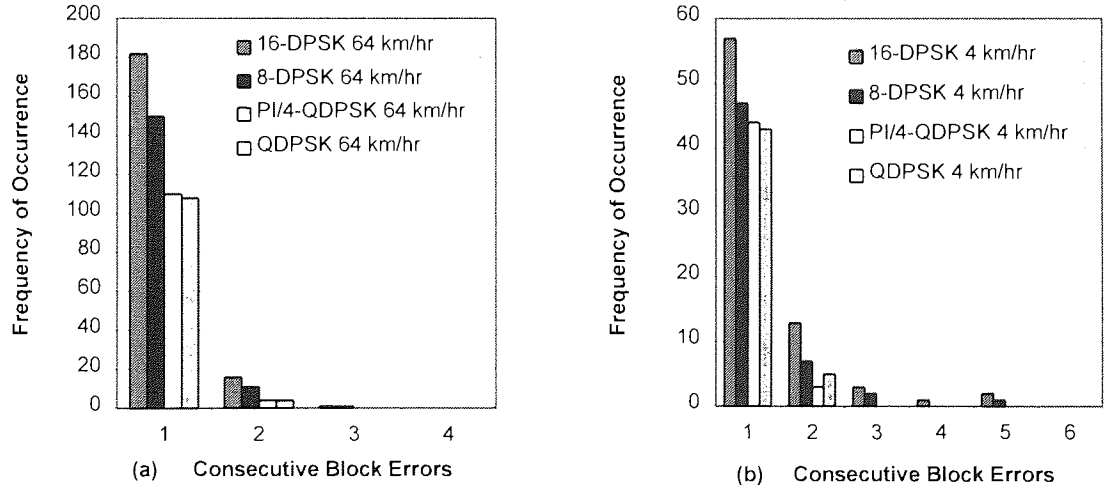


Figure 7-10 Consecutive Block Errors

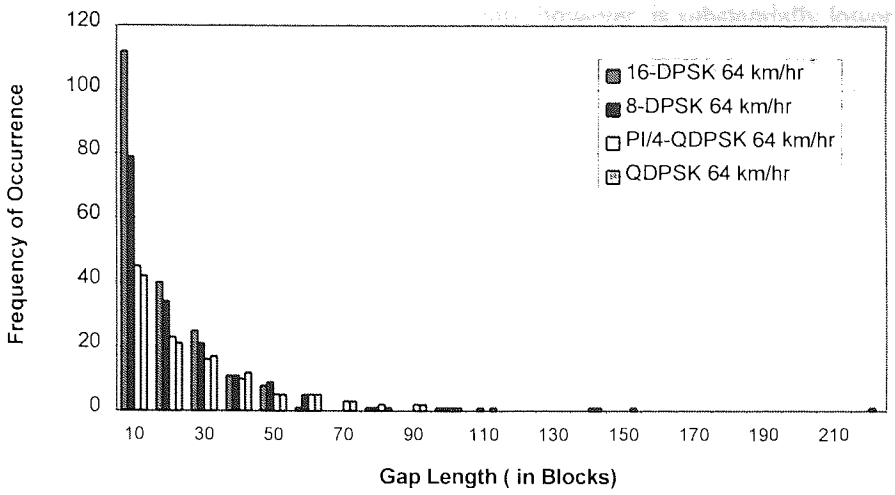


Figure 7-11 Errored Block Gap Distribution at 64 km/hr

The longest block bursts were measured for the 16-DPSK modulation at 4 km/hr. These bursts span five time frames. At a walking speed of 4 km/hr (1.11 m/s) one 10 ms time frame is equivalent to 0.011 m. The carrier wavelength at 1.9 GHz is equivalent to 1.58 m. Therefore the five block span is equivalent to 0.055 m which is nearly half the carrier wavelength. This is as expected since signal fading varies over small distances of the order of half a wavelength [Cox85] and hence fades rarely span more than half a wavelength.

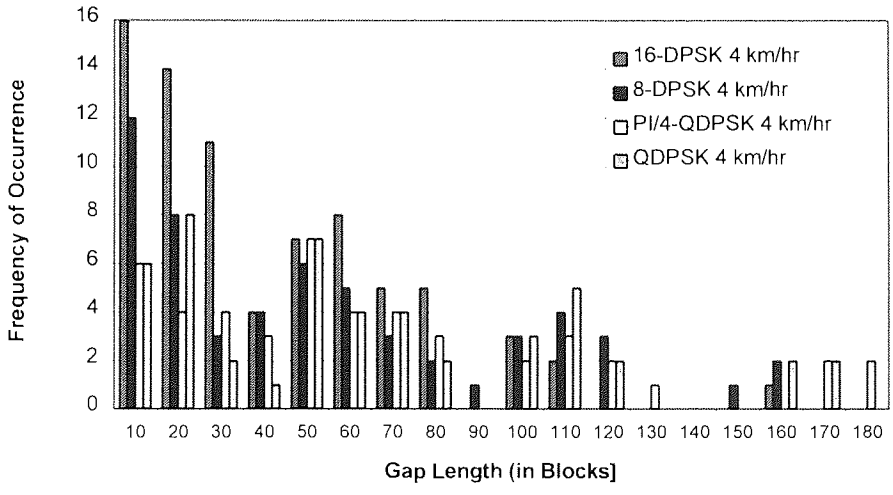


Figure 7-12 Errored Block Gap Distribution at 4 km/hr

The Errored Block Gap distributions are shown in Figure 7-11 and Figure 7-12 below. The distribution of the errored-block gap at 64 km/hr shown in Figure 7-11 can be well modelled by a negative exponential distribution. This suggests that the errored block arrival can be modelled as a Poisson stochastic process, *i.e.* errored-block events occur randomly. At 4 km/hr, the errored-block gap distribution depicted in Figure 7-12 is not well modelled by a negative exponential distribution. There are more long gaps between errored-blocks, indicating that the packet error arrivals, at lower velocity,

are bursty in nature. The number of simulated error events, however, is substantially lower at the lower speed so that the error margins are significantly higher than in the previous case. A longer simulation run is necessary to establish conclusively the nature of the burst gap distribution.

7.4 Video Transmission over Rayleigh fading channels

To study the transmission of H.261 coded streams over the wireless network described above, the colour QCIF Miss America sequence was first down-sampled to 10 fps and then coded using the IHC coder at a fixed bit rate of 32 kbit/s. Two sequences were generated. The first sequence was coded in normal mode with force updating implemented at the rate of one macroblock per frame. The second sequence was coded more robustly. The first frame was intracoded using the fixed-length DC coefficient mode proposed in section 5.6.2 and force updated at the rate of three frames per GOB per frame, thereby refreshing the frame with intracoded macroblocks every eleven video frames. Thus the second sequence has much better temporal error recovery capabilities, and it can recover from macroblock errors in just over one second.

The penalty incurred because of the forced updating is small as shown in Figure 7-13. The average PSNR of the two sequences is 39 dB for the normally updated sequence - QM32_11, and 38.2 dB for the force updated sequence QM32_23. The graph shows that the first frame in the QM32_23 is coded with a significantly lower PSNR, but the sequence recovers quickly. Thus, the lower resolution of this first robust frame is not very visible.

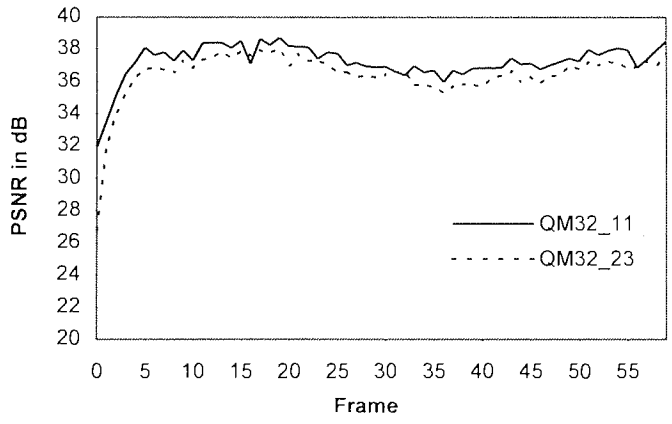


Figure 7-13 Forced Updating - PSNR Penalty

Velocity	π/4 QDPSK	16-DPSK
4 km/hr	Stream A	Stream C
64 km/hr	Stream B	Stream D

Table 7-2 Error Stream Summary

These two video streams were corrupted using four of the error streams simulated and analysed in section 7.3 above. The characteristics of the streams, labelled from A to D, are tabulated in Table 7-2.

7.5 Unprotected H.261 streams on the Rayleigh fading Channel

The QM32_11 video sequence was coded in normal H.261 mode and then corrupted using stream A. This stream contains $1.5x10^6$ bits whereas the coded video sequence contains $2x10^5$ bits. Thus stream A was split into five equal sections and each was used to corrupt the sequence. The video sequence was decoded five times using the IHC decoder without error correction or detection and the average MSE, the subjective usefulness rating (SUR), the number of bit errors, and the number of errored frames were noted. A summary of the results is tabulated in Table 7-3 below.

Stream	MSE	SUR	Errored Bits	Errored Frames
Aa	16.5	3+	142	5
Ab	240.0	3-	178	10
Ac	6.8	4	223	14
Ad	10.8	4	329	11
Ae	8.8	5-	107	8

Table 7-3 Summary of results - QM32_11 corrupted by Stream A

The results show that the five stream A segments (a to e) contain an unequal number of bits in error, from 107 for the lowest (Ae) to 329 bits for the highest (Ad). The average number of errored bits is 196 bits per segment yielding an average BER of $0.98x10^{-3}$.

The highest MSE successfully predicts the lowest SUR score, and the other MSE values are well correlated to the SUR scores as well. The average SUR score is 3.7 indicating that without error correction or detection, the stream can still be decoded usefully, and that the visible errors are not severe. This is in sharp contrast with the results obtained on the AWGN channel, where the same BER renders the unprotected sequences nearly unusable, with a SQR score close to 2. This arises because of the bursty nature of the errors on the Rayleigh fading channel. It is noted that on average, only 9.6 frames in 65 encountered bit errors on the Rayleigh fading channel. Of the five segments only one caused marginal sequence usefulness, and then only because the first sensitive frame was corrupted

On a random error channel with bit error probability P_e, the probability of zero bit errors in a N bit codeword, P_C, is given by:-

$$P_C=(1-P_e)^N \qquad (7\text{-}1)$$

Substituting $P_e=10^{-3}$ and N=3200 for the number of bits in a frame, the probability of zero errors within a frame is only 0.041, that is only 4 frames out of a hundred are error free on average.

The MSE per frame curves for the five sequences are plotted in Figure 7-14 (a) below. Clearly the Ab stream produces the worse MSE results and also the worse SUR scores. This stream was chosen

to test the robustness of the error control mechanisms in the robust IHC decoder as shown in the next section.

7.6 Decoding H.261 streams using FEC and error detection

The MSE per frame for the QM32_11 sequence decoded using FEC, and using Error Detection and Discarding are shown in Figure 7-14 (b). The Ab curve is also plotted as reference. The Error Detection results are obtained by decoding the stream using the BCH code to detect and discard errored blocks using implicit concealment. The FEC results are obtained by decoding the stream using the BCH(63,36,5) codewords in the data link layer. The 20 CRC in Figure 7-6 is used to detect when the BCH decoder commits a decoding error, in which case the 160 bit packet is discarded.

The stream decoded using the Error Detection option suffered from a partial GOB wipe out error in the first frame and although there was continuous improvement, the overall SUR score was 3-, barely sufficient. This is reflected in the MSE curve in Figure 7-14 (b) which shows that the first frame is severely corrupted but the recovery is constant albeit rather slow. In fact most frames have a worse MSE than without using error control at all, until the corrupted GOB is recovered, following which there is a substantial improvement using detection and discard. This happens because of the discarded data following error detection.

The SUR rating for the sequence decoded using the FEC was above 4. Four error bursts (with more than 64 bits each) defeated the correction capability of the low rate BCH code and led to some picture degradation. This result is significant because it shows that even the strong BCH code is defeated quite frequently on the Rayleigh fading channel and other countermeasures will have to be adopted to reduce the errored periods.

The results show that when a segment suffers severe errors, Error Detection and Discard does not achieve very good results if the first intracoded frame is corrupted in a normal (non-robust) H.261 stream. The decoder does recover steadily, however, and subsequent errors do not cause the same severe picture quality reduction. Therefore periodic intracoded frames should be avoided because these are more vulnerable to errors and generate three or more times the bit rate of intercoded frames.

When FEC is used, the sequence is decoded error-free most of the time. However error bursts do occur which defeat the powerful FEC code used in the data link layer. The FEC code can be improved by interleaving over a number of time slots. This is investigated in section 7-10. Another way to improve the picture quality is to improve the inherent error recovery of the video stream, *e.g.* by using macroblock level cyclic refresh using periodic forced updating. This last option is investigated in the next section.

7.7 Decoding H.261 with Forced Updating and fixed length DC frames

The same stream Ab was then used to test the resilience of QM32_23 stream using the Force Updating and the fixed length DC intraframe mechanisms, proposed in chapter 5, to improve the

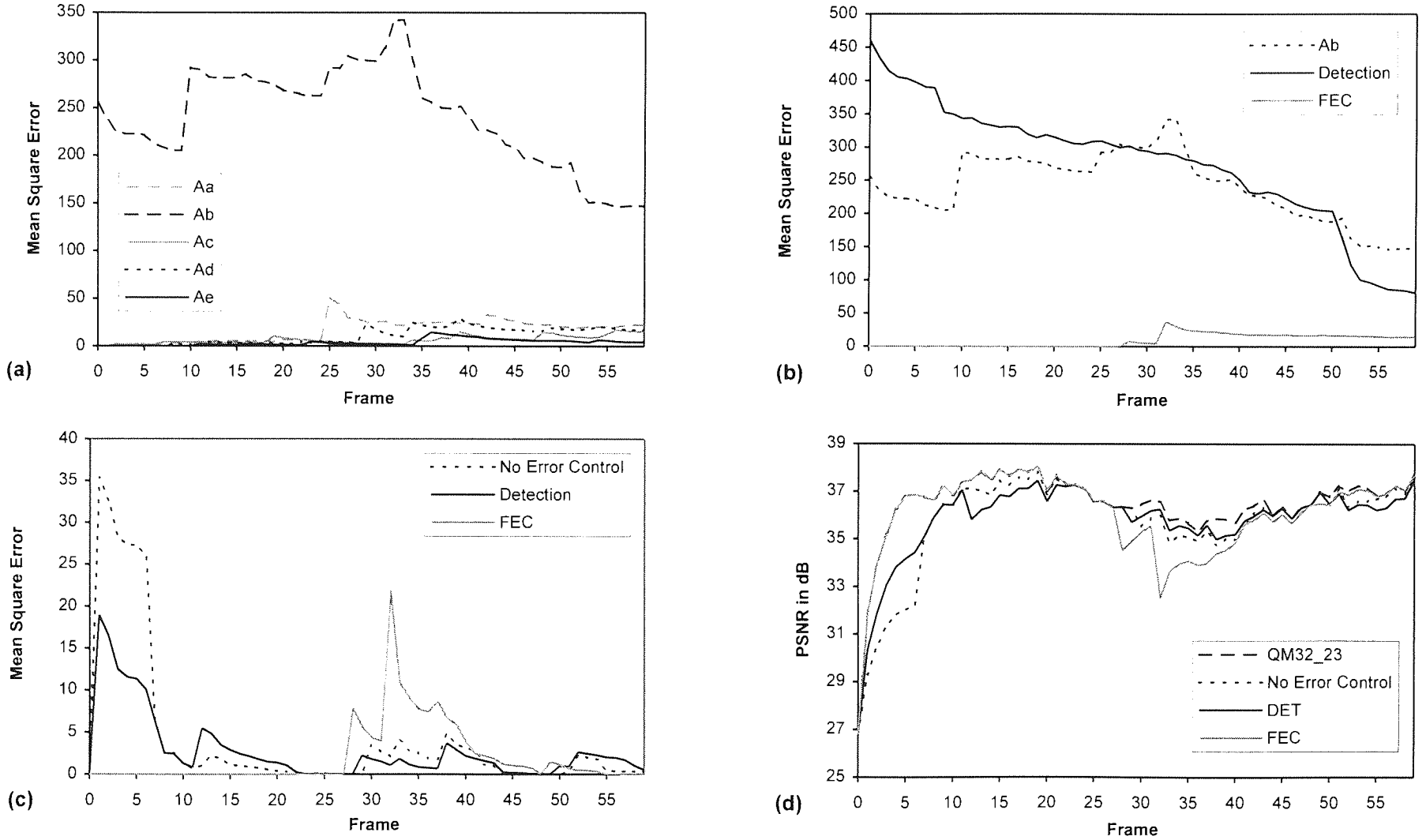


Figure 7.14 QM32 sequence in Rayleigh Fading - QDPSK modulation at 4 km/hr. (a) MSE for QM32_11 without Error Control (b) MSE for QM32_11 with Error Control (c) MSE for QM32_23 with Error Control (d) PSNR for QM32_23 with Error Control.

inherent error recovery capability of the H.261 stream.

When the stream was decoded without using error control, the start of the sequence was corrupted again, but this time the error was only visible for a short duration and the whole sequence was awarded an SUR score above four.

The reconstruction quality was significantly better when the stream was decoded with error detection and discard. One GOB suffered some macroblock errors in the first frames but there was very quick recovery. There were no visible errors for most of the sequence and the SUR score was above four.

When the corrupted sequence was decoded using the FEC decoder there were very few perceivable errors and the SUR score was close to 5.

The MSEs for the above three cases are plotted in Figure 7-14 (c). The results show that the second frame in the sequence decoded without error control suffered from a severe error. The error detection and discard policy contained prevented the appearance of several chequered macroblock errors evident in the previous case. The error burst in the second frame defeated the FEC code as well, but the errors were not very visible. The initial error was recovered by the eleventh frame, using the error control schemes. The error recovery took twice as long without error control.

It should be noted that the first frame was devoid of errors. This was due to the use of the fixed-length DC intraframe, which coded the first frame robustly, using less bits. Therefore there was less probability of an error hitting this frame. When errors do occur, however, there is a good probability that the errors hit the fixed length DC coefficient fields without the risk of losing video stream synchronisation.

An interesting result emerges from Figure 7-14 (c). Following the 26th frame the sequence decoded using the FEC suffered from three strong error bursts which defeated the FEC and led to a MSE which was higher than that of the other two decoded sequences.

The luminance PSNR for the uncorrupted QM32_23 sequence is plotted in Figure 7-14 (d) alongside the decoded PSNR using the FEC codes, error detection and discard and without using error control. The Results show that the three corrupted sequence can recover very rapidly following a severe error. The FEC decoded sequence enjoys the best PSNR, however this is obtained using twice the bit rate of the other two decoding schemes.

The above results indicate that it is possible to transmit QCIF H.261 coded video at the rate of 32 kbit/s over a slowly Rayleigh fading channel, and recover a video sequence which is still useful at its worse.

By using the proposed Forced Updating and DC-intracoded frame mechanisms, the video can be recovered with very little visible degradation using an error detection and discard policy, with picture quality nearly as good as that recovered using a powerful half rate FEC code.

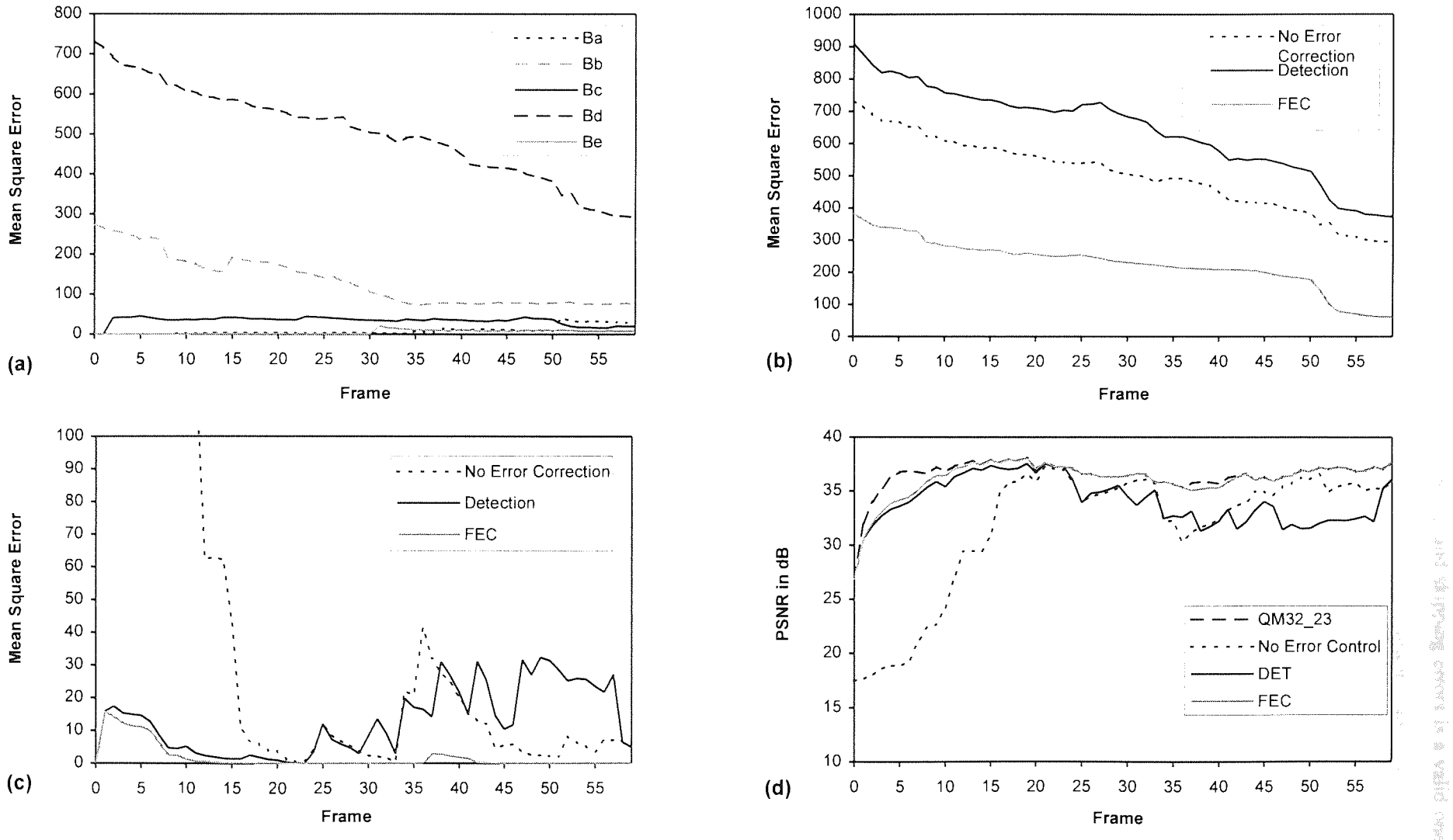


Figure 7.15 QM32 sequence in Rayleigh Fading - QDPSK Modulation 64 km/hr. (a) MSE for QM32_11 without Error Control, (b) MSE for QM32_11 with Error Control, (c) MSE for QM32_23 with error control, (d) PSNR for QM32_23 with error control.

These results indicate that the policy of detecting and skipping errors is a valid one on slow Rayleigh fading channels, because the errors occur in bursts, hit fewer frames and allow forced updating to recover from the visible errors that occur. (Similar results have been obtained using other sequences. Space and time limitations prevent their inclusion in this thesis. These results will be published elsewhere in due course.)

FEC achieves the best results, and can guarantee unimpaired picture quality most of the time. Error bursts are frequently intense, however, and defeat the low rate FEC codes used, although interleaving over a number of time slots should reduce the residual error rate.

The results also confirm the importance of the first intracoded frame and the success of the DC-intracoded frame mechanism in mitigating bit error effects in this sensitive frame.

7.8 The effect of increasing the mobile speed

The above tests were repeated at a higher simulated mobile speed - 64 km/hr, using Stream B. The results are shown in Figure 7-15 (a) to (d). The results for the QM32_11 stream decoded without error control are summarised in Table 7-4.

Stream	MSE	SUR	Errored Bits	Errored Frames
Ba	8.9	4	126	13
Bb	133.0	3-	166	13
Bc	3.8	3	271	18
Bd	502.0	2+	259	16
Be	5.3	5-	109	9

Table 7-4 Summary of Results - QM32_11 corrupted by Stream B

The average MSE and SUR scores for the five decoded B streams are 130.6 and 3.3 respectively. In comparison, the Stream A scores were 56.7 and 3.7 respectively. These results show clearly that the higher mobile speed results in a decreased picture quality. This can be explained by referring to Figure 7-8 and Figure 7-9. These show that the packet error rate increases at the higher speed and hence more frames are hit by errors, leading to an increase in MSE. This is confirmed by the average number of frames in error which is 9.6 for the stream A and 13.8 for stream B.

The worse decoded sequence was obtained using error stream Ad as shown in Figure 7-15 (a). This stream was then used to corrupt the QM32_23 sequence again,. The corrupted sequence was decoded without using error control, using error detection and discard, and using the FEC code. The results in Figure 7-14 (b) and Figure 7-15 (b) indicate that the decoded MSE is worse at 64 km/hr than at 4 km/hr. The average MSEs for streams decoded with detection and FEC are 642 and 226 at 64 km/hr, and 274 and 9.4 at 4 km/hr. The SUR scores are also higher; 3.5 versus 3 at the higher speed.

The high MSE experienced by the FEC decoded sequence shows that the FEC can be defeated at high speeds and leads to very visible errors (SUR=3.5). It was expected, however, that the average MSE should decrease for FEC decoding as the mobile speed increases. This should happen because, as

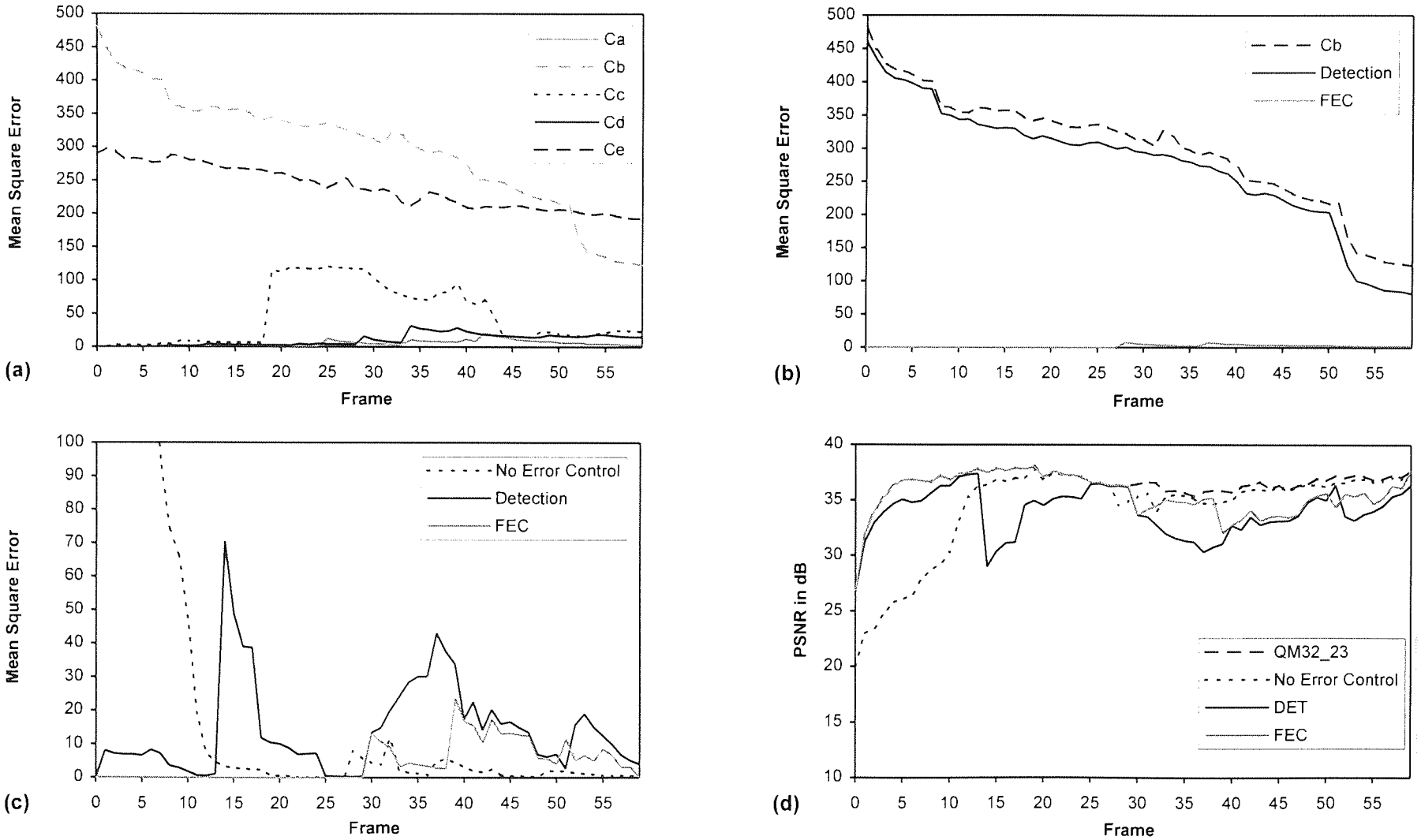


Figure 7-16 QM32 sequence in Rayleigh Fading - 16-DPSK modulation at 4 km/hr. (a) MSE for QM32_11 without Error Control (b) MSE for QM32_11 with Error Control (c) MSE for QM32_23 with Error Control (d) PSNR for QM32_23 with Error Control.

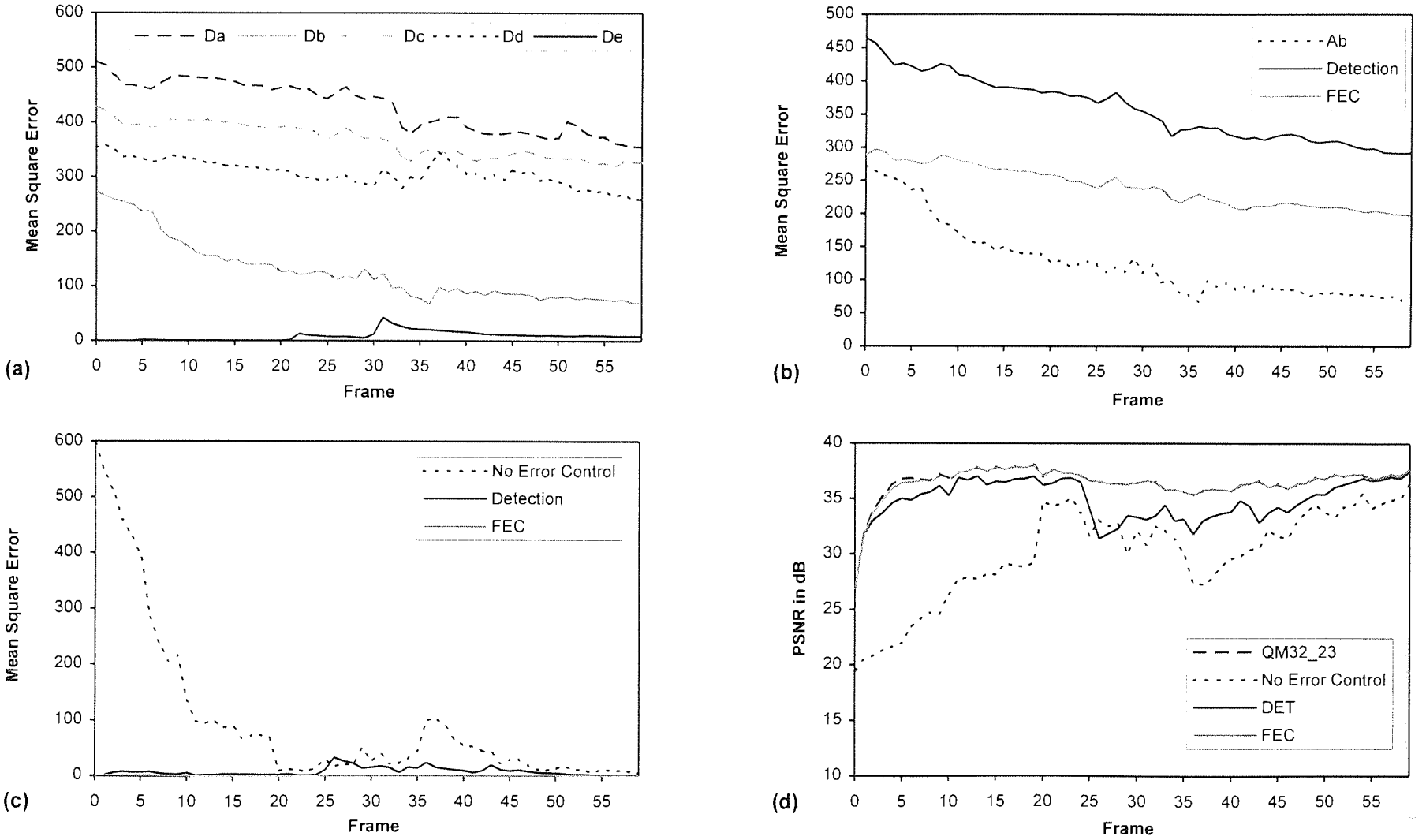


Figure 7.17 QM32 sequence in Rayleigh Fading - 16-DPSK Modulation at 64 km/hr (a) MSE for QM32_11 without Error Control (b) MSE for QM32_11 with Error Control (c) MSE for QM32_23 with Error Control (d) PSNR for QM32_23 with Error Control.

is evident from Figure 7-8 and Figure 7-9 the BER per block decreases at the higher speed so that the FEC code is more likely to decode errored packets correctly. This anomalous result may be due to the fact that the MSE has not been averaged over a long enough sequence to yield the correct result. It is also noted that the error is again due to corruption of the first intracoded frame.

The robust QM32_23 sequence, was then decoded under the same conditions. The results in Figure 7-15 (c) and (d) are clearly worse than the 4 km/hr results in Figure 7-14 (c) and (d) when decoding the sequence without error control and with error detection and discard. The MSE for the FEC decoded sequence is lower at the higher speed now, as anticipated above.

The results show that the sequences can still be decoded successfully at the higher vehicle speed of 64 km/hr. The robust sequences decoded using FEC can yield better results at 64 km/hr than at 4 km/hr as anticipated. The sequence decoded using error detection and discard performed worse at the higher speed because more frames were affected. However the decoded stream was still useful (SUR=4).

The decoder using detection and discard in conjunction with the DC intracoded frame achieves better results than a normal H.261 decoder, when decoding the robust H.261 stream. This emerges from all the (c) curves in Figure 7-14 to Figure 7-17.

7.9 The effect of increasing the modulation level

The effect of increasing the modulation level was studied by repeating the above series of tests on the same QCIF sequences using streams C and D for 16-DPSK at 4 km/hr and 64 km/hr. The two error streams were again split into five sub-streams (a)-(e) and used to corrupt the QM32_11 sequence. The results are shown in Figure 7-16 (a)-(b) and Figure 7-17 (a)-(b) and tabulated in Table 7-5 and Table 7-6.

Stream	MSE	SUR	Errored Bits	Errored Frames
Ca	5.1	4+	156	11
Cb	298.0	2+	194	10
Cc	46.5	3+	245	20
Cd	10.3	4+	298	16
Ce	239.6	3	137	14

Table 7-5 Summary of Results - QM32_11 corrupted by Stream C.

Stream	MSE	SUR	Errored Bits	Errored Frames
Da	429	2+	185	19
Db	127	3	140	16
Dc	367	3-	314	19
Dd	307	2+	248	21
De	9	5-	87	11

Table 7-6 Summary of Results - QM32_11 corrupted by Stream D.

A comparison of the results in Table 7-3 to Table 7-6 shows that the unprotected streams perform worse at the higher modulation rates as expected. The average MSE for QDPSK at 4 km/hr is 69.55 which increases to 150 for 16-DPSK at the same speed. The average SUR for QDPSK at this speed, is marginally higher than that for 16-DPSK - 3.7 against 3.6. The same degradation in picture quality is also evident at the higher vehicle speed. The MSEs for QDPSK and 16-DPSK at 64 km/hr are 171 and 309 respectively, whereas the relative SUR scores are 3.3 and 3.

The two worse subjective results were obtained using streams Cb and Dd. These were used to test the robust decoder as explained above using the QM32_23 video sequence. The results for these streams are summarised in Table 7-7 below.

Stream	Detection MSE	Detection SUR	FEC MSE	FEC SUR
QDPSK 4 km/hr Ab	2.74	4+	1.99	5
QDPSK 64 km/hr Bd	12.8	4	1.68	5
16-DPSK 4 km/hr Cb	14.0	4	4.22	4+
16-DPSK 64 km/hr Dd	7.63	4	0.16	5

Table 7-7 Summary of Results - QM32_23 sequence

The objective and subjective results in Table 7-7 indicate that the picture quality deteriorated when the modulation levels were increased from 4 (QDPSK) to 16 (16-DPSK) for both error control techniques (Detection and FEC) at 4 km/hr. On the other hand the picture quality improved when the modulation levels were increased for the same sequence decoded using the same error control techniques but at the higher speed of 64 km/hr. These results are inconclusive and indicate that more results are needed to generalise the results to other segments in the same video sequence, and thence to other sequences.

7.10 Interleaved FEC

The classic approach to combat burst errors is to employ interleaving to randomise the burst errors and then employ a FEC code which is designed to combat the average BER after interleaving [Lin83]. Without interleaving, the FEC code would have to cater for the peak BER endured during a burst, which is greater than the average BER, and would thus have to be more complex and less efficient.

A simple interleaving scheme involving an interleave buffer is depicted in Figure 7-18. The k, n bit codewords to be transmitted are stacked in the k rows of the rectangular transmit buffer and then shifted out column by column. At the receiver the columns are stored side by side in the de-interleaving

buffer and then read out in serial row fashion. The number of codewords stacked **k** is called the interleaving depth [Lin83].

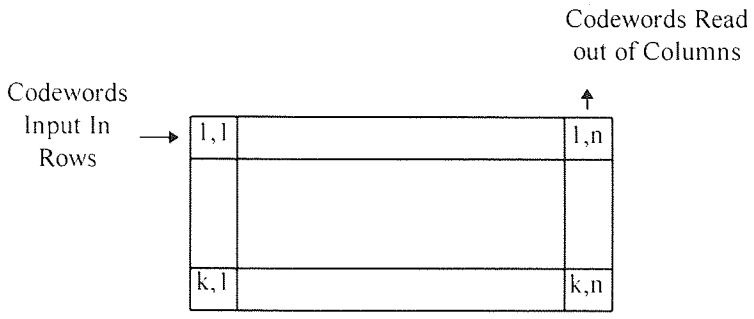


Figure 7-18 Interleave Buffer

Interleaving is used frequently to combat burst errors on radio channels [Webb90, Bergman90]. A computer program has been written to investigate the use of interleaving with the FEC packet format described above. The program reads in the error files generated by the modem simulator, simulates the interleave buffer and the t error correcting block code, and outputs the resulting bit stream and error correction/failure statistics. The output stream can then be used to corrupt the video streams.

The program was used to investigate the interleave depth required to achieve zero un-corrupted packets at the output of the decoder. The de-interleave buffer size was set to 630 bits to accept two 315 bit FEC packets side by side, since two time slots are needed per MAC frame to transmit 32 kbit/s. The channel error files were then read into the de-interleaver, de-interleaved and decoded using bounded distance decoding. The number of packets still in error was noted whilst increasing the interleave depth k. When k was set to five, none of the output packets were in error. Thus by using this interleaving depth in conjunction with the FEC codes simulated, error free transmission of the video sequence can be achieved *using the generated error sequences*.

7.11 Summary of Conclusions

This chapter uses the simulation infrastructure developed in the previous chapters to study the transmission of video over narrowband Rayleigh fading channels. A wireless network is proposed similar to several cordless networks introduced in chapters 2 and 3. The wireless network link is then simulated at physical, MAC and data link layers to study the transmission of unprotected and robust H.261 coded video streams at a BER of 10^{-3}, which is typical for voice oriented wireless networks.

The simulated network link uses TDM on the down-link, circuit switched TDMA on the up-link and staggered FDD. The frame rate is 10 ms, and each frame contains 8 time-slots. Each time slot consists of 240 symbols, which are equivalent to 480 bits with QDPSK and 860 bits with 16-DPSK. Alternatively the 16-DPSK modem can support 16 time-slots of 480 bits each.

Two data link packets have been proposed; one to be used with error detection oriented, error control techniques and the other to be used with BCH(63,36,5) FEC codes. The basic detection packet format accepts a data payload of 320 bits and uses the rest for CRC and the header field. The FEC packet format packs five BCH(63,36,5) codewords into a 315 bit payload field. The 180 information bits consist of 160 video bits and a 20 bit CRC. The codewords in one time slot can be interleaved to disperse an error burst over all five codewords to make best use of their random error correction capability. Note that the above packet structures are supported by a transcoder which uses the raw bearer channels of the MAC layer.

The carrier frequency is 1.9 GHz and rotation invariant, differential M-ary Phase Shift Keying modems have been simulated to transmit video information over the radio link at two simulated vehicle speeds - 4 km/hr and 64 km/hr. The transmission baud rate is 192 kbaud, and it is assumed that the propagation delay spread is low enough for the channel to be modelled as Rayleigh fading. Furthermore CCI is ignored, so that the results obtained are for a non-CCI environment as in [Stedman92, Hanzo94, Streit94].

The link was simulated and packet error statistics were collected and analysed. The results show that for the above TDM MAC format, the packet errors are more bursty on the 4 km/hr channel than on the 64 km/hr channel as expected. The average number of consecutive packets in error is low for both speeds. The lowest is 1.035 for QDPSK at 64 km/hr and the highest 1.395 for 16-DPSK at 4 km/hr. Multiple packet bursts are common, however, with 5-packet error bursts detected at 4 km/hr for 16-DPSK. The packet error bursts have been found to occur more or less randomly at the higher speed, but there is evidence of a non-Poisson arrival distribution at the lower speeds.

The transmission of video over the radio link was simulated using a 'normal' H.261 video sequence with intrinsic force updating at one macroblock per frame, and a 'robust' sequence with force updating increased to three macroblocks per GOB and using the fixed length DC intracoded frame described in chapter 5. The QCIF, colour Miss America sequence was used to generate both sequences at a video frame rate of 10 fps and a coded rate of 32 kbit/s.

These two sequences were then corrupted using four error files generated by simulating QDPSK and 16-DPSK modems at 4 and 64 km/hr. The corrupted sequences were then decoded using the robust IHC decoder - without using error control, using error detection and discard, and using a modified robust decoder based on the IHC decoder but which can use the powerful FEC codewords inserted by the transcoder.

The first important conclusion was that the sequences can be decoded *usefully* at a bit error rate of 10^{-3} using the robust IHC decoder without the use of error control. In this context useful means that although the images may be impaired, the image context is conserved and the picture quality does not deteriorate catastrophically in an MSE sense. The average SUR scores for the unprotected sequences were 3.7 and 3.3 for QDPSK at 4 km/hr and 64 km/hr, and 3.6 and 3 for 16-DPSK at 4

km/hr and 64 km/hr. The respective mean square errors were 70, 171, 150 and 309. Therefore the MSE scores predict the SUR in the correct order. *Note that both objective and subjective results are averaged over the whole error correction stream (30 seconds long).*

The above results are in sharp contrast with the results obtained on the random error channel, were the BER of 10^{-3} caused the unprotected streams to fail catastrophically. This can be explained by noting that the errors on the radio link occur in bursts such that substantially fewer frames are hit by errors than on the random error channel at the same BER.

When the robust streams were decoded using the robust IHC decoder with error detection and discarding, the SUR scores rose to 4.5 and 4 for QDPSK at 4 km/hr and 64 km/hr, and to 4.2 and 4 for 16-DPSK at the same speeds. The corresponding average mean square errors were 2.74, 12.8, 14 and 7.63. *These results are for the worst effected 6 second segment in the stream.* It was noted that the DC intracoded frame was very effective in avoiding the highly visible and objectionable errors in the first frame.

When the robust streams were decoded using the low rate FEC option, the SUR scores rose to 5 for all decoded streams except the 16-DPSK 4 km/hr stream which scored 4.5. The corresponding mean square errors were 1.99, 1.68, 4.22 and 0.16 *for the worst effected 6 second segment.* Thus the streams were decoded practically without any visible errors. It must be pointed out, however, that errors still defeated the low rate FEC codes, with 4, 2, 4 and 1 frames hit by error bursts in error streams Ab, Bd, Cb and Dd respectively.

It has been shown that the reconstruction quality deteriorates with increasing vehicle speed, unless FEC is used, in which case the residual BER decreases with increase in speed because the error bursts are less intense. This result holds for both QDPSK and 16-DPSK.

Interleaving can be incorporated in the simulation framework by pre-processing the error files prior to corrupting the video sequences. The interleaved and corrected error files have been analysed to predict the interleave depth needed to guarantee no residual errors using the BCH(63,56,5) codewords. The minimum interleave depth required was found to be five time-frames. This is equivalent to a 50 ms delay either side of the link and the total delay of 100 ms is not considered excessive. The interleave delay, however, may have to be increased at lower speeds. Alternatively, other techniques such as frequency hopping can be used to break up the error bursts at low speed [Webb90].

When used in conjunction with the FEC protected packets, the radio link is essentially error-free for long durations and can deliver very good quality video. However this is achieved by using half rate codes and the interleave buffers.

Thus it has been shown that H.261 coded video can be transmitted over the proposed network at an average BER of 10^{-3} by using the robust decoders developed in chapter 5 and modified slightly to use more powerful, low rate BCH FEC codes. It is possible to trade bandwidth efficiency and video quality on the simulated narrowband Rayleigh fading channel as follows:-

- An unprotected H.261 stream without error control overhead can be transmitted and recovered with useful reconstruction albeit with visible impairments every few tens of seconds.
- A robust video stream incorporating forced updating and the fixed-length DC intraframe can be transmitted and recovered using the robust decoders with significantly less noticeable picture impairments occurring at the same frequency. The error control overhead is less than 7% because only high-rate error detection codes are used.
- The same robust video stream can be transmitted with hardly any visible impairments over a 30 second period, by using the low rate FEC BCH(63,36,5) code with 50% error control overhead.
- If ten consecutive FEC coded time slots are interleaved over five MAC frames (two time slots per frame), the video streams are decoded without any residual error. The coding efficiency is unchanged but an additional delay, equivalent to ten frame periods, is introduced in the simplex transmission chain.

Chapter 8

Error Control by Retransmission for Video Transmission over Wireless Networks

In the previous chapter, it was shown that the errors on a radio link are bursty in nature. When errors occur the severity of the bursts is such that even low rate FEC codes fail to correct a significant number of errors. When the classical technique of interleaving is applied to randomise the errors, the interleaving depth is limited by the delay sensitivity of the video data. If the interleave depth is small, interleaving may aggravate matters by spreading errors to otherwise uncorrupted blocks. At the same time, the error rate is relatively low between error bursts, and the low rate FEC codes do not use bandwidth efficiently.

In this chapter, feedback techniques are investigated as a class of error control techniques which can inherently adapt to the error conditions on the channel and hence achieve the high reliability required by video streams in a bandwidth-efficient manner. Feedback techniques conserve bandwidth by employing high throughput, high reliability error detection codes during the 'good' phase of the channel, and deploy low throughput corrective measures when the channel performance deteriorates.

8.1 Introduction

Automatic Repeat Query (ARQ) techniques are a class of efficient and highly reliable error control techniques for use on full duplex circuits where a feedback path is available. ARQ is widely used for computer and data communications, where high reliability is imperative and some delay penalty can be tolerated given the non real-time nature of the transacted data [Tannenbaum89]. However ARQ has never been popular in real time applications such as voice transmission, where the associated delay, especially on links with long round trip delay, has been considered detrimental to the intelligibility of speech. A further drawback in the application of retransmission to voice are the low bit rates associated with voice. Then a relatively long time is necessary to collate a sufficient number of speech bits into a packet such that the additional overhead bits required (e.g. synchronisation, header, control and parity fields) do not drastically reduce the transmission efficiency.

The first limitation does not exist in the case of microcellular and picocellular transmission. The round trip delays involved are minimal provided the medium access control (MAC) is designed with care as shown below. The second limiting factor is also absent when one considers low and medium rate digitally coded video, where the data for one video frame occupies a relatively high number of packets.

Systems employing feedback to *refresh* corrupted frames have been proposed to limit temporal error propagation in video transmission [Wada89] but these do not employ *retransmission*. Recently McDonald *et al.* at British Telecom Research Laboratories (BTRL) proposed an ARQ error control scheme for the transmission of H.261 coded video over a DECT link [MacDonald92]. One of their main conclusions was that ARQ was preferable to interleaved FEC on the DECT channel, since the error bursts are severe, and defeat the interleaved FEC, resulting in errors spreading to otherwise uncorrupted blocks. However the issue of potentially unbounded delay was left un-tackled.

This chapter investigates the applicability and performance of link level ARQ techniques for digital video transport. Basic and hybrid ARQ techniques are introduced first and then their application to mobile radio links is reviewed from the literature. The BTRL ARQ system is analysed and several shortcomings are identified. The main problem is identified as the retransmission buffer control on the down link. A number of solutions are proposed to solve this problem based on flow control and bandwidth-on-demand techniques. Two flow control techniques are investigated next. The investigation leads to a proposal of a simple transcoder based system for transmitting FBR H.261 video over a radio link. The scheme combines a simple Stop-And-Wait ARQ protocol with a flow control technique and simulation results are presented confirming the validity of the technique on Rayleigh fading channels with and without CCI.

A second, more complex, ARQ scheme is proposed which combines a Selective Retransmission ARQ scheme with a bandwidth-on-demand technique for the retransmission channel. Simulation results are presented for this scheme operating on a third-generation wireless network radio link perturbed by ISI.

A novel decoder, called the Trace-Back Decoder (TBD), is then proposed. The TBD combines selective retransmission ARQ with error concealment to achieve a low-delay, fast-recovery error control technique which can deliver very good quality video independent of the video source coding algorithm.

A summary of the results is presented in the final section.

8.2 Basic ARQ techniques

The basic elements of any retransmission scheme are shown in Figure 8-1. The most obvious requirement is a feedback channel. On full duplex links, such as considered for conversational video communications, the feedback channel can be implemented easily, for example by utilising a small part of the reverse channel bandwidth.

In an ARQ system, a channel code with good error detection properties is applied to the transmitted data blocks to enable the receiver to detect erroneous codewords. The receiver notifies the transmitter of correctly received codewords via the feedback channel. When an error is detected the

receiver instructs the transmitter to re-transmit the codeword, by appropriate signalling on the feedback channel. Retransmission usually continues until the codeword is received correctly.

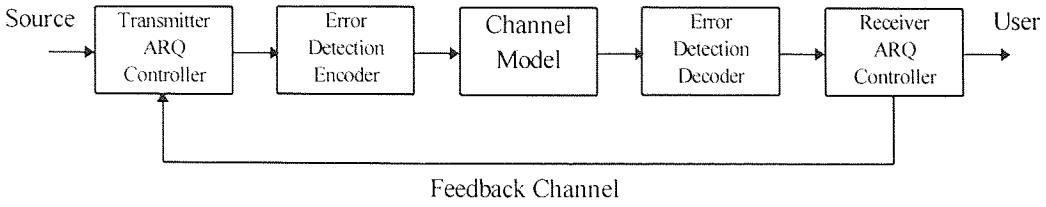


Figure 8-1 Block diagram of a basic ARQ error control system

The throughput efficiency of an ARQ system is defined as the ratio of the average number of information bits per second accepted by the receiver to the maximum data transmission rate on the channel [Lin83]. It is a key figure of merit of ARQ techniques. The throughput efficiency depends on block size, bit error rate, round trip delay and the retransmission protocols operating in the transmitter and receiver ARQ controllers [Lin83].

The smaller the size of the packet the lower the probability of corruption and hence the smaller the retransmission overhead. However the combined header, error detection code and reverse signalling channel overhead is proportionately larger for smaller packet sizes. Thus an optimal packet size can be determined which maximises the channel utilisation.

The parity field is quite shorter than for FEC codes, since the objective is to detect the presence of errors. It has been shown that the probability of undetected error P_u is not strongly dependent on the message block size and is proportional to 2^{-n} where n is the number of parity bits [Lin83]. Hence even for small n the probability of undetected error is very small. For example with n equals to 16, 24, and 32 the probability of undetected error is 1.53×10^{-5}, 5.96×10^{-8} and 2.33×10^{-10} respectively. Thus a modest number of parity bits can guarantee the very low residual BER necessary for video.

There are three basic ARQ protocols, Stop-And-Wait (SAW), Go-Back-N (GBN) and Selective Retransmission. (SR). These basic techniques can combined with Forward Error Correction techniques to improve the throughput efficiency. The key concepts behind the techniques are introduced below. A more detailed introduction can be found in [Lin83] and [Lin84] and in the references therein.

8.2.1 Stop-And-Wait ARQ

In the Stop and Wait (SAW) protocol the transmitter waits for the acknowledgement following every packet transmission. On circuits with long round trip delay, the channel utilisation can be very low [Lin83]. The protocol, however, always delivers the packets in the correct order. If an acknowledgement is lost the transmitter times out and retransmits the packet assuming it has not been received correctly. The receiver must discard duplicate packets. This scheme is very simple to

implement and also efficient since only two packet identifiers are necessary to resolve duplicate packets [Tannenbaum89].

8.2.2 Go-Back-N ARQ

When the transmission delay is large, the throughput of the basic SAW technique is low and the Go-Back-N (GBN) protocol achieves a higher throughput [Lin83]. In GBN the transmitter transmits packets continuously without waiting for acknowledgements for the transmitted packets. The receiver sends back an acknowledgement (ACK) for every correctly received packet. The transmitter thus receives an acknowledgement for a packet one round trip delay later, during which time (N-1) more packets have also been transmitted. If a packet is received in error, the receiver discards it and the following (N-1) packets and feeds back a negative acknowledgement (NAK). When the transmitter receives a NAK it traces back to the indicated packet and retransmits the N consecutive packets. The transmitter must store all packets awaiting acknowledgement but the receiver stores only one packet at a time. The protocol delivers the packets in the correct sequence barring any undetectable errors.

8.2.3 Selective Retransmission ARQ

If the transmission rate and the round trip delay are high, the number of packets N discarded in the GBN protocol becomes high and the throughput efficiency drops. In Selective Retransmission (SR) the transmitter transmits packets continuously as in GBN. When the receiver detects an erroneous packet it sends a NAK, otherwise it acknowledges the packet with an ACK. When the transmitter receives a valid ACK it removes the packet from the retransmission buffer. If a NAK is received or a packet time out is reached (in the event of corrupted ACK messages), the packet is retransmitted.

The receiver must buffer correctly received packets which are preceded by errored packets awaiting retransmission, to deliver the packets in the correct order. Practical SR ARQ systems use a finite receiver buffer and achieve a lower throughput than theoretical systems using an infinite buffer size.

GBN and SR can be combined to deliver a throughput which is between that of GBN and SR with infinite buffer size [Lin83].

8.2.4 Hybrid ARQ

When the probability of bit error is high, the probability of packet failure increases and the throughput decreases [Lin83]. It is then advantageous to deploy Hybrid techniques which combine any of the above ARQ techniques with FEC as shown in Figure 8-2. The FEC code is used to correct the majority of low weight error patterns and the error detection code is used to detect FEC code failure. This combination of techniques results in improved throughput.

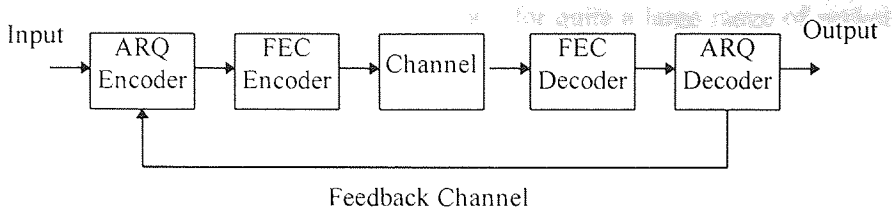


Figure 8-2 Generic Hybrid ARQ/FEC Technique.

There are two types of hybrid ARQ techniques called Type-I and Type-II. In Type-I hybrid techniques each re-transmitted packet is decoded independently. In Type-II techniques the first retransmission is an invertible code [Lin83] which, if received correctly, will deliver the packet. If it is received in error, it can be used with the previously transmitted packet to form a strong low rate code which has a high probability of decoding the packet correctly. The transmitter alternates between the two kinds of packets during repeated retransmissions.

The Type-I Hybrid throughput is lower than the ARQ technique alone for low probabilities, but is substantially better at medium and high probability of error [Lin83]. Type-II Hybrid schemes have the same throughput efficiency of the independent ARQ techniques at low error probabilities and perform better than Type-I hybrid techniques at high error rates, at the cost of increased complexity. At medium probability of errors Type-I Hybrid techniques may outperform Type-II techniques [Lin83].

8.3 ARQ techniques for data transmission over fading radio channels.

ARQ techniques were first applied to telegraph networks in the early 1940s. A large number of the telegraph circuits used HF radio links which fade notoriously and the two character RQ (for Repeat Query) sequence was introduced to request a retransmission of corrupted messages.

In 1970 Abramson introduced the ALHOA system, which transmitted computer data over a radio network and featured ARQ as a prominent Error Control Protocol [Tannenbaum89]. A number of studies on ARQ for mobile data transmission were published in the late seventies, early eighties and are referenced in [Comroe84].

In 1984 Comroe and Costello [Comroe84] explored the problem of reliable data transmission over land mobile channels. They characterised the channel as a Rayleigh fading channel and studied throughput and reliability of various ARQ techniques at different mobile speeds and with different packet sizes.

Siew and Goodman [Siew89] used a mathematical model of a Rayleigh fading channel to develop formulas for data link throughput and delay. The formulas can be used to yield numerical results for these two key performance parameters. They found that throughput efficiency increases with transmission rate, as the mobile speed decreases and as the fade margin increases. The peak throughput efficiency for a 256 kbit/s system, at 900 MHz and a vehicle speed of 20 Km/hr is 0.835 and the optimum packet size is 95 bytes (from Table II in [Siew89]). Figure 7 in [Siew89] reveals that at this

vehicle speed the throughput efficiency is quite constant for quite a large range of packet sizes. Also the optimum packet size is strongly dependent on the vehicle speed.

Recently, Chuang [Chaung90] and Chang [Chang91] studied the throughput of GBN and SR ARQ on a Rayleigh fading channel. Chuang used simulation techniques to study the two protocols in Rayleigh fading channels at two vehicle speeds, using two branch diversity. The throughput efficiency of SR was found to be higher than GBN for both speeds. With two branch diversity, the throughput efficiency reached 0.9 at 0.5 mi/h. The carrier frequency was 1800 MHz, using coherently detected QPSK, with an average E_b/N_0 of 10dB. The 161 bit packets were transmitted in one time slot on a TDMA system with 40 time slots per frame of 16 ms duration. Chang's paper develops a semi-analytical model based on fade and interfade statistics, derived from the Jakes Rayleigh channel fading simulator, which avoids the analytic difficulties in deriving results based on realistic system assumptions. Chang derives estimations which compare favourably with Chuang's simulation results.

The DECT system can support high bit rate applications by aggregating time slots in symmetrical or asymmetrical fashion. The throughput performance of the DECT multibearer system using various ARQ algorithms has been studied by Wong *et al.* [Wong94]. The throughput for symmetrical multibearer circuits is above 0.85 for up to 8 aggregated time slots using SR-ARQ, in 0.1 normalised delay spread. When the channel deteriorates, Antenna diversity can be deployed to keep the throughput at high levels.

The main implications of this brief review are that ARQ techniques can achieve very good throughputs at the transmission rates contemplated in the system described in chapter 7. What remains to be seen is whether these throughputs are realisable at delays which are compatible with real time video transmission.

8.4 Feedback techniques for video transmission.

As already mentioned the use of ARQ for real-time services including video, has not been considered very favourably in the literature. Several authors have considered using feedback information to recover from transmission errors, however the feedback information is not used to retransmit the data but to refresh the errored portion of the image [Wada89, Horst93].

A simple technique in this class, which can be used with an unmodified H.261 coder, uses the BCH (511,493,2) code to detect errors in the current frame. When an error is detected the previous frame is frozen and an update request is issued via the fast update mechanism originally intended for multipoint (multiparty) operation. When the encoder receives the update request, it codes the next frame in intramode and issues the picture release signal in the picture header. When the receiver receives the picture release signal it resumes decoding and displaying the video data.

Wada described an elegant though elaborate system which uses feedback information to recover from packet errors on ATM networks [Wada89]. The feedback information is used to

determine the area of the image that needs to be replenished in intracoded mode. Only the corrupted data is refreshed thereby reducing the retransmission overhead.

Recently MacDonald *et al.* at British Telecom Research Labs. (BTRL) published a seminal paper, proposing a link level ARQ system for transmitting H261 video over DECT [MacDonald92]. The proposed scheme is discussed in section 8.5.

8.4.1 ARQ for Video Transmission over Mobile Radio links.

The main objection to the use of ARQ techniques for real time services is the round trip delay. On cellular mobile radio networks, particularly microcellular and picocellular networks the round trip delay will be shown to be low and hence the right conditions exist for the viable implementation of a real time ARQ scheme for video.

The round trip delay on wireless networks is due to the radio propagation delay and the MAC protocol delay. In these systems the radio cells are small and the radio path round trip propagation delay is very low (less than 10us typically). The biggest component of the round trip delay is then due to the MAC protocol in force.

Consider the DECT MAC structure shown in Figure 8-3 which supports twenty four time slots using a time division multiple access (TDMA) protocol [Goodman91a]. Twelve full duplex, 32 kbit/s voice channels are provided for voice or data communications. Time division duplex (TDD) is employed with the first twelve time slots in the frame reserved for base station to mobile transmission, and the last twelve time slots catering for the reverse direction. The forward and reverse time slots are paired such that TS 1 is used for one direction and TS 13 is used for the opposite direction. The time slots in DECT, however, can be used in either direction at any time and asymmetrical channels are supported as well.

If an ARQ protocol were to be implemented on a DECT-like multiplex, then the ARQ feedback for a packet could be received twelve time slots or 5 ms later, which is a very small round trip delay. If the video service uses one time slot (32 kbit/s) then the acknowledgement is received before the transmission of the next packet and a simple stop and wait protocol would suffice in this situation.

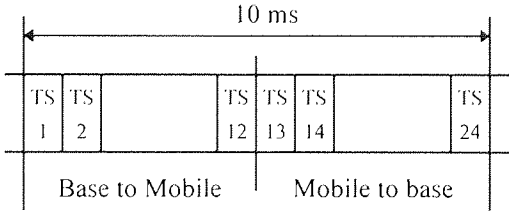


Figure 8-3 DECT MAC TDD frame structure.

Furthermore, the video frame structure is inherently compatible with ARQ schemes. Each DECT time slot contains 320 data bits. In a 64 Kbit/s videophone stream, if the video frame rate is 10 frames per second each video frame consists of 6400 bits on average which require 20 timeslots to

transmit. Thus if the first time slot is lost there would ample time for the slot to be re-transmitted before the frame is completely decoded and displayed. Thus, if extra retransmission bandwidth is available, ARQ can be used to recover lost packets in time for display without incurring any delay penalty with respect to the scheduled frame display time.

8.4.2 Multiple Access Control considerations

The MAC protocol has a strong influence on the implementation and performance of ARQ techniques. Consider the TDMA/TDD multiplex scheme adopted in DECT and shown in Figure 8-3. In this case the strict alternation of forward and reverse packets means that a simple SAW ARQ protocol will be as efficient as the more complex continuous ARQ schemes. If a data channel greater than 32 kbit/s is required in DECT this is achieved by using multiple time slots. Then two or more packets arrive at the receiver before they can be acknowledged. This makes it impossible to implement a SAW ARQ protocol unless the two packets are treated as one.

Consider now a TDMA/FDD system, a simple SAW protocol can be implemented on systems which stagger the forward and reverse packets as shown in Figure 7-4. If multiple time slots are to be used to increase the bit rate than the configuration in Figure 8-10 should be used to support a SAW protocol.

The performance of ARQ systems depends on the MAC adopted as well. For example if a TDMA/TDD system is used, then the packet loss on the forward and reverse links are correlated. If a TDMA/FDD system is used, then the forward and reverse directions are uncorrelated. This gives rise to different delay performance for identical ARQ protocols implemented on different MAC multiplexes.

8.4.3 Forward and reverse channels

The down link (base station to mobile) is more critical with respect to the application of ARQ to video transport. On the up link, the current state of the retransmission buffer can be fed back to the rate control algorithm. The video output rate can then be reduced to allow retransmission backlogs to clear and prevent monotonically increasing delays, even on fixed bit rate links. Alternatively, rather than retransmit the error packets, the encoder could re-code the effected data in intramode in the next transmitted frame. On the down link, this close coupling between the transcoder and the far end video codec does not exist and the design of appropriate retransmission buffer control protocols becomes a major consideration.

8.5 The BTRL ARQ System

The experimental ARQ system studied by MacDonald *et al.* at BTRL is shown in Figure 8-4 below. A Reed Solomon (RS) (63,59,2) FEC code was employed. This code can correct two six-bit symbols when used for error correction. When used for error detection only, the twenty four parity bits

guarantee a very low probability of undetected errors. The code can also be used to correct and detect errors, but the probability of undetected errors is significantly higher in this case.

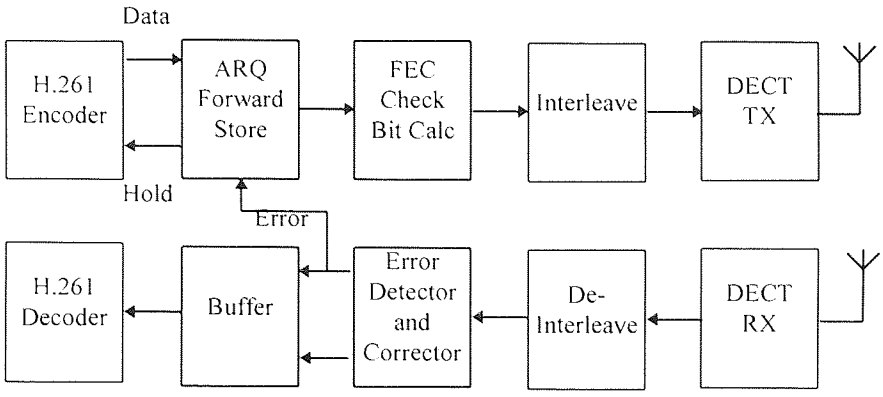


Figure 8-4 The BTRL ARQ system for H.261 transmission over DECT

When the FEC was used to correct errors only, with the ARQ system turned off, the picture quality degraded substantially. The use of the interleaver caused further degradation as the error bursts defeated the RS code and caused errors to spread to otherwise uncorrupted blocks.

The FEC code was then configured to correct and detect errors, and the detected errors were used to retransmit corrupted packets using a Selective-Retransmit ARQ technique, effectively forming a Type-I Hybrid ARQ scheme. The severe error bursts, however, caused many undetectable decoding errors and the picture quality degradation was again marked.

The best results were obtained using Selective-Retransmit ARQ, with the FEC code used to detect errors only. In this mode of operation the probability of error detection failure is very small, in the region of 2^{-24}, so that provided the retransmit buffers are prevented from overflowing, which would lead to lost packets, the picture quality should be largely unchanged. Buffer overflow was prevented by the use of the HOLD signals in Figure 8-4. However the HOLD control to the far-end encoder cannot be implemented on a real network.

A number of problems were identified in this paper, but were left un-tackled. These include the issue of the far-end video encoder output bitrate control to prevent monotonically increasing retransmit buffer delay, the bit rate mismatch between DECT and non-DECT video codecs and audio-video synchronisation. The remainder of this chapter investigates ARQ techniques which solve the first two problems.

8.6 Retransmission Buffer Control

A transcoder based ARQ system is shown in Figure 8-5. The transcoder receives the video information from the far end H.261 coder, and is in control of the bit rate of the video stream output to the ARQ controller. The ARQ controller contains the retransmission buffer modelled as in Figure 8-6.

Then the radio channel determines the buffer output bit rate R_o and the Transcoder the buffer input bit rate R_i.

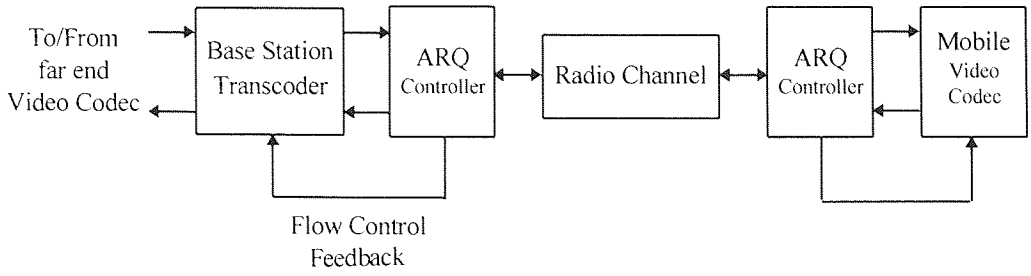


Figure 8-5 A transcoder based ARQ system

The H.261 video codec generates a fixed bit rate output. Thus the input data rate into the retransmission buffer is constant. Then each packet retransmission causes the retransmission buffer queue to grow by one packet. If the Packet Error Rate (PER) is high e.g. above 10^{-3}, the retransmission buffer delay grows without bound.

This problem may be modelled simply as a buffer (Figure 8-6) with unequal input and output rates, R_i and R_o. The input rate R_i is constant for standard H.261 codecs, whereas the output rate R_o may be nominally constant, but varies with retransmission requirements. Then queuing theory considerations require average R_o greater than average R_i for a stable buffer system [Kleinrock75]. The problem is that the R_o variations are difficult to predict and vary widely with propagation conditions and mobile speed. At the same time, the average R_o cannot be set too high as this would lead to bandwidth inefficiencies. Clearly some form of adaptive system is called for.

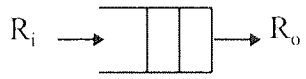


Figure 8-6 Simple Buffer System

Given a fixed buffer size, buffer overflow can be prevented by varying either R_i or R_o. Schemes which vary R_i depend on flow control of the video codec or transcoder output, whereas schemes which varying Ro are basically bandwidth-on-demand techniques.

8.6.1 Flow Control Techniques.

A straightforward flow control technique would be to feed back the state of the retransmission buffer to the far-end video encoder (which is attached directly to the WAN), which would incorporate this in its video coding control algorithm, in addition to the feedback from the transmit buffer in the coder itself. The variable transmission delay between the retransmission buffer and the far-end encoder complicates the design of closed loop flow control systems and may lead to video quality instabilities.

If the far-end video codec is a standard H.261 codec, then the technique described in section 8.4 above, can be used to implement a basic form of stop-start flow control. When the retransmit buffer is approaching overflow it signals to the transcoder which completes the delivery of packets belonging to the current frame and starts discarding frames. As soon as the re-transmit buffer is about to empty, the transcoder issues a fast update request and resumes delivery of the packets to the ARQ system as soon as the freeze picture release signal is detected. Thus each flow control event is accompanied by an intracoded frame which leads to a visible loss of temporal resolution in the received video which may not be acceptable.

If a full complexity transcoder is contemplated, the flow control issue becomes less critical because the transcoder encoder and the retransmit buffer would be closely coupled and the same solutions applicable on the up-link would be available. The transcoders contemplated in this thesis are low complexity, low-delay transcoders but flow control can be applied in this case as well as shown below.

A general transcoder flow control technique reduces R_i by discarding video data. Four levels of discard are possible; frame discard, GOB discard, macroblock discard and coefficient discard. In frame discard the transcoder discards whole frames as the ARQ retransmission buffer approaches overflow. The technique is the least complex but raises issues of picture quality degradation and recovery from the introduced errors. GOB discard is also simple to implement and raises the same issues as frame discard. Macroblock discarding is more complex to implement since differential macroblock headers have to be re-coded. With a properly implemented concealment mechanism, the introduced picture degradations would be markedly lower than for the other two schemes. Coefficient discard can be implemented by discarding a fixed number of coefficients from each block in the frame. If at least one coefficient per block is retained, the macroblock headers need not be changed, further simplifying the transcoder implementation. This technique has the least visible discard induced picture degradations and is less complex than macroblock discard. Frame and coefficient discard flow control techniques are analysed below.

8.6.1.1 Flow Control by discarding Frames

In this technique the transcoder reduces R_i by discarding complete frames. Thus when the transcoder estimates that retransmission buffer overflow is imminent, the transcoder will discard the next frame. This allows the buffer to empty but may introduce long term visible effects.

To study this possibility, the colour QM48 sequence coded at 48 kbits/s was decoded and the 38 th frame in the sequence was skipped. This frame has the largest interframe difference in the sequence. The frame contains 3791 bits and its discard reduces the input packet rate by more than eleven, 320-bit packets. The mean square error without forced updating (FU) is shown in Figure 8-7. The MSE is seen to reduce rapidly following the discarded frame but then reaches a rather high MSE

plateau. When the same frame is discarded from the same sequence coded with forced updating (at the rate of three macroblocks per GOB) the recovery is even stronger and falls rapidly within 5 frames, reaching completion within fifteen frames.

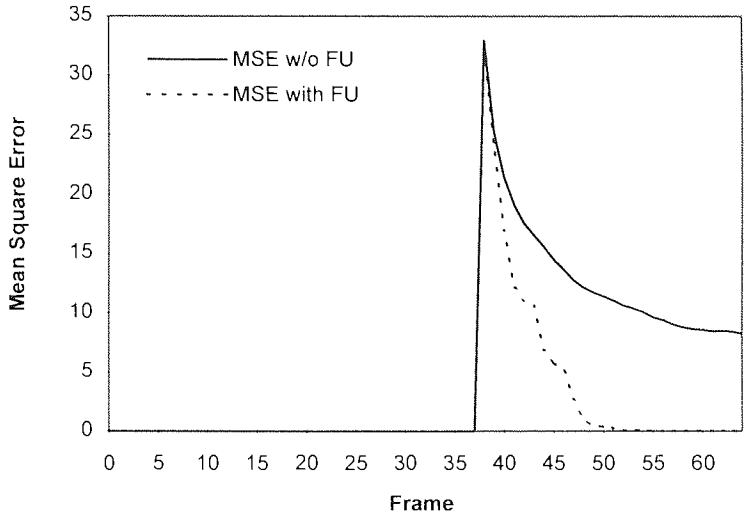


Figure 8-7 MSE with Frame Discard

The IHC decoder conceals the skipped frame by repeating the previous frame. This is quite visible at the frame in question (12.5 frames per second) and would be even more obvious at lower frame rates. The visible impact is not very objectionable.

The subsequent reduction in quality is not very obvious, unless the original and the errored sequence are viewed side by side. The forced updating technique can achieve complete error recovery inside one second. This indicates that flow control by frame discard works, and is very easy to implement.

8.6.1.2 Flow Control by discarding Coefficients

In this technique the transcoder lowers the instantaneous data rate R_i by discarding transform coefficients. The scheme still retains low complexity, because it can operate at VLC stream level. The main concern with this scheme is that once the coefficients are discarded, they are lost from the video decoder prediction loop, and lead to monotonically increasing image degradation.

The scheme was tested by inducing the simulated transcoder to discard coefficients in the 31 st to 38 th frames using the sequences described in section 8.6.1.1. The number of coefficients that can be discarded is programmable. Coefficients are discarded only if at least one coefficient would remain in the block. Also the *n* discarded coefficients are the trailing *n* coefficients in the zigzag scan.

The mean square error for the QM48 sequence without forced updating (w/o FU) is shown in Figure 8-8. Here one coefficient is discarded per macroblock between frames 31 and 38 inclusive. The total bit rate reduction achieved is 4767 bits. Comparing the MSE with that for Frame discarding in

Figure 8-7, it is obvious that the MSE is an order of magnitude lower with coefficient discard. Also there is rapid error recovery following the 38th frame and the error plateau reached is also an order of magnitude below that in Figure 8-7.

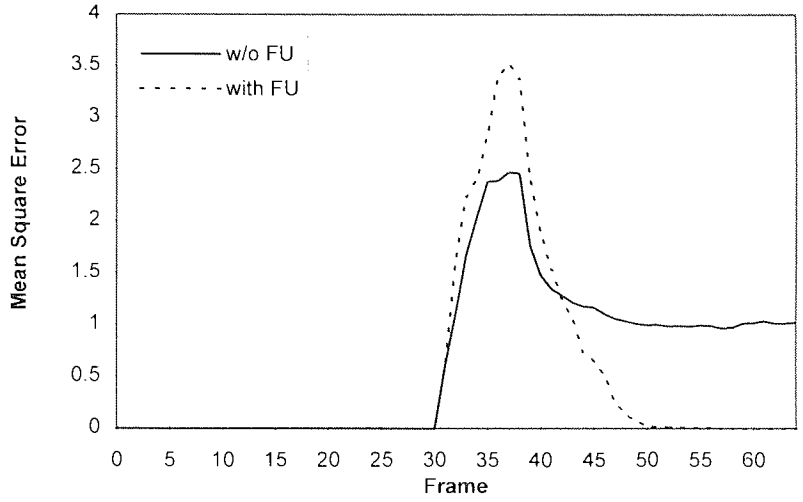


Figure 8-8 MSE with Coefficient Discard.

When the sequence with forced updating is transcoded, the coefficient discard achieves a frame rate reduction of 4078 bits. The MSE behaviour is shown in Figure 8-8 (with FU). The MSE is slightly higher than without forced updating but the error recovery is rapid and completes within ten frames. The MSE is also an order of magnitude lower than that with Frame discard. Thus coefficient discard achieves the same frame rate reduction as frame discard, and by distributing the rate reduction over more frames, the frame square error is an order of magnitude lower. The video quality reduction is difficult to perceive, even for a trained observer. The better results obtained by Coefficient Discard is achieved at the cost of increased complexity. The integration of the technique with the ARQ buffer control is also more complex and is explored in more detail below.

The success of the technique suggests the possibility of continuously discarding one or more coefficients per block per frame, as a means of transparent transmission rate reduction (with respect to the far end video codec). This possibility is investigated in Figure 8-9 below. The figure shows the resultant mean square error when one and two coefficients are discarded per block, with and without using forced updating, for the same QM48 sequence considered above. The bit rate reductions achieved were 39314 and 62280 with one and two discarded coefficients, without forced updating, and 34535 and 54869 with forced updating. This is equivalent to a reduction in bit rate between 13% and 25%.

The graphs show that there is a stable reduction in image PSNR, and that catastrophic quality degradation does not materialise. Thus the techniques can be used to achieve constant bit rate reduction. Subjectively, there is a distinct reduction in quality. The first intracoded frame contributes a lot towards this degradation. This is evident from Figure 8-9, where the first frame suffers the largest

MSE which is recovered in subsequent frames. Recovery is quicker for the forced updated sequences. Thus it is better to avoid discarding coefficients from the first intracoded frame.

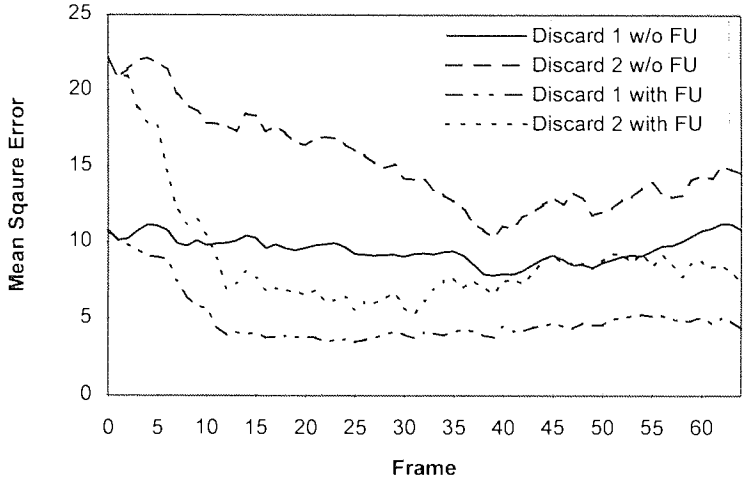


Figure 8-9 MSE with Continuous Coefficient Discarding

8.6.2 Bandwidth-on-Demand Techniques

The second option to prevent buffer overflow and reduce retransmission delay, is to increase R_o as necessary. This is achieved by making use of bandwidth-on-demand techniques to provide the extra retransmission bandwidth. Five such techniques are proposed and discussed in this section. The first is a variable rate system with extra packets allocated dynamically by the MAC. The second technique involves variable rate modulation to increase the transmission rate without MAC packet reallocation. The third option is to use a shared retransmission channel common to all links, with retransmission packets from all channels multiplexed and broadcast via the one common channel. The fourth technique creates additional downlink bandwidth by pre-empting up-link packets. The fifth technique uses down-link packet pre-emption, with the more important downlink packets pre-empting less important down-link packets.

8.6.2.1 Variable Rate MAC

The natural way of providing bandwidth on demand is to use variable rate MAC techniques. The Packet Reservation Multiple Access (PRMA) MAC [Goodman89] is an example of such techniques which has been proposed for video transmission systems [Zammit93, Steele94]. Another important protocol to emerge recently is the CSMA/CA technique proposed by the IEEE 8-12 wireless LAN standardisation committee [Links94]. A common feature of these MAC techniques is that they can accommodate isochronous channels, with the possibility of re-negotiating isochronous bandwidth requirements during the data transmission phase. The main advantage of using this technique is that all

the video packets get through eventually so that there is no degradation of image quality in either direction of transmission.

8.6.2.2 Variable Rate Modulation

Variable rate modulation has been proposed as an anti-multipath technique to increase the transmitted rate on wireless networks [Acampora87, Zhang90]. Variable rate modulation can provide extra bandwidth during advantageous propagation phases by switching to a higher rate modulation. Thus if the normal transmission rate can be accommodated by QDPSK, following a burst error event, the required retransmission bandwidth can be accommodated by switching to 8-DPSK, or 16-DPSK until the retransmission backlog is cleared. The rate change signalling information can be accommodated in a header field which is always transmitted with the lowest (*e.g.* QDPSK) rate modulation.

8.6.2.3 Shared Retransmission Channel

A wireless network is usually designed so that only a low number of users experience fading simultaneously. For example Acampora and Winters proposed a network with only a 10^{-4} probability of a user encountering a short term BER higher than 10^{-4} [Acampora87b]. Thus it is possible to reserve one or more time slots in both directions for the retransmission of packets and which are shared by many channels.

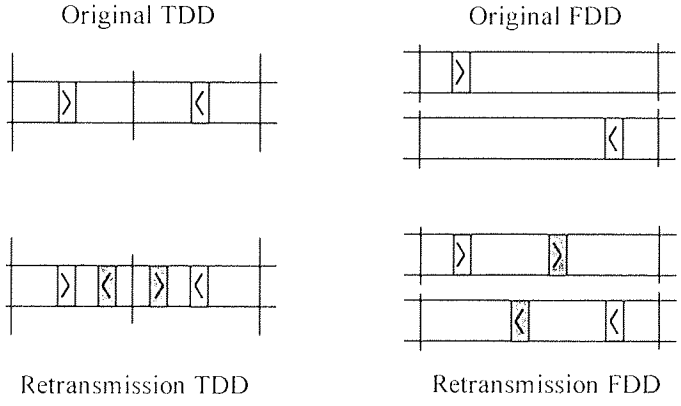


Figure 8-10 Retransmission Bandwidth Allocation

When a packet is corrupted on the downlink, the mobile station issues a retransmission request and begins scanning the known retransmission time slots. Upon reception of a retransmission request, the base station requests the use of the retransmission channel. When the retransmission channel becomes available, the base station transmits on this channel *in addition* to its allocated time slot. The extra time slots should be allocated in pairs, one on the uplink and one on the down link. If a downlink packet is lost this also means that an up-link acknowledgement packet is lost, such that the extra retransmission bandwidth is required in both directions. If the extra time slots are allocated so that these

occur *in between* the normal paired time slot allocations, as shown in Figure 8-10 for TDD and FDD, then these do not interfere with the ARQ protocol, even if a simple SAW protocol is implemented. Note that in the TDD case the MAC must be capable of supporting bi-directional packets in the normal TDD allocation, as in DECT for example.

8.6.2.4 Up link packet pre-emption

The additional retransmission bandwidth on the down-link can be implemented by pre-empting up-link packets. The idea is that the up-link is more robust to packet loss because of the close proximity of the retransmission buffer and the video coder in the mobile, as described above. Thus the mobile video codec buffer control would incorporate the retransmission buffer occupancy in its closed loop transmission buffer control algorithm. This technique solves the down link problem by sacrificing some image quality on the uplink. This scheme is most easily implemented with TDD MACs which support bi-directional forward and reverse channels as in DECT.

8.6.2.5 Down link packet pre-emption

If the data on the down link can be separated into two or more streams, with one stream carrying the essential structural elements and basic frame data, then each stream can be transported via independent packets. In this case retransmission bandwidth for the essential stream packets can be made available at the expense of the other streams' packets, by pre-empting these packets as necessary. This pre-emption scheme is more attractive than that in section 8.6.2.4 above, as the implementation is relatively simple, and does not require bi-directional time slots. This technique can be viewed as an instance of unequal error protection, and is studied extensively in the next chapter.

8.7 SAW ARQ for video transmission

In the previous sections it has been shown that a properly implemented TDMA MAC multiplex can support a simple SAW protocol. The use of such a basic protocol for video transmission is investigated next. Particular attention is given to the retransmission buffer control scheme. A flow control technique is proposed in this section based on the coefficient discard mechanism described in section 8.6.1 above.

8.7.1 System Description

The block diagram of the ARQ system investigated is shown in Figure 8-5. It consists of a transcoder which interfaces the mobile and fixed networks. The far-end video codec is assumed to be an H.261 compliant codec but which implements cyclic refresh at 3 macroblocks per GOB per frame, and uses the fixed length DCT intracoded frame described previously.

The wireless network is assumed to be the same as that in chapter 7, *i.e.* it is an FDD/TDMA based network with 10 ms frames and 240 symbol packets. The error detection data link packet format

is assumed in this section, and the 320 bit payload is taken up by the video data, which is generated at a fixed bit rate of 32 kbit/s. Then only one time slot per frame is required.

The transcoder performs the following functions:-

- It terminates the H.261 error framing at the fixed network side.
- It collects, buffers and re-packs the video data into the MAC error-detection data format.
- It implements flow control by discarding DCT coefficients.

The ARQ controller implements the SAW protocol. It buffers the data from the Transcoder pending confirmation of receipt from the mobile video codec. It also receives video packets from the mobile and extracts the reverse signalling channel. This section concentrates on the down link, since this is the critical part *from an ARQ point of view.*

8.7.2 The SAW protocol

The stop and wait protocol used in the simulation is based on protocol 4 in [Tannenbaum89]. The protocol allows full duplex transmission on one circuit. The data and control frames travelling in the same direction are multiplexed into one packet in piggyback fashion, which allows better use of the available bandwidth. The piggybacked control field for this protocol costs only one bit in the packet header.

The problem of how long the receiver has to wait for an acknowledgement identified in [Tannenbaum89], does not exist in this case since the MAC structure is such that packets travel on the down and up link in sequence. Then the synchronous nature of the MAC means that the packet arrival times are deterministic.

The protocol used belongs to the class of Sliding Window Protocols [Tannenbaum89] which are highly robust and can withstand any combination of errors and time outs without losing or duplicating data packets. In these protocols each transmitted packet contains a sequence number, which in this case need only be one bit long, since only two sequence numbers are required. The packet header also contains a one bit control field for acknowledging packets in the reverse direction.

It is assumed, without loss of generality, that the base station starts transmitting first. The protocol fails if both mobile and base stations start transmitting simultaneously [Tannenbaum89] but this cannot happen with the MAC multiplex under consideration.

The base station controller fetches the next packet to transmit, attaches the sequence number to the packet, transmits it and stores it in the retransmit buffer pending confirmation of correct delivery. The receiver knows which packet sequence number to expect, which can only be '0' or '1' in this case. If the checksum of the received packet is valid and the packet sequence number is correct, the packet is delivered to the video decoder and the receiver schedules an acknowledgement (ACK) on the feedback channel which is a copy of the correctly received packet sequence number. If the checksum fails or if

the PSN is not that expected the packet is discarded by the receiver and a negative acknowledgement (NAK) is sent in the form of the last correctly received PSN.

The feedback signals are transmitted on the next uplink packet. If the transmitter receives an ACK signal, *i.e.* the PSN number in the control field matches that one it is trying to send, then the transmitter removes the packet from the retransmit buffer and sends the next packet. If an NAK is received, the uplink packet is corrupted or if it is lost, the transmitter must retransmit the packet again.

The problem in applying this simple protocol is due to the monotonically increasing delay caused by each retransmission. The accumulated delay $\tau(t)$ at time t is equal to the total number of attempted retransmissions up to time t, multiplied by the frame duration. Thus unless the transcoder can control the flow of packets reaching it from the far end video encoder, the retransmission buffer will eventually overflow. Alternatively, as has been argued above, the transmission rate can be increased temporarily to clear the backlog, avoiding buffer overflow and meeting the delay requirements.

This problem is not as acute on the reverse direction. The close coupling between the retransmit buffer and the video encoder means that it is relatively easy to incorporate the retransmission buffer in the video buffer control algorithm, or even consider it as an extension of the same buffer. Then when the retransmission buffer fills the video codec reduces the rate automatically to prevent buffer overflow. This technique has been investigated by Watanabe *et al.* who have studied the buffer control function to optimise the subjective video quality [Watanabe94].

8.7.3 Flow Control

Flow control is achieved by feeding back information about the retransmission buffer occupancy to the transcoder. Whenever there are packets to be retransmitted, the transcoder enforces flow control by discarding a fixed number of coefficients from the tail of the DCT coefficient zigzag scan. Coefficients are only discarded if at least one coefficient remains in the block. In this way it is not necessary to change any macroblock header information, further simplifying the transcoder. The number of coefficients discarded is a control parameter of the flow control algorithm. The algorithm may discard a fixed number of coefficients or else it may implement a dynamic discard policy, to optimise some aspect of the system, *e.g.* to maximise video quality or minimise the retransmission buffer size. In the experiments reported below, the flow control discards a fixed number of coefficients per block per frame.

8.7.4 Simulation

The system in Figure 8-11 was simulated in C. The error streams generated in chapter 7 were used to corrupt the video data on the forward channel and the signalling information on the feedback channel. The ARQ controller is modelled by the retransmission buffer.

The simulation outputs the retransmission buffer size at each frame, and the number of discarded coefficients. The coefficient discard information is fed separately into the IHC decoder to

simulate the effects of flow control on the picture quality. The MSE and transmission rate reduction information are collected during this phase.

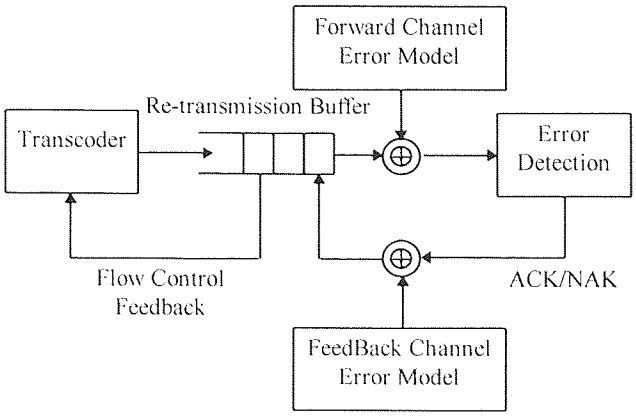


Figure 8-11 Flow control / SAW ARQ system model

8.7.5 Results - Flow control on a Rayleigh fading channel

The four error streams A, B, C and D, generated on a Rayleigh fading channel as described in chapter 7, were used to corrupt the forward and feedback packets in the model described above. The flow control algorithm was set to discard one DCT coefficient per block per frame, whenever the retransmission buffer contains data to be re-transmitted.

The resulting buffer occupancy with the four streams is shown in Figure 8-12. The results show that the flow control scheme can successfully control the buffer size for the four cases. The maximum buffer size is large, and the maximum retransmission delay is equal to 8 time slots or 80 ms. This is within the 100 ms buffer size usually employed with interleaving schemes which are usually in the region of one video frame period [Stedman93, Hanzo94]. The results also indicate that the 16-DPSK streams require larger retransmission buffers to prevent overflow and packet loss.

The 16-DPSK stream at 4 km/hr requires the largest buffer of the four sequences. The discard information for this sequence, from the 1500 time slot onwards, was used to control the coefficient discard of the IHC based transcoder, when decoding the robust QM32_23 sequence. The rate bit rate reduction in all cases was above the required 320 bits per frame using one-coefficient discard. The PSNR of the decoded sequence is shown in Figure 8-13. As can be seen the PSNR reduction is minimal. This is confirmed by subjective results since it is difficult to spot the differences between the two sequences, except in a side by side comparison test.

Thus the flow control mechanism in conjunction with the simple SAW protocol can support robust H.261 video transmission over a Rayleigh fading link *efficiently and effectively*. There are no visible impairments except some occasional loss of resolution. This good performance is obtained using high rate error detecting codes, and the throughput is twice that of the interleaved FEC technique in chapter seven.

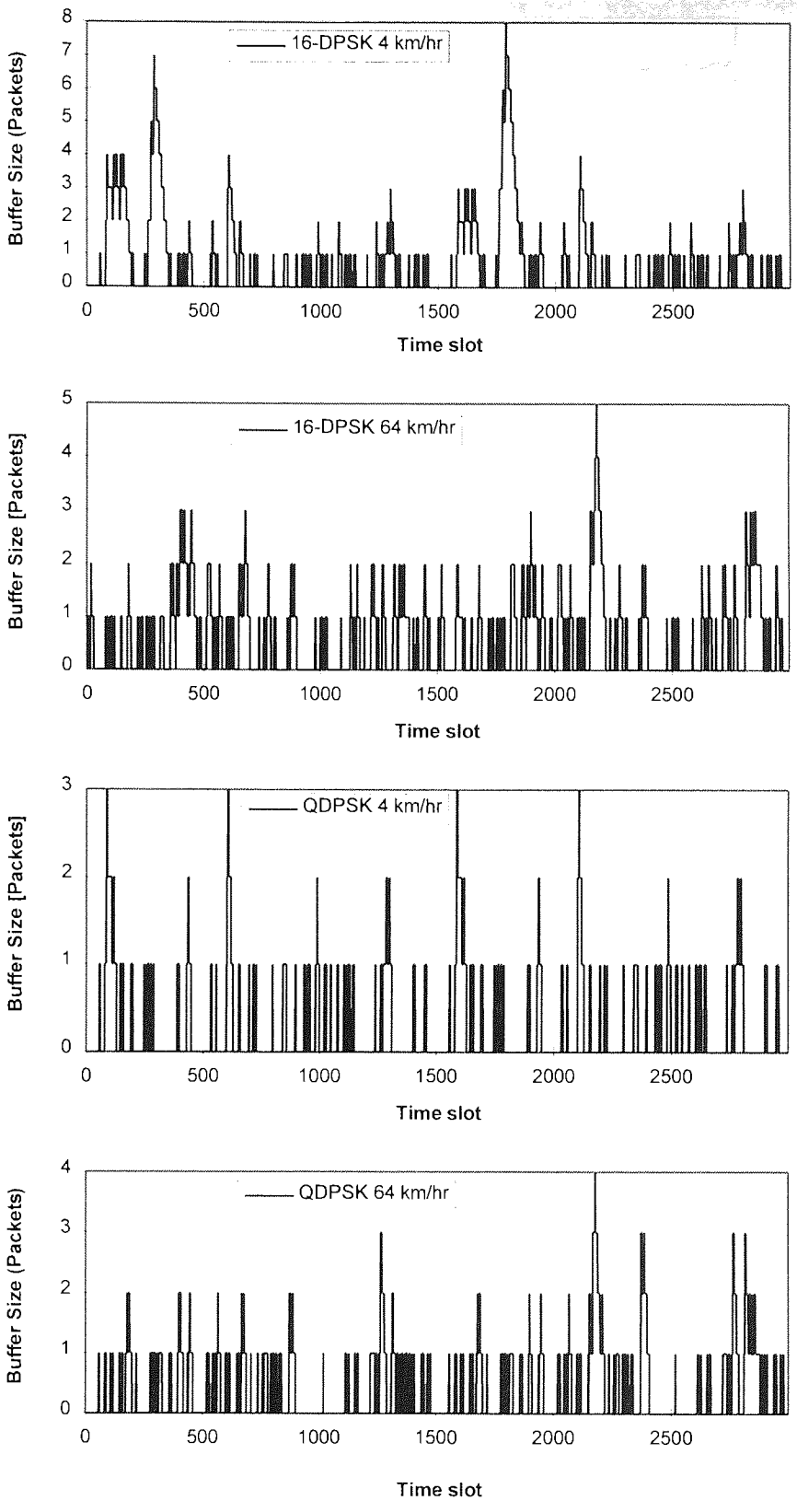


Figure 8-12 Down Link buffer occupancy using SAW ARQ [Tannenbaum89 - Protocol 4], transmitted over an FDD, TDMA wireless network with transcoder -based flow control. The channel is Rayleigh fading and the average E_b/N_o is such that the bit error rate is 10^{-3} for each of the modulations.

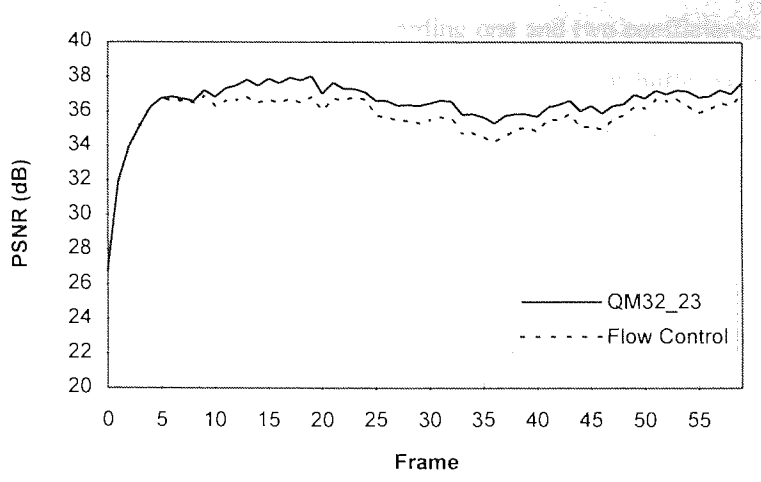


Figure 8-13 PSNR results for SAW ARQ with Flow Control in Rayleigh fading.
16-DPSK modulation, at a simulated speed of 4 km/hr

The retransmission delay introduced by the system is of the same order as that required by interleavers in FEC techniques. The retransmission scheme introduces some delay jitter in the stream delivered to the video codecs. This can be eliminated by introducing a dejittering buffer which need not be above 80 ms in this case.

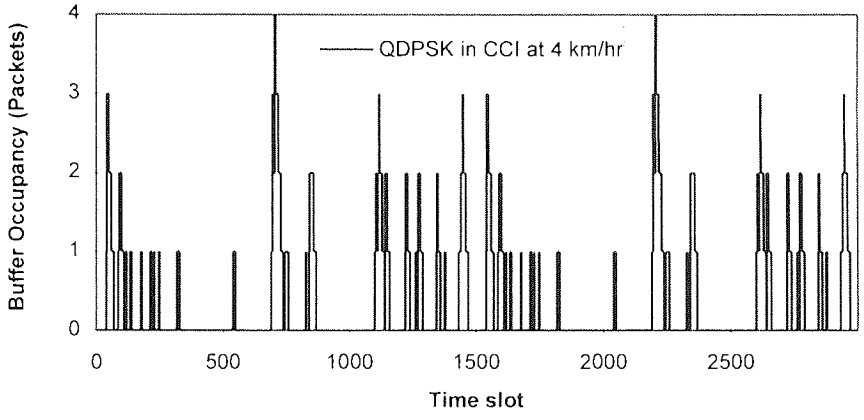


Figure 8-14 Buffer Occupancy with SAW ARQ and Flow Control - CCI channel
QDPSK with second order, post-detection selection diversity, at a mobile speed of 4 km/hr

8.7.6 Results - Flow control on a Rayleigh fading channel with CCI

The flow control technique was next tested on the Rayleigh fading channel with CCI. The error stream used was generated using QDPSK modulation scheme, at 4 km/hr and a CIR of 15 dB. When the flow control set to discard one and two coefficients, the buffer size grew without bound. The flow control managed to bound the buffer size only when three coefficients were discarded, but the maximum buffer size was high - 27 time-slots equivalent to 270 ms.

Further investigations revealed that the packet error rate on the radio links was 12 %. Thus 24 % of transmitted packets had to be retransmitted on average since both corrupted data and acknowledgement packets precipitate a retransmission request (the probability of acknowledgement

packet error was the same as data packet error). Discarding one and two coefficients, reduced the input rate R_i by only 10% and 20% which led to the unbounded increase in buffer size. Discarding three coefficients reduced R_i by 30% and resulted in a bound buffer size. When three coefficients are discarded, however, the picture quality degrades substantially and it was concluded that the flow control scheme fails on the CCI channel at a CIR of 15dB.

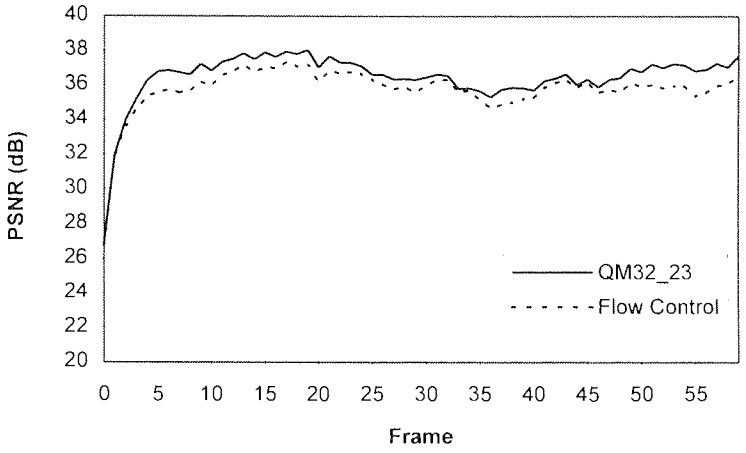


Figure 8-15 PSNR results for SAW ARQ with flow control on a CCI channel
QDPSK with second order, post-detection selection diversity, at a mobile speed of 4 km/hr

Second-order postdetection selection diversity [Parsons89] was then deployed to counter the effects of cochannel inteference. The packet error rate dropped to 1.4% as predicted by theory since the expected packet error rate with second order diversity is approximately the square of the packet error rate without diversity [Parsons89]. The flow control scheme bounded the buffer size tightly as shown in Figure 8-14 and the maximum buffer size was 4 packets on the Rayleigh fading channel with 15 dB CCI.

When the generated discard information was used with the IHD decoder to decode the QM32-23 sequence, the bit rate reductions were marginally above the required rate reductions. The reduction in video quality was marginal as well as can be seen from Figure 8-15. Again the difference in image quality is difficult to detect unless a side by side comparison test is performed. Hence flow control by coefficient discard has been show to allow video transmission even on this difficult, though often encountered, channel.

8.8 SR ARQ for video transmission

The last section investigated buffer control by flow control. This section investigates buffer control by bandwidth-on-demand. A more complex SR ARQ protocol is implemented. The main difference between this system and the previous one is that, when properly designed, the SR ARQ can deliver the video without any impairments.

8.8.1 Selective Retransmission ARQ System - Model and Description

A model of the proposed Selective Retransmission (SR) ARQ system is shown in Figure 8-16. The transcoder has the same functions as that in section 8.7.1 except that it has no role to play in controlling the ARQ buffer, unlike in the previous case. It is up to the ARQ control algorithm to deliver the data packets expediently and without loss.

The forward, reverse and retransmission channels are modelled as shown in Figure 8-17. The Error Detection Code (EDC) encoder codes the video data packets using a block code with good error detection properties. Then the error detection codewords are further protected by the FEC encoder. The FEC codewords are processed by one of the radio channel models described previously. The received codewords are decoded by the FEC decoder which corrects many low-weight error patterns introduced by the radio channel model. The decoded word is then decoded by the EDC decoder which traps decoding errors committed by the FEC decoder. The EDC decoder issues an ACK or NAK feedback signal depending on whether the data packet is received correctly or not.

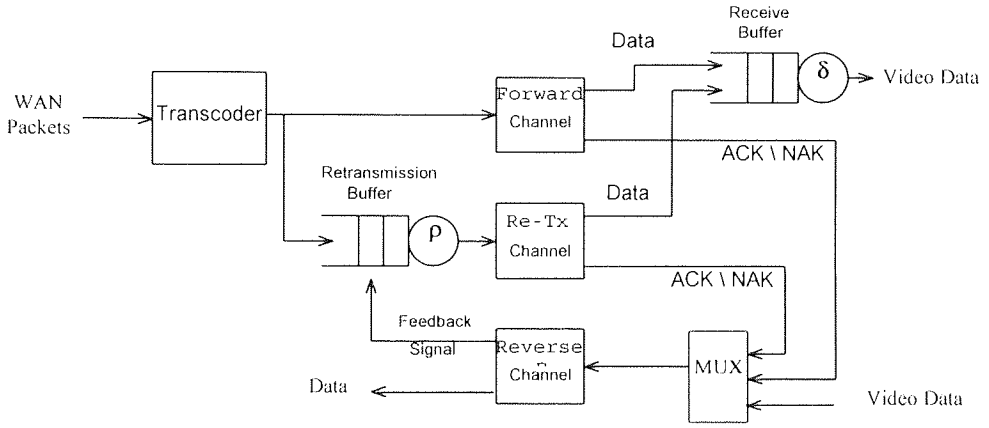


Figure 8-16 SR ARQ System model

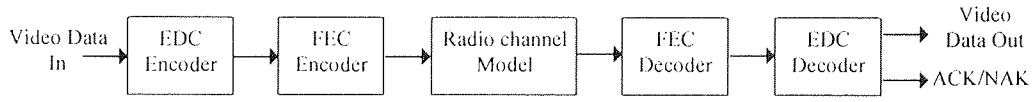


Figure 8-17 Channel Model for the SR ARQ system

8.8.2 SR ARQ Algorithm

The transcoder transmits video packets over the forward channel, and these packets are stored in the retransmission buffer pending acknowledgement. If the packet is received correctly it is deposited in the receive buffer and an acknowledgement (ACK) is issued on the reverse feedback channel. A packet remains in the receive buffer as long as there are previous packets which have not been received correctly yet or until a preset time out δ expires. The acknowledgement is transmitted over the feedback channel which may corrupt it using the associated channel error file. When the re-transmission buffer controller receives an packet acknowledgement it removes the packet from the

retransmit buffer. If a packet is negatively acknowledged it is scheduled for retransmission. If a packet is not transmitted successfully within the time out period δ, it is discarded from the retransmission buffer.

The novelty of the scheme is that retransmitted packets are retransmitted over a separate retransmission channel (Re-Tx) shared with other channels. All channels with scheduled retransmission packets contend for the services of this channel. The retransmission channel availability is modelled as a uniformly distributed random variable with probability ρ.

Retransmitted packets are decoded and acknowledged in the same manner as normal packets. The retransmitted packets' feedback signals use the normal reverse channel, which is always available. Thus the MAC must support the forward and reverse TDMA channels using either TDD or FDD, and the dynamically allocated retransmission channel.

8.8.3 System and Channel Model

The system was modelled as in section 7, i.e. it is a TDM/TDMA/TDD system operating at 192 kbaud. The channel error files were generated as described in chapter six. The modulation was QDPSK and the simulated mobile speed was 4 km/hr. The channel was modelled as Rayleigh fading with cochannel interference. The CIR was set to 15 dB and the AWGN and the ISI were assumed to be negligible. Second order, post-detection, selection diversity was deployed to mitigate the effect of the CCI.

The simulation accepts up to three different channel error files. If the error files are the same than the simulated system uses TDD. An FDD system uses different channel error files for the up link and the down link, with the retransmission channel having access to both. Alternatively the retransmission channel may be on a third, diversity, channel.

8.8.4 Results

The object of the reported simulation was to investigate the relationship between the three key parameters of the SR ARQ technique described, namely the receive buffer delay δ measured in time slots, the retransmission channel availability ρ, and the probability of packet loss P_e due to time out. The receiver buffer is organised as a circular buffer from which packets are retrieved in sequence every time slot, after the first δ time slots have been received. The circular buffer size and the packet sequence number (PSN) order coincide, so that all the packets with the same PSN occupy the same receiver buffer location. Then a time out error occurs if a packet is overly delayed and is not available when it is scheduled to be read out of the receive buffer. Thus the larger the initial access delay - δ (which is equivalent to the time out period), the longer the retransmission window. The access delay represents the maximum delay introduced by the system and packets which are delayed more than δ are discarded by the system. This is a fundamental property of the implemented *real time* continuous ARQ system.

In the reported simulation the forward and reverse channels were on diversity channels. The shared retransmission channel was simulated on a third diversity channel. The retransmission channel availability ρ was varied from 0.5 to 1 in steps of 0.1. An availability of 1 signifies that the retransmission is always available; 0.5 implies an even chance of finding the channel free. The receive buffer delay (and time out δ) was varied from 1 to 20. The residual packet error rate was monitored and the results are plotted in **Figure 8-18**.

The packet loss rate P_e on the link was high at 2.3%. When the channel was always available for re-transmission, the residual packet error rate was zero when the receive buffer delay was set above 11 time slots which is equivalent to 110 ms.

P_e increases as ρ decreases because more packets are overly delayed and violate the retransmission time out period δ. However a 200 ms receive buffer cancels packet loss when the availability is above 0.7. Thus the proposed technique can achieve packet-error free delivery with an absolute delay δ which is a function of the retransmission channel availability for a particular radio channel model. When the availability is lower than 0.7 and the receive buffer is set to 200 ms some packets are lost. However the residual packet error rate is low and is only 0.014% at 0.5 availability.

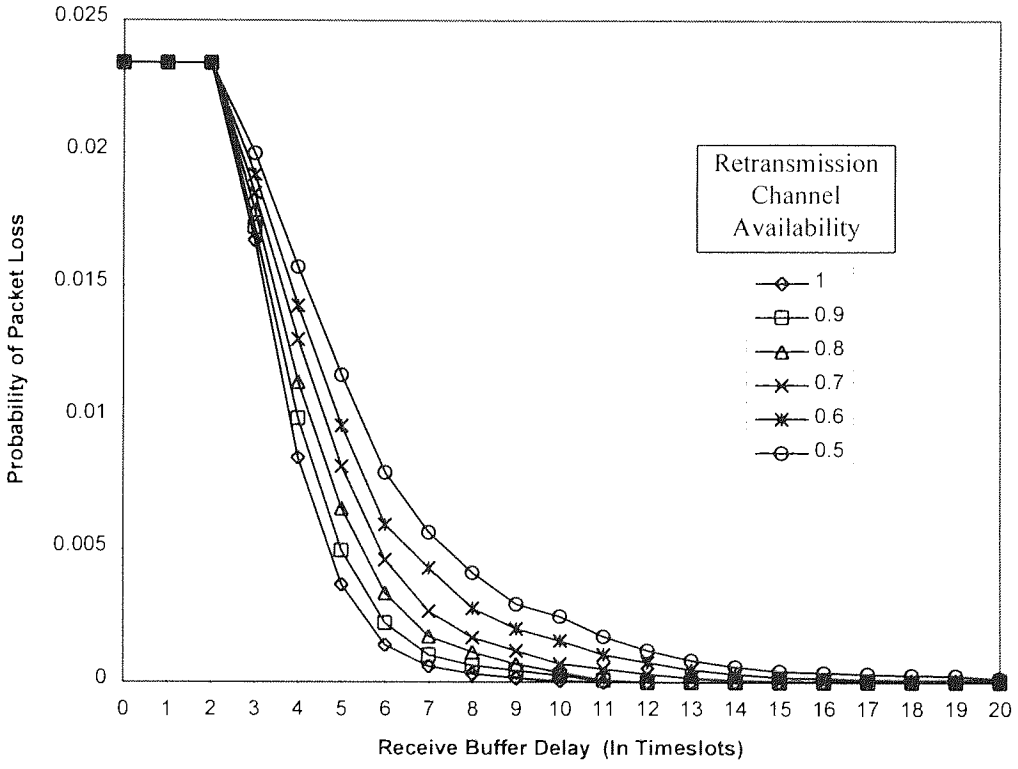


Figure 8-18 Probability of Packet loss P_e for SR ARQ (with diversity)

Thus SR ARQ allows unimpaired video to be transmitted over the radio link with acceptable delay, provided the retransmission channel is available. Even when this channel is only available half the time, setting the system delay to 200 ms reduces the packet error rate to around one packet per 10000 which allows video to be transmitted with very little degradation, if any.

8.9 The Trace-Back Decoder

The receive buffer is necessary in SR ARQ to deliver the packets in sequence but it introduces a delay in the transmission chain which can be substantial if the retransmission channel availability is low. The Trace-Back decoder (TBD) is an attempt to eliminate the delay associated with the retransmission buffer in the continuous ARQ scheme described above.

The concept behind the TBD is simple. Consider two real-time decoders which decode the same video stream, but using a different strategy. The first decoder decodes all the packets it receives immediately and discards all out of sequence packets. It also enforces a concealment strategy to hide reconstruction defects caused by the lost data. The second decoder decodes the data in the correct order, waiting for all the delayed packets to arrive in sequence.

Consider now the case where the first packet in a sequence is delayed. Then the second decoder buffers all the subsequent data packets it receives. As soon as the first delayed packet arrives, *the second decoder has all the data it needs to decode the stream up to the same point in time as the first decoder*. Then the second decoder can catch up with the first decoder if it can decode the data packets fast enough. If this is possible the first decoder copies the error-free reference memory from the second decoder and hence wipes out all the residual errors in its memory, achieving perfect temporal error recovery. The question is whether a decoder operating at twice the decoding speed of the first, say, can decode the data fast enough to catch up with the first decoder.

Alternatively the second decoder may attempt to decode all the data in the stream which does not reference the data from any delayed block, then when the delayed data arrives the decoder does not have to decode the whole sequence. This strategy makes it very possible for the second decoder to catch up with the first.

The Trace-Back Decoder combines the two decoders into one. It decodes the data which it receives as best as it can, but when retarded data arrives, the TBD *traces back* to the three dimensional co-ordinates (x,y and time) of the retarded block and reconstructs all the affected erroneous blocks.

Thus the retransmission delay is eliminated because the Trace-Back Decoder continues to decode the video data in the presence of packet delay, using error concealment for the current frame. Fast recovery is possible because the Trace-Back Decoder can use the out of sequence retransmitted packets to recover a perfect picture usually within one frame. Thus error concealment, combined with the fast error recovery result in quasi-invisible degradation in the presence of heavy packet loss.

The design of a TBD is quite complex due to the recursive nature of the searches that have to be performed to trace the affected data. Some of the design issues involved will be discussed in a future publicat ion.

8.10 Summary of conclusions

In this chapter it has been shown that error control by retransmission allows video services to be delivered effectively and efficiently over wireless networks and that these techniques should be

considered as viable alternatives to FEC techniques for the provision of real-time video services on wireless networks, since their adaptive nature allows very efficient use of the scarce radio spectrum.

ARQ techniques were investigated following the publication of a seminal paper by MacDonald *et al.* describing a prototype H.261 video transmission system for DECT, wherein they demonstrated the success of the prototype system but reported a number of problems with practical implementation in a real-world environment.

First, an analysis of the suitability of retransmission techniques for video transmission over wireless networks was undertaken. The main conclusion was that ARQ error control techniques are suitable for the proposed application, provided they are used across the radio link and not end-to-end across the fixed WAN. This conclusion was reached because the round trip delay on wireless networks is very low, and one video frame is transmitted using tens of radio link packets such that retransmitted packets can still arrive on time to be displayed in the correct frame without causing picture degradation.

It was also argued that it is easier to apply retransmission techniques on the up link because the encoder smoothing buffer can be closely coupled with the retransmission buffer. The video encoder buffer control can then vary the encoder parameters to optimise the video quality depending on the instantaneous state of both buffers. This close coupling is not possible on the down link.

An analysis of the BTRL prototype system revealed that the main problem in applying ARQ to real time video transport was the difficulty in controlling the bit rate of the far-end FBR video encoder to prevent monotonically increasing retransmission delay or retransmission buffer overflow. Two main classes of techniques were proposed to solve this problem; flow control techniques and bandwidth-on-demand techniques. The techniques are best implemented by using a transcoder at the interface between the fixed and the wireless networks.

Four flow control mechanisms were proposed based on low complexity transcoders to transmit H.261 coded data over a wireless network in a transparent manner. The four mechanisms monitor the occupancy of the retransmission buffer and when it starts to fill up the transcoder discards video information in an unobtrusive manner. The transcoder can discard whole frames, whole GOBs, whole macroblocks or DCT coefficients. Frame and GOB discard strategies are the least complex to implement but generate the worse image degradations. Macroblock discard results in less visible impairments, especially if concealment techniques are used to minimise the visual impact of the discarded macroblocks. However MB discard is the most complex technique to implement because it requires the macroblock headers of undiscarded blocks to be modified. DCT coefficient discard is less complex to implement than MB discard and results in the lowest picture quality degradation.

Tests showed that discarding frames had the desired effect of reducing the bit rate but the recovery time was excessive unless the special forced updating mode described in the previous chapters was used by the far end encoder. Flow control by discarding DCT coefficients was found to reduce the bit rate into the retransmission buffer whilst causing minimal visible degradations provided no

coefficients from the first intracoded frame are discarded. DCT coefficient discarding does not require complex transcoding because it only involves segmenting the variable length bit stream into codewords and discarding portions of the stream. It can also be used to reduce the bit rate permanently, to match disparate bit rates, by continuously discarding coefficients which leads to a constant and stable loss of picture quality.

The DCT coefficient flow control was then combined with a simple SAW ARQ system implemented on a TDM/TDMA/FDD wireless network. The system gave good results on Rayleigh fading channels, including those suffering from strong CCI, provided second order diversity is employed in the latter case. The retransmission buffer delay jitter was bounded to less than 100ms and the picture quality degradation was not easy to detect even on a 32 kbit/s coded video stream.

Bandwidth-on-demand techniques have the potential to deliver an unaltered picture quality but require more support from the wireless network infrastructure, particularly the MAC layer. Five such techniques were proposed:-

- Variable rate Modulation
- Variable rate MAC
- Shared feedback channels
- Up link packet pre-emption
- Down Link packet pre-emption.

The first technique is implemented in the physical layer, whereas the remaining four techniques are implemented in the MAC layer.

A real-time ARQ system was proposed, based on SR ARQ but which discards overly delayed packets. The retransmitted packets use a dynamically allocated retransmission channel shared with other video links. The performance of the system was studied on a TDM/TDMA/FDD 384 kbit/s digital wireless network, with QDPSK modulation, second order selection diversity, on a Rayleigh fading channel with average E_b/N_0 of 15 dB, at a simulated mobile speed of 4 km/hr. It was demonstrated that the technique can transport video error-free, provided the receive buffer time-out is sufficiently high to recuperate delayed packets. The required receive buffer delay was found to be strongly dependent on the availability of the shared retransmission channel. When the availability was set to 1, the time out period and hence the system delay was 110 ms, which compares favourably with the delay introduced by interleave buffers required by FEC techniques.

Finally, the concept of a Trace-Back Decoder was introduced. The TBD can decode retransmitted packets which are delayed and out of sequence. A proportion of errored packets are retransmitted in time to be decoded and displayed in the appropriate frame. Retransmitted packets which are delayed beyond their scheduled display time are not discarded but are decoded and replace

the concealed macroblocks in the reference frame. In this way errored macroblocks do not propagate for more than one or two frames at most. Thus there is no system delay penalty due to the retransmission scheme.

Chapter 9
Layered Video Coding for Wireless Networks

Layered video coding is an unequal error protection technique which has been proposed recently to mitigate the effect of cell loss on packet video in Asynchronous Transfer Mode (ATM) networks [Ghanbari89]. This chapter investigates video transport over wireless networks using layered video coding techniques.

Transmission errors on a wireless network result in packet loss as in ATM networks. The packet loss mechanisms on wireless networks are compared with those used on ATM networks.

The problem of providing a guaranteed channel with very low probability of error is identified as the major challenge in applying Layered coding to wireless networks. The techniques developed in the previous chapters can, however, be called upon to deliver the required robust, guaranteed channel. ARQ techniques are shown to be particularly attractive to fill this role. It is shown that the layered coding and the ARQ techniques can be merged synergistically in a wireless environment. The ARQ system is used to provide the robust, error-free, guaranteed channel required by the layer coding algorithm, whereas the layered coding enhancement channel provides readily available retransmission capacity for the ARQ technique.

Two-layered video transport systems are proposed in this chapter. The first system transmits fixed bit rate H.261 coded video over a wireless network using a layered transcoder.

The second proposal is a three-layer coding algorithm for end-to-end packet video transport system over a third generation Wireless Access Network (WAcN). The SR ARQ technique is used to protect the guaranteed channel.

The simulated layered codecs and transcodecs are based on the IHC video codecs described in chapter four. These are then used together with the simulation infrastructure developed in the previous chapters, to assess the validity of key aspects of the proposals.

9.1 Introduction.

The unequal error protection technique recognises the fact that not all the bits in a coded bit stream contribute equally to the reconstructed information quality [Modestino81]. Therefore, bits which are identified to contribute more to the information fidelity are better protected than the less significant bits. This divides the data to be split into layers, only one of which needs to be coded at the lowest rate, thereby improving the bandwidth efficiency.

Unequal error protection is used to protect digitally coded speech in wireless networks. In the GSM system, for example, the speech coder delivers 260 bits per 20ms [GSM]. The first 50 bits must be discarded if any of the bits are in error and a (53,50) block code is used for error detection purposes. The resulting 53 bit codeword is combined with the next 132 bits and then convolutionally encoded using a powerful half rate code, producing 378 bits. These bits are combined with the remaining unprotected 78 bits to form the 456 bit message block.

In 1982 Carr *et al.* used unequal error protection to transmit videoconference signals over a noisy satellite channel by giving special protection to the line and field synchronization signals [Ghanbari91]. Then in 1989, Ghanbari [Ghanbari89] applied unequal error protection, in the form of layered coding, to solve the problem of video packet loss in ATM networks.

The CCITT Broadband ISDN network uses ATM techniques and can support VBR services including video [Verbiest88]. There is, thus, the opportunity to transmit compressed video with constant quality by removing the FBR control system and transmitting the VBR output of the source coding algorithm directly [Verbiest88]. ATM networks, however, suffer from cell loss during congestion periods and this creates problems when transporting the vulnerable coded-video-streams.

Ghanbari proposed a two-layer coding technique to solve this problem [Ghanbari89]. Layered coding is based on unequal error protection, and operates by packing the important header information, motion vectors and essential video information bits into base layer packets, which are transmitted over a guaranteed channel with priority. The remaining video information is transmitted over a non-guaranteed enhancement channel. During congestion guaranteed packets are given priority over enhancement packets which are the first to be discarded by the system. Two-layer coding can also be regarded as a joint source/channel coding technique since the channel priority scheme is matched to the two tier resilience of the coding algorithm.

This chapter is structured as follows. Two layered coding techniques proposed by Ghanbari [Ghanbari89] and Kishino *et al.* [Kishino89] are described in the next section. Then the layered coding technique is studied with respect to its adoption for video transmission over wireless networks. A two-layer, transcoder-based, system is then proposed to transmit FBR H.261 coded video over a wireless network. The system is analysed and compared with the single-layer SAW ARQ scheme proposed in the previous chapter. Next, a layered error control system is proposed for transmitting video over a third generation wireless access network. The validity of key aspects of the proposal are investigated by simulation using the simulation infrastructure developed in the previous chapters and the layered video codecs developed from the IHC programs. The main conclusions are summarised in the final section.

9.2 Layered Coding.

In layered coding for ATM networks [Ghanbari91], the unequal error protection concept is applied by first partitioning the data into layers, with one layer, the base layer, containing all the data

required to decode a low-resolution image sequence. The remaining high resolution information, is coded separately. Unequal error protection is then applied by buffering base layer information with priority during congestion. Thus the unequal error protection is implemented by packet handling policies in the ATM network nodes and not by applying low- rate error correction codes.

The layered coding technique need not necessarily start from a single layer. In the two-layer system shown in Figure 9-1 the two coders code the video data independently. The Layer One coder codes the video at low resolution with a lower bit rate than the high resolution coder. The low resolution stream is transmitted over a priority channel (or a better protected channel). The high resolution decoder is selected when the second layer video stream is uncorrupted and can be decoded successfully otherwise the decoder switches to the more robust, lower resolution, Layer One stream.

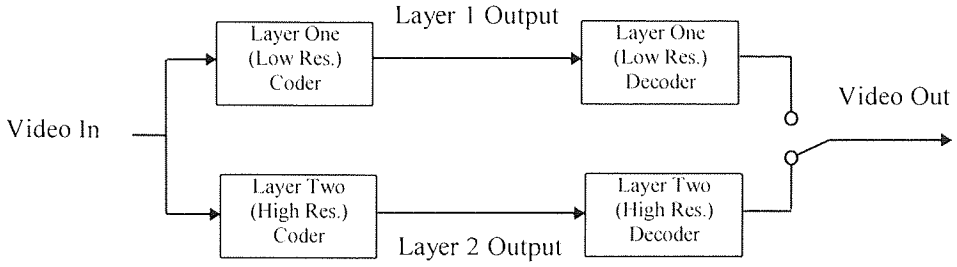


Figure 9-1 Two Independent Layer Codec System

Although the above conceptual layered coding technique is technically feasible, it is not particularly efficient as the same data is coded transmitted more than once, at different resolutions. The two-layer coding scheme proposed by M. Ghanbari for variable bit rate transmission over ATM networks, shown in Figure 9-2, is more efficient [Ghanbari89]. The input video is coded in two streams, one called the guaranteed packets stream and the other, the enhancement packets stream. It is essential for the guaranteed layer data to arrive at the destination, and hence it is assigned the highest priority. Data from the enhancement layer is less important and can be assigned a lower priority. The network treats higher priority packets preferentially. If traffic congestion leads to queue buffer overflows, the low priority packets will be discarded first. This technique is more efficient than the two independent codec system, because the output of the first layer coder is subtracted from the original video before recoding at the second, higher resolution, thus reducing the bit rate in the enhanced stream.

The scheme proposed by Ghanbari has been shown to perform well with a guaranteed channel rate as low as 10 to 20% of the total bit rate. The guaranteed channel operates ideally at a fixed bit rate to maximize efficiency. Therefore enhancement packets are diverted as required to pad the guaranteed stream during periods of underflow. This means that the enhancement bit stream fluctuates even more, a desirable attribute for channel sharing in a VBR environment such as ATM, for which the scheme was originally proposed.

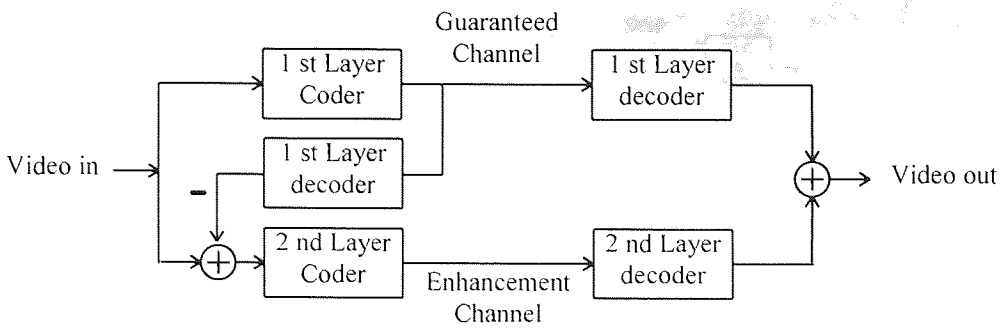


Figure 9-2 Generic Two-Layer Video Coding System.

Ghanbari used a H.261 codec for the guaranteed layer, modified to provide lower than usual bit rates (15Kbit/s) by coarser quantization of the coefficients. The enhancement layer was generated by requantizing the DCT coefficients with a finer quantization step size (Q_2) as shown in Figure 9-3. The overall bit rate was 110-120 kbit/s when coding CIF image sequences. The two-layer scheme was shown to be very robust to packet loss and performed well with an enhancement cell loss rate of 10%. At scene changes the cell loss effects become visible but only persist for one frame period. Thus two-layer coding will solve the cell loss problem.

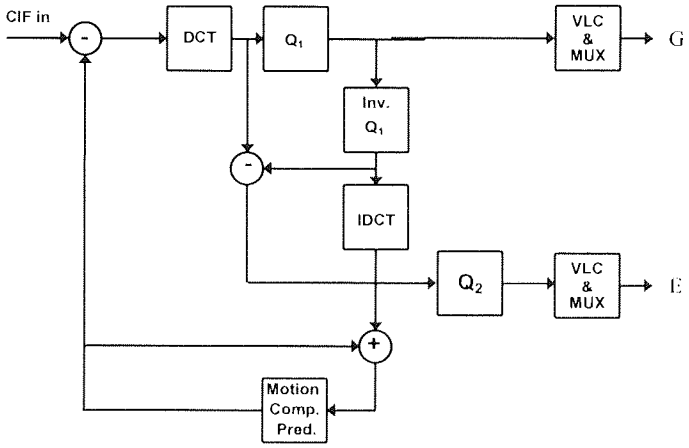


Figure 9-3 An H.261 Based Two-Layer Video Coder [Ghanbari89]

The above technique, however, generates bit rates which may be more than 100% higher than those of a single-layer coding producing the same (unerrored) video quality. Ghanbari and Seferidis found that by coding the enhanced layer with an interframe coder, using either interframe DPCM with a leak factor α or a second H.261 coder, the total bit rate was brought to within 10% of the single-layer coder for the same video quality, but retained the robustness of the two-layer scheme [Ghanbari93].

Kishino *et al.* proposed a two-layer coding scheme based on a hybrid DPCM/DCT coder as shown in Figure 9-4 [Kishino89]. In this case the two layers are built by splitting the quantized coefficients in a Most Significant Part (MSP) and a Least Significant Part (LSP). The DC and lower coefficients make up the MSP stream and the higher frequency AC coefficients, the LSP stream.

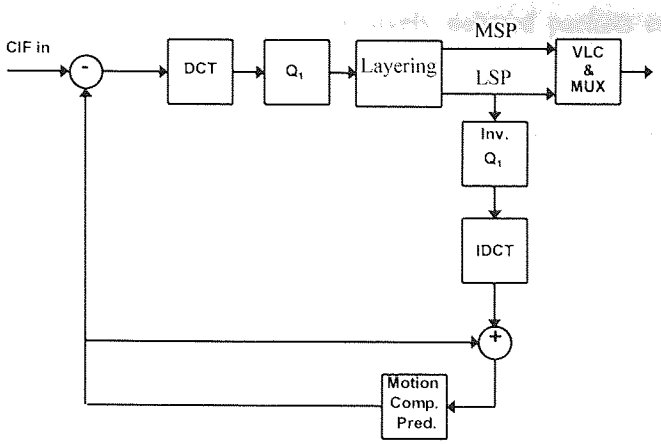


Figure 9-4. A Hybrid DPCM/DCT Two-Layer Video Coder [Kishino89]

Kishino *et al.* proposed two schemes for selecting the coefficient cut point between the two streams, one based on the coded bit length and the other on the SNR. The scheme was found to perform well even with LSP packet loss rates up to 25%.

9.3 Adapting Layered Coding Techniques For Wireless Networks.

Video data transport on the radio link is invariably packet based. Fixed sized packets are used in all the digital mobile and cordless wireless networks described in chapter two, although wireless LANs tend to adopt a variable packet size. B-ISDN uses fixed-size 53-octet packets, with five bytes reserved for header information.

There are two main similarities between the B-ISDN ATM and the radio link, packet transport systems. The first is that whole packets are lost in both cases. On ATM links packets are discarded during congestion. On radio links burst errors frequently render the data useless and the whole packet is best discarded. T he probability of packet header corruption leading the packet loss, however, is much higher than on B-ISDN ATM links. The second similarity is that both packet loss mechanisms are bursty in nature.

These similarities imply that the error control techniques developed for packet video transport over ATM may also be relevant to packet video transport over the radio link in wireless networks. In particular the layer coding technique developed for ATM networks promises particular advantages when applied to wireless networks.

Firstly, the fact that layer coded video can tolerate a substantial amount of packet corruption eases the stringent requirements of nearly zero packet loss in unprotected one-layer codecs. Thus the error control techniques applied to the enhancement stream need not be so strict, potentially leading to more efficient utilization of the scarce radio bandwidth.

Secondly, layered coding can reduce the isochronous bandwidth requirement since the enhancement packets can share bandwidth with other asynchronous traffic. Asynchronous data-oriented channels target packet loss at the expense of delay. The asynchronous enhancement layer

packets would be more delay sensitive, but excessively delayed packets can be discarded without undue loss of quality.

Another advantage is that the efficient ARQ techniques proposed previously can be readily employed to implement the guaranteed channel. The retransmission bandwidth is then readily available by pre-empting enhancement packets, if required.

Also, its adoption would ensure compatibility with the error control schemes being developed for B-ISDN networks. Since many wireless networks will ultimately lie at the periphery of global tethered B-ISDN networks, this would allow the development of a single error control technique suitable for both networks. Obviously this would simplify the network and codec design.

Finally, it is anticipated that layer coding would facilitate handover in wireless networks, by reducing the handover bandwidth requirement down to the base layer bandwidth. Initially, the enhancement cells may not all be accommodated on the new channel, and may have to continue to be delivered over the pre-handover channel with corruption, until bandwidth becomes available on the new base station.

There is one major obstacle in implementing layer coding techniques on wireless links. On tethered B-ISDN networks the guaranteed stream can be created by implementing a priority scheme, so that priority packets are treated preferentially during congestion. On wireless networks, packet loss is not due to congestion. Packet loss occurs due to the inherent nature of the radio propagation channel. Thus it is not a simple matter to guarantee base layer packets. Base layer packets have the same intolerance to packet loss as single-layered video codecs. Thus, the countermeasures proposed in the previous chapters for single-layer codecs, need to be applied to the base layer coder to create the guaranteed channel.

Another problem is that the packet loss rates on wireless networks may be substantially higher than on ATM networks. Thus the layer coding techniques developed for wireless networks may require different optimizations than those proposed for ATM.

The layer coding algorithms of interest here are those based on hybrid predictive DPCM/DCT codecs. A number of two-layer coders have been proposed for B-ISDN networks based on the H.261 standard [Ghanbari89, Morrison91, Tubaro91]. Two H.261 based layered codecs are investigated; a two- layered coder similar to the one described by Kishino *et al.* and a three-layer codec similar to the two-layer codecs described in [Ghanbari89] and [Morrison91].

9.4 A Two-Layer Video Transport System For Wireless Networks.

The ARQ techniques proposed in the previous chapter have been shown to be suitable for protect-ing coded video transmission on fading radio links. However the retransmitted packets create a need for an on-demand retransmission bandwidth. Layered coding, on the other hand, separates the video data into essential and non-essential streams, but requires a robust channel for the essential data - the guaranteed channel. There is thus the opportunity of combining the two techniques synergistically;

the ARQ technique providing the robust, guaranteed channel, and the enhancement packets providing a readily available supply of retransmission bandwidth. This observation was first made by Zammit and Carpenter [Zammit93] and used to transport packet video over a third generation wireless network (to be described below). Khansari and Mermelstein [Khansari93] later proposed a similar technique for video transport over a CDMA based wireless network.

9.4.1 System Description.

Two-Layer coding can be combined with ARQ to implement a system for accessing fixed bit rate coded H.261 video over a TDMA wireless network. The down link for such a system is shown in Figure 9-5. The up-link can be provided using the same technique or using the combined smoothing/retransmission buffer algorithm referred to in section 8.4.3. Only the down link is considered here.

The far-end coder generates a normal H.261 fixed bit rate stream for transport over a fixed bit rate network such as ISDN. The transcoder terminates the H.261 error correction framing layer and thus prevents end-to-end loss of synchronisation in this layer. The transcoder then separates the video stream into two, a guaranteed stream containing the essential headers and video data, and an enhancement stream containing the non-essential video data. This part of the transcoder will be described in detail below.

The Guaranteed Channel is protected by a SR ARQ scheme similar to the one considered in section 8.8. The transcoder uses two independent channels from the TDMA multiplex. The feedback signalling channels are provided in piggyback fashion on the reverse links as described above.

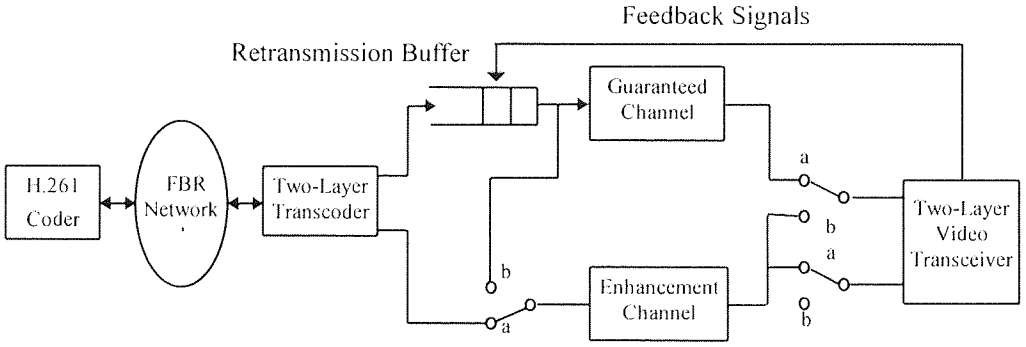


Figure 9-5 Down-Link Model For Two-Layer Video Transport.

In normal operation the three switches in Figure 9-5 are in position a. Then the guaranteed packets travel through the retransmit buffer and the guaranteed channel, and the enhancement packets through the enhancement channel. The two packet streams are combined by the two-layer decoder and decoded with perfect reconstruction in the absence of channel errors. When the guaranteed stream packets are corrupted these are retransmitted over both the guaranteed and enhancement channel, as long as the re-transmit buffer contains unacknowledged packets. To retransmit a guaranteed packet

over the enhancement channel, the three switches are switched to position b pre-empting the next enhancement packet.

The buffer size does not increase monotonically using this technique. The retransmission mechanism creates a packet arrival delay jitter which depends on the combined packet loss statistics on the forward and reverse channels. These in turn depend on the radio propagation environment and on the MAC protocol, especially on the type of duplex employed and whether the channels operate over diversity paths.

9.4.2 Two-Layer Transcoder.

The transcoder must operate on the output of a single-layer H.261 stream and produce two unequal priority streams. The technique used to achieve this is similar to that proposed by Kishino *et al.* and described above. The low complexity transcoder operates on the video stream at VLC level. The Picture, GOB and Macroblock VLC headers are copied unaltered to the guaranteed stream. The block coefficients' variable length codes are split into two groups; the guaranteed coefficients group and the enhancement group as shown in Figure 9-6.

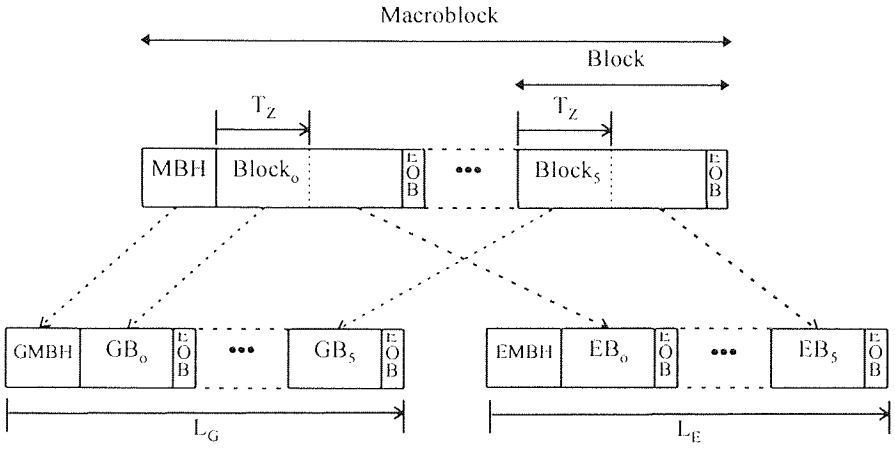


Figure 9-6 Generating The Guaranteed And Enhancement Macroblocks

The guaranteed group contains the first T_Z coefficients in the H.261 zigzag scan. The remaining coefficients are grouped together into the enhancement coefficients group. Each group in each block is terminated by the EOB field (01_b). The transcoder must determine T_Z so that separation into two groups results in the required number of guaranteed bits L_G and enhancement bits L_E. Thus, if the strategy adopted requires equal sized streams, T_Z is selected such that $L_G = L_E$. This partitioning may be undertaken using both open loop and closed loop techniques. In the latter case, the running total L_G and L_E for the frame are used to fine tune the selection of T_Z. This is equivalent to the buffer control algorithm in normal coders, but in this case the operation may be carried out at macroblock level and does not introduce significant delays.

The guaranteed coefficient blocks are then redirected to the guaranteed stream. The reconstructed guaranteed stream is thus H.261 compliant and can be decoded using any H.261 decoder.

The enhancement blocks are grouped into macroblocks which are preceded by enhancement stream macroblock headers consisting of an absolute macroblock address field (fixed length coded), and a Coded Block Pattern field with VLC codes identical to the H.261 specification. The CBP field is required because some blocks will contain less coefficients than T_Z and will be transmitted entirely within the guaranteed stream. The enhancement stream macroblocks are partitioned into frames and GOBs using a picture and GOB header fields identical to the H.261 specification.

9.4.3 Two-Layer Decoder.

The decoder receives and buffers guaranteed and enhancement stream packets in two separate buffers. The decoder fetches the next macroblock data from the guaranteed stream and decodes it to quantized coefficients level. The decoder then searches in the enhancement stream for additional macroblock data. If it is found, it is decoded to quantized coefficient level. The two sets of macroblock quantized coefficients are then combined and decoded. If the enhancement data is not found in the enhancement stream the decoder proceeds to decode the guaranteed macroblock without the enhancement coefficients.

The coefficient data lost in enhancement packets, is lost from within the prediction loop and will continue to remain visible until the area is refreshed by forced updating. However the lost data is very well concealed by the guaranteed data, which is sufficient by itself to deliver a meaningful display, clear of highly visible artefacts or picture break-ups. Thus the scheme inherently exhibits strong macroblock-level error concealment. Error recovery by forced updating or other means can also be implemented, but the recovery times are less severe because of the acceptable quality of the base layer.

9.4.4 Analysis.

When the system described above was analysed further, the following points emerged.

(a) The creation of the low priority, enhancement layer, introduces a bit rate overhead, in the form of the enhancement layer picture, GOB and macroblock headers. The headers are required to limit spatial error propagation by providing re-synchronization points following variable length codeword corruption.

(b) The system still requires cyclic refresh to recover from temporal error propagation because lost blocks represent lost data in the reference frame.

(c) The system behaviour is very similar to that of the SAW ARQ technique with flow control described in chapter 8. Packet errors on the radio link require retransmission and lead to the pre-emption of low priority data, which is selected identically in the two cases.

Hence it was realised that the proposed system had potentially the same performance of the flow control system, but had a higher bit rate overhead. Furthermore, the SR ARQ protocol is more complex to implement than the simple SAW ARQ protocol. Thus it was concluded that the flow

control system is preferable, unless the MAC protocol precludes the use of the simple SAW ARQ technique.

9.5 *Layered Video Transport Over A Third-Generation Wireless Access Network.*

In the previous section it was shown that layered transcoder systems using SR ARQ to protect the guaranteed channel have potentially the same performance of the single-layer SAW ARQ protected transcoders combined with simple flow control. This section investigates the transport of ATM video over a third generation wireless network. It is assumed that a layered video coder is used on the ATM network to protect it from cell loss caused by the ATM network. Then the arguments presented above, with regards to the overhead incurred in providing the enhancement stream in the layered transcoder, are no longer valid. The issue of protecting the guaranteed stream remains, however.

If a single-layer SAW ARQ system with flow control is used to protect the guaranteed layer, the coefficient discard causes a loss of resolution in the base layer. If a single-layer SR ARQ technique is used, this resolution reduction does not occur. Hence there is now a potential advantage in using a SR ARQ system instead of SAW ARQ to protect the sensitive base layer video data. If SR ARQ is used, the enhancement data packets can be pre-empted to provide a readily available supply of re-transmission bandwidth. This section investigates the use of SR ARQ in combination with layered video coding and studies the reduction in picture quality which results from pre-empting the enhancement packets.

9.5.1 A Third-Generation Wireless Access Network

Layered video transport is studied in the context of a third generation Wireless Access Network (WAcN) [Zammit93]. The network is assumed to be similar to the Cellular Packet Switch (CPS) proposed by Goodman [Goodman90, Goodman91] and is shown in Figure 9-7.

The network consists of a number of base stations providing microcellular radio coverage to wireless terminals within 100m (typically). The Base Stations (BS) communicate with each other and with the Cellular Control Unit (CCU) across a Metropolitan Area Network (MAN) backbone which supports data rates in the region of 150Mbit/s. User, signalling and system control data is transmitted over the network in packets. The user data packets can accommodate one ATM B-ISDN cell. The BS and CCU units connect to the MAN via BS and CCU interface units respectively (BIU and CIU). The Trunk Interface Unit (TIU) interfaces the CPS to the B-ISDN. The wireless mobile video user accesses the fixed network through the WAcN by using a Wireless Interface Unit (WIU) - a radio based transceiver.

The Wireless Network protocol layers of interest concern the error control schemes in the data link control layer and the physical resource allocation of the logical channels. Routing of the data and signalling packets is performed by the network layer and is not considered here.

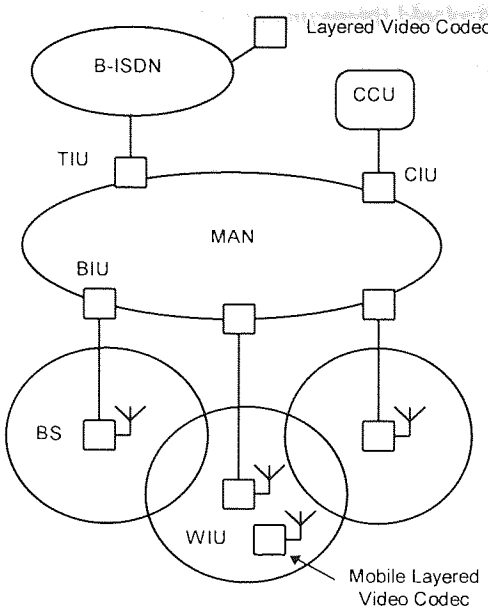


Figure 9-7 Third-Generation Wireless Access Network

Of particular interest to video communications is the Medium Access Control (MAC) scheme used since this will relate directly to important transport characteristics, such as delay. The MAC adopted employs TDM on the down link (BS to WIU), Packet Reservation Multiple Access (PRMA) [Goodman89] on the uplink (WIU to BS) to support voice, data and enhancement video packets, and isochronous TDMA for the guaranteed video channels. Frequency division duplex is used. Both the up links and the down links are assumed to operate at 10 Mbit/s with 10ms frames.

9.5.2 A Three-Layer Video Codec

A three-layer video codec is proposed for use in the hybrid fixed/wireless network environment. The block diagram of the three-layered coder is shown in Figure 9-8. It is similar to the one proposed by Ghanbari [Ghanbari89] and Morrison and Beaumont [Morrion91]. The main difference is that the enhancement layer has been split into two to make the enhancement layer more robust to cell loss and cell pre-emption. The codec was simulated in software and is part of the IHC suite of programs described in chapter four.

The enhancement layer is derived by requantization of the difference between the quantizer input and the inverse quantizer output using a finer quantizer [Morrison91]. This technique eliminates the need of one DCT block in the encoder and generates a compressed enhancement stream, which is useful for bandwidth conservation on the WAcN.

The buffer control scheme adjusts the quantization step to generate video data at a constant bit rate. The output of the smoothing buffer is packetized to form the guaranteed data stream G and is assigned to the high priority ATM channel. The guaranteed channel data rate is set to some fraction of the total mean bit rate produced by the coder, one half say.

A fixed quantizer is used to requantize the enhancement blocks using the quantization tables in the H.261 specification. The requantized enhancement macroblocks are coded and multiplexed with the macroblock header information. Alternate macroblocks are routed to two separate packetizers and combined with Picture and Group of Blocks headers to form two autonomous enhancement streams, E1 and E2, with approximately the same rate. The picture and GOB headers in the enhancement stream are identical to the H.261 specification. The enhancement stream macroblock header contains a 9 bit (7 bit for QCIF) absolute macroblock address and a coded block pattern VLC field with codewords identical to the H.261 CBP table.

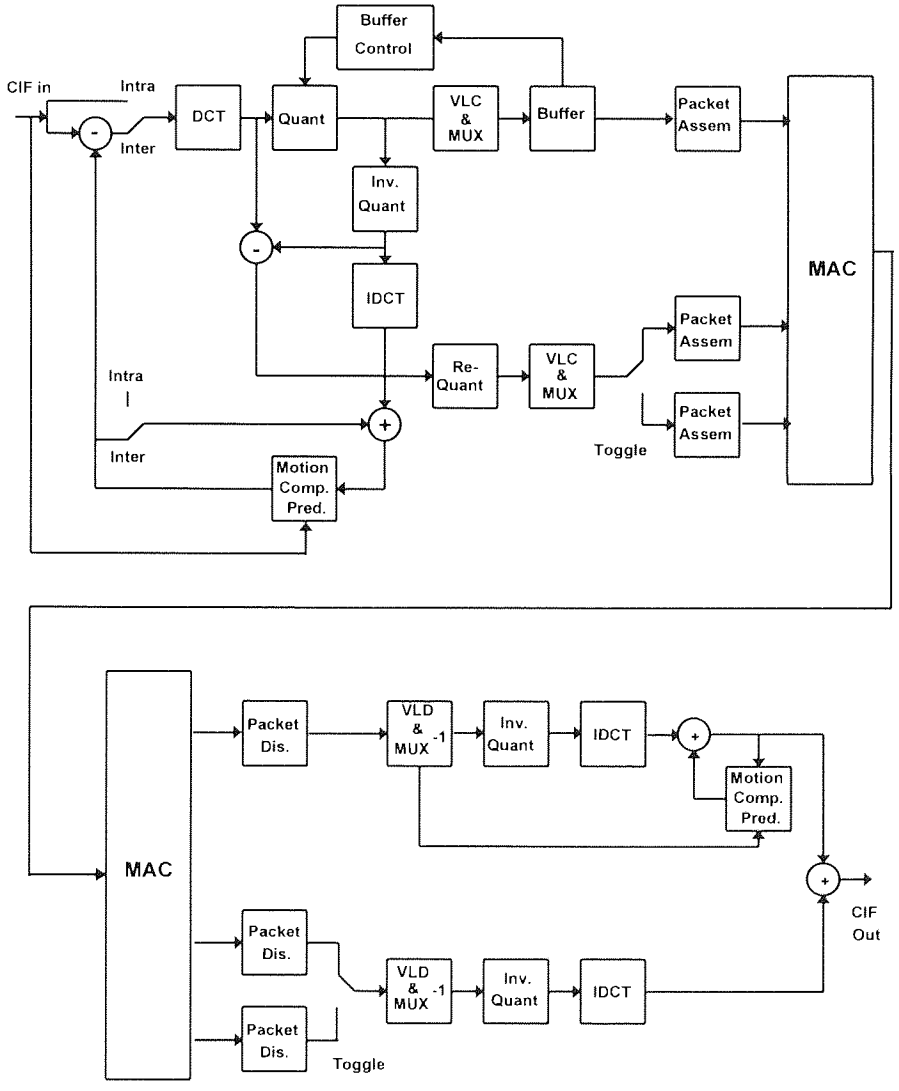


Figure 9-8 A Three-Layer Video Codec

A lost or pre-empted enhancement stream packet causes the loss of synchronization in the E1 or E2 streams until the occurrence of the next GOB header. If only one enhancement stream is lost, the data in the other stream can be used to generate concealment data by interpolation.

9.5.3 Error Control Sub-System

The proposed error control system is a joint source/channel coding scheme based on the three-layer codec described above and fast retransmission of lost or corrupted guaranteed channel packets across the radio link, to meet the low bit error rate required by the guaranteed stream.

The error control sub-system block diagram is shown in Figure 9-9 and is used for both up and down links. The ATM G, E1 and E2 packets are re-packetized (RE-PAK), replacing the ATM packet headers with wireless network specific headers. The wireless network packets are then passed to the data link layer.

The data link layer protects the enhancement packets using forward error correcting (FEC) block codes. The two enhancement streams are transmitted over independent MAC channels ($E1_F$ and $E2_F$). The two enhancement channels may be provided by different base stations at any one time. Either of the two streams may be partially pre-empted by the guaranteed stream packets during handover on a congested wireless network.

The guaranteed layer packets (G) are transmitted over an isochronous channel (G_F) and must arrive at the video decoder with very little residual error. A hybrid type-I SR ARQ scheme is used to meet the low BER requirement. Cascaded error detection and FEC codes (EDC/FEC) have been used in the simulation.

The ARQ scheme has been modelled by the SR ARQ system model described in section 8.8.1 and shown in Figures 8-16 and 8-17. The retransmission channel (R_{TX}) is implemented by pre-empting one or both of the enhancement packets. If the enhancement packets are allocated an isochronous channel, then the availability ρ (Figure 8-16) is equal to unity, meaning that the retransmission channel is always available. If the channels are allocated dynamically then the availability ρ is fractional.

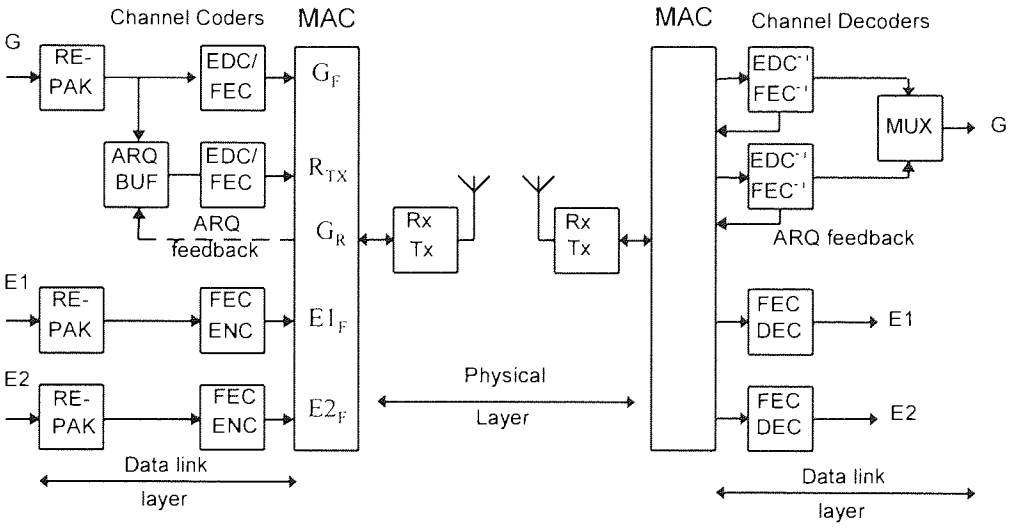


Figure 9-9 Block Diagram Of The Hybrid ARQ-FEC Error Control Scheme

9.5.4 MAC Requirements

The proposed error control scheme requires two classes of channels which must be provided by the third generation wireless network. It is proposed that a third generation MAC protocol layer should provide three classes of channels to support multi-media services including video services. The three channel classes are defined below.

Class 1 channels are isochronous, have the highest priority and are protected by the hybrid ARQ scheme in the data link layer. These channels carry the guaranteed video packets with retransmission over lower class channels.

Class 2 channels have the same priority as class 1 but use a dynamic channel allocation MAC protocol, such as PRMA, to support voice channels.

Class 3 channels have the lowest priority and also use a dynamic channel allocation MAC to support data traffic, the two video enhancement streams and retransmission packets.

The MAC must implement a reliable ARQ signalling channel using fields in the forward and reverse guaranteed packets. The MAC scheme has to be capable of re-allocating an enhancement channel to transmit guaranteed channel packets should there not be enough free packets during handover. Finally the WIU must be capable of communicating with two base stations simultaneously.

9.5.5 Third generation MAC

A third-generation MAC is proposed in this section, which can support real-time video services. The third generation MAC is based on a TDM/TDMA/FDD scheme. The MAC supports both isochronous and asynchronous traffic. The synchronous channels are assigned fixed time slots on both the up link and the down link. The remaining time slots are available for asynchronous traffic and may use PRMA as suggested in [Zammit93]. The normal forward and reverse channels are allocated isochronous time slots. The retransmission channel is allocated time slots dynamically from the asynchronous time slot pool.

The proposed third generation MAC supports a nominal bit rate of 4 Mbit/s. There are 100 frames per second in the multiplex with 80 time slots, each with 512 bits. On time slot supports a bearer data channel of 32 kbit/s. Multibearer services are possible by aggregating time slots.

The time slots on the down link and the up link are staggered to better support retransmission schemes. The time slots in a multibearer channel are not assigned contiguously, but are spaced out over the entire frame, with up link and down link time slots interleaved one-to-one, again to facilitate retransmission schemes.

9.5.6 Simulation Results

The channel was simulated using QDPSK at a speed of 4 km/hr. The symbol rate was 2 Mbaud resulting in a symbol period of 488 ns. The rms delay spread of the channel was 50 ns resulting in a

normalised delay spread of 0.1. The channel therefore suffers from substantial intersymbol interference and second-order post-detection selection diversity is deployed to mitigate the effects of the ISI. The AWGN and CCI are ignored in the simulation.

The salesman sequence was coded using a three-layer video codec with a 320 kbit/s allocated to the fixed bit rate guaranteed channel and a further allocation of 320 kbit/s for the enhancement layers. The bit rate for 100 frames is shown in Figure 9-11. The guaranteed channel was transported using 10 isochronous time slots per 10 ms frame on the TDM down-link multiplex. These time slots are separated by 8 timeslots from each other. The uplink timeslots are allocated on the reverse TDMA isochronous channel allocations. The up-link timeslots are interleaved between the down link timeslots. The enhancement time slots are allocated dynamically but a fixed allocation was assumed in the simulation. The retransmission time slots can pre-empt the enhancement packets as required.

The retransmission channel availability was varied from 1 to 0.5, the receive buffer was varied from 10 to 200 timeslots in steps of 10 timeslots and the residual packet error rate was measured. The results are plotted in Figure 9-10. Each timeslot period is equivalent to 1 ms. The results show that when the availability of the retransmission buffer is high (above 0.8) the packet error rate is reduced to zero with a receive buffer delay below 200 ms. Again this compares favourably with the interleave buffers used with FEC codes. The retransmission channel is used for 13% of the time. So the enhancement stream cell loss rate is 13%.

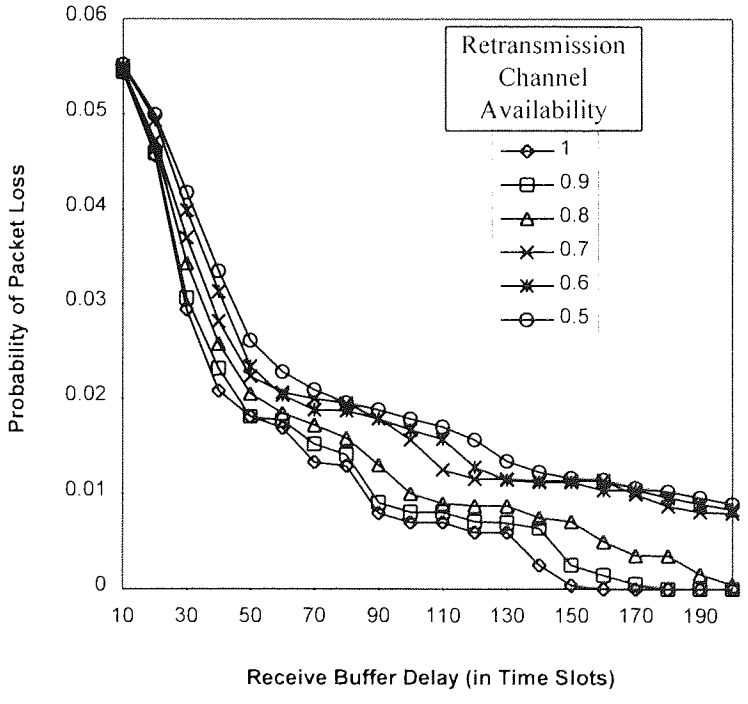


Figure 9-10 Receive Buffer Delay in 0.1 Delay Spread

The base layer salesman sequence delivers very good quality as shown in Figure 9-12 so that enhancement packet loss is difficult to detect, sometimes even with 100% enhancement packet loss

since the PSNR of the two streams is so close. Thus the overall system performance is acceptable from a video quality point of view, but the enhancement stream is not contributing much to video quality.

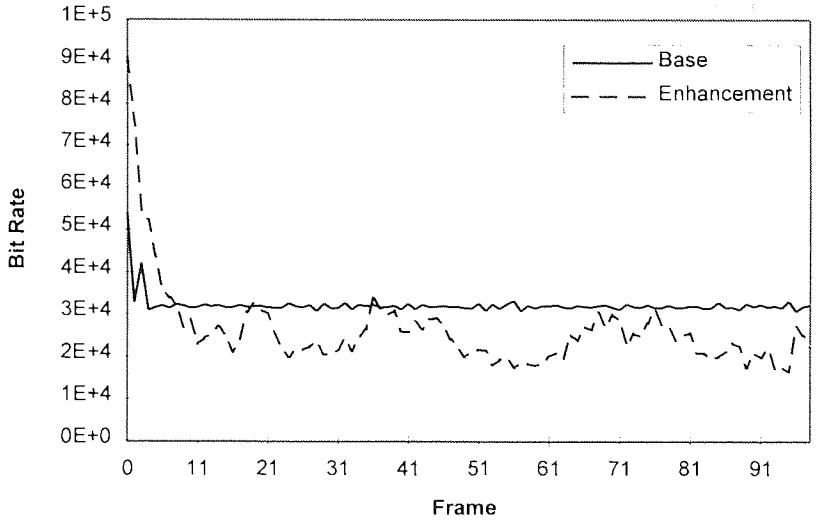


Figure 9-11 Layered Coder Bit Rate Division

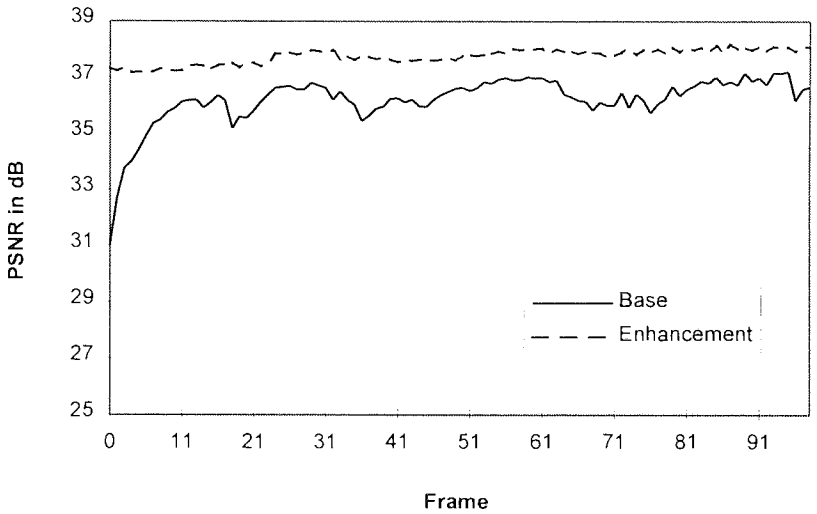


Figure 9-12 Layered Coder - PSNR for Base and Enhancement Layers

9.6 Conclusions

This chapter investigated the suitability of layered coding techniques for robust video transport over wireless networks.

First, the key concepts behind layered video coding were reviewed, using two-layered video codecs proposed in the literature for robust transport over ATM networks, as examples. The following key concepts were identified.

- The unequal error protection nature of the layered coding concept.
- The partitioning of video information into prioritised layers.
- The need to give extra protection to the base layer information.

The error characteristics of ATM and wireless networks were compared. The main difference is that ATM networks can control which packets to discard, whereas packet loss on a wireless network is purely random. This has important repercussions on the adoption of layered coding techniques for wireless networks, because it makes it more difficult to implement a guaranteed channel.

The following advantages were anticipated to accrue from the use of layered coding techniques:-

- A substantial reduction in error control related overheads, leading to substantial bandwidth efficiency gains
- A potentially synergistic merger with ARQ techniques.
- Compatibility with future ATM networks if layered coding is adopted on these networks.
- Relaxation of the stringent network functional requirements, such as handover.

In the previous chapter, ARQ techniques were found to be suitable for video transport over bursty channels. Furthermore, it was noted that ARQ and layered coding seem to be eminently compatible, with ARQ protecting the guaranteed channel and the enhancement layer providing a readily available supply of re-transmission time slots. This led to the proposal of two hybrid ARQ/Layered Coding systems (ARQ/LC).

The first system was proposed to transport H.261 compliant video streams over wireless networks. The system consists of a low complexity transcoder which uses the technique first proposed by Kishino *et al.* [Kishino89] to divide the stream into two layers. The division is engineered so that loss of the low priority packets does not led to excessive picture distortion. Then the enhancement (low priority) packets can be used by the SR ARQ system to re-transmit un-acknowledged base-layer packets. In this way, the base layer bandwidth can increase 'on-demand' when the extra throughput is needed to retransmit packets.

The creation of the enhancement layer introduces an enhancement stream overhead. The enhancement stream packets must be robust to packet errors because they are subject to packet loss. This leads to a further increase in the error control overhead. When the hybrid SR ARQ/LC technique was compared to the single-layer SAW ARQ with Flow control, it was found the latter had the same reconstruction quality of the former but had less overhead and is relatively simple to implement. Thus it is concluded that SAW ARQ is a more attractive option when the MAC multiplex can support it. The SR ARQ/LC technique, on the other hand, can be applied to all MAC multiplexes and does achieve the same reconstruction quality of the single-layer system, which was shown to be very good in the previous chapter.

The second layered coding system that was proposed, allows access to ATM packet video services over a third generation wireless network. In this case the ATM based video codecs are assumed to use the layered coding, such that there need not be an extra coding overhead across the

wired-to-tetherless interface.

A third generation network was proposed to provide the context in which to study the validity of the proposed algorithms. The third generation network is based on the Cellular Packet Switch (Goodman89) and can support multimedia traffic.

A three-layer video codec was then proposed based on two-layer video codecs proposed in the literature [Ghanbari89, Morrison91]. The two enhancement streams can be used to conceal errors when packets are lost on one of the streams.

An error control technique was then proposed to protect the guaranteed video layer, using SR ARQ. The Data Link error control system can pre-empt enhancement time slots to re-transmit base layer packets.

The proposed technique was found to give acceptable results in diverse conditions and it was noted that the layered coding error control mechanism inherently adapts to the error conditions. This good picture quality was achieved without resorting to low-rate codes and hence improves bandwidth efficiency. Diversity systems can also be used to improve the system performance significantly, as shown in the previous chapter.

The simulated SR-ARQ system has been shown to allow a simple, effective and efficient interface to packet video networks. However the scheme requires further testing, especially to optimise the base layer / enhancement layer bit rate allocations and to optimise rate versus distortion.

Chapter 10
Conclusions and Further Work

10.1 Introduction

The contributions in this thesis concern the transmission of video over wireless networks and lie at the region of overlap of three fields; Digital Video Coding, Wireless Networks and Error Control. Background material on these three areas, of relevance to this thesis, is presented in chapter two.

Previous work in the following areas was reviewed in chapter three:-

- Video Coding Standards
- Wireless Networks for Video Transmission
- Error control for video transmission
    - Generic Techniques
    - Techniques for Packet Video
    - Techniques for Digital Terrestrial Television Broadcasting (DTTB)
- Video transmission over wireless networks

The main conclusions of this chapter were:-

⇒ The identification of the hybrid interframe DPCM/DCT algorithm for further study, because of its predominant position in current and proposed video coding standards.

⇒ The identification of second generation digital wireless networks with adequate bandwidth to support video services.

⇒ The identification of common error control techniques proposed for digital video transmission, broadly classified into interframe and intraframe techniques, to counter spatial and temporal error propagation.

⇒ The identification of Layered Coding, proposed to counter cell loss in ATM networks, as a potential candidate for error control on wireless networks.

⇒ The identification of the proposed use of ARQ for H-261 transmission over DECT.

The H-261 video codec was chosen as the base video coding algorithm for the simulation studies in this thesis, mainly because it was the only "conversational", interactive, standardised video algorithm available. Furthermore it is also very similar to the MPEG-1, MPEG-2 and H.263 so that the results obtained for H-261 should also be of relevance to these closely related standards.

Second-Generation digital cordless networks were found to have adequate bandwidth to support video services. However these were designed for voice and data services which have different

properties from video. Digital video required low delay similar to voice and was thought to require low BER similar to data. Hence it was necessary to study the effective and efficient transmission of video over these networks.

Second-Generation mobile networks can support low bit rate video services up to 32 kbit/s, provided time slots can be aggregated. Third-Generation networks are expected to provide support for isochronous services including video but these networks are still in the design phase.

The second generation wireless networks identified capable of supporting video services use TDM/TDMA multiplex with FDD or TDD duplex. Furthermore it was anticipated that a number of third-generation wireless networks would also adopt TDM and TDMA because of their well understood characteristics, their flexibility and good support for variable bit rate services, which make them suitable for multimedia services. Therefore the wireless networks studied employ TDM/TDMA and FDD or TDD.

The need to use bandwidth efficiently and the introduction of medium rate wireless services, such as video, led to the selection of multi-level modulation techniques for further study. The class of low complexity, rotation invariant, differential M-ary PSK techniques with M = 4, 8 and 16 were used in the analysis and simulations. QDPSK (M = 4) is already used in a number of second generation wireless networks as $\pi/4$ shifted QDPSK, and 4-QAM has been proposed for the third generation European UMTS network.

Second-Generation cordless networks have been designed to operate with low complexity receivers, and this approach is adopted here as well: only low complexity countermeasures are considered e.g. antenna diversity.

Two error control techniques were identified for further study, Layered Coding and ARQ. Layered coding is an unequal error protection technique developed for video transmission over ATM networks which are prone to cell loss during congestion periods. It was anticipated that packet loss on radio links would be similar to cell loss on ATM networks in that whole packets would be corrupted and lost. It was thus anticipated that layered coding techniques could be successfully applied to wireless networks as well.

The second error control technique identified was error control by retransmission or ARQ. ARQ is not usually considered suitable for real time services such as voice and video. Recently, however, a prototype system at BTRL was used to transmit H.261 coded video successfully over a DECT testbed using ARQ [Macdonald92].

Furthermore, the two error control techniques were thought to be compatible and it was anticipated that the two techniques could be merged in a synergistic manner.

10.2 Simulation Framework

The study of video transmission techniques requires the proposed techniques to be compared and assessed subjectively because there is no generally accepted method of predicting subjective results

from objective measures. This led to the requirement of simulating the whole video transmission chain; feeding in the video at one end, source coding the frames, transmitting the encoded stream over the simulated radio channel models, decoding the received streams and displaying the recovered video in real time.

A major contribution in this thesis was the development of a simulation framework for the study and analysis of radio transmission over wireless networks. The simulation environment consists of a number of program modules coded in C and C++ which perform the following functions:-

10.2.1 Video source coding and transcoding

The video source coding and transcoding algorithms developed are either H.261 compatible or are derived from this standard. All codecs are based on the motion compensated, hybrid interframe DPCM/DCT algorithm which is the source coding algorithm in H.261, MPEG-1, MPEG-2 and probably H.263.

The simulated transcoders are modified H.261 codecs which can generate and decode both H.261 compliant and non-compliant coded video streams. The transcoders generate video streams which are more robust to transmission errors on wireless networks than normal H.261 streams.

10.2.2 Real-time video display

The QCIF decoders can decode the streams in real time on fast workstations, but the CIF decoders decode frames substantially slower than real time. A program module was thus developed to display video in real-time. The maximum frame rate is 25 fps when displaying sequences from random access memory.

10.2.3 Digital radio modem simulation

The digital radio modems simulated can be modelled as QAM schemes. Both IF and base-band systems can be simulated. The following modulators and demodulators were simulated:-

M-PSK M=2,4,8,16 ...

M-DPSK M=2,4,8,16 ...

$\pi/4$ QDPSK

These modulations can be demodulated coherently and differentially. Perfect carrier recovery and symbol clock recovery are assumed in the simulations, however the program modules allow the study of imperfect carrier and symbol timing recovery.

10.2.4 Radio channel modelling

Four radio channel modules can be used; an AWGN channel, a Rayleigh fading channel, a Rayleigh fading channel with cochannel interference and a frequency selective channel. The latter uses the two-ray model, where each ray is Rayleigh fading. The simulations produce channel error files

which are used to corrupt the video streams.

10.2.5 Countermeasures

The simulations allow the inclusion of antenna diversity techniques by proper manipulation of the channel error files. Using the same techniques, it is possible to simulate other countermeasures such as frequency hopping and variable rate modulation.

10.2.6 Channel coding

Two error control techniques were simulated: Forward Error Correction codes and ARQ protocols. Generic Block FEC codes were simulated and decoded using bounded distance techniques. Two ARQ protocols were simulated; a simple SAW protocol and a real-time Selective Retransmission protocol proposed for wireless environments.

10.2.7 Statistical analysis

The error streams can be fed into a statistical analysis module to estimate the first-order and second-order statistics. The error burst and error gap statistics were of particular importance in this thesis.

10.2.8 Burst error generation

A module was coded to generate burst errors based on the empirical distribution functions derived from the channel error files. The burst error modules use the fade and interfade distributions using a modified Gilbert-Elliot Model [GEM] as described in [Jeruchim93].

The simulation framework is divided in two distinct parts; the video simulation program suite and the radio simulation program suite. Video and radio system simulations are usually performed independently. The two program suites interface via the error vector files produced by the radio link simulations which are then used to corrupt the video streams.

Both sets of simulations were validated extensively using known theoretical results, published empirical results and, in the case of the video codecs, by exchanging video streams with other public domain codecs as these became available.

The waveform simulations were found to be very time consuming. It is preferable to use the waveform simulations to derive bit-level and/or packet-level burst error generators, as these reduce simulation time and allow longer simulations to be executed.

The simulation framework was then used to study and analyse the transmission of hybrid interframe DPCM/DCT coded video over high BER AWGN channels and wireless network links.

10.2.9 Subjective and objective video quality measures

Both subjective and objective results are reported in this thesis. The objective measures used were the mean square error, the peak signal-to-noise ratio, the number of corrupted macroblocks per

frame and the error pipe length histogram. The latter two measures are proposed in this thesis. The former measure measures the spatial diffusion of errors in a frame and the latter, the temporal error propagation of corrupted macroblocks.

10.3 Hybrid Interframe DPCM / DCT in high BER

Initially, the resilience of H.261 video codecs to high bit error rates was tested on a random error binary symmetric channel. Although the wireless network channels considered in this thesis are not well modelled as random error channels, the analysis of the behaviour of the H.261 codec on this channel revealed certain shortcomings in the standard and allowed effective modifications to be developed.

The error resilience of the H.261 codec was investigated by analysing the source coding algorithm, the video multiplex coder and the transmission coder separately.

The components of the source coding algorithm were identified and treated separately. It was argued that a zonal coefficient selection scheme would have been more robust than the threshold selection scheme used in the recommendation. Although the zonal selection scheme is less efficient than threshold selection, it is much less sensitive to transmission errors since the transform coefficient field boundaries are known before transmission.

The lack of an adequate temporal error prorogation countermeasure for the interframe DPCM codec was identified as the weakest part in the source coding algorithm. Conditional replenishment increases the sensitivity to errors by requiring the transmission of block addresses. Motion compensation has been shown to contribute to both temporal and spatial error propagation.

The video multiplex coder was identified to be a major cause of concern. The frequent use of differential encoding followed by variable length coding leaves the video multiplex stream very vulnerable to transmission errors. Absolute encoding is used to facilitate mid-stream synchronisation but it is not very effective at high BER. The GBSC re-synchronisation mechanism was found to be inadequate for QCIF sequences.

The performance of the transmission codec was found to be unsatisfactory at a BER above 10^{-3} as typically encountered on wireless networks. The FEC framing mechanism caused concern because loss of synchronisation leads to the loss of the video stream for several hundred milliseconds.

Techniques for implementing robust H.261 decoders were discussed. These techniques were used to implement the IHC decoder. The IHD uses the BCH code in the H.261 standard for forward error correction or for error detection. The limited VMC code violation detection capability is also used to detect decoding errors. When an error is detected, the BCH codeword is discarded and the IHD seeks the nearest PSC or GBSC re-synchronisation point. The IHD conceals the resulting errors by displaying previous-frame data. In intrinsic mode, the IHD also displays any data decoded prior to the error detection event. In extrinsic mode, the IHD replaces all the data in the GOB in which the error detection event occurs.

The robust decoder was tested in high BER conditions. It has been shown that a properly designed H.261 decoder can survive around one bit in error per transmitted frame, without making use of the general BCH code or error concealment, but protecting the PH and GOBH header fields. The decoder traps field errors and skips to the next GOB when errors are detected. However the error detection is very unreliable since the detected error events are very few, due to the complete nature of the variable length codes. Occasionally, highly visible errors occur but at this error rate these do not propagate catastrophically for the tested sequences and are kept in check by the forced updating of the unaltered algorithm. Although some errors are highly objectionable, the decoded image is still perfectly intelligible. When the error rate is increased to 10 bit errors per transmitted frame, the visible errors coalesce rapidly and intelligibility is lost.

When the BCH code is used to detect errors, error detection is highly reliable. The decoder still attempts to decode the PH and GOBH fields in corrupted blocks, but it will then discard the decoded macroblock data and skip to the next non-corrupted GOB.

When the BER is 10^{-4} the FEC code can correct most of the error patterns and the three sequences were decoded without errors. When the BER is 10^{-3} the FEC decoder commits a significant number of decoding errors and there is visible picture degradation. The sequences can be decoded usefully, however, and there is no catastrophic picture degradation.

It was argued that the error resilience of H.261 can be improved by employing H.261-compliant, H.261-compatible and non-H.261-compatible techniques. H.261-compliant techniques follow the standard to the letter. Compatible techniques can be decoded using H.261 compliant decoders but require special decoders, with additional intelligence, to capitalise on the improved decoding robustness. Non-compatible techniques can be implemented using low-complexity, low-delay transcoders to interface the robust codecs to H.261 compliant codecs.

Five H.261 compliant or compatible techniques were proposed to improve the resilience of H.261.

(i) Header protection.

The header protection technique uses the PSPARE and GSPARE header fields which are ignored by normal H.261 decoders. It is proposed to use enough SPARE fields to implement a FEC code 'signature'. Then probable corrupted picture and GOB headers can be located by correlation and positive identification achieved by proper decoding of the FEC code 'signature'. This technique is H.261 compatible but not H.261 compliant since the use of the SPARE fields in prohibited in H.261.

(ii) Fixed length DC intraframes.

It was noted that corruption of the first intracoded frame led to very visible and persistent image degradation. The fixed length DC intraframe was proposed to implement a robust intraframe mode. The technique codes the first frame using fixed-length eight-bit DC coefficients only, such that all the fields in the frame are fixed length and synchronisation cannot be lost, even when the DC

coefficients are corrupted. The error resilience is also improved because the frame format uses the lowest number of bits possible and hence stands a better chance of avoiding corruption. Results confirm the robustness of the technique. The initial reduction in picture quality is generally recovered in the subsequent frame. The combined resilience of the first two frames is much improved with the proposed technique.

(iii) Cyclic Refresh using Periodic Intracoding.

Error recovery using periodic intracoding was shown to be inadequate for fixed bit rate video codecs because the intraframes consume three to five times the bit rate of interframes and lead to very visible temporal resolution reduction, especially at low frame rates. Also the intracoded frames are not very efficient when compared to MPEG-1, say.

(iv) Cyclic Refresh using Forced Updating.

Forced updating over and above the H.261 recommendation was found to provide very good error recovery capabilities. The good temporal error recovery is also enjoyed by normal H.261 decoders. Force updating three macroblocks per GOB per frame was found to be a good compromise between picture quality reduction and error recovery capabilities. At this rate, the picture usually recovers completely from errors within eleven frames.

(v) Error Recovery using Fast Update Request.

H.261 incorporates a fast update request signal which can be asserted by third parties wishing to join a multi-party conversation. This signal can be used by a decoder to force the far end encoder to code the next frame in intramode to recover from an error. It was found that this technique is very sensitive to BER in the region of 10^{-4} to 10^{-3} on a random error channel. It is anticipated, however, that the technique would be more successful on a burst error channel at these error rates.

It was argued that the best way to access H.261 video services over a wireless network may be to use a robust non-H.261 compliant mobile video codec and to use transcoders to interface the video stream with H.261 codecs on the fixed network. The two main issues are, then, transcoder complexity and transcoding delay. Complexity should be low to reduce cost. The transcoding delay should be low to prevent excessive communications delays which are detrimental to speech intelligibility.

Four low-complexity, low-delay techniques were proposed.

(i) Improved FEC codes

The optimal solution on an AWGN channel is to replace the H.261 BCH code with a more robust FEC code. However this solution may not be optimal if the channel has memory and the system complexity and decoding delay are an important issue.

(ii) Structured Packing

Structured packing is another low-delay, low-complexity option. The technique adds absolute pointers to video data packets to allow re-synchronisation at points other than the picture and GOB headers, and thus allows better concealment strategies to be adopted. The first macroblock header in

each packet has to be identified by a pointer. The differentially coded fields in the macroblock, i.e. the macroblock address and the motion vector fields, must be augmented with absolute information.

(iii) More frequent QCIF re-synchronisation points

For QCIF sequences the GOB re-synchronisation points do not offer adequate spatial error protection since the loss of one GOB signifies the loss of 33% of one frame. Therefore it is proposed to include a new synchronisation point at the beginning of macroblock rows.

(iv) Fixed length Macroblocks.

The fourth proposal is to re-code the DCT coefficients using a number of fixed length block sizes to allow synchronisation at macroblock boundaries. This is the most complex proposal and may result in some coding inefficiencies.

10.4 Hybrid Interframe DPCM / DCT on Rayleigh fading channels

The robust video codecs and transcoders developed on the AWGN channels were then tested on the narrowband, Rayleigh fading channel using the simulation infrastructure developed in the previous chapters.

A wireless network was proposed similar to several cordless networks introduced in chapters two and three. The wireless network link was then simulated at the physical, MAC and data link layers to study the transmission of unprotected and robust H.261 coded video streams at a BER of 10^{-3}, which is typical for current, voice-oriented wireless networks.

At the physical layer rotation invariant, differential M-ary Phase Shift Keying modems have been simulated to transmit video information over the radio link at two simulated vehicle speeds - 4 km/hr and 64 km/hr. The transmission baud rate is 192 kbaud, and it is assumed that the propagation delay spread is low enough for the channel to be modelled as Rayleigh fading. Furthermore CCI is ignored, so that the results obtained are for a non-CCI environment as in [Stedman92, Hanzo94, Streit94]. The simulated carrier frequency was 1.9 GHz.

At the MAC level the simulated wireless link uses TDM on the down-link, circuit switched TDMA on the up-link and staggered FDD. The frame rate is 10 ms, and each frame contains 8 time-slots. Each time slot consists of 240 symbols, which are equivalent to 480 bits with QDPSK and 860 bits with 16-DPSK. Alternatively the 16-DPSK modem can support 16 time-slots of 480 bits each.

Two data link layer packet formats have been proposed. The first is intended for error detection oriented protocols, and the second to be used with BCH(63,36,5) FEC codes. The error detection packet format consists of a data payload of 320 bits and uses the remaining bits for CRC and the header field. The FEC packet format packs five BCH(63,36,5) codewords into a 315 bit payload field. The 180 information bits consist of 160 video bits and a 20 bit CRC. The codewords in one time slot can be interleaved to disperse an error burst over all five codewords to make best use of their random error correction capability. It should be noted that the above packet structures are supported by a transcoder which uses the raw bearer channels of the MAC layer.

Initial packet error statistics showed that for the above TDM MAC format, the packet errors are more bursty on the 4 km/hr channel than on the 64 km/hr channel. The average number of consecutive packets in error is low for both speeds. The lowest is 1.035 for QDPSK at 64 km/hr and the highest 1.395 for 16-DPSK at 4 km/hr. Multiple packet bursts are common, however, with 5-packet error bursts detected at 4 km/hr for 16-DPSK. The packet error bursts have been found to occur more or less randomly at the higher speed, but there is some evidence of a non-Poisson arrival distribution at the lower speeds.

The transmission of video over the radio link was then studied by simulation. The QCIF, colour Miss America sequence was used to generate two H.261 video sequences at a video frame rate of 10 fps and a coded rate of 32 kbit/s. The first sequence uses intrinsic forced updating at one macroblock per frame. The second 'robust' sequence uses forced updating at the increased rate of three macroblocks per GOB per frame and the fixed length DC intracoded frame described in chapter five.

These two sequences were then corrupted, using four error files generated by simulating QDPSK and 16-DPSK modems at 4 and 64 km/hr, and decoded using the robust IHC decoder under three conditions

(i) without using error control,

(ii) using error detection and discard,

(iii) and using a transcoder with more powerful FEC codes.

The first important conclusion was that the sequences can be decoded *usefully,* without catastrophic quality deterioration, at a bit error rate of 10^{-3} using the robust IHC decoder without the use of error control. The SUR scores for the unprotected sequences, averaged over the whole 30 second error file, were 3.7 and 3.3 for QDPSK at 4 km/hr and 64 km/hr, and 3.6 and 3 for 16-DPSK at 4 km/hr and 64 km/hr. These results are in sharp contrast with the results obtained on the random error channel, were the BER of 10^{-3} caused the unprotected streams to fail catastrophically. This can be explained by noting that the errors on the radio link occur in bursts such that substantially fewer frames are hit by errors than on the random error channel at the same BER.

When the robust streams were decoded using the robust IHC decoder with error detection and discarding, the SUR scores, *for the worse six-second stream segment*, rose to 4.5 and 4 for QDPSK at 4 km/hr and 64 km/hr, and to 4.2 and 4 for 16-DPSK at the same speeds. It was noted that the DC intracoded frame was very effective in avoiding the highly visible and objectionable errors in the first frame.

When the robust streams were decoded using the low rate FEC option, the SUR scores, *for the worse six-second stream segment,* rose to 5 for all decoded streams except the 16-DPSK 4 km/hr stream which scored 4.5. Although the streams were decoded practically without any visible errors it was noted that errors still defeated the low rate FEC codes, with 4, 2, 4 and 1 frames hit by error bursts in worst error streams Ab, Bd, Cb and Dd respectively.

It has been shown that the reconstruction quality deteriorates with increasing vehicle speed, unless FEC is used, in which case the residual BER decreases with increase in speed because the error bursts are less intense. This result holds for both QDPSK and 16-DPSK.

The classical technique of Interleaved FEC codes to combat burst errors was also studied. The minimum interleave depth required to guarantee no residual errors using the BCH(63,56,5) codewords, was found to be five time-slots. This is equivalent to a total delay of 100 ms which is not considered excessive. This delay may, however, need to be increased at lower speeds. Other techniques, such as frequency hopping, can be used to break up the error bursts at low speed. Further work is needed to study the performance of these techniques.

To summarise, it has been shown that H.261 coded video can be transmitted over the proposed network at an average BER of 10^{-3} by using the robust decoders developed for the AWGN channel and modified slightly to use more powerful, low rate BCH FEC codes. Furthermore, it is possible to compromise between bandwidth efficiency and video quality on the simulated narrowband Rayleigh fading channel such that :-

- An unprotected H.261 stream without error control overhead can be transmitted and recovered with useful reconstruction albeit with visible impairments every few tens of seconds.
- A robust video stream incorporating forced updating and the fixed-length DC intraframe can be transmitted and recovered using the robust decoders with significantly less noticeable picture impairments occurring at the same frequency. The error control overhead is less than 7% because only high-rate error detection codes are used.
- The same robust video stream can be transmitted with hardly any visible impairments over a 30 second period, by using the low rate FEC BCH(63,36,5) code with 50% error control overhead.
- If ten consecutive FEC coded time slots are interleaved over five MAC frames (two time slots per frame), the video streams are decoded without any residual error. The coding efficiency is unchanged but an additional delay, equivalent to ten MAC frame periods, is introduced in the simplex transmission chain.

10.5 Error Control by Retransmission for Video Transmission

In the study of video transmission over the Rayleigh fading channel, it was found that the severe error bursts defeat low-rate FEC codes regularly. The residual channel errors can be removed by deploying interleaving but this introduces transmission delays of the order of 100 ms per radio link. This led to an investigation of ARQ techniques as a class of error control techniques which allow better bandwidth efficiency with the same order of transmission delay.

The main conclusion of an analysis, conducted to study the suitability of retransmission techniques for video transmission over wireless networks, was that ARQ error control techniques are suitable for the proposed application, provided they are used across the radio link and not end-to-end across the fixed WAN. This is possible because the round trip delay on wireless networks is very low, and one video frame is transmitted using tens of radio link packets. Then, retransmitted packets can still arrive on time to be displayed in the correct frame without causing picture degradation.

It was also argued that it is easier to apply retransmission techniques on the up link because the encoder smoothing buffer can be closely coupled with the retransmission buffer. The video encoder buffer control can then vary the encoder parameters to optimise the video quality depending on the instantaneous state of both buffers. This close coupling is not possible on the down link.

ARQ for video transmission was first proposed by MacDonald *et. al.* at BTRL [MacDonald92]. Their prototype H.261 video transmission system for DECT demonstrated the potential of ARQ but revealed a number of problems with practical implementation in a real-world environment. An analysis of the BTRL prototype system revealed that the main problem in applying ARQ to real-time video transport was the difficulty in controlling the far-end (across the WAN) FBR video encoder bit rate to prevent monotonically increasing retransmission delay or retransmission buffer overflow.

Two main classes of techniques were proposed to solve this problem, namely flow control techniques and bandwidth-on-demand techniques. It was argued that the techniques are best implemented using a transcoder at the interface between the fixed and the wireless networks.

Four **Flow Control** mechanisms were proposed, based on low complexity transcoders to transmit H.261 coded data over a wireless network in a transparent manner. The four mechanisms monitor the occupancy of the retransmission buffer and when it approaches overflow the transcoder discards video information in an unobtrusive manner. The transcoder can discard whole frames, whole GOBs, whole macroblocks or DCT coefficients. Frame and GOB discard strategies are the least complex to implement but generate the worse image degradations. Macroblock discard results in less visible impairments, especially if concealment techniques are used to minimise the visual impact of the discarded macroblocks. MB discard, however, is the most complex technique to implement because it requires the macroblock headers of un-discarded blocks to be modified. DCT coefficient discard is less complex to implement than MB discard and results in the lowest picture quality degradation.

Tests showed that discarding frames had the desired effect of reducing the bit rate but the recovery time was excessive unless the special forced updating mode, described in the previous chapters, was used by the far-end encoder. Flow control by discarding DCT coefficients was found to reduce the bit rate into the retransmission buffer whilst causing minimal visible degradations, provided no coefficients from the first intracoded frame are discarded. DCT coefficient discarding does not require complex transcoding because it only involves segmenting the variable length bit stream into

codewords and discarding portions of the stream. It can also be used to reduce the bit rate permanently, to match disparate bit rates, by continuously discarding coefficients which leads to a stable loss of picture quality.

The DCT coefficient flow control was then combined with a simple SAW ARQ system implemented on a TDM/TDMA/FDD wireless network. The system gave good results on Rayleigh fading channels, including those suffering from strong CCI, provided second order diversity was employed in the latter case. The retransmission buffer delay jitter was bounded to less than 100ms and the picture quality degradation was not easy to detect even on a 32 kbit/s coded video stream.

Bandwidth-on-demand techniques were shown to be able to deliver an unaltered picture quality but required more support from the wireless network infrastructure, particularly the MAC layer. One physical layer technique and four MAC-layer techniques were proposed:-

- Variable rate Modulation (Physical Layer)
- Variable rate MAC
- Shared feedback channels
- Up link packet pre-emption
- Down Link packet pre-emption.

A bandwidth-on-demand system, based on a shared feedback channel, was then studied on a TDM/TDMA/FDD 384 kbit/s digital wireless network. A QDPSK modem with second order selection diversity was used in the physical layer. The channel was Rayleigh fading with average E_b/N_o of 15 dB at a simulated mobile speed of 4 km/hr.

A real-time ARQ system was then proposed, based on SR ARQ but which discards overly delayed packets. The retransmitted packets use a dynamically allocated retransmission channel shared with other video links. It was demonstrated that the technique can transport video error-free, provided the receive buffer timeout is sufficiently high to recuperate delayed packets. The required receive buffer delay was found to be strongly dependent on the availability of the shared retransmission channel. When the availability was set to 1, the time out period and hence the system delay was 110 ms, which compares favourably with the delay introduced by interleave buffers required by FEC techniques.

Thus it has been shown that error control by retransmission allows video services to be delivered effectively and efficiently over wireless networks and that these techniques should be considered as viable alternatives to FEC techniques for the provision of real-time video services on wireless networks, since their adaptive nature allows very efficient use of the scarce radio spectrum.

Finally, the concept of a Trace-Back Decoder was introduced. The TBD can decode retransmitted packets which are delayed and out of sequence. A proportion of errored packets are retransmitted in time to be decoded and displayed in the appropriate frame. Retransmitted packets

which are delayed beyond their scheduled display time are not discarded but are decoded and replace the concealed macroblocks in the reference frame. In this way errored macroblocks do not propagate for more than one or two frames at most. Thus there is no system delay penalty due to the retransmission scheme.

10.6 Layered Coding

Layered video coding has been proposed recently to mitigate the effect of cell loss on packet video transmission over ATM networks. The transmission errors on wireless networks result in packet loss as in ATM networks. It was thus natural to investigate the suitability of layered coding techniques for robust video transport over wireless networks.

Layered coding is an unequal error protection scheme. The video data is partitioned into prioritised layers. The base layer is essential for the image sequence reconstruction and is given extra protection. Thus the layered coding can also be regarded as a joint source-channel coding scheme. The following advantages were expected from the use of layered coding techniques:-

- A substantial reduction in error control related overheads, leading to substantial bandwidth efficiency gains
- Compatibility with future ATM networks, if layered coding is adopted on these networks.
- Relaxation of the overheads imposed by the stringent, network, functional requirements, such as handover.

A comparison of ATM and wireless networks revealed that packet loss in ATM networks is largely limited to low priority packets, whereas packet loss on a wireless network is purely random. It is thus more difficult to implement a guaranteed channel on wireless networks.

In the previous chapter, ARQ techniques were found to be suitable for video transport over bursty channels, especially if some form of bandwidth-on-demand is available. It was then conjectured that ARQ and layered coding could be merged synergistically, with ARQ protecting the guaranteed channel and the enhancement layer providing a readily available supply of re-transmission time slots. This led to the proposal of two hybrid ARQ/Layered Coding systems (ARQ/LC).

The first system was proposed to transport H.261 compliant video streams. The system consists of a low complexity transcoder which uses the technique first proposed by Kishino *et. al.* [Kishino89] to divide the stream into two layers such that loss of the low priority packets does not lead to excessive picture distortion. Then the enhancement (low priority) packets can be used by the SR ARQ system to re-transmit un-acknowledged base-layer packets. In this way, the base layer bandwidth can increase 'on-demand' when the extra throughput is needed to retransmit packets.

It was found that this hybrid SR ARQ/LC technique resulted in the same reconstruction quality as that of the single-layer SAW ARQ scheme with flow control studied in the previous chapter, because both discard data as part of the error control strategy. The layered coding technique, however, requires more overhead since the creation of the enhancement layer, which must be robust to packet loss,

introduces an enhancement stream overhead. It is also more complex to implement. Thus it is concluded that SAW ARQ is a more attractive option when the MAC multiplex can support it. The SR ARQ/LC technique, on the other hand, can be applied to all MAC multiplexes and does achieve the same reconstruction quality of the single-layer system, which was shown to be very good in the previous chapter.

The second Layered Coding scheme proposed, facilitates access to ATM packet video services over a third generation wireless network. In this case the ATM based video codecs are assumed to use layered coding, such that there is no extra overhead in the transcoder as in the previous case.

A third generation network was proposed to provide the context against which to study the validity of the proposed algorithms. The third generation network is based on the Cellular Packet Switch [Goodman89] and can support multimedia traffic.

A three-layer video codec was then proposed based on two-layer video codecs proposed in the literature [Ghanbari89, Morrison91]. The two enhancement streams can be used to conceal errors when packets are lost on one of the streams.

An error control technique was then proposed to protect the guaranteed video layer, using SR ARQ. The Data Link error control system can pre-empt enhancement time slots to re-transmit base layer packets.

The proposed technique was found to give acceptable results in diverse conditions and it was noted that the layered coding error control mechanism inherently adapts to the error condition. This good picture quality was achieved without resorting to low rate codes, and hence uses bandwidth efficiently. Diversity systems can also be used to improve the system performance significantly, as shown in the previous chapter.

The simulated SR-ARQ layered coding system has been shown to allow a simple, effective and efficient interface to packet video networks. However the scheme requires further testing, especially to optimise the base layer / enhancement layer bit rate allocations optimise rate versus distortion.

10.7 Further Work

The work in this thesis is amongst the first to address the problem of video transmission over wireless networks. The options available for study were numerous because of this, but only a few of the possibilities could be tackled.

In this thesis it has been shown that the motion compensated, hybrid interframe DPCM/DCT algorithm can be transmitted successfully over current and future wireless networks. The work in this thesis has also shown that the use of retransmission can achieve better bandwidth efficiency than using interleaved, low-rate FEC codes and without a delay penalty with respect to the latter.

It is thus recommended to study the application of these techniques to similar coding schemes. In particular MPEG-1 and H.263 have been targeted as candidate techniques for further study.

The proposed ARQ scheme is source-coder independent and should be applicable to other source coding algorithms to be proposed during the MPEG-4 competitive design phase [Reader95]. It is recommended that ARQ techniques should be studied for new source coding algorithms as well, and to compare the improvement in efficiency (if any) with respect to conventional interleaved FEC techniques.

Only low complexity BCH block codes were studied in this thesis. To allow a more balanced comparison between FEC and ARQ techniques, more work is required on the use of RS codes which are amongst the best known burst error correction block codes. Rate Punctured Codes and soft decision decoding techniques for block codes should also allow efficiency improvements over the reported FEC results.

Transcoders have been used liberally in this thesis. It has been shown that transcoders can be designed to be both simple to implement and have a low transcoding delay. It is thus recommended that this technique be adopted for wireless networks to allow fixed network video services to be accessed transparently over wireless networks. This needs to be investigated further and each wireless network may require a different encoder scheme to be economically feasible.

The use of layered coding with ARQ has been studied and shown to have promise. Further work is required on the layered coding aspect to devise layered source coding techniques which are as efficient as single layer techniques (if this is at all possible).

The results obtained in this thesis suggest that third-generation, multimedia-oriented wireless networks should have support for real-time retransmission schemes. The inclusion of these techniques would result in the efficient use of bandwidth because the ARQ schemes have been shown to be very bandwidth efficient. A possible third-generation MAC scheme has been proposed in this thesis but more work is needed to design MAC protocols which can support the error control schemes proposed in this thesis.

Parts of the simulation framework can be implemented in hardware. The advent of programmable hardware, in the form of programmable gate arrays, can be exploited in the simulation environment. A programmable hardware architecture can be developed to support generic video coding algorithms. This also opens up the possibility of designing systems in which the video codec is not fixed but can be downloaded as part of the videophone or videoconference initialisation phase. This allows different coding algorithms to be used on different networks and network combinations, to optimise video quality and/or bandwidth efficiency.

The waveform simulations are very time consuming. It is recommended that a program is undertaken to generate bit and packet level error patterns for current and proposed wireless networks. Hardware simulators can then be used, *e.g.* using programmable hardware, to simulate the whole video transmission chain in real-time.

Time constraints precluded the testing of all the proposed schemes on other channels of practical interest, *e.g.* channels employing spread-spectrum and CDMA. These need to be studied. Also, only low-complexity countermeasures were of interest in the thesis and a number of more complex techniques need to be tested. These include equalisation and variable rate modulation, to name but a few.

Time and budgetary constraints limited subjective testing to informal sessions. However it is understood that formal tests will also have to be conducted to provide more unbiased evidence on which to assess the techniques proposed in the thesis.

Finally, more work is needed to develop the Trace Back Decoder concept. Initial results were encouraging but the implementation details still have to be worked out in detail.

Publications

S. Zammit and G.F. Carpenter, "Packet Video Coding for Wireless Access Networks", IEEE Visicom93, 5th. Intl. Workshop on Packet Video, Berlin, Mar. 22-23 1993, G2.1-G2.4.

S. Zammit and G.F. Carpenter, "Low-complexity, low-delay H.261/MPEG-1 Transcoder ", accepted for publication in the Picture Coding Symposium 1996.

In Preparation

S. Zammit and G.F. Carpenter, "Digital video transmission over Rayleigh fading channels"

S. Zammit and G.F. Carpenter, "ARQ techniques for digital video transmission over fading channels"

S. Zammit and G.F. Carpenter, "Layered video coding for video transmission over fading channels"

S. Zammit and G.F. Carpenter, "The Trace-Back-N video codec"

References

[Acampora87a] A.S. Acampora and J.H. Winters, "A Wireless Network for Wide-Band Indoor Communications", IEEE Journal on Selected Areas in Commun., Vol. 5, No. 5, Jun. 1987, pp. 796-805.

[Acampora87b] A.S. Acampora and J.H. Winters, "System Applications for Wireless Indoor Communications", IEEE Commun. Mag., Vol. 25, No. 8, Aug. 1987, pp. 11-20.

[Adachi92] F. Adachi and M. Sawahashi, "Performance analysis of various 16 level modulation schemes for Rayleigh fading channels", IEE Electron. Lett., Vol. 28, No. 17, Aug. 1992, pp. 1579-1581.

[Adams92] D.E. Adams and C.R. Frank, "WARC Embraces PCN", IEEE Commun. Mag., Vol. 30, No. 6, Jun. 1992, pp 44-47.

[Ahl92] K.A. Ahl, "Radio for the Local Loop", Communications International, Nov. 1992, pp. 59-64.

[Ahmed74] N. Ahmed, T. Nataraja and K. Rao, "Discrete Cosine Transform", IEEE Trans. Commun., Vol. 23, No. 1, Jan. 1974, pp. 90-93.

[Anastassiou94] D. Anastassiou, "Digital Television", Proc. IEEE, Vol. 82, No. 4, Apr. 1994, pp. 510-519.

[Arai88] Y. Arai, T. Agui and M. Nakajima, "A fast DCT-SQ scheme for images", Trans. IEICE, Vol. 71, 1988, pp. 1095-1097.

[Argenti93] F. Argenti, G. Benelli and C. Kutufa, "Transmission of Wavelet Transform Coded TV Images on a Noisy Channel", IEEE International Conference on Communications 93, Geneva, 23-26 May 1993, Ch. 367, pp. 386-390.

[Arguello71] R.J. Arguello, H.R. Sellner and A.A. Stuller, "The effect of channel errors in the differential pulse-code-modulation transmission of sampled imagery", IEEE Trans. Commun., Vol. 19, 1971, pp. 926-933.

[Armbruster90] H. Armbruster and G. Arndt, "Broadband Communication and its realisation with B-ISDN", IEEE Communications Mag., Vol. 28, No. 4, Apr. 1990, pp. 66-71.

[Baier94] 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 in Commun., Vol. 12, No. 4, May 1994, pp. 733-743.

[Barnsley88] M.F. Barnsley and A.D. Sloan, "A better way to compress images", Byte Magazine, Jan. 1988, pp. 215-223.

[Barry91] J.R. Barry, J.M. Kahn, E.A. Lee and D.G. Messerschmitt, "High-Speed Nondirective Optical Communication for Wireless Networks", IEEE Network Mag., Vol. 5, No. 6, Nov. 1991, pp. 44-54.

[Belzer94] B. Belzer, J. Liao and J.D. Villasenor, "Adaptive Video Coding for Mobile Wireless Networks", IEEE ICIP-94, Austin, Texas, 13-16 Nov. 1994, Ch. 213, pp. 972-976.

[Bergman] M.L. Bergman and P.G. Farrell, "A Comparison of Block Code Performance for Mobile Radio Channels", IEE Colloquium on University Research in Mobile Radio, IEE Conference Digest 1990/143, 2 Nov. 1990, pp. 8/1-8/6.

[Bito93] J.Bito, "A simple model for the loss process in the cell stream of variable bit rate video sources", IEEE Visicom93, 5th. Intl. Workshop on Packet Video, Berlin, Mar. 22-23 1993 C4.1-C4.5.

[Black92] S. Black, "An Overview of Standards work for Radio LANs", IEE Colloquium on Radio LANs, IEE Digest 1992/104, May 1992, pp. 1-8.

[Brewster91] R.L. Brewster and R.S. Jalal, "Transmission of Graphic Image Data to Mobile Terminals", IEE Colloquium on GSM and PCN Enhanced Mobile Services, 30th Jan. 1991, Digest No. 1991/023, pp. 9/1-9/4.

[Brofferio89] S.C. Brofferio, "Object-background image model for predictive video coding", IEEE Trans. Commun, Vol. 37, No. 12, Dec. 1989, pp. 1391-1394.

[BT] Geoff Morrison, "Researching Video Coding Techniques", Private Communication.

[Burt81] P.J. Burt and E.H. Adelson, "The Laplacian pyramid as a compact image code", IEEE Trans. Commun., Vol. 31, No. 12, Dec. 1981, pp. 532-540.

[Caffario90] C. Caffario, F. Rocca and S. Tubaro, "Motion compensated image interpolation", IEEE Trans. Commun., Vol. 38, No. 2, Feb. 1990, pp. 215-222.

[Callendar94] M.H. Callendar, "Future Public Land Mobile Telecommunication Systems", IEEE Personal Communications, Vol. 1, No. 3, Fourth Quarter 1994, pp. 18-22.

[Candy71] J.C. Candy, M.A. Franke, B.G. Haskell and F.W. Mounts, "Transmitting television as clusters of frame-to-frame differences", Bell System Technical Journal, Vol. 50, No. 6, Jul. 1971, pp. 1889-1917.

[Capellini85] V. Capellini, "Data compression and error control techniques with applications", Academic Press, London, 1985.

[Cassereau89] P.M. Cassereau , D.H. Staelin and G.De Jager, "Encoding of images based on a lapped orthogonal transform", IEEE Trans. Commun. Vol. 37, No. 2, Feb. 1989, pp. 189-193.

[CCIR601] CCIR Recommendation 601, "Encoding Parameters of Digital Television for Studios".

[CCITT-H.120] CCITT Recommendation H.120, "Codecs for videoconferencing using primary digital group transmission", Blue Book, Vol. III, Fascicle III.6.

[CCITT-H.261] CCITT Recommendation H.261, "Video Codec for Audiovisual Services at p x 64 kbit/s", Geneva 1990.

[CCITT-H.320] CCITT Recommendation H.320, "Narrow-band visual telephone systems and terminal equipment", Geneva 1990.

[Chang90] L.F. Chang and P.T. Porter, "Data Services in a TDMA Digital Portable Radio System", IEEE Globecom 1990, San Diego, 2-5 Dec. 1990, pp. 480-484.

[Chang91] L.F. Chang, "Throughput estimation of ARQ protocols for a Rayleigh fading channel using fade and interfade duration statistics", IEEE Trans. on Vehicular Tech., Vol. 40, No. 1, Feb. 1991, pp. 223-229.

[Chennakeshu93] S. Chennakeshu and G.J. Saulnier, "Differential Detection of PI/4-Shifted-DQPSK for Digital Cellular Radio", IEEE Trans. on Vehicular Tech., Vol. 42, No. 1, Feb. 1993, pp. 46-57.

[Chia92] S. Chia, "The Universal Mobile Telecommunication System", IEEE Communications Mag., Vol. 30, No. 12, Dec. 1992, pp. 54-63.

[Chia93] T. Chia, P.J. Coventry, D.J. Parish and J.W.R. Griffiths, "Motion JPEG on a Network and the Treatment of Video Cell Losses", IEEE Visicom 93, Fifth Intl. Workshop on Packet Video, Berlin, Mar. 22-23, 1993, pp. D1.1- D1.6.

[Chin89] H.S. Chin, J.W. Goodge and D.J. Parish, "Statistics of viewphone type pictures", IEEE Journal on Selected Areas in Commun., Vol. 7, No. 5, Jun. 1989, pp. 826-832.

[Chu87] T.S. Chu. and M.J. Gans, " High Speed Infrared Local Wireless Communication", IEEE Communications Mag., Vol. 25, No. 8, Aug. 1987, pp. 4-10.

[Chuang87] J.C.I. Chunag, "The Effects of Time Delay Spread on Portable Radio Communications Channels with Digital Modulation", IEEE Journal on Selected Areas in Commun., Vol. 5, No. 5, Jun. 1987, pp. 879-889.

[Chuang89] J.C.I Chuang, "The Effects of Delay Spread on 2-PSK, 4-PSK, 8-PSK and 16-QAM in a Portable Environment", IEEE Trans. on Vehicular Technology, Vol. 32, No. 2, May 1989, pp. 43-45.

[Chuang90] J.C.I. Chuang, "Comparison of two ARQ protocols in a Rayleigh fading channel", IEEE Trans. on Vehicular Tech., Vol. 39, No. 4, Nov. 90, pp. 367-373.

[Comroe84] R.A. Comroe and D.J. Costello Jr., "ARQ schemes for data transmission in mobile radio systems", IEEE Trans. on Vehicular Tech., Vol. 33, No. 3, Aug. 84, pp. 88-97.

[Connor73] D.J. Connor, "Techniques for reducing the visibility of transmission errors in digitally encoded video signals", IEEE Trans. Commun., Vol. 21, No. 3, June 1973, pp. 695-706.

[Cox85] D.C. Cox, "Universal Portable Radio Communications", IEEE. Trans. on Vehicular Tech., Vol. 34, No. 3, Aug. 1985, pp. 117-121.

[Cox92] D.C. Cox, "Wireless Network Access for Personal Communications", IEEE Communications Mag., Vol. 30, No. 12, Dec. 1992, pp. 96-115.

[Cox95] D.C. Cox, "Wireless Personal Communications: What Is It?", IEEE Personal Communications, Vol. 2, No. 2, Apr. 1995, pp. 20-35.

[Dent93] P. Dent, G.E. Bottomley and T. Croft, "Jakes Fading Model Revisited", IEE Electron. Lett., Vol. 29, No. 13, 1993, pp. 1162-1163.

[Devasirvatham87] D.M.J. Devasirvatham, "Multipath Time Delay Spread in the Digital Portable Radio Environment", IEEE Communications Mag., Vol. 25, No. 6, Jun. 1987, pp. 13-22.

[Dunlop95a] J. Dunlop, J. Irvine and P. Cosimini, "Link Adaptation in ATDMA", IEE Electron. Lett., Vol. 31, No. 7, Mar. 1995, pp. 528-529.

[Dunlop95b] J. Dunlop, J. Irvine, D. Roberston and P. Cosimini, "Performance of a Statistically Multiplexed Access Mechanism for a TDMA Radio Interface", IEEE Personal Communications, Vol. 2, No. 3, Jun. 1995, pp. 56-64.

[Farvardin87] N. Farvardin and V. Vaishampayan, "Optimal quantizer design for noisy channel: An approach to combined source-channel coding", IEEE Trans. Inform. Theory, Vol. 33, No. 6 , Nov. 1987, pp. 827-838.

[Farvardin90] N. Farvardin, "A study of vector quantization for noisy channels", IEEE Trans. Inform. Theory, Vol. 36, No. 4 , Jul. 1990, pp. 799-809.

[Ferguson84] J.J. Ferguson and J.H. Rabinowitz, "Self-synchronising Huffman codes", IEEE Trans. Inform. Theory, Vol. 30, No. 4 , Jul. 1984, pp. 687-693.

[Flatman94] A. Flatman, "Wireless LANs: developments in technology and standards", IEE Computing and Control Engineering Journal, Oct. 1994, pp. 219-224.

[Forchheimer89] R. Forchheimer and T. Kronander, "Image Coding - From Waveforms to Animation", IEEE Trans. on Acoustics, Speech and Signal Process., Vol. 37, No. 12, Dec. 1989, pp. 2008 - 2023.

[Fritchman67] B.D. Fritchman, "A Binary Channel Characterisation using Partitioned Markov Chains", IEEE. Trans. of Information Theory, Vol. 13, No. 2, Apr. 1967, pp. 221-227.

[Fung93] V. Fung, T.S. Rappaport and B. Thoma, "Bit Error Simulation for $\pi/4$ DQPSK Mobile Radio Communications using Two-Ray and Measurement-Based Impulse Response Models", IEEE Journal on Selected Areas in Commun., Vol. 11, No. 3, Apr. 1993, pp. 393-405.

[Gardiner90] J. Gardiner and W. Tuttlebee, "The future of Personal Communications", in *Cordless Telecommunications in Europe*, W.H.W. Tuttlebee (Ed.), Springer-Verlag, 1990, pp. 233-250.

[Ghanbari89] M. Ghanbari, 'Two layer coding of video signals for VBR networks", Journal on Selected Areas in Commun., Vol. 7, No. 5, Jun. 1989, pp. 771-781.

[Ghanbari90a] M. Ghanbari, "The cross-search algorithm for motion estimation", IEEE Trans. on Commun., Vol. 38, No. 7, Jul. 1990, pp. 950-953.

[Ghanbari91] M. Ghanbari, "Packet Video", in *Image Processing*, edited by D. Pearson, McGraw-Hill, London, U.K., 1991.

[Ghanbari92] M. Ghanbari, "An adapted H.261 two-layer video codec for ATM networks", IEEE Trans. on Commun., Vol. 40, No. 9, Sep. 1992, pp. 1481-1490.

[Ghanbari93] M. Ghanbari and V. Seferidis, "Efficient Two-Layer Coding Techniques", IEEE Visicom 93, Fifth Intl. Workshop on Packet Video, Berlin, Mar. 22-23, 1993, pp. B2.1- B2.5.

[Gharavi89] H. Gharavi, "Differential sub-band coding of video signals", IEEE ICASSP89, Intl. Conf. on Acoustics, Speech and Signal Processing 1989, pp. M9.4.1-M9.4.5.

[Gharavi90] H. Gharavi and M. Mills, "Blockmatching motion estimation algorithms - New results", IEEE Transactions on Circuits and Systems, Vol. 37, No. 5, May 1990, pp. 649-651.

[Gilbert60] E.N. Gilbert, "Capacity of a Burst-Noise Channel," Bell System Technical Journal, Vol. 39, No. 9, Sep. 1960, pp. 1253-1266.

[Gold94] E.M. Gold, "PCs rewrite the rules for videoconferencing", Data Communications, Mar. 1994, pp. 95-103.

[Gonzalez87] R.C. Gonzalez and P. Winz, *Digital Image Processing*, Addison-Wesely, Reading, Massachusetts, 1987.

[Goodman89] D.J. Goodman, R.A. Valenzuela, K.T. Gayliard and B. Ramamurthi, "Packet Reservation Multiple Access for local Wireless Communications", IEEE Trans. on Commun., Vol. 37, No. 8, Aug. 1989, pp. 885-890.

[Goodman90a] D.J. Goodman, "Cellular Packet Communications," IEEE Trans. on Commun., Vol. 38, No. 8, Aug. 1990, pp. 31-40.

[Goodman90b] D.J. Goodman, "Packet transmission and switching in advanced wireless information networks", IEEE ICC90, Atlanta, 16-19 Apr. 1990, pp. 1473-1478.

[Goodman91a] D.J. Goodman, "Trends in cellular and cordless communications", IEEE Commun. Mag., Vol. 29, No. 6, Jun. 1991, pp. 31-40.

[Goodman91b] D.J. Goodman, "Second Generation Wireless Information Networks", IEEE Trans. on Vehicular Tech., Vol. 40, No. 2, May 1991, pp. 366-374.

[Goodman91c] D.J. Goodman and S.X. Wei., "Efficiency of packet reservation multiple access," IEEE Trans. Vehicular Trans., Vol. 40, No. 1, Feb. 1991, pp. 170-176.

[Gray84] Gray R.M., "Vector Quantization", IEEE Acoustics, Speech and Signal Process. Mag., Apr. 1984, pp 4-29.

[Grillo95] D. Grillo, N. Metzner and E.D. Murray, "Testbeds for Assessing the Performance of a TDMA-Based Radio Access Design for UMTS", IEEE Personal Communications , Vol. 2, No. 2, Apr. 1995, pp. 36-45.

[GSM] References in M. Rahnema, "Overview of the GSM system and protocol Architecture", IEEE Comms. Mag., April 1993, pp. 92-100.

[Halsall92] F. Halsall, "An investigation into next generation cordless technology for Data Communication within a building", IEE Colloquium on Radio LANs, IEE Digest 1992/104, May 1992, pp. 1-6.

[Hamano93] T. Hamano, K.Sakai, E. Morimatsu and K. Matsuda, "Cell-loss compensation scheme based on MPEG2 Test Model 2", IEEE Visicom93, 5th. Intl. Workshop on Packet Video, Berlin, Mar. 22-23 1993, pp. C5.1-C5.6.

[Hanzo94a] L. Hanzo, R. Stedman, R. Steele and J.C.S. Cheung, "A Mobile Speech, Video and Data Transceiver Scheme", IEEE VTC94, Stockholm, Sweden, 8-10 June 1994, Ch. 389, pp. 452-456.

[Hanzo94b] L. Hanzo, J.C.S. Cheung, R. Steele and W.T. Webb, "A Packet Reservation Multiple Access Assisted Cordless Telecommunication Scheme", IEEE Trans. on Vehicular Tech., Vol. 43, No. 2, May 1994, pp. 234-244.

[Harasaki94] H. Harasaki and M. Yano, "Interleaved forward Error-Correction for Variable Bit-Rate Video Coding", IEEE Supercomm/ICC94, New Orleans, Los Angeles, 1-5 May 1994, Ch. 344, pp. 225-229.

[Hashemi93] H. Hashemi, "The Indoor Radio Propagation Channel", Proc. of the IEEE, Vol. 81, No. 7, Jul. 1993, pp. 943-968.

[Hayes91] V. Hayes, "Standardisation efforts for Wireless LANs", IEEE Network Mag., Vol. 5, No. 6, Nov. 1991, pp. 19-20.

[Helard91] J.F. Helard and B. Lefloch, "Trellis Coded Orthogonal Frequency-Division Multiplexing for Digital Video Transmission", IEEE Globecom91, Phoenix, Arizona, 2-5 Dec. 1991, Ch. 400, pp. 785-791.

[Hemami93] S. Hemami and T.H.Y. Meng, "Spatial and Temporal Video Reconstruction for Non-Layered Transmission", IEEE Visicom 93, Fifth Intl. Workshop on Packet Video, Berlin, Mar. 22-23, 1993, pp. D2.1- D2.6.

[Hepper90] Hepper and Dietmar, "Efficiency analysis and application of uncovered background prediction in a low bit rate image coder", IEEE Trans. Commun., Vol. 38, No. 9, Sep. 1990, pp. 1578-1584.

[Horst93] R. Ter Horst, F. Hoeksema, G. Heidemann and H. Tattje, "Efficient ATM Transmission with 1-layer H.261 Videocodecs", IEEE Visicom 93, 5th

International Workshop on Packet Video, Berlin, Mar. 22-23 1993, pp. D 4.1-D 4.8.

[Hung93] A.C. Hung, *PVRG P64-Codec 1.1*, Stanford Portable Video Radio Group, Stanford University, 1st March 1993.

[Hung94] A.C. Hung and Th.Y. Meng, "A comparison of fast inverse discrete cosine transform algorithms", IEEE Multimedia Systems, Vol. 2, No. 5, 1994, pp. 204-217.

[Hulyalkar93] S.M. Hulyalkar, Y.S. Ho, K.S. Challapali, D.A. Bryan, C. Basile, H. White, D. Wilson and B. Bhatt, "Advanced Digital HDTV Transmission System for Terrestrial Video Simulcasting", IEEE Journal on Selected Areas in Commun., Vol. 11, No. 1, Jan. 1993, pp. 119-125.

[Jain81] A.K. Jain, "Image Data Compression: A Review", Proc. IEEE, Vol. 69, No. 3, Mar. 1981, pp. 349-389.

[Jakes74] W.C. Jakes (Ed.), *Microwave Mobile Communications*, J. Wiley & Sons, Inc., New York, 1974.

[Jalali91] R.S. Jalali, "Transmission of image data over digital networks involving mobile terminals", Ph.D. Thesis, University of Aston in Birmingham, U.K., 1990.

[Jeruchim92] M.C. Jeruchim, P. Balaban and K.S. Shanmugan, *Simulation of Communication Systems*, Plenum Press, New York, 1992.

[Kahn94] J.M. Kahn, J.R. Barry, M.D. Audeh, J.H. Carruthers, W.J. Krause and G.W. Marsh, "Non-directed Infrared links for high-capacity wireless LANs", IEEE Personal Communications, Vol. 1, No. 2, Second Quarter 1994, pp. 12-25.

[Karlsson88] G. Karlsson and M. Vetterli, "Subband Coding of Video for Packet Networks", Optical Engineering, Vol. 27, No. 7, Jul. 1988, pp. 574-586.

[Karlsson89] G. Karlsson and M. Vetterli, "Packet Video and Its Integration into the Network Architecture", IEEE Journal on Selected Areas in Commun., Vol. 7, No. 5, Jun. 1989, pp. 739-751.

[Katz94] R.H. Katz, "Adaptation and Mobility in Wireless Information Systems", IEEE Personal Communications, Vol. 1, No. 1, First Quarter 1994, pp. 6-17.

[Khansari94a] M. Khansari, A. Jalali, E. Dubois and P. Mermelstein, " Robust low bit-rate video transmission over wireless access systems", IEEE Supercomm/ICC94, Conference Proc., 1994, pp. 571-575.

[Khansari94b] M. Khansari, A. Zakauddin, W.Y. Chan, E. Dubois and P. Mermelstein, " Approaches to layered coding for dual-rate wireless video transmission", IEEE ICIP-94, Austin, Texas, 13-16 Nov. 1994, pp. 258-262.

[Kishimoto91] R. Kishimoto and K. Irie, "HDTV Transmission system in an ATM-based network", Signal Processing: Image Communication, Vol. 3, No. 2-3, Sep. 1991, pp. 111-122.

[Kishino89] F. Kishino, K. Manabe, Y. Hayashi, and H. Yasuda, "Variable Bit Rate Coding of Video Signals for ATM Networks", IEEE Journal on Selected Areas in Commun., Vol. 7, No. 5, Jun. 1989, pp. 801-806.

[Kleinrock75] L. Kleinrock, *Queuing Systems Vol. I: Theory*, J. Wiley and Sons, New York, 1985.

[Kucar91] A.D. Kucar, "Mobile Radio: An Overview", IEEE Communications Mag., Vol. 29, No. 11, Nov. 1991, pp. 72-85.

[Kunt85] M. Kunt, A. Ikonomopoulos and M. Kocher, "Second Generation Image Coding Techniques", Proc. IEEE, Vol. 73, No. 4, Apr. 1985, pp. 549-574.

[LeGall91] A.D. Le Gall, "MPEG: A Video Compression Standard for Multimedia Applications", Communications of the ACM, Vol. 34, No. 4, Apr. 1991, pp. 47-58.

[Lettera89] C. Lettera and L. Masera, "Foreground/background segmentation in videotelephony", Signal Processing: Image Communications, Vol. 1, No. 2, Oct. 1989, pp. 181-189.

[Li87] V.O.K. Li, "Multiple Access Communications Networks", IEEE Communications Mag., Vol. 25, No. 6, Jun. 1987, pp. 41-48.

[Lim90] J.S. Lim, *Two Dimensional Signal and Image Processing*, Prentice-Hall, 1990.

[Lin83] S. Lin and D.J. Costello Jr., *Error Control Coding: Fundamentals and Applications*, Prentice-Hall, Englewood Cliffs, New Jersey, 1983.

[Lin84] S. Lin, D.J. Costello Jr. and M.J. Miller, "Automatic-Repeat-Request Error-Control Schemes," IEEE Commun. Mag., Vol. 22, No. 12, Dec. 1984, pp. 5-17.

[Lindsey73] W.C. Lindsey and M.K. Simon, *Telecommunication System Engineering*, Prentice Hall, New Jersey 1973.

[Links94] C. Links, W. Diepstraten and V. Hayes, "Universal Wireless LANs", Byte May 1994, pp. 99-108.

[Lodge91] J.H. Lodge, "Mobile Satellite Communications Systems: Towards Global Personal Communications", IEEE Communications, Vol. 31, No. 11, Nov. 1991, pp. 24-31.

[MacDonald92] N. MacDonald, "Transmission of compressed video over radio links", SPIE Annual Conf. on Visual Communications and Image Processing '92, Boston, MA, 18-20 Nov. 1992, Conf. Proc. Vol. 1818, pp. 1484-1488.

[Maxted85] J.C. Maxted and J.P. Robinson, "Error Recovery for variable length codes", IEEE Trans. Inform. Theory, Vol. 31, No. 6, Nov. 1985, pp. 794-801.

[Meng94] T.H. Meng, E.K. Tsern, A.C. Hung, S.S. Hemami and B.M.Gordon, "Video compression for Wireless Communications", Annual Symposium on Wireless Personal Comms., Virginia Tech. Campus, Blacksburg, VA, 1994, Ch. 23, pp. 101-117.

[Mermelstein93] P. Mermelstein, A. Jalali and H. Leib, "Integrated Services on Wireless Multiple Access Networks", IEEE ICC93, Geneva, Switzerland, 23-26 May 1993, Conf. Proc. Vol. 3, pp. 863-867.

[Michelson85] A.M. Michelson and A.H. Levesque, *Error-Control Techniques for Digital Communication*, J. Wiley & Sons, Inc., New York, 1985.

[Mickos94] R. Mickos and L. Onural, "Simulation of A DCT Based Very-Low Bit-Rate Codec for Mobile Video Coding", IEEE 7th Mediterranean Electrotechnical Conference, Antalya, Turkey, 12-14 Apr. 1994, Ch. 343, pp. 203-206.

[Mitchell81] O.R. Mitchell and A.J. Tabatai, "Channel Error Recovery for Transform Image Coding", IEEE Trans. Commun., Vol. 29, No. 12, Dec. 1981, pp. 1754-1762.

[Modestino79] J.W. Modestino and D.G. Daut, "Combined source-channel coding of images", Trans. Commun., Vol. 27, No. 11, Nov. 1979, pp. 1644-1659.

[Modestino81] J.W. Modestino , D.G. Daut, and A.L. Vickers, "Combined source-channel coding of images using the block cosine transform", IEEE Trans. Commun., Vol. 29, No. 9, Sep. 1981, pp. 1261-1274.

[Morrison91] D.G. Morrison and D. Beaumont, "Two-layer video coding for ATM networks", Signal Processing: Image Communications, Vol. 3, No. 2, Jun. 1991, pp. 179-195.

[Morrison94] D.G. Morrison, "Low Delay Video Transcoders for Multimedia Interworking", Private Communication.

[Morrsion90] D.G. Morrison, "Variable bit rate video coding for asynchronous transfer mode networks", British Telecom Technology Journal, Vol. 8, No. 3, Mar. 1990, pp. 70-80.

[Mounts69] F.W. Mounts, "A video encoding system using conditional picture-element replenishment", Bell System Technical Journal, Vol. 48, No. 7, Sep. 1969, pp. 2545-2554.

[MPEG1] ISO/IEC DIS 11172, "Coding of Moving Pictures and Associated Audio for Digital Storage Media up to about 1.5 Mbit/s", Oct. 1992.

[MPEG2] ISO/IEC DIS 1318-2, "Video part", March 1994.

[Musmann85] H.G. Musmann and P. Pirsch, "Advances in Picture Coding", Proc. IEEE, Vol. 73, No. 4, Apr. 1985, pp. 525-548.

[Nasarbadi89] N.M. Nasarbadi, and R.A. King, "Image coding using vector quantization - A review", IEEE Trans. Commun., Vol. 37, No. 6, Jun. 1989, pp. 957-971.

[Netravali80] A.N. Netravali and J.O. Limb, "Picture Coding: A Review", Proc. IEEE, Vol. 68, No. 3, Mar. 1980, pp. 366-406.

[Netravali88] A.N. Netravali and B.G. Haskell, *Digital Pictures: Representation and Compression*, Plenum Press, New York, 1988.

[Ngan82] K.N. Ngan and R. Steele, "Enhancement of PCM and DPCM images corrupted by transmission errors", IEEE Trans. Commun., Vol. 30, No. 1, Jan. 1982, pp. 257-265.

[Nix91] A.R. Nix, R.J. Castle and J.P. McGeehan, "The application of 16 APSK to mobile fading channels", Proc. of the IEE 6th Intl. Conf. on Mobile Radio and Personal Commun., Warwick UK, 9-11 Dec. 1991, pp. 233-240.

[Nix92] A.R. Nix and J.P. McGeehan, "High Speed Data Transmission formats for Radio LANs," IEE Colloquium on Radio LANs, Digest 1992/104, May 1992, pp. 3/1-3/6.

[Nomura91] M. Nomura, T. Fujii and N. Ohta, "Layered coding for ATM based video distribution systems", Signal Processing: Image Communication, Vol. 3, No. 4, Sep. 1991, pp. 301-311.

[Owen90a] F.C. Owen and C.D. Pudney, "DECT - Integrated Services for Cordless Telecommunications", Proc . of the 5th IEE Intl. Conf. on Mobile Radio and Personal Commun., IEE Conf. Digest No. 315, pp. 152-156.

[Owen90b] F. Owen, "Cordless data communications", in *Cordless Telecommunications in Europe*, W.H.W. Tuttlebee (Ed.), Springer-Verlag, 1990.

[Padgett95] J.E. Padgett, C.G. Gunther and T. Hattori, "Overview of Wireless Personal Communications", Proc. of the IEEE, Vol. 33, No. 1, Jan. 1995, pp. 28-41.

[Pahlavan94] K. Pahlavan and A.H. Levesque, "Wireless Data Communications", Proc. of the IEEE, Vol. 82, No. 9, 1994, pp. 1398-1430.

[Parsons89] J.D. Parsons and J.G. Gardiner, *Mobile Communication Systems*, Blackie, Glasgow and London, 1989.

[Pearson90] D. Pearson, "Packet Video", IEE Review, Vol. 36, No. 8, Sep. 1990, pp. 315-318.

[Pelz94] R.M. Pelz, "An Unequal Error Protected px8 Kbit/S Video Transmission for DECT", IEEE VTC 1994, Stockholm, Sweden, 8-10 Jun. 1994, pp. 1020-1024.

[Pereira93] F. Pereira, D. Cortez and P. Nunes, "Mobile Videotelephone Communications - the CCITT H.261 Chances", SPIE Vol. 1977, Conf. on Video Communications and PACS for Medical Applications, Berlin, Germany, 5-8 Apr. 1993, pp. 168-179.

[Pereira94] R. Pereira, "A Mobile Audiovisual Terminal for the DECT System", IEEE 7th Mediterranean Electrotechnical Conference, Antalya, Turkey, 12-14 Apr. 1994, Ch. 343, pp. 28-31.

[Potter92] A.R. Potter, "Implementation of PCNs Using DCS1800", IEEE Communications Mag., Vol. 30, No. 12, Dec. 1992, pp. 32-37.

[Pratt78] W.K. Pratt, *Digital Image Processing*, J. Wiley & Sons, Inc., New York, 1978.

[Proakis89] J.G. Proakis, *Digital Communications*, McGraw Hill, New York, 1989.

[RACE2072] RACE project 2072, "Mobile Audio Visual Terminal - MAVT".

[Rahnema93] M. Rahnema, "Overview of the GSM system and protocol Architecture", IEEE Comms. Mag., April 1993, pp. 92-100.

[Ramchandran93] K. Ramchandran, A. Ortega, K. M. U and Martin Vetterli, "Multiresolution Broadcast for Digital HDTV Using Joint Source/Channel Coding", IEEE Journal on Selected Areas in Commun., Vol. 11, No. 1, Jan. 1993, pp. 6-23.

[Rappaport91a] T.S. Rappaport, S.Y. Seidel and K. Takamizawa, "Statistical Channel Impulse Response Models for Factory and Open Plan Building Radio Communciation System Design", IEEE Trans. on Commun., Vol. 39, No. 5, May 1991, pp. 794-806.

[Rappaport91b] T.S. Rappaport, "Wireless Personal Communications: Trends and Challenges", IEEE Antennas and Propagation. Mag., Vol. 33, No. 5, Oct. 1991, pp. 19-29.

[Rappaport91c] T.S. Rappaport and V. Fung, "Simulation of Bit Error Performance of FSK, BPSK, Pi/4 DQPSK in Flat Fading Indoor Radio Channels Using a Measurement-Based Channel Model.", IEEE Trans. on Vehicular Tech., Vol. 40, No. 4, Nov. 1991, pp. 731-739.

[Rappaport91d] T.S. Rappaport, "The Wireless Revolution", IEEE Commun. Mag., Vol. 29, No. 11, Nov. 1991, pp. 52-71.

[Raychaudhuri93] D. Raychaudhuri and N. Wilson, "Multimedia Transport in Next-Generation Personal Communication Networks", IEEE ICC93, Geneva, Switzerland, 23-26 May 1993, Ch. 367, pp. 858-862.

[Raychaudhuri94a] D. Raychaudhuri, "Video and Multimedia Transport over Packet Media", pp. 9-18, in B. Jabbari (Ed.), *Worldwide Advances in Communication Networks*, Plenum Press, New York, 1994.

[Raychaudhuri94b] D. Raychaudhuri, "ATM Based Transport Architecture for Multiservices Wireless Personal Communication-Networks", IEEE Supercomm/ICC 94,

New Orleans, Los Angeles, 1-5 May 1994, Ch. 344, pp. 559-565.

[Reader95] C. Reader, "MPEG-4 - Coding for content, interactivity, and universal accessibility", Optical Engineering, Vol. 35, No. 1, 1996, pp. 104-108.

[RFC1889] RFC1889, "RTP: A Transport Protocol for Real Time Applications", IESG, Nov. 22, 1995.

[Rocca69] F. Rocca, "Television bandwidth compression utilising frame-to-frame correlation and movement compensation", Symp. Picture Bandwidth Compression, M.I.T. Cambridge, MA, 1969.

[Roser93] M. Roser, J. Caballero, P. Villegas and M. Simon, "Study of the influence of cell loss in the MPEG scheme", Visicom 93, IEEE 5th. Intl. Workshop On Packet Video, Berlin 1993, pp. C.3.1-C.3.

[Saleh87] A.A.M. Saleh, A.J. Rustako Jr. and R.S. Roman, "Distributed Antennas for Indoor Radio Communications", IEEE Trans. on Commun., Vol. 35, No. 12, Dec. 1987, pp. 1245-1251.

[Saleh91] A.A.M. Saleh, A.J. Rustako, L.J. Cimini, G.J. Owens and R.S. Roman, "An experimental TDMA Indoor Radio Communications System using Slow Frequency Hopping and coding", IEEE Trans. on Commun., Vol. 39, No. 1, Jan. 1991, pp. 152-162.

[Schultes90] G. Schultes, E. Bonek, P. Weger and W. Herzog, "Basic Performance of a Direct Conversion DECT Receiver", IEEE Electron. Lett., Vol. 26, No. 21, Oct. 1990, pp. 1746-1748.

[Schultes92] G. Schultes and I. Crhon, "Measured Performance of DECT Transmission in Low Dispersive Indoor Radio Channel", IEE Electron. Lett., Vol. 28, No. 17, Aug. 1992, pp. 1625-1626.

[Scorse93] J.F. Scorse, "Sending Images through the Air: Image Transmission by Radio", Advanced Imaging, Mar. 1993, pp. 38-41.

[Seferidis93] V. Seferidis, M. Ghanbari and D.E. Pearson, "Packet Video Quality Assessment", IEEE Visicom 93, Fifth Intl. Workshop on Packet Video, Berlin, Mar. 22-23, 1993, pp. E3.1- E3.4.

[Shannon64] C.E. Shannon and W. Weaver, *The Mathematical Theory of Communication*, The University of Illinois Press, 1964.

[Sherif94] M. H. Sherif, "The Metamorphosis of the Public Switched Telecommunications Network", in Standards Report, IEEE Commun. Mag., Vol. 32, No. 1, Jan. 1994, pp. 14 -15.

[Siew89] C.K. Siew and D.J. Goodman, "Packet Data-Transmission Over Mobile Radio Channels", IEEE Transactions on Vehicular Technology, Vol. 38, No. 2, May 1989, pp. 95-101.

[Siracusa93] R.J. Siracusa and J. Zdepski, "Flexible and Robust Packet Transport for Digital HDTV", IEEE Journal on Selected Areas in Commun., Vol. 11, No. 1, Jan. 1993, pp. 88-97.

[Sklar83] B. Sklar, "A structured overview of digital communications-A tutorial review", IEEE Commun. Mag., Part I - Vol. 21, Aug. 1983, pp. 4-17, Part II Vol. 21, Oct. 1983, pp. 6-21.

[Stallings90] W. Stallings, "CCITT standards foreshadow B-ISDN", Telecommunications, May 1990, pp. 89-95.

[Stedman92] R. Stedman, R. Steele, H. Gharavi and L. Hanzo, "A 22 kBd Mobile Video Telephone Scheme", IEEE VTC 92, Denver, CO, 10-13 May 1992, pp. 251-254.

[Steele79] R. Steele, D.J. Goodman and D.J. McGonegal, "DPCM with forced updating and partial correction of transmission errors", Bell Syst. Tech. Journal, Vol. 58, No. 3, 1979, pp. 721-728.

[Steele79] R. Steele, D.J. Goodman and D.J. McGonegal, "A difference detection and correction scheme for combating DPCM transmission errors", IEEE Trans. Commun., Vol. 27, No. 1, Jan. 1979, pp. 252-255.

[Steele89] R. Steele, "The Cellular Environment of Lightweight Handheld Portables", IEEE Commun. Mag., Vol. 27, No. 7, Jul. 1989, pp. 20-29.

[Steele92] R. Steel, "An update on Personal Communications", Guest Editorial, IEEE Commun. Mag., Vol. 30, No. 12, Dec. 1992, pp. 30-31.

[Steele94] R. Steele, "The Evolution of Personal Communications", IEEE Personal Communications, Vol. 1, No. 2, Second Quarter 1994, pp. 6-11.

[Stenger89] L. Stenger, "Digital Coding of television signals - CCIR activities for standardisation", Signal Processing: Image Communication, Vol. 1, No. 1, Jun. 1989, pp. 29-43.

[Streit94] J. Streit and L. Hanzo, "A Fractal Video Communicator", IEEE VTC 94, Stockholm, Sweden, 8-10 June 1994, pp. 1030-1034.

[Tanenbaum89] A.S. Tanenbaum, *Computer Networks*, Prentice Hall, Englewood Cliffs, N.J., 1989.

[Thoma89] R. Thoma and M. Bierling, "Motion compensated interpolation considering covered and uncovered background", Signal Processing: Image Communication, Vol. 1, No. 2, Oct. 1989, pp. 191-212.

[Tubaro90] S. Tubaro, "Hybrid image coder with vector quantizer", Signal Processing: Image Communication, Vol. 2, No. 1, May 1990, pp. 95-104.

[Tubaro91] S. Tubaro, "A two layers coding scheme for ATM networks", Signal Processing: Image Communication, Vol. 2, No. 3, Oct. 1990, pp. 129-141.

[Turletti92] T. Turletti, *Inria Vision System Manual*, INRIA, Sophia Antipolis, France, 6th May 1992.

[Tuttlebee90] W.H.W. Tuttlebee (Ed.), *Cordless Telecommunications in Europe*, Springer-Verlag, London, 1990.

[Tuttlebee92] W.H.W. Tuttlebee, "Cordless Personal Communications", IEEE Commun. Mag., Vol. 32, No. 12, Dec. 1992, pp. 42-53.

[Vembu95] S. Vembu, S. Verdu and Y. Steinberg, "The Source-channel separation theorem revisited", IEEE Trans. on Information Theory, Vol. 41, No. 1, Jan. 1995, pp. 44-54.

[Verbiest88] W. Verbiest, L. Pinnoo and B. Voeten, "The Impact of the ATM Concept on Video Coding", IEEE Journal on Selected Areas in Commun., Vol. 6, No. 9, Dec. 1988, pp. 1623-1632.

[Voran92] S.D. Voran and S. Wolf, "The Development and Evaluation of An Objective Video Quality Assessment System that Emulates Human Viewing Panels", IEE IBC-92, Amsterdam, Netherlands, 3-7 Jul. 1992, pp. 504-508

[Wada89] M. Wada, "Selective Recovery Of Video Packet Loss Using Error Concealment", IEEE Journal on Selected Areas in Commun., Vol. 7, No. 5, Jun. 1989, pp. 807-814.

[Watanabe94] K. Watanabe, S. Masaki, M Shinoda and A. Kurobe, "A Study on Transmission of Low Bit-Rate Coded Video over Radio Links", IEEE VTC 1994, Stockholm, Sweden, 8-10 June 1994, pp. 1025-1029.

[Webb90] W.T. Webb, L. Hanzo and R. Steele, "Speech and Data Transmission via QAM for Mobile Personal Communications", IEE Colloquium on University Research in Mobile Radio, IEE Conference Digest 1990/143, pp. 7/1-7/6.

[Webb91] W.T. Webb and R. Steele, "A 32 Mbit/s mobile radio link", IEE Sixth Intl. Conf. of Mobile Radio and Personal Communications, Warwick, U.K., 1991, pp. 31-38.

[Webb92] W.T. Webb, "Modulation Methods for PCNs", IEEE Communications Mag., Vol. 30, No. 12, Dec. 1992, pp. 90-95.

[Wong94] P. Wong, G. Schultes, A. Lasa and F. Halsal, "Performance of ARQ Protocols on Multibearer Connections in the DECT System", IEEE VTC 1994, Stockholm, Sweden, 8-10 June 1994, pp. 1398-1401.

[Woods86] J.W. Woods and S.D. O'Neil, "Subband coding of Images", IEEE Trans. Acoustics, Speech and Signal Process., Vol. 34, No. 5, Oct. 1986, pp. 1278-1287.

[Wu94] W.W. Wu, E.F. Miller, W.L. Pritchard and R.L. Pickholtz, "Mobile Satellite Communications", Proc. of the IEEE, Vol. 82 No. 9, 1994, pp. 1431-1448.

[Yamamoto89] Y. Yamamoto and T. Wright, "Error Performance in Evolving Digital Networks including ISDNs", IEEE Commun. Mag., Vol. 27, No. 4, Apr. 1989, pp. 12-18.

[Yasuda89] Yasuda, "Standardisation on multimedia coding in ISO", Signal Processing: Image Communication, Vol. 1, No. 1, Jun. 1989, pp. 3-16.

[Yin93] L. Yin and S. Yu, "An effective anti-interference method for videoconference image coding", IEEE Visicom 93, 5th. Intl. Workshop On Packet Video, Berlin 1993, C.8.1-C.8.4.

[Yoshida88] S. Yoshida, F. Ikegami and T. Takeuchi, "Causes of Burst Errors in Multipath Fading Channel", IEEE Trans. on Commun., Vol. 36, No. 1, Jan. 1988, pp. 107-113.

[Young79] W.R. Young, "Advanced Mobile Phone Service: Introduction, Background and Objectives", BSTJ, Vol. 58, No. 1, Jan 1979, pp. 1-14.

[Young92] K. Young, "A WARC on the wild side of the spectrum", Communications Networks, May 1992, p. 23.

[Zager90] K. Zager and A. Gersho, "Pseudo-Gray coding", IEEE Trans. Commun., Vol. 29, No. 12, Dec. 1990, pp. 2147-2158.

[Zaghloul91] H. Zaghloul, M. Fattouche, G. Morrison and D. Tholl, "Comparison of indoor propagation channel characteristics at different frequencies", IEE Electron. Lett., Vol. 27, No. 22, Oct. 1991, pp. 2077-2079.

[Zammit89] Zammit S., "Parallel computer architectures for real time image processing", MSc. Thesis, University of Aston, U.K., Sept. 1989.

[Zammit91] S. Zammit, "Digital Image Communications", Ph.D. Transfer Report, University of Aston, U.K., 1991.

[Zammit93] S. Zammit and G.F. Carpenter, "Packet Video Coding for Wireless Access Networks", IEEE Visicom93, 5th. Intl. Workshop on Packet Video, Berlin, Mar. 22-23 1993 G2.1-G2.4.

[Zammit96] S. Zammit and G. Carpenter, "Transcoding Between H.261 and MPEG-1 streams", PCS96, IEEE Picture Coding Symposium 1996, accepted for publication.

[Zhang90] K. Zhang and K. Pahlavan, "Integrated voice/data system for mobile indoor radio networks", IEEE Trans. on Vehicular Tech., Vol. 39, No. 1, Feb. 1990, pp. 75-82.

[Zhang94] Y.Q. Zhang, Y.J. Liu and R.L. Pickholtz, " Layered Image Transmission over Cellular Radio Channels", IEEE Trans. Vehicular Tech., Vol. 43, No. 3, Aug. 1994, pp. 786-793.

[Zubrzycki94] J.T. Zubrzycki, "MBS - A Wireless Network for Digital Video", IEE IBC94, Amsterdam, Netherlands, 16-20 Sep. 1994, pp. 266-271.

Appendix One
Report on Current
Wireless Telecommunications Networks

1. Wireless Network Taxonomy

This report summaries the salient characteristics of the current, major, wireless networks based on a literature survey [Goodman91, Kucar91, Cox92, Potter92, Katz94, Steele94, Padgett95, Cox95]. A taxonomy has been developed to present Wireless Networks (WNs) in a coherent manner. The taxonomy shown in Figure A.1 distinguishes between Wide Area Wireless Networks (WAWNs) and Local Area Wireless Networks (LAWNs) at the highest level. WAWNs offer nation-wide and global coverage, whereas LAWNs cover a local area. LAWNs may also be used to access global wide area networks, but the planning and execution of this is left to the end user.

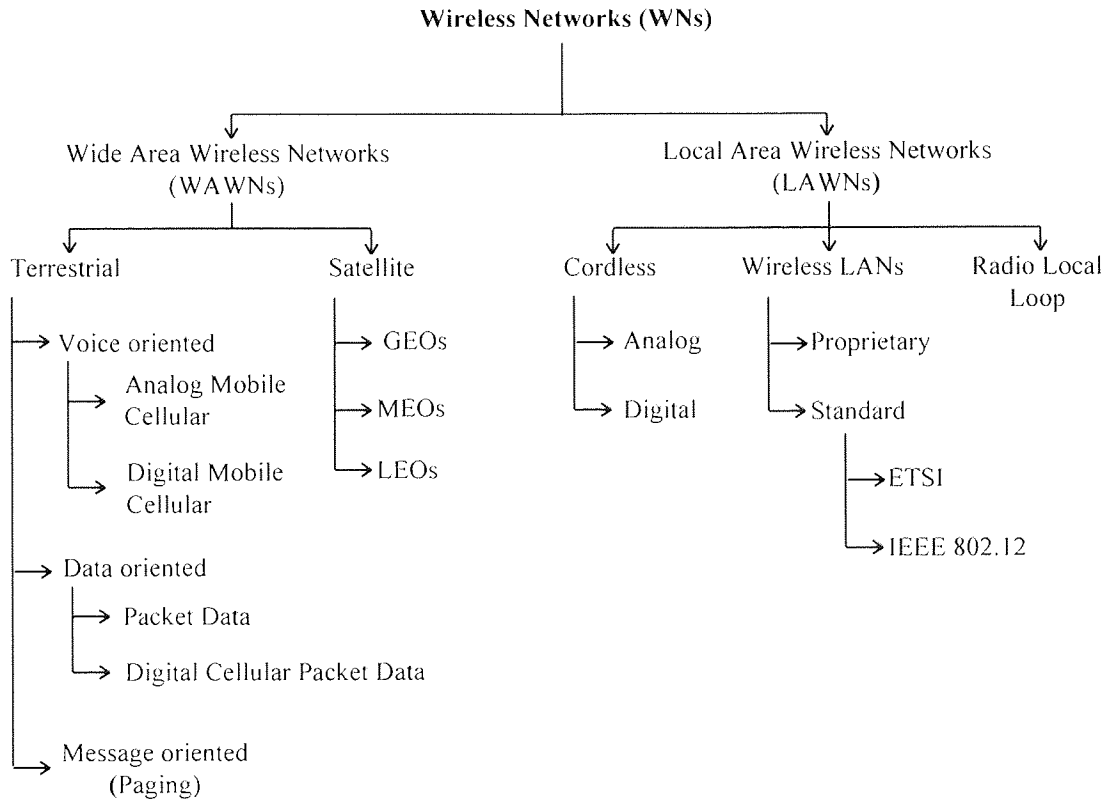


Figure A.1 Wireless Network Taxonomy

WAWNs can be divided into Terrestrial and Satellite based mobile networks. Satellite mobile networks [Lodge91, Wu94] may be classified according to their orbit into Geostationary Earth Orbit (GEO), Medium Earth Orbit (GEO) and Low Earth Orbit (LEO). LEO systems can be further classified

into 'Little' LEO and 'Big' LEO to reflect the size of the satellites used in the two classes.

Terrestrial based WAWNs can be classified by their main (predominant) service offering namely Voice, Data or Message, though this does not preclude mixed service offerings (*e.g.* Fax services over Analog Mobile Cellular networks). Voice based Terrestrial WAWNs can be further divided into Analog Cellular and Digital Cellular mobile networks. Analog networks are first-generation networks whereas digital networks are second-generation networks.

Data oriented WAWNs can be classified as packet data networks and digital cellular packet data networks. The former include custom networks such as ARDIS, MOBITEX and TETRA [Pahlavan94] which provide an essentially packet switched data service. The latter use spare capacity on analog cellular networks to overlay packet switched service (e.g. Cellular Digital Packet Data). As already mentioned, voice oriented networks may also support asynchronous data and facsimile transmission.

LAWNs are subdivided into three classes; Cordless Telephones, Wireless Local Area Networks and Radio Local Loops. Cordless Telephones (CTs) are subdivided into Analog CTs and Digital CTs with Analog CTs belonging to the first generation and Digital CTs to the second. Wireless LANs (WLANs) can be proprietary or standards based. Two bodies are developing WLAN standards; the European Telecommunications Standards Institute (ETSI) and the IEEE committee 802.11 [Hayes91, Black92]. Radio Local Loops are designed to offer competition to fixed line local loops by using 'Lamppost' base stations to deliver basic voice and data services to the home [Padgett95].

The salient Physical layer, MAC layer and Link layer characteristics of the current, major wireless networks are summarised below. Only networks which can support real-time, interactive communications services, such as voice or video, are considered.

2. Wide Area Wireless Networks

Wide Area Wireless Networks (WAWNs) extended telecommunications services to mobile users on a national and global scale and comprise both analog and digital cellular networks.

2.1 Analog Cellular Networks

Although commercial mobile telephone services have been available since the late 1940's, the first analog mobile system, with the now ubiquitous cellular frequency reuse system to improve user capacity, was announced by AT&T in the USA in the late 70s [Young79]. The Advanced Mobile Phone Service (AMPS) uses hundreds of base stations, to provide wide coverage, interconnected using point-to-point links to switching and call processing centres. Frequency allocation and licensing problems delayed the introduction of AMPS until October 1983. AMPS now operates in 50 MHz of bandwidth between 824 MHz and 894 MHz. The bandwidth is divided into two separate transmit and receive bands (FDD) separated by 45 MHz. FDMA is used, with one channel per 30 kHz channel. The cellular system uses a 7-cell frequency resuse pattern with 96 channels per cell.

The AMPS system was followed closely by the Nordic Mobile Telephone (NMT) system, the Total Access System (TACS), and the Extended European TACS (ETACS) system in Europe, JTACS in Japan and similar systems in the rest of world.

2.2 Digital Cellular Networks

The success of Analog Cellular Systems eventually resulted in saturated systems in some markets and led to the design of second generation WNs with greater capacities. This second generation was to be fully digital to capitalise on the robustness of digital transport, voice compression, and the economies of scale allowed by Very Large Scale Integration (VLSI) technologies in integrated circuit design.

Currently five major digital cellular systems exist. The Global System for Mobile (GSM) is the most advanced second generation system, and has been installed in numerous countries. DCS1800 systems, a variant of GSM with smaller cell sizes and operating in a higher part of the spectrum, have been deployed in the UK and elsewhere recently. In the USA the IS-54 system provides three or six digital voice channels in one AMPS analog channel. A second system - IS-95, originally proposed and designed by QUALCOMM and based on spread spectrum and CDMA technology, has been standardised and is currently trialing in the USA. The Personal Digital Cellular (PDC) system has been deployed in Japan and is similar to the IS-54 system in the USA.

2.2.1 GSM

GSM [Rahnema93] is the second-generation digital pan-European cellular system specified by ETSI. It increases present capacities by 2-3 times. The system operates in the 900 MHz band. It uses TDMA with 8 timeslots (16 timeslots for later implementation) and frequency division duplex with 890-915 MHz mobile transmit and 935-960 MHz base transmit. GSM uses GMSK modulation with a BT product of 0.3. The modulation rate is 270.8333 kbit/s per TDMA channel or 33.9 kbit/s (16.9 kbit/s) per full-rate (half-rate) user channel. An equaliser is specified to combat ISI. Frequency hopping can also be used to combat the effects of multipath.

Five classes of users are envisaged; 3 classes of handhelds (0.8 W, 2 W and 5 W) and two vehicle or portable (8 W and 20 W) transceivers. GSM allows international roaming and has been implemented in many countries both within and outside of Europe.

GSM is designed to handle integrated voice/data services. Speech is coded using Residually Excited Linear Predictive (RELP) codecs. The full-rate codec bit rate is 13 kbit/s (6.5 kbit/s for the half-rate codec). Voice activity detection is used to increase battery life and reduce interference. Voice is protected using a half-rate convolutional code and a 40 ms interleaver.

Data transmission at up to 9600 baud full duplex is supported by the full-rate codec (4800 baud by the half-rate codec). Data is protected using FEC codes.

2.2.2 DCS1800

The DCS1800 system [Potter92] is based on the GSM standard but operates at 1.8 GHz in 150 MHz of spectrum. A number of systems have already been deployed in the UK and elsewhere. DCS-1800 operates with smaller cell sizes than GSM, hence it can use lower transmit powers and has been optimised for hand-held operation.

2.2.3 IS-54

The U.S. Digital Cellular (USDC) system was designed to increase cellular capacity in large cities. USDC is compatible with AMPS and the dual mode AMPS/USDC system has been standardised by the Telecommunications Industry Association (TIA) and the Electronics Industry Association (EIA) as Interim Standard IS-54 [Padgett95].

USDC supports three full-rate and six half-rate users per 30 kHz AMPS channel. The system uses TDMA/FDD and $\pi/4$ DQPSK modems. The forward and reverse channel data rate is 48.6 kbit/s which results in a spectrum efficiency of 1.62 bps/Hz. Although no equaliser is specified, it is expected that one will be required in general. Speech is coded at 7.95 kbit/s using a Vector Sum Excited Linear Predictive (VSELP) codec which rises to 13 kbit/s with FEC (full-rate).

2.2.4 IS-95

The TIA in the U.S. standardised a CDMA-based Digital Cellular System, originally proposed by Qualcomm, as interim standard IS-95 [Padgett95]. IS-95 was also designed to coexist with AMPS and allows dual mode operation.

IS-95 uses CDMA/FDD with a simplex link bandwidth of 1.25 MHz and a channel chip rate of 1.2288 Mchip/s. Spread Spectrum CDMA techniques are used on both the forward and reverse channels, although the implementation details are different. The system supports 63 voice channels at a maximum bit rate of 9.6 kbit/s. Rake receivers are employed by both mobile and base stations.

The user data rate in IS-95 changes in real-time. The Qualcomm Code Excited Linear Predictive (QCELP) coder operates at a maximum rate of 9.6 kbit/s but falls back to 1200 kbit/s during silent periods. Speech data is protected using a 1/2 rate convolutional code and a 20 ms block interleaver on the forward channel. A 1/3 rate convolutional code is used on the reverse channel.

2.2.5 PDC

The Personal Digital Cellular (PDC) system used in Japan is very similar to the IS-54 system used in the US [Padgett95]. Both use TDMA/FDMA/FDD, $\pi/4$ DQPSK modems and adaptive equalisers. The Japanese system, however, operates with a carrier separation of 25 kHz instead of 30 kHz and the channel data rate is slightly lower at 42 kbit/s. The VSELP codec operates at 6.7 kbit/s and is protected using a convolutional code. A half-rate codec is also available.

3. Local Area Wireless Networks

Local Area Wireless Networks (LAWNs) comprise Cordless Telephone Systems (CTS), Wireless Local Area Networks (WLANs) and Radio Local Loops (RLL). The former offer wireless access to integrated services, WLANs are wireless extensions to computer LANs, and the latter provide wireless connectivity in the local loop.

3.1 Cordless Telephone Systems

Cordless telephone systems [Tuttlebee90, Tuttlebee92] are short range systems and allow telephone calls to be originated and received in the vicinity (50-300 m) of a base station in the home and office, or outdoors close to telepoints. Cordless Telephones (CTs) allow mobility, although the roaming areas are non-contiguous and usually small, maybe covering one building. CTs are smaller and cheaper than cellular sets and are better poised to reach the mass markets.

3.1.1 CT1

CT1 refers to the analog domestic/office cordless telephone system [Padgett95]. A base station (BS) plugs into the phone socket and communicates with a mobile station (MS) or handset. CT1 uses analog transmissions with FDD. Very few frequencies are available to CT1 and calls may interfere with each other. There is also a lack of security because of the analog nature of the signal. The handset may only communicate with its own base station and the range is very short, typically 50-100 m.

3.1.2 CT2

CT2 is a digital extension of CT1 and follows the UCT concept [Tuttlebee92, Padgett95]. The concept calls for a personal handset to be used in three situations; at home within the range of a domestic base station, at work in a cordless PBX environment, and outdoors in the vicinity of base stations (also called Telepoints). The original CT2 system did not allow roaming between adjacent microcells within the office environment. A variation of CT2 called CT2 Plus allows incoming calls at telepoints and cordless PBX roaming.

CT2 uses 4 MHz of bandwidth between 864-868 MHz. It uses frequency division multiple access with time division duplex (FDMA/TDD) to allow communication between the BS and the MS at up to 40 m in built up areas (200 m line of sight), using 40 channels separated by 100 kHz. The system uses GFSK modulation with a BT product of 0.3. Transmitter powers as low as 10 mW are adequate allowing the handset to be small, light and run longer between charges.

CT2 does not bypass the local PSTN loop. Voice is ADPCM coded and transmitted at 32 kbit/s. Technically, the 32 kbit/s can be used to carry data although the bit rate would be lower to allow for error detection and correction protocols. Unfortunately it is not easy to increase the throughput as this would mean using more transmitters and receivers.

The Common Air Interface (CAI) was introduced by the DTI in the UK to describe the minimum requirements for a telepoint service such that a handset would be able to operate on all four licensed CT2 networks [Tuttlebee92]. It also delineated the requirements for a business environment.

The CAI allows a handset to interface to the PSTN or to ISDN. It specifies Base Station to Handset set-ups as well as in the reverse direction. The billing function is specified for a public access point. Data transmission is not precluded by CAI. A memorandum of understanding was signed by several countries in the US and in Europe, which paved the way for the introduction of CT2 systems outside the UK.

3.1.3 DECT

The Digital European Cordless Telecommunications (DECT) system follows the UCT concept [Tuttlebee90, Tuttlebee92, Padgett95]. It was commissioned by CEPT, and standardised by ETSI. The DECT standard provides a single solution for all applications within the cordless field: office systems, Telepoint applications, residential telephones and WLANs.

DECT uses microcells and has a seamless handover facility which allows full roaming in the cordless PBX environment. A DECT handset can be used domestically and in a telepoint environment where it can receive, as well as originate, calls.

DECT uses time division multiple access with time division duplex (TDMA/TDD). Up to 20 MHz of bandwidth in the 1.88-1.9 GHz range is divided into 11, 1.728 MHz channels. Each channel supports 24 timeslots in a 10 ms frame. The first twelve slots are used for BS to MS transmission and the last twelve slots for the reverse direction. The timeslots are paired to provide full duplex channels, carrying ADPCM coded speech at 32 kbit/s. A handset uses only two out of the twenty four timeslots and is thus active for one twelfth of the time. The other time is spent looking for better channels and transacting handovers. Cell handover is seamless and channels are allocated dynamically.

DECT has advantages - Dynamic Call Allocation (DCA) allows full roaming with seamless cell handovers; call set-up whenever in proximity of a base station; the ability to build a network freely without the need to re-plan for expansion; data handling is better than for CT2 since DECT was designed for speech and data in the first place. DECT is more expensive than CT2 having to operate at a higher frequency.

Digital transmission is possible via cordless modems allowing fax machines and computers communications to be used. DECT can aggregate timeslots to support data communications at up to 384 kbit/s full duplex.

3.1.4 PACS

The Personal Access System (PACS) was proposed by Bellcore in 1992 as a third generation Personal Communications System (PCS) [Cox92, Padgett95]. PACS is a development of the Wireless

Access System (WACS) described in section 3.3 below. PACS provides wireless access to integrated services in a 500 m radius around base stations which are interconnected using wired links.

The PACS radio interface consists of multiple 300 kHz channels in the 1-3 GHz range. Each channel supports a maximum bit rate of 384 kbit/s. The modulation is $\pi/4$ DQPSK. The channels are split into 2.5 ms frames, with eight timeslots per frame. Each timeslot provides a bearer channel data rate of 32 kbit/s. The system can use both TDD and FDD. Voice is coded using ADPCM, and is protected using a CRC code.

3.1.5 PHS

The Personal Handyphone System (PHS) was standardised by the Telecommunications Technical Committee of Japan for microcellular and indoor use [Padgett95].

PHS uses TDMA/FDMA/TDD with four full duplex channels on each 300 kHz channel. The modems use $\pi/4$ QPSK with coherent demodulation. The channel data rate is 384 kbit/s and each of the four, full duplex channels, implemented using two timeslots out of eight in a frame, support a bearer data rate of 32 kbit/s. The speech codec uses 32 kbit/s ADPCM and a CRC code for protection. It is thus seen that PHS is very similar to the TDD version of PACS.

3.2 Wireless Local Area Networks

Wireless LANs (WLANs) extend wireless access to the LAN environment. Two types of architectures are being contemplated: wireless LANs with a fixed infrastructure, and wireless, peer-to-peer networks for ad-hoc groups. Wireless LANs usually operate in an indoor environment and the target rates are high, usually above 1 Mbit/s. Both infra-red and radio based wireless LANs are available. A number of these have been described in the literature [Hayes91, Barry91, Flatman94].

3.2.1 Standards

A number of Wireless LAN standards are in various stages of completion in the USA, Europe and Japan [Hayes91, Barry91, Flatman94].

The **IEEE 802.11 committee** in the USA is developing standards for WLANs which can also coexist with existing wired LANs, at bit rates exceeding 1 Mb/s. A conceptual architecture and reference model for Wireless LANs has been developed [Links94]. Initially, IEEE 802.11 has concentrated on defining a single medium-independent Media Access Control (MAC) specification. The Distributed Foundation Wireless MAC (DFWMAC) uses Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA).

Three classes of traffic are supported. In decreasing priority order, these are:- signalling and contention-free responses, station polling for time-bounded services, and asynchronous contention packets. These classes are separated by imposing a minimum deferment period before transmission following the detection of a free medium, which increases with decreasing priority.

A protocol on top of CSMA/CA provides for contention-free, time-bounded services by providing a contention-free period every superframe period. The support for time-bounded packets allows the provision of real time, audio-visual services over IEEE 802.11 Wireless LANs. The DFWMAC supports a number of physical layer (PHY) specifications. PHY layers are being specified for infrared media, the 915 MHz, 5.8 GHz, and 2.4 GHz radio ISM bands, the 1.9 GHz PCS band and the 5.2 GHz band targeted for HIPERLAN in Europe (see below).

The **European Telecommunications Standards Institute (ETSI)** is developing Wireless LAN standards in Europe [Black92]. Three categories of Wireless LANs were identified:-

Category 1 Medium rate WLANs with typical terminal bit rates in the region of 200 kbit/s and system density less than 1 Mbit/s/hectare/floor. Target applications include non-critical performance data communications, data capture, control and telecommand services. Initial deployment will probably be in the 2.4 GHz ISM band.

Category 2 High rate WLANs with typical 2 Mbit/s terminal bit rates and system density between 3 and 10 Mbit/s/hectare/floor. DECT is a typical offering in this category.

Category 3 Very high performance WLANs with terminal rates above 10 Mbit/s and a system density of 100-1000 Mbit/s/hectare/floor. A HIgh PERformance LAN (HIPERLAN) is being developed by ETSI in this category. The 17.1 - 17.3 GHz band and the 5.150-5.250 GHz band are being studied for HIPERLAN.

Two wireless LAN standards have been developed in Japan, one for medium rates (256 kbit/s to 2 Mb/s) and the other for high rates (>10 Mb/s) [Padgett95]. The former operates in the 2.4 GHz ISM band, whereas the higher rate WLANs operate near 18 GHz and use Quadrature Amplitude Modulation (QAM), QPSK or 4 level FSK.

3.3 Radio Local Loops

In 1992 Bellcore proposed the Wireless Access Communications System (WACS) to provide wireless access to integrated services, such as voice, video and data [Cox92, Padgett95,]. The main target application was the implementation of Radio Local Loops (RLLs) to provide wireless connectivity to the fixed PTT networks.

In the WACS version of the RLL, small base stations separated by about 600 m provide digital wireless access to voice services at 32 kbit/s. The radio interface uses narrowband TDMA with FDMA and FDD. Each simplex channel supports a link transmission rate of 500 kbit/s and a user payload of 320 kbit/s. The user channels are supported via ten timeslots per 2 ms frame. Coherently demodulated QPSK and two-branch diversity at both the transmitter and the receiver, allow 32 kbit/s ADPCM speech transmission without the need of equalisation or FEC. Timeslots can be aggregated to support bearer data rates from 8 kbit/s to 320 kbit/s [Cox92].

Appendix Two - IHC Functionality

THE IHC CODER

```
H261 BASED VIDEO CODERS - (C) SAVIOUR ZAMMIT
COLOUR CODER V 3.1

A - SET CODER PARAMS        N - DISPLAY SEQ COL.
B -                         O -
C - STORE TxBUF             P -
D - TOGGLE DEBUG [0]        Q - QUIT
E -                         R -
F -                         S - DISPLAY SEQ MONO
G - VBR CODER               T - LP FILTER    [0]
H - FBR CODER               U - HP FILTER    [0]
I - LAYERED CODER           V - VIEW
J -                         W -
K - LOAD FRAME              X -
L -                         Y -
M -                         Z - SET FILES

SELECTION ?
```

The **Main Menu** provides access to six sub-menus (options A,G,H,I,V, and Z) besides eight direct functions.

Option C Store the generated code to a file specified through the SET FILES menu.

Option D Toggles the Debug flag which allows debug information to be displayed. In Debug mode the VLC codes generated are stored to a file as text strings.

Option K Load a frame specified via the SET FILES menu into the IMG buffer.

Option N Display the sequence specified via SET FILES menu in 24 bit RGB colour.

Option S Display the sequence specified via the SET FILES menu as monochrome components.

Option T Low pass filter and subsample CIF frames to generate QCIF frames (QCIF mode only).

Option U Post-processing QCIF high pass filter option.

```
PARAMETERS MENU

[0] MOTION ESTIMATION ?   YES

[1] AUTO QUANTIZER ?      NO
[2] MSE with AUTO Quant   0

[3] VBR QUANTIZER         8
[4] Layered QUANTIZER     5

[5] SEQUENCE LENGTH       70
[6] BIT RATE PER FRAME    10240
[7] FRAME DIFFERENCE      1
[8] FIRST FRAME           0

[9] THRESHOLD             0
[A] UPDATE MODE           1
[B] UPDATE WIDTH          3
[C] COLOUR [1] MONO [0]   1
[D] INTRA DC              0

[Q] QUIT

SELECTION ?
```

The **Parameters Menu** controls the major coding parameters.

Option 0 Toggles Motion Estimation on and off.

Option 1 Selects Auto Quantizer (VBR) mode in which the MB Quantizer is chosen to meet a prescribed MSE.

Option 2 Input the MB target MSE for Auto Quantizer mode.

Option 3 Input the quantizer for VBR mode.

Option 4 Input the quantizer for Layers 2 and 3 in Layered mode.

Option 5 Input the Sequence Length in frames.

Option 6 Bit rate per frame in FBR mode.

Option 7 Skipped frames.

Option 8 First frame to code in the sequence.

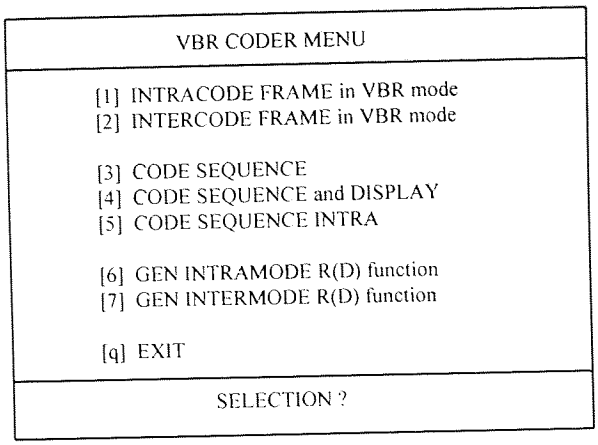
Option 9 Input the threshold value T. If the number of TCs after Quantization is less than T, the block is not coded.

Option A Input the Update mode. Mode[1] force updates N consecutive macroblocks per frame. Mode [2] updates N consecutive macroblocks per GOB.

Option B Input N for use with option A.

Option C Code frames in colour or monochrome mode.

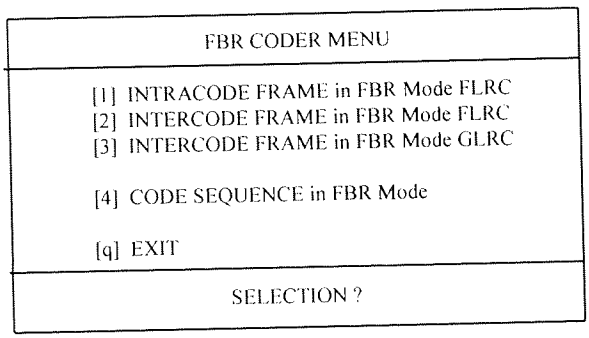
Option D If set to [1], codes the first frame in fixed length DC mode.



The **Variable Bit Rate Coder Menu** controls all VBR coding modes and uses a fixed quantizer per sequence to generate fixed quality, variable bit rate streams.

Option 1 Code IMG buffer in intramode.
Option 2 Code IMG buffer in intermode.
Option 3 Code the sequence specified via SET FILES menu and store to file automatically. The first frame is Intracoded and subsequent frames are Intercoded.
Option 4 As option 3 but with RGB display.
Option 5 Code the sequence in Intramode.

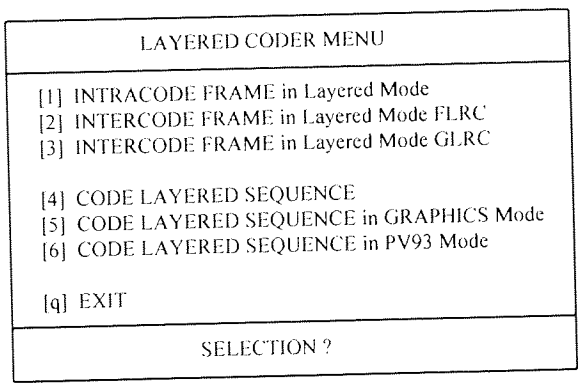
Option 6 Code IMG repeatedly in Intramode varying Q from 1 to 31 and store R(D) information.
Option 7 Intracode the first frame with Q=8 and then code the next frame in Intermode varying Q from 1 to 31. R(D) information is stored to file.



The **Fixed Bit Rate Coder Menu** allows frames to be coded with a fixed bit rate per frame.

Option 1 Intracode the frame in the IMG buffer in FBR mode using Frame Level Rate Control i.e. determine one Q parameter to best meet the target bit rate.
Option 2 Intercode the frame in the IMG buffer in FBR mode using Frame Level Rate Control.

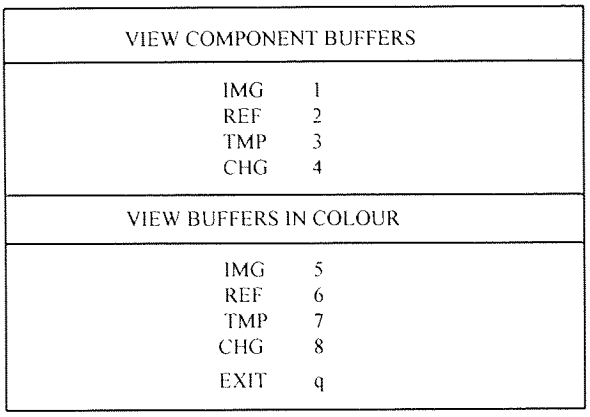
Option 3 Intercode the frame in the IMG buffer in FBR mode using GOB Level Rate Control i.e. determine a new Q per GOB to best meet the target bit rate for the frame.
Option 4 Code the whole sequence in FBR mode. Intracode the first frame with FLRC, intercode the next frame with FLRC and subsequent frames with GLRC.



The **Layered Coder Menu** codes frames and sequences using a three layer coder described in chapter 9.

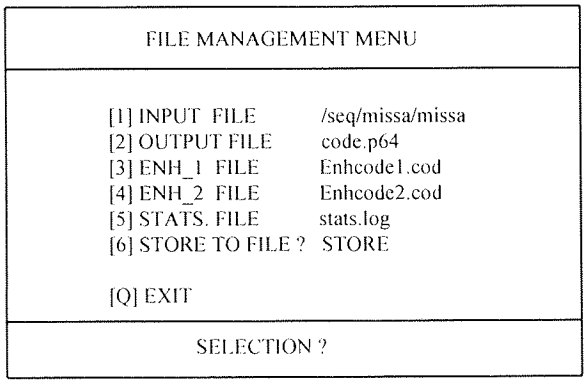
Option 1 Intracode the frame in the IMG buffer in layered mode.
Option 2 Intercode the frame with a fixed bit rate base layer using FLRC.
Option 3 Intercode the frame with fixed bit rate base layer using GLRC.
Option 4 Code a whole sequence using the current three-layer algorithm.

Option 5 As option 4 but with RGB display.
Option 6 Code a sequence using the three-layer algorithm used in [Zammit93].



The **View Menu** allows access to the main frame buffers in the IHC coder. In normal operation the IMG buffer contains the current frame being coded, the REF buffer contains the previously coded frame as decoded locally, the TMP buffer contains the partial results of the currently coded frame and the CHG buffer is used to store the data of those blocks that are conditionally replenished during the coding operation.

The colour mode displays frames in 24 bit RGB colour. The monochrome modes display the Y,U,V components in monochrome.



The **File Management Menu** allows the specification of the input and output video data and statistics files.

Option 1 Sets the template of the input file. The first frame to load consists of the three files:-
/seq/missa/missa0.Y,
/seq/missa/missa0.U,
/seq/missa/missa0.V.

Option 2 Sets the output H.261 stream file for one layer coders and the base layer file for the three layer coder.

Option 3 Sets the output filename for Layer 2 for the three layer coder.

Option 4 Sets the output filename for Layer 3 for the three layer coder

Option 5 Sets the output filename for the Statistics files.

Option 6 When set to STORE, the coders store all generated files to disk automatically.

The IHC Decoder

The **Main Menu** provides access to three sub menus (Options V,X,Z), fourteen decode options, and four miscellaneous commands

```
H261 BASED VIDEO DECODER - (c) SAVIOUR ZAMMIT
Decodes and Displays colour Sequences
V 1.3

A - TEST DECODE                 N - VLC DECODE
B - DECODE FRAME DC             O - ( NO LOG )
C -                             P -
D - DECODE 1st      GR          Q - QUIT
E - DECODE NEXT     GR          R -
F - CONTINUE        GR          S - SHOW SEQUENCE
G - DECODE @        GR          T - DEBUG          [0]
H - DECODE SEQ      GR          U - CHECK          [0]
I - DECODE SEQ                  V - VIEW & SAVE
J - STORE SEQ                   W -
K - DECODE STREAM               X - SIMULATE ERRORS
L - DECODE LC                   Y -
M - DECODE LC       GR          Z - SET PARAMS

INPUT SELECTION
```

Option A In the X windows (UNIX) version, the user can select any macroblock using the cursor, and trace the decoding operations via a 16 by 16 element array display.

Option B Decodes a fixed length DC Intraframe without the possibility of loss of synchronization.

Option D Open the H.261 stream and decode the first Intracoded frame.

Option E Decode the next frame from an open stream.

Option F Decode the sequence from the current open position.

Option G Decode the sequence starting from an arbitrary position.

Option H Decode the sequence in normal mode. Display the frames to screen.

Option I Decode the sequence without display. Useful to catch decode errors in DEBUG mode.

Option J Decode the seqence to hard disk. Catch DEBUG messages but do not display frames.

Option K Locate the picture and GOB headers in a sequence.

Option L Decode a three-layer coded sequence.

Option M Decode a three-layer coded sequence and display.

Option N Decode the stream to VLC level. Generates comprehensive statistics and stores them in a statistics file. (Very Fast).

Option O As option N but does not store to file. Displays macroblock type statistics.

Option Q Quit the Program.

Option S Display a sequence in 24 bit RGB colour.

Option T Toggle debug information generation. Used with non graphics modes to debug stream and/or program.

Option U With the debug option eneabled, enables the generation of VLC codewords and pauses after every macroblock.

```
SAVE AND DISPLAY MENU

[1]  Display rIMG
[2]  Display rREF
[3]  Display rTMP
[4]  Display rCHG
[5]  Save rIMG
[6]  Save rREF
[7]  Save rTMP
[8]  Save rCHG
[0]  EXIT

SELECTION ?
```

The **Save and Display Menu** allows the main frame buffers to be viewed and saved to harddisk.

SIMULATE ERRORS MENU	
[1] Blank MB Row,Col	
[2] Skip Frame	1
[3] Error Mode	OFF
[4] Error Rate 1 in	1000
[5] Error File Name	error.file
[6] Error Packet Size	511
[7] Conceal	OFF
[8] Discard	ON
[9] BCH Mode	OFF
[a] Block Size N	511
[b] Word Size K	493
[c] Correct T	2
[d] Min. Dist. D.	5
[e] Protect Headers	OFF
[f] Fixed Len. DC Frame	OFF
[g] Transcoder Select	0
[h] Transcoder TrT	0
[i] Transcoder Zone	15
INPUT SELECTION	

The **Simulate Errors Menu** encompasses some of the most important options in the IHC codec. It simulates the transmit coder, decoder and a random error channel. It has provisions to read in error files and corrupt the stream as it is decoding.

Option 1 Set a macroblock to zero to investigate temporal error propagation.

Option 2 Skip the first frame to investigate error recovery.

Option 3 Set error mode. [0] No errors, [1] random errors, [2] read in error vectors from error file.

Option 4 Sets error rate in error mode [1].

Option 5 Sets input error filename for use with error mode 2.

Option 6 Sets number of error bits read from the error file, per error file access.

Option 7 Toggles error concealment (GOB level) on and off.

Option 8 Toggles block discard on and off.

Option 9 Sets BCH mode. [0] Ignore BCH, [1] use BCH for detection only, [2] use BCH for correction and detection.

Option A Sets BCH (n,k,t) block size n.

Option B Sets information word length k.

Option C Sets error correction capability t.

Option D Sets the minimum distance of the BCH code.

Option E Prevents errors from corrupting the Picutre and GOB headers.

Option F Instructs the decoder to use fixed word length decode for the first frame.

Option G Turn transcoder simulation on/off.

Option H Prevents coefficient discard from blocks with TrT or less coefficients.

Option I Discards coeffient C_{ij} if (i+j)>Zone.

The **Files and Parameters Menu** allows the various input and output files to be specified, and permits the entry of global parameters.

FILE & PARAMS MENU
[1] Code File = ma128.261
[2] EnhFile1 File = Enhcode1.cod
[3] Input File = /dosg/seq/miss/missa
[4] Output File = seq
[5] Frame File = frame.out
[6] Sequence Len = 70
[7] Save Stats Flag = 0
[8] Save OutPut Flag = 0
[9] COLOUR = 1
[A] nGOBs = 12
[Q] Quit
Selection

Option 1 Specify H.261 input stream name.

Option 2 Specify the Layered coder enhancement layer filename.

Option 3 Specify the input file template for sequence display.

Option 4 Specify the filename for decoded sequence storage.

Option 5 Specify the filename for single frame storage.

Option 6 Set the number of frames to decode.

Option 7 Enable statistics storage to disk.

Option 8 Enable video data storage to disk.

Option 9 Sets Color [1] or monochrome [0] mode

Option A Sets the number of GOBs to 12 for CIF frames, 3 for QCIF, and 10 to enable some 10 GOB frame sequences to be decoded correctly.

Appendix Three
Error Trapping Conditions

An examination of the video multiplex fixed and variable length fields, led to the identification of the following error condition traps. These can be used by the decoder to detect errors independent of the BCH decoder.

For each field in the H.261 video multiplex field, bit errors may or may not be detectable as listed below. That is, given that the decoder is decoding a **known** field, and the field is corrupted, the error may or may not be detectable as listed below.

PSC Errors in the PSC are detectable.

TR All 5 bit patterns allowed so that bit errors cannot be detected, unless the TR sequence is known beforehand (which is not the case usually).

PTYPE All 6 bit patterns allowed. However B6 and B5 are spare and should be set to zero.

PEI Should be 0, but future use of this bit is possible.

PSPARE Undefined, so all received patterns are correct if PEI=1 is allowed.

GBSC Errors in the GBSC are detectable.

GN All 4 bit patterns are allowed except 0000, which is used in PSC. GNs must occur in sequence. Out of sequence GNs imply an error in the current GN or indicate a skipped GOB header.

GQUANT All 5 bit patterns allowed except 00000 which indicates an error.

GEI Should be 0, but future use of this bit is possible.

GSPARE Undefined, so all received patterns are correct if GEI=1 is allowed.

MBA Complete decoding with up to five preceding zeros.
After six zeros:-
00 indicates an error if next byte is not equal to 0x01 (i.e. a start code).
01 indicates an error if the next tribits are not equal to 111 (MB stuffing).
10 indicates an error
11 is correct

MTYPE Complete decoding with up to nine preceding zeros.
An error is detected if there are 10 or more preceding zeros.

MQUANT Fixed length. All five bit patterns valid except 00000.

MVD Complete decoding with up to five preceding zeros.
Following six zeros:-
00 indicates an error
01 indicates an error
10 indicates an error
11 is correct except when next dibit is equal to 00.

CBP Complete decoding with up to seven preceding zeros.
Eight preceding zeros indicate an error.

TCOEFF Complete decoding with up to eight preceding zeros.
Nine preceding zeros imply an error.

When decoding DC coefficient 0000 0000 and 1000 0000 are forbidden.

In addition the following decoding computations can result in error detection:-

Computation of MBA: MBA should be in the range 0<MBA<34

Computation of MVD: MVs should be in the range -16<MV<16

Computation of Run in Zig Zag de-scanning: Result should be < 65;

Appendix Four
Abbreviations and Acronyms

A

AAN	Arai, Agui and Nakajima
ACF	Autocorrelation Function
ACI	Adjacent Channel Interference
ACK	Acknowledgement
ADPCM	Adaptive DPCM
AMPS	Advanced Mobile Phone System
ANL	Access Node Level
ARDIS	Advance Radio Data Information System
ARQ	Automatic Repeat Query
ATDMA	Advanced TDMA
ATM	Asynchronous Transfer Mode.
AWGN	Additive White Gaussian Noise

B

BISDN	Broadband ISDN
BCH	Bose, Chaudhuri, Hocquenghem
BEP	Bit Error Probability
BER	Bit Error Rate
BIU	BS Interface Unit
BLK	Block
BPSK	Binary Phase Shift Keying
BRA	Basic Rate Access
BS	Base Station
BSC	Binary Symmetric Channel
BTRL	British Telecom Research Laboratories

C

CAI	Common Air Interface
CBP	Cell Block Pattern
CBR	Constant Bit Rate
CCI	Co-Channel Interference
CCIR	Consultative Committee for International Radiocommunications
CCITT	International Telegraph and Telephone Consultative Committee
CCU	Cellular Control Unit
CD-ROM	Compact Disk - Read Only Memory
CDF	Cumulative Distribution Function
CDM	Code Division Multiplexing
CDMA	Code Division Multiple Access
CIF	Common Intermediate Format
CIR	Carrier to Interference Ratio
CIU	CCU Interface Unit
CNR	Carrier-to-Noise Ratio
CODIT	Code Division Testbed
COFDM	Coded Orthogonal Frequency Division Multiplexing
CPS	Cellular Packet Switch
CRC	Cyclic Redundancy Check

CSMA	Carrier Sense Multiple Access
CT	Cordless Telephone
CT1,CT2,CT3	Cordless Telephone 1,2 or 3

D

DAPSK	Differential Amplitude Phase Shift Keying
DCA	Dynamic Channel Allocation
DCS1800	Digital Communications System at 1800 MHz
DCT	Discrete Cosine Transform.
DECT	Digital European Cordless Telecommunications
DOS	Disk Operating System
DPCM	Differential Pulse Code Modulation
DPSK	Differential Phase Shift Keying
DQPSK	Differential Quadrature Phase Shift Keying
DTI	Department of Trade and Industry
DTTB	Digital Terrestrial TV Broadcasting
DWHT	Discrete Walsh Hadamard Transform

E

EDC	Error Detection Code
EOB	End of Block field.
EPL	Error Pipe Length
ERC	European Radio-Communications Committee
ERMES	European Radio Message System
ETACS	Extended Total Access Communications System
ETSI	European Telecommunications Standard Institute

F

FBR	Fixed Bit Rate.
FCC	Federal Communications Commission, Inc.
FDD	Frequency Division Duplex
FDM	Frequency Division Multiplexing
FDMA	Frequency Division Multiple Access
FEC	Forward Error Correction
FIR	Finite Impulse Response
FPLMTS	Future Public Land Mobile Telephone System
fps	frames per second.
FSK	Frequency Shift Keying
FTTH	Fibre To The Home
FU	Forced Updating

G

GBN	Go-Back-N
GBSC	Group of Blocks Start Code
GEI	Group of Blocks Extension Field Indicator
GEM	Gilbert Elliot Model
GEO	Geostationary Earth Orbit
GFSK	Gaussian filtered FSK
GMSK	Gaussian Minimum Shift Keying
GOB	Group of Blocks
GOBH	Group of blocks Header.
GQUANT	GOB Quantizer

GN	GOB number
GSM	Global System for Mobile communication
GSM900	GSM at 900 MHz
GSPARE	GOB Spare field

H

H.261	CCITT recommendation - Video Codec for Audio-visual Services at px64 kbit/s
HDTV	High Definition Television
HIPERLAN	High Performance LAN
HRD	Hypothetical Reference Decoder
HVS	Human Vision System

I

IBCN	Integrated Broadband Communications Network
IDCT	Inverse DCT
IEC	International Electrotechnical Committee
IEEE	Institute of Electrical and Electronics Engineers
IETF	Internet Engineering Task Force
IHC	Interactive H.261 Codec
IHD	Interactive H.261 Decoder
INMARSAT	International Maritime Satellite Organisation
INTER	Between frames
INTRA	Within the same frame.
IS-54	EIA Interim Standard for U.S. Digital Cellular (USDC)
IS-95	EIA Interim Standard for U.S. CDMA
ISC	Image Sequence Coding
ISDN	Integrated Services Digital Network
ISE	Intraframe Source Encoder
ISI	Intersymbol Interference
ISM	Industrial, Scientific and Medical
ISO	International Standards Organisation
ITU	International Telecommunications Union
IVS	INRIA Videoconferencing System

J

JDC	Japanese Digital Cellular
JPEG	Joint Picture Expert Group

K

KLT	Karhunen-Loeve Transform

L

LAN	Local Area Network
LAWN	Local Area Wireless Network
LC	Layered Coding
LEO	Low Earth Orbit
LOS	Line of Sight
LOT	Lapped Orthogonal Transform
LSB	Least Significant Bit
LSP	Least Significant Part

M

MAC	Medium Access Control
MAD	Mean Absolute Difference
MAN	Metropolitan Area Network
MAVT	Mobile Audio Visual Terminal
MB	Macroblock
MBA	Macroblock Address
MBE	Macroblocks in Error
MBH	Macroblock Header
MBS	Mobile Broadband System
MC	Motion Compensation
MCFI	Motion Compensated Frame Interpolation
MCP	Motion Compensated Prediction
MDPSK	M-ary Differential Phase Shift Keying
MDR	Multiservice Dynamic Reservation
MEO	Medium Earth Orbit
MEP	Macroblock Error Pipe
MOS	Mean Opinion Score
MPEG	Motion Picture Expert Group
MPSK	M-ary Phase Shift Keying
MQAUNT	Macroblock Quantizer
MS	Mobile Station
MSB	Most Significant Bit
MSD	Mean Square Distance
MSE	Mean Square Error
MSK	Minimum Shift Keying
MSP	Most Significant Part
MTBCL	Mean Time Between Cell Loss
MTYPE	Macroblock Type
MV	Motion Vector
MVD	Differential Motion Vector Field (H.261)
MVx	Horizontal Motion vector component.
MVy	Vertical Motion vector component.

N

NAK	Negative Acknowledgement
nMC	No Motion Compensation
NMT	Nordic Mobile Telephone
NTSC	National Television Standards Committee

O

OFDM	Orthogonal Frequency Division Multiplexing
OSI	Open Systems Interconnection

P

p64	PVRG's H.261 public domain software codec
PACS	Personal Access Communication System
PAL	Phase Alternation Line-by-Line
PAM	Pulse Amplitude Modulation
PBX	Private Branch Exchange
PCM	Pulse Code Modulation

PCN	Personal Communication Network
PCS	Personal Communication System
PEI	Picture Header Extension field Indicator
PER	Packet Error Rate
PHS	Personal Handy Phone System
PLVQ	Pyramid Lattice Vector Quantization
PMR	Private Mobile Radio
PMV	Predicted with Motion Vectors
PMVnQC	Predicted with Motion Vectors but no Quantized Coefficients
PnMV	Predicted with no Motion Vectors
POCSAG	Post Office Code Standard Advisory Group
PRMA	Packet Reservation Multiple Access
PSC	Picture Start Code (H.261)
PSD	Power Spectral Density
PSK	Phase Shift Keying
PSN	Packet Sequence Number
PSNR	Peak Signal-to-Noise Ratio
PSPARE	Spare field in Picture Header (H.261)
PSTN	Public Switched Telephone Network
PTYPE	Picture Type (H.261)
PVRG	Portable Video Radio Group (Stanford University)

Q

Q	Quantization parameter.
QAM	Quadrature Amplitude Modulation
QCIF	Quarter Common Intermediate Format
QDPSK	Quadrature Differential Phase Shift Keying
QPSK	Quadrature Phase Shift Keying

R

RACE	Research on Advanced Communications in Europe
RCPC	Rate Compatible Punctured Convolutional Codes
RGB	Red, Green and Blue
RMS	Root Mean Square
RPC	Radio Propagation Channel
RS	Reed Solomon code.

S

SAW	Stop-And-Wait (ARQ)
SBC	Sub-Band Coding
SEP	Symbol Error Probability
SER	Symbol Error Rate
SIR	Signal-to-Interference Ratio
SNR	Signal-to-Noise Ratio
SQR	Subjective Quality Rating
SUR	Subjective Usefulness Rating

T

TACS	Total Access Communications System
TBD	Trace Back Decoder
TCM	Trellis Coded Modulation
TDD	Time Division Duplex

TDM	Time Division Multiplexing
TDMA	Time Division Multiple Access
THL	Transport Highway Level
TIA	Telecommunications Industry Association
TIU	Trunk Interface Unit
TR	Temporal Reference

U

UMT	Universal Mobile Telephone
UMTS	Universal Mobile Telecommunication System

V

VBR	Variable Bit Rate.
VLC	Variable Length Code
VLD	Variable Length Decoder
VLSI	Very Large Scale Integration
VMC	Video Multiplex Coder
VMS	Video Multiplex Stream
VQ	Vector Quantization
VSB	Vestigial Side-Band

W

WACS	Wireless Access Communication System
WAN	Wide Area Network
WARC	World Administrative Radio Conference
WIN	Wireless Indoor Network
WIU	Wireless Interface Unit
WLAN	Wireless Local Area Network
WN	Wireless Network