Multimedia Communications over Mobile Packet Networks

S. N. Fabri

Submitted for the Degree of Doctor of Philosophy from the University of Surrey

Centre for Communication Systems Research
School of Electronics, Computing and Mathematics
University of Surrey
Guildford, Surrey GU2 7XH, UK

July 2001

© S.N. Fabri 2001
Summary

This thesis describes several concepts associated with the transmission of multimedia services over mobile radio access networks. The error performance and traffic requirements of real-time video transmission over the General Packet Radio Services (GPRS) access network and its successor Enhanced-GPRS is examined. In view of this, video error resilience techniques which exploit channel prioritisation mechanisms are introduced with a view to increasing the robustness of received video sequences encoded with MPEG-4 to channel errors. These include stream prioritisation using unequal error protection and region-of-interest prioritisation for use in multiparty communications and streaming applications. A new forward-error correction scheme for EGPRS which uses iterative serially-concatenated convolutional-Reed Solomon codes is designed and is shown to significantly enhance the error performance for real-time services. A study of the use of backward error correction mechanisms when transmitting streaming multimedia services is carried out, and a new retransmission scheme is introduced to increase the capacity of the radio access network when supporting streaming services.

Key words: Telecommunications, Mobile Communications, Multimedia, Video

Email: s.fabri@eim.surrey.ac.uk

WWW: http://www.eim.surrey.ac.uk/ccsr
Acknowledgments

I would like to thank my family, supervisors, and friends for their patience and support.
## Contents

Summary ................................................................................................................................................... i  

Acknowledgments ................................................................................................................................ ii  

Contents ................................................................................................................................................. iii  

Glossary of Terms .................................................................................................................................. vii  

1 Introduction .......................................................................................................................................... 1  
  1.1 Objectives .................................................................................................................................... 2  
  1.2 Thesis Overview ...................................................................................................................... 4  

2 Enabling Technologies ....................................................................................................................... 6  
  2.1 Multimedia Coding .................................................................................................................... 6  
     2.1.1 Video Representation .......................................................................................................... 6  
  2.2 Video compression technology ................................................................................................. 7  
     2.2.1 MPEG-4 Video and Systems ................................................................................................ 8  
     2.2.2 Error Resilience .................................................................................................................. 8  
     2.2.2.1 H.263 Error Resilience ................................................................................................. 9  
     2.2.2.2 MPEG-4 Error Resilience ............................................................................................ 9  
     2.2.2.3 Adaptive Intra Refresh (AIR) ..................................................................................... 11  
     2.2.3 Error Concealment ............................................................................................................ 11  
     2.2.4 Video transport mechanisms ............................................................................................ 11  
     2.2.5 Assessment of Received Video Quality ......................................................................... 12  
     2.2.6 Real-Time Transport Protocol ......................................................................................... 13  
  2.3 Packet Mobile Communication Systems .................................................................................. 14  
     2.3.1 3GPP Service Requirements ............................................................................................ 15  
     2.3.2 GPRS and HSCSD ............................................................................................................. 18  
     2.3.3 EDGE – Enhanced Data Rates for GSM Evolution ..................................................... 21  
     2.3.4 The GSM/EDGE Radio Access Network (GERAN) ..................................................... 22  
  2.4 Support of voice over packet networks .................................................................................. 22  
  2.5 Conclusion ................................................................................................................................. 25  

3 Design and development of GPRS and EGPRS radio access simulator ............................ 27  
  3.1 Introduction .............................................................................................................................. 27  
  3.2 GPRS Model Description ........................................................................................................ 27
4.3.2.1 Simulation Conditions................................................................................................64
4.3.2.2 Effect of bit errors ......................................................................................................65
4.3.2.3 Channel Coding Scheme Selection..........................................................................67
4.4 Video Communications over EGPRS ....................................................................................70
4.4.1 Traffic Characteristics ......................................................................................................70
4.4.2 Error Performance .............................................................................................................74
4.5 Conclusions...............................................................................................................................76
5 Video Prioritisation Techniques ..................................................................................................78
5.1 Introduction...............................................................................................................................78
5.2 Intra-Stream Video Stream Prioritisation ...............................................................................78
5.2.1 Prioritisation Technique .................................................................................................80
5.2.2 Use in Mobile Networks ................................................................................................81
5.2.3 System Performance.........................................................................................................82
5.3 Prioritisation for multi-party conferencing.............................................................................85
5.3.1 System Architecture.......................................................................................................86
5.3.2 Traffic Characteristics ...................................................................................................87
5.3.3 Prioritisation on multiple bearer channels ......................................................................89
5.3.4 Scene Segmentation ..........................................................................................................91
5.3.4.1 Segmentation algorithm for videoconferencing applications.................................92
5.3.4.2 Performance of scene segmentation ..........................................................................92
5.4 Differential Quantisation .........................................................................................................94
5.4.1 Video Conferencing.......................................................................................................95
5.4.2 Football Sequences ........................................................................................................96
5.4.2.1 Segmentation of Football Sequences ........................................................................97
5.5 Conclusion...............................................................................................................................101
6 EGPRS bearer channel optimisation for real-time video services......................................102
6.1 Overview.................................................................................................................................102
6.2 Limitations of EGPRS bearers ..............................................................................................102
6.3 Channel Coding Scheme Requirements...............................................................................103
6.4 Selection of Channel Coding Techniques ............................................................................104
6.4.1 Operation of the Maximum Aposteriori Algorithm.....................................................106
6.4.2 Soft Decision Decoding of Reed Solomon Codes.......................................................108
6.4.3 Design of an Iterative Decoder.....................................................................................110
6.4.4 Comparison between MAP and Viterbi Decoder ....................................................111
6.4.5 Interleaver Design..........................................................................................................113
6.4.6 Comparison of Soft and Hard Decision decoding ......................................................116
Glossary of Terms

3GPP 3rd Generation Partnership Project
8-PSK 8-Phase Shift Keying
AIR Adaptive Intra Refresh
ATM Asynchronous Transfer Mode
AVO Audio-Visual Object
AWGN Additive White Gaussian Noise
BCH Bose-Chaudhuri-Hocquenghem
BER Bit Error Rate
BSSGP Base Station System GPRS Protocol
BTS Base Transceiver Station
CIF Common Intermediate Format
DCT Discrete Cosine Transform
DMIF Delivery Multimedia Integration Framework
DQ Differential Quantisation
EDGE Enhanced Data Rates for GSM Evolution
EFR Enhanced Full Rate
EGPRS Enhanced GPRS
ETSI European Telecommunications Standardisation Institute
FIFO First-in, First-out
GERAN GSM/EDGE Radio Access Network
GGSN Gateway GPRS Support Node
GMSK Gaussian Minimum Shift Keying
GPRS General Packet Radio Service
GSM Groupe Speciale Mobile or Global System for Mobile Communications
GTP Gateway Tunnelling Protocol
GW   Gateway
HDTV  High-Definition Television
HSCSD High Speed Circuit Switched Data
HTTP  Hypertext Transfer Protocol
IETF  Internet Engineering Task Force
IFH   Ideal Frequency Hopping
IMT-2000 International Mobile Telecommunications-2000
IP    Internet Protocol
ISDN  Integrated Services Digital Network
ITU   International Telecommunications Union
LLC   Link Layer Control
MAC   Medium Access Control
MAP   Maximum Aposteriori
MB    Macroblock
MCS   Modulation-Coding Scheme
MCU   Multipoint Control Unit
MLSE  Mean Least Squares Error
MP3   MPEG-1 Layer 3
MPEG  Motion Picture Experts Group
MSC   Mobile Switching Centre
NFH   No Frequency Hopping
PDTCH Packet Data Traffic Channel
PDU   Protocol Data Unit
PRMA  Packet Reservation Multiple Access
PSNR  Peak Signal to Noise Ratio
PSTN  Public Switched Telephone Network
QCIF  Quarter Common Intermediate Format
QoS   Quality of Service
Glossary of Terms

RAB Radio Access Bearer
RLC Radio Link Control
RNS Radio Network Subsystem
RS Reed-Solomon
RTP Real-time Transport Protocol
RVLC Reversible Variable-Length Codewords
SAACH Slow Associated Control Channel
SAD Sum of Absolute Differences
SGSN Serving GPRS Support Node
SMG Special Mobile Group
SNDCP Sub-Network Dependent Convergence Protocol
SNR Signal-to-Noise Ratio
SOVA Soft-Output Viterbi Algorithm
SSRC Synchronisation Source
TBF Temporary Block Flow
TCH Traffic Channel
TCP Transmission Control Protocol
TDMA Time Division Multiple Access
TFI Temporary Flow Identification
TS Timestamp or Timeslot
TU Typical Urban
UDP User Datagram Protocol
UEP Unequal Error Protection
UMTS Universal Mobile Telecommunications System
UTRAN UMTS Terrestrial Radio Access Network
VoIP Voice over IP
WCDMA Wideband CDMA
Introduction

Chapter 1

1 Introduction

The scope of the work described in this thesis is summarised appropriately enough in its title, 'Multimedia Communications over Mobile Packet Networks'. Multimedia is one of the more over-used and misused terms in communications engineering. A search on the IEE's INSPEC database of scientific publications returned over 22,300 documents published in the previous five years only. Multimedia means different things to different people. The Cambridge Dictionary of International English defines it as the use of a combination of moving and still pictures, sound, music and words, esp. in computers or entertainment. A slightly more loose interpretation of the term is adopted in this thesis, where multimedia communications will be used to refer to the transmission of any of video, audio, speech or images, both jointly as well as separately. As such, it is used as a collective word, encompassing the different media types, rather than specifically requiring the use of more than one media type simultaneously. What however distinguishes multimedia communications from any other form of data communications is that the received information is converted immediately into a non-textual form for playback by an appropriate application to which it is intrinsically linked.

The drive towards extending the range of services that may be offered to the end-user is set to increase the role played by multimedia communications technology in future mobile communications systems. Indeed, communications is no longer restricted to voice-only telephony services, but is now due to encompass a wide range of different services all being offered concurrently and in real-time to the user. These services are expected to include high-quality audio, speech, video, computer-generated graphics and animations, together with a whole host of applications and applets which will be used to integrate all these media components together so as to provide new value-added services.

The adopted definition on multimedia communications raises two characteristic features that will be a recurring theme throughout this thesis. One is the immediacy of use of the received information. Once transmitted, there are certain time constraints which must be met for the transmitted information to retain relevance. In conversational speech applications, such delay
constraints are very tight, whereas in broadcast television-type services, a latency of up to a few seconds may be acceptable. The second feature is perceptual quality of the received information. The quality of received multimedia streams is not best characterised by how many bits are incorrectly received, but rather on more subjective measures such as intelligibility and faithfulness of reproduction.

The second key term in the title is 'Mobile'. This refers to the capability of a user to access a range of communications and information services independently of his location. This means that a subscriber to a set of services is no longer tied down to his service access point, such as his fixed telephone line. Mobility may be implemented in either of two ways. The first, 'pseudo-mobile' approach is to allow a user to access information related to his subscription to communications services at any of a large number of geographically distributed fixed access points. The second approach is to make use of wireless access technologies to allow the user to retain connection with the host network while on the move, without having to physically connect to the network at some access point.

These two trends in communications, namely that of greatly extending the breadth of services and of releasing the user from his tether are however in sharp conflict with each other. On one hand, an enhancement to the range of services on offer requires an increase in the applications’ throughput requirements. On the other hand, the need for mobility entails the use of time-varying error-prone radio access channels that are throughput-limited. The resulting situation is one in which the requirements of these trends are pulling in opposite directions, as if in a tug-of-war. This project therefore aims to provide the engineer's approach to solving some of the problems, namely to move the goalposts, or to continue the tug-of-war analogy, to lengthen the rope, thereby allowing both teams to gain ground. A number of techniques which will allow for more application-sensitive networks and more network-sensitive applications will be developed and examined.

### 1.1 Objectives

The main objective of this thesis is to demonstrate that if end-to-end mobile multimedia communication systems are to be designed in such a way so as to offer flexibility of services to the user while maintaining network and spectrum efficiency, a near-holistic design approach must
be taken, one which encompasses the needs of both the transmission channels as well as of the end-applications.

This will be achieved by examining the requirements placed upon multimedia communications applications by the characteristics of the mobile radio channel. The applications scenarios that will be evaluated will be primarily conversational-type services which due to their low-latency requirements place the most stringent demands upon the underlying access network. However one-way streaming services will also be described, and a comparative study will be carried out. Although multimedia communications applications may broadly be classified into three main categories, speech, audio and video, the most demanding in terms of throughput requirements and the susceptibility of compressed bitstreams to channel errors is undoubtedly video communications. Consequently the work described in this thesis will be focussed on meeting the demands and requirements for 'acceptable' quality video communications, as a system which can support real-time video services will also be able to act as a bearer for speech and audio.

The first objective of the work described in this thesis to design techniques to exploit the inherent diversity in the subjective importance of different sections of compressed media will be examined, with a view to increasing the robustness of the transmitted information, as well as improving the perceptual quality of the received streams. The design requirements of these prioritisation techniques are that they must be capable of being employed over standard radio bearers of 3rd Generation mobile communications systems. In addition, they should require minor, if at all any, modifications to the media compression syntaxes. These constraints should facilitate adoption of media prioritisation techniques over mobile networks, as well as preventing a proliferation of dissimilar and incompatible media compression techniques.

The second main objective is to design optimised radio bearers for transmitting multimedia information over packet radio channels. Mobile packet access offers the greatest service flexibility as IP can act as a common service interface across a wide range of underlying networks. An all-IP model has consequently been selected for the long-term evolution of 3G networks, and the work described in this thesis will focus upon the performance of mobile packet access. Research carried out by the author has already demonstrated the viability and efficiency improvements that can be obtained in providing voice services over mobile packet radio bearers [FABR-99]. Experiments will be carried out to determine the resource allocation requirements and the effect of interference upon received quality for video transmission over standard packet radio access bearers. In particular, the properties and suitability of the physical link layer and data link layers for the
transmission of multimedia services will be examined. Enhancements to the schemes examined will be proposed, with emphasis being made on the forward and backward error correction mechanisms employed.

1.2 Thesis Overview

The background to the work described in this thesis is given in Chapter 2, where the enabling technologies, particularly the multimedia compression standards and communications techniques employed, as well as the mobile communications systems which act as radio bearers are described. Research carried out into the development of suitable techniques for the transmission of voice over packet networks is also introduced.

The mobile packet communications system employed for the experiments carried out in this thesis was ETSI's General Packet Radio Service (GPRS) and its successor Enhanced GPRS (EGPRS) which is a TDMA access technology for the ITU's IMT-2000 3rd Generation mobile communications programme. Chapter 3 describes the design of a GPRS and EGPRS Physical Link Layer model for use in multimedia communications experiments. Both models are validated against reference performance results given to the relevant standardisation workgroups. An appropriate data flow model for the protocol layers across the GPRS radio access interface is described, as is the design of a real-time E/GPRS radio access emulator implementation for use in conjunction with real-time multimedia communications applications.

The bit error patterns obtained under several different propagation conditions are used in Chapter 4 to assess the received quality of video streams compressed with MPEG-4 as transmitted over GPRS and EGPRS radio access networks. The effect of error resilience tools in reducing the effect of channel errors upon the visual quality of the received sequences is examined, as is the role played by rate control mechanisms in adapting to the throughput constraints laid down by the radio bearers. The video quality that can be achieved at different resource allocation levels and at various levels of channel interference is determined.

Chapter 5 investigates the use of media stream prioritisation as a technique for enhancing the robustness and perceptual quality of received video streams. Three main techniques are proposed. The first technique applies the principle of unequal error protection to MPEG-4 bitstreams by
exploiting the data partitioning tools available to MPEG-4. This allows for the more sensitive parameters within the video syntax to be transmitted over different UMTS bearer channels. The two other prioritisation techniques proposed exploit the dissimilar levels of importance of different areas of a received video scene. Two different approaches are used, one which employs the object-oriented coding properties of MPEG-4 in order to represent different areas of a scene as different Audio-Visual objects, and another which applies different levels of quantisation resolution directly to different areas of the scene. These two techniques are shown to provide significant perceptual quality improvements for videoconferencing applications and delivery of sports footage.

Enhancements to the EGPRS radio data bearers for the support of real-time services are introduced in Chapter 6. It is demonstrated that although the rate punctured convolutional codes perform efficiently for the transfer of data when using backward error correction, the coding gain at the residual error rates required to support video services is not sufficient to allow for the transmission of video services at interference levels comparable to those which can support speech services. A new concatenated convolutional-Reed Solomon (RS) code is introduced which makes use of the same convolutional codes as employed by EGPRS, thereby allowing for it to be applied retroactively to terminals without modifications to the relevant channel coding specifications. Significant improvements in the coding gain are observed when using a concatenated Maximum a Posteriori (MAP) convolutional decoder with a soft-input soft-output Dorsch RS decoder, particularly when iterative decoding is used. The resulting improvements in the received video quality are also examined.

Chapter 7 examines the issues involved in providing streaming media services as compared to real-time conversational services. The quality that can be obtained in transmitting audio and video services over EGPRS access networks is examined, and a new retransmission scheme which allows for some residual errors to be present in the received packets is introduced. This is shown to increase the effective throughput of the radio bearer channel at given interference conditions, while providing only a slight degradation in quality. An enhancement to this scheme, referred to as differential thresholding is shown to provide improved performance when coupled with the stream prioritisation techniques developed in Chapter 3, and is shown to improve the received quality of both audio and video streams.
Chapter 2

2 Enabling Technologies

2.1 Multimedia Coding

The requirements of multimedia communications systems place two major demands upon the performance of underlying access networks. In the case of transmission of video sequences, high levels of subjective visual quality are achieved only at relatively high data rates, while real-time interactive applications such as videoconferencing or virtual meeting rooms require strict end-to-end delay quality of service guarantees. Multimedia coding technology provides techniques for relaxing these requirements on the transmission systems by offering two main solutions. On one hand, much research is being carried out into increasing the transmission efficiency and reducing the storage requirements of video and audio sequences by eliminating much of the redundancy in media sequences. In the meantime, other parallel research activities are being carried out into minimising the effect of network and communications constraints on the quality of decoded media streams. These areas of activities are introduced in this Chapter.

2.1.1 Video Representation

For many years video technology was based on analogue storage and transmission technologies, with all the associated quality and processing drawbacks associated with analogue signals. Digital video, on the other hand provides for a more robust representation of sequences and allows for greatly enhanced storage, manipulation and transmission capabilities. There has however been one serious bottleneck restricting the implementation of video-capable devices on computers and communications systems, namely the large amount of information inherent to a video sequence. Compared to digital audio signals, video sequences are notoriously memory-hungry. For example, the ITU-R 601 standard for digital video representation for use in broadcast TV environments specifies a frame size of $720 \times 576$ pixels at a temporal rate of 50 fields/second. This video format requires a raw data rate of 165 Mbps, and the higher quality High-Definition TV standard ($1440 \times 1050$ pels) produces a rate in excess of 540 Mbps. These rates are several orders of magnitude in excess of the capabilities of any mobile or wireless access technologies, and even using the lower quality Common Intermediate Format which uses a frame size of $360 \times 288$ at a rate of 30 frames/second requires a rate of approximately 36 Mbps. These figures clearly
demonstrate the importance of efficient video compression algorithms which can reduce throughput levels to those that can be supported by mobile networks.

2.2 Video compression technology

Although a wide range of compression technologies have been developed for the efficient representation of video sequences, the video coding scene is currently dominated by two major standards which are acceptance in the communications industry. The ITU-T has developed the H.263 set of standards, currently in the H.263++ guise, which essentially describe a video coding scheme for use in low bitrate communications.[ITU-00, RJK-96]. The International Standards Organisation (ISO) has completed version 2 of the MPEG-4 audio-visual codec.[ISO-N2202, SCHA-98] These two standards employ very similar techniques to reduce the throughput requirements to transmit video sequences. Both systems employ a hybrid DCT structure in which 8x8 blocks are encoded using the Discrete Cosine Transform. This transform compacts the energy of the blocks into the lower frequency components, thereby allowing for the higher frequency components to be discarded during the quantisation process with relatively little visual degradation. In both systems, prediction is used in most of the frames so that it is the energy difference which is transmitted, rather than the absolute value of each frame is encoded, thereby providing for more efficient compression. In addition, motion estimation on a 16x16 macroblock (MB) basis is used to determine by how much each Macroblock has moved between consecutive frames. This allows for prediction to be assisted by motion compensation, thereby reducing the number of bits required to encode the difference between successive blocks. Variable length codewords are then employed to perform lossless entropy coding on the output bitstream. This process can be seen in Figure 2.1.

![Image of DCT-based video encoder](https://example.com/dct_encoder.png)

Figure 2.1 Layout of DCT-based video encoder
2.2.1 MPEG-4 Video and Systems

Although the video coding aspects of MPEG-4 are very similar to H.263, the scope of MPEG-4 is much wider, and as may be argued, much more complicated. MPEG-4 is an object-based codec, in which the basic unit from which audio-visual scenes are composited is the audio-visual (AV) objects. These may be video objects, such as a section of a scene or computer-generated graphics, or audio objects such as speech or a background soundtrack. The methods in which these objects are assembled and multiplexed into one or several bitstreams and then demultiplexed and assembled at the decoder is described at the system level of MPEG-4 [ISO-N2201]. The MPEG-4 systems level specifies a two-layer multiplex system. AV objects are encoded into elementary streams. Elementary streams with the same QoS requirements are then combined to form FlexMux streams in the FlexMux layer, which also contains mechanisms for clock recovery and synchronisation with other AVOs. These streams are then forwarded to the TransMux (transport multiplex) layer for transmission over different networks. This layer is not specified within the context of MPEG-4, but existing transport protocols such as UDP/IP and ATM (AAL-5) may be used. In order to aid integration between these two layers, the Delivery Multimedia Integration Framework (DMIF) is being developed to allow different MPEG-4 terminals to be able to communicate with each other using a common transport protocol. The system that will be investigated within this thesis is that described in [LEAN-99] where methods for encapsulating MPEG-4 streams within RTP payloads are described. Potential applications for the use of the object-oriented coding technique are described in a Chapter 5.

2.2.2 Error Resilience

When transmitting information in a packet-based mobile environment, two main sources of error occur. Packet loss caused by header corruption and delay due to congestion in the core switching network result in the occasional loss of large, contiguous sections of the information bitstream, while fading, interference and multipath effects in the mobile channel result in individual bit errors in the received signal. In addition, increasing the compression performance of video coding algorithms inevitably results in the reduction of the redundancy present in the encoded bitstreams, a process which has the undesirable side-effect of increasing the susceptibility of the media streams to errors. Variable-length coding renders the compressed bitstream very brittle to channel errors as the decoder can easily lose synchronisation and be unable to recognise the break points between successive frames. Predictive coding and motion compensation, on the other hand, have the effect of propagating the effect of errors, both in time as well as spatially. Several techniques have been proposed and have been implemented to tackle these undesirable effects. The most basic error-resilience option is to periodically insert INTRA-coded frames into the bitstream.
These do not make use of any predictive information, and consequently prevent the propagation of errors throughout the video sequence.

### 2.2.2.1 H.263 Error Resilience

Wenger et al [WENG-98] describe the improvements in error resilience performance of H.263 offered by the annexes added in Version 2 of the recommendation (H.263+). These optional modes include the use of a BCH forward error correction checksum for error detection and use of independent segment decoding to limit the temporal propagation of errors. It is also possible to select which frame to use as a reference for prediction. When back-channel signalling is used, the encoder can keep track of the last correctly-decoded picture at the decoder, and can produce a predicted frame from an uncorrupted frame should any frame at the decoder be corrupted. This helps to minimise the propagation of errors. These error-resilience techniques are also used within the context of H.324/M (Mobile) [ITU-98] which provides a framework for multiparty multimedia communications in a circuit-switched environment which is susceptible to channel errors [RABI-98]. The error resilience performance of H.263 was further improved with the addition of extra annexes in the H.263++ release [ITU-00]. These included the Data Partitioned Slice Mode which separates the motion vector data from the DCT coefficients in a way similar to that employed by MPEG-4 and the use of header repetition techniques which allow the decoder to receive and recover the header information from a previous picture in the event of data loss or corruption.

### 2.2.2.2 MPEG-4 Error Resilience

Although both MPEG-4 and H.263 are based on the same compression technology, MPEG-4 initially offered much more flexibility and a more comprehensive suite of tools than could be found in the first two releases of H.263, particularly in terms of error resilience capability. Although arguably, this does make MPEG-4 rather more complicated than H.263, in the long-term it is much more likely to be adopted for the codec of choice in several applications. This is demonstrated by the fact that most of the leading vendors of video compression software for use in mobile devices [PACK-01], [HANT-01] have selected, or have indicated an intention to select MPEG-4 as the video compression technology of choice.

One of MPEG-4’s most effective error resilience tool is the use of regular resynchronisation markers throughout the video bitstream, dividing it into video packets of practically equal size, each containing an integer number of macroblocks. Due to the variable length nature of the
encoder output, this method leaves some sections of the picture more vulnerable to error than other parts. Packetisation succeeds in limiting the effect of synchronisation loss, and also, by varying packet sizes, can provide multi-rate error resilience. Further improvements in robustness can be achieved by separating the data into two parts. As described in [TALL-98], the different types of data in MPEG-4 differ in their degree of sensitivity to error. Loss of some data types causes greater distortion at the decoder than the loss of other types. The less sensitive texture data is separated from the more sensitive shape and motion data, and a resynchronisation code is inserted between. This technique is referred to as data partitioning and the resulting video frame layout is shown in Figure 2.2. This second resynchronisation code is called a Motion Marker in INTER frames, or a DC Marker in INTRA frames. Any errors occurring in the second partition can usually be successfully concealed, resulting in little visible distortion. As texture data usually makes up the majority of each packet, data partitioning allows errors to occur in a large part of the packet with relatively benign results.

A single error in a variable length texture code may result in the remaining codes in a data partitioned packet becoming unreadable. If a single error occurs at the beginning of a texture partition, the rest of that data must be discarded. MPEG-4 can represent texture data by means of Reversible Variable Length Codewords (or RVLCs), which can be read either forwards or backwards. Thus, if an error occurs in a packet, the decoder can search for the next resynchronisation code and read backwards. In the case of a single error only the macroblock in which the error occurred need be discarded.

![Figure 2.2 MPEG-4 Data Partitioning](image)

![Figure 2.3 Use of Reversible Codewords](image)
2.2.2.3 Adaptive Intra Refresh (AIR)

As already described, MPEG-4 makes use of periodic non-predicted (INTRA) frames to prevent propagation of errors throughout a sequence. Unfortunately, INTRA frames are, on average, much larger than predictively-coded frames. This leads to a number of problems with respect to error resilience. Firstly, the periodic peaks in bit rate can result in a certain amount of data being dropped following an INTRA frame due to link congestion, leading to data loss even in conditions where there are no channel errors. Secondly, INTRA frames are much more sensitive to errors as if an INTRA block is incorrectly decoded and displayed, the visual distortion will be quite obvious and will seriously degrade picture quality. If a single error is found in an INTRA packet it is common to discard all of the data to decrease the probability of displaying an incorrectly decoded Intra block. One technique used to get around these problems is the Adaptive Intra Refresh (AIR) method [WORR-00]. This involves coding a fixed number of INTRA macroblocks in each frame, eliminating the need for Intraframes. This means that the burden of refreshing all the image area in a video sequence is spread throughout all the frames, thereby producing a much smoother output rate, and enhanced robustness to errors.

2.2.3 Error Concealment

Error resilience is only effective in reducing the perceptual effects of channel errors if effective detection and concealment of received errors can take place. [WANG-98] and [SHIR-00] both describe a range of error concealment techniques which may be used to reduce the distortion caused by channel errors. Of particular relevance is the use of postprocessing techniques at the decoder. These include using motion-compensated temporal prediction, which involves replacing a damaged macroblock with its corresponding error-free version in the previous frame taking into account any motion that may have occurred between frames. Other techniques exploit the smoothness properties of most video signals by a block-based energy minimisation technique or by using spatial or frequency domain interpolation.

2.2.4 Video transport mechanisms

Resilience mechanisms to reduce the effect of channel errors or packet loss are only a subset of the suite of tools required to transmit video sequences across lossy, time-varying channels. In general it is required to adapt the throughput of the transmitted sequence to the properties of the communications link which may be subjected to congestion or high channel errors. Two main approaches have been proposed. The encoder and decoder may be designed so as to be scalable [RADH-99]. This requires the encoder to decompose a compressed image into several layers,
including a base layer which contains the lowest quality of the sequence that may be decoded at the receiver end, while all other layers contain spatial or temporal enhancements to the base layer. This allows the network to selectively discard the enhancement layers in the event of congestion or any other form of throughput restrictions, without compromising the decoder’s ability to reconstruct the original video sequence. Another approach, which does not assume any extra capability on behalf of the codecs employs a technique known as transcoding [IWAS-00]. This involves placing an application gateway within the network capable of manipulating the rate of the received video bitstream without actually decoding and re-encoding the compressed signal. These rate manipulation techniques are only effective if coupled with source rate control mechanisms. [SONG-99] describes an enhanced rate-control algorithm for use in packet-switched environments which exploits the inherent motion information within the video sequence to allow for effective source rate manipulation. Other techniques such as that proposed by Liu et al [LIU-98] combine rate control mechanisms with underlying backward error-correction mechanisms such as selective ARQ schemes to produce low-latency rate tracking suitable for use in mobile channels. Although integrated source-channel coding and unequal error protection have been proposed and described in several papers [BATR-98], [HORN-99], the role between channel coding and syntax prioritisation will be examined in more detail in Chapters 5 and 6.

2.2.5 Assessment of Received Video Quality

The metric used to measure the quality of received video quality under different encoding and transmission conditions is the Peak Signal to Noise Ratio (PSNR) as expressed in decibels. This measure is taken on a frame-by-frame basis. The peak signal for each frame consists of the sum of the highest possible luminance value for each pixel in a frame. The PSNR is then obtained by dividing this value by the sum of the differences between the luminance values of corresponding pixel positions in the two sequences. This means that higher quality sequences are those with higher PSNR values, as that would indicate a low difference (or distortion measure) between signals. The PSNR can be expressed as a function of the Root-Mean Squared Error (RMSE) as shown below.

$$\begin{align*}
\text{PSNR} &= 20 \log_{10} \left( \frac{2^{n-1}}{\text{RMSE}} \right) \\
\text{RMSE} &= \sqrt{\frac{1}{M \times N} \sum_{i=1}^{M} \sum_{j=1}^{N} [f(i,j) - \hat{f}(i,j)]^2}
\end{align*}$$

- (2.1)
This expression is true for an $n$-bit image of size $N$ by $M$ pixels.

Although PSNR measurements are used as the standard metric for assessing video sequence quality, they cannot accurately characterise the nature of the distortions and artefacts introduced by channel errors and compression algorithms. For this reason studies are being carried out into the design of predictive models of the human perception of quality so as to be able to develop improved automated quality assessment techniques [MYER-00].

### 2.2.6 Real-Time Transport Protocol

As described in the Introduction, the use of IP as a service interface and layer 3 protocol will be used in all the work described in this thesis. IP networks however do not guarantee delivery of packets, and neither do they provide any guarantee on the time of arrival properties of the packets. This means that not only will the interarrival time of packets vary, but depending on the network involved, it is also likely that packets may arrive out of order. This means that in order to be able to transmit real-time information, some transport-layer functionality must be overlaid on the network layer to provide timing information from which streaming audio may be reconstructed.

One such protocol, which is gaining popularity is the IETF Real Time Transport Protocol (RTP) defined in [RFC-1889]. RTP provides end-to-end network transport functions suitable for applications transmitting real-time data, such as audio, video or telemetry information over multicast or unicast networks. These functions include payload type identification, sequence numbering, timestamping and delivery monitoring. Typically, applications run RTP over UDP (User Datagram Protocol) rather than TCP (Transport Control Protocol), as the latter’s end-to-end retransmission schemes are inappropriate for real-time applications due to the great delays required by even a single retransmission. Note that this does not exclude the use of repeat-request systems at the data-link layer between adjoining nodes, as these may be able to improve the reliability of the network without adding an excessive delay.
RTP provides a number of services which are useful when transmitting multimedia information over an IP network. As packets may be reordered or even lost, the timestamp information and sequence number in the RTP header may be used to allow receivers to reconstruct the timing produced by the source so that the media packets are contiguously played out at the rate at which they were produced at the source. In using RTP to transport speech packets, the timestamps will be obtained from the sampling clock of the input device. In addition the sequence numbers can be used to reorder frames at the receiver and determine which frames have gone missing and therefore perform error concealment by using lost frame reconstruction. The Payload Type field in the RTP header identifies the RTP payload and determines its interpretation by the application. This feature could be used in multi-rate coding schemes where different codecs are used according to the changing channel conditions and variations in the network load. The RTP header is shown in Figure 2.4. Note that the SSRC identifier represents the RTP source port and each RTP session has a unique SSRC address. As is obvious from this diagram, one of the main disadvantages of using RTP is the size of its header, which when combined with the small size of packets required to encapsulate real-time speech, corresponds to a considerable overhead, especially when including the UDP/IP headers which give a total of 40 bytes of headers at the network layer and above.

2.3 Packet Mobile Communication Systems

The services offered by the currently-available GSM 2nd generation mobile communications system are restricted to conversational narrowband speech services and data access at speeds up to 9.6 kbit/s [ALAN-94]. These capabilities are clearly not sufficient to provide real-time
multimedia services including simultaneous video and audio transmission. Access to the network is only available in circuit-switched mode, making it particularly inefficient for services such as email access and web browsing which have very bursty traffic characteristics.

2.3.1 3GPP Service Requirements

These restrictions are the main driving force behind the standardisation efforts into 3rd Generation mobile systems. UMTS (Universal Mobile Telecommunications System) is ETSI's (now 3GPP) candidate system for the ITU's IMT-2000 family of 3G systems. The service requirements and architecture for UMTS are specified in Technical Specification documents TS 22.105 [3GPP-22.105] and TS 23.107 [3GPP-23.107]. One of the most important differences between the service description of 3G systems and that of GSM lies in the definition of Radio Access Bearer (RAB) services. These provide the capability for information transfer between access points and only involve low-layer functions and are characterised by a set of end-to-end characteristics with requirements on Quality of Service. This contrasts with the traditional use of teleservices in GSM such as end-to-end speech telephony or the Short Message Service (SMS) in which the service, or end-application is defined on an end-to-end basis.

This decoupling of the application layer from the bearer services allows for a far greater range of services to be offered over UMTS than has previously been possible over current mobile networks. In order to allow for greatest service flexibility, when negotiating the characteristics of a bearer service both the network's transfer capabilities must be defined as well as the information quality characteristics. These are defined in terms of the maximum transfer delay, the delay variation, the bit error ratio and the data rate between two access points in a given period of time. The relationship between the services to be offered and their corresponding QoS requirements are shown in Figure 2.5. Current 3GPP service requirements stipulate that a terminal should be able to support conversations combining two or more media components between several parties and connections [3GPP-22.101]. These services will be divided into interactive and distribution services. Interactive services will include real-time bi-directional conversational services such as video conferencing, messaging services where combined audio, video, text and data messages may be forwarded to the user's mailbox. Retrieval services will also be allowed, thereby enabling the user to access information from a multimedia information centre. The second category of services, distribution services, can be visualised as a form of television-type service which is broadcast on particular channels and is accessible to all subscribed users.
It is immediately obvious that the service characteristics described above are very similar to the high-end Internet applications currently available. The Internet is the only global communications infrastructure allowing access to a wide range of information archives, with increasing availability of access to multimedia services, such as video conferencing and streaming video. This was acknowledged in [UMTS-2260], where Internet access is seen as being the main user of UMTS data services, and given the ever-increasing dominance of multimedia content within the Internet, of multimedia services within UMTS. Applications which future mobile systems will have to support will include distributed 3D gaming, multiparty conferencing in virtual rooms, telepresence, location-sensitive information services and Internet TV and radio services.
### Table 2.1 3GPP QoS Requirements - from [3GPP-22.105]

<table>
<thead>
<tr>
<th>Operating environment</th>
<th>Real Time (Constant Delay)</th>
<th>Non Real Time (Variable Delay)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>BER/Max Transfer Delay</td>
<td>BER/Max Transfer Delay</td>
</tr>
<tr>
<td><strong>Satellite</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Terminal relative speed to ground up to 1000 km/h for plane)</td>
<td>Max Transfer Delay less than 400 ms</td>
<td>Max Transfer Delay 1200 ms or more (Note 2)</td>
</tr>
<tr>
<td></td>
<td>BER 10-3 - 10-7 (Note 1)</td>
<td>BER = 10-5 to 10-8</td>
</tr>
<tr>
<td><strong>Rural outdoor</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Terminal relative speed to ground up to 500 km/h) (Note 3)</td>
<td>Max Transfer Delay 20 - 300 ms</td>
<td>Max Transfer Delay 150 ms or more (Note 2)</td>
</tr>
<tr>
<td></td>
<td>BER 10-3 - 10-7 (Note 1)</td>
<td>BER = 10-5 to 10-8</td>
</tr>
<tr>
<td><strong>Urban/Suburban outdoor</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Terminal relative speed to ground up to 120 km/h)</td>
<td>Max Transfer Delay 20 - 300 ms</td>
<td>Max Transfer Delay 150 ms or more (Note 2)</td>
</tr>
<tr>
<td></td>
<td>BER 10-3 - 10-7 (Note 1)</td>
<td>BER = 10-5 to 10-8</td>
</tr>
<tr>
<td><strong>Indoor/Low range outdoor</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Terminal relative speed to ground up to 10 km/h)</td>
<td>Max Transfer Delay 20 - 300 ms</td>
<td>Max Transfer Delay 150 ms or more (Note 2)</td>
</tr>
<tr>
<td></td>
<td>BER 10-3 - 10-7 (Note 1)</td>
<td>BER = 10-5 to 10-8</td>
</tr>
</tbody>
</table>

**NOTE 1:** There is likely to be a compromise between BER and delay.

**NOTE 2:** The Max Transfer Delay should be here regarded as the target value for 95% of the data.

**NOTE 3:** The value of 500 km/h as the maximum speed to be supported in the rural outdoor environment was selected in order to provide service on high speed vehicles (e.g. trains). This is not meant to be the typical value for this environment (250 km/h is more typical).

In order to be able to support these demanding applications, UMTS will have to provide much higher bit rates than are provided on current systems. The 3GPP Services and Systems Aspects Technical Specifications document [3GPP-22.105] requires that a single terminal may have access to a total of 144 kbit/s in a satellite radio or rural environments, up to 384 kbit/s in urban radio environments and 2048 kbit/s in indoor and low-range pico-cells. In addition, QoS classes for UMTS are divided into two major categories. Non-real time services will have to meet very stringent BER requirements, with error rates as low as $10^{-8}$, while real-time, constant delay systems may experience errors at rates as high as $10^{-3}$. The QoS requirements in different reception environments are given in Table 2.1.

The set of requirements for such a wide range of envisaged services are catered for by the definition of four main types of Radio Access Bearers. These are outlined in Table 2.2. The required network QoS requirements of the applications envisaged to be used over UMTS are expected to fall within these four main categories.
Enabling Technologies

<table>
<thead>
<tr>
<th>Traffic class</th>
<th>Conversational class</th>
<th>Streaming class</th>
<th>Interactive class</th>
<th>Background class</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>conversational RT</td>
<td>streaming RT</td>
<td>Interactive best effort</td>
<td>Background best effort</td>
</tr>
<tr>
<td>Fundamental characteristics</td>
<td>- Preserve time relation (variation) between information entities of the stream</td>
<td>- Preserve time relation (variation) between information entities of the same stream</td>
<td>- Request response pattern</td>
<td>- Destination is not expecting the data within a certain time</td>
</tr>
<tr>
<td></td>
<td>Conversational pattern (stringent and low delay)</td>
<td></td>
<td>- Preserve payload content</td>
<td>- Preserve payload content</td>
</tr>
<tr>
<td>Example of the application</td>
<td>- voice</td>
<td>- streaming video</td>
<td>- Web browsing</td>
<td>- background download of emails</td>
</tr>
</tbody>
</table>

Table 2.2 3GPP UMTS Service Classes

2.3.2 GPRS and HSCSD

In order to bridge the gulf that currently exists between GSM and UMTS, which is not scheduled for widespread deployment until at least 2003, ETSI developed a family of mobile systems that build on the current TDMA technology, enhancing its capabilities with a view to forming an evolutionary pathway to UMTS. Phase 2+ of the GSM standardisation process has introduced two complementary new mobile networks, HSCSD (High Speed Circuit Switched Data) [GSM-02.34] and GPRS (General Packet Radio Service).

GPRS offers an end-to-end packet transfer network [3GPP-22.060], and as such is the most suitable mechanism for forwarding packetised multimedia streams across the mobile radio access link. GPRS makes use of the same radio interface as the other GSM services, and so may be used together with the currently available services. In order to achieve such coexistence, GPRS introduces a number of new channels, including the Packet Data Traffic Channels (PDTCHs) which are used to transfer the user information. There exist four channel coding schemes used for protecting user information across the radio interface, ranging from a half-rate convolutional code to a transparent scheme, carrying out error detection only [GSM-05.03].

<table>
<thead>
<tr>
<th>Scheme</th>
<th>Code Rate</th>
<th>Radio Block payload (bits)</th>
<th>Data Rate kb/s</th>
</tr>
</thead>
<tbody>
<tr>
<td>CS-1</td>
<td>1/2</td>
<td>181</td>
<td>9.05</td>
</tr>
<tr>
<td>CS-2</td>
<td>=2/3</td>
<td>268</td>
<td>13.4</td>
</tr>
<tr>
<td>CS-3</td>
<td>=3/4</td>
<td>312</td>
<td>15.6</td>
</tr>
<tr>
<td>CS-4</td>
<td>1</td>
<td>428</td>
<td>21.4</td>
</tr>
</tbody>
</table>

Table 2.3 GPRS Channel Coding Schemes

The core network structure of GPRS is radically different from that used in GSM, as it consists of a number of nodes (referred to as GSNs - GPRS Support Nodes) which are essentially packet-based routers [CAI-97]. Within the GPRS network, protocol data units are encapsulated at the
Enabling Technologies

originating GSN and extracted at the destination GSN. In between, these two GSNs, IP is used as the backbone network protocol to transfer PDUs. This means that should the end terminals be using IP-based applications, the protocol architecture would involve the transmission of IP over IP. This is clear from the GPRS user plane architecture as shown in Figure 2.6.

![Figure 2.6 GPRS Core Network Architecture](image)

When discussing GPRS performance, it is important to consider the access and core network capabilities separately, as the restraints that they impose upon end-to-end quality of service are radically different. The GPRS access network, is essentially the same as the current GSM speech
access network, as cells which currently provide GSM speech services, may dedicate timeslots and carriers to GPRS data services. In fact, the voice and data services may be multiplexed seamlessly within the same carriers. Although current GPRS standards allow for up to 8-slot multiplexing per terminal, the first release of GPRS will support only 3-slot multislotting. At CS-2, for example, this will mean that each terminal will be limited to about 33 kbit/s. The capacity restrictions across the radio interface also affect the delay performance across the radio interface. It has been shown in [LARS-98] that the access delay due to channel contention and congestion in GPRS access networks, far exceeds that caused by the effect of channel interference on the GPRS ARQ (Automatic Repeat Request) system. As the GPRS core network is based on IPv4, it can only offer best-effort service classes. As can be seen from the delay and reliability service classes shown in Table 2.4 [3GPP-22.060], even the best delay class can only offer a one-way transfer delay of between 0.5s and 2s, which is greatly in excess of the 200ms round-trip delay necessary for interactive services.

<table>
<thead>
<tr>
<th>Delay Class</th>
<th>SDU size: 128 octets</th>
<th>SDU size: 1024 octets</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean Transfer Delay (sec)</td>
<td>95 percentile Delay (sec)</td>
</tr>
<tr>
<td>1. (Predictive)</td>
<td>&lt; 0.5</td>
<td>&lt; 1.5</td>
</tr>
<tr>
<td>2. (Predictive)</td>
<td>&lt; 5</td>
<td>&lt; 25</td>
</tr>
<tr>
<td>3. (Predictive)</td>
<td>&lt; 50</td>
<td>&lt; 250</td>
</tr>
<tr>
<td>4. (Best Effort)</td>
<td>Unspecified</td>
<td></td>
</tr>
</tbody>
</table>

Table 2.4 GPRS Service Classes

The sister service to GPRS, HSCSD provides a high-speed circuit-switched mobile communications access by allowing a single user to use multiple timeslots at the same time. The same data traffic channels as used in GSM are employed in HSCSD using identical channel coding schemes in such a way that the 4.8 kbit/s, 9.6 kbit/s and 14.4 kbit/s air interface rates offered by TCH/F4.8, TCH/F9.6 and TCH/F14.4 can be combined in a number of ways to produce throughput capabilities of up to 57.6 kbit/s [GSM-02.34] This restriction is placed by the upper limit of the circuit-switched trunking capacity. [GSM-03.34]. Although HSCSD has the advantage of being able to provide real end-to-end delay guarantees which are suitable for conversational services, it does not allow for flexible packet-based access across the radio interface. No further investigation of HSCSD will therefore be carried in this thesis.
2.3.3 EDGE – Enhanced Data Rates for GSM Evolution

Although it is accepted that GPRS should have a considerable impact on the types of data services that can be offered over a mobile network, it still has two major limitations. First, as it employs the same access architecture as GSM systems, the throughput of GPRS access channels is limited to the amount of information bits that can be transmitted using the GMSK carrier at a symbol rate of 270 ksym/second. Secondly, as already described, GPRS currently offers only very poor end-to-end delay guarantees. These limitations are being addressed, both directly and indirectly, in the standardisation work being carried out in the development of EDGE, Enhanced Data Rates for GSM Evolution [FURU-98].

EDGE introduces a new modulation and coding scheme to allow for an increase in the throughput capacity of GSM systems. In order to retain compatibility with current GSM GMSK systems, the symbol rate of 270 ksym/sec is retained, but a 3 bit/symbol 8-PSK modulation scheme is employed. This increases the radio data rate per timeslot from 22.8 kbit/s to 59.2 kbit/s, virtually a three-fold increase. The resulting evolution of GPRS is known as Enhanced-GPRS (EGPRS). These schemes employ both GSMK and 8-PSK modulation methods so as to be able to provide optimum throughput in a wide variety of C/I conditions. In fact, a gain in throughput is obtained with 8-PSK only at high C/I conditions experienced by about 50% of users in a typical urban cell [SMG2-144/99]. EGPRS is based on the same concept as GPRS, and indeed, most of specifications regarding the core network and much of the radio network remain unchanged. However, the different transmission characteristics of 8-PSK-modulated signals requires the use of much more efficient and rapidly-acting link adaptation techniques, together with different block structures to accommodate the necessary modifications in the protocols.
2.3.4 The GSM/EDGE Radio Access Network (GERAN)

The first phase of the EDGE standardisation process which defined the modulation and access schemes was completed in the first half of 2000. Current work is focusing upon developing systems for real-time services over packet-switched networks by adopting many of the features of the UMTS Terrestrial Access Network (UTRAN) and adopting the UMTS service classes and bearer definitions [ERIK-00]. The service definitions and classes described in Section 2.3.1 should therefore be directly applicable to GERAN. The main enhancement that Release 4 and Release 5 of GERAN [3GPP-43.051] provide over E/GPRS is the provision of real-time services in packet-switched environments. This is achieved by adopting the UMTS Core Network Interface, referred to as Iu-ps, and the Radio Access Bearer concept described in Section 2.3.1. This allows the GSM/EDGE radio interface to connect both to a legacy 2G GPRS core network as originally defined for GPRS, as well as to the UMTS core network as shown in Figure 2.7. The main attraction of GERAN over WCDMA is that it allows for a migration path for GSM network operators to support 3G services over the current radio infrastructure using spectrum already allocated to them, while allowing for operation over 200kHz carriers.

![Figure 2.8 GERAN Reference Architecture](image)

2.4 Support of voice over packet networks

In order to meet the UMTS service requirements as outlined above, the GSM/EDGE radio access network must be able to support real-time voice services in packet-switched environments. Operating in packet-switched mode at the radio interface should allow for the exploitation of the bursty characteristics of human speech, thereby increasing radio capacity by implementing
statistical multiplexing techniques. None of the GPRS service classes shown in Table 2.4 provide for quality of service levels sufficient for the provision of low-latency packet speech services. For this reason, much research has been carried out into the evolution of GPRS and EGPRS to be able to provide speech bearers which provide at least the same quality and efficiency as can be achieved with GSM.

In work carried out by the author [FABR-99], a proposed pathway for GPRS evolution for the support of voice services was described. The GPRS uplink Medium Access Control (MAC) operation [3GPP-04.60] was enhanced so as to allow for the prioritisation of speech over data traffic. The Voice over GPRS implementation proposed for supporting voice services required that the master channel be left free for control signalling only and not be allowed to contain any voice or data traffic, preventing any unacceptable access grant delays. The remaining seven slave channels in a given carrier were then divided into Packet Data Traffic Channels dedicated to voice services and others dedicated to data services. The system being proposed does not allow a PDTCH to share voice and data services, as the delay requirements are radically different. The operation of the RLC/MAC protocol is summarised in Figure 2.7. The most important concept of this scheme is its ability to make use of the bursty nature of human voice, while requiring a minimum of signalling overhead. Channel contention only occurs at the beginning of each talkspurt, after which the mobile terminal retains exclusive use of the PDTCH and only releases it when voice activity for that particular talkspurt ceases. This access protocol is an adaptation of the PRMA system described in [GOOD-89]. This makes use of the concatenation of LLC frames as allowed by the current GPRS standards [3GPP-04.64]. Other enhancements required included a rationalisation of the GPRS data payload so as to allow for the encapsulation of EFR-coded speech [SALA-97] within a single GPRS timeslot. This was achieved by introducing a new channel coding scheme modelled upon the GSM TCH/EFS channel which makes use of unequal error protection to optimise the forward error correction to match the error susceptibility characteristics of the compressed speech.

The schemes proposed were shown to provide similar subjective quality in interference-limited channel conditions to the quality that can be obtained using circuit-switched operation, while providing for a significant improvement in the traffic capacity over a single carrier by employing statistical multiplexing. This was demonstrated by carrying out Mean Opinion Scores (MOS) testing on 16 subjects as shown in Figure 2.11. In fact, when seven traffic channels are allocated to voice services, it was observed that 10 users can be accommodated while sustaining a packet loss rate due to the contention mechanism of less than 1%. This represents a statistical
Enabling Technologies

Multiplexing gain of 1.428. Figure 2.10 shows the statistical multiplexing gain, i.e. the number of voice users divided by the number of required channels, as a function of the number of traffic channels allocated to voice traffic. The effects of the proposed access mechanism was also shown to have a negligible effect upon the delay performance of the data traffic.

![Figure 2.9 MAC operation for Voice over GPRS - from [FABR-99]](image)

Figure 2.10 Statistical multiplexing gain obtained using Voice over GPRS
Further research has since been carried out on providing voice services over packet-switched networks, with emphasis being placed on EGPRS. Wu et al [WU-00] showed that if appropriate header protection mechanisms are implemented, significant increases in system capacity may be achieved for Voice over EDGE using the EFR codec when operating in a 3/9 or 4/12 frequency reuse pattern. These results were supported by Samaras et al [SAMA-00] who proposed methods for determining the capacity of packet-switched cellular networks to support voice users. The results obtained indicated that when using 3/9 and 4/12 reuse patterns, packet-switched voice provides greater capacity than circuit-switched services under most propagation conditions. The quality obtained may be further improved by using adaptive codecs such as ETSI’s Adaptive Multirate (AMR) [EKU-99] codec. This was demonstrated in [ERIK-00], where it was shown that an EDGE-AMR combination could provide the same voice quality at a C/I of 4dB as can be achieved by GSM at 10dB.

2.5 Conclusion

This Chapter has given an overview of the technologies and concepts involved in providing multimedia communications over mobile access links. In particular, the role of media compression technologies was discussed, with emphasis being placed on the effect of high error rates on such schemes. GPRS and EGPRS were chosen as the access networks for use in the experiments described in this thesis, as they provide a suitable migration pathway from GSM networks to UMTS. In addition, E/GPRS offers end-to-end packet transfer capabilities, which if
tailored for the demands of real-time services, have been shown to provide significant performance and flexibility advantages over equivalent circuit-switched access networks in the provision of voice services. The remainder of this thesis will therefore examine whether similar advantages may be obtained when transmitting other forms of media services, and which techniques may be used to reconcile the conflicting requirements between the quality of service requirements of multimedia communication systems and the radio access characteristics of mobile packet networks.
Chapter 3

3 Design and development of GPRS and EGPRS radio access simulator

3.1 Introduction

The most critical aspect of transmitting real-time multimedia information over packet-switched mobile networks is the poor error characteristics and bandwidth scarcity of the mobile channel whose throughput and error characteristics are time-varying. This Chapter examines the characteristics of the physical link layers, and details of simulation models used to carry out the experiments described in the rest of the Chapter are described. The performance of the simulator model is validated by comparisons with figures presented in the relevant ETSI SMG (Special Mobile Group) Workgroups, and a set of E/GPRS interference performance results are presented. The resulting error patterns are used in conjunction with an E/GPRS radio interface protocol model which represents the data flow across the E/GPRS protocol layers at the physical interface. This is used to assess the effect of bit and block errors upon different layers in the protocol stack, and consequently upon the received video quality at the mobile terminal. Experiments are carried out to assess the requirements of the radio interface parameters to support 'acceptable' quality video services. These include resource allocation requirements, and the carrier-to-interference levels that need to be maintained.

3.2 GPRS Model Description

The simulator model described in this Chapter characterises the GPRS and EGPRS physical link layer functionality. The main issues tackled are forward error correction, modulation, transmission over fading channels, equalisation and reception and detection of correctable and uncorrectable errors. Power control mechanisms are not implemented in this model.
3.2.1 GPRS Model Overview

The model developed simulates the physical layer characteristics of the channel between a GPRS mobile terminal and a base station. An outline of the simulator is shown in Figure 3.1. It can be seen that the transmitted signal is subjected to a multipath fast fading environment and a single co-channel interferer. In addition, an additive white gaussian noise source is present at the receiver. This allows for the bit error and block error characteristics to be determined for a range of carrier-to-interference (C/I) ratios and signal-to-noise ratios (Eb/N0) using the different GPRS and EGPRS channel coding schemes. Note that a static C/I profile is implemented in this model and no shadowing or slow-fading effects are described in the simulator model. However these may be easily implemented by concatenating the data sets describing the channel bit error characteristics of different, static carrier-to-interference levels. The simulator model was built using the COSSAP stream simulation environment. The modulator and receiver block employed for the GMSK schemes was the standard COSSAP block found in the GSM Reference Design Kit. A decision-feedback equaliser [PROK-95] was used for the 8-PSK modulation-coding.
schemes\(^1\). All other elements of the simulation model were either created by the author or adapted from other lower-level COSSAP blocks.

### 3.2.2 Channel Coding

GPRS employs four channel protection schemes, offering flexibility in the degree of protection and traffic capacity available to the user. Varying the channel coding scheme allows for an optimisation of the throughput across the radio interface as the channel quality varies. The channel coding schemes are defined in GSM 05.03 [GSM-05.03] and are outlined in this section.

#### 3.2.2.1 Packet Data Block Type 1 (CS-1)

The coding scheme used for CS-1 is the same as used for the GSM SACCH (Slow Associated Control Channel). 184 source data bits are coded using a shortened binary cyclic (FIRE) code to give a block of 224 bits, to which 4 tail bits are added. These 228 bits are then encoded using the \(\frac{1}{2}\) rate convolutional encoder used to protected Class 1 bits in the TCH/FS speech channel, thereby producing a block of 456 coded bits. The code polynomial is shown below.

\[
G_0 = 1 + D^3 + D^4
\]

\[
G_1 = 1 + D + D^3 + D^4
\]

#### 3.2.2.2 Packet Data Block Type 2 (CS-2)

The data is forwarded to the encoder in fixed-sized blocks of 271 bits. The first three bits, representing the Uplink Status Flag (USF) are precoded into six bits, and a further 16 parity bits and 4 tail bits are added to the information block. The resulting 294 bits are encoded with the same \(\frac{1}{2}\) rate code as used in CS-1, thereby giving a block of 588 bits which are reduced to 456 bits by means of puncturing.

#### 3.2.2.3 Packet Data Block Type 3 (CS-3)

In this coding scheme the information bits are forwarded to the encoder in blocks of 315 bits. As in scheme CS-2, the three USF bits are precoded into six bits, and 16 parity and 4 tail bits are added, to form a block of 338 bits. Convolutional encoding using the CS-1 encoder produces a block of 676 bits which are reduced to 456 bits by puncturing.

---

\(^1\) The EDGE demodulator and equaliser were designed, implemented and validated by Cyril Valadon, another researcher with CCSR.
3.2.2.4 Packet Data Block Type 4 (CS-4)

This is a transparent scheme in which no forward error correction is carried out on the information bits. The three USF bits are block coded into 12 bits for protection, while 16 parity bits are added to the end of the 431-bit information block for error detection purposes. Once again, the result is a 456 bit radio block.

3.2.3 Interleaving

The GPRS Packet Data Channels are interleaved using a block rectangular scheme identical to that used for the SACCH [GSM-05.03]. In this scheme, the depth of interleaving is four blocks deep. This means that each data block is spread over four consecutive TDMA frames, which is equivalent to 18.46ms.

3.2.4 Channel Decoding

The channel decoder uses a soft-decision Viterbi algorithm for decoding the convolutional codes. In scheme CS-1, the FIRE code is used for both error correction and detection, whereas in codes CS-2, CS-3 and CS-4, the cyclic codes are used for error detection purposes only.

3.3 EGPRS Model Description

The structure of the EGPRS physical link layer model used is similar to that discussed thus far for the GPRS simulator.

3.3.1 EGPRS Channel Coding Schemes

EGPRS supports nine joint modulation-coding schemes, referred to as schemes MCS-1 to MCS-9. MCS-1 to MCS-4 are based on the same GMSK modulation scheme as used in GSM speech services and GPRS PDTCHs, whereas schemes MCS-5 through MCS-9 employ a higher-rate 8-PSK modulation scheme [3GPP-05.04]. A summary of the coding parameters for the EGPRS coding schemes are given in Table 3.1. One very noticeable difference between the schemes used in EGPRS and those employed by the GPRS PDTCHs is that the radio block headers are encoded separately from the data payload. Another difference is that in schemes MCS-7, MCS-8, and MCS-9, two RLC/MAC blocks are inserted into a single radio block. In GPRS a one-to-one block mapping is always maintained. The coding schemes and block formatting for uplink and downlink scenarios vary slightly. In the work presented in this report, schemes and results for the downlink are shown only. There is negligible difference in error performance for the uplink and downlink of the same scheme. As can be seen in Table 3.1, the schemes are divided into three families.
Each of these families has a different unit of payload (37 & 34 for A, 28 for B and 22 for C respectively) into which the data payload is divided. Each radio block may contain either two or four of these payload units, and in the case where four units are present, the data is split into two RLC/MAC blocks. The details of the MCS-1 and MCS-5 scheme which form the basis of all the other punctured codes, and which may therefore be regarded as being representative of the other schemes, are shown below.

<table>
<thead>
<tr>
<th>Scheme</th>
<th>Code rate</th>
<th>Header Code rate</th>
<th>Modulation</th>
<th>RLC blocks per Radio Block (20ms)</th>
<th>Raw Data within one Radio Block</th>
<th>Family</th>
<th>BCS</th>
<th>Tail payload</th>
<th>HCS</th>
<th>Data rate kb/s</th>
</tr>
</thead>
<tbody>
<tr>
<td>MCS-9</td>
<td>1.0</td>
<td>0.36</td>
<td>8-PSK</td>
<td>2</td>
<td>2x592</td>
<td>A</td>
<td>2x12</td>
<td>2x6</td>
<td>8</td>
<td>59.2</td>
</tr>
<tr>
<td>MCS-8</td>
<td>0.92</td>
<td>0.36</td>
<td>8-PSK</td>
<td>2</td>
<td>2x544</td>
<td>A</td>
<td></td>
<td></td>
<td></td>
<td>54.4</td>
</tr>
<tr>
<td>MCS-7</td>
<td>0.76</td>
<td>0.36</td>
<td>8-PSK</td>
<td>2</td>
<td>2x448</td>
<td>B</td>
<td></td>
<td></td>
<td></td>
<td>44.8</td>
</tr>
<tr>
<td>MCS-6</td>
<td>0.49</td>
<td>1/3</td>
<td>8-PSK</td>
<td>1</td>
<td>592</td>
<td>A</td>
<td>12</td>
<td>6</td>
<td></td>
<td>29.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>27.2</td>
</tr>
<tr>
<td>MCS-5</td>
<td>0.37</td>
<td>1/3</td>
<td>8-PSK</td>
<td>1</td>
<td>448</td>
<td>B</td>
<td></td>
<td></td>
<td></td>
<td>22.4</td>
</tr>
<tr>
<td>MCS-4</td>
<td>1.0</td>
<td>0.53</td>
<td>GMSK</td>
<td>1</td>
<td>352</td>
<td>C</td>
<td></td>
<td></td>
<td></td>
<td>17.6</td>
</tr>
<tr>
<td>MCS-3</td>
<td>0.80</td>
<td>0.53</td>
<td>1</td>
<td>296</td>
<td>272+24</td>
<td>A</td>
<td></td>
<td></td>
<td></td>
<td>14.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>13.6</td>
</tr>
<tr>
<td>MCS-2</td>
<td>0.66</td>
<td>0.53</td>
<td></td>
<td>1</td>
<td>224</td>
<td>B</td>
<td></td>
<td></td>
<td></td>
<td>11.2</td>
</tr>
<tr>
<td>MCS-1</td>
<td>0.53</td>
<td>0.53</td>
<td></td>
<td>1</td>
<td>176</td>
<td>C</td>
<td></td>
<td></td>
<td></td>
<td>8.8</td>
</tr>
</tbody>
</table>

**Table 3.1 EGPRS Channel Coding Schemes [extracted from GSM 03.64]**

### 3.3.2 MCS-1

**Block Size:** 209 bits  
**Information Payload:** 176 bits

Both the header and the payload sections of the radio block are encoded using a 1/3 rate convolutional mother code which has a constraint length of 7. The polynomials used to generate the codewords are:

\[
G_4 = 1 + D^3 + D^5 + D^6
\]

\[
G_7 = 1 + D + D^3 + D^5 + D^6
\]

\[
G_5 = 1 + D + D^4 + D^5
\]  

- (3.2)

Both the header and payload sections are punctured so as to produce a ½ rate convolutional code. An eight-bit parity check is added to header section to provide for error detection, while a twelve-bit check is added to the data payload.
3.3.3 MCS-5

Block Size: 478 bits

Information Payload: 448 bits (56 octets)

An eight-bit cyclic redundancy check sequence is added to the 25-bit header for error detection. The resulting block is then encoded using a 1/3 rate convolutional code defined by the polynomials:

\[
G_4 = 1 + D^2 + D^3 + D^5 + D^6 \\
G_7 = 1 + D + D^2 + D^3 + D^6 \\
G_5 = 1 + D + D^3 + D^6
\] - (3.3)

This results in a 99-bit block to which one spare bit is added. The three-bit USF is mapped to a 36-bit sequence. The burst mapping and interleaving mechanisms ensure this 36-bit precoded USF is spread evenly over the four bursts containing the MCS-5 radio block. A 12-bit CRC and six tailing bits (all zero) are added to the data payload as in MCS-1. Two puncturing schemes are used to produce an output code rate for the user data of 0.37.
3.3.4 Propagation Model

The channel model used in the simulator follows the description of the GSM mobile radio multipath propagation model described in GSM 05.05 [3GPP-05.05]. In this model, it is assumed that the mobile radio environment is dispersive, with several reflectors and scatterers and different distances from the line-of-sight path between the mobile terminal and the base station. For this reason, the transmitted signal may reach the receiver via a number of distinct paths, each having different delays and amplitudes. This phenomenon is best described as a power-delay profile of the propagation environment. This is essentially a number individual taps representing a single beam, and has a gain which varies with time according to fast-fading characteristics. GSM 05.05 defines three such multipath propagation models for use in simulation models, namely the typical cases for rural area (RA), hilly terrain (HT) and urban area (TU). All of these three models are implemented in the simulator, and the delay-spread characteristics for the urban area model are shown below:

<table>
<thead>
<tr>
<th>Tap Number</th>
<th>Relative Time (µs) (1)</th>
<th>Relative Time (µs) (2)</th>
<th>Average Relative Power (dB) (1)</th>
<th>Average Relative Power (dB) (2)</th>
<th>Doppler Spectrum</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.0</td>
<td>0.0</td>
<td>-3.0</td>
<td>-3.0</td>
<td>Classical</td>
</tr>
<tr>
<td>2</td>
<td>0.2</td>
<td>0.2</td>
<td>-0.0</td>
<td>-0.0</td>
<td>Classical</td>
</tr>
<tr>
<td>3</td>
<td>0.5</td>
<td>0.6</td>
<td>-2.0</td>
<td>-2.0</td>
<td>Classical</td>
</tr>
<tr>
<td>4</td>
<td>1.6</td>
<td>1.6</td>
<td>-6.0</td>
<td>-6.0</td>
<td>Classical</td>
</tr>
<tr>
<td>5</td>
<td>2.3</td>
<td>2.4</td>
<td>-8.0</td>
<td>-8.0</td>
<td>Classical</td>
</tr>
<tr>
<td>6</td>
<td>5.0</td>
<td>5.0</td>
<td>-10.0</td>
<td>-10.0</td>
<td>Classical</td>
</tr>
</tbody>
</table>

Table 3.2 Typical Case for Urban Area (Tux) - from [3GPP-05.05]
Within each path followed by the transmitted waves, there occurs narrowband fading. The first-order statistics of such fading are described by the Rayleigh distribution, whereas the second-order statistics are characterised by the classical Doppler spectrum. [RAPP-96] This describes the spread of frequencies that occurs when there is a relative difference in velocity between the mobile terminal and the base station and is a function of speed of the mobile terminal and of the carrier wavelength. In the simulator model, two main propagation models are used, one for ideal frequency hopping and one for no frequency hopping. When no frequency hopping is used, it is assumed that the radio bursts are transmitted on a single carrier with continuous second-order fast fading characteristics. This means, that if a burst is currently experiencing a fade in received power, then it is likely that subsequent bursts may experience similar fading. However, when slow frequency hopping is used, the mobile station will transmit or receive on a fixed frequency for one timeslot (~577µs) and then must hop before the timeslot in the next TDMA frame. This provides interference diversity at the receiver as all bursts experience fading characteristics with greatly reduced correlation with the characteristics of any previous bursts. The simulator model implements ideal frequency hopping by using four uncorrelated Rayleigh processes which are applied alternately to each successive burst [KELL-97]. As the interleaving depth of the PDTCH is only four bursts, then the fading characteristics of all bursts within a single radio block are uncorrelated thereby emulating ideal frequency hopping.

In addition to fading and multipath characteristics, the received signal is corrupted by co-channel interference and noise at the receiver. The interference is simulated by a single co-channel carrier which is added to the transmitted carrier. Eb/No characteristics are represented by an additive white Gaussian noise source added at the receiver.

3.3.5 Simulator Parameters

Table 3.3 is a list of all the parameters that are user-definable, either by modifying the parameters of COSSAP hierarchical models, by changing the building blocks which constitute the model or by using different schematics.
<table>
<thead>
<tr>
<th>Parameter</th>
<th>Settings</th>
</tr>
</thead>
</table>
| Channel Coding Scheme Supported   | GPRS PDTCH CS-1, CS-2, CS-3, CS-4  
HSCSD TCH/F9.6, TCH/F14.4, EGPRS  
PDTCH MCS-1, MCS-2, MCS-3, MCS-4,  
MCS-5, MCS-6, MCS-7                                                               |
| Interleaving                      | Block Rectangular over 4 frames for GPRS  
As specified in GSM 05.01 for EGPRS                                                   |
| Training Sequence Codes           | 8 codes available                                                                                                                          |
| Modulation                        | GMSK, 8-PSK                                                                                                                               |
| Interference Characteristics      | User definable static C/I ratio for single co-channel interferer. May also be disabled. No frequency offset                                    |
| Fading Characteristics            | Rayleigh fading for each path (Rice for one component of RA). Fading varies during one burst.                                                |
| Multipath Characteristics         | TU, RA, HT propagation environments supported as in GSM 05.05                                                                         |
| Transmission Capabilities         | User definable. Can simulate no frequency hopping and ideal frequency hopping (no correlation between successive bursts)                   |
| Mobile terminal velocity          | User definable. Static -> 250 km/hr (for 900MHz)                                                                                           |
| Carrier Frequency                 | User definable to 900MHz or 1800MHz                                                                                                        |
| Antenna Characteristics           | 0dB gain for both transmitter and receiver. No antenna diversity.                                                                          |
| Signal-to-Noise Characteristics   | AWGN source at receiver. User definable $E_b/N_0$ ratio                                                                                 |
| Burst Recovery                    | Synchronisation based on the cross-correlation properties of the training sequence                                                        |
| Equaliser                         | 16-state soft output MLSE Equaliser for GMSK  
16-state Decision-feedback MLSE equaliser for 8-PSK                                                                                  |
| Channel Decoding                  | Soft-decision Viterbi convolutional decoder. Fire correction and detection for CS-1 and CRC detection for CS-2, CS-3, CS-4, and  
MCS-1 to MCS-9                                                                       |
| Performance Measures              | Bit Error Patterns and Block Error Patterns                                                                                               |
| Simulation Length                 | User definable. Most experiments run for 15000 blocks/timeslot                                                                        |

Table 3.3 E/GPRS Simulator Parameters
3.4 Simulator Validation and Results

3.4.1 Context – GPRS Model

Reference performance figures for the GSM physical channel are given in GSM 05.05 [3GPP-05.05]. These allow for the setting of reference transmitter and receiver performance figures for nominal error rates, sensitivity levels, interference levels and erroneous frame indication performance levels. Techniques for specifying reference interference performance of GPRS are tackled in Annex L of GSM 05.50 [GSM-05.50] where methods on how to report GPRS performance in GSM 05.05 are described. As GPRS is designed to provide error-free transmission in bursty error environments, the most appropriate metric for evaluating channel error performance of the different channel coding schemes is the block loss error ratio (BLER). This gives the rate at which uncorrectable errors are detected in a decoded RLC/MAC radio block, which essentially translates into the rate of discarded blocks. In [GSM-05.50], two methods for reporting GPRS performance are proposed. One is to evaluate the BLER and C/I ratios for all coding schemes corresponding to the ranges of highest throughput. The second is to evaluate the C/I ratios for a fixed reference BLER value of 10% for coding schemes CS-1 through CS-4 so as to try to maximise the throughput performance. For all tests a 2dB implementation margin is included when quoting C/I values. As the physical link layer simulator model described in this Chapter does not implement any RLC/MAC functionality, the latter approach is selected for validating the GPRS simulator performance. This is the method adopted in GSM 05.05.

<table>
<thead>
<tr>
<th>Coding Scheme</th>
<th>C/I at BLER=10%</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>GSM 900 TU50 - ideal FH</td>
</tr>
<tr>
<td>CS-1</td>
<td>9 dB</td>
</tr>
<tr>
<td>CS-2</td>
<td>13.8 dB</td>
</tr>
<tr>
<td>CS-3</td>
<td>16 dB</td>
</tr>
<tr>
<td>CS-4</td>
<td>23 dB</td>
</tr>
</tbody>
</table>

Table 3.4 Reference Interference Performance Values - source: [3GPP-05.50]
Implementation margin of 2dB included

3.4.2 GPRS Model Validation at TU50 IFH 900MHz

As mentioned above, Annex L of [GSM-05.50] presents reference interference performance figures for GPRS systems operating at 900 MHz using the TU50 multipath fading channel model. Table 3.5 shows a comparison of the results obtained using the simulation model designed for these experiments, which will be referred to as the CCSR model, and those quoted in [GSM-05.05].

Chapter 4 Page 36
Design and Development of GPRS and EGPRS Radio Access Simulator

### Table 3.5 Comparison of reference performance at TU50 IFH 900MHz

<table>
<thead>
<tr>
<th>Coding Scheme</th>
<th>C/I at BLER=10%</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>[GSM-05.50]</td>
</tr>
<tr>
<td>CS-1</td>
<td>9 dB</td>
</tr>
<tr>
<td>CS-2</td>
<td>13.8 dB</td>
</tr>
<tr>
<td>CS-3</td>
<td>16 dB</td>
</tr>
<tr>
<td>CS-4</td>
<td>23 dB</td>
</tr>
</tbody>
</table>

Conditions: Ideal Frequency Hopping, Receiver Noise Floor Eb/No=28dB 2dB implementation margin assumed

From the results in Table 3.5 it can be seen that particularly for coding schemes CS-2 and CS-3 the performance of the CCSR model is about 1.5dB better than the reference value suggested in Annex L of GSM 05.50. The difference in performance for schemes CS-1 and CS-4 is less pronounced at around 0.5dB. However, comparison of BLER interference performance traces obtained using the CCSR model with those presented in Annex P of the same document [GSM 05.05] Annex P which give the results obtained by Ericsson, show that the performance of the two simulated systems are virtually identical. This is shown in Figure 3.4.

### 3.4.3 GPRS Model Verification at TU1.5 No FH 1800MHz

[GSM 05.05 -P] also presents proposals for GPRS reference performance results for the TU3 multipath model at 1800 MHz with no frequency hopping implemented. The results of the CCSR model at a BLER of 10% under these conditions are shown in Table 3.6.

<table>
<thead>
<tr>
<th>Coding Scheme</th>
<th>C/I at BLER=10%</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>[GSM 05. 50]</td>
</tr>
<tr>
<td>CS-1</td>
<td>13 dB</td>
</tr>
<tr>
<td>CS-2</td>
<td>15 dB</td>
</tr>
<tr>
<td>CS-3</td>
<td>16 dB</td>
</tr>
<tr>
<td>CS-4</td>
<td>19.3 dB</td>
</tr>
</tbody>
</table>

Table 3.6 Comparison of reference performance at TU1.5 No FH 1800MHz

Conditions: No Frequency Hopping, Receiver Noise Floor Eb/No=28dB 2dB implementation margin assumed

The results in Table 3.6 clearly show that under the propagation conditions described in TU1.5, the performance of the CCSR model at a resulting BLER of 10% is very close to the reference interference performance levels specified in [GSM 05.50] Annex L. In fact, the obtained C/I ratio differs by no more than 0.5 dB for CS-1 and CS-3. No suitable results for comparison were obtained from CS-4 as simulations were carried up to C/I=18 dB, at which value the resulting BLER was still in excess of 10%. The performance traces under these conditions can be seen in...
Figure 3.5, where the results obtained with the CCSR model closely match those obtained by Ericsson in [GSM 05.50] Annex P.

### 3.4.4 GPRS Model Verification at TU1.5 Ideal FH 1800MHz

Results obtained using the CCSR Simulator were compared with those specified in GSM 05.05 and are shown below:

<table>
<thead>
<tr>
<th>Coding Scheme</th>
<th>C/I at BLER=10%</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>[3GPP 05. 05]</td>
</tr>
<tr>
<td>CS-1</td>
<td>9 dB</td>
</tr>
<tr>
<td>CS-2</td>
<td>13 dB</td>
</tr>
<tr>
<td>CS-3</td>
<td>15 dB</td>
</tr>
<tr>
<td>CS-4</td>
<td>23 dB</td>
</tr>
</tbody>
</table>

Table 3.7 Comparison of reference performance at TU1.5 IFH 1800MHz

**Conditions:** No Frequency Hopping, Receiver Noise Floor Eb/No=28dB
2dB implementation margin assumed

![GPRS Interference Performance TU50 IFH 900MHz](image)

Figure 3.4 GPRS Interference Performance TU50 IFH 900MHz
Figure 3.5 GPRS Interference Performance TU1.5 no FH 1800MHz

Figure 3.6 GPRS Interference Performance TU1.5 no FH 1800MHz
3.4.5 Comments and Conclusions

The simulations were run for lengths equivalent to $2 \times 10^6$ information bits, which is equivalent to roughly 10800 RLC/MAC blocks for CS-1 and 4600 RLC/MAC blocks for CS-4 coding schemes. This length was chosen for two main reasons. This length of information bits can be used to provide a continuous channel error pattern for a 64 kbit/s video stream for over 30 seconds before looping over to the beginning of the error sequence. As will be demonstrated in Chapter 4, the bursty nature of the GSM fading channel, coupled with the differing visual susceptibility to errors of different parameters that constitute the MPEG-4 bitstream require that sufficiently long error patterns are used in order to obtain meaningful results.

The performance of the CCSR model was validated for a variety of conditions. In particular, the model was seen to give accurate results particularly with respect to variations in interference levels, transmission modes (frequency hopping enabled/disabled) and carrier frequency. Although the results for GPRS were presented as block error values, the output from the simulators characterises the physical layer link in terms of both bit and block error patterns. Although the results obtained with the designed model closely match the quoted reference performance levels, it must be appreciated that much of the system performance relies on implementation-dependent factors. This is particularly true for the receiver’s correlator and equaliser, and to a lesser extent,
Design and Development of GPRS and EGPRS Radio Access Simulator

on the channel decoding mechanisms. These factors lead to variations in GPRS physical layer performance figures released by different manufacturers of up to 2 dB [GSM-05.50]. In addition, it must be noted, that the assumption made in [GSM-05.05] that the TU1.5 1800MHz and TU3 900MHz propagation models are identical is retained in thesis.

3.5 EGPRS Physical Link Layer Simulator

As for the GPRS PDTCHs, the reference performance of EGPRS links is specified in terms of the carrier-to-interference ratio or energy per modulated bit required to obtain a 10% radio block error rate. A block is considered erroneous in the simulation when any of the following occur:

- Uncorrectable bit errors in the data field after decoding (including CRC bits)
- Uncorrectable bit errors in the header field after decoding (including CRC bits)
- Erroneous decoded stealing flag code word

Tdocs [SMG2-1566/99] and [SMG2-561/99] also specify that erroneous modulation detection, which is also referred to as blind detection error should be simulated. This type of error was not included in the simulation models used, as different models were used for the GMSK and 8-PSK modulation schemes, with the receivers knowing a priori the type of modulation to expect. It is not expected that this assumption would cause excessive deviation from the required results.

When examining the results obtained by different manufacturers and specified in [SMG2-401/99] and [SMG2-1565/99], it can be seen that the specified values for the 8-PSK schemes vary widely. For example, in the co-channel interference case at TU1.5 no FH at 1800 MHz given in [SMG2-401/99], there exists a 4.9dB difference between the worst and best quoted value for MCS-5 and of 5.1 for MCS-6. In [SMG2-060/00], fewer results from fewer manufacturers are given, but there still remains a spread of around 3dB in the given values. As a result, average-based reference performance values are given in [SMG2-401/99] and [SMG2-060/00], which are seen to be quite similar to each other. The performance of the CCSR simulation model was therefore compared with these average figures. The variance in values of the GMSK coding schemes (MCS-1 to MCS-4) is shown to be considerably lower than for the 8-PSK schemes. This is probably due to the greater maturity of GMSK-based technology for fading channels as compared to 8-PSK.
### 3.5.1 EGPRS Validation at TU 50 NFH 900 MHz

The EGPRS models were simulated for co-channel interference with a single interferer being used. When examining the results obtained in [SMG2-561/99], it was seen that MCS-8 and MCS-9 could not reach 10% BLER even at C/I values in excess of 30dB. For this reason, reference performance figures at 30% BLER were used in these cases. As will be shown in this thesis, channel error rates which result in BLER values of around 10% are considerably too high to produce acceptable video quality. For this reason MCS-8 and MCS-9 were not used in the tests carried out. The results obtained are shown in Tables 3.8 through 3.11.

<table>
<thead>
<tr>
<th></th>
<th>SMG2-401/99</th>
<th>SMG2-086/00</th>
<th>CCSR</th>
</tr>
</thead>
<tbody>
<tr>
<td>MCS-1</td>
<td>8.5</td>
<td>-</td>
<td>9</td>
</tr>
<tr>
<td>MCS-2</td>
<td>10.5</td>
<td>-</td>
<td>12</td>
</tr>
<tr>
<td>MCS-3</td>
<td>15.0</td>
<td>-</td>
<td>17</td>
</tr>
<tr>
<td>MCS-4</td>
<td>20.0</td>
<td>-</td>
<td>20.5</td>
</tr>
<tr>
<td>MCS-5</td>
<td>15.5</td>
<td>15.5</td>
<td>15.3</td>
</tr>
<tr>
<td>MCS-6</td>
<td>18.0</td>
<td>18.0</td>
<td>18.5</td>
</tr>
<tr>
<td>MCS-7</td>
<td>23.0</td>
<td>25.0</td>
<td>24.0</td>
</tr>
</tbody>
</table>

Table 3.8

The reference figures given in [SMG2-401/99] and [SMG2-086/00], which are both published by the EDGE drafting group are seen to be extremely similar to each other. Under these propagation conditions, the CCSR model is seen to produce results to within 0.5 dB for all coding schemes, with the exception of MCS-2, where the performance of the CCSR model is inferior by 1.5dB, and MCS-3, where a discrepancy of 2dB is noted. At MCS-7, and MCS-5, the CCSR model's performance is superior to the reference figures.

### 3.5.2 EGPRS Validation at TU50 NFH 1800MHz

<table>
<thead>
<tr>
<th></th>
<th>SMG2-060/00</th>
<th>SMG2-564/99</th>
<th>CCSR</th>
</tr>
</thead>
<tbody>
<tr>
<td>MCS-1</td>
<td>-</td>
<td>-</td>
<td>8</td>
</tr>
<tr>
<td>MCS-2</td>
<td>-</td>
<td>-</td>
<td>12</td>
</tr>
<tr>
<td>MCS-3</td>
<td>-</td>
<td>-</td>
<td>17</td>
</tr>
<tr>
<td>MCS-4</td>
<td>-</td>
<td>-</td>
<td>21</td>
</tr>
<tr>
<td>MCS-5</td>
<td>15</td>
<td>13.0</td>
<td>16</td>
</tr>
<tr>
<td>MCS-6</td>
<td>18</td>
<td>15.5</td>
<td>24</td>
</tr>
<tr>
<td>MCS-7</td>
<td>27.5</td>
<td>21.5</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 3.9

Reference performance figures for MCS-1 to MCS-4 for these propagation conditions were not available in the given references. However the performance is fairly similar to the equivalent results obtained at TU50 NFH 900MHz, indicating that the performance is close to the expected values. Indeed, comparing the values for CS-1 to CS-4 for GSM 900MHz and DCS 1800 MHz at TU50 NFH in [GSM-05.05], it can be seen that the reference figures at 10% BLER do not vary by
more than 1dB. There does however exist a considerable discrepancy between the results for 
MCS-5 through MCS-7 as given in [SMG2-564/99] and [SMG2-060/00], where differences of up 
to 6dB can be seen. Although the CCSR model performs well at MCS-5, giving results of within 
1dB of the value quoted in [SMG2-564/99], codes MCS-6 and MCS-7 perform considerably 
worse. In fact at MCS-7 a figure of 10% BLER was not achieved at all.

3.5.3 Validation at TU1.5 NFH 1800MHz

<table>
<thead>
<tr>
<th></th>
<th>SMG2-086/00</th>
<th>SMG2-401/99</th>
<th>CCSR</th>
</tr>
</thead>
<tbody>
<tr>
<td>MCS-1</td>
<td>-</td>
<td>11.0</td>
<td>11.3</td>
</tr>
<tr>
<td>MCS-2</td>
<td>-</td>
<td>13.0</td>
<td>13</td>
</tr>
<tr>
<td>MCS-3</td>
<td>-</td>
<td>14.5</td>
<td>16</td>
</tr>
<tr>
<td>MCS-4</td>
<td>-</td>
<td>17.0</td>
<td>18</td>
</tr>
<tr>
<td>MCS-5</td>
<td>19.5</td>
<td>19.0</td>
<td>15.2</td>
</tr>
<tr>
<td>MCS-6</td>
<td>21.5</td>
<td>21.0</td>
<td>21</td>
</tr>
<tr>
<td>MCS-7</td>
<td>26.5</td>
<td>24.0</td>
<td>24</td>
</tr>
</tbody>
</table>

Table 3.10

The CCSR EGPRS model performs to within 1dB at all C/I ratios as the reference figures except 
for MCS-3 where the discrepancy is of 1.5dB. Indeed at MCS-5, the CCSR model exceeds the 
values given in [SMG2-401/99] and [SMG2-086/00] by about 4dB. It is evident that at these 
conditions, the CCSR model performs very similarly to the reference models.

3.5.4 Validation at TU1.5 IFH 1800MHz

<table>
<thead>
<tr>
<th></th>
<th>SMG2-086/00</th>
<th>SMG2-401/99</th>
<th>CCSR</th>
</tr>
</thead>
<tbody>
<tr>
<td>MCS-1</td>
<td>-</td>
<td>7.5</td>
<td>9.0</td>
</tr>
<tr>
<td>MCS-2</td>
<td>-</td>
<td>10.0</td>
<td>11.5</td>
</tr>
<tr>
<td>MCS-3</td>
<td>-</td>
<td>14.5</td>
<td>16.5</td>
</tr>
<tr>
<td>MCS-4</td>
<td>-</td>
<td>19.5</td>
<td>20.0</td>
</tr>
<tr>
<td>MCS-5</td>
<td>14.5</td>
<td>14.0</td>
<td>15.3</td>
</tr>
<tr>
<td>MCS-6</td>
<td>17.0</td>
<td>17.0</td>
<td>18.5</td>
</tr>
<tr>
<td>MCS-7</td>
<td>23.5</td>
<td>22.5</td>
<td>22.0</td>
</tr>
</tbody>
</table>

Table 3.11

When using ideal frequency hopping, the CCSR model on average performs around 1.5 dB worse 
than the reference figures. One reason for this is that in order to simplify implementation and 
reduce the complexity of the simulation models, although the degree of correlation between 
consecutive bursts is greatly reduced when compared to the non-frequency hopping case, the 
decorrelation is not perfect. This slightly reduces the efficacy of the interleaving mechanisms. 
However the difference between the obtained values and the reference figures are only in excess 
of 1.5dB at MCS-3, and are less than 1dB for MCS-4, MCS-5 and MCS-7. The performance of 
the model under these conditions can consequently be considered as being adequate.
Figure 3.8 GMSK EGPRS Interference Performance TU50 NFH 1800MHz

Figure 3.9 8-PSK EGPRS Interference Performance TU50 NFH 1800MHz
Figure 3.10 GMSK EGPRS Interference Performance TU1.5 IFH 1800MHz

Figure 3.11 8-PSK EGPRS Interference Performance TU1.5 IFH 1800MHz
Figure 3.12 GMSK EGPRS Interference Performance TU1.5 NFH 1800MHz

Figure 3.13 8-PSK EGPRS Interference Performance TU1.5 NFH 1800MHz
3.5.5 Comments and Conclusions

The EGPRS physical link layer model was validated for four propagation conditions, all of which assumed the Typical Urban multipath model. Both the GMSK and 8-PSK modulator/demodulator structures were seen to give the expected results, and the relative differences between the performance of the two receiver structures as compared to the reference values were rather small.

There does however exist a considerable difference between the performance of the designed model and that of the better performing models given in [SMG2-401/99], particularly for the 8-PSK modulation-coding schemes. As already mentioned, there exists a spread of around 4-5dB in the figures given for different receiver implementations. There may be various reasons for this, including possible differences in the way the propagation models are simulated. However, the most likely reason lies within the different implementation strategies of the receivers. There exist several techniques for carrying out channel and noise estimation, and implementing equalisation at the receiver. Some methods can adapt dynamically to differing channel conditions, whereas others perform optimally under certain conditions, and not so well under others. The differences in the 8-PSK values are greater than for GMSK where the technology is fairly stable and consolidated.

As a result, although the CCSR model matches the reference performance figures or comes very close to them under practically all conditions tested, they should be regarded to a certain extent as the worst case performance figures. In fact as already described, the CCSR model used a custom-made EDGE receiver mechanism, which although employs a very effective non-linear direct-feedback equaliser [PROK-95], cannot be considered as being the most optimal solution. In particular, no automatic frequency correction mechanisms were implemented in the receiver.

The comparison with the reference performance figures is however only part of the story. The 10% BLER figure chosen for measuring performance as compared to reference values was selected on the basis of being around the position where optimal throughput is achieved when operating with block retransmissions [GSM-05.50]. However as will be explained in Chapter 5, real-time services require information integrity without use of retransmissions and consequently the error performance requirements are much more stringent. Typically, error rates in the order of $10^{-3}$ to $10^{-4}$ are required for video communications. For this reason, the performance of the various Modulation-Coding schemes at lower BLER values and relatively high C/I values are more critical than they would otherwise be for typical data transfer applications. The equaliser used for
the GMSK modulation schemes is a 16-state Viterbi equaliser. Examination of results above C/I values of around 15-20dB show a considerable deviation from the quoted figures. Figure 3.14 shows the BER sensitivity performance of the GMSK receiver for the TU50 multipath model at 900MHz and 1800MHz. As these figures display raw error rates with no forward error correction, the results are not affected by frequency hopping. For $E_b/N_0$ values below 18dB, the equaliser used in the CCSR model outperforms the reference figures, but then levels off to a higher asymptotic BER value than quoted in [8]. The probable reason for such a deviation is that the equaliser was optimised for operation of the speech channels and low-bitrate data. These typically operate at C/I values below 12dB. Indeed differences in bit-error rates below $10^{-3}$ have a negligible effect on speech quality or data throughput using the TCH/9.6 or TCH/4.8 channels. This difference is consequently more conspicuous when using schemes MCS-3 and MCS-4, which for real-time services, are the coding schemes that would operate at such interference levels when using GMSK. The major deviations are visible for MCS-3, where for example, the BLER values at C/I=20dB for TU50 NFH 1800MHz are 0.05 and 0.015 respectively. A similar trend is visible when examining the sensitivity performance of the 8-PSK receiver in the TU50 multipath propagation conditions. At low Eb/No values, while inferior to the reference performance figures, the performance of the equaliser used in these experiments is quite close to the reference performance figures. However, at low noise values, this discrepancy increases considerably. There is also a large difference between the performance at 900MHz and at 1800MHz, where the difference asymptotic bitrates is nearly in the order of one magnitude. This is caused by the greater susceptibility of 8-PSK to intersymbol interference, and the greater symbol spreading that occurs at 1800MHz, particularly at high mobile speeds.

The residual bit error patterns obtained for both GPRS and EGPRS are nevertheless suitable for use in the multimedia transmission experiments to be described in the following Chapters, as they exhibit relative performance figures between coding schemes which are consistent with the relative power of the schemes and with the results quoted in the references examined in this Chapter. Moreover, the obtained results were shown to display a high degree of correlation with the performance results given by several manufacturers.
3.6 E/GPRS Radio Interface Data Flow Model

The design of the EGPRS physical link layer model was restricted to examining the effects of varying channel conditions upon bits exiting the channel decoders. In order to carry out more extensive and detailed examinations of the effect of channel errors upon end-applications, such as video coding implementations a GPRS data flow simulator was implemented. The model was implemented in Matlab as this language provides a rapid development environment and
Design and Development of GPRS and EGPRS Radio Access Simulator

comprehensive data analysis tools. The layers implemented included an MPEG-4 video codec with rate control functionality, RTP/UDP/IP transport layers and GPRS SNDC, LLC, RLC/MAC layer protocols. This layout is shown in Figure 3.16. It must be emphasised that only the data flow properties of the protocols have been implemented in this model. This means that none of the protocol signalling mechanisms have actually been included in the model, but only the resulting effect on header sizes, packet and stream segmentation procedures and flow control effects. For example, when describing the RTP layer, sequence numbering is not actually implemented, but only its effect on the resulting RTP-PDU header size is modelled in the simulator.

The application layer consists of an traffic source emulating an MPEG-4 video codec which employs error resilience functionality as described in [TALL-98] and rate control mechanisms which place an upper limit on the output throughput from the encoder as calculated on a frame-by-frame basis. The maximum allowable throughput is set according to the resources allocated across the radio interface. The output from the MPEG-4 codec is forwarded to the transport layers in units of discrete video frames, which will be referred to as video packets. These packets are then split up into transport layer PDUs according to the maximum IP packet size defined by the user. Each packet is then encapsulated into an independent RTP/UDP/IP [RFC-1889] packet for forwarding down to the GPRS network. Header compression [BURM-00], [CASN-01] which is a user-definable feature of this simulator model, is implemented at the transport layer, although is not actually used in the experiments described in this report. This means that the compression protocol that is employed in the end terminals must be supported in all the intermediate nodes in the core network, which potentially includes the Internet. This is hardly a realistic scenario, and a more appropriate implementation would be to include such functionality in the SNDC layer [3GPP-04.65]. This would allow for headers to be compressed for transmission across the GPRS radio interface, only to be restored to their initial size at the SGSN for forwarding across the remainder of the network. However, as this model is solely concerned with the performance across the Um interface, the location of the compression algorithm has no effect upon any results obtained.

At the SNDC layer, the 8-bit SNDC header [3GPP-04.65] is added to the transport packet before forwarding on to the LLC layer. Here a 24-bit frame header and 24-bit frame check sequence is added to each LLC-PDU. A check is also carried out to ensure that no LLC frames exceed the maximum size of 1520 octets specified in [3GPP-04.64]. The error control and flow control functions of the LLC layer are not implemented in this model. The LLC frames are then passed on to the RLC/MAC layer, where they are encapsulated into radio blocks according to the forward
error correction scheme selected. In practice the choice of coding scheme depends upon the carrier-to-interference ratio at the terminal and the resulting throughput that can be sustained at that C/I level using the different channel coding schemes. The major side-effect of varying the protection afforded to the user data is the resulting modification in the size of the GPRS radio blocks in terms of the number of information bits per block. The model therefore segments the incoming PDU into data payloads for the output radio blocks according to the selected channel coding scheme and forwards these blocks to a FIFO buffer. Once in the output buffer, the blocks wait for one of the timeslots allocated to their associated terminal to become available and are then transmitted over the given timeslots. The model allows any number of timeslots from 1 to 8 to be allocated to the source terminal.

This layered model design provides for error occurrences at the physical layer to be mapped onto the actual application-layer payload. Each channel error can therefore be mapped onto an individual video information bit or the header or checksum section of any protocol in the GPRS stack.
Design and Development of GPRS and EGPRS Radio Access Simulator

**Application Layer**: MPEG-4 Codec

**Video Packet**

**Transport Layer**: Rate Control: Some Video Packets Discarded

**Segmentation**

**Transport PDU**

**Add Headers**

**RTP-UDP-IP**

**IP PDU**

**Header Compression (optional)**

**IP PDU**

**GPRS SNDC Layer**

**SNDC Payload**

**SNDC Header**

**GPRS LLC Layer**

**LLC Payload**

**LLC Header**

**GPRS RLC/MAC Layer**: Select Channel Coding Scheme

**CS-1**  **CS-2**  **CS-3**  **CS-4**

**RLC/MAC block**  **RLC/MAC block**  **RLC/MAC block**  **RLC/MAC block**

**Physical Link Layer**: Select Number of Timeslots

1  2  ...  8

**Figure 3.16 GPRS Data Flow Model**
3.7 Real-time Emulator

The GPRS and EGPRS physical link layer simulation model is extremely computationally intensive, largely due to the modelling of the multipath propagation model, where a Rayleigh fading filter is used to represent each pathway. For example, a simulation representing data encoded using the CS-1 scheme at TU1.5 IFH 1800MHz at a carrier-to-interference ratio of 15dB run on a 296MHz Ultra SPARC processor, runs at a average rate of 138 information bits per second. Although the exact processing speed depends upon several factors, including the propagation conditions modelled, the C/I ratio present and the modulation-coding scheme used, the obtained rates are way below those necessary to support a real-time simulation environment. Chapter 4 will show that when transmitting video sequences over GPRS, two- or three-slot operation is required.

![Figure 3.17 GPRS Radio Access Emulator Structure](image)

In order to create a real-time testing environment for video communications applications, a real-time emulator was built using Visual C++ for Microsoft Windows. The emulator implemented the data flow model described in Section 3.6 and allowed for up to 8-slot allocation. The emulator program made use of a table look-up method in order to allow for real-time emulation. Data sets of bit-error patterns at the Physical Link Layer were created with the E/GPRS simulator for a wide range of interference and propagation conditions for each coding scheme. These were then used by the real-time emulator and fed into the GPRS Radio Interface data flow model described in Figure 3.16. Multi-threaded programming techniques were used to build the model, and a Graphical User Interface was designed to allow for interactive manipulation of the Coding Scheme, interference level, carrier frequency, timeslot allocation and frequency hopping capability. The emulator was used in conjunction with a real-time MPEG-4 video encoder/decoder application using RTP and is shown in Figure 3.18.
3.8 Conclusion

This Chapter has described the design and validation of EGPRS and GPRS physical link layer simulation model. It was seen that the GPRS models were seen to give performance which closely matched the GSM reference performance figures. Although the performance of the EGPRS model was seen to match the figures given by different terminal manufacturers when operating at low terminal velocities, a significant divergence from the reference figures was obtained using the TU50 propagation models at 1800MHz. This may be attributable to the increased Doppler spread at high terminal velocities and carrier frequencies, and corresponding limitations in the equaliser and receiver architecture in dealing with the resulting intersymbol interference. However at lower terminal velocity figures and carrier frequencies, the simulator model closely matched the reference performance figures, and may therefore be considered as being suitable for use in the media transmission experiments to be described in the next Chapters.
Chapter 4

4 Real-time Video Communications over GPRS and EGPRS

4.1 Introduction

The most demanding form of communications class in terms of the demands and requirements that must be satisfied by the radio interface of a mobile communications system is the real-time, or conversational service class. ITU-T recommendation G.114 [ITU-T-G.114] lays down maximum one-way delays of 150ms for telephony services. As video communications will have to be aligned with a corresponding audio channel for carrying the speech information, the delay requirements for the video information must satisfy these same requirements [3GPP-23.107]. Similar QoS requirements are specified in the UMTS Service Description documents. It has been shown in [PARA-99, FABR-99] that when transmitting speech over GPRS, such stringent delay requirements may be met as long as the same Temporary Block Flow (TBF) is maintained, and that no retransmissions are employed. The GPRS uplink MAC protocol requires that channel contention and allocation be carried out at the beginning of each new group of Layer 3 packets, referred to as a TBF. As opposed to speech, the video traffic characteristics cannot be modelled by a two state ON-OFF Markov model. Instead, the packet generation rate is typically constant, determined by the video frame rate, with the main variation being restricted to the length of the individual packets. This allows for transmission over the E/GPPRS radio interfaces to require limited TBF initialisations, thereby facilitating real-time operation. The limitation on the use of backward error correction does however have significant implications on the quality of video that can be received in a mobile environment. This Chapter will examine the main issues concerned with real-time video over GPRS and EGPRS, with particular emphasis being placed on the resource allocation requirements and on the effect of co-channel interference upon the received video quality. The effect of the characteristics of the mobile radio channel upon the transmitted video sequences will be assessed experimentally using the Physical Link Layer simulator described in Chapter 3.
4.2 Source Sequences

Three different video sequences were used for determining the capacity requirements and error performance of MPEG-4 video over the GPRS physical link layer. Two of these, Foreman and Carphone are standard ITU-T sequences used for the evaluation of video coding compression and error-resilience performance. Although essentially video-conferencing type sequences, they place rather severe requirements upon the compression algorithms, as they describe high-activity scenes. The Foreman sequence, in particular is very demanding, as the camera is non-stationary and pans around the scene. In addition to these two standard sequences, a third, Akiyo with Background, or just Akiyo for short, is used. This is a modification of yet another standard ITU-T sequence, in which a non-moving background is replaced by a scene of a moving crowd. This is an extremely high-activity scene, which as will be seen, is very difficult to compress without losing much spatial or temporal detail. Complex, high-detail sequences were used rather than simpler, low-activity head and shoulder sequences, so as to allow for non-rate adaptive systems to be engineered around the expected worst-case conditions.

4.2.1 Source Traffic Characteristics

As described in Chapter 2 one of the most effective ways of preventing the propagation of errors in encoded video sequences is the regular insertion of INTRA-coded frames which do not make use of any information from previously transmitted frames. Unfortunately, this has the disadvantage of making the traffic characteristics of a video sequence extremely bursty, as a much larger number of bits are required to obtain the same quality levels as INTER-coded (predictively-coded) frames. Figure 4.1 shows the traffic characteristics of the Akiyo sequence when encoded using a fixed quantiser of value 20. It was observed that the average number of bits required to represent an INTRA-coded frame is roughly three times that required for a predictively coded frame. These characteristics mean that on average, the subsequent two video frames generated by the encoder will have to be discarded as they would have been made obsolete by newer frames by the time new channel capacity is available. This sequence was coded at 10 frames/second, thereby giving an extra delay of approximately 200ms every INTRA-frame. At an INTRA-spacing of 10 frames, this represents a loss of 20% of encoded frames, which is clearly unacceptable. When encoding using a fixed quantiser, it is also noticed that the spatial quality of the frames in terms of the Peak Signal-to-Noise Ratio (PSNR) throughout the sequence remains relatively constant, with slight peaks at each INTRA-coded frame.
These traffic characteristics clearly demonstrate the need for rate control mechanisms. One method, adopted by the encoder implementation used in these experiments is to vary the quantiser value used to truncate the higher-frequency DCT coefficients according to the required target bitrate and the number of bits available to encode a particular frame. When the rate control
algorithm is implemented, it is seen that the number of bits necessary to represent INTRA-frames is greatly reduced, and the variance of the video frame size is greatly decreased, an improvement which consolidates with time. However this improvement is obtained at the expense of received quality. As can be seen in Figure 4.2, the spatial quality of the encoded sequence now varies much more than for the case with a fixed quantiser, and the quality level actually takes a dip at INTRA-coded frames. This is because a very coarse quantiser must be adopted to ensure that the bit allocation is not exceeded when no prediction information is employed. This variability is however preferable to the jitter caused by the disparity in size between Intra- and Inter-coded frames when no rate control is used. For this reason, all the experiments described in this report will use video sequences using the rate control mechanisms enabled.

4.3 Video Communications over GPRS

The MPEG-4 video encoder was operated with the rate control algorithm enabled and set to the desired level. Each video frame was encapsulated into a separate RTP packet, unless a maximum RTP-PDU size of 512 octets (4096 bits) was reached, in which case the video frame segmentation is carried out. This value was chosen because although an increase in average IP packet size results in a more efficient usage of the available throughput by reducing the protocol overheads, excessively large packets will be more vulnerable to information loss due to header corruption. It is also assumed that 96 bits are required for the RTP header, 32 bits for the UDP header, 192 bits for the IP header, 8 bits for the SNDC-PDU header, 24 bits for the LLC header and a further 24 for the LLC Frame Check Sequence.

4.3.1 GPRS Traffic Capacity

The payload available in a GPRS RLC/MAC block depends upon the channel coding scheme used. As can be seen in Table 4.1, the data rate of the RLC/MAC data payload, i.e. the rate presented to the LLC layer, varies from 8 kbit/s for CS-1 to 20.35 kbit/s for CS-4. The available throughput to a single terminal will be multiples of these rates, depending upon the multislotting capabilities of the terminal. As it is envisaged that the CS-1 and CS-2 schemes will be used for video applications, so emphasis will be made on these two schemes.

<table>
<thead>
<tr>
<th>Scheme</th>
<th>Code Rate</th>
<th>Pre-coded USF</th>
<th>Data Payload</th>
<th>BCS</th>
<th>Tail</th>
<th>Radio Block Size (headers+data)</th>
<th>Data Rate kb/s</th>
</tr>
</thead>
<tbody>
<tr>
<td>CS-1</td>
<td>1/2</td>
<td>3</td>
<td>160</td>
<td>40</td>
<td>4</td>
<td>181</td>
<td>8.0</td>
</tr>
<tr>
<td>CS-2</td>
<td>2/3</td>
<td>6</td>
<td>247</td>
<td>16</td>
<td>4</td>
<td>268</td>
<td>12.35</td>
</tr>
<tr>
<td>CS-3</td>
<td>3/4</td>
<td>6</td>
<td>291</td>
<td>16</td>
<td>4</td>
<td>312</td>
<td>14.55</td>
</tr>
<tr>
<td>CS-4</td>
<td>1</td>
<td>12</td>
<td>407</td>
<td>16</td>
<td>-</td>
<td>428</td>
<td>20.35</td>
</tr>
</tbody>
</table>

Table 4.1 GPRS Channel Coding Schemes
As described, the values of the data rates given in Table 4.1 represent the throughput at which LLC-PDUs are transmitted across the U_m and G_b interfaces to the Serving GPRS Support Node (SGSN) over the RLC/MAC and BSSGP interfaces. In the work described in this Chapter, it is assumed that traffic bottlenecks are caused solely by restrictions in timeslot allocation across the U_m interface, and never due to restrictions placed by the BSSGP layer. When considering the protocol stack across the U_m interface as described in Chapter 3, it can be seen that in addition to the video information, the RLC/MAC data payload contains header and other related signalling overheads from the LLC, SNDC, IP, UDP and RTP layers. The presence of these overheads will reduce the true throughput available at the application layer, which in the case being studied, is the MPEG-4 encoder.

In order to determine the throughput per timeslot available at coding schemes CS-1, CS-2, CS-3, the Akiyo and Foreman video sequences were encoded at a number of different source rates and passed through the GPRS Data Flow Simulator so as to determine the average number of source bits transmitted in each RLC/MAC block. This was done so as to be able to determine the output rate that must be set at the source video decoder for the combination of timeslot allocation and channel coding used.

![Protocol Efficiency of GPRS Stack](image)

Figure 4.3 GPRS Protocol Efficiency
Analysis of the protocol efficiency in the payload of the RLC/MAC blocks for the operating conditions described above shows that when encoding at both 5 frames/s as well as at 10 frames/s the efficiency is in excess of 88%. This means that fewer than 15% of the bits in the payload of a radio block are used up by header information belonging to the overlying protocols. The efficiency is generally seen to increase together with source rate, an indication of the increasing size of LLC PDUs. Similar experiments carried out for sequences encoded at 10 frames/s show a variation in efficiency from around 89% to 90%. This, together with the significant differences between the Akiyo and Foreman sequences, clearly indicates that a reduction in the data rate per timeslot as seen by the video encoder of 15% is enough to compensate for all the protocol overheads. The newly-computed radio block data rates are given in Table 4.2.

<table>
<thead>
<tr>
<th>Timeslots</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>CS-1</td>
<td>6800</td>
<td>13600</td>
<td>20400</td>
<td>27200</td>
<td>34000</td>
<td>40800</td>
<td>47600</td>
<td>54400</td>
</tr>
<tr>
<td>CS-2</td>
<td>10500</td>
<td>21000</td>
<td>31500</td>
<td>42000</td>
<td>52500</td>
<td>63000</td>
<td>73500</td>
<td>84000</td>
</tr>
<tr>
<td>CS-3</td>
<td>12200</td>
<td>24400</td>
<td>36600</td>
<td>48800</td>
<td>61000</td>
<td>73200</td>
<td>85400</td>
<td>97600</td>
</tr>
<tr>
<td>CS-4</td>
<td>17200</td>
<td>34400</td>
<td>51600</td>
<td>68800</td>
<td>86000</td>
<td>103200</td>
<td>120400</td>
<td>137600</td>
</tr>
</tbody>
</table>

Table 4.2 Source Throughput for different Allocation Schemes (units in bit/s)

In order to assess the image quality that can be achieved at different rates, as obtained by the different timeslot/coding-scheme combinations shown in Table 4.2, the Akiyo and Foreman sequences were encoded using full-error tools enabled at a range of rates from 20kbit/s to 84kbit/s. The frame rates employed were 5 and 10 frames/second each. These rates were chosen because experiments show that higher rates cannot be supported at the restricted throughput available over the GPRS physical channel. These rates are sufficient to provide acceptable quality video communications. Figures 4.4 and 4.5 show the spatial quality of these two sequences encoded at 5 and 10 kbit/s. At both rates, it is seen that quality of the Foreman sequence is constantly between 3 and 4dB better than that of the Akiyo sequence. This is due to the higher levels of activity in the latter sequence.

These results can be used to determine the spatial quality of received video streams transmitted over GPRS PDTCHs using different levels of multislotting. The Akiyo and Foreman sequences were encoded for all the output rates shown in Table 4.2 at both 5 and 10 frames/sec. In Figure 4.6, it can be seen that at 5 fps, under conditions of no channel errors, at least two slots are required to encode video sequences, and this is only possible using codes CS-2 and CS-3. If channel conditions dictate that CS-1 must be used, then 3-slot operation is necessary. It can also be seen that the differences in radio block payload sizes results in a significant quality deficit.
Design and Development of GPRS and EGPRS Radio Access Simulator

when operating at CS-1 as compared to CS-2 or CS-3 whereas there is little visible difference between the quality achieved with CS-2 and CS-3. The results obtained when transmitting 10 frames/sec are similar to those observed at 5 frames/sec. However, once again, it can be seen that the throughput advantage of codes CS-2 and CS-3 allow them to support this frame rate using only 3 timeslots, whereas 5 timeslots are required using CS-1. It can also be seen that there is little difference between codes CS-2 and CS-3, particularly at lower levels of multislotting.

![Quality performance of MPEG-4 at 10 frames/s](image1)

**Figure 4.4 Quality performance of MPEG-4 at 10 frames/s**

![Quality performance of MPEG-4 at 5 frames/s](image2)

**Figure 4.5 Quality performance of MPEG-4 at 5 frames/s**
The spatial quality of a video sequence is not the only factor that must be taken into consideration when assessing the performance of video coding mechanisms. Video sequences are three-dimensional signals, in which the temporal resolution and quality play an important part in determining whether a received signal is acceptable or not. Under many circumstances, particularly at lower rates, the MPEG-4 rate control mechanisms is forced to discard some frames and not encode them so as not to exceed the stipulated output rate. This causes temporary
‘freezing’ of the sequence at the receiver as an expected frame is not received, and the receiver has to wait for the following video frame. This jitter, or variation in the rate of display at the receiver terminal is subjectively very annoying and therefore efforts must be made to reduce its occurrence as much as possible. In Figure 4.8 it can be observed that an increase in available throughput not only increases the spatial quality of the sequences, but also brings about an improvement in the temporal quality of the same received sequences.

The converse is also true, and is possibly more relevant to the relatively low bitrates made available by GPRS. We can see that although decent PSNR values for sequences at 5fps are achieved using 2 slots at CS-2 and CS-3, these levels can only be achieved if 35% and 20% of source frames are not transmitted. This results in a considerable degradation in sequence quality. Use of 3-slot operation reduces these values to between 5% and 1% which can be considered as being acceptable. A similar situation can be seen to occur when transmitting at 10fps. Although it is possible to support this rate using 3 timeslots at CS-2 and CS-3, these rates are only sustainable by discarding between 30 and 35% of the source frames. In fact, loss rates inferior to 5% are only possible at throughputs in excess of 50 kbit/s, representing at least 5 slots using CS-2 and CS-3 and is unobtainable using CS-1.

![Figure 4.8 Frames dropped by MPEG-4 rate control](image)
4.3.2 Error Performance

One of the most critical factors in providing video communications over mobile channels such as the GPRS PDTCH is the high levels of errors that occur on the link. Traditionally, compressed video has been extremely susceptible to bit and packet errors as any loss or corruption of data could easily lead to the loss of synchronisation of the variable-rate packets. In order to counter this problem, several error resilience techniques have been proposed and implemented for use in different coding schemes as described in Chapter 3. As a result of these techniques, superior image quality may be obtained by using corrupted video information, rather than discarding all the bits in such blocks. This is because although it may be ascertained that total data integrity has been compromised, in general the information available in corrupted data blocks allows for a greater improvement in received quality than the potential distortion caused by those bits in error.

For this reason, in the simulations described in this section, it is assumed that GPRS Reliability Class 5 (see Table 2.5) is used. This specifies use of unacknowledged operation in GTP with unacknowledged operation in the LLC layer with no data protection enabled. The RLC Block Operation is also set to unacknowledged mode. This means that corrupted RLC blocks are forwarded to the LLC layer, which then also forwards corrupted LLC frames, and frames in which data has been lost.

4.3.2.1 Simulation Conditions

The following tests were carried out using the bit error patterns generated using the CCSR GPRS Physical Link Layer Simulator described in Chapter 3 for the following conditions. No implementation loss was included.

- TU 50 Ideal Frequency Hopping 1800 MHz
- TU1.5 Ideal Frequency Hopping 1800 MHz
- TU1.5 No Frequency Hopping 1800 MHz

Each point in the experiments, corresponding to a particular channel condition was determined by carrying out 10 runs each of the Akiyo, Carphone and Foreman sequences using randomly selected starting positions for the bit error files. This is equivalent to averaging out the measured PSNR values of 1790 frames, and was found to be sufficient to provide reliable results.
4.3.2.2 Effect of bit errors

In Figure 4.9 the average PSNR values for sequences protected using schemes CS-1, CS-2 and CS-3 when corrupted by bit errors characteristic of the TU50 IFH propagation conditions are shown. It can be seen that for the CS-1 scheme the maximum PSNR achievable is approached at a C/I value of around 14dB. For the CS-2 and CS-3, such levels are only reached at C/I values in excess of 19dB, although for most operating conditions below this point, the CS-2 scheme gives results superior to those obtained with CS-3 by about 2.5dB.

A similar situation can be seen when examining the results obtained using the TU1.5 IFH propagation model. Although once again the CS-1 code gives maximal results at around 14dB, there is now a more noticeable difference between the CS-2 and CS-3 schemes. In fact, the difference between these two schemes varies between 3.5 and 4dB, while the asymptotic limits are seen to be reached at around 18.5dB and above 21dB respectively. However, a completely different situation can be seen when operating in the TU1.5 NFH propagation model. Under these circumstances, there is very little difference between the different channel coding schemes. In fact, for the same given throughput, the CS-1 code gives only between ¾dB and 1½dB improvement over CS-2, while there is virtually no difference between the CS-2 and CS-3 schemes.
Design and Development of GPRS and EGPRS Radio Access Simulator

Error Performance at GPRS PDTCH TU50 IFH 1800MHz

Figure 4.9

Error Performance at GPRS PDTCH TU1.5 IFH 1800MHz

Figure 4.10
4.3.2.3 Channel Coding Scheme Selection

The experiments described thus far clearly demonstrate the compromise that must be made between payload capacity and error-correcting capability when selecting the appropriate channel coding scheme for video transmission. It is necessary to assess the variation in received quality for a given number of timeslots under different C/I ratios so as to allow for a determination of the operation range of each coding scheme in terms of received C/I. It must be remembered that the criteria for the selection of coding schemes for real-time applications are different than for data transfer applications where data integrity must be maintained and consequently the delay limits allow for the use of backward error correction mechanisms. The optimum choice of coding scheme is that which provides the highest throughput of error-free data (after error detection and retransmission) at the LLC layer at a given C/I ratio. On the other hand, in real-time multimedia communications the optimum coding scheme is that which provides the best subjective quality results at the particular C/I conditions being tested.

Figures 4.12-14 show the averaged PSNR results for the Akiyo and Foreman sequences using 3-slots. It must be noted that the PSNR is calculated between the received video frame and the equivalent original frame, with no concessions made for frames being discarded by the source rate
control mechanism. This means that when frames are discarded, a misalignment occurs between the source and receiver, thereby affecting the computed PSNR values. It is felt that for the length of sequences used in these tests, the reduction in PSNR caused by such misalignment gives a useful indication of the reduction in quality due to the effect of lost frames. While it is most definitely not the optimum way of providing an objective metric of spatial and temporal quality, it can nonetheless serve to give a useful indication of the received video quality.

As expected, when using the TU1.5 IFH and TU50 IFH models, the CS-1 code gave optimal performance under high interference levels, although when the channel conditions allowed for use of the CS-2 code, a very noticeable improvement in quality over that obtainable using CS-1 was observed. This difference is partly attributable to the severe throughput limitations of three-slot operation. In fact, using 3 slots at CS-1 only allows for a source throughput of 20 kbit/s. Allocating four slots to the user will somewhat reduce this difference. The video quality that can be obtained under the TU1.5 propagation conditions with no frequency hopping is much worse than the two models just considered. In fact, enabling frequency hopping results in an improvement in PSNR of between 4 and 5 dB at 1.5 km/hr. If no frequency hopping is implemented, and it is seen that acceptable quality is only achieved at extremely high C/I values. In fact, whereas a PSNR value of 26dB is achieved at around 16dB for TU1.5 IFH and TU50 IFH, a C/I ratio in excess of 21dB is necessary to achieve a similar performance at TU 1.5NFH.

![Figure 4.12 Rate-Adaptive Performance with 3 slots TU50 IFH 1800MHz](image-url)
Figure 4.13 Rate-Adaptive Performance with 3 slots TU1.5 NFH 1800MHz

Figure 4.14 Rate-Adaptive Performance with 3 slots TU1.5 IFH 1800MHz
Table 4.4 Operational Scenarios for Video over GPRS

<table>
<thead>
<tr>
<th>Channel Model</th>
<th>CS-1</th>
<th>CS-2</th>
<th>CS-3</th>
</tr>
</thead>
<tbody>
<tr>
<td>TU50 IFH 1800MHz</td>
<td>&lt;15dB</td>
<td>15 - 21 dB</td>
<td>&gt;21 dB</td>
</tr>
<tr>
<td>TU1.5 IFH 1800MHz</td>
<td>&lt;15dB</td>
<td>15 - 22 dB</td>
<td>&gt;22 dB</td>
</tr>
<tr>
<td>TU1.5 NFH 1800MHz</td>
<td>&lt;14dB</td>
<td>14 - 23 dB</td>
<td>&gt;23 dB</td>
</tr>
</tbody>
</table>

It is immediately obvious from the above table that transmitting video in real-time (i.e. without retransmissions) requires low interference levels, which are considerably more demanding than for ordinary data transfer applications. Typically a C/I ratio of at least 14dB (with no implementation or antenna loss included) is required, while at low mobile terminal speeds, frequency hopping greatly enhances system performance.

4.4 Video Communications over EGPRS

As outlined in Chapter 3, EGPRS allows for a considerable increase in throughput availability to a single user, given enough traffic availability and benign interference conditions. This means that it is possible to provide video services with higher data rates than is possible with GPRS. In Section 4.3.1, it was shown that on average, assuming no header compression, a protocol efficiency of between 88% and 90% is achieved. In order to allow for a fair comparison with the results obtained for GPRS, an 85% efficiency level is assumed. The resulting throughput allocation per timeslot at the application level, or as seen by the video codec, is shown in Table 4.5.

<table>
<thead>
<tr>
<th>Scheme</th>
<th>1 TS</th>
<th>2 TS</th>
<th>3 TS</th>
<th>4 TS</th>
<th>5 TS</th>
<th>6 TS</th>
<th>7 TS</th>
<th>8 TS</th>
</tr>
</thead>
<tbody>
<tr>
<td>MCS-1</td>
<td>7.5</td>
<td>15</td>
<td>22.5</td>
<td>30</td>
<td>37.5</td>
<td>45</td>
<td>52.5</td>
<td>60</td>
</tr>
<tr>
<td>MCS-2</td>
<td>9.6</td>
<td>19.2</td>
<td>28.8</td>
<td>38.4</td>
<td>48</td>
<td>57.6</td>
<td>67.2</td>
<td>76.8</td>
</tr>
<tr>
<td>MCS-3</td>
<td>12.6</td>
<td>25.2</td>
<td>37.8</td>
<td>50.4</td>
<td>63</td>
<td>75.6</td>
<td>88.2</td>
<td>100.8</td>
</tr>
<tr>
<td>MCS-4</td>
<td>15</td>
<td>30</td>
<td>45</td>
<td>60</td>
<td>75</td>
<td>90</td>
<td>105</td>
<td>120</td>
</tr>
<tr>
<td>MCS-5</td>
<td>19</td>
<td>38</td>
<td>57</td>
<td>76</td>
<td>95</td>
<td>114</td>
<td>133</td>
<td>152</td>
</tr>
<tr>
<td>MCS-6</td>
<td>25.2</td>
<td>50.4</td>
<td>75.6</td>
<td>100.8</td>
<td>126</td>
<td>151.2</td>
<td>176.4</td>
<td>201.6</td>
</tr>
<tr>
<td>MCS-7</td>
<td>38</td>
<td>76</td>
<td>114</td>
<td>152</td>
<td>190</td>
<td>228</td>
<td>266</td>
<td>304</td>
</tr>
<tr>
<td>MCS-8</td>
<td>46.2</td>
<td>92.4</td>
<td>138.6</td>
<td>184.8</td>
<td>231</td>
<td>277.2</td>
<td>323.4</td>
<td>369.6</td>
</tr>
<tr>
<td>MCS-9</td>
<td>50.3</td>
<td>100.6</td>
<td>150.9</td>
<td>201.2</td>
<td>251.5</td>
<td>301.8</td>
<td>352.1</td>
<td>402.4</td>
</tr>
</tbody>
</table>

Table 4.5 EGPRS Multislotting capacity for video (kbit/s)

The source throughput capacity for a single channel are seen to vary from 7.5 kbit/s for MCS-1 to 50 kbit/s for MCS-9. This means that there is a much greater spread in available throughput values for video services over EGPRS, assuming that the bit-error conditions are met.

4.4.1 Traffic Characteristics

Experiments were carried out to determine the range of quality levels that can be supported by EGPRS. The sequences used were those introduced in Section 4.2.
The received video quality for the error-free case of the three sequences are shown in Figures 4.15 and 4.16 for 5 and 10 frames/second respectively. Temporal distortion caused by frame dropping due to rate control limitations is represented by lower PSNR values brought about by frame mismatch between the source and decoded sequences. For all bitrates in both sets of sequences, it can be seen that the PSNR values in descending order are Salesman, Foreman and Akiyo as observed in Section 4.1. What is more interesting however, is relationship between the received
quality and throughput. At 5fps, there is a rapid degradation in quality at bit rates below 40kbit/s, particularly for the Akiyo sequence caused by frame dropping at the encoder. There then follows a fairly steady increase in received quality as a function of bitrate, until a throughput of around 140 kbit/s. At around this level, the quantiser at the encoder for the Carphone sequence is set at such a level where essentially no or very little quantisation takes place, and most of the DCT coefficients representing the video sequence are transmitted. As, there is considerably more activity in the other two sequences, predictive coding is not as efficient as for Carphone. This means that there is more texture information requiring quantisation, and so such a level is not achieved until higher bitrates. A levelling of the gradient at these relatively high throughputs does however take place. The video traces exhibit similar properties when encoded at 10fps. There are however some differences. There is a severe degradation below around 50kbit/s, particularly for the Akiyo sequence and encoding at rates below 30kbit/s yields extremely poor results. The levelling off of the PSNR traces is very slight at throughputs of up to 200 kbit/s.

In order to put these results into perspective, tests were carried out to evaluate the video quality that can be achieved using the EGPRS data channels using different modulation-coding schemes. In Figures 4.17 and 4.18, the obtained PSNR values for the averaged three sequences for error-free transmission from using schemes MCS-1 through MCS-9 are shown for up to 5-channel multislotting. From these results it can be seen that there exists a greater quality difference between using one slot and two slots, than between any other multi-slotting combinations. In fact, as the number of slots is increased, the relative difference between two adjacent multislotting settings decreases. When transmitting sequences at 5 frames/second at least three slots are required to transmit video information at MCS-1, and two slots are the minimum for schemes MCS-2, MCS-3 and MCS-4. The 8-PSK data channels (MCS-5 upwards) can all support video using a single slot only. In addition, it is also evident that a progression in modulation-coding scheme between MCS-1 & MCS-2 and between MCS-2 and MCS-3 result in a greater increase than that observed between other adjacent schemes.

Similar results were obtained for the MPEG-4 quality traces for video sequences encoded at 10fps. As expected, the PSNR values for given timeslot-coding scheme combinations was approximately 2dB lower than corresponding values for sequences encoded at 5fps. In addition, it was seen that sequences can be encoded at the higher rate of 10fps using a single slot only if channel conditions allow for use of scheme MCS-7 or higher. The availability of only two slots allow for the operation at schemes MCS-5, and MCS-6, whereas three slots are necessary to
provide 10fps video using schemes MCS-2 through MCS-4. If channel conditions are such that it is necessary to use MCS-1, then four-slot operation is required.

Figure 4.17 Received Video Quality at 5 frames/second

Figure 4.18 Received Video Quality at 10 frames/second
4.4.2 Error Performance

In order to determine the effect of propagation conditions upon video sequences encoded with MPEG-4, the three sequences used previously were encoded for operation at three-slot usage for coding schemes MCS-1 to MCS-6. Full error-resilience tools were enabled with INTRA-frame spacing set to 10, video packet size set to 600 bits and Reversible Codewords and Data Partitioning enabled. Simulations were carried out at TU1.5 IFH, TU1.5 NFH, and TU50 NFH, all at a carrier frequency of 1800MHz. The results of these experiments are shown in Figures 4.19 to 4.21. The experiments were repeated for ten times for each sequence so as to ensure that meaningful averages were obtained.

When operating at TU1.5 IFH, it is seen that MCS-1 gives better performance than MCS-2 at all C/I values up to around 20 dB. At this value however, MCS-5 begins to provide a superior video quality than either of the two schemes. MCS-6 does not match this performance until at least a C/I value of 30dB. It is also seen that the MCS-3 model considerably under-performs all other codes, as does MCS-4, whose results are not displayed on the graph.

A similar trend can be seen when again operating using the TU 1.5 multipath model with frequency hopping disabled. In this case the performance is markedly inferior to the frequency-hopping case. Indeed, at a C/I of 15dB, the maximum PSNR is 19dB compared to 24dB for FH, and at 20dB, it is 23dB compared to 25dB. Such results stem from the differences in BER and BLER performance when using frequency hopping at low velocities. In such cases, the fading is very slow, bringing about high degrees of correlation between successive bursts, which the block-diagonal interleaving with a depth of 4 bursts cannot remove sufficiently to allow for optimal decoding by the Viterbi decoder. However, as seen in the non-hopping case, MCS-1 and MCS-2 provide the best performance until about 23-24dB, with MCS-5 and MCS-6 providing the best quality at higher C/I values.

The results at TU 50 NFH are however somewhat different. At low C/I values the results for MCS-1 and MCS-2 are similar to those obtained at TU1.5 IFH, although MCS-2 outperforms MCS-1 at C/I ratios in excess of around 17dB. The main deviation from the TU 1.5 models lies in the performance of the 8-PSK schemes. At TU 50, MCS-5 does not provide better video quality than MCS-1 and MCS-2 at C/I values of up to around 30dB, way in excess of the crossover figure for TU 1.5. The reason for this is the greater susceptibility of 8-PSK to intersymbol interference, a phenomenon which is more dominant at high speeds.
Figure 4.19 5fps Video Quality at TU50 NFH 1800MHz

Figure 4.20 5fps Video Quality at TU1.5 IFH 1800MHz
4.5 Conclusions

The results, in particular those describing the error performance, highlight a number of key points regarding the provision of real-time video services over GPRS and EGPRS. The most important of these, is the difference between providing services which retain data integrity at the expense of delay, and the provision of delay-sensitive error-tolerant services. This is an issue which has been already highlighted elsewhere in this report, but is crucial for the analysis of the obtained results. Whereas when providing data services, high error rates can be tolerated due to the use of hybrid retransmission mechanisms (typically block error rates of around 10%), when transmitting real-time media services, much lower error rates are required. In fact, a ballpark figure of around $10^{-3}$ is the highest bit error rate that can be sustained by a DCT-based codec such as that used in MPEG-4. This means that switching between channel coding schemes must occur at considerably higher C/I and $E_b/N_0$ values than would otherwise be required. How to guarantee these operation points will be a major issue in providing multimedia services.

As a result of this, the MPEG-4 performance simulations carried out used the ‘tail-end’ of the BLER curves shown in the validation section of this report. These areas were those, which for reasons already outlined, showed the maximum deviation from the expected results. One can therefore safely assume that the error figures shown are pessimistic and the actual performance
may be somewhat better. This will be achieved by superior receiver performance than obtained in these experiments at high carrier-to-interference ratios, and further advances in video error-resilience and concealment technologies. Irrespective of the actual figures that may be obtained from further experimentation, certain results will still hold. It is highly unlikely that the use of MCS-3 and MCS-4 will provide any useful contribution when transmitting real-time services. At the required error rates, not only does MCS-5 outperform them, but also provides a considerable throughput increase. The throughput advantage of MCS-3 over MCS-2 is marginal, and is unlikely to provide any noticeable quality improvement, even at the highest C/I values. The same argument applies for the use of schemes MCS-7, MCS-8 and MCS-9. Not only are they too vulnerable to errors, particularly when high Doppler spreads are present (high terminal velocities), but the relative increase in throughput capability over the more powerful schemes (MCS-5, MCS-6) does not justify the increase in error rates.

The sequence quality results have been measured in terms of the Peak Signal-to-Noise Ratio. Although this is the standard means of judging video quality, it has several weaknesses, the most serious of which is its one-dimensional status. As highlighted several times in this and previous reports, video, particularly sequences that have been transmitted over error-prone channels can be characterised by several parameters, including frame size, resolution, spatial detail, frame rate, chrominance accuracy, and extent of corruption. The end perceptual quality of a video sequence depends upon all these parameters, and so although it may be useful as a comparative metric, it is extremely difficult to represent the characteristics of video using a single parameter. In particular, at the crossover points in the error performance graphs, although it can be seen that two schemes result in very similar PSNR values, the characteristics of the sequences are very different. For this reason, the subjective opinion of the visual quality of video sequences is extremely important when determining the source-channel coding allocation schemes.
Chapter 5

5 Video Prioritisation Techniques

5.1 Introduction

The work described in Chapter 4 highlighted the two main issues involved in transmitting real-time video sequences over mobile packet networks such as GPRS and EGPRS, namely the limited throughput availability and the high error rates present on the mobile channel. The use of error resilience tools such as data partitioning, reversible codewords was shown to provide significant improvements in the quality that can be achieved under given error conditions, particularly with bit error rates in the order of $10^{-3}$ to $10^{-4}$. This Chapter will introduce new error resilience and video traffic management techniques that will exploit a feature that is common to all compressed multimedia, namely the differing levels of importance of different parts of video information. It will be shown that there are broadly two main aspects of this property that may be exploited by stream prioritisation algorithms. The first property is that of unequal error susceptibility to channel errors of different sections of the bitstream syntax. The second category of differentiation is more subtle, but no less important and effective in the quest to reduce the effect of channel errors. It deals with the classification of different spatial areas of video frames into regions of high or low priority according to subjective, or perceptual criteria. These techniques will be shown to provide significant improvements in the transmission efficiency of video sequences as well as improving the subjective quality of the received sequences.

5.2 Intra-Stream Video Stream Prioritisation

A large variety of techniques have been proposed to attempt to reduce the bit error sensitivity of compressed video data to corruption. These have been discussed in detail Chapter 2. However, as shown in Chapter 4, these tools on their own are still insufficiently powerful for use in mobile channels under high interference conditions.

One technique capable of delivering good results is Unequal Error Protection (UEP). As data partitioning places critical data at the beginning of each packet MPEG-4 can benefit significantly from such an approach. The scheme introduced by Rabiner et al [RABI-98] protects fixed lengths
of data with different strength convolutional codes, with data at the start of the packet receiving
the greatest protection. However, as more motion occurs in the scene, the amount of important
motion data at the beginning of the packet grows in size. This results in some of the motion
information receiving less protection than required. A similar method for improving error
resilience is the prioritisation of different parts of the video bitstream by sending the data as two
separate streams. This enables the encoder to demand that the network send the data using
channels with different priorities, allocating more important data to more reliable channels. In
[ARAV-96] MPEG-2 was prioritised in a number of different ways for an ATM network.
Methods of prioritisation included data partitioning, spatial scalability, and SNR scalability. The
work was carried out for high bandwidths, assuming no losses on the highly prioritised base layer.
Over a mobile channel it may not be possible to guarantee such a situation, both in terms of
bandwidth and base layer loss ratio. The method of data partitioning is different from that used in
MPEG-4, while the tests involved cell loss, without corruption of the data inside each cell.

Although the approach to implementing Unequal Error Protection in MPEG-4 described in
[RABI-98] gives a marked improvement in system performance, it suffers from three main
drawbacks. As the channel protection mechanisms are closely integrated with the video
compression formatting, this approach is suitable mainly for circuit-switched applications as it
does not provide for decoupling between the network layers and the higher transport and
application layers. In fact implementing such a scheme at the application layer would leave all
network and transport layer headers unprotected. A second drawback is that it makes no provision
for the occurrence of packet loss. Although noisy channels generally result in high bit error rates,
these same channel errors also cause frame and packet erasures, mainly dictated by the operation
of error check mechanisms and the loss of payload information caused by the corruption of packet
information. The UEP mechanism does not provide any protection against such packet loss.
However, the most severe drawback of this approach is the limitations it places on application
interoperability. In order to be effective, channel correction mechanisms must be carried out at the
physical link layer, where soft-decision information can be used to obtain around 3dB advantage
over hard-bit decoding. In order to implement UEP as described in [RABI-98], each underlying
network must be tailored for the properties of MPEG-4. Not only does this approach greatly
hinder transmission across diverse heterogeneous networks, but it also reduces the scope for
developing enhanced services, as improving the application capability also requires upgrading all
the underlying networks.
The approach to providing Unequal Error Protection being presented in this Chapter addresses these issues. In this scheme, prioritisation is implemented by sending the data as two separate streams. This allows for the encoder application to allocate a higher priority to the more error-sensitive data and transmit it over higher-quality more reliable channels. The proposed scheme employs a data partitioning approach to prioritisation using MPEG-4 which is then tested over GPRS PDTCH channels.

### 5.2.1 Prioritisation Technique

Data partitioning in MPEG-4 divides the video information into two sections. Header, motion, and shape data is coded in the first partition, while the less important texture information is placed in the second partition. Figure 2.2 shows this arrangement, although it should be noted that the code separating the two partitions is different from the code at the beginning of the packet. It is a simple task to produce two streams from this arrangement, using the first partition as the high priority stream and the second partition as a lower priority stream.

A major problem when transmitting a single video sequence using a number of streams, is that of synchronisation loss. It is important to ensure that there is some way by which the decoder can determine whether it is decoding data corresponding to the same frame and packet in both streams. To this end, the proposed scheme exploits the time stamping feature of MPEG-4. Time stamps are present in the Video Object Plane (VOP) header. It is therefore unnecessary to place any extra time stamps on the first partition stream. For the second partition stream, Header Extension information is added at the start of every frame. This includes time stamps, and the VOP type. VOP start codes are also placed at the beginning of every frame (see Figure 5.1a).

![Diagram](image.png)

**Figure 5.1** Extra data added to second partition stream to ensure stream synchronisation
In order to ensure synchronisation within a frame, the macroblock number of the first macroblock contained in the packet is added to each packet header in the second partition stream (it is already present in the first partition).\(^2\) This arrangement is shown in Figure 5.2b. The extra overhead incurred by this technique depends upon the frame rate, packet size (700 bits in the experiments described here), and bit-rate. For the bitstreams tested, the extra overhead generated amounted to approximately 10-13% of the overall rate. This overhead was calculated using the following information:

- 320 bits for the RTP/UDP/IP header
- 47 bits for the second partition VOP header (this is an average)

QCIF video requires 7 bits for "MB num" in the second partition packet header (implies 24 header bits per packet)

### 5.2.2 Use in Mobile Networks

This arrangement of video information into different streams will allow for true Unequal Protection mechanisms to be employed for MPEG-4 in a UMTS scenario. As described in Chapter 2, long-term viability of any of the IMT-2000 candidates is ensured by a decoupling between higher-level functions such as source encoding mechanisms and the lower-layer network-dependent functionality. This is implemented by the use of Radio Access Bearers. The key property which will be exploited to implement separate streams transmission is the requirement that a single terminal should have access to more than one bearer channel, with each separate channel being able to offer different levels of Quality of Service.

In the scheme being proposed, each stream will be transported over a different mobile bearer channel as offered by the underlying network. While this scheme allows for the error-protection mechanisms to be tailored to the specific formatting of the MPEG-4 video bitstream syntax, it also offers independence from the type of access technology that is used. As long as the bearer channels can meet the QoS parameters set by the application, then the prioritisation scheme is completely independent from the underlying network. This facilitates interoperability between similarly configured terminals communicating via different networks. QoS management will be carried out by integrating the end-to-end requirements set by the end applications with the

---

This technique was designed and implemented by Stewart Worrall, as were all the safeguards designed to prevent loss of synchronisation\(^2\)
mechanisms employed by the underlying networks. In this scenario, UEP is implemented by transmitting the two separate streams over different bearer channels meeting different QoS levels, thereby optimising performance over the access networks. This is achieved without integrating the higher-layers with network-layer mechanisms, thereby guaranteeing interworking functionality.

![Diagram](image)

**Figure 5.2 Separate streams operation for MPEG-4**

### 5.2.3 System Performance

The stream prioritisation scheme described above was tested over the simulated GPRS mobile access channels described in Channel 3. The video packets were encapsulated into RTP packets and transmitted using the UDP transport protocol. Each video frame was placed into a single RTP/UDP/IP packet. This means that a combined transport-network header overhead of 320 bits was required for every transmitted video frame. It was assumed that an error occurring anywhere within the combined RTP/UDP/IP header is uncorrectable and results in the contents of the entire packet being discarded. Errors were assumed to occur only within the radio access part of the network, with no packets being lost due to congestion or other mechanisms. The radio conditions simulated were for an interference-limited scenario, which is the most common operating scenario.
for mobile terminals. The noise at the receiver was set at 26dB, and the propagation conditions were those specified in GSM 05.05 as TU50 Ideal Frequency Hopping at 900MHz.

The results are shown in Figures 5.4 and 5.5 where each point on the graphs corresponds to the average PSNR taken from decoding the same sequence 24 times. The video sequences were coded so that the combined source-channel coded bit-rate was approximately 64 kbit/sec. Suzie and Foreman were the standard sequences selected for testing. Both are in QCIF format. Suzie was coded at 10 frames/sec, while Foreman was coded at 7.5 frames/sec. The coded sequences had INTRA frames inserted every two seconds for Foreman, and every three seconds for Suzie. The decoder employs simple error concealment upon encountering errors, which involves replacing incorrect or missing macroblocks with their corresponding motion compensated macroblocks. This is accomplished in the decoder implementation by setting the prediction errors for the motion vectors and pixel values to zero for each affected macroblock. This is a standard technique employed in many decoder implementations. The effectiveness of the error concealment scheme can be seen in the results in Table 5.1. Simulations were run as described above for the non-prioritised case, but with the error concealment functionality turned off at the decoder. Without error concealment, picture quality is completely unacceptable. However, turning error concealment on gives results that are reasonably acceptable.

Figure 5.3 PSNR of Suzie for single stream and prioritised stream transmission
The objective results in Figure 5.3 and Figure 5.4 show that clear benefits may be obtained by using the technique. Of course, the additional overhead generated from using prioritised streams leads to a small reduction in error-free quality. In relatively clean channel conditions, single stream encoding will often give superior performance. This is demonstrated by the results for 18 dB C/I for Suzie. However, such good channel conditions are unlikely in real life. It should be noted that the “crossover” in the results shown for Foreman, also occurs in the results for Suzie. The Suzie sequence’s results cross between 18 and 21 dB, which are, as previously stated, conditions that are unlikely to be encountered when using a real network.

Differences in crossover position are due to the amount of motion in each sequence. The high degree of motion present in the Foreman sequence means that, on average, the first partition of each foreman packet takes up a greater proportion of the packet than it does for Suzie. Therefore, texture data must occupy less space in order to achieve the same bitrate as Suzie. This leads to the use of coarser quantisation, resulting in lower quality video. The effect is exacerbated using prioritised streams, because the size of the first partition with respect to the second is increased by

---

**Table 5.1: Effectiveness of error concealment for Suzie, with transmission over a 15 dB C/I channel, with coding scheme CS 2.**

<table>
<thead>
<tr>
<th>Error Concealment</th>
<th>PSNR (dB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>29.26</td>
</tr>
<tr>
<td>No</td>
<td>23.58</td>
</tr>
</tbody>
</table>

---
the difference in convolutional codes applied to each stream. Thus, the reduction in quality from using the proposed scheme is much greater for Foreman than it is for Suzie. Subjective assessments also tend to reveal performance improvements from using prioritised streams. Figure 5.5 shows a frame from the middle of the Suzie sequence when subjected to a mobile channel with a C/I of 12 dB. This frame is shown because the high degree of movement at this point makes this part of the sequence highly sensitive to error. Although degradation is still visible when using prioritised streams, the quality is much more acceptable than when a single stream is used.

(a) Single stream transmission  (b) Prioritised stream transmission

Figure 5.5 Subjective test results for Suzie

5.3 Prioritisation for multi-party conferencing

The work described in this Section addresses two main problems that are encountered when implementing multiparty communications over mobile networks. In heterogeneous networks, and when operating over packet-based radio links that are affected by fading, the throughput available to each end-user will be different, and for the case of mobile users may be time-variant. In such situations, two approaches are normally available. One is to determine the lowest common bitrate available to all users, and have each user transmit at that level. This is clearly a sub-optimal solution. An alternative approach, and the one introduced in this Section is to utilise a central node which will either employ video rate transcoding or manipulation of scaleable streams so as to vary the bitrate of the video streams being forwarded to each user according the channel capacity available to each participant. The second main issue is that in bandwidth-limited error-prone channels, it is desirable to allocate the greatest share of the throughput available to any one end-terminal to the video stream being transmitted from the participant that is currently active, i.e. speaking, or to any other transmitting terminal that the end-user selects.
5.3.1 System Architecture

The systems being considered in this thesis are based on the centralised multipoint H.323 conference configuration. H.323 is an umbrella recommendation [ITU-H.323] which includes a set of standards for multimedia compression, signalling and transport over packet-switched networks that do not provide end-to-end QoS guarantees. H.323 itself is part of a wider family of recommendations for multimedia communications over a variety of underlying networks, including H.324 for PSTN, and H.320 for ISDN. As the role of the Internet as a supplier of multimedia content and as a common service interface for cross-network communications expands, so does the market penetration of H.323-compliant equipment and software.

In the network topology being examined, a centralised architecture is assumed. In this setup, all terminals send their media streams to a multipoint control unit (MCU) which then distributes selected or mixed media streams back to the terminals. In this mode of operation, the information is unicast to all the different users. Although this is more demanding on the core network than the multicast mode of operation, this is very important for the application of the multimedia processing techniques being investigated in this paper. As all terminals connected to the conference via a mobile link will be subjected to time-varying channels, it is important that the media streams be tailored for the capabilities of the access link of each individual terminal. This means that each terminal may occupy the full link capacity allocated to it for both reception and transmission of the media streams. The Multipoint Processor (MP) within the MCU will be responsible for carrying out these stream manipulation techniques. Such techniques will involve rate and resolution transcoding described in Chapter 2. These processes manipulate the video sequences by directly accessing and modifying the coded video parameters without resorting to decoding and re-encoding. Not only are these techniques much more computationally efficient, but they also have been shown to provide considerably improved video quality.
As already described, MPEG-4 operates in a very similar manner to the ITU-T's recommendation H.263 for low bitrate video coding, in that it is a DCT-based block coder which employs motion prediction. However, MPEG-4 is object-oriented, meaning that different sections of a scene may be encoded separately and then decoded and composited together at the decoder to form the final image. Video objects may be of various types, and the composition process is described by a VRML-type description. The user may, however, vary the nature in which different video objects are assembled, or even select which objects to display and which to discard. The use of MPEG-4's video-object coding properties to implementing prioritisation mechanisms will be described in this Section.

![Network Architecture for Object-oriented Multiparty Communications](image)

**Figure 5.6 Network Architecture for Object-oriented Multiparty Communications**

### 5.3.2 Traffic Characteristics

The demanding throughput requirements of most multimedia applications and the spectral scarcity and variability of mobile access links means that making the most efficient use of available radio resources should be a priority for any mobile multimedia communications system. The object-oriented architecture outlined above may be employed towards this task. Consider a scenario where five people are participating in multi-party communications, and at least one person is accessing the conference via a 64kbit/s mobile link on a terminal with a QCIF-sized screen (176x144 pixels). The multipoint processor in the MCU will receive the video streams from all five sources, and will downsample the sequences to forward to the mobile. As the available throughput is restricted, the size of each scene is reduced to 88x72 pixels. This size is sufficiently small to allow for the four 10 frames/sec video sequences to be transmitted in over a single 64 kbit/s link, which is assumed to be error-free. There now exist two options, either transmit a
composited image as a single Video Object (VO) or transmit them independently as four separate VOs. In either case, the user on the mobile link will receive a video sequence containing all the scenes multiplexed onto a single scene. This is shown in Figure 5.7.

The traffic characteristics of video formatted for optimal transmission over radio links are very bursty, with very visible peaks of Intra-coded frames. Figure 5.8 shows the traffic characteristics of the combined, single-VO sequence and the trace of the sum of the bitrates required by the situation where each downsamped sequence is encoded as a separate object. When multiple video objects are used, the Intra-coded frames may be transmitted out of sync with each other, thereby reducing the burstiness of the overall traffic on a single link, even though the overall bitrate of the two streams are similar at 67.6kbit/s for the composited scene and 74.4kbit/s for the sum of the separately-coded VOs. This smoothing out of the required traffic provides better perceptual quality at the receiver as fewer frames are discarded by the rate control mechanism for exceeding the capacity of the data link, while more efficient use is made of the link.

<table>
<thead>
<tr>
<th>Scheme</th>
<th>Link Usage (69.6kbit/s)</th>
<th>Discarded Video Frames</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Composited VO</td>
<td>85.8%</td>
<td>16%</td>
</tr>
<tr>
<td>4 Separate VOs</td>
<td>91.7%</td>
<td>1.96%</td>
</tr>
</tbody>
</table>

Table 5.2 Improvement in Traffic Characteristics using Multiple VOs
This gain in efficiency and reduction of jitter is not achieved at the expense of spatial quality. In fact, as can be seen in Table 5.3, the average PSNRs for the individually-coded sequences are higher than for the single VO composited at the MP. Moreover, this technique may be combined with the rate control mechanisms described in Chapter 4 in order to obtain greater control and flexibility over the total throughput.

<table>
<thead>
<tr>
<th>Sequence</th>
<th>Peak Signal-to-Noise Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>VO-1 'carphone'</td>
<td>30.91 dB</td>
</tr>
<tr>
<td>VO-2 'suzie'</td>
<td>27.68 dB</td>
</tr>
<tr>
<td>VO-3 'foreman'</td>
<td>34.40 dB</td>
</tr>
<tr>
<td>VO-4 'miss america'</td>
<td>35.16 dB</td>
</tr>
<tr>
<td>Average VOs 1-4</td>
<td>32.03 dB</td>
</tr>
<tr>
<td>Single VO (176 x 144)</td>
<td>31.06 dB</td>
</tr>
</tbody>
</table>

Table 5.3. PSNR Values

5.3.3 Prioritisation on multiple bearer channels

One of the main advantages of transmitting multiple video objects to a single mobile multimedia terminal is to be able to exploit multiple bearer channels offering different service classes. As described previously, UMTS specifications allow a single terminal to simultaneously access multiple bearers. Using multiple video objects allows the user or the MCU to decide which object should be transmitted with the greatest quality. In this way, high-priority bitstreams may be
transmitted over bearer channels offering superior bit error rates by employing differing rates of channel protection.

In order to assess the suitability of such prioritisation schemes, a variation of the technique described in Section 5.2 is employed. For example, assume that a user is participating in a multi-party conference with four other people via a mobile link with an average channel bit rate (including forward channel protection) of 96kbit/s. The user wishes to view all four participants simultaneously on his mobile terminal which is however restricted to resolutions of 176×144 pixels (QCIF format). As both the channel capacity and screen display are limited, all sequences are initially downsampl ed to 88×72 pixels. When encoded using all MPEG-4 error resilience tools enabled and by applying Intra frames every 15 transmitted frames, the obtained PSNR for the encoded sequence representing all four source scenes combined as shown in Figure 5.7 is 31.06 dB at a source rate of 63.8 kbit/s. If a 2/3 rate convolutional code is applied, the combined source and channel coding requirement is 95.7 kbit/s which occupies the entire available throughput.

Alternatively, if the spatially-downsampled sequences are encoded as separate objects, much more flexibility is possible. If, for example, the voice-switching software in the MCU indicates that the scene of most interest is the Foreman sequence, differing levels of channel protection and source rate allocation can be added to each sequence. This allocation scheme is shown in Table 5.4. In this scheme, it can be seen that foreman receives a larger share of the source bits and is offered greater channel protection. This allows for a greater spatial quality and less susceptibility to errors. This can be seen from the PSNR traces of the downsampled Foreman sequences decoded after being corrupted by channel errors in Figure 5.9 as well as being perceptually of much higher quality. The average PSNR for the four received VOs is 28.16 dB, which is considerably higher than the situation where the scenes are encoded as a single video object, which has a received PSNR of 26.17dB. This difference is also very visible when viewing the decoded sequences (Figure 5.10).

<table>
<thead>
<tr>
<th>Sequence</th>
<th>Source Bit rate</th>
<th>Channel Coding</th>
</tr>
</thead>
<tbody>
<tr>
<td>VO-1 'earphone'</td>
<td>11.6 kbit/s</td>
<td>×1.5=17.40 kbit/s</td>
</tr>
<tr>
<td>VO-2 'suzie'</td>
<td>14.3 kbit/s</td>
<td>×1.5=21.45 kbit/s</td>
</tr>
<tr>
<td>VO-3 'foreman'</td>
<td>20.1 kbit/s</td>
<td>×2.0=40.18 kbit/s</td>
</tr>
<tr>
<td>VO-4 'miss america'</td>
<td>10.3 kbit/s</td>
<td>×1.5=15.45 kbit/s</td>
</tr>
<tr>
<td>Total</td>
<td>56.3 kbit/s</td>
<td>94.48 kbit/s</td>
</tr>
</tbody>
</table>

Table 5.4 Rate Allocation for multiple VOs
5.3.4 Scene Segmentation

The use of Video Object prioritisation may be taken one step further by applying different priority levels to different parts of a scene originating from a single participant. In videoconferencing scenarios, the most obvious application of such techniques is the extraction of the participant, or foreground, from the background, which may be stationary or moving. Traditionally, image segmentation has been considered as being excessively computationally-intensive for real-time applications. Indeed, it is extremely complex to extract a logically-contiguous region of interest from a scene containing much spatial and temporal detail. Rather than to try to perfectly extract...
the shape of the participant and remove the background or superimpose him upon a different scene, the aim of the scene prioritisation scheme is to separate the participant from his background with the intention of transmitting the two scenes at different quality levels. As both the background and the participant will always be transmitted, albeit at different levels of quality or protection, it is expected that the visible quality degradation caused by segmentation misclassification should be fairly minor.

5.3.4.1 Segmentation algorithm for videoconferencing applications

The technique used to carry out the segmentation exploits the semantic information embedded in the MPEG-4 bitstream representation of a video sequence. It is assumed that the foreground, or region of interest, can be characterised by two main features, namely it is close to the centre of the scene, and has on average, greater temporal activity than the rest of the scene. The Sum of the Absolute Differences (SAD) is a metric used to determine the level of activity between the same macroblock in successive frames. The SAD is available to the source encoder during motion estimation, but may easily be extracted by the MP during the rate transcoding process.

The segmentation algorithm operates on a macroblock (MB) level, meaning that a 176 by 144 pixel sequence is divided into 11 by 9 macroblocks of size 16x16 pixels. The centre macroblock is selected by the algorithm as being the seed block, namely the first block for inclusion in the foreground. A metric is then computed for all neighbouring MBs consisting of the normalised SAD of the macroblock added to a factor representing its distance from the centre of the scene. The macroblock with the highest metric is then selected and added to the foreground. The subsequent macroblock is then selected from all the neighbours of the included blocks. This process continues until the number of MBs included in the foreground reaches a predetermined threshold.

5.3.4.2 Performance of scene segmentation

The algorithm was tested on the Salesman QCIF sequence, which contains a stationary background. When averaged over 100 frames, it was observed that 5.3% of the macroblocks containing foreground information were misclassified as belonging to the background. Although the rate of misclassification of background macroblocks as foreground ones was 10.78%, this leads to fewer visual artefacts than misclassification of foreground macroblocks. Similar tests were carried out for the Foreman sequence, where although there is considerably more activity in the scene, including camera panning, the misclassification of foreground blocks drops to 3.2%.
These results indicate that although the segmentation algorithm is susceptible to misclassification errors, on average approximately 95% of high priority foreground macroblocks are correctly identified, which is sufficiently high for meaningful results on traffic profiles and corresponding relative quality levels to be obtained.

![Figure 5.11 Bit allocation to foreground and background](image)

The *Salesman* sequence was encoded at 64.7 kbit/s and on decoding, the spatial quality in terms of the PSNR between the original video sequence and the decoded one was found to be 34.35 dB. When the sequence was segmented into two distinct video objects, the encoding rate for the foreground and background was set to 55 kbit/s and 10 kbit/s respectively. The achieved rates were 55.6 kbit/s and 9.5 kbit/s, giving a total of 65.1 kbit/s, slightly higher than obtained when encoding as a single video object. The traffic traces can be seen in Figure 5.11. However, the PSNR values for the foreground and the background were 36.48 dB and 29.74 dB respectively. This indicates that the spatial quality of the area of interest is much higher than the background, even though there is higher levels of temporal activity. There is no discernible evidence of the border between the region of activity and the background. However, as evident in Figure 5.12, much more detail is visible in the foreground than in the background.
5.4 Differential Quantisation

Although MPEG-4’s capability of manipulating several video objects simultaneously offers great flexibility in stream management, there is one significant drawback when carrying out intra-scene segmentation. The representation of the contours of each video object is carried in the shape information, which is very susceptible to channel errors. Although the concealment of errors in the texture information of MPEG-4 coded video has been vastly investigated in the frameworks of MPEG-2 and H.263 which both use motion-compensated DCT-based prediction, error resilience and concealment techniques for shape information is as yet an under-researched area. Shirani et al [SHIR-00] propose the use of a maximum a posteriori (MAP) estimation algorithm for the restoration of missing shape information based upon the observed image data. Although this approach is shown to be effective at high block error rates, the algorithm requires the use of an iterative estimation procedure for each block within the video frame which is computationally rather intensive. This may pose a problem in the implementation on mobile terminals, where battery life is limited.

An alternative scheme which does not require the use of multiple VOPs, and thereby sidesteps the error susceptibility problem posed by the use of shape information is therefore introduced. The MPEG-4 standard divides a scene into macroblocks 16x16 pixels in size for carrying out motion estimation. It is also possible to dictate the quantisation parameter, i.e. the level of spatial detail of the texture information, individually for each macroblock. This feature of the coding syntax can be exploited to apply different levels of spatial detail to different areas of the scene, in a technique...
which will be referred to as Differential Quantisation (DQ). The criteria for determining the priority of each area for the purposes of throughput allocation are dependent upon the application the codec is being used for. The use of Differential Encoding in two different scenarios is outlined in the next sections.

5.4.1 Video Conferencing

The segmentation algorithm introduced in Section 5.3.4 can be applied to the DQ technique. Experiments were carried out on the Foreman sequence. The segmentation algorithm consistently identified correctly the face region, which was found to contain approximately 24 macroblocks out of the 99 MBs which constitute the scene. In order to allow for a high quantiser differential between foreground and background MBs, an adaptation of the MPEG-4 Adaptive Intra-Refresh scheme was used [WORR-00]. This involves coding a fixed number of Intra blocks in each frame, which eliminates the need for non-predictively-encoded Intra frames. The frame rate was set to 5fps, while the target bitrate was 40 kbit/s. Each frame was set to contain 8 Intra MBs within the foreground of each scene, and another 2 in the background. The Intra MBs are set to have a finer quantiser than the Inter MBs, (QP set to 10 rather than 16) so that the spatial quality of the foreground is considerably higher than the background, even though there is more motion. This can be seen in Figure 5.13 and 5.14, where the average PSNR for the foreground is 30.7dB and 31.1dB for the cases without and with DQ respectively. As this foreground macroblocks contain a higher proportion of non-predictively-coded INTRA macroblocks, this area is also inherently more robust to channel errors.
5.4.2 Football Sequences

One of the most demanding applications for low-bitrate video compression is the efficient representation of sports footage. These types of video sequences are characterised by high-activity, rapid movement, camera panning, and the necessity of high-resolution display of some
areas of the scene, such as the participants in the sport. All these contribute towards making the task of encoding such sequences at bitrates suitable for transmission over mobile links notoriously difficult. In order to achieve high levels of perceptual quality at the decoder, video sequences containing football footage must satisfy a number of criteria. The most important of these is that there is a high temporal resolution and sufficient spatial detail in the areas of high interest. The quality of the scene around the area where the action takes place should be sufficiently detailed for the viewer to be able to easily discern what is happening. The approach taken in improving the quality of the video sequences exploits all these features as found in footage generated for broadcast transmission. The scene was divided into four regions, namely, in descending order of importance, the football itself, the players and officials, the pitch and the stands.

5.4.2.1 Segmentation of Football Sequences

In order to minimise computational complexity, the segmentation and classification algorithm exploited the characteristics of football sequences. The source video images were manipulated in the hue-saturation-value domain, in which the hue roughly represents the colour of the pixel being transmitted. The pitch area was identified as being all pixels whose hue value is within 2% of the median hue value of the image. This approach was found to be robust to variations in shadowing on the pitch, as these are represented by changes in saturation values, whereas the hue remains relatively unvaried. The histogram of the hue component of a typical scene clearly shows the contribution of the colour of the playing field, which is represented by a prominent peak. The main weakness of this technique lies in the sub-sampling of the chrominance components of the YUV format used as input to the video codec, which reduces the resolution of the hue component. The remaining areas are then deemed to belong to players and officials on the field, field markings and other paraphernalia and crowds on the terraces. The last of these were found to be the easiest to extract, as whenever a camera is aimed at the playing field lengthwise from a height above the field as is generally used in broadcast-quality footage, perspective dictates that the border of the pitch always enters the field of view from one of the four corners. This allows for all contiguous blocks of non-pitch areas at the corners of the scene to be classified as belonging to the terraces. All remaining macroblocks are classified as being either players or officials, with the exception of the ball, which is identified by carrying out a least-means-error search with a mask representing the football.
Video Prioritisation Techniques

Figure 5.15 shows the accuracy levels achieved by the classification for a typical football sequence. The trace can be easily divided into two distinct sections. A large portion of the frames 0 to 100 of the sequence consists of the goal area of the pitch and the spectators behind it. In this region approximately 18% of the macroblocks classified as belonging to the stand/crowd area actually contain high-importance information, either the goal area and goal keeper, or a player in front of the crowd. This high level of misclassification results in an accuracy rate of correctly identifying the players and officials of only 57.1%. In the remaining 200 frames tested, the accuracy rate is increased to 89.9%, whereas the accuracy rate for tracking the ball is 89% over the entire sequence. These results clearly show that whereas sufficiently high levels of correct decisions are made in certain portions of the sequence to allow for effective macro-block prioritisation, under certain conditions, and arguably in the most critical parts of the sequence, the segmentation algorithm fails to produce acceptable results. Several other alternative more effect image classification algorithms have been proposed for in the literature [CHOI-97], [OHNO-00], [VAND-97] which could be applied to improve the algorithm. In particular, machine vision techniques based on the Hough Transform could be used to reliably delineate the pitch limits, and object tracking should produce more effective in identifying the players and ball. The algorithm used in this thesis could however be expected to work more reliably if segmentation and classification were to be carried out on the larger (768 by 576 pixels of PAL systems) source video sequence before downsampling to QCIF for transmission over the mobile channel. Restrictions in the hardware used for video grabbing did not support such frame sizes at the
required frame rate. However, the performance of the segmentation algorithm used is sufficient to determine what quality improvements may obtained using Differential Quantisation.

The efficacy of the Differential Quantisation technique used can be seen by examining some of the obtained results. Figure 5.16 shows the result of the segmentation as performed upon a frame extracted from one of the sequences used. It can be seen, that the players are correctly identified, as is the crowd. The sequence was then encoded at a rate of 100 kbit/s at 10 frames/sec both with and without the Differential quantisation scheme enabled. A close-up view of one of the players clearly shows the improvement in the spatial quality of the player. Examination of the PSNR traces of the four regions, clearly emphasises the performance advantages. At the same rate, when using a single quantisation parameter, it can be seen that the spatial quality of the players and the football is consistently below the average PSNR for the whole sequence and well, below that of the pitch itself, which is fairly featureless. When DiffQuant is enabled, the highest quality macroblocks are those containing the highest-priority information, namely the players and the football. In fact, it can be seen that although when Differential Quantisation is used, the average PSNR is reduced from 25.3dB to 24.8dB, this is accompanied by a 2.6dB increase in the quality of the players and the football. The quality loss is restricted to a 1dB degradation in the pitch and stands areas.

![Figure 5.16 Performance of Segmentation and Classification Algorithm](image)

<table>
<thead>
<tr>
<th>Pitch</th>
<th>Players</th>
<th>Crowd</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diff. Quant.</td>
<td>26.6 dB</td>
<td>27.3 dB</td>
<td>20.9 dB</td>
</tr>
<tr>
<td>Standard MPEG-4</td>
<td>27.5 dB</td>
<td>24.7 dB</td>
<td>21.9 dB</td>
</tr>
</tbody>
</table>

Table 5.5 Peak Signal-to-Noise Ratios for Football Sequence
Figure 5.17 Comparison between standard MPEG-4 (bottom) and MPEG-4 with DQ (top)
5.5 Conclusion

The work described in this Chapter has presented three different techniques for the prioritisation of video information for the transmission over mobile channels. Stream prioritisation was shown to allow for unequal error protection to be implemented in a way that is transparent to the underlying networks and consequently can be used across diverse networks. Test results showed that the proposed technique provides considerable improvements in quality when transmitting over a mobile channel.

The perceptual quality in multiparty communications scenarios was shown to be improved by using a proposed network topology based on centralised H.323 multipoint communications for the implementation of multi-party conferencing in mobile networks. The use of object-based prioritisation schemes has been shown to give a noticeable perceptual quality improvement in received video signals under bandwidth-limited and noisy channels. The use of scene segmentation also provides improvements in spectral efficiency. This is achieved by tailoring either the throughput, and consequently the level of detail, or the amount of protection afforded to each region of interest according to set levels of importance. The problems encountered when using shape coding were solved by using the capability of the MPEG-4 syntax to be adapted to apply different levels of spatial detail to different areas of a video scene. This feature was used to enhance the perceptual quality of low-bitrate video sequences as used in video-conferencing applications and streaming video applications. The Differential Quantisation technique introduced will facilitate the implementation and enhance the quality of low bitrate video services for use in mobile and fixed low-bitrate environments as it may be used with any standard MPEG-4-compliant decoder.
Chapter 6

6 EGPRS bearer channel optimisation for real-time video services

6.1 Overview

The video transmission experiments and the video prioritisation techniques introduced in Chapters 4 and 5 clearly demonstrate the limitations of the standard GPRS and EGPRS data bearers for transporting video services. This Chapter will introduce an enhanced forward error correction scheme for use in real-time services. The code will consist of a concatenated convolutional inner code and an outer Reed-Solomon code. A soft-in soft-out Maximum Aposteriori (MAP) decoder will be used for decoding the convolutional code, while the Dorsch algorithm will be implemented to decode the RS code. An iterative convolutional-RS decoder will be shown to further enhance the error-correcting capabilities of the multi-level code. This code structure will be used to create three new modulation coding schemes for use as bearers for real-time services, and their performance will be compared to that of the standard EGPRS channel coding schemes. This Chapter will therefore introduce the two decoding algorithms used, proposed a serially-concatenated decoding structure, and examine the performance of the newly-designed bearers.

6.2 Limitations of EGPRS bearers

The results obtained in Chapter 4 clearly show the limitations of the EGPRS physical link layer for the support of real-time conversational-style multimedia services. Although efficient levels of throughput can be achieved across a wide range of carrier-to-interference ratios when operating in the acknowledge mode of transmission with ARQ enabled, the performance when supporting real-time video services is quite poor. Figure 4.20 shows that for the transfer of QCIF video using 3 timeslot allocation per user over the TU1.5 IFH 1800MHz channel can only allow for a PSNR level of 27dB at a carrier-to-interference level of 23dB. Assuming a minimum acceptable PSNR quality threshold of 25dB, the minimum C/I level required to support video applications with all standard MPEG-4 error resilience tools is approximately 18dB. This is well in excess of the...
interference levels of 7-8dB required to support Full Rate and Enhanced Full Rate speech in GSM over the TCH/FS channel.

Chapter 5 examined tools and techniques that could be used to enhance the video quality by increasing the inherent error-resilience of the transmitted streams and by prioritising the video information according to subjective quality levels. Although the techniques proposed all contributed some degree of improvement in the received quality, the recorded improvements were incremental in nature over the quality levels previously recorded. The main problem is that even employing state-of-the-art error resilience schemes, the maximum bit error rate allowable to support ‘acceptable’ quality video services is in the order of $10^{-3}$. This contrasts sharply with the bit error rates at which the optimal transfer rate of EGPRS radio blocks across the radio interface takes place, which in Chapter 7 will be shown to be approximately $10^1$ in Rayleigh fading bursty channels with multipath propagation. Using the EGPRS MCS-5 and MCS-6 coding schemes, a BER of $10^{-3}$ is only achieved at 22dB and 28dB respectively. This is usually well in excess of the operating interference level that would be supported in a cellular environment and exceeds the equivalent operating regions for packet data transfer using the same codes by approximately 13dB. This slow decay towards low bit error rates of the punctured convolutional codes used by EGPRS MCS-5 and MCS-6 is the main cause of the poor performance of real-time services of the EDGE radio interface.

Although it is desirable to improve the subjective performance of a multimedia application operating over a mobile link solely by enhancing the performance of the application without modifying the underlying radio access network protocols, the results shown in Figure 3.11 clearly show that the forward error protection mechanisms are inadequate at target error rates of between $10^{-3}$ and $10^{-4}$. Finding an alternative channel coding approach for real-time multimedia applications over the EGPRS radio channels is the objective of the work described in this Chapter.

### 6.3 Channel Coding Scheme Requirements

The forward error correction mechanisms used in EGPRS were described in Chapter 3 and are specified in [3GPP-05.03]. Essentially, all schemes make use of a 1/3 rate convolutional code of constraint length 7 with different puncturing patterns being used to obtain the desired output rates. The RLC/MAC headers are protected separately from the payload which also contains a cyclic redundancy check for error detection purposes. A wide range of different error control coding
schemes have been developed in the past few years. These schemes can be described by means of a similarly wide range of characteristics, including the type of algorithm, encoder/decoder complexity, performance under different channel conditions, the coding rate and resulting coding gain at different conditions. The required coding scheme should ideally be backwards compatible with the existing EGPRS convolutional encoder, or alternatively comprise the mother code within it. This would ease migration to the new coding scheme, and facilitate handsets to be backwards compatible with the current EGPRS standards. Additionally it is desirable that the new code should be designed so as to allow handsets designed conforming to current EGPRS standards to be able to correctly decode the received codewords. The method to be adopted should produce the greatest coding gain in the region of $10^3$ to $10^4$ bit error rate and should be particularly suited to dealing with bursty errors and Rayleigh fading environments.

6.4 Selection of Channel Coding Techniques

The most exciting developments in channel coding technology in recent years has been the discovery of iterative, or turbo, decoding techniques [BARB-99], [RYAN-01]. These employ two or more different codes arranged in such a way so as to allow for significant improvements in coding gain. Traditionally turbo codes employed two recursive systematic convolutional codes concatenated in parallel [BERR-96], although now, the term has been extended to include a wide range of different iterative concatenated codes. Iterative decoding was developed during the research into code concatenation, where codes with different error characteristics are concatenated so as to result in a powerful multilevel code without the complexity implications of long codewords or long constraint lengths. The performance of such concatenated decoding structures was greatly enhanced by the development of soft-input-soft-output (SISO) decoding algorithms such as the Soft Output Viterbi Algorithm (SOVA) and the Maximum Aposteriori (MAP) Algorithm which allowed decoders to operate, and output reliability measures of received bits, rather than simple +1, -1 values. These techniques were found to be of particular value in code concatenation, as they allowed for much more information to be passed on between decoders. In general, there are two main methods of concatenation, namely parallel and serial concatenation. In the solution to be adopted in this Chapter, serial concatenation will be used, with the EGPRS 1/3 rate convolutional code being the inner code, with the outer code being yet to be determined. Not only will this will facilitate the implementation of new coding schemes as being enhancements of the existing protection mechanisms, but research shows that serial concatenation can outperform parallel structures [BENE-96b]. In addition, parallel concatenation schemes may only reasonably be applied to two similar, but complementary codes, such as the use of two convolutional codes.
The design of parallel block-convolutional codes is, to say the least, rather problematic, and its efficacy is doubtful.

Although most of the initial work into iterative decoding was carried out using convolutional codes, there has been increasing interest in the development of 'block-turbo' codes. In particular, the performance of Reed-Solomon codes within iterative structures has been investigated. Reed-Solomon codes were first described in a June 1960 paper in the *SIAM Journal on Applied Mathematics* by Reed and Solomon and are an extremely powerful set of non-binary codes and are generally regarded as being a special case of BCH codes [WICK-95]. Rather than constructing codewords out of bits, as in all binary codes, RS codes make use of symbols from the Galois field \( GF(q) \), and must be of length \( n = q - 1 \). It has been shown however, that Reed-Solomon codes are maximum-distance separable (MDS), which means that a \((n, k)\) RS code has a minimum distance exactly equal to \((n - k + 1)\). This gives all RS codes exceptionally good error correction properties as shown in [WICK-95], and it can be shown that RS code of length \( q - 1 \) is capable of correcting any combination of \( t \) or fewer errors with no more than \( 2r \) parity check digits. Reed-Solomon codes were shown [WICK-92] to provide powerful error correction capabilities over Rayleigh fading channels and have been shown to be particularly effective in bursty error environments.

Iterative decoding of block codes was introduced by Aitsab *et al* [AITS-96] in the form of two-dimensional product Reed-Solomon codes. Product block codes form an efficient way building very long block codes by using two or more short codes. Their performance was described in [PYND-98] and has shown to perform within 0.8dB of Shannon's limit for both Gaussian and Rayleigh fading environments. Block codes have also been used as parts of multilevel channel coding schemes. Sim *et al* [SIM-98] make use of concatenated punctured convolutional codes and RS codes for protection of ATM cells in wireless environments. Aitsab *et al* extended the idea by developing a recursive structure combining a SOVA convolutional decoder and a SISO RS decoder [AITS-97]. This algorithm is shown to achieve error rates of \( 10^{-4} \) at a value of \( E_b/N_0 \) of approximately 2.5dB in a Gaussian channel. This compares with values in the order of 2.4dB - 3.2dB for RS product codes as found by Sweeney *et al* [SWEE-00]. An adaptation of this decoder structure will be adopted as not only has it been shown to produce significant coding gains at the desired error rates, but it can be constructed incorporating the EGPRS punctured convolutional code to produce a variable-rate iterative coding mechanism. In the following sections of this Chapter, the two main building blocks of the decoder, namely the SISO convolutional decoder and the SISO RS decoder will be described, followed by details of the actual codes used, and of the structure of the overall concatenated decoder.
6.4.1 Operation of the Maximum Aposteriori Algorithm

The maximum aposteriori (MAP) Algorithm, also known as the BCJR algorithm after the initials of its inventors was introduced in 1974 by Bahl et al [BAHL-74] for the optimal estimation of the states or outputs of a Markov process in the presence of white noise. Unlike the much-used Viterbi algorithm, it is an inherently soft-input soft-output algorithm, as a MAP algorithm operates by determining the probability $P[s_i \rightarrow s_{i+1} | y]$ of each state transition in a trellis given the noisy observation vector $y$. The decoder may take hard decision on the output by deciding that the output bit $x_i = +1$ if $P(x_i = +1 | \bar{y}) > P(x_i = -1 | \bar{y})$. This decision is normally taken on the basis of the sign of the log a posteriori probability ratio which is defined as

$$LLR(x_i) = \ln \left( \frac{P(x_i = +1 | \bar{y})}{P(x_i = -1 | \bar{y})} \right) \quad - (6.1)$$

The dynamics of time-invariant convolutional codes are completely specified by a single trellis section which describes the transitions, or edges, between the states of the trellis at successive time instants $i$ and $i + 1$. In order for a code to be uniquely decodable, given an initial trellis state there must be a one-to-one correspondence between input sequences and corresponding state changes. This trellis characterisation of the code allows for the log likelihood ratio to be expressed as

$$LLR(x_i) = \ln \left( \frac{\sum_{s_i} P[s_i \rightarrow s_{i+1} | \bar{y}]}{P(\bar{y})} \cdot \frac{\sum_{s_i} P[s_i \rightarrow s_{i+1} | \bar{y}]}{P(\bar{y})} \right) \quad - (6.2)$$

where $s_i$ is the state of the encoder at time $i$, $S^+$ is the set of ordered pairs $(s', s)$ corresponding to all state transitions caused by the data input $m_i = +1$, $S$ is the corresponding set for the input $m_i = -1$, and $\bar{y}$ is the received observation vector. In this case, it will be the output of the equaliser. This means that the output is expressed in terms of the relatively likelihood of all state transitions produced by a message input of +1, given the received vector, as compared to the corresponding state transitions for an input of -1.
Bahl et al make use of the fact that a convolutional code is a Markov process to derive a solution to the above expression. As the $P(y)$ terms cancel out, the numerator can be expressed in the following way:

$$P[s_i \rightarrow s_{i+1}, \hat{y}] = \alpha(s_i) \gamma(s_i \rightarrow s_{i+1}) \beta(s_{i+1})$$ - (6.3)

This represents the probability that the current state is $s_i$ given the noisy observation vector $y_0, \ldots, y_{L-1}$. As the initial state of the encoder is assumed to be zero, $\alpha_0(0) = 1$ and $\alpha_0(s \neq 0) = 0$.

$\alpha(s_i)$ may be computed recursively as

$$\alpha(s_i) = \sum_{s_{i-1} \in S} \alpha(s_{i-1}) \gamma(s_{i-1} \rightarrow s_i)$$ - (6.4)

Similarly, the probabilities $\beta(s_i)$ which are defined as

$$\beta(s_i) = P[(y_{i+1}, \ldots, y_{L-1})|s_{i+1}]$$

may be computed in a backward recursion as

$$\beta(s_i) = \sum_{s_{i+1} \in S} \beta(s_{i+1}) \gamma(s_i \rightarrow s_{i+1})$$ - (6.5)

with boundary conditions $\beta_L(0) = 1$ and $\beta_L(s \neq 0) = 0$, meaning that the encoder is expected to end in state 0 after $L$ input bits and must consequently be properly terminated.

[RYAN-01] shows that forward recursion of the MAP can be expressed as:

$$\alpha_i = \sum_{s_{i-1}} \gamma_i(s_{i-1} \rightarrow s_i) \alpha(s_{i-1})$$

and the backward recursion as

$$\beta_{i-1} = \sum_{s_{i+1}} \gamma_i(s_i \rightarrow s_{i+1}) \alpha(s_{i+1})$$ - (6.7)
The branch metric, \( \gamma \) represents the probability of the state transition from \( s_i \) to \( s_{i+1} \), given that the current state is \( s_i \) and can be expressed as:

\[
\gamma(s_i \rightarrow s_{i+1}) = P[s_{i+1}|y_i|s_i] \tag{6.8}
\]

This represents the probability that the final state is \( s_{i+1} \), with the starting state not being considered. Once again, applying Markov principles, the Branch Metric can also be expressed as

\[
\gamma(s_i \rightarrow s_{i+1}) = P[s_{i+1}|s_i]P[y_i|s_i \rightarrow s_{i+1}] = P[m_i]P[y_i|x_i] \tag{6.8}
\]

\( P[y_i|x_i] \) is a function of the modulation scheme used and the channel characteristics and is the Euclidean distance between the received (noisy) output and all possible encoder outputs. \( P[m_i] \) is the a priori information fed to the decoder. For the 1st iteration there is no a priori information available to the decoder, and consequently \( P[m_i = 1] \) is equal to 0.5. However, when used as part of an iterative decoder, the outputs of other sections of the decoder may be used to provide an estimate of the expected output.

**6.4.2 Soft Decision Decoding of Reed Solomon Codes**

Full maximum-likelihood soft-decision decoding of linear block codes is prohibitively complex for all but the simplest codes [SWEE-00]. Much research has therefore gone into the development of sub-optimal algorithms which approach maximum-likelihood performance, while achieving much reduced complexity. Two algorithms, in particular have shown to produce good results at low complexity, namely the Chase algorithm [CHAS-72] and the Dorsch algorithm [DORS-74].

The Chase algorithm is a repetitive algebraic method which uses the soft decision information to identify low reliability symbols. A series of algebraic decodings with alterations (referred to as test patterns) are carried out to the low reliability symbols, in the hope that some errors will be corrected, and that the algebraic decoder will be able to correct the remaining errors. The best solution is found by comparing the Euclidean distances from the received sequence to each of the decoded solution. This algorithm has been extended by Pyndiah et al [PYND-98] to produce a soft decision output. In work carried out by [SWEE-00], it was shown that the Dorsch algorithm is better suited to decoding block codes with the objective of producing a soft output.
The Dorsch Algorithm is considered as being a repetitive algebraic erasure-filling decoder. Assuming \( n \) code symbols and \( k \) information symbols, soft information is used to choose a high reliability information set of \( k \) symbols, with the remaining \( n - k \) symbols being erased. Test patterns are then generated on the information set, each time re-encoding to generate a codeword. Essentially, this means that the decoder aims to select the most reliable set of bits subject to the constraint that they must form an information set. Each decoding attempt is actually a re-encoding operation which involves relatively little operation once the re-ordering has been carried out. The decoding stops when a 'decoding threshold' has been crossed or the maximum time allocated has run out. The main advantage of the Dorsch algorithm is its complexity. Whereas the Chase algorithm requires an algebraic decoder, the Dorsch algorithm has only to sort the received bits once according to their reliability. Although exact comparisons of complexity are difficult, whereas the complexity of the Chase algorithm is proportional to the number of decoding attempts, the Dorsch algorithm has a overhead associated with sorting, but then only a low increase in complexity for each further decoding attempt. In addition, each new re-encoding attempt is guaranteed to produce an output whereas the Chase algorithm may fail to produce an output. This has been shown [SWEE-00] to produce better soft output decisions. For these reasons the Dorsch algorithm was used to decode the Reed-Solomon codes.\(^3\)

The soft decision output is computed following decoding of the received RS codewords. In a method introduced in [AITS-98], the soft decision is regarded as being a reliability measure of the decision made. The codeword with the minimum Euclidean distance to the received sequence \( r = (r_1, r_2, ..., r_n) \) is minimal for all codewords is the maximum likelihood solution. Assuming that the soft decision output \( d = (d_1, d_2, ..., d_n) \) is this solution, and there exists another solution \( c = (c_1, c_2, ..., c_n) \) which can be considered as a competing codeword of \( d \) at a minimum Euclidean distance from \( r \) with \( c_i \neq d_i \), then it can be shown that the soft output is given by the following equation:

\[
 r_i' = \left( \frac{\sum_{i=1}^{n} (r_i - c_i) - \sum_{i=1}^{n} (r_i - d_i)}{4} \right) d_i
\]  

\(^{(6.9)}\)

This expression essentially compares the difference between the decoded solution and the 'second-best' solution which is also close in terms of its Euclidean distance to the received

\(^3\) The implementation of the Dorsch algorithm employed was designed and implemented by Steve Wesemeyer and David Burgess
sequence. On some occasions, such a solution may not be found within the given number of decoding attempts dictated by the required complexity. In these situations, an approximation is used based on the channel and code characteristics. In the method described in [AITS-98], the following formula is used:

\[ r_i' = \beta(m) \times d_i \]

where \( \beta(m) \geq 0 \) and \( m \) is the \( m \)th half iteration.

The value for \( \beta \) typically increases with each iteration, reflecting an increase in confidence of the estimate of the soft information.

### 6.4.3 Design of an Iterative Decoder

As described above, a MAP decoder was used to decode the convolutional code. The 1/3 rate punctured code specified in [3GPP-05.03] was used as an inner code. This has constraint length 7 and is specified by the following generator polynomials:

\[
\begin{align*}
G_0 &= 1 + D^2 + D^3 + D^5 + D^6 \\
G_1 &= 1 + D + D^2 + D^3 + D^6 \\
G_2 &= 1 + D + D^2 + D^6
\end{align*}
\] - (6.10)

As interleaving is used between the inner and outer codes, to spread the errors which the MAP decoder cannot correct, a deinterleaving block must be used between the two decoders. The LLR estimates are then fed into the RS decoder which in turn produces a soft output of the decoder codewords \( D = [d_1, d_2...d_n] \). This represents the \textit{a priori} information that should be fed into the MAP decoder in successive iterations as it contains the refinements to the soft decision output of the MAP decoder as produced by the RS decoder. The \textit{a priori} information during the first iteration is obviously set to zero (in terms of LLR), except for the tailing bits, which are all set to \(-1\) in order to guarantee correct termination. Aitsab \textit{et al} [AITS-98] showed that the LLR output of the convolutional decoder can be written as:

\[
\text{LLR}(x_i) = \frac{2}{\sigma^2} y_i + w_i
\] - (6.11)

where \( w_i \) is a function of the redundant information as introduced by the encoder and \( \sigma^2 \) is the variance of the channel noise. In general, \( w_i \) has the same sign as the information bit \( m_i \) and does
not depend on the decoder input $y_i$. This characteristic may be used to improve the estimate of the LLR($x_i$) with each iteration. Assuming that at iteration $p$, the RS decoder produces an estimate of the input codewords $Z_p$, then the LLR generated by the inner decoder is now:

$$LLR(m_i) = \frac{2}{\sigma^2} y_i + \frac{2}{\sigma^2} z_{i}^{p-1} + \omega_i$$

where $\sigma^2$ is the variance of the extrinsic a priori information $Z^{p-1}$ provided by the RS decoder in the previous iteration. This information should not be forwarded for RS decoding and must therefore be subtracted from the output of the MAP decoder. A scaling factor $\gamma$ is used to represent the variance of the intrinsic information. Pyndiah et al recommend setting $\gamma_p$ to the same value as the coefficient $\alpha_p$ which is used to reduce the influence of $d_i$ in the first iterations, where the residual BER is expected to be relatively high and consequently not very reliable. The values recommended for $\alpha_p$ were [0.6, 0.8, 1,1]. These values are dependent upon the combination of channel conditions and modulation scheme used, and experimentation carried out in developing this decoder has shown that scaling factors of [0.8, 0.8, 0.8, 0.8] produce better results in the multipath fading channels used in testing.

**Figure 6.1 Iterative Decoder for concatenated Convolutional-RS codes**

6.4.4 Comparison between MAP and Viterbi Decoder

The main objective in designing a new set channel coding schemes for EGPRS is to be able to extend the carrier-to-interference envelope that would produce residual BER levels low enough to
support acceptable quality video communications. The differences in BER performance, and consequently in video quality that should be measured in these experiments must be the result of the inherent differences between coding methods, rather than between differences between decoding implantations acting upon the same code. As described above, the Maximum Aposteriori algorithm was used instead of the Viterbi decoder employed in the experiments described in Chapters 3 and 4. A comparison of the performance of these two decoders operating on the MCS-5 scheme was carried out under the TU1.5 1800MHz multipath channel model using ideal frequency hopping. This will be the channel propagation environment used for all the experiments described in this chapter. The two algorithms were shown to provide very similar performance when producing hard decision outputs, with the Viterbi decoder providing marginally better performance (approximately 0.3dB) at C/I levels below 12dB, while being outperformed by up to 0.5dB at C/I levels above 21 dB by the MAP decoder. In the region between these values, the difference between the two decoders is minimal. These results are in line with the findings of [PIET-94], where it was shown that for most values of E_b/N_0, the Viterbi and MAP algorithms can be shown to provide similar performance. This means that whatever improvements may be obtained using the proposed concatenation structure will not be caused by

![Figure 6.2: Comparison between performance of Viterbi and MAP decoders](image)

**Figure 6.2** Comparison between performance of Viterbi and MAP decoders
6.4.5 Interleaver Design

Convolutional codes and most block codes in common use have random error correcting capabilities. This means that optimal error correcting performance is obtained when the errors are randomly distributed throughout the blocks or sequences, as the occurrence of several errors in a ‘burst’, or clumped together, is likely to cause a failure in the decoding mechanism. Reed-Solomon codes, do however have a limited bursty-error correction capability, as each codeword is constructed from code symbols over GF(2^p), rather than from individual bits. Although this property reduces the effect of bursty errors, if the corrupted received bits are confined to within the boundaries of a single symbol, RS codes are still sensitive to the occurrence of several symbol errors within a single codeword. The burstiness of a received error pattern can be reduced by employing an interleaver, which mixes up the symbols from several codewords so that the symbols from any given codeword are well separated during transmission. When the codewords are reconstructed by the interleaver, error burst introduced by the channel are broken up and spread across several block words. In essence, an ideal interleaver creates a random channel from a bursty one.

The EGPRS channel protection specifications require the use of block-diagonal interleaving following the forward error correction mechanisms. This reduces the effect of the burstiness of the fading channel, by spreading bit errors throughout the length of the interleaved frames. Experiments described in [BARB-99] demonstrate the effectiveness of these interleaving mechanisms in the propagation environments being investigated. The EGPRS interleaver is however only effective in optimising the performance of the inner convolutional code. As the decoded output of the convolutional code is also bursty, another interleaver must be used between the inner and outer codes. Two alternative forms of interleaver were used, a block-symbol interleaver and a block-bit interleaver. In both types of interleaver, the coded data stream is read into a \((n \times m)\) interleaver row-by-row, as shown in Figure 6.3. The interleaver contents are then read out in a column-by-column fashion, meaning that any two adjacent symbols at the input are separated by \((n - 1)\) other symbols at the output. The only difference between the two interleaver structures is that, as the name implies, the symbol-block interleaver retains the symbol boundaries, and the data stream is interleaved on a symbol basis, whereas the bit-block interleaver breaks up the symbols and interleaves at a bit level.
The performance of bit and symbol interleavers was examined using a concatenated MCS-5 – Reed Solomon code structure. As described in Chapter 3, the MCS-5 coding scheme has an information payload of 480 bits per radio block, with a further 6 bits being used for tailing. The implementation of the Dorsch algorithm employed in this study was capable of decoding RS codes over GF($2^4$), which requires a symbol length of 4 and a block length of $(15 \times 4) = 60$ bits for the $[15,11]$ code used in these experiments. In order to simplify both the encoding and decoding implementations, it was required that an integer number of MCS-5 data payloads should fit into a single interleaver data block. If, for example, an interleaving depth of 8 were to be used, this would result in interleaver block size of $(8 \times 60) = 480$ bits, which is exactly the size of a single MCS-5 radio block. Consequently, the interleaver depths used were integer multiples of 8, and the results of the concatenated code using various interleaver combinations as compared to the performance of the MCS-5 code using the MAP decoder are shown in Figure 6.4. The Reed-
Solomon decoder was set a maximum number of decoding attempts of 500, as [SWEE-00] showed that this provided the most effective compromise between error-correcting capability and implementation complexity.

The results clearly show that for a concatenated convolution-RS code using the Dorsch algorithm, the bit interleaver provides much better performance at all interleaving depths. At a depth of 56 and a BER of 10^{-3}, the difference between the symbol and bit interleavers is approximately 4 dB. In fact, it can be seen that at the same BER level of 10^{-3}, the performance improvement using the concatenated coding scheme with a 56-bit deep interleaver is only of 1.3 dB as compared to that obtained using the MAP decoder only. This is extended to 5.3 dB for the symbol-based interleaver. Even a 24-deep bit interleaver outperforms the 56-deep symbol interleaver.

These results are in line with those obtained by Sweeney et al, where a gain of approximately 1 dB was obtained for an iterated product RS code between symbol and bit interleavers at a bit error rate of 10^{-3} using a Gaussian channel. The reason for the advantage of bit interleavers lies in the nature of the Dorsch algorithm. As this is essentially an algorithm for dealing with binary codes, non-binary codes defined over GF(2^p) such as RS codes, have to be treated as binary codes by using a binary representation of the field elements. This means that there is no longer an advantage to be obtained by retaining symbol integrity while interleaving, as is the case for most hard-decision RS decoding methods, thereby resulting in improved performance for the bit-ordered interleaver.

The use of an interleaver depth of 56 requires 3360 bits to be buffered, which at two-slot operation implies an extra delay of 70 ms. This is reduced to 46.7 ms when three slots are allocated. As the schemes being examined are being designed for use in real-time conversational applications, it is required that any coding gains without introducing excessive extra delay. For this reason, an interleaving depth of 60 will be the upper limit for the interleavers to be used in the proposed solution.
6.4.6 Comparison of Soft and Hard Decision decoding

Decoding of the Reed-Solomon code is carried out using the Dorsch algorithm because of its ability to make use of soft input information in the decoding process. The effectiveness of transfer soft information between the MAP decoder and RS decoder is shown in Figure 6.5, where the decoding performance of an non-iterative decoder using soft information transfer between the inner and outer decoders is compared to the performance of an equivalent decoder which employs a hard-input Dorsch decoder. The concatenated decoder is seen to provide a significant advantage over the non-concatenated MAP decoder only if the transfer of soft information is enabled. The implementation of the MAP decoder used in these experiments outputs the reliability measures in terms of the Likelihood Ratios, whereas the RS decoder quantises the probability measure that the received bit is a 1. The following expression is used to convert between the two methods of representing the soft information:

\[ P(x_i = +1 | \hat{y}) = \frac{LR(x_i)}{1 + LR(x_i)} \]  

- (6.12)

Figure 6.5 clearly indicates that the soft information is being effectively transferred between the two successive decoding blocks, as a coding gain of 5dB is seen between the soft concatenated
codec and the MAP MCS-5 decoder. The gain is however negligible if hard information bits are transferred between decoders. Although an improvement due to soft decision decoding was expected, the difference in performance between the two methods is surprising, particularly as concatenation has not shown to provide any advantage whatsoever when carrying out hard decision RS decoding using a [15,11] code.

In order to assess whether the poor performance of hard decision concatenation is the general case for RS codes of length 15, or whether it is a property of the Dorsch algorithm being employed, tests were carried out with a slightly modified decoding structure. A hard decision RS decoder based on the Euclidean / continued fraction algorithm was used to decode symbol-ordered output from a hard-output Viterbi decoder as used in the experiments described in Chapter 3. Two outer codes with similar rates were used, namely the [15,11] code used thus far, and a longer [63,45] rate 0.714 code. The performance of the [15,11] concatenated code is similar to that obtained using the Dorsch algorithm, where a BER of $10^{-3}$ is obtained at a C/I ratio of approximately 21dB. The [63,45] code does however significantly outperform the shorter code when hard decision decoding is used. In fact, it can be seen that when using a symbol-based interleaver of depth 8, which is roughly equivalent in length to the 56-deep interleavers used for the GF(16) code, the error performance matches very closely that of the [15,11] code with soft decision decoding. The performance advantage of the longer code lies in the larger minimum distance and consequently in the stronger error correcting capabilities. It can be shown that the [63,45] and [15,11] codes
used can correct 2 and 9 errors respectively. These results show that for hard-decision concatenation in the channel environments used, a code word length of 15 is inadequate, irrespectively of the actual decoding algorithm used, and does not provide any significant performance improvement. A block length of 63, however, gives an improvement of 4.5 dB at $10^{-3}$. If soft decision decoding is enabled, the shorter [15,11] RS code exceeds the performance of the longer code by approximately 0.4dB. Soft decision decoding of the longer codes is prohibitively complex, unless product codes are used to create long codes from shorter constituent ones.

![Figure 6.6 Comparison between concatenation with soft [15,11] and hard [63,45] RS decoders](image)

**6.4.7 Iterative decoding**

The main reason for employing soft-decision output estimation at the RS decoder is to be able to construct an iterative decoder, the structure of which has been described previously. Pyndiah et al determined that the optimal compromise between decoding performance and computational complexity was obtained when 4 iterations were used. Given a Gaussian channel, an improvement of approximately 0.6dB was observed between the concatenated coder using a single iteration, and one using 4 iterations. Experiments carried out as part of this investigation confirm that a notable improvement is indeed obtained when iterative decoding is used. At a BER of $10^{-3}$, an improvement of 1.3dB is observed, increasing to 1.8dB at $10^{-4}$. These results show, that although iterative decoding does significantly reduce the residual error rate, the greatest reduction in errors is observed when migrating from a single convolutional decoder to an iterative structure. The
results obtained in these experiments used a constant level of \( \alpha \) of 0.8 for all iterations. Although Pyndiah et al. reported optimal performance at values of \([0.6, 0.8, 1, 1]\) for the AWGN channel, constant values of 0.8, provided better results in the given channel conditions. The optimal values of these parameters depend upon the particular combination of modulation scheme and channel used. In the varying channel conditions experienced by mobile terminals, different sets of parameters can be seen to work best under different conditions, and it is probable that only adaptive values will give optimal performance across the board. This should be an area for further research.

![Figure 6.7 Effect of Iterative Decoding](image)

### 6.5 Design of EGPRS Coding Scheme for Video Applications

The results that have been obtained so far in this Chapter clearly indicate that a combination of convolutional and Reed-Solomon decoding carried out in an iterative way can produce significantly better performance than that obtained using the standard EGPRS schemes, particularly in the \(10^3 - 10^4\) region, in which mobile video communications systems operate. An alternative range of coding schemes using this encode-decode structure is therefore proposed for use in multimedia communications. Three new coding schemes are introduced, which will be referred to as EMCS-1 (Enhanced), EMCS-2 and EMCS-3. A diagrammatic layout of the data flow between the different sections of the encoder/decoder pair is shown in Figure 6.8.
Figure 6.8 Data Flow of Proposed Concatenated Coding Schemes
The schemes designed are based on the standard EGPRS channel coding schemes MCS-5 and MCS-6 with slight modifications. EMCS-1 makes use of a (7,15) code of rate 0.47. In order to simplify the mapping of RS blocks of length 15 into the MCS-5 payload which consists of 448 bits, an interleaver of length 56 blocks was used. This allows the resulting 3360 bits to be encapsulated into exactly 7½ MCS-5 radio blocks. The resulting concatenated scheme has an overall rate of 0.173. The second scheme, EMCS-2 makes use of the same mapping as used in EMCS-1, with the interleaved RS blocks being carried by 7½ MCS-5 radio blocks. The overall code is however less powerful, as the same (11,15) RS code described in Sections 6.3, is used instead of the (7,15) code employed by EMCS-1. This produces a code with an overall rate of 0.2713.

Unlike the first two schemes, EMCS-3 uses the less powerful MCS-6 convolutional code as its inner code. This means that an alternative mapping arrangement must be used to accommodate the RS blocks into the EGPRS radio blocks. The scheme chosen makes use of an interleaver of depth 60 blocks. Assuming an extended MCS-6 payload of 600 bits, rather than 592 bits, this allows the resulting 3600 bits to be easily mapped onto 6 radio blocks. This code has an overall rate of 0.360. The properties of these three new codes are summarised in Table 6.1.

<table>
<thead>
<tr>
<th>Coding Scheme</th>
<th>Number of RS blocks</th>
<th>EGPRS inner code</th>
<th>Number of Radio blocks</th>
<th>Number of source bits/radio block</th>
<th>Code Rate</th>
<th>Bitrate (kbit/s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>EMCS-1</td>
<td>56</td>
<td>MCS-5</td>
<td>7.5</td>
<td>209</td>
<td>0.173</td>
<td>10.45</td>
</tr>
<tr>
<td>EMCS-2</td>
<td>56</td>
<td>MCS-5</td>
<td>7.5</td>
<td>328.5</td>
<td>0.271</td>
<td>16.42</td>
</tr>
<tr>
<td>EMCS-3</td>
<td>60</td>
<td>MCS-6</td>
<td>6</td>
<td>440</td>
<td>0.360</td>
<td>22.00</td>
</tr>
</tbody>
</table>

Table 6.1 Properties of enhanced modulation-coding schemes

6.5.1 Results

The performance of the coding schemes described above was evaluated under the TU 1.5IFH multipath propagation conditions used previously in this Chapter. The iterative decoding algorithm described in Section 6.3.3 was used with 4 iterations, and 500 decoding attempts for the Dorsch algorithm. Soft decision decoding was employed for both the convolutional code, as well as for the Reed-Solomon code. The residual BER traces are shown in Figure 6.9.
Schemes EMCS-1 and EMCS-2 are constructed from the concatenation of the MCS-5 convolutional code and Reed-Solomon codes of differing strengths. When decoded iteratively, this combination is seen to provide very significant improvements in performance, particularly at lower values of residual BER. At a bit error rate of $10^{-3}$, the improvement is between the error performance of MCS-5 and of the two concatenated codes are 10.4 dB and 7.4 dB for EMCS-1 and EMCS-2 respectively. Although these are indeed large improvements in performance, it must be remembered that the overall rate of these concatenated codes are 0.173 and 0.271 as compared to 0.37 for MCS-5. EMCS-3, however uses MCS-6 as an inner code and has an overall rate of 0.36, which is very similar to that of MCS-5. Nonetheless, this code exhibits an improvement of 3.4 dB at a BER of $10^{-3}$ and of approximately 5.9 dB at $10^{-4}$. The sharp cut-off typical of Reed-Solomon codes is seen to produce significantly enhanced performance at the optimal BER levels for the transmission of video information. This improvement is however not seen across all operating conditions. In fact, at higher residual error rates, no or very little improvement can be observed, and the extra overhead introduced by the concatenation of the Reed-Solomon code is wasted. These results clearly show that the characteristics of the convolutional codes and Reed-Solomon codes are complementary, and when concatenated and decoded iteratively, can be used to produce an extremely powerful code, particularly at low BER levels.
6.5.2 Performance of Video Transmission

Although the residual bit error rate is a useful metric in assessing the efficacy of channel coding schemes, it is not the sole standard by which the performance of forward error correction schemes should be assessed. Of equal importance in many applications, is the error event rate, or the residual error characteristics. In general, channel errors at the decoder may be split broadly into two categories. The first, and more benign, are errors that are within the error-correcting capabilities of the code being used, and which therefore can be properly corrected by the decoder, while the second category refers to the channel errors that exceed the code’s error-correcting capability, and consequently result in uncorrected errors. This has particular implications when operating under bursty channel conditions, as for a given error rate, bursty channels contain far fewer error events than random channels, and in general, there is a much higher ratio of uncorrectable errors to errors that can be corrected by the decoder. This means that the properties and characteristics of the particular forward error correction employed needs to be assessed with a view to the application it will be employed for, such as packet data transmission with backward error correction, or real-time multimedia communications.

For this reason, the performance of the coding schemes developed in this Chapter was tested in a video communications scenario. The Foreman sequence used in the experiments described in Chapter 4 was encoded using MPEG-4 with full error resilience tools enabled. Data partitioning and the use of Reversible codewords was enabled, while the INTRA frame spacing was set to 10 frames and the video packet size was set at 700 bits. The sequence was encoded at an output frame rate of 5 frames/second with rate control enabled so as to allow for a source throughput equal to the capacity of the selected modulation-coding scheme selected with three timeslots allocated to the mobile terminal. The results of these experiments can be seen in Figure 6.10.

The main objective of introducing this new set of channel coding schemes was to reduce the carrier-to-interference requirements necessary to support real-time video communication services. Figure 6.10 shows the quality levels in terms of the decoded Peak Signal-to-Noise Ratios (PSNR) that can be obtained using the three coding schemes at a range of C/I levels. The significant improvement that is obtained using the concatenated coding scheme is immediately apparent. Assuming a PSNR value of 27dB as the minimum value that would allow for acceptable quality video services, then using the standard EGPRS schemes, a carrier-to-interference ratio of 21dB is required. This is reduced to 10.4dB when using EMCS-1, an improvement in excess of 10dB. Figure 6.11 shows the variation in C/I gain obtained at different PSNR levels by using the new proposed schemes as compared to the EGPRS coding schemes. In this diagram it can be observed
that the difference in performance between the two sets of schemes increases as the required video quality improves.

Figure 6.10 Performance of Video Transmission over Enhanced Radio Bearers

Figure 6.11 C/I gain obtained using enhanced radio bearers
Similar improvements are observed when examining the quality levels obtainable at different C/I values. For example, at a C/I ratio of 21dB, MCS-1 provides a quality level of 27dB, whereas EMCS-3 supports a PSNR value of 32.1dB. This enhancement in quality is observed across the entire range of interference levels, decreasing to within a PSNR difference of less than 2dB only at C/I values in excess of 28dB. This reduction in quality difference can be attributed to the fact, that when using EMCS-3, errors are reduced to effectively zero (i.e. to a level that results in negligible video quality distortion) at a C/I of approximately 21dB, and a levelling off in PSNR values is observed beyond this point. It is also interesting to observe that the concatenated codes produce a very sharp cut-off in video quality as the operating C/I level is decreased, whereas the convolutional codes produce a far more gradual roll-off. This is in line with BER traces observed in Figure 6.9, where it was shown that the code concatenation produces far steeper curves at low BER values than the EGPRS codes.

6.5.3 Application-level Channel Protection

The results shown in Figures 6.9 are for an iterative decoder with soft decision decoding, which was demonstrated to provide the best results given the code structure used. This requires the channel decoding to be carried out at the physical layer so as to be able to use soft information which would not be available at any other layer of the protocol stack. However, one of the main aims of the new set of forward-error correction mechanisms was for backwards-compatibility with existing standards. In other words, the extra parity bits introduced by code concatenation could equally as well be decoded at the physical layer as well as at the application layer. Application-level decoding would take place if a standard EGPRS terminal was being used, which would only recognise the punctured convolutional codes specified in [3GPP-05.03], with the remaining RS section of the concatenated code being decoded in a non-iterative fashion by the application currently operating on the terminal. A proposed data flow structure which allows for application-level encoding/decoding is shown for the source terminal in Figure 6.12.

In order not to cause any modification to the channel coding specifications, three main elements are added to the video packet, or whatever type of multimedia information is being transmitted. The first two are the RS parity bits and the interleaver as have been described throughout this Chapter. In addition, a packet starter and termination code is added at the beginning and end of the interleaved information. These should contain a unique bit pattern of sufficient length to allow for a minimal risk of the same pattern occurring within the data payload, and their function is to indicate to the joint convolutional-RS decoder where the RS-encoded data starts and ends. In this
way, standard EGPRS channel decoding is carried out over the headers and LLC frame check sequence, while all application-level data is also protected by the RS code. If application-level decoding is used, the starter and termination codes are ignored.

![Diagram of data flow](image)

Figure 6.12 Data Flow for Application-layer decoding

Although this scheme allows for backwards compatibility with terminals whose channel protection mechanisms are specified in [3GPP-05.03], it must be noted that the performance of application-level decoding of RS codes of length 15 is not expected to provide any significant coding advantage as was shown in Section 6.3.6. In fact, if RS codes are to be implemented at the source encoder/decoder, longer codes of length 63 are likely to be much more effective, although they introduce significant decoding complexity.
6.6 Conclusions and Further Work

This Chapter has introduced a concatenated Reed-Solomon coding scheme for the transport of real-time video sequences over EGPRS. An iterative decoder containing a Maximum Aposteriori convolutional decoder and a soft-output RS decoder based on the Dorsch algorithm was shown to produce the best decoding capabilities for the code used. When compared to standard EGPRS schemes, improvements of up to 12dB were observed when operating in interference-limited environments.

The code concatenation employed was only shown to provide significant improvements when iterative decoding was used, thereby implying operation at the physical link layer. Experiments showed that longer-length codes must be used to obtain comparable performance using the transfer of hard-decision information. For this reason, more research needs to be carried out into the development of coding strategies that would allow for effective application-layer decoding. A method for ensuring backwards-compatibility with current EGPRS schemes was introduced, but this was shown only to provide optimal performance with iterative decoding. A technique which has much potential for improving the error resilience of MPEG-4 video is the insertion of user-defined data as described by Worrall et al [WORR-00b]. This method allows for the introduction of extraneous information in the MPEG-4 while still retaining backwards-compatibility with standard decoders. As Reed-Solomon codes are systematic codes, the parity bits could be used as the ‘extraneous’ information, allowing for customised codecs to make use of the RS decoded information, while standard decoders would ignore the extra bits.

The work carried out has focussed on the use of one particular coding strategy. The 3GPP UMTS standards specify the use of turbo codes constructed from convolutional codes, and these have been shown to provide good performance under a wide range of channel conditions [MELI-00], including AWGN channels and Rayleigh fading channels. A more thorough comparison between these schemes and those discussed in this thesis would determine which approach is the most suitable for multimedia applications. Although the experiments carried out were restricted to determining the performance of the codes with respect to the transport of video streams, the schemes developed may be equally applicable to speech services. The provision of conversational speech over GERAN is currently the subject of much research and the techniques proposed could make a valid contribution to the field.
Chapter 7

7 Streaming multimedia services

7.1 Introduction

This Chapter will examine the performance of the transmission of streaming multimedia services over EGPRS access networks. The characteristics of streaming services, and their resulting QoS service requirements will be seen to differ from those of the conversational services described thus far in this thesis, the main difference being a relaxation in the end-to-end delay constraints. This Chapter will therefore examine whether this characteristic, combined with the error resilience properties of compressed streaming media services may be exploited to create a dedicated streaming services access radio bearer for EGPRS which can provide for a greater throughput at given interference levels than is allowed by current EGPRS specifications.

7.2 Characteristics of streaming media services

The work described in Chapters 4 through 6 was largely concerned with the performance of conversational multimedia services. As discussed, these are characterised by stringent delay constraints, which generally precludes the use of backward error correction, be it on an end-to-end basis, or at a link level. This is however only one of three main categories of multimedia delivery, the other two being streaming services and multimedia messaging. The latter, by definition are delay-insensitive, and consequently operates in similar ways to email on the Internet or messaging in current cellular systems. The requirements of streaming video, on the other hand are more complicated and consequently should be discussed separately.

Streaming media technology is being developed primarily for use on the Internet where its popularity has been rapidly increasing, both in terms of the number of hosts providing such services as well as the proportion of Internet traffic consisting of media streams. These services are mainly provided by technology offered by Microsoft [MICR-01] and Real Networks [REAL-01], who currently dominate the streaming media market. Both solutions are fairly similar and are
based on similar techniques. They both allow for the transmission of either archived "on-demand" footage, as well as live events, similar in concept to that used in the broadcast industry. The latter is also referred to as Web-TV. The general architecture of such streaming media servers is shown in Figure 7.1.

As mentioned there are a number of differences between conversational and streaming services. The most fundamental one is that unlike real-time interactive services such as video-conferencing, streaming media is capable of employing retransmission techniques to recover lost information. This is made possible by the relaxation of the end-to-end delay requirements, which are typically in the order of 5-10 seconds for streaming media. This means that should packets containing audio or video information be lost, then the client (i.e. the media player or decoder) can ask for a retransmission of the corrupted or lost packets. Should these replacement packets arrive before the corrupted or lost frames were due to be played, then the retransmitted packets are used and no perceptible distortion takes place. The leeway or time duration that can be tolerated between packet loss or corruption and reception of a replacement packet depends upon the size of the buffer at the receiver. This transmission strategy means that a number of transport protocols may be used. These are highlighted below.

TCP provides a reliable connection-oriented transport service for data applications over IP. It makes use of retransmissions to ensure data integrity. In order to reduce and try to prevent network congestion, the TCP employs a congestion control algorithm which uses a 'slow-start'. Congestion is assumed to have occurred when a timeout occurs. This leads to two main disadvantages when trying to transmit real-time services mobile links. The start-up period and congestion control algorithm does not allow for optimal use of what is a scarce resource. In addition in mobile networks, lost packets are at least just as likely to be a result of corruption caused by the wireless link as it is to be caused by network corruption. Several studies show that TCP consequently under-performs in wireless links. [XYL-01]
UDP provides the most efficient network throughput as the server does not rely on feedback from the receiving client. In addition, as it does not employ retransmissions it also allows for the lowest latency. This can have a very positive impact on the user (player) experience. UDP can be combined with an application-layer retransmission mechanism which forwards packets which have been indicated as having been lost by the client application. In this way, the only end-to-end retransmission schemes occur at the application level, rather than at the transport level. This does not preclude use of retransmissions at a link layer, such as RLC in GPRS. The major downsides are that many firewalls do not allow UDP traffic, thereby cutting out a potential audience.

HTTP + TCP -- The Windows Media server can also support HTTP-based control commands along with TCP-based data delivery. This combination has the benefit of working with all firewalls that let Web traffic through (port 80) and provides much more control than a standard Web server, but also adds some overhead to the raw TCP stream that decreases scalability.

Multicast – Many media server platforms can also support IP Multicast protocols to allow very efficient delivery of streaming content to large numbers of users. Multicast enables hundreds or thousands of users to play a single stream, but will only work on networks with Multicast-enabled routers. Multicast is becoming prevalent on corporate networks, but is still very rare on the Internet.

This streaming video paradigm has a number of further implications for operation over mobile networks such as GPRS/EGPRS as compared to operation over wired networks. One of the most serious is the presence of bit errors in addition to lost packets due to congestion. If using TCP, this implies that there will be a high throughput usage due to transport-layer retransmission. Similarly, use of application-layer retransmissions over UDP risks making sub-optimal use of the radio resources. TCP-based services will be able to be supported transparently over mobile networks with no apparent quality deficit as compared to fixed-line access. When using UDP, performance will depend upon the use of the UDP checksum, and possibly on the error-resilience properties of the video codecs used. Error resilience is currently only an issue in streaming video over the Internet in as far as lost packets are concerned rather than corrupted packets. Depending upon the mode of operation, the latter may be the more dominant source of errors, because as demonstrated in work carried out in Chapter 4, error detection and recovery of the contents of a corrupted packet results in better received quality than would be possible if the packet is discarded. This Chapter will examine these issues and propose enhanced schemes for the support of streaming services over EGPRS.
7.3 Requirements of streaming media service class

In Chapter 2 it was seen that GPRS specifies service classes in terms of Reliability and Delay classes [3GPP-22.060]. These classes were designed primarily for the specification of service levels for non-time-critical applications. These include a minimum average delay limit of 2 seconds for packets of size 1024 octets. Similarly, the highest allowable packet loss or corruption rate is in the order of $10^{-2}$. These service classes are clearly unsuitable for multimedia applications, as not only do the values specified not reflect the requirements of multimedia codecs as outlined in this thesis, but the actual parameters used are unsuitable for characterising real-time and streaming applications.

<table>
<thead>
<tr>
<th>Traffic class</th>
<th>Description</th>
<th>Streaming class</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maximum bitrate (kbit/s)</td>
<td>Upper limit on data transfer rate across UTRAN/GERAN</td>
<td>$&lt; 2048 (1) (2)$</td>
</tr>
<tr>
<td>Delivery order</td>
<td>Yes/No</td>
<td></td>
</tr>
<tr>
<td>Maximum SDU size (octets)</td>
<td>&lt;=1500 or 1502 (4)</td>
<td></td>
</tr>
<tr>
<td>SDU format information</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Delivery of erroneous SDUs</td>
<td>Specifies whether packets corrupted by errors are forwarded</td>
<td>Yes/No/- (6)</td>
</tr>
<tr>
<td>Residual BER</td>
<td>Rate of undetected bit errors</td>
<td>$5 \times 10^{-2}, 10^{-2}, 5 \times 10^{-3}, 10^{-3}, 10^{-4}, 10^{-5}, 10^{-6}$</td>
</tr>
<tr>
<td>SDU error ratio</td>
<td>Fraction of packets containing detected errors or lost</td>
<td>$10^{-1}, 10^{-2}, 7 \times 10^{-3}, 10^{-3}, 10^{-4}, 10^{-5}$</td>
</tr>
<tr>
<td>Transfer delay (ms)</td>
<td>Maximum delivery delay for 95% of all packets across the radio interface</td>
<td>250 – maximum value</td>
</tr>
<tr>
<td>Guaranteed bit rate (kbit/s)</td>
<td></td>
<td>$&lt; 2048 (1) (2)$</td>
</tr>
<tr>
<td>Traffic handling priority</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Allocation/Retention priority</td>
<td>1,2,3 (8)</td>
<td>1,2,3 (8)</td>
</tr>
</tbody>
</table>

Table 7.1 3G Service description for streaming class - from [3GPP-23.107]

The requirements for a streaming service class are laid out in 3GPP TS.107, [3GPP-23.107] where the timing relation between successive packets transferred across the radio interface is regarded as being as important as the residual error rates in ensuring acceptable received quality. In practice, the playout application buffers the output media stream in order to remove any variation in inter-packet arrival time, and consequently the maximum time variation allowable is determined by the time-alignment capabilities of the decoding application. Although streaming applications generally require a constant playout rate of multimedia content, variations in buffering capability of different players, together with differences in error robustness, mean that the actual QOS
parameters required to support ‘acceptable’ quality streaming services may vary greatly, and the actual figures may only be determined experimentally for each streaming media solution.

Although the specifications laid out in Table 7.1 describe the service requirements for UTRAN, as GERAN is required to meet the same service requirements, the given parameters are suitable for assessing the quality and suitability of EGPRS to support streaming services. The critical requirements laid out in TS 23.107 for streaming services is for the timing relation between successive packets transferred across the radio interface to be preserved, as a variation in the playout timing of the contents of incoming packets can produce a significant degradation in the perceived quality of service. As mentioned, most of the other main parameters given in Table 7.1 may take a wide range of values. This, once again, is to accommodate the expected wide range of application capabilities. It stands to reason that a more error-resilient application will require less stringent residual BER and SDU loss rate requirements than a less robust application.

7.4 Description of E/GPRS RLC/MAC performance in downlink

The Radio Link Control [3GPP-04.60] protocol in the EGPRS stack is primarily responsible for providing the transmission of data blocks across the radio interface at the required level of service. The RLC layer at the transmit end accepts Link Layer Control (LLC) PDUs from the LLC layer and segments them into radio blocks of the correct size for transmission using the selected modulation-coding scheme. Similarly, the RLC/MAC control messages are segmented into RLC/MAC control blocks. These data and control blocks are then forwarded on to the Medium Access Control (MAC) layer for transmission over the radio interface.

In order to ensure that the requirements of the Service Classes specified in [3GPP-22.105] are met, Backward Error Correction (BEC) in the form of repeat-request systems are implemented. The nature of the particular scheme used is determined by whether GPRS or EGPRS is implemented, and in the case of the latter, one of two alternative schemes may be used. In either case the data transfer process takes place through a Temporary Block Flow (TBF). This is a physical connection between two Radio Resource Entities (namely a mobile terminal and its corresponding base station) for the unidirectional transfer of LLC PDUs on the packet data physical channels. The TBF is allocated resources on one, or if multislotting is enabled, more PDCHs and may carry one or more LLC PDUs. As its very name implies, the TBF is temporary and is maintained only for the duration of data transfer. In practice, this normally refers to the time taken to forward a single IP packet, or more if queuing takes place at the transmission node.
Streaming Multimedia Services

TBFs are identified by means of a Temporary Flow Identity (TFI) which uniquely represents a unique TBF in a given transmission direction. As all TFI values in a particular instance in time are different, the TFI not only specifies a block flow, but it also refers to a unique source or destination mobile terminal. For this reason, the network refers to terminals by the TFIIs they are currently supporting.

7.4.1 Acknowledged mode of GPRS RLC/MAC operation

As this Chapter is concerned with the performance of retransmission mechanisms in the support of streaming media services, the description of the RLC performance will be limited to the Acknowledged mode of operation. In GPRS, backward error control is implemented in the form of a selective ARQ mechanism within a single TBF. The sending side transmits blocks within a window and the receiving side sends correct transmission acknowledgement (Ack messages) or incorrect transmission acknowledgement (Nack messages) according to whether errors occur or not. With the receipt of each correctly transmitted acknowledgement, the transmission window is brought forward. The contents of all data blocks pending acknowledgement are stored in the sending side until such acknowledgement is received. In the case of erroneous blocks, the incorrectly received blocks are retransmitted.

7.4.2 Acknowledged mode of EGPRS RLC/MAC operation

EGPRS offers a second mode of retransmission, namely a type II hybrid ARQ system in addition to the type 1 selective ARQ as used in GPRS. In EGPRS, all retransmissions are restricted to the scope of a single TFI, in which all radio blocks are uniquely identified by their Block Sequence Number. Where EGPRS differs significantly from GPRS lies in the fact that the retransmitted block may not necessarily employ the same coding scheme as used by the original transmission of the same block, but may rather use another modulation-coding scheme belonging to the same family (see Table 3.1). Alternatively the same scheme with a different bit puncturing pattern may be used.

When Incremental Redundancy (IR) is used, the initial transmissions employ a fairly powerful scheme, say MCS-5 or MCS-6. If errors occur in the transmission of these blocks, the erroneous block is stored in the receiver, and the retransmissions are encoded with a less powerful code of the same family. In this example, codes MCS-7 and MCS-9 may be used respectively as like codes MCS-5 and MCS-6 they too belong to families B and A. As all schemes employ the same mother convolutional code, with different puncturing patterns, the two separately received
Streaming Multimedia Services

RLC/MAC blocks are combined and jointly decoded. As the schemes MCS-7 and MCS-9 both encapsulate two RLC/MAC blocks with different block numbers within a single radio block, a single radio block transmission may correct two erroneously received blocks.

An alternative method of implementing type II hybrid ARQ is to initially transmit the radio blocks with a particular modulation-coding scheme, with a set puncturing pattern, say P1. Should a block be incorrectly decoded, the same scheme is used in the block retransmission, only using a different puncturing pattern, which in this case would be P2. As the puncturing patterns complement each other, in that different bit positions are retained for transmission in each of the three different patterns, when combined, they result in a considerably more powerful convolutional code. This process is repeated until the block is correctly decoded.

7.5 Transfer of RLC/MAC blocks

As described above, the RLC protocol makes use of a transmission window to control the transmission rate of data blocks when operating in the acknowledged mode. The state of the send and receive windows governing such transmission are defined by a number of variables which are the topic of interest in this section. The window size in GPRS is fixed at 64, whereas in EGPRS, this may vary from 64 to 1024 and is dependent upon the number of timeslots allocated and the direction of transmission. In order to achieve the highest throughput possible, it is generally recommended that the largest window size allowable should be used. In addition, a constant referred to as the Sequence Number Space (SNS) whose value is 128 for GPRS and 2048 for EGPRS represents the maximum possible number of blocks for which it is possible to retain state information.

Each send point maintains two key variables linked with the transmission of data blocks, namely the send state variable V(S), and the acknowledge state variable V(A). V(S) represents the sequence number of the RLC data block due to be transmitted and may take any value from zero to SNS – 1, whereas V(A) contains the sequence number of the oldest block still waiting acknowledgement. The relative values of these two variables control the operation of the sliding window. An acknowledge state array V(B) is used to store state information of transmitted blocks. On transmission, each block is assigned a PENDING_ACK state, which is then changed to ACKED or NACKED according to the value of the acknowledgement returned from the receiver. If V(S)-V(A) is equal to the send window size, or in other words, the number of blocks pending acknowledgement is equal to the window size, then the transmit window is halted and no new
data blocks are sent. If there are no new data blocks waiting to be sent or if the transmit window is stalled, the data blocks pending waiting acknowledgement will be preemptively retransmitted, starting with the oldest block first. This scheme is optional for EGPRS terminals transmitting in the uplink direction.

7.6 Downlink E/GPRS simulation model

An EGPRS/GPRS radio link transmission model was designed with OPNET to simulate the performance of the retransmission schemes with respect to the provision of streaming media services to mobile terminals. The model describes a single cell operation in which there is a single base station and up to 200 mobile terminals at any one time, which may be either accessing non-real-time data services or streaming media services. Only one carrier is supported by the Base Station, giving eight timeslots for multiplexing the downlink data streams. Each mobile terminal can support any multislotting capability up to a maximum of eight slots. All nine modulation-coding schemes and block formats are supported. The bit error patterns obtained at the physical link layer are obtained from the physical link layer model described in Chapter 3. This link model was integrated within a streaming media simulation model. This consisted of a single server capable of generating media traffic (either video or audio traffic patterns) as well as web traffic. The generated packets are forwarded onto the nodes constituting the Serving GPRS Support Node (SGSN) and the Base Station Subsystem (BSS), where the protocols responsible for the downlink transmission of IP packets across the radio interface are situated. Incremental redundancy is not implemented.

Figure 7.2 Streaming media over EGPRS simulation model
7.6.1 Source Traffic Node

In a GPRS downlink transmission scenario, all packets to terminals in a single cell are routed to the said cell by the Serving GPRS Support Node. For this reason, the protocol entities within the BSS see all the traffic as originating from this single node, and consequently can be regarded as being the source of all traffic intended for that particular cell. In the simulation model developed, this node is referred to as the Source Traffic Node and can support two main categories of traffic, namely non-real-time data, in which case a WWW traffic model using a Pareto distribution is used, or streaming media. When a mobile terminal wishes to initiate a data session, the appropriate signaling information is sent to Source Traffic Node. This then spawns a dynamic child process which will provide the required traffic flow addressed to the correct mobile terminal.

7.6.2 RLC/MAC Operation

As the performance of downlink streaming media services in terms of capacity requirements and resulting user quality depends mainly upon the operation and configuration of the RLC/MAC protocols, the method of operation of these protocols within the developed simulation model will be described in detail. On arrival of an LLC PDU at the RLC node, a request is sent to the destination mobile terminal in order to determine its multislotting capability. This configuration allows for each terminal in the cell to be individually assigned different operational configurations. When the requested information is received, the LLC PDUs are inserted into the input queue of the RLC processor node. This is referred to as QUEUE_RLC_IN in Figure 7.3.

When the RLC node detects the arrival of a new LLC_PDU, it checks the address of the destination terminal and checks if a Temporary Block Flow is open for that terminal. If this is not the case, a new RLC entity responsible for the transmission of blocks between the BSS and the mobile terminal will be created in the form of a child process. The TFI of the newly created TBF will be added to a list of currently maintained TBFs, and the LLC PDU will be inserted into the input queue (QUEUE_TBF_x) of the transmitting RLC entity. The RLC node may support up to 200 distinct child processes. If however, a TBF for the destination terminal is already in existence, then the incoming PDU will be added to the queue of the relevant child process.
Streaming Multimedia Services

The RLC entities thus created implement the downlink RLC protocol described in Section 7.3.2. Each mobile terminal may be configured according to a number of options, including whether Incremental Redundancy is to be enabled and on the choice of modulation-coding schemes. Each LLC PDU is inserted in turn into a segmentation buffer, where it is segmented into radio blocks of the appropriate size and then forwarded to the MAC node according to the operation of the RLC retransmission mechanisms. If all transmitted blocks have been correctly acknowledged by the peer process in the mobile terminal, and there are no remaining PDUs for transmission across the radio interface, the process signals its parent process for its destruction. In this way, the number of open TBFs is restricted to the number of block flows in which there actually exist packets to be transmitted. This reduces the demand for resources across the radio interface. Once a radio block is forwarded by the RLC entity, the MAC node inserts it into one of the eight subqueues present in the node. Each of these queues contains blocks to be transmitted in a particular timeslot, and the downlink allocation scheme employed in this implementation is one which ensures that the blocks are inserted in the shortest queue available, as long as the multislotting limit of the destination terminal is not exceeded.

Figure 7.3 Implementation of EGPRS RLC/MAC Operation
7.6.3 Selective Retransmission

The retransmission schemes described above were designed to satisfy the GPRS service classes. When operating in acknowledged mode of transmission, these classes specify that data integrity must be achieved and therefore specify limits on the end-to-end latency of the system rather than on residual bit error rates. This is because in general data communication scenarios, it is assumed that all information bits are equally important, and consequently all must be received and interpreted correctly. Although this full recovery of data inherently provides the best quality of service in terms of the residual error rates, the unlimited upper bound on the number of retransmissions required to achieve such data integrity means that considerable radio channel resources are required. It will be shown that such schemes are inefficient for the support of streaming services.

Work described in Chapters 4 and 5 shows that compressed media, be it video or audio, while being fairly brittle with respect to channel errors, can however tolerate some degree of corruption without overly degrading the received media quality. This means that multimedia information is inherently more robust to channel errors than general-purpose data, due to the presence of specific error detection and recovery techniques in the latter. This feature can be used to significantly increase the capacity of the EGPRS radio access network in supporting streaming media. Whereas the current RLC retransmission schemes allow for unlimited retransmissions in order to meet the service class requirements, the alternative scheme developed, known as Partial or Selective Retransmission restricts the number of retransmissions to that necessary to meet a specific error rate, or until a certain throughput limit is reached. Rather than carrying out retransmissions for all corrupted RLC blocks, with no upper limit on the total transfer delay, the proposed scheme sets an upper limit by which all retransmissions must have taken place. If this threshold is exceeded, the corrupt RLC block is forwarded.

The proposed scheme therefore increases the throughput available to the radio bearer at the expense of the residual bit error rate. This allows for an increase in the range of interference conditions under which a given throughput level may be sustained. An improvement in the spectral efficiency may however only be achieved if the subjective improvement brought about by the increased throughput is greater than any degradation caused by the increase in residual errors.

This solution effectively targets what may be considered as being the hybrid nature of streaming media services. Whereas conversational services are inherently delay-intolerant and thus cannot
employ most forms of backward error correction mechanisms, streaming services can tolerate a significant delay, sometimes up to a number of seconds (typically 5-10 seconds) for buffering received packets. This makes their service requirements more akin to the data services described above. However, unlike these data services and in common with conversational services, a certain limited residual error rate can be tolerated. The limited backward error protection provided by the selective retransmission scheme is therefore tailored for these specific service requirements.

Selective retransmission may be implemented at either the transmitting or receiving end of a transmission link. In either case, a timer is maintained storing amount of time elapsed since the LLC PDU which is found in the radio block payload was initially forwarded from the LLC layer to the RLC/MAC layer. This will be referred to as $T_e$. When implemented at the transmitting end, prior to forwarding a radio block for retransmission, the RLC entity checks whether $T_e$ exceeds a predefined threshold $T_d$. If this is the case, then retransmission will not take place, and the transmission window will be updated accordingly. The transmitter will however still need to inform the receiver that transmission of the radio block in question has been aborted. The alternative scenario, and the one implemented in these experiments involves the receiver monitoring the elapsed time $T_e$ as compared with the set threshold. Once again, if this threshold is exceeded, no retransmission will be requested. In this case, no extra signalling is required, as the receiver simply has to send an ACK message to the transmitter instead of the NACK message it would otherwise have sent.

### 7.6.4 Differential Thresholding

As outlined in Chapter 5 a compressed media bitstream often contains information segments which have different levels of subjective importance and consequently have varying degrees of susceptibility to channel errors. This characteristic of compressed media may be exploited by a further enhancement to the EGPRS downlink selective retransmission introduced in this Chapter. The technique will be referred to as **Differential Thresholding** and essentially exploits the partial recovery mechanisms introduced by the selective retransmission algorithm so as to favour the QoS parameters of those sections of the compressed data stream which are deemed as being subjectively more important. In this scheme, the compressed audio or video information is split into two separate streams in much the same way as described in Chapter 5. When the LLC PDUs containing the media information is received at the transmitting RLC entity, the priority tag is read and each radio block is assigned a **HIGH** or **LOW** priority tag according to the priority of the contents of the received blocks. Should the radio blocks arrive at the receiver RLC entity...
intact without any channel errors, the contents of the payload are forwarded to a resegmentation buffer in very much the same way as specified in the current EGPRS standards. If, on the other hand, a detectable channel error does take place, the selective retransmission algorithm is employed with one significant improvement. The threshold $T_e$ is set at a different level for the two priority streams, with the higher priority packets being assigned a larger value of $T_e$. This ensures, that a greater portion of the resources available to the RLC entity are allocated to the high priority packets. It will be shown that in practice a value of $T_e$ equal to zero for the lower priority packets ensures the highest differential in residual bit error rates, and consequently, in the best overall perceptual quality. This setting restricts all channel resources available for retransmission to the high priority packets, effectively allowing for backward error correction only to be carried out on high priority packets. Although this has the unavoidable effect of reducing the error performance of the low-priority streams, it effectively improves the performance of the subjectively more important packets.

![Diagram](image)

**Figure 7.4 Operation of Selective Retransmission with Differential Thresholding**

### 7.7 Performance Evaluation

In order to validate the RLC streaming algorithms described above, two media compression algorithms are used as case studies. One is the MPEG-4 video compression standard as described in the previous chapters. The second, and the one which will be used most extensively in this Chapter is an audio compression algorithm [FARR-01]. Audio streaming lends itself particularly well to transmission over mobile networks because not only is it inherently more robust to channel errors than video transmission, but can easily be represented by constant-bitrate
representations, therefore allowing for easier and more transparent packet encapsulation. In addition, as will be seen, the throughput requirements of compressed audio are less demanding than those of compressed video and can be accommodated into a single EGPRS timeslot.

7.7.1 Backward-Adaptive Scaleable Audio Codec

The audio coding scheme used is one developed at the University of Surrey for use in general-purpose audio storage and communications systems. The algorithm at the core of the compression technology used is scaleable not only across a wide range of bit rates, but also across a set of other parameters, such as signal bandwidth and implementation complexity. Other distinguishing features include low delay and inherent robustness of the bit stream syntax to channel errors. The coder operates in two modes. The first mode, which results in an output rate of 16 kbit/s is suitable for encoding wideband signals with a bandwidth of 8kHz with a rate of 1 bit/sample. This is achieved by splitting the signal into two independent sub-bands of 0-4kHz and 4-8kHz which are encoded using psychoacoustically-weighted encoding algorithms which exploit the short-term correlations between samples. A gradual increase in the output quality can be obtained by embedding additional layers in the bit stream at the expense of a higher bit rate. Speech and audio signals of higher bandwidths (> 8kHz) are efficiently represented by the second operative mode which splits the 16kHz signal into four independent sub-bands which can then be represented at a variety of bit-rates ranging from 16kb/s to 64kb/s. When operating at higher rates (>16kbit/s), the enhancement layers, representing higher-quality signal enhancements, may be stripped off to leave a lower-quality, yet syntactically-intact base layer. The resulting obtained audio quality is comparable to that produced by MPEG-1 layer 3 (MP3).

![Perceptual Quality of Audio Codec](image)

**Figure 7.5 Quality Performance of Audio Codec**
7.7.2 EGPRS Capacity for Streaming Audio Services

The three lowest throughput options, namely the 16 kbit/s, 24 kbit/s and 32 kbit/s options will be used in these experiments because they can be transmitted within a single EGPRS slot when using schemes MCS-5, MCS-6, and MCS-7 respectively. The throughput capacity of these schemes, in terms of the radio block payload, are 22,400 kbit/s, 29,600 kbit/s and 44,800 kbit/s respectively. The actual throughput required at the radio block layer must however include all packet headers, and is dependent on the number of audio frames encapsulated within a single packet. The RTP/UDP/IP packet transfer system described in Chapter 4 is used, which contributes an overhead of 320 bits per transmitted packet. As the audio codec makes use of very small frame sizes (a new frame is generated every 30ms), a high degree of protocol efficiency may be achieved by including a large number of frames within a single packet.

<table>
<thead>
<tr>
<th>Throughput</th>
<th>Frame size (bits)</th>
<th>Frames/packet</th>
<th>Packet size (bits)</th>
</tr>
</thead>
<tbody>
<tr>
<td>16 kbit/s</td>
<td>48</td>
<td>80</td>
<td>3840</td>
</tr>
<tr>
<td>24 kbit/s</td>
<td>72</td>
<td>80</td>
<td>5760</td>
</tr>
<tr>
<td>32 kbit/s</td>
<td>96</td>
<td>80</td>
<td>7680</td>
</tr>
</tbody>
</table>

Table 7.2 Packetisation of Audio Information

When using these encapsulation schemes, and assuming that 320 bits are required for the RTP/UDP/IP headers, 8 bits for the SNDC headers and a further 48 bits for the LLC header and frame check sequence, the resulting throughput requirements become 17,567 kbit/s, 2,5567 kbit/s and 33,567 kbit/s respectively.

The capacity of EGPRS in error-prone environments was tested by carrying out simulations using error patterns describing the Typical Urban propagation model at 1800 MHz and a mobile velocity of 1.5 km/hr with ideal frequency hopping enabled. Five mobile terminals were simulated for a duration of 100 seconds, each receiving data representative of a fixed rate audio source at 64 kbit/s. The source packet sizes were 5375 bits, and the inter-packet delivery time was 0.084 sec. The resulting throughput of error-free blocks, also referred to as the Goodput, across the radio interface is shown in Figure 7.6. These results clearly show the link-adaptation capabilities of EGPRS, with the progressively less powerful modulation-coding schemes offering increased goodput as the interference conditions improve. This is not strictly the case with the GMSK-based schemes, as the MCS-1 schemes gives higher throughput rates, than the other GMSK schemes, until a C/I value of around 9dB, whence MCS-5 gives better performance. The 8-PSK schemes simulated, namely MCS-5 through MCS-7, do however follow a clear progression in throughput capability, with MCS-5 giving the best performance up to a C/I value of 12dB, with MCS-6...
providing the optimum performance up to 18.5dB and MCS-7 beyond that. It must be noted that these results do not make use of incremental redundancy and hold only for the specific TU1.5 1800IFH propagation model. As the raw bit error rates for the different coding schemes have been shown to depend significantly upon the nature of the propagation model (see Chapter 3), the specific switching points and throughput figures will vary according to the model used.

Comparing the throughput requirements of the audio codec as laid out in Table 7.2 to the results shown in Figure 7.6, the lower bounds of operation of the audio streaming setup in interference-limited environments can be obtained. The MCS-6 scheme can provide sufficient throughput for supporting the 16 kbit/s rate down to a C/I value of 14.8dB. MCS-7 can however be used to transport 24 kbit/s audio at interference levels down to 19.5dB, and at 32 kbit/s at C/I values in excess of 23.2dB.

### 7.7.3 Subjective quality of received audio

In assessing the quality of the received audio, the signal to noise ratio between the original, i.e. uncompressed, and received audio sequences is used. All experiments are carried out on four different sequences called Police, Catatonia, Susan, and Organ, which represent a mix of rock music, vocals and instrument music. The error-free SNR figures for the sequences are given in Table 7.3.
These results clearly show the improvement in received audio quality that can be obtained by transmitting the audio sequences at higher bitrates. The quality of the received audio can be expressed as a function of interference levels by mapping the SNR results onto the EGPRS capacity results. This is shown in Figure 7.7. The flat traces at 16, 24 and 32 kbit/s indicate that no packet loss occurs. This is true if no delay is assumed to occur in the core network, and if a playout buffer of 5 seconds is employed at the receiver. On average, when operating at the most demanding conditions, namely at 16kbit/s at a C/I of 15dB, the end-to-end delay is 0.45s, with a peak of 3.09s over a transmission period of 400s.

![Figure 7.7 Perceptual performance of single-slot audio streaming over EGPRS](image)

**Table 7.3 SNR values for test audio sequences**

<table>
<thead>
<tr>
<th>Sequence</th>
<th>16 kbit/s</th>
<th>24 kbit/s</th>
<th>32 kbit/s</th>
</tr>
</thead>
<tbody>
<tr>
<td>Police</td>
<td>14.2dB</td>
<td>17.8dB</td>
<td>19.8dB</td>
</tr>
<tr>
<td>Catatonia</td>
<td>10.9dB</td>
<td>14.1dB</td>
<td>16.2dB</td>
</tr>
<tr>
<td>Susan</td>
<td>14.9dB</td>
<td>18.3dB</td>
<td>20.3dB</td>
</tr>
<tr>
<td>Organ</td>
<td>18.4dB</td>
<td>23.5dB</td>
<td>26.3dB</td>
</tr>
<tr>
<td>AVERAGE</td>
<td>14.6dB</td>
<td>18.4dB</td>
<td>20.7dB</td>
</tr>
</tbody>
</table>

### 7.7.4 Performance of Selective Retransmission

The Selective Retransmission algorithm described in Section 7.5.4 aims to increase the throughput of the bearer channels at the expense of the bit error rates on the same channels. The algorithm was implemented using schemes MCS-5, MCS-6 and MCS-7 at the source audio traffic rates of 16, 24 and 32 kbit/s respectively. Figure 7.8 shows the residual error rate of the audio stream forwarded to the receiver using the selective thresholding mechanism. It can be seen that
an increase in interference results in a corresponding increase in the bit error rate. Non-zero values for the residual bit-error rate are obtained whenever the required throughput to ensure complete protection using backward-error-correction mechanisms exceeds the available throughput. The traces in Figure 7.8 show the upper limits of the BER traces, as the error rates reduce to zero beyond the end of each trace. The selective retransmission scheme therefore increases the throughput capacity of the channel at the expense of introducing errors.

![Figure 7.8 Residual bit error rate for selective retransmission](image)

A corresponding increase in the average block transmission delay, as measured from the time of arrival of the corresponding LLC PDU at the transmitting RLC entity to the time of arrival of the radio block at the receiver, is also observed. These results show that a delay budget of at least 1.5 seconds is required in order to accommodate the incoming radio blocks with incurring any additional packet loss. In practice, a figure of 4 seconds would be typical, as the results shown in Figure 7.9 show the average delay values.
7.7.5 Performance of Differential Thresholding

As described above, the selective retransmission algorithm can be enhanced by prioritising the streams, and using different retransmission thresholds. Experiments were carried out at MCS-5, MCS-6 and MCS-7 using threshold levels of 3s and 0s for the high and low priority streams respectively. In practice this means that backward error protection is restricted to the high priority streams, with no ARQ being applied to the lower priority information. The results shown in Figures 7.10-12 clearly show that a considerable differentiation in residual bit error rates can be obtained between the high priority and lower priority streams. The experiments were carried out onto segmented audio streams, where the bits were divided according to which frequency band they represent, with the lower 0-4 kHz band representing the high priority information, and the higher band being classified as being subjectively less important. The bit allocations per packet are given in Table 7.4.

<table>
<thead>
<tr>
<th>Audio codec rate</th>
<th>Lower Band (0-4kHz)</th>
<th>Upper Band (4-8kHz)</th>
<th>Inter-packet spacing</th>
</tr>
</thead>
<tbody>
<tr>
<td>16 kbit/s</td>
<td>2560</td>
<td>1280</td>
<td>0.24s</td>
</tr>
<tr>
<td>24 kbit/s</td>
<td>3840</td>
<td>1920</td>
<td>0.24s</td>
</tr>
<tr>
<td>32 kbit/s</td>
<td>5120</td>
<td>2560</td>
<td>0.24s</td>
</tr>
</tbody>
</table>

Table 7.4 Packetisation scheme for differential thresholding of compressed audio

As the differential thresholding scheme requires the information belonging to the two different schemes to be encapsulated into different radio blocks, it is also necessary to transmit the
information describing the two frequency bands in different LLC PDUs, and consequently in different IP packets. This however results in a considerable increase in header overhead, particularly when the RTP/UDP/IP headers are duplicated, and consequently reduces the effective available throughput, thereby reducing the efficacy of the algorithm. In the results shown, joint RTP/UDP/IP compression as introduced in [CASN-00] is used, in which the required joint RTP/UDP/IP header requirements are reduced to 32 bits.

In all three cases, the reduction in bit error rate of the high-priority streams is obtained at the expense of the lower-priority stream. For example, when using the MCS-5 scheme, a reduction in the residual error rate of one order of magnitude at an interference level of 9dB, and of a factor of 512.5 at 12dB is obtained at the expense of a doubling of the low-priority bit error rate. However at C/I values in excess of 12dB, no errors are obtained for high-priority information. In all three cases examined, the difference between the residual error rates of the two streams increases as the channel quality improves. This is brought about by an increase in throughput that may be used for backward error correction. In fact, very little gain is to be obtained by employing differential thersholding at the given rates below 15dB for MCS-6 and MCS-7 and below 9dB for MCS-5.

![Figure 7.10 Performance of Differential Thresholding for MCS-5 at 16kbit/s](image)
Figure 7.11 Performance of Differential Thresholding for MCS-6 at 24kbit/s

Figure 7.12 Performance of Differential Thresholding for MCS-7 at 32 kbit/s

7.7.6 Perceptual Performance of Differential Thresholding

The effectiveness of the scheme introduced above may only be truly assessed by examining the output of the audio decoder. Figure 7.13 shows the obtained outputs at 16, 24 and 32 kbit/s streaming audio using the selective thresholding algorithm both with, and without differential thresholding enabled.
The results in Figure 7.13 clearly show that an improvement in audio quality can be achieved using differential thresholding. At low error rates, the difference is insignificant, but in the useful areas of the performance envelope, an improvement of between 1dB and 3dB is obtained. In particular, at interference levels between 12dB and 15dB, an improvement of between 2.4 and 1.3 is obtained, while at 18dB, an improvement of 2.9dB is recorded. These improvements are all perceptually noticeable, and they significantly improve the performance envelope of the streaming audio system. This may be compared to the results obtained in Figure 7.7 describing the audio performance obtained by employing the standard EGPRS ARQ schemes. This comparison is shown in Figure 7.14, where the performance of streaming audio without any backward error correction is also shown.

It can clearly be seen that retransmissions across the radio interface bring about a significant performance advantage. In fact when using standard EGPRS ARQ, the same audio quality level can be achieved at a carrier-to-interference ratio between 1dB and 6dB below the required level when no retransmissions are employed. The lowest C/I value which can be used to support audio services using standard EGPRS ARQ is 14.9dB, where a SNR value of 14.6dB is obtained. Without using ARQ, the minimum interference level required to support such audio quality is increased to 21dB. The performance advantage of the EGPRS ARQ schemes can be further increased by employing the selective retransmission with differential thresholding. This improves...
the minimum C/I ratio necessary to support audio services by about 3dB, with only a 1dB reduction in SNR. Although noticeable, this degradation may be considered as being acceptable. There is a 2.2dB gain at the upper end of the envelope, where the maximum quality that can be supported by the 32kbit/s codec is reached at approximately 21dB rather than at 23.2 dB using the standard EGPRS protocols.

![Figure 7.14 Comparison between audio performance of standard EGPRS and EGPRS with Diffquant](image)

It must be noted that the perceptual quality gain at the audio decoder when using the selective retransmission schemes are obtained by allowing for bit errors to be forwarded on by the receiving RLC entity on to the audio decoder. Although the decoder used has properties which make it inherently robust to errors, such as its fixed frame length, no specific error detection or recovery techniques are implemented. If an error detection algorithm is included, together with suitable concealment techniques, it may reasonably expected that a further improvement in the decoder performance envelope may be achieved.

### 7.8 Video Streaming

As already discussed, the characteristics of compressed video and compressed audio in terms of error performance are rather similar, in that both media types exhibit a degree of robustness to channel errors. This means that the selective retransmission techniques described, should in principle be applicable to the transmission of streaming video sequences. Unlike audio
compression algorithms, the output of most video coding techniques consists of variable-sized frames, the actual size of each individual frame being determined by the level of temporal activity and spatial detail within that section of the video sequence, as well as of the level of quality of the encoded frame. This means that most video codecs can operate at a wide range of output bitrates. In order to allow for an easy comparison between the operation of audio and video streaming, the video output rates are set to approximately twice the rates used in the streaming audio experiments. Two slot operation is required for the support of the streaming video services using schemes MCS-5, MCS-6 and MCS-7. The source data rates used, and the resulting quality levels in terms of the Peak Signal-to-Noise Ratios (psnr) are shown.

<table>
<thead>
<tr>
<th>Data rate</th>
<th>Sequence Quality</th>
</tr>
</thead>
<tbody>
<tr>
<td>33.2 kbit/s - MCS-5</td>
<td>29.33 dB</td>
</tr>
<tr>
<td>49.2 kbit/s - MCS-6</td>
<td>31.57 dB</td>
</tr>
<tr>
<td>64.6 kbit/s - MCS-7</td>
<td>32.797 dB</td>
</tr>
</tbody>
</table>

Table 7.5 Video quality levels

The selective retransmission techniques using differential thresholding was tested for the transmission of video sequences using the separate streams partitioning technique described introduced in Chapter 5. This separates an MPEG-4 compressed stream into two discrete streams, one containing mainly the motion information, and the second stream containing the more robust texture information. When retransmissions are enabled, the minimum C/I ratio that is required to sustain a video throughput of 32kbit/s over two EGPRS slots is 16.4dB. Selective retransmissions with differential thresholding reduces this value to approximately 14dB, with little or no degradation in received quality. Similarly, the maximum quality level that can be achieved at 64kbit/s is only obtained at a C/I level of 23dB when using standard EGPRS retransmissions, as compared with 21.5dB for the differential thresholding. This improvement is however not observed at all positions of the performance envelope. At 16.5dB, the standard EGPRS retransmission schemes outperform the proposed techniques. This is caused by the fact that splitting the video information into two discrete streams, doubles the overhead required to support the transmission of the video streams.
When comparing the performance envelope of the two streams, it can be seen that the lowest C/I values which can support the streaming media using EGPRS retransmissions at the given slot allocation schemes are 16.4 and 14.6 dB for video and audio respectively. These values are reduced to approximately 14dB and 12dB when selective retransmission with differential thresholding is applied. This shows that although, as expected, audio displays more robustness to errors than compressed video, and can be transmitted at acceptable quality over poorer channel conditions, a reduction of approximately 2.5dB in the required signal-to-noise ratios required to support both services is observed.

7.9 Conclusion

This Chapter has introduced a novel backward error correction scheme for streaming media which exploits the characteristics of compressed multimedia information. This scheme was tested for audio and video transmission, and was seen to extend the performance envelope by approximately 2dB and provide a more graceful degradation between operation at different coding schemes. The performance of streaming audio services were seen to exhibit a more graceful degradation, due to the increased inherent error-resilience of audio compression schemes. In fact, the effectiveness of this scheme depends upon the error resilience of the media compression algorithm used to channel bit errors.
Although the use of fixed-length frames and short propagation of errors by prediction algorithms reduces the susceptibility of the audio codec to channel errors, the decoder used did not employ any specific error detection or concealment algorithms. Use of a more error-robust decoding algorithm should therefore result in significant values of carrier power required to support streaming audio services. In addition, the simulation model used in the experiments did not implement incremental redundancy (IR), which is required to be employed by all EGPRS handsets in downlink [3GPP-04.60]. Incremental redundancy is expected to lend itself well to the partial retransmission scheme described, particularly as it provides significant enhancements to the efficacy of the backward error correction schemes, and will consequently provide enhanced quality differentiation when employing differential thresholding. Moreover, further work should examine the relationship between the selective retransmission algorithm with transport layer protocols and on enhancing the integration between the RLC/MAC link adaptation and media rate adaptation.
Chapter 8

8 Conclusion

8.1 Overview

This thesis has examined and investigated several issues concerned with the transmission of multimedia services over mobile access networks. The main finding of the work carried out is that although decoupling of the mobile bearer channels from the services that will be offered over the radio access network will undoubtedly lead to a much wider range of services than is currently available, optimum radio efficiency may only be obtained by closely tailoring the radio access protocols to the requirements of the application layer. This optimisation must not be restricted to the just the physical link layer but must be carried out at all levels of the protocol stack across the radio interface. Conversely, the applications must also be designed around the requirements of the access network.

It was shown that the standard EGPRS and GPRS radio channels are not suitable for supporting real-time multimedia services, be they voice, speech or audio. The lower-bound of carrier-to-interference levels that can support video services was found to be approximately 13dB, which is approximately 6 decibels higher than the lower bound for speech services. Even at this rate, the average throughput per timeslot is less than 10 kbit/s. Such poor performance is wholly attributable to the fact that the GPRS and EGPRS bearers were designed for supporting generic data services such as email, web browsing, file transfer etc. In such application environments, throughputs in excess of 16kbit/s may easily be sustained, supporting the argument that the E/GPRS protocols and system were not poorly designed, but merely optimised for use in conjunction with-link level retransmission schemes, rather than as a bearer for real-time or delay-critical services. This thesis has shown that by modifying the characteristics of the data bearers to match the application requirements, significant improvements in system performance may be achieved. This was demonstrated by the design of optimised bearers for real-time video communications and for streaming media applications.
The second main area of investigation was focused upon improving video representation technology for transmission across mobile networks. It was shown that the differing levels of importance of different parts of a compressed video syntax can be exploited to make most efficient use of the available radio resources. High priority information may be given preferential treatment by either simply apportioning a greater share of the available throughput, or else by assigning greater channel protection. The latter may be in the form of more powerful error detection and correction mechanisms, or in more effective backward-error correction. Two main forms of information prioritisation were found to be effective in improving the quality of received media streams. Different parameters included within the syntax of compressed media have different levels of subjective importance or susceptibility to channel errors, and can therefore form the basis of unequal protection mechanisms. Similarly, not all areas of a scene represented by a video sequence may be perceived as being equally important. If region classification algorithms are used, this property may be exploited to enhance the quality of received video sequences.

The remainder of this Chapter will review the main achievements described and outline the author's view of the future of mobile multimedia communications.

8.2 Novelty

The main achievements of the work described in this thesis are summarised below.

1 Comprehensive mapping of video services onto GPRS/EGPRS data bearers

A thorough study of the transmission parameters required to transmit real-time MPEG-4 video services over GPRS and EGPRS radio access bearers was carried out. The traffic characteristics of source video sequences encoded at different rates were examined as was the received video quality at different output bitrates, allowing for a determination of the multislotting requirements for different levels of service. The performance improvements that can be obtained by employing the standard MPEG-4 error resilience tools and rate control mechanisms was assessed. A physical-link layer simulation model of the GPRS and EGPRS packet data channels was designed together with a GPRS radio interface data flow model. This was used to assess the quality of received video in interference-limited channels under various propagation conditions. A real-time emulator was built to allow for interactive testing.
2 Scheme for prioritised transmission of video over multiple data bearer channels

An unequal protection mechanism which makes use of the data portioning video frame structure used by MPEG-4 was introduced. This scheme allows for the video information contained within a single frame to be split into two separate segments which can be encapsulated into different IP packets and transmitted over different channels. The scheme was designed so as to be able to be used over UMTS bearer channels without any customisation of the radio links. Extra data was added to the MPEG-4 syntax to ensure resynchronisation of the two separate schemes. This technique was shown to provide an improvement of between 1 and 2dB over GPRS channels.

3 Adaptation of object-oriented video representation techniques for scene prioritisation

MPEG-4's video-object representation tools were used to develop object-oriented scene prioritisation mechanisms. Two techniques were proposed. An inter-scene prioritisation scheme for use in multi-party conferencing applications was described for use in H.323 environments. This was shown to provide effective traffic management capabilities as well allowing for throughput resources and to be allocated to the perceptually more important scenes. This technique was also shown to be useful in enhancing the quality of a single received video stream if used in conjunction with an appropriate scene segmentation algorithm, by dividing a videoconferencing scene into areas of high and low perceptual importance.

4 Development of Differential Quantisation techniques for intra-scene prioritisation

An alternative scheme which does not require the use of multiple video objects, and which consequently does not make use of any shape information was developed to enhance the transmitted video quality. This involved assigning different priority levels to different macroblocks within a video frame and allocating differing shares of the available resources to the different macroblocks according to their importance by varying the level of the DCT quantiser. This technique, which was named Differential Quantisation was used with two region classification algorithms designed for this purpose, one for separating the speaker from the background in video communications scenarios, and another for use in the transmission of sports footage. In the latter case 3-level prioritisation was implemented.

5 Design of Enhanced Data Bearers for real-time video communications

Three enhanced data bearer channels with rates of 10.45, 16.42 and 22.00kbit/s were designed for use in EGPRS radio access networks. These made use of a concatenated Convolution – Reed
Solomon coding scheme. Soft-in, soft-out algorithms were implemented for both decoders, which when used with an interleaver of appropriate depth were shown to provide significant improvement over standard EGPRS modulation-coding schemes at the error rates required to support video communications, particularly when implemented with an iterative joint convolutional-RS decoding structure. The efficacy of the proposed schemes in acting as bearers for video transmission was compared to the standard EGPRS coding schemes and were shown to produce a relative coding gain of between 4 and 12 dB in interference-limited channels. The use of packet starter and termination codes was introduced to allow for the scheme to be backwards-compatible with the current EGPRS schemes, while allowing also for application-level decoding.

6 Link-level retransmission mechanisms optimised for streaming media

A novel retransmission technique was developed for use in supporting streaming multimedia applications. Selective or partial retransmission sets an upper limit to the number of retransmissions allowed. This effectively, increases the throughput of the data link, at the expense of introducing a non-zero residual error. This scheme was shown to be adaptable to implementing stream prioritisation by simply setting different timer thresholds for different priority levels. The effectiveness of the proposed schemes for the transmission of audio and video sequences was demonstrated.

8.3 Lessons Learnt

The work described in this thesis clearly shows the benefit of adopting a system-level view when designing multimedia communications systems. The poor quality of decoded MPEG-4 streams that were subjected to errors that occur over the downlink GPRS and EGPRS access channels clearly demonstrated both the error susceptibility and throughput requirements of compressed video streams, as well as the high residual error rates of the E/GPRS channels when operating in the unacknowledged mode of transmission. Less specifically, these results demonstrate the need to design radio access bearers to meet the specific quality of service requirements dictated by the multimedia applications. The prioritisation techniques introduced in Chapter 5 clearly show the benefits of designing multimedia applications that exploit the different characteristics of the bearer channels. Conversely, the forward error correction scheme introduced in Chapter 6, and the partial retransmission mechanisms designed for streaming bearers clearly demonstrate the performance advantage that may be obtained by tailoring the quality of service characteristics to meet the demands of the multimedia applications.
8.4 Future work and the way ahead

This thesis should be treated as a discussion of work-in-progress and a snapshot of what is currently state-of-the-art in the field of mobile multimedia communications. It was never the scope of this work to produce the definitive guide or solution to the technological problems described, but merely to serve as a pointer to the most important issues and to highlight some of the potential solutions. This Section will describe some of the issues that remain be tackled.

The more immediate research objectives are concerned with the work carried out in optimising the bearer channels as described in Chapters 6 and 7. When designing an enhanced bearer for real-time services, research was focused upon the use of concatenated convolutional-RS coding structures. This approach had one main drawback, namely that the length of the block codes used was insufficient to provide any significant improvement in the residual error rates when only hard-decision information is available from the convolutional decoder. The obtained solution is consequently most effective only if implemented as a physical link layer solution. No techniques currently exist to allow for soft information to be conveyed across the protocol stack up to the application layer, where further error detection may be added as part of a multi-level error correction scheme. Two main approaches therefore exist for solving this problem. One is to develop more efficient ways of carrying out soft-decision decoding on longer RS codes. The use of product block codes is one promising technique by which short RS codes are used to construct what are effectively longer code structures. The alternative would be to provide for fully reconfigurable terminals which allow for upgrades to the channel decoding algorithms, possibly including interfaces for soft information in between protocol layers.

Chapter 7 examined the issues involved in designing a new backward-error correction scheme for use in streaming multimedia applications. The main characteristic that distinguished the selective retransmission scheme proposed from other similar mechanisms was that delay was one of the major design criteria used in constructing the scheme. Further research needs to be carried out, particularly in investigating the relationship between the proposed scheme and the link adaptation and incremental redundancy mechanisms as used by EGPRS. Furthermore, the experiments described in this thesis were concerned with employing only link-level retransmission over the radio access network. The interaction between partial retransmission and end-to-end backward error correction needs to be investigated fully so as to be able to ascertain the optimum retransmission strategy. Although delay constraints dictated that backward error correction be used solely for streaming applications, the use of fast, hybrid ARQ mechanisms for delay-critical real-time applications should be examined.
The optimisation of the video coding algorithms for use in mobile environments was discussed in Chapter 5. In the experiments discussed, the proposed solutions were approached individually. Research needs to be carried out into which prioritisation schemes are most appropriate for different operational scenarios. The role that may be played by video object representation techniques needs to be further investigated, particularly in view of the advances being made to the error resilience of shape coding techniques. The classification algorithms used in this thesis, while being effective enough to demonstrate the usefulness of region prioritisation, are nevertheless sub-optimal and are far from representing what is state-of-the-art. The use of computationally-efficient algorithms which can be adapted to use in different scenarios needs to be studied.

The main weakness of most of the research carried out to date on mobile video communications, is that it has been too focused on the properties of the actual video compression algorithms, and on the performance when subjected to various channel constraints. More work needs to be carried out into the integration and synchronisation of various media types, in particular voice and video, when carrying out communications over a number of different radio bearers. Several issues in the development of stream management techniques for multiparty scenarios, and the implications of the deployment of portable multimedia applications remain to be resolved.

In the longer term, two apparently fundamental tenets of mobile multimedia communications need to be reassessed, namely that the information will be digital, and that it will transported in IP packets over IP networks. The latter statement is likely to remain true for the foreseeable future in the core network. New high-speed switching mechanisms are allowing for more efficient and higher-volume IP routing capabilities, and this is leading to a significant increase in the use of Voice over IP technologies. However, the necessity to transmit full IP packets over mobile access interfaces needs to be re-examined. For example, in mobile Voice over IP scenarios, the bulky protocol headers down the GPRS protocol stack when using the Real-Time Transport Protocol are so inefficient, that depending upon the packetisation scheme used, a greater proportion of the available throughput in packet voice applications may be used to convey the various header information, rather than the actual speech. Much research is already underway into developing efficient header compression mechanisms as well as carrying out header stripping and reconstruction over the radio link when supporting packet voice services. It is the author’s opinion that such efforts should be taken further and dispensing with much of the IP/UDP/RTP information over the mobile radio link, as well as rationalising the network-specific link layer protocols. These should be replaced, at least for multimedia and low-latency applications, by a
light-weight stack capable of supporting the two main classes of multimedia services, namely conversational and streaming services at the different QoS levels required to support different media types. Suitable RTP/UDP/IP packets may be reconstructed at edge proxies for transmission over the core network to fixed-line devices or other similarly-configured proxies. It is envisaged that such an approach would result in a common multimedia payload format for use across several underlying access networks, acting as a universal service interface for mobile multimedia services in much the same way as IP acts as a common service interface for all Internet applications. The cell format as used in ATM transmission is inherently more suitable for multimedia services, as it can provide proper quality of service guarantees, while still allowing for flexible service operation, such as the support of variable bitrate operation. Although ATM-based technology has recently fallen out of favour when compared with IP, it can certainly act as a model for a multimedia transmission technology for use in wireless environments. As IP will remain the technology of choice in the core network, appropriate service interfaces, and border gateways will have to be specified and implemented.

The second, more fundamental issue that needs to be addressed is on the merits of digital transmission of multimedia services. Although the ease of representation of data by two-state storage devices as offered by electronic, optical and magnetic technologies mean that it is unlikely that information will be stored in any other way than in a digital format, at least not until the advent of some revolutionary storage medium, the case is not so clear-cut for the transmission of multimedia information over lossy channels. The main advantage of digital transmission is that, when coupled with appropriate error correction mechanisms, it allows for a perfect reconstruction of the original information. Although such authenticity is a cardinal requirement of general data transfer applications which may be used as bearers of critical information such as medical or financial data, it is somewhat less useful in multimedia communications. In these applications, not only do delay constraints place limits on the amount of error correction that may be efficiently carried out, but also perceptual quality of service supplants data integrity as the main criterion for evaluating quality of service. This thesis has demonstrated how error resilience techniques ensure that some useful information is obtained from packets or radio blocks containing bits in errors, and how the effective throughput of radio channels and the resulting received quality may be increased by accepting non-zero residual error rates.

Similarly, recent research into the use of soft-decision decoding of variable-length codes which have been corrupted by channel errors indicates that significant improvements in the quality of decoded media streams may be obtained by combining the soft information received with some
knowledge of the previously-received sections of the media stream. Just as soft information about the received bits in the receiver allows for more efficient channel decoding, it is expected that the soft output from a channel decoder will be employed to enhance the visual, or audio quality of received video streams. Sampled digital representations of soft information may well be tomorrow’s analogue signals, replacing the use of hard-decision bits in the receiver terminals of multimedia communications systems from the physical layer right up to the application layer.
Publications


References


[3GPP-04.64] 3GPP, "Technical Specification 3rd Generation Partnership Project; Technical Specification Group Core Network; Digital cellular telecommunications system (Phase 2+); General Packet Radio Service (GPRS); Mobile Station - Serving GPRS Support Node (MS-SGSN) Logical Link Control (LLC) layer specification, (Release 1999)", 3GPP TS 04.64 V8.6.0, December 2000.

[3GPP-04.65] 3GPP, "Technical Specification 3rd Generation Partnership Project; Technical Specification Group Core Network; Digital cellular telecommunications system (Phase 2+); General Packet Radio Service (GPRS); Mobile Station (MS) - Serving GPRS Support Node (SGSN); Subnetwork Dependent Convergence Protocol (SNDCP), (Release 1999)", 3GPP TS 04.65 V8.1.0, September 2000.


References


References


References


[GSM-02.34] “Digital cellular telecommunications system (Phase 2+); High Speed Circuit Switched Data (HSCSD); Stage 1”, GSM 02.34, V6.0.0, 1999.


References


References


[SMG2-060/00] Tdoc SMG2 060/00, “Change Request GSM 05.05: EGPRS Receiver Performance for MS DCS 1800 and PCS 1900”, Ericsson, Motorola, Nokia, January 2000.


References


