University of Southampton
Faculty of Engineering

Wavelet Video Compression for Iterative Wireless Transceivers

by

Anh Quang Pham
MEng.

A thesis submitted in partial fulfilment of the
requirements for the award of Doctor of Philosophy
at the University of Southampton

SUPERVISOR:

Professor Lajos Hanzo
Dip Ing, MSc, PhD, FIEEE, DSc, FREng
Chair of Telecommunications

Dr Lie-Liang Yang
BEng, MEng, PhD, SIEEE
Reader

This thesis is dedicated to:
My wife and family in Vietnam.

UNIVERSITY OF SOUTHAMPTON

FACULTY OF ENGINEERING AND APPLIED SCIENCE
DEPARTMENT OF ELECTRONICS AND COMPUTER SCIENCE

ABSTRACT

Wavelet Video Compression and Communications

by Anh Quang Pham

Over the last decades, there has been an increasing demand for multimedia services provided over both fixed and wireless channels. Designing robust image or video transmission systems for these multimedia services is hence of high importance.

For the sake of exploring the feasibility of providing video services for mobile users, this thesis investigates the design of wireless video communication systems using the British Broadcasting Corporation's recent proprietary wavelet-based video codec referred to as the Dirac codec. More specifically, the Dirac video codec is capable of achieving an approximately two-fold bit rate reduction over MPEG-2 when compressing high definition video (e.g. 1920x1080 pixels), while achieving a high reconstructed video quality at a low bit-rate. The Dirac video-encoded bitstream is subjected to a rigorous error sensitivity investigation for the sake of assisting us in contriving various joint source-channel coding (JSCC) and decoding schemes for wireless videophones.

Unequal Error Protection (UEP) is an attractive technique of implementing JSCC. Based on the Dirac video codec's bit sensitivity studies, an Unequal Error Protection (UEP) scheme using turbo-equalized Irregular Convolutional Codes (IRCCs) was designed. Our UEP scheme allocates a lower channel coding rate to highly sensitive bits and a higher channel coding rate to less sensitive bits, while keeping the coding-rate of the UEP the same as that of the Equal Error Protection (EEP) benchmarker.

Furthermore, an Iterative Source-Channel Decoding (ISCD) scheme, which exploits the residual redundancy left in the source encoded bitstream for the sake of improving the attainable system performance was investigated. Hence, a novel ISCD scheme employing a specific bit-to-symbol mapping scheme referred to as Over-Complete Mapping (OCM) was proposed. The proposed scheme benefits both from the residual redundancy inherent in the source encoded bitstream as well as from the intentional redundancy imposed by the specific over-complete source mapping. This allows us to design an attractive video transmission scheme having a high error resilience at a reasonable complexity and a low

delay.

The research reported in this thesis was concluded with the design of unequal error protection irregular over-complete mapping for wavelet video telephony using iterative source and channel decoding. The philosophy of this video telephone scheme is that we exploit as much redundancy inherent in the Dirac-encoded bitstream as possible for improving the system's BER performance, while protecting the more sensitive portions of the Dirac video-encoded sequence with a lower OCM rate. This design dilemma was addressed in the context of a novel Irregular Over-Complete Mapping (IrOCM) scheme constituted by the amalgam of appropriately weighted different-rate OCM schemes. Moreover, in order to accommodate the variations of the relative frequency of the K-bit source symbols used for modelling the Probability Density Function (PDF) of the different video-encoded bitstreams, a near-instantaneously adaptive transceiver for wireless video telephony was designed. This adaptive transceiver allows our wireless video telephone scheme to attain iterative detection gains without increasing the system's delay, regardless of the choice of the source codec - even for speech and audio codecs, and hence the novel OCM scheme is widely applicable.

Acknowledgements

I would like to express my deep gratitude to Professor Lajos Hanzo and Dr Lie-Liang Yang, my supervisors, for their constant encouragement, guidance, kindness and patience. This project would not have progressed far without their involvement, especially when things were not going to plan. All their help, advice, ideas and encouragement are greatly appreciated and I am very thankful for their wonderful friendship.

I am also thankful to Dr Bee Leong Yeap for his help through the early months of chaos and confusion. I would also like to take this opportunity to thank all my previous and current colleagues in the Communications Group. They created such a great research environment and were always ready to lend a helping hand. Special thanks to Jin Wang and Robert Maunder for sharing their knowledge in channel coding. I am grateful to Andreas Wolfgang, Soon Xin Ng with whom I enjoyed working in our fruitful collaborations. Warmest thanks also Noor Shamsiah Othman and all other colleagues and staff. The assistance provided by Denise Harvey and Rebecca Earl, the Communications Group secretaries, is gratefully acknowledged. In addition, I offer my heartfelt thanks to all my friends in Southampton and elsewhere, for their constant kind wishes and the good times that we shared together.

I also owe sincere thanks to my family especially my parents, my beloved wife and my children in Vietnam for their patience, love, encouragement and supporting me in all my decisions. Without you all, this thesis would never have been completed.

Southampton, April, 2008

List of Publications

1. **A. Q. Pham, N. S. Othman, L. L. Yang, and L. Hanzo**, "EXIT-Chart Optimized Block Codes for Wireless Video Telephony", Accepted for publication in the IEEE Transactions on Circuits and Systems for Video Technology.

2. **A. Q. Pham, J. Wang, L. L. Yang, and L. Hanzo**, "An Iterative Detection Aided Unequal Error Protection Wavelet Video Scheme Using Irregular Convolutional Codes", In Proceedings of IEEE VTC'06 Spring, pp. 2484-2488, Melbourne, Australia.

3. **A. Q. Pham, J. Wang, L. L. Yang, and L. Hanzo**, "An Iterative Detection Aided Irregular Convolutional Coded Wavelet Videophone Scheme Using Reversible Variable-Length Codes and Map Equalization", in Proceedings of IEEE 66th Vehicular Technology Conference, pp. 2404-2408, Dublin, Ireland, 22-25 April 2007.

4. **A. Q. Pham, L. L. Yang, and L. Hanzo**, "Joint Optimization of Iterative Source and Channel Decoding Using Over-Complete Source-Mapping", in Proceedings of IEEE 66th Vehicular Technology Conference VTC2007-Fall, pp. 1072-1076, Baltimore, USA, 30 September-3 October 2007.

5. **A. Q. Pham, L. L. Yang, and L. Hanzo**, "Unequal Error Protection Irregular Over-Complete Mapping for Wavelet Coded Wireless Video Telephony Using Iterative Source and Channel Decoding", in Proceedings of the IEEE Wireless Communications and Networking Conference (WCNC-2008), pp. 1251-1255, Las Vegas, Nevada, USA, 31 March-3 April 2008.

6. **N. S. Othman, M. El-Hajjar, A. Q. Pham, O. Alamri, S. X. Ng and L. Hanzo**, "Over-Complete Source-Mapping Aided AMR-WB MIMO Transceiver Using Three-Stage Iterative Detection", in Proceedings of the 2008 IEEE International Conference on Communications (ICC 2008), CD-ROM, Beijing, China, 19-23 May 2008.

7. **Nasruminallah, M. El-Hajjar, N. S. Othman, A. Q. Pham and L. Hanzo**, "Over-Complete Mapping Aided, Soft-Bit Assisted Iterative Unequal Error Protection H.264 Joint Source and Channel Decoding", accepted for the 68th IEEE Vehicular Technology Conference, Calgary, Canada, 21 - 24 September, 2008.

Contents

Chapter 1

Introduction

1.1 Research Motivation

Contemporary communication techniques have pervaded our daily life. A fundamental issue in communication theory is how to communicate information between two geographically separated locations with adequate quality. The ultimate aim of source coding is the removal of redundancy from the source information, while channel coding intentionally incorporates controlled redundancy into the source-coded bitstream, in order to mitigate the effects of channel noise. According to Shannon's source and channel coding separation theorem [1], these operations may be carried out separately, following separate optimization of each component. However, this statement holds only under ideal conditions, when transmitting for example lossless entropy-coded infinite-length source-coded data, protected by a potentially infinite-length channel code over Gaussian channels. This system would impose an unlimited delay. Unfortunately, most practical communication systems do not meet these requirements.

Significant progress was made throughout the last decades in optimizing each individual component of communication systems. Accordingly, considerable efforts have been invested also in the design and development of efficient video compression schemes and standards [2–5]. For the sake of achieving high compression, most standardized video codecs, such as for example the MPEG4 [2, 3] and H.264 [6] schemes, aim for removing the predictable video contents inherent both within as well as between consecutive video frames. The inter-frame redundancy is predicted and removed with the aid of motion

compensation. The residual intra-frame redundancy inherent in the Motion-Compensated Error Residual (MCER) is typically transformed to the frequency domain using the Discrete Cosine Transform (DCT) and the resultant DCT coefficients are usually further entropy-coded, in order to produce the compressed bitstream. These algorithms provide a high compression, however, the compressed signal becomes highly vulnerable to transmission errors, which is particularly detrimental when communicating over hostile wireless networks.

Furthermore, as argued in the preface of [7], the design of wireless video systems hinges on numerous trade-offs, including factors such as the achievable compression ratio, the implementation complexity, the encoding delay, the robustness against transmission errors, etc. In this spirit, the British Broadcasting Corporation (BBC) developed an attractive wavelet-based design alternative referred to as Dirac[1], which may be employed without incurring any license-fee [8]. This wavelet video codec is capable of achieving an approximately two-fold bit rate reduction over MPEG-2 [9, 10], when compressing high-definition video (e.g. 1920x1080 pixels), while achieving a high reconstructed video quality at a low bit-rate.

Encouraged by the Dirac codec's and the H.264 codec's attractive video compression performance, we focus our attention on the design of wireless video communication systems using these two codecs as our source codec. With respect to the design of communication systems, in the preface of [7], the authors showed that in the context of limited-complexity, limited-delay, lossy source codecs, which exploit the psycho-visual and psycho-acoustic properties of human perception, Shannon's lessons have limited applicability. This is particularly so, when the real-time, interactive video system experiences bursty errors inflicted by dispersive fading channels, rather than independent random errors. Hence the joint design of source coding and channel coding is expected to yield the best end-to-end system performance [11].

This observation has led to Joint Source-Channel Coding and decoding (JSCC), which motivates the research work presented in this thesis. Unequal Error Protection (UEP) is one of the popular means of implementing JSCC. A typical video-encoded sequence exhibits different bit error sensitivities for different bit positions. If the channel encoder can exploit the knowledge of the source bit error sensitivity, UEP can perform better

[1]The British Broadcasting Corporation's Research and Development Department contrived the Dirac concept in January 2003 and at the time of writing it is still under development.

than Equal Error Protection (EEP), since UEP protects the more sensitive bits of the video-encoded sequence with the aid of a lower channel coding rate. By contrast, the less sensitive bits of the video-encoded sequence are protected using a higher channel coding rate, while keeping the overall channel coding rate fixed. However, in practice, an ideal UEP design, which adapts the channel coding rate in real-time to the source coding is very complex and hence this design issue remains a challenging research topic.

Furthermore, although source codecs try to remove redundancy from the source, there is often substantial residual redundancy left in the source encoded bitstreams when using finite-delay, finite-complexity codecs. This residual redundancy may not be readily exploited by conventional separate source-channel decoders. In this thesis, we will demonstrate how the residual redundancy inherent in the video-encoded sequence can be utilized by our iterative source-channel decoders. Moreover, since there is limited source redundancy left in the source-encoded bitstream when using efficient source encoders, the softbit-source decoder may have a limited extrinsic information contribution, which may result in negligible system performance improvements. In this scenario, we may opt for partitioning the total available bit-rate budget between the source and channel codecs, and may deliberately increase the inherent redundancy in the video-encoded sequence with the aid of a novel technique proposed in this treatise, which we refer to as over-complete mapping. We will demonstrate that the error correcting capability of the resultant Iterative Source-Channel Decoding (ISCD) scheme has been significantly improved.

As a further investigation, unequal error protection aided Irregular Over-Complete Mapping (IrOCM) schemes are designed for wireless video telephony using iterative source and channel decoding. The philosophy of our video telephony scheme is that we aim for exploiting as much redundancy inherent in the Dirac-encoded bitstream as possible for improving the system's BER performance, while protecting the more sensitive portions of the Dirac video-encoded sequence with the aid of a lower mapping rate. More specifically, the IrOCM is optimized to match the characteristics of both the Dirac codec and those of the channel codec, so that the UEP scheme has attractive iterative decoding convergence properties, while maintaining unequal error protection capabilities matched to the requirements of the wavelet encoded video sequence. Moreover, we also design and investigate a near-instantaneously adaptive transceiver for wireless video telephony, which allows our video scheme to attain iterative detection gains without increasing the system's delay, regardless of the choice of source codec - even for speech and audio codecs,

and hence our design principles becomes widely applicable.

1.2 Related Background

The classic communication system of in Figure 1.1 consists of a source encoder and decoder pair, a channel encoder and decoder pair, as well as a modulator and demodulator. The source encoder and decoder pair is commonly referred to as the source codec module, whereas the channel encoding and decoding pair is usually referred to as the channel codec while the modulator and demodulator module as the modem. Over the past decades, the separate design of the two coding modules was justified by Shannon's source and channel coding separation theorem [1]. However, It was demonstrated that their joint design is capable of improving the attainable system performance for practical communication systems [11]. The intention of this section is to provide the required research background on source coding, channel coding, on the comparison between Separate Source and Channel Coding (SSCC) as well as Joint Source-Channel Coding (JSCC), and the fundamental principles of Iterative Source and Channel Decoding (ISCD).

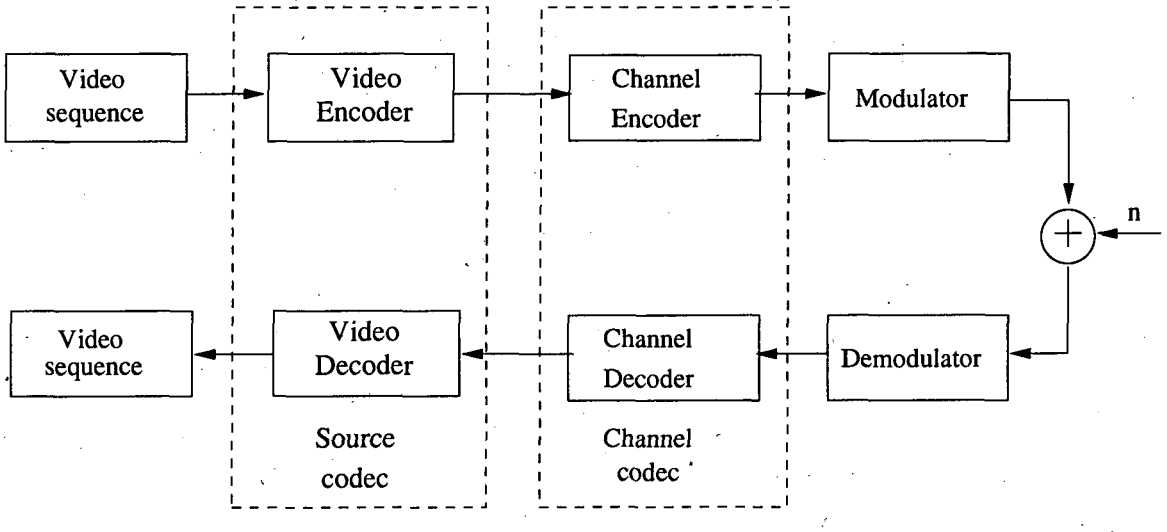


Figure 1.1: Block diagram a classic communication system.

1.2.1 Overview of Video Coding

In a communication system, the ultimate aim of source coding is the removal of redundancy from the source information. This, therefore, reduces the number of bits required for representing the information contained within the source. Achieving the best possible compression ratio requires not only an understanding of the nature of the source signal in its binary representation, but also how we as humans interpret the information that the data represents. In the forthcoming sections we will focus our attention on the most important techniques widely used in the existing international image and video standards, which will be employed in our simulations in the subsequent chapters.

1.2.1.1 Background

Digitized video signals may be viewed as a three-dimensional array of color pixels. Two of the dimensions represent the horizontal and vertical spatial directions of the moving pictures, while the third dimension represents the time domain. A video frame represents a set of pixels that correspond to a *'snap-shot'* at a single point in time.

The first stage of video compression is the *quantizer*, which rounds the continuous valued analogue video signal amplitudes arriving from the camera to a set of suitably chosen discrete values referred to as the quantization levels. The discrete pixel values are then represented by bit patterns, which are communicated to the decoder. Owing to the rounding process of the quantizer the decompressed pixel values become different from the original analogue ones, hence the quantizer is the element responsible for striking a trade-off between the bit/information rate and video distortion. The theory and optimal design of scalar quantizers under different constraints has been widely studied [12], resulting in different categories of quantizers such as uniform [13] and Lloyd-Max [14] quantizers.

The extension of scalar quantization to the joint quantization of several pixels, which is referred to as vector quantization (VQ) [15–18], represented a major step towards approaching the rate-distortion bounds derived for dependent sources. Naturally, VQs increase both the delay and complexity. Gersho [19] proposed these principles for image and video compression in 1982.

An alternative technique of exploiting the dependencies or correlation of video signals is constituted by predictive or differential compression strategies. In predictive compres-

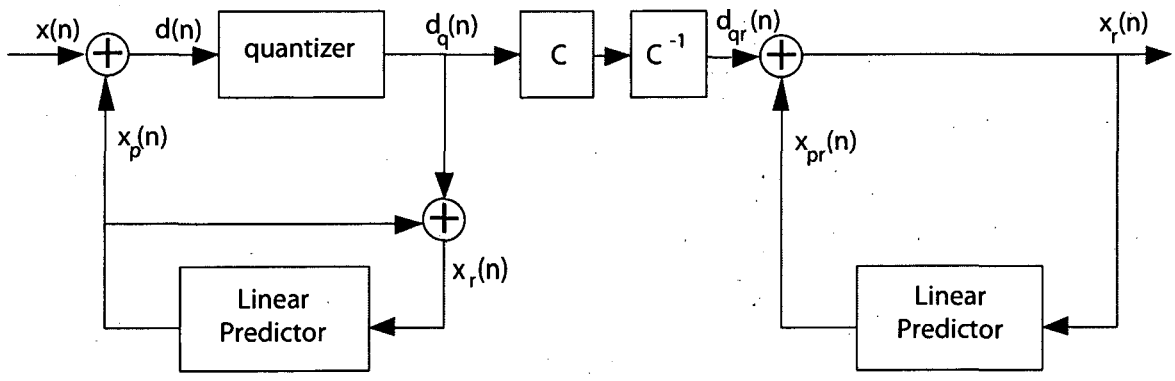


Figure 1.2: Basic Differential PCM (DPCM) compression.

sion, the signal amplitudes are predicted on the basis of the neighboring signal amplitudes. In order for the decoder to be able to exactly reproduce a pixel value predicted by the encoder, the prediction mechanism has to be fed with already quantized pixel amplitudes. This leads to the basic predictive compression technique referred to as Differential Pulse-Code Modulation (DPCM), which is illustrated in Figure 1.2. The linear predictor produces the predicted signal $x_p(n)$ from the reconstructed signal $x_r(n)$ according to:

$$x_p(n) = \sum_{j=1}^{J} a_j \cdot x_r(n-j), \tag{1.1}$$

where a_j denotes the prediction coefficients, for $j = 1, 2, ..., J$. The prediction coefficients are calculated by expressing the Mean Squared Error (MSE) expression of

$$MSE = E[x(n) - \sum_{j=1}^{J} a_j \cdot x_r(n-j)]$$

between the original and compressed signal and setting its partial derivative with respect to a_j to zero, which gives J equations for the J coefficients. The extension from the above one-dimensional ($1D$) prediction model to $2D$ is straightforward. The first standardized video codec, which was developed by the European COST211 project and standardized by the International Telecommunication Union's study group T (ITU-T) as the H.120 standard in the early 1980s, uses spatial DPCM of the video frames, operating at a 2 Mbit/s compressed bit rate. The resultant video quality is far from unimpaired because (i) the temporal correlation of the consecutive video frames is not exploited; (ii) pixel-by-pixel scalar quantization is used instead of more efficient VQ schemes. In order to

improve the achievable video quality, two research and development directions were vigorously pursued, namely block-based transform coding [20–23], which aims for exploiting the spatial correlation inherent within a frame and motion estimation as well as motion compensation [24–28] in order to exploit the temporal correlation of consecutive frames along the motion trajectories. These developments led to the design of block-based image codecs and motion-compensated block-based video codecs during the late 1980s, which form the foundation of today's image and video compression standards such as JPEG, MPEG2, MPEG4 as well as H.263/H.264.

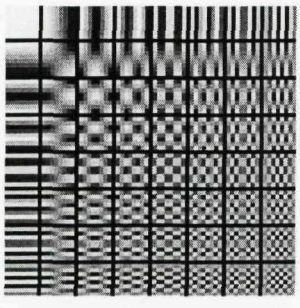


Figure 1.3: Basis functions of the DCT (8x8 blocks of pixels).

During the late 1980s, a number of block-based transform coding proposals designed for video conferencing were submitted to ITU-T [29–32]. Except for one, all the proposals were based on the Discrete Cosine Transform(DCT). In parallel to the ITU-T's investigations during the period of 1984-1988, the Joint Photographic Experts Group (JPEG) was aiming for the compression of still images. They also opted for DCT-based compression of 8x8 pixel blocks. The decision of the JPEG group undoubtedly influenced the ITU-T in opting for the 8x8-pixel DCT for spatial decorrelation as a basis for their video compression standard known as H.261 [32].

The DCT decomposes a block of image pixels into a set of orthogonal basis functions, typically referred to as basis images in image and video compression. The benefit of this decomposition is that provided that the 8x8-pixel video tiles have similar luminance, say a given gray colour, then only the first DCT coefficient is high. Accordingly, the higher-frequency DCT coefficients may be set to zero, which substantially reduces the required bit rate. The 64 basis images of the 8x8-pixel DCT are shown in Figure 1.3. The weights of the individual DCT basis images, which are referred to as DCT coefficients, are quantized, entropy encoded and transmitted to the decoder. The reconstruction of the original 8x8-pixel video tile may be viewed as superimposing all the 64 basis images of Figure 1.3 on each other, after they were weighted by the quantized DCT coefficients.

An alternative to the DCT-based video decomposition is a subband or wavelet decomposition [20, 33, 34]. Since these schemes are conceptually somewhat more complex, their philosophy will be developed more gradually. Efficient implementations of subband/wavelet decomposition now exist, using for example so-called lifting schemes [35,36], and a variety of ways has been found for making quantization of subband/wavelet coefficients adaptive, so that the perceptually most important visual information is represented at the best possible quality. Subband/wavelet decompositions are currently used in the JPEG2000 image compression standard [5], and in the audio compression standards such as MPEG-1 Audio Layer 3 [37] commonly referred to as MP3.

Due to the popularity of motion-compensated DCT systems [30–32, 38], motion estimation techniques became well developed, yielding both theoretical concepts and a wide variety of practical motion-estimation algorithms [28, 39, 40]. In a video compression context typically a temporal DPCM coding scheme is used, where a motion-compensated block-based prediction of the current (8x8)-pixel video tile is created based on the previous video frame. The difference between the motion-compensated prediction and the actual pixel value results in MCER frame, which is spatially compressed using intra-frame compression and sent to the decoder. In video compression, the motion estimators are typically constituted by conceptually simple block-based searching procedures [7, 40]. However, in terms of their computational demands they may be extremely complex, when evaluating the correlation of the (8x8)=64-pixel video tiles for a high number of motion-vectors [7]. When designing computationally efficient block-based motion estimators often hierarchical reduced-search block-matching motion estimators were favoured [7]. The first standardized motion-compensated DCT-based video codec designed for video conferencing is known as the *H*.261 scheme [32], which can be configured to operate at bit rates between 384 kbit/s and 1.15 Mbit/s.

In the early 1990s, the ISO Moving Picture Experts Group (MPEG) started investigating compression techniques for storage of video on CD-ROMs. The resultant standard, known as MPEG-1 [31], has been very successful. MPEG-1 encoders and decoders/players are widely used on multimedia computers and for video playback in Asia. Since MPEG-1 was not designed for efficient compression for interlaced signals, its successor, namely MPEG-2 [38] became the standard for the broadcasting and storage of digital standard TV signals. The ISO MPEG-2 standard is also known as the ITU-T H.262 standard.

Following the success of MPEG-2, the development in compression techniques pursued

four different objectives:

- Achieving an increased compression ratio at a given quality. This has resulted in the H.263 and H.264 video compression standards [4]. Alternative proprietary compression schemes also exist, such as RealVideo [41] or x264 [42] etc. Although the resultant bit streams are incompatible with any standard, the heart of the underlying compression scheme is still a motion-compensated DCT-based encoder.

- Supporting application in Internet-based or wireless communication scenarios, where communication channel may corrupt the compressed bit stream in various ways. Hence robust, joint source-channel coding schemes [43] were developed [44–46].

- Designing perceptually optimized multi-rate codecs for employment in High Speed Downlink Packet Access (HSDPA) systems, where the channel quality determines the number of video bits that can be delivered by the adaptive modulation and coding aided HSDPA modem.

- Developing region or object-based compression. This is the most revolutionary step in moving away from the rectangular-tile-aided DCT-based compression philosophy. The basic unit for motion estimation and decorrelation is no longer an (8x8)-pixel or (16x16)-pixel blocks, but an arbitrarily-shaped area of pixels that is homogeneous or correlated in a more meaningful way [47]. The MPEG-4 standardization [3] has contributed significantly to the research of region/object-based image and video compression.

1.2.1.2 Image and Video Coding Standards

Naturally, the most promising research results in advanced image and video compression techniques have found their way into practical standards. The evolution of video compression standards over the past half-a-century is summarized in Figure 1.4. The International Standard Organization (ISO) (jointly with the International Electrotechnical Commission (IEC), also known as ISO/IEC), and the International Telecommunication Union (ITU) are the major international standardization bodies responsible for image and video coding standards. While the ISO/IEC proposed the family of the Joint Photographic Experts Group (JPEG) and the Moving Pictures Expert Groups (MPEG) coding standards, such

as JPEG [48], MPEG-1 [31] and MPEG-2 [38], ITU is in charge of the ITU-T recommendations H.26x for video conferencing and telephony such as H.261 [32], H.263 [49] and H264 [4].

Each coding standard aims at its specifically targeted application. For instance, the JPEG standard targets still image compression applications, whereas the family of MPEG standards targets video compression applications. The difference between image and video coding standards lies in the fact that video coding utilizes not only intra-frame redundancy, but also inter-frame redundancy. For reasons of space economy, here we only briefly review the three other coding standards representing the state-of-the art in image and video compression, i.e., the MPEG standards the JPEG-2000 standard [5] and the ITU-T recommendation H.264.

1. **The MPEG Standards**

   The "Moving Pictures Experts Group" (MPEG) was established in 1988 [50] within the International Standard Organization's (ISO) Steering Group (SG), which was responsible for the encoding of video and audio. The MPEG group commenced the development of the MPEG-1 standard in 1988, released the MPEG-1 standard in 1993 and embarked on the standardization of the MPEG-2 scheme in 1990 [51]. The MPEG-1 standard was mainly targeted at CD-ROM applications dedicated to recording video at bit rates of up to 1.5Mb/s [51]. By contrast, the MPEG-2 standard was designed for substantially higher quality, namely for audiovisual applications such as today's home entertainment systems and digital broadcasting systems requiring video bit rates between 2Mb/s and 30Mb/s. The MPEG-4 standardization process was initiated in 1994 - with the mandate of standardizing algorithms for audio-visual coding in multimedia applications, while allowing for interactivity, and supporting high compression as well as universal accessibility and portability of both the audio and video content.

   MPEG-4 is a standard used primarily to compress audio and video (AV) signals. Introduced in late 1998, it is constituted by a group of audio and video coding standards and related technology agreed upon by the ISO/IEC MPEG under the formal standard ISO/IEC 14496. Diverse applications of the MPEG-4 standard can be found in Internet streaming media and CD distribution, videophones and broadcast television, all of which benefit from compressing the AV stream. Additional

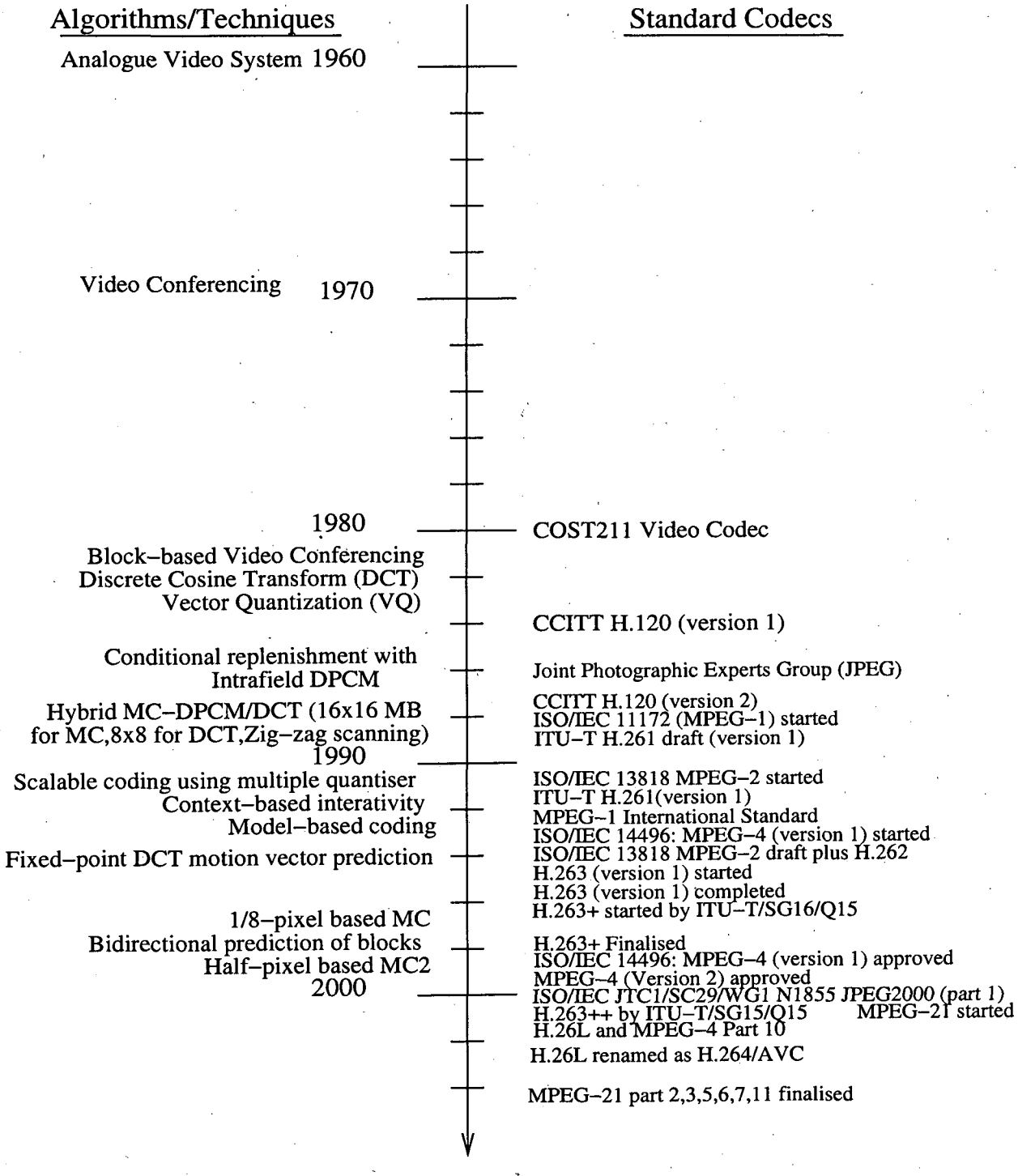


Figure 1.4: An overview of video compression standards [7].

compression techniques and coding profiles were added in the form of two amendments of the MPEG-4 Version 1 standard, leading to Version 2 during late 2001. The operating bit rates targeted by the MPEG-4 video standard are between 5 and 64kbit/s in the context of mobile or Public Switched Telephone Network (PSTN) based video applications, spanning up to 4 Mbit/s for digital TV broadcast applications, and even to rates in excess of 100Mbit/s in High Definition TV (HDTV) studio applications. Some of the key features of the MPEG-4 video standard, which are superior in comparison to the previous video coding standards, can be listed as follows:

- The core compression techniques are based on those of the ITU-T H.263 standard, which are more efficient than those of the MPEG-1 and MPEG-2 video compression schemes. Efficient compression of progressive and interlaced video sequences as well as optional additional tools were introduced for the sake of further improving the attainable compression efficiency.
- Coding of video objects, having both rectangular shapes and irregularly shaped object. This is a new concept in the context of standard-based video coding and enables the independent encoding of both foreground and background objects in a video scene.
- Support for error-resilient transmission over hostile networks. A range of error resilience enhancement techniques were included in the MPEG-4 codec for the sake of assisting the decoder in recovering from transmission errors and for maintaining a successful video connection in an error-prone network. Furthermore, the scalable coding techniques are capable of supporting flexible transmission at a range of desired coded bitrates.
- Coding of still images within the same framework as full-motion video sequences.
- Coding of animated visual objects, such as 2D and 3D computer-generated polygonal meshes, animated objects, etc.
- Coding for specialist applications, such as very high quality medical applications, etc. In this application maintaining a high visual quality is more important than attaining a high compression.

2. **JPEG-2000 Standard**

JPEG-2000 [5] is a wavelet-based image compression standard. It was created by the Joint Photographic Experts Group committee with the intention of superseding their original discrete cosine transform-based JPEG standard. The JPEG-2000 standard achieves higher coding efficiency in comparison to the baseline JPEG standard without generating the characteristic 'blocky and blurry' artifacts of the original DCT-based JPEG standard. Part 1 of JPEG 2000 has been published as an ISO standard, ISO/IEC 15444-1:2000.

The targeted applications include Internet imaging, wireless image transmission, digital imaging, remote sensing, medical imagery, etc. Some of the key features offered by the JPEG2000 standard are summarized as follows [52]: (1) Superior low bit-rate performance; (2) Lossless and lossy compression; (3) Random codestream access and processing; (4) Robustness to bit errors; (5) Region-of-interest (ROI) coding; and (6) Open architecture.

The JPEG-2000 standard takes advantage of a few recent research advances in image compression in order to achieve the above features. The major notable characteristics of the new standard include: (1) The DCT is replaced with the Discrete Wavelet Transform (DWT); (2) Embedded scalar quantisation using a dead-zone quantizer [53]; (3) Embedded block coding with optimized truncation (EBCOT); (4) Huffman coding is replaced by context based adaptive binary arithmetic coding [54, 55]. Although the JPEG-2000 standard provides a superior compression ratio in comparison to other still image compression standards, this is achieved at the expense of an increased memory and computation complexity.

3. **The H.26L/H.264 Standard [4]**

The H.264 [4] coding standard section activity were initiated in 1997 by the ITU-T Video Coding Experts Group (VCEG) together with the ISO/IEC Moving Picture Experts Group (MPEG) as the product of a collective partnership known as the Joint Video Team (JVT), following the earlier development work of the ITU-T as a VCEG project referred to as H.26L [56–58]. The prime objective of the H.264/AVC project was to create a standard that would be capable of providing good video quality at substantially lower bit rates (e.g., half or less) than previous standards would require (e.g., relative to MPEG-2, H.263, or MPEG-4 Part 2) without substantially increasing the implementational complexity. An additional goal was to provide sufficient flexibility for the standard to be applied in a wide variety of applications

(e.g., for both low and high bit rates, and for low and high resolution video) and to communicate over a wide variety of networks and systems (e.g., for broadcast, DVD storage, RTP/IP packet networks, and ITU-T multimedia telephony systems) [59].

H.264/AVC/MPEG-4 Part 10 contains a number of novel features that allow us to compress video more effectively than the older standards and to provide a higher grade of flexibility for applications in a wide variety of networks and propagation environments. In particular, some of its key features include:

- *Multi-picture and inter-picture prediction including the following features [59, 60]*:
  - Using previously encoded pictures as motion-compensation references in a significantly more flexible way than in past standards, allowing up to 32 reference pictures to be used in some cases. By contrast, in prior standards, typically a single reference frame was used for the P-picture, while in case of B-pictures two reference frames were employed. This particular feature typically allows useful improvements to be achieved in terms of both the bit rate and video quality in most video scenes. However, in specific types of scenes, such as for example scenes with rapid repetitive flashing or back-and-forth scene cuts, it allows a very significant bit rate reduction.
  - Variable block-size based motion compensation (VBSMC) using block sizes as large as 16x16 and as small as 4x4 pixels, enabling a flexible segmentation of moving regions.
  - Six-tap filtering for the derivation of half-pixel luminance sample predictions, in order to lessen the aliasing effects and hence to provide sharper images.
  - Quarter-pixel precision for motion compensation, enabling a precise description of the displacements of moving areas. For the chroma components the resolution is typically halved both vertically and horizontally, therefore the motion compensation precision becomes as accurate as one-eighth pixel.
  - Weighted prediction, allowing an encoder to specify the use of a multiplicative scaling factor and additive offset, when performing motion compensation, and providing a significant performance benefit in specific cases, such as fade-to-black, fade-in and cross-fade transitions.

- *Spatial prediction of the intra-coded frames on the basis of the neighboring blocks* rather than predicting only the "DC" component, as in MPEG-2 Part 2 or only the transform coefficients as in H.263+ and MPEG-4 Part 2 [59].
- *Lossless macroblock coding features, including*:
  - A lossless PCM macroblock representation mode in which the video pixels are represented directly, allowing perfect representation of specific regions, while still allowing a strict limit to be imposed on the number of coded bits for each macroblock.
  - An enhanced lossless macroblock representation mode allowing near-perfect reconstruction of specific regions, while potentially using substantially fewer bits than the PCM mode (not supported in all profiles).
- *Flexible interlaced-scan video coding features* (not supported in all profiles), including:
  - Macroblock-adaptive frame versus field (MBAFF) coding [59] allows 16x16-pixel macroblocks of pictures coded in the frame mode to be coded in field mode using the structure of a macroblock pair [59].
  - Picture-adaptive frame versus field coding [59] (PAFF or PicAFF) allowing a freely-selected mixture of pictures to be coded as MBAFF frames with pictures coded as individual single fields (half frames) of interlaced video.
- *New transform design features*, including [59]:
  - An integer (4x4)-pixel spatial block transform (conceptually similar to the well-known DCT design, but simplified and designed to avoid any further processing-induced errors to be imposed) allows the scheme to avoid the finite-precision induced "ringing" often observed in prior codec designs.
  - An exact-match based integer (8x8)-pixel spatial block transform which is conceptually similar to the well-known DCT design, but simplified and again designed to provide exactly-specified decoding, which is not supported in all coding profiles, allowing highly correlated regions to be compressed more efficiently than with the aid of the (4x4)-pixel block transform.
  - Adaptive encoder selection for the (4x4)-pixel and (8x8)-pixel transform block sizes for the integer transform based operation (not supported in all

coding profiles).

- A secondary Hadamard transform performed on the "DC" coefficients of the primary spatial transform (for the chrominance DC coefficients and also for the luminance coefficients in a specific case) in order to increase the achievable compression in smooth picture regions.

- *A quantization design, including [59]*:
  - Quantization step size control for the sake of flexible bit rate management by the encoders and for attaining a simplified inverse-quantization scaling.
  - Customized quantization scaling matrices selected by the encoder for the sake of perceptually motivated quantization (not supported in all coding profiles).

- An in-loop deblocking filter, which helps prevent the blocking artifacts common to other DCT-based image compression techniques.

- *An entropy coding design, including [59]*:
  - Context-based Adaptive Binary Arithmetic Coding [54] (CABAC), which is a sophisticated technique designed for losslessly compressing the syntax elements in the video stream by exploiting the knowledge of the probabilities of all the syntax elements in a given context (not supported in all coding profiles).
  - Context-Adaptive Variable-Length Coding (CAVLC) [59, 61, 62], which is a lower-complexity alternative to the above-mentioned CABAC technique designed for the coding of quantized transform coefficient values. Although it has a lower complexity than CABAC, CAVLC is more elaborate and more efficient than the methods typically used to code coefficients in previous video codec designs.
  - A simple and highly structured variable length coding (VLC) technique designed for many of the syntax elements not coded by CABAC or CAVLC, which is referred to as Exponential-Golomb coding [59, 62].

- *Packet-loss resilient features, including*:
  - A network abstraction layer (NAL) [59] definition allowing the same video syntax to be used in many network environments, including features such as sequence parameter sets (SPSs) and picture parameter sets (PPSs) that

provide a higher robustness and flexibility than those provided in prior designs.

- Flexible macroblock ordering (FMO) [59, 60] (also known as slice groups, which are not supported in all coding profiles) and arbitrary slice ordering (ASO), which are techniques designed for restructuring the ordering of the representation of the fundamental regions, such as macroblocks in pictures. Typically the FMO and ASO techniques are considered to be error/loss robustness enhancement features, but FMO and ASO can also be used for other purposes.
- Data partitioning (DP) [59], a feature separating more important and less important syntax elements into different transmission packets for unequal error protection (UEP), but DP can also be used for other types of error/loss robustness enhancement techniques (not supported in all coding profiles).
- The generation of redundant slices (RS) [59] also constitutes an error/loss robustness enhancement feature, allowing an encoder to send an extra representation of a particularly important picture region (typically at a lower resolution) that can be used, if the primary representation is corrupted or lost (not supported in all coding profiles).
- Frame numbering constitutes a feature that allows the creation of "subsequences" (enabling temporal scalability by optional inclusion of extra pictures between other pictures), and the detection as well as concealment of the loss of entire pictures (which may occur due to network-induced packet losses or channel errors).

- Switching slices (called SP and SI slices and not supported in all coding profiles), which allow an encoder to direct a decoder to jump into an ongoing video stream for reasons, such as video streaming induced bit rate switching operations and "trick mode" operation. When a decoder jumps into the middle of a video stream using the SP/SI feature, it can get an exact match to the decoded pictures at that location in the video stream, despite using different pictures (or no pictures at all) as motion-compensation references prior to the switch.
- A simple automatic process was designed for preventing the accidental emula-

tion of start codes, which are special sequences of bits embedded in the coded data that allow random access to the bitstream and support the recovery of byte alignment in systems that may lose byte synchronization.

- Supplemental enhancement information (SEI) [59] and video usability information (VUI) [63], which constitute extra information that can be inserted into the bitstream to enhance the versatility of the video used for a wide variety of purposes.
- Support of monochrome mode, 4:2:0, 4:2:2, and 4:4:4 color sampling structures [63](depending on the selected coding profile).
- Support of different pixel precisions, ranging from 8 to 14 bits per video sample (depending on the selected coding profile).
- The ability to encode individual color planes as distinct pictures based on their own slice structures, macroblock modes, motion vectors, etc., and allowing encoders to be designed with a simple parallel implementational structure (supported only in the three 4:4:4-capable coding profiles)
- Picture order count, a feature that assists in keeping the ordering of the pictures and the values of samples in the absence of timing information (allowing timing information to be carried and controlled/changed separately by a system without affecting the decoded picture content).

These techniques, along with several others, assist the H.264 codec to perform significantly better than any prior standard, under a wide variety of circumstances in a wide variety of application environments.

1.2.2 Video Communication

1.2.2.1 Error Control Coding

Channel coding [64] is another major component of the wireless video communication system depicted in Figure 1.1. Regarding the transmission of image and video signals through wired or wireless channels, there are two main types of error control techniques concerned with the protection of error sensitive video signals, namely Automatic Repeat reQuest (ARQ) [65,66] and Forward Error Correction (FEC) [64], respectively. The ARQ technique retransmits each video frame, until it is either correctly received or the errors

persist beyond a predetermined number of retransmissions. The family of ARQ techniques perform well in circumstances, where the bit error probability of the channel is relatively low, but in Additive White Gaussian Noise (AWGN) channels each retransmitted frame may experience the same errors and hence hybrid-ARQ Type-II systems [67, 68] tend to transmit additional redundancy for the channel decoders, until error-free reception is achieved. However, in heavily error-prone environments, ARQ wastes a high amount valuable channel bandwidth and hence may cause traffic congestion in packet-switched networks, such as the Internet. Another disadvantage of ARQs is that they require a feedback channel. FEC coding [64], on the other hand, adds intentional redundancy to the information source in order to increase the achievable transmission reliability without ARQs. At the receiver end, the channel decoders exploit the redundancy in order to correct the transmission errors. By controlling the amount of added redundancy, adaptive FEC coding is capable of reliable operation over transmission channels for medium to severely noisy conditions without ARQs.

Again, a sophisticated channel coding module invoked in a robust video communication system may incorporate both FEC and ARQs, provided that the ARQ-delay does not affect lip-synchronization. Missing or corrupted packets may be recovered at the receiver, as long as a sufficiently high fraction of packets is received without errors [7,69,70]. FEC is also widely used for providing unequal error protection (UEP), where the more vulnerable bits are protected by stronger FEC codes. Recent work has addressed the problem of how much redundancy should be added and distributed across differently prioritized data partitions [71]. In addition to FEC codes, typically interleaving is also employed for providing enhanced protection [72] by near uniformly spreading the effects of error across FEC blocks. Unlike fixed-mode FEC schemes, ARQ intrinsically adapts to the varying channel conditions and hence in many applications tends to be more efficient. However, in the context of real-time communication and low-latency streaming the latency introduced by ARQ may become a major concern, unless very short-duration transmission blocks are retransmitted, such as the blocks used by the ARQ-regime of the High-Speed Downlink Packet Access (HSDPA) system [73,74]. Layered or scalable coding, combined with transmission prioritisation, is another effective approach devised for providing an increased error resilience [75]. In a layered scheme, the source signal is encoded by generating more than one different-significance groups or layers, with the base layer containing the most essential information required for media reconstruction at an acceptable quality, while

the enhancement layer(s) contain information that may be invoked for reconstruction at an enhanced quality. At high packet loss rates, the more important, more strongly protected layers can still be recovered, while the less important layers might not. Commonly used layered techniques may be categorized into temporal scalability, spatial scalability, Signal-to-Noise Ratio (SNR) scalability [76], data partitioning [4], or any combinations of these. Layered scalable coding has been widely employed for video streaming over best-effort networks, including the Internet and wireless networks [77]. Different layers can be transmitted under the control of a built-in prioritization mechanism without network support, such as the UEP scheme mentioned above, or using network architectures capable of providing a variety of different Qualities of Service (QoS). A scheme designed for optimal Intra/Inter-mode selection has recently been proposed for scalable coding, in order to limit the inter-frame error propagation inflicted by packet losses. Another scheme devised for adaptive bit rate allocation in the context of scalable coding was presented in [78]. Layered scalable coding has become part of various established video coding standards, such as for example the members of the MPEG [2,38] and H.264 [4] codec family.

Dogan, Cellatoglu, Uyguroglu, Sadka and Kondoz [78] reported promising results on adopting the MPEG-4 codec for wireless applications by exploiting the rate control features of video transcoders and combined them with error resilient General Packet Radio Service (GPRS) type mobile access networks. Further important contributions in the area of joint source and channel coding entail the development of the scheme that offers the end to end joint optimization of source coding and channel coding/modulation schemes communicating over wireless links, for example those by Kliewer and Thobaben [79].

It is not the intention of this thesis to provide a comprehensive coverage of all the known error control codes. Rather, we provide a brief introduction of turbo decoding techniques that are extensively used in the simulations throughout this thesis.

1.2.2.2 Turbo Principle

The method of iteratively decoding a concatenated channel code by exchanging soft information between two decoding components using the turbo principle was first presented in [80]. Of all practical error correction methods known to date, initially turbo decoding techniques were shown to come closest to approaching the Shannon limit, the theoretical limit of maximum information transmission rate over a noisy channel. However, in recent

years numerous other codecs were shown to approach these limits [7, 64, 81]. The key insight is that instead of producing a stream of binary symbols from the received signal with the aid of hard-decisions, the decoder can be designed to produce a likelihood measure for each bit. The main motivation of this thesis is to design iterative detection aided video tranceivers for wireless channels.

1.3 Organization of Thesis

In general terms, the focus of this thesis is on the design of wireless video communication systems using the BBC's Dirac codec [8, 82] as a source codec, although some results are also provided for the H.264 codec [83]. The remainder of thesis is organized as follows:

- Commencing with an overview of the BBC's Dirac Wavelet video codec, Chapter 2 describes the Dirac codec's compression algorithms in detail. At the end of chapter 2, the PSNR objective video quality measure was used for characterizing the codec and the relative frequency of occurrence for the various Dirac video codec parameters was quantified in support of our further elaborations in the forthcoming chapters.

- Chapter 3 has been divided into two parts, the first half characterizes the effects of transmission errors on the Dirac video-encoded sequence, quantifying the sensitivity of the various bits, while the second part elaborates a the design of a novel Unequal Error Protection (UEP) scheme [82], namely on iterative detection aided unequal error protection using Irregular Convolutional Codes (IRCCs) for protecting the Dirac video-encoded sequence. The proposed UEP scheme is capable of attaining a near-unimpaired video quality for channel Signal-to-Noise Ratios (SNRs) in excess of about 4.5dB for transmission over a five-path dispersive AWGN channel. It was demonstrated that the proposed scheme outperforms the Equal Error Protection aided (EEP) benchmarker scheme by about 1dB, which protects the whole coded bitstream by rate-1/2 channel coding.

- In Chapter 4, we focus our attention on the design of a powerful ISCD scheme employing a so-called bit-to-symbol over-complete mapping scheme [84]. First of all, a softbit source decoding technique designed for Iterative Source-Channel Decoding (ISCD) was investigated. This is followed by the study of an ISCD scheme employing a specific bit-to-symbol over-complete mapper, which benefits both from the

residual redundancy inherent in the source encoded bitstream as well as from the intentional redundancy imposed by the proposed over-complete source mapping. We characterize the entire system's robustness against transmission errors, as well as the implementational complexity of the OverComplete Mapping-assisted Softbit Source Decoding (OCM-SSD) scheme. Finally, the effects of different interleaver delays on the OCM-SSD's performance is quantified by analyzing the EXIT characteristics of the ISCD scheme. The experimental results provided in this chapter demonstrated that at the same overall system code-rate of 0.5, the ISCD scheme designed with the aid of our over-complete mapper outperforms the ISCD scheme using no mapping by about 2dB, when communicating over the uncorrelated Rayleigh fading channel.

- In Chapter 5, a further developed wavelet video video telephony system was proposed and investigated. We first designed an unequal error protection based irregular over-complete mapper for the Dirac codec while using iterative source and channel decoding, followed by the study of a near-instantaneously adaptive transceiver, which is capable of quantifying the probability density function of the video-encoded sequence. This allows our wireless video telephony scheme to attain iterative detection gains without increasing the system's delay, regardless of the choice of the source codec - even for speech and audio codecs, and hence makes our technique widely applicable.

- The summary of the thesis and an outlook on future work are given in Chapter 6.

The inter-dependence of the chapters is illustrated in Figure 1.5. Having provided an overview of the thesis, let us now commence our discourse with a brief overview of the Dirac video codec.

1.4 Novel Contributions

In summary, the novel contributions of this thesis are:

- Quantifying the relative frequency of occurrence for the various Dirac video codec parameters and investigating their bit-sensitivity allowed us to design a novel UEP scheme using iterative detection aided IRCCs for protecting the Dirac video-encoded sequence [82].

- A powerful ISCD scheme employing a novel OCM scheme was designed [84, 85].
- The necessary and sufficient condition for ensuring the iterative decoder's convergence to an infinitesimally low Bit Error Ratio (BER) was formulated in [84, 85].
- The overall system's robustness against transmission errors as well as the implementational complexity of the OCM-assisted Softbit Source Decoding (OCM-SSD) scheme was characterized [85].
- The effects of interleaver delay on the OCM-SSD's performance was quantified using EXIT charts [85].
- The trade-offs between the achievable decoding performance and the mapping rate were quantified [85].
- An Irregular Over-Complete Mapping (IrOCM) was designed [86].
- A near-instantaneously adaptive transceiver, which is capable of quantifying the probability density function of the video-encoded sequence was proposed.

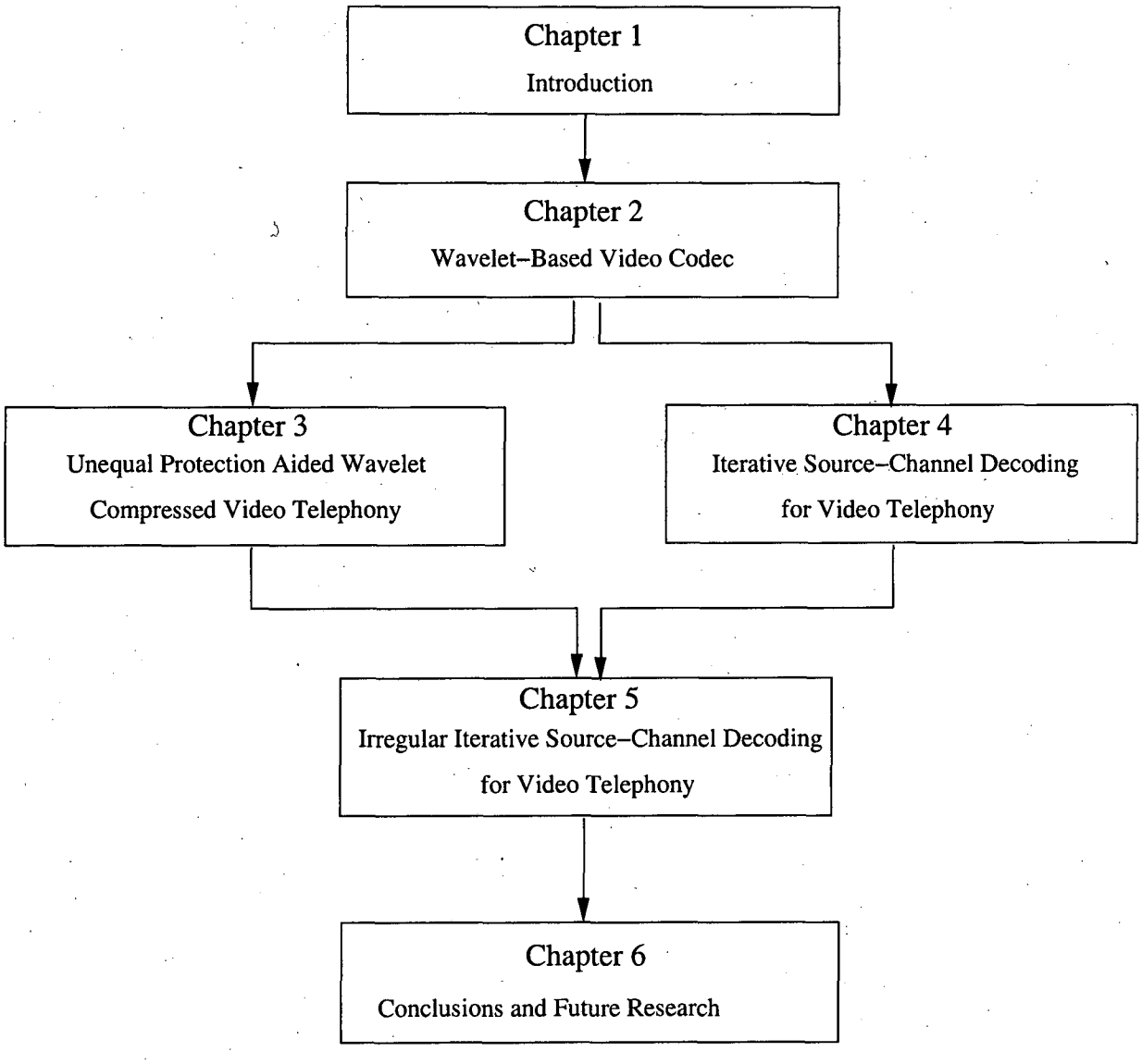


Figure 1.5: Inter-dependence of the thesis chapters.

Chapter 2

Wavelet-Based Video Codec

2.1 Introduction

Wavelet-based video compression has been successfully developed by Joint Photographic Expert Group (JPEG) for the JPEG2000 standard [5]. Its advantages have been demonstrated in both image processing as well as in video compression [20,33,87–89] and hence it was also adopted by the British Broadcasting Corporation's recent proprietary video codec referred to as Dirac[1]. The Dirac compression algorithm is capable of achieving an approximately two-fold bit rate reduction over MPEG-2 [38], when compressing high definition video (e.g. 1920x1080 pixels). The design of the Dirac codec hinges on striking an attractive trade-off, amongst numerous codec parameters, including factors such as the achievable compression ratio, the implementational complexity, the encoding delay, etc [7]. Recently, it has been further developed to optimize its performance for supporting a wide range of applications, including the Internet and mobile streaming, delivery of standard-definition and High-Definition Television (HDTV), digital television and cinema production as well as distribution low-power devices and embedded applications.

The basic philosophy of two-dimensional subband or wavelet-based video coding is that the video frame or the Motion Compensated Error Residual (MCER) frame is decomposed into a number of subbands containing the different-frequency representations of the original video or MCER frame. The frame decomposition process is typically carried out at a number of consecutive levels for the sake of producing different-frequency

[1]British Broadcasting Corporation Research and Development Department developed Dirac from January 2003 and it is still in stage of development

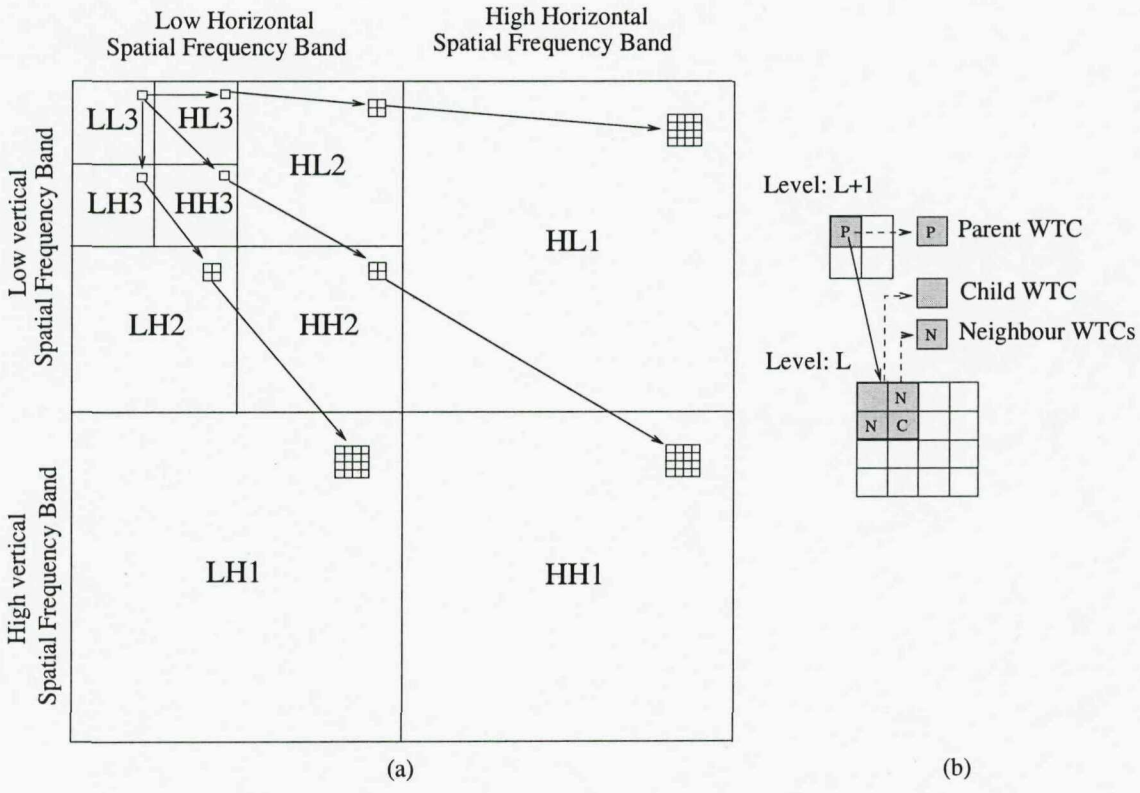


Figure 2.1: Three-level wavelet decomposition illustrating the parent-child relationship in the spatial frequency domain.

subbands, for example as outlined on page 629 of [7], where a video frame is decomposed into 10-subbands in a 3-level decomposition process. More specifically, at the first decomposition level a frame is split into four 'quadrants', as seen in Figure 2.1. Each quadrant represents a High (H) or a Low (L) spatial frequency band in the horizontal as well as vertical direction of the original full-band frame. For example, the quadrant HH1 contains the high-horizontal and high-vertical frequency information of the original full-band frame. By contrast, the HL1 and LH1 quadrants indicate that the high horizontal but low vertical and the low horizontal but high vertical frequency components have been retained. All these frame-quadrants appear to have a quarter of the original frame's size because they are subsampled by a factor of two owing to having half the bandwidth. In the second-level wavelet decomposition, the subjectively most important LL1 quadrant is split further into four subbands, namely HH2, HL2, LH2 and LL2. This frequency band based splitting allows the designer to selectively allocate the source encoding bits to those particular frequency bands, where they result in the most significant improvements of video quality. The Dirac video codec used Daubechies' so-called 9/7 biorthogonal fil-

ters outlined in [35], where 9 and 7 denote the number of coefficients in the analysis and synthesis filters.

To elaborate a little further, the Dirac video codec is a hybrid video compression scheme utilizing wavelet transforms [20, 34] and motion compensation. A key element of its flexibility is the employment of the multi-resolution wavelet transform for compressing pictures and motion-compensated residuals, which allows the Dirac video codec to be used across a wide range of resolutions.

Following the above brief overview of the wavelet decomposition, a more detailed portrayal of the Dirac video codec will be provided in Section 2.2, whilst a range of efficient techniques invoked in the Dirac video codec is presented in Section 2.3. The subjective video quality performance of the Dirac video codec, the effect of the various coding parameters as well as our experimental results are presented in Section 2.4. For the sake of a concise treatment, we assume that readers are familiar with the wavelet transform, which can be read in [20, 33, 87–89] and [35, 90, 91].

2.2 Overview of Wavelet Video Codec

The Dirac video codec can operate on pictures scanned at various rates (e.g. at 15, 25 or 30 frames per second) and encoded in the 8-bit (Y, U, V) format[2]. The codec supports several sampling modes, such as: 411[3], 420[4], 422[5], or 444[6] mode, and can operate on the different video resolution: the 176x144 pixels Quarter Common Intermediate Format (QCIF), the 352x288 pixels Common Intermediate Format (CIF), the 1280x720 pixels High Definition Video (HD720), and 1920x1080 pixels HD1080. Figure 2.4 depicts the spatial relationship between luminance and color difference samples in 420 (Y, U, V).

Overall, the Dirac video codec is a classic motion-compensated hybrid codec. The block diagram of encoding algorithm, as seen in Figure 2.2, consists of five main elements:

i. The wavelet transform of the video frame.

[2]The Y component represents the luminance of the pixel, and the U and V components represent the corresponding blue and red colors.

[3]The color components (U, V) are sampled at quarter the rate of the luminance on horizontal scan line, whereas on vertical scan line at the same rate of the luminance

[4]The color components (U, V) are sampled at quarter of the luminance rate on each scan line

[5]The color components (U, V) are sampled at half of the luminance rate on the horizontal scan line, whereas on the vertical scan line at the same rate as the luminance

[6]The color components (U, V) are sampled at the same rate as the luminance

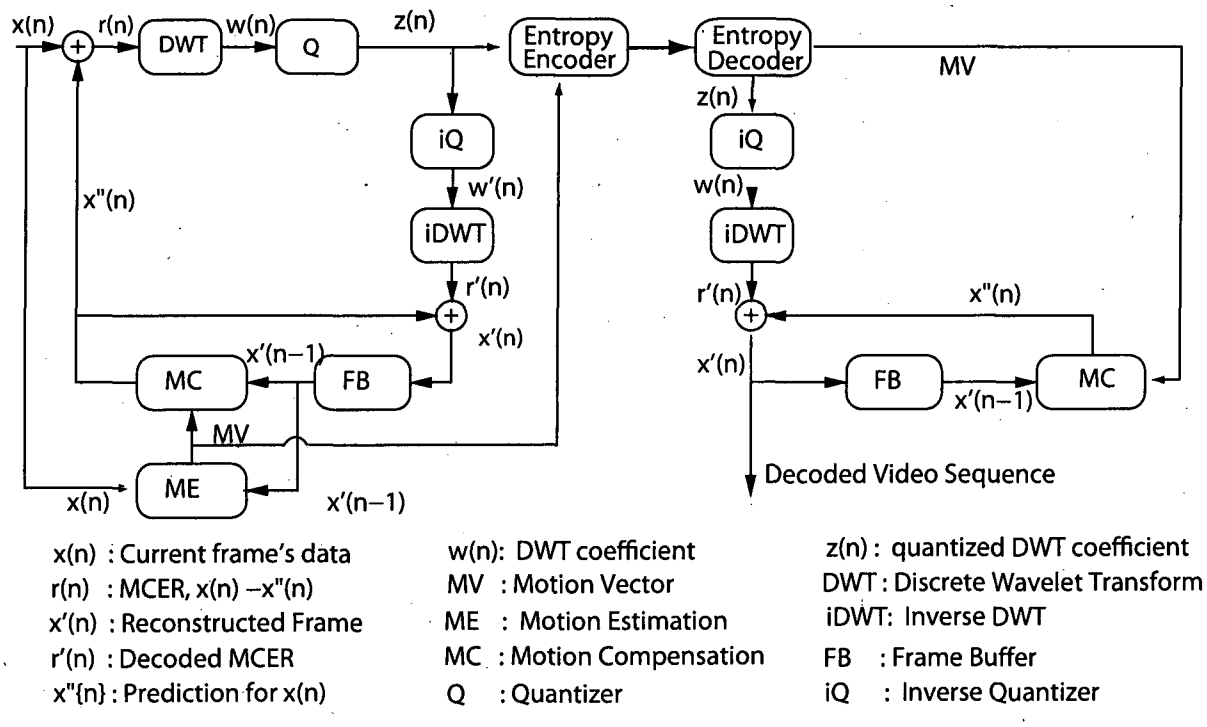


Figure 2.2: Block diagram of the Dirac video codec.

ii. The quantization of the wavelet transform's coefficients.

iii. Lossless entropy coding is applied to both the wavelet transform's coefficients and motion vectors.

iv. Motion Estimation (ME) involves finding matches for each block from previously reconstructed frames by dividing the frame into (MxN)-pixel blocks[7], trading off accuracy with motion vector bit rate.

v. Motion Compensation (MC) invokes the appropriately displaced corresponding block of the previously reconstructed frames to predict the current block, so that the number of bits required for encoding the Motion-Compensated Error Residual (MCER) is minimized.

As in most existing video compression standards, in the Dirac video codec each video frame is encoded in either intra-frame mode or in inter-frame mode. In the intra-frame mode, the original video frame is directly encoded with the aid of the wavelet transform.

[7]Depending on the position in the data frame, the block used for motion estimation may be a 10x10 pixel, 10x12 pixel, 12x10 pixel or 12x12 pixel block.

By contrast, in the inter-frame mode, the MCER is generated with reference to the previous decoded frame by using ME and MC.

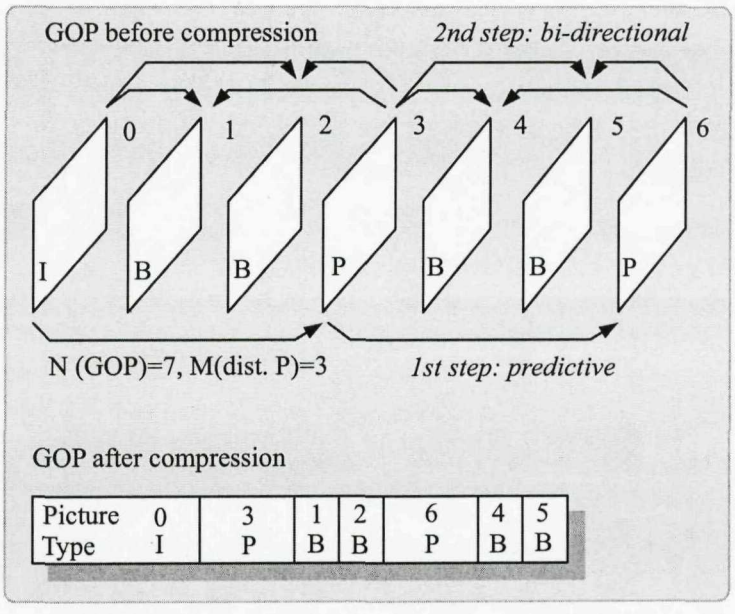


Figure 2.3: Temporal frame prediction structure.

Note that the Dirac video codec achieves a high compression ratio by using bidirectional ME/MC in almost all pictures (frames) of the video sequence. Since bidirectionally encoded pictures require near-future pictures for MC, reference pictures are periodically included, as exemplified by the Group Of Pictures (GOP) structure shown in Figure 2.3 and detailed below.

In the Dirac video codec, the GOP contains an I-frame, followed by 10 P-frames and 22 B-frames which will be detailed in Section 2.2.1. The I-frame is compressed as a completely independent frame, thus only the spatial redundancy is exploited for compression. For P- and B-frames, the temporal redundancy is also exploited, where P-Frames use a single temporal reference frame, namely the previous I- or P-frame. By contrast, B-frames use both previous and future reference frames, where the reference frames are the I-frame and P-frames. The top of Figure 2.3 shows the display order of the pictures in the video stream. Furthermore, the GOP structure seen at the bottom of the figure indicates the compression order of the frames in the GOP. The B-pictures are most compressed and hence are never used as a reference for other pictures. A GOP implicitly defines the processing order of the video frames. Since the N^{th}-frame refers to a future reference frame, it cannot be (en/de)coded before this reference frame has been received

and processed by the encoder/decoder. Therefore, the video frames are transmitted for example in the order 'IPBB' instead of their 'IBBP' *display order*.

Again, Figure 2.2 portrays the simplified block diagram of the Dirac codec. Observe that the structure of the decoding process is identical to that of the encoder's local decoder. Motion estimation and motion compensation is the most important process both in the video encoder and decoder in terms of achieving a high video compression. Motion estimation generates the motion vectors on the basis of identifying the most likely position within the previous video frame, where the current video block has originated from, as it moves along a certain motion trajectory in the consecutive video frames. This motion-estimation process involves allocating a certain search area in the previous frame and then sliding the current block over this search area in an effort to find the position of highest correlation. Once this position has been identified, the Motion-Compensated Error Residual (MCER) is formed by subtracting the two blocks from each other. Let us now investigate each individual block of the Dirac video codec seen in Figure 2.2 by discussing both the intra-frame coded mode and the inter-frame coded mode.

2.2.1 The Intra-frame Coded Mode

In the intra-frame coded mode, the frame in a video sequence is encoded without reference to any past or future frames. As seen in the left part of Figure 2.2, each of the Y, U, and V component of a frame is directly transformed using the 3-level wavelet decomposition [20, 87, 89] of Figure 2.1 in order to produce a set of Discrete Wavelet Transform (DWT) coefficients $W(n)$ for 10 subbands, as seen in Section 2.1. This decomposition uses the so-called fast lifting scheme [35,36] of Daubechies' 9/7 filters [34,91] detailed in [35,36]. The coefficients in the lowest frequency LL subband are treated differently from the coefficients in the remaining subbands. More specifically, the lowest LL subband's coefficients are encoded by using a differential component encoding method, employing the previously encoded neighbor coefficients as references. Both the differentially encoded LL subband's coefficients and the remaining subbands' coefficients are quantized by a quantizer having a dead-zone [53] in order to produce the quantized DWT coefficients $Z(n)$ and then the coefficients $Z(n)$ are entropy coded with the aid of Context-based Adaptive Binary Arithmetic Coding (CABAC) [54]. More details on the lowest LL subband prediction, dead-zone quantization, and adaptive arithmetic coding will be provided in Sections 2.3.2

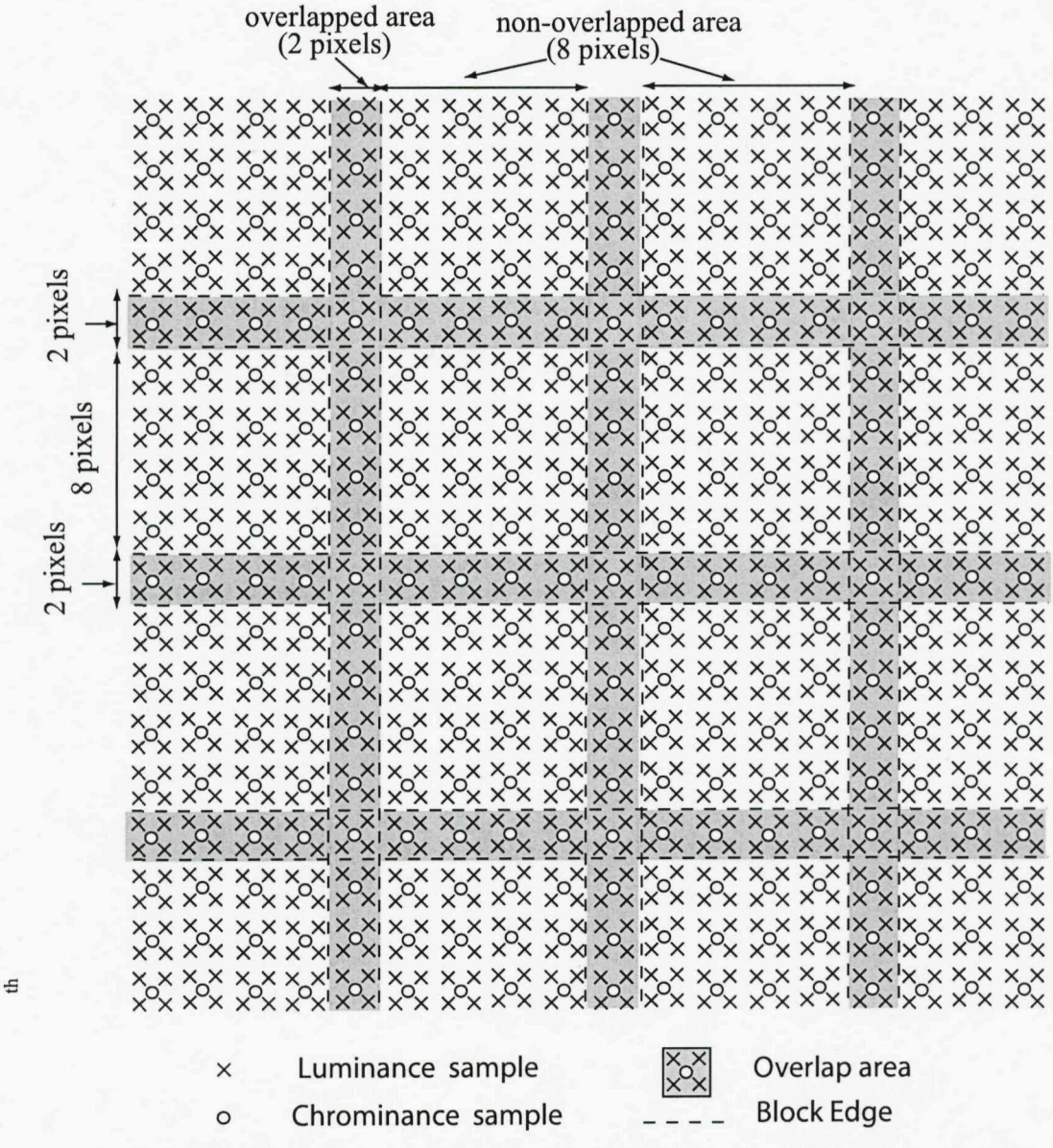


Figure 2.4: Positioning of luminance and chrominance samples in a macroblock containing sixteen luminance blocks in the YUV420 data format, in which the chrominance components (U, V) are sampled at quarter of the luminance rate on each scan line.

and 2.3.3.

2.2.2 The Inter-frame Coded Mode

In the **inter-frame mode**, the MCER, determined with respect to the previous decoded frames, is decoded using either forward Predicted (P) frame or Bi-directionally (B) predicted frames as seen in Figure 2.3. More specifically, the original video frame's Y, U and V component is divided into overlapped blocks. Sixteen adjacent 12x12-pixel luminance blocks and sixteen 6x6-pixel color difference blocks represented in the YUV420 format are then combined to form a so-called MacroBlock (MB), which has 52x44 luminance pixels and 26x22x2 chrominance pixels. In this mode, a predicted macroblock PM is formed with the aid of motion-compensation using one or more reference frame(s) $X'(n-1)$ on a MB basis. The Motion Vectors (MVs) are quantized, CABAC encoded as we will detail in Section 2.3.3 and transmitted to the receiver. The predicted macroblock PM (or $x'(n)$ in Figure 2.2) is then subtracted on a pixel-by-pixel basis from its motion-shifted best-matching counterpart in the reference frame for the sake of producing the Motion Compensated Error Residual (MCER) or difference $r(n)$. Finally, the MCER is Discrete Wavelet Transformed (DWT), quantized, and entropy encoded in the same way as in the intra-coded mode. Note that the reference frame must have already been encoded, reconstructed and stored in the reconstruction frame buffer.

In the video reconstruction path of Figure 2.2, after entropy decoding, the quantized coefficients $Z(n)$ are inverse quantized in the block iQ in order to produce the DWT's coefficients $W'(n)$ and inverse transformed in the block $iDWT$ for producing the reconstructed MCER macroblock $r'(n)$. Note that this is not identical to the original difference macroblock $r'(n)$ because the quantization process introduces granular effects and therefore $r'(n)$ is an approximate replica of $r(n)$. Finally, the predicted macroblock $x'(n)$ is then added to $r'(n)$ in Figure 2.2 for creating an approximate replica of the original macroblock, namely the reconstructed macroblock $x'(n)$.

2.3 Encoding within the Wavelet-Bands

Having described the top-level encoding and decoding processes, in this section we investigate the novel features of the Dirac video codec in more detail. Some examples of

the novel techniques first applied in this particular codec are constituted by the spatial prediction within the lowest LL subband, dead-zone based rate-distortion optimization aided quantizer [53], motion compensation using overlapped block, context-based adaptive binary arithmetic coding, etc.

2.3.1 Lowest LL Subband Prediction

Corresponding to the average luminance intensity of the Y, U or V component considered, the encoded lowest LL subband typically generates a high number of bits although it represents a small picture area owing to three-level decomposition, since no temporal and spatial redundancy removal is used as part of the encoding process regardless, whether I, P or B-frames are considered. In order to increase the achievable compression efficiency of the intra-frame coding process in the context of the Dirac video codec, the spatial correlation between the adjacent coefficients of the lowest LL subband[8] in the Intra-frame coded mode is exploited to decorrelate the related data.

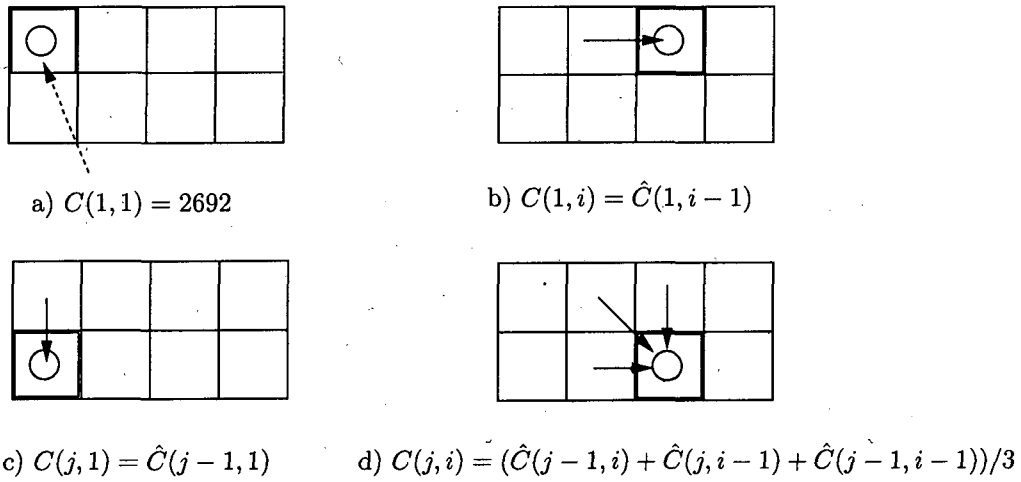


Figure 2.5: Coefficient prediction in the lowest LL subband.

As seen in Figure 2.5, in the lowest LL subband, the coefficients are scanned on a line by line basis. Typically the coefficient $c(i,j)$ and its neighbor coefficients are closely correlated. Hence we can create a predictor for the coefficient $c(i,j)$ at position (i,j) using any neighbouring encoded coefficients $c'(k,l)$, which are naturally located to the left and above the current coefficient $c(i,j)$. More explicitly, the coefficient $c(i,j)$ at position (i,j)

[8]The size of the lowest LL subband depends on the the resolution of the video clip. For example, when compressing (176x144)-pixel QCIF video clips, the lowest LL subband has the size of (22x18) pixels.

is predicted as follows [92]:

$$c(i,j) = \begin{cases} c & \text{if i=1, and j=1} \\ c'(i,j-1) & \text{if i=1, and } j > 1 \\ c'(i-1,j) & \text{if } i > 1, \text{ and j=1} \\ \dfrac{(c'(i-1,j-1)+c'(i,j-1)+c'(i-1,j)}{3} & \text{if } i > 1, \text{ and } j > 1 \end{cases} \tag{2.1}$$

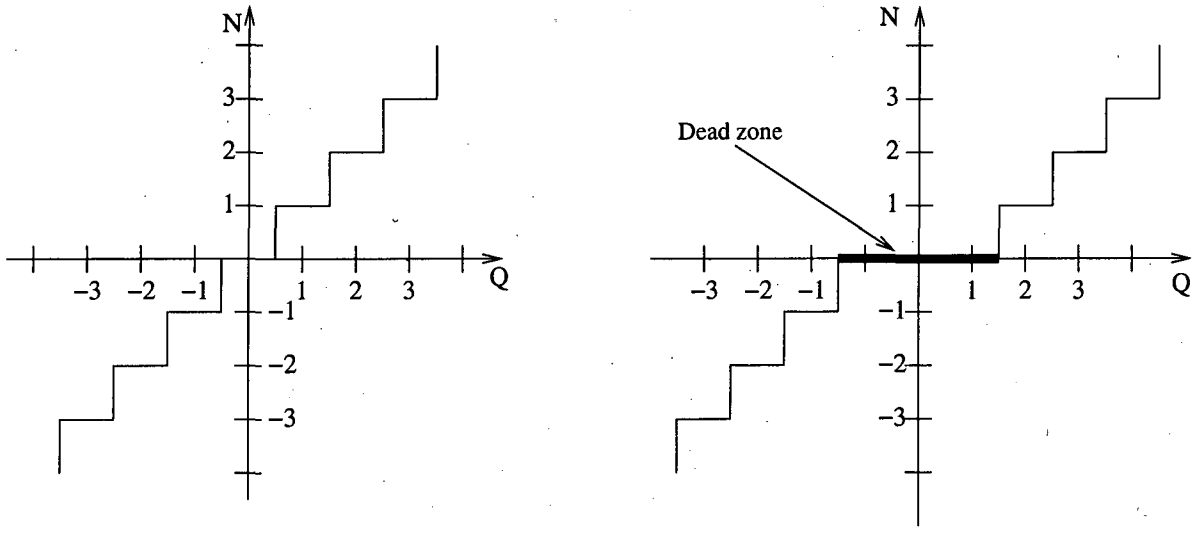
where the constant c represents the predicted value of the Direct Current (DC) coefficient of the lowest LL subband. It is worth noting that the choice of constant c does not substantially affect the bit-rate and the quality of the reconstructed frame. After prediction, the difference between the original value and its prediction is quantized and conveyed to the decoder. For the sake of reconstructing the corresponding DC value and predicting the next lowest LL subband coefficient, the predicted value must be added again to the quantized difference in order to produce the reconstructed coefficient $c'(k,l)$ of Equation 2.1. This process is applied to all the coefficients. Because prediction is inter-twined with quantization, it is also essential for coding the lowest LL subband.

2.3.2 Dead-Zone Based Quantization

Having transformed the YUV component data using the DWT, each subband's coefficients are quantized by using a uniform quantizer having dead-zone. The simple dead-zone based uniform quantizer is depicted in Figure 2.6.

The advantage of the dead-zone based quantizer is two-fold. Firstly, it sets the low input values, which typically represent camera-noise, to zero and hence reduces the bitrate. Secondly, it facilitates a simple implementation, where the input value is simply divided by the step-size and truncated to the closest smaller integer.

Typically, the transformed coefficients representing the output signals of the wavelet subbands have a distribution with a mean close to zero and exhibiting a fairly rapid decay for larger values. Hence these values are more likely to occur in the lower half of a bin, rather than in the upper half. Therefore, the reconstructed value at the output of the quantizer is not half-way between the decision boundary. Explicitly, instead of 0.5, a smaller value of 0.375 is added to the lower decision boundary to reflect this bias, which



a) Scalar Quantizer having a step–size of Q

a) Dead–zone Quantizer having a dead–zone of 3 step–sizes

Figure 2.6: Quantizers having a step-size of Q.

gives an improved performance in practice. As a result, in the Dirac video codec, the corresponding reconstructed value of $\tilde{v}$ is given by (an integer approximation to):

$$\tilde{v} = \begin{cases} 0 & \text{if } N = 0 \\ (N + 0.375) \cdot Q & \text{if } N > 0 \\ (N - 0.375) \cdot Q & \text{if } N < 0. \end{cases}$$

This reconstructed value is then used by the encoder in order to produce the locally decoded YUV component data, which is identical to what the remote decoder would produce in the absence of transmission errors after decoding the quantized value N.

Note that the Dirac video codec uses a total of 60 different quantizers with step sizes ranging from 1 to 27554. The list of quantizer indeces and step sizes of the corresponding quantizers is shown in Table 2.1. Based on the specific video quality and bitrate requirements, the encoder will choose a suitable quantizer for each subband in order to quantize its coefficients. The quantizer index and the quantized coefficients are then transmitted to the decoder in order to ensure that both the encoder and decoder use the same quantizer to quantize/inverse quantize the subband considered.

index	step size	index	step size	index	step size	index	step size
0	1	15	13	30	181	45	2435
1	1	16	16	31	215	46	2896
2	1	17	19	32	256	47	3444
3	1	18	22	33	304	48	4096
4	2	19	26	34	362	49	4870
5	2	20	32	35	430	50	5792
6	2	21	38	36	512	51	6888
7	3	22	45	37	608	52	8192
8	4	23	53	38	724	53	9741
9	4	24	64	39	861	54	11585
10	5	25	76	40	1024	55	13777
11	6	26	90	41	1217	56	16384
12	8	27	107	42	1448	57	19483
13	9	28	128	43	1722	58	23170
14	11	29	152	44	2048	59	27554

Table 2.1: The list of dead-zone based quantizers and their step sizes used by the Dirac video codec.

2.3.3 Context-Based Adaptive Binary Arithmetic Coding

Context-based Adaptive Binary Arithmetic Coding (CABAC) [54] represents an entropy coding [16,93] approach that exploits both the time-varying statistics of the diverse video parameters, such as the quantized WTC coefficients and the motion vectors generated by the video codec as well as the statistical dependencies between them for the sake of achieving a further bit-rate reduction. More specifically, in order to increase the coding efficiency of arithmetic coding, the underlying probability model is adapted to the time-variant statistics of the consecutive video frames, using a process referred to as context modelling. CABAC has three distinct features:

1. The context model provides estimates of the conditional probabilities of the source codec's symbols, such as the quantized DWT coefficients, motion vectors, etc.

2. The arithmetic code provides non-integer number of encoded bits/symbol, which allows CABAC to attain a potentially improved compression performance.

3. Adaptive arithmetic coding allows the entropy codec to adapt itself to non-stationary symbol statistics.

As a benefit of its advantageous properties, the CABAC scheme was adopted by the ITU-T/ISO/IEC video coding standard, namely the H.264/AVC scheme [4] as well as by the British Broadcasting Corporation's recent proprietary video codec referred to as the Dirac video codec [8]. The strategy for encoding the quantized wavelet coefficients and motion vectors follows the Dirac video codec's design philosophy initially presented in [94] and later refined in [54,88]. More explicitly, the CABAC process of the Dirac video codec is based on three different stages, namely the so-called binarization [54], context modelling [54,59] and adaptive arithmetic coding [54,93], as detailed in Figure 2.7.

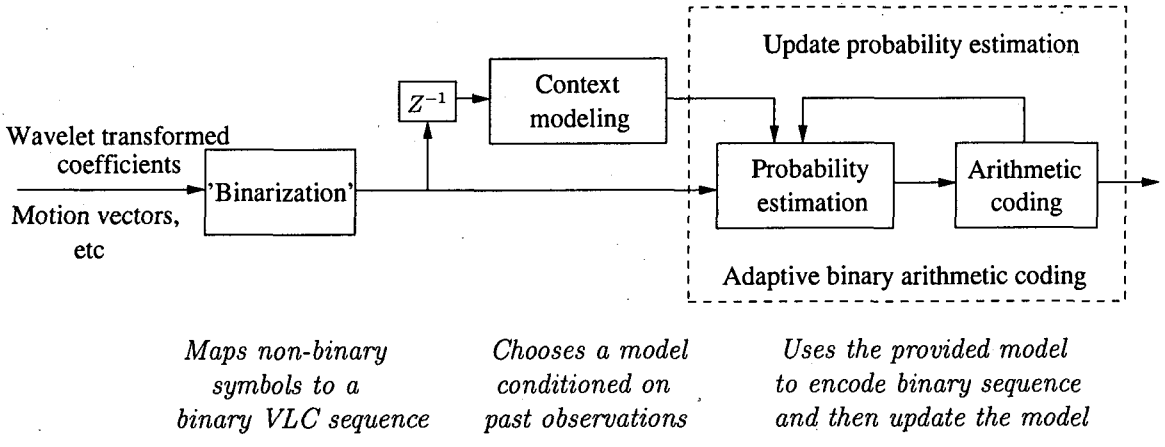


Figure 2.7: Block diagram of the context-based adaptive binary arithmetic coding used by the Dirac video codec.

2.3.4 Motion Estimation and Motion Compensation

The motion estimation/compensation scheme used in the Dirac video codec is similar to that in [95], which relies on a simple model of rectangular block displacements. More specifically, the Dirac video codec employs a hierarchical motion vector estimation [96], which proceeds from the higher level to the lower ones using a pyramid-like hierarchical search structure in order to reduce the computational complexity, as a benefit of having reduced image sizes at the higher levels, as seen in Figure 2.8.

The hierarchical images are constructed by appropriate subsampling[9] and up-sampling

[9]The subsampling process is carried out by applying low-pass filter having the following set of 12 filter coefficients
$(a_{-5}, a_{-4}, a_{-3}, a_{-2}, a_{-1}, a_0, a_1, a_2, a_3, a_4, a_5, a_6) = (0.015625, -0.015625, -0.03125, 0.015625, 0.1796875, 0.3359375, 0.3359375, 0.1796875, 0.015625, -0.03125, -0.015625, 0.015625)$.

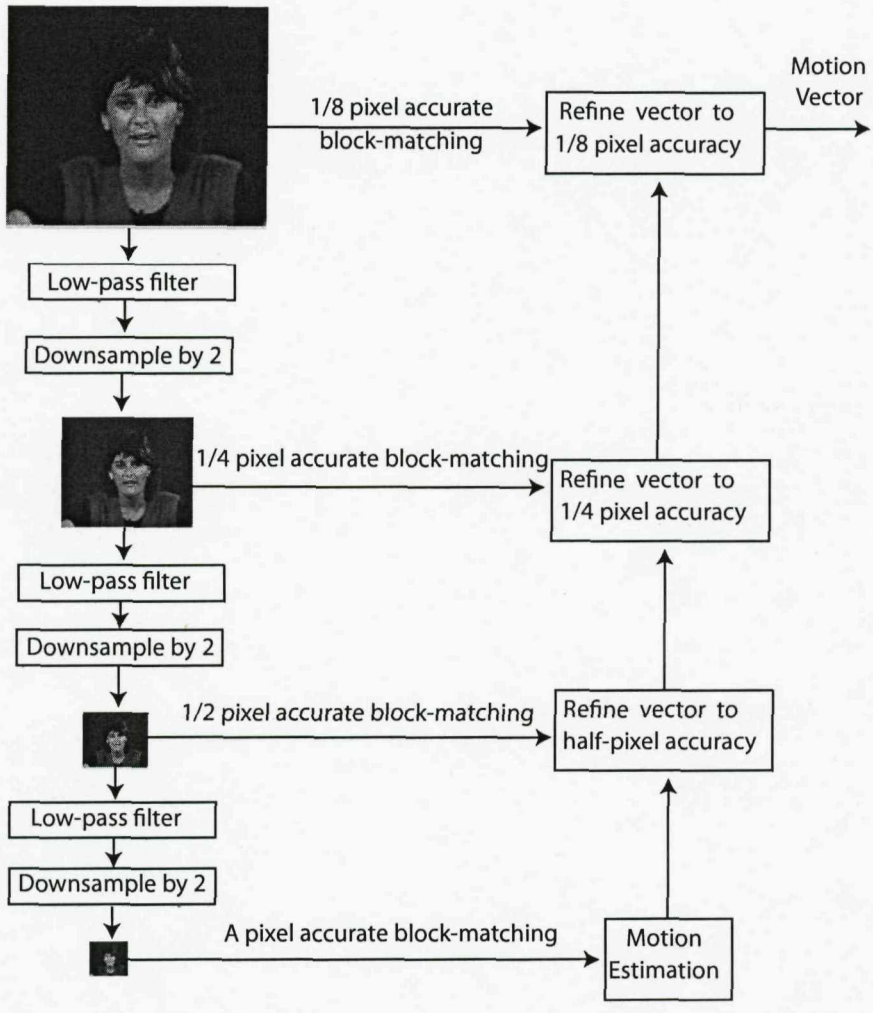


Figure 2.8: The hierarchical motion estimation structure used by the Dirac video codec.

[10] processes. In the Dirac codec, the sum of absolute differences (SAD) [97], which is defined for two blocks of samples X,Y as $SAD(X,Y) = \sum_{i,j} |X_{i,j} - Y_{i,j}|$, is used as the block matching metric. In order to find the best motion vector, the SAD calculations have to be carried out at nine surrounding points at each hierarchical level. More specifically, once the pixel-accuracy based motion vectors have been determined, each block will have an associated motion vector (V_0, W_0), where V_0 and W_0 are multiples of eight, which reflects that initial ME is carried out at a coarse 8-pixel-resolution. The double-resolution accuracy vectors are found by finding the best match in the displacement of set of (V_0, W_0) and its eight neighbours: $(V_0+4, W_0+4), (V_0, W_0+4), (V_0-4, W_0+4), (V_0+4, W_0), (V_0-$

[10]The up-sampling filter used in the Dirac video codec is determined by the following set of 12 filter coefficients
$(a_{-5}, a_{-4}, a_{-3}, a_{-2}, a_{-1}, a_0, a_1, a_2, a_3, a_4, a_5, a_6) = (-0.00390625, 0.015625, -0.04296875, 0.09765625, -0.21875, 0.65234375, 0.65234375, -0.21875, 0.09765625, -0.04296875, 0.015625, -0.00390625)$.

$4, W_0), (V_0+4, W_0-4), (V_0, W_0-4), (V_0-4, W_0-4)$. This in turn produces a new best MV (V_1, W_1), which provides a guide for the 4-pixel-resolution refinement, and so on. This process is illustrated in Figure 2.9.

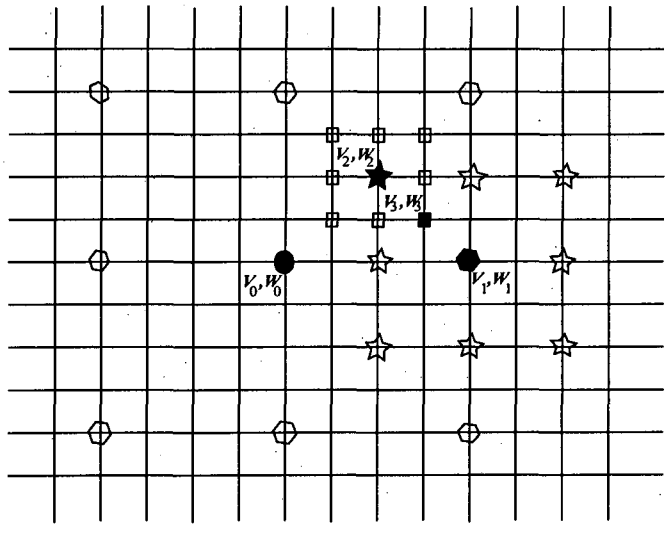


Figure 2.9: Motion vector refinement used by the Dirac.

Moreover, the adjacent blocks are expected to have similar MVs [98, 99] and this is reflected in the Dirac codec by imposing a certain degree of smoothening on the ME process. More specifically, the grade of smoothness in the MV field is determined by the difference between the candidate motion vector of a block and the median of the previously computed neighboring MVs. Since the blocks are estimated in a left-to-right, top-to-bottom raster-scan order, the MVs of the blocks to the left and above the current block are available for calculating the median of the neighbouring MVs. The MVs chosen for computing the local median value of the neighbouring MVs are $MV_{Top,Left}$, MV_{Top} and MV_{Left} in Figure 2.10.

The final metric used during ME is then computed as a weighted combination of the SAD and the median MV value. For a vector $V = (V_x, V_y)$, which maps the current block X to a block $Y = V(X)$ in the reference frame, the ME error E_{ME} is calculated as:

$$E_{ME} = SAD(X, Y) + \lambda \sum_{i \in neighbouring MVs} (|V_x - V_{i,x}| + |V_y - V_{i,y}|), \qquad (2.2)$$

where λ is a coding parameter used to control the trade-off between imposing smoothness on the motion vector field and the accuracy of the match. When λ is high, the neighbouring MVs depicted in Figure 2.10 dominate the final value of E_{ME} and hence the specific

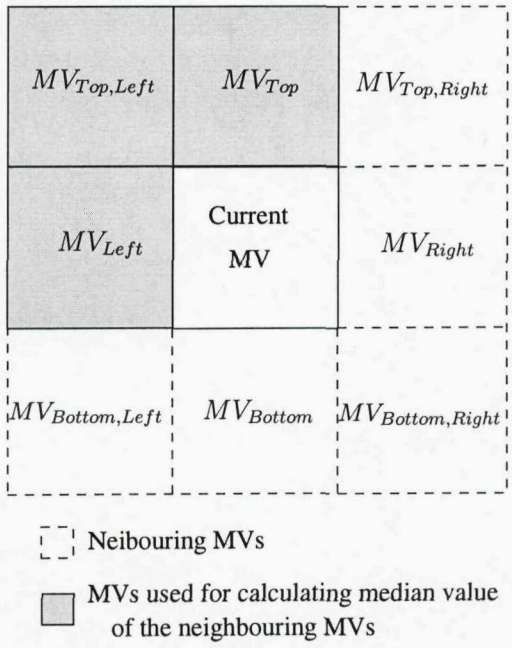


Figure 2.10: Neighboring MVs available in raster-scan order for calculating the median value of the neighbouring MVs.

motion vector, which gives the smallest metric is simply the one which is closest to its neighbours. By contrast, when λ is low, the value of the metric E_{ME} is dominated by the SAD term and hence the particular MV deemed to be the best will simply be the one that gives the best match for that block, regardless of its deviations from the surrounding MVs. For median λ values, varying degrees of MV field smoothness can be achieved.

2.3.5 Overlapped Block Motion Compensation

The motion prediction error of the luminance signal is constituted by the difference between the (12x12)-pixel overlapped blocks of the current frame and the corresponding overlapped blocks having MV-based displaced locations in the reference frame using the MVs that have been estimated in the first stage of motion estimation for the corresponding (8x8)-pixel core block. These (8x8)-pixel core blocks constituted the centre of the extended (12x12)-pixel overlapping blocks, which were windowed. The Dirac video codec employs the "raised-cosine" window function given by:

$$w(m,n) = w_m.w_n, \qquad w_n = \frac{1}{2}\left[1 - cos\frac{2\pi n}{N}\right] \tag{2.3}$$

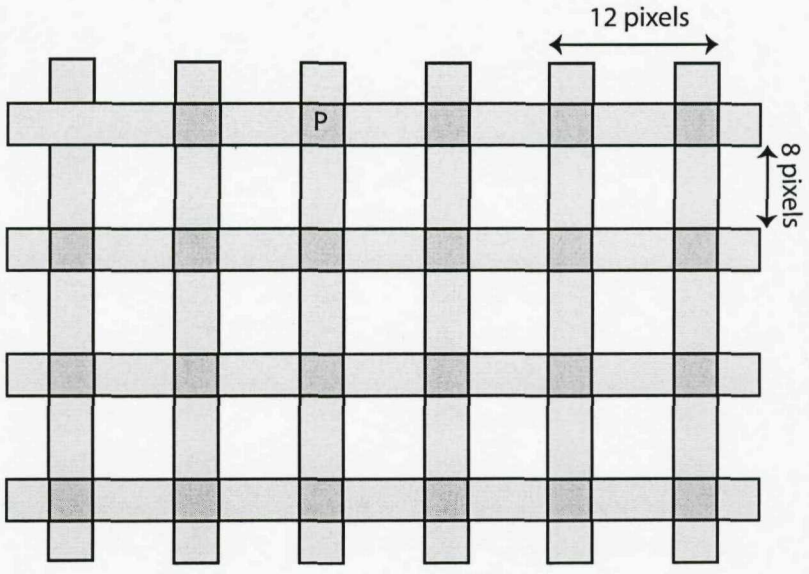


Figure 2.11: Overlapping (12x12)-pixel blocks.

for all $(m, n) \in [0, N]$x$[0, N]$ in order to avoid the frequency-domain Gibbs-oscillation encourtered in case of rectangular windowing, i.e in the absence of the above-mentioned tapered windowing. For QCIF video $N = 12$ was chosen in Equation 2.3 for the luminance component, resulting in (12x12)-pixel blocks centered over a "core" block of (8x8) pixels.

Let us now briefly consider the motion compensation algorithm used in the Dirac codec, which is based on a separable raised-cosine mask. As seen in Figure 2.11, given that a pixel $p = p(x, y, t)$ of frame index t may fall in a core block or in the overlapping area of up to four blocks if it lies at the corner of a block, each block containing the pixel p has a predicting block within the reference frame, as selected by the motion estimation process. The predicted value $\tilde{p}(x, y, t)$ of a pixel $p = p(x, y, t)$ is given by the weighted sum of all the corresponding pixels in the same position of up to four blocks used for prediction, where the corresponding blocks are identified by the motion vectors (V_i, W_i)[11]:

$$\tilde{p}(x, y, t) = \sum_i w_i p(x - V_i, y - W_i, t'), \qquad \sum_i w_i = 1. \tag{2.4}$$

[11]The sum of the weights will always be 1.

2.4 Effect of Coding Parameters on the Video Quality

2.4.1 Objective Video Quality

In this section, the objective video quality of the Dirac video codec is investigated. Quantifying the video quality is a challenging task, because numerous factors may affect the results. The most widely used measure is the Peak Signal to Noise Ratio (PSNR) [7], which is measured on a logarithmic scale expressed in dBs and depends on the normalized mean squared error (MSE) between the original and the reconstructed frames as well as on the maximum pixel value of the image or video frame. More specifically, when the image/video signal is represented using linear PCM having n bits per image/video sample, the highest possible pixel value is $(2^n - 1)$, yielding:

$$PSNR = 10 \log_{10} \frac{(2^n - 1)^2}{MSE}. \tag{2.5}$$

The PSNR may be conveniently calculated. However, its evaluation requires the knowledge of the unimpaired original image or video signal for comparison, which may not be available.

Unfortunately the PSNR defined above does not always constitute a reliable image quality measure. For example, if the received image is shifted by one pixel compared to the original image, the human eye would hardly notice any difference, while the PSNR objective measure would indicate a more substantial degradation in quality. Nonetheless, owing to its appealings implicity, in this report the PSNR will be used as the predominant image quality measure.

In our experiments the well-known video sequence *"Miss_America"* scanned at 30 frames/s was used. Our experiments were carried out using QCIF 420 YUV format color video sequences under the following conditions:

- An Intra-frame (I) coded picture was transmitted at the beginning of the sequence and then transmitted after each 33 frames.
- Two Bi-directionally predicted pictures (B) have been inserted between the Intra-frame coded picture and a Predicted picture (P), or between two successive P-

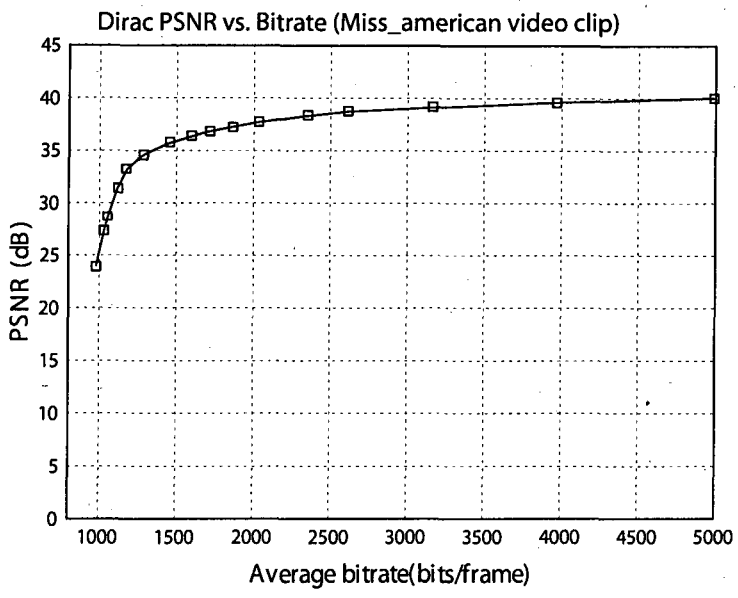


Figure 2.12: Image quality (PSNR) versus coded bit-rate performance of the Dirac codec for the *"Miss_America"* video sequence scanned at 30 frames/s.

pictures. More specifically, the frame-sequence had the following syntax: I,*B*,*B*,P,*B*,*B*,P.....

- The video sequence length of the *"Miss_America"* clip was 150 frames.

Our experiments were performed at a wide range of target bitrates. The results of these simulations are shown in Figures 2.12, which characterizes the average PSNR versus bit-rate performance. As seen in Figure 2.12, the image quality expressed on a logarithmic dB scale in terms of PSNR obeyed a logarithmic-like shape, as the bit rate increased. Naturally, on a linear MSE scale this would correspond to a near-linear relationship.

The number of bits required by each frame in the first 40 frames of (144x176)-pixel *"Miss_America"* video sequence corresponding to the average PSNR of 37.5dB, is illustrated in Figure 2.13, while the corresponding subjective video quality of frame 4 is demonstrated in Figure 2.14. As seen in Figure 2.13, the number of required bits per video frame depends on the encoding mode. The bit-rate contribution of I-frames was around 11.000 bits/frame, which is significantly higher than that of the P-frames averaging around 2.500 bits/frame, while that of the B-frames was about 1.800 bits/frame. This shows the effect of GOP structure on the achievable compression efficiency. More explicitly, the more frames we have in the GOP, the higher the codec's compression efficiency and vice versa.

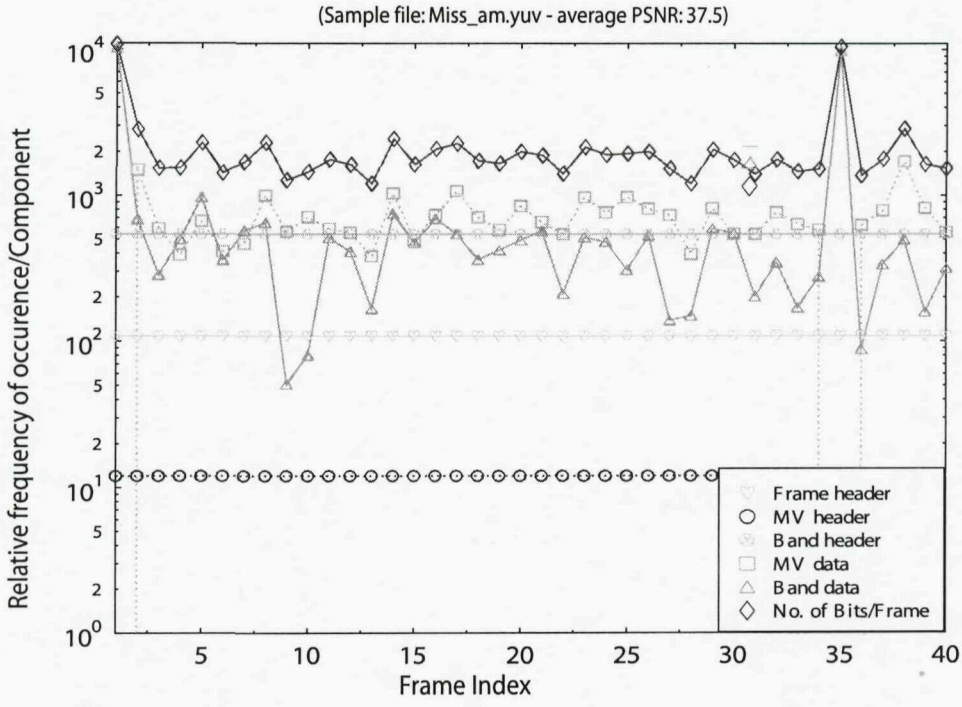


Figure 2.13: Number of bits used by each frame of the *"Miss_America"* video sequence in the first 40 frames.

Figure 2.14: Subjective video quality of frame 4 at 2279 bits/frame encoded by the Dirac video codec using the (144x176)-pixel QCIF *"Miss_America"* sequence.

2.4.2 Effects of Quantizer Stepsize

In this section we investigate the effect of the quantizer step size employed in the Dirac video codec on the attainable video quality. In video compression, the video parameter quantization is carried out by the quantizer in the encoder, while an 'inverse quantizer' is used in the decoder. A critical parameter is quantizer step size. If the step size is large, typically a low-rate compressed bit stream is generated. However, the reconstructed values provide a crude, low-quality representation of the original signal. By contrast, if the step size is small, the reconstructed values match the original signal more closely at the cost of a reduced compression efficiency. The effect of different quantizer step sizes is demonstrated in Figure 2.15 in terms of the achievable video PSNR, while the associated subjective video quality is illustrated in Figure 2.16.

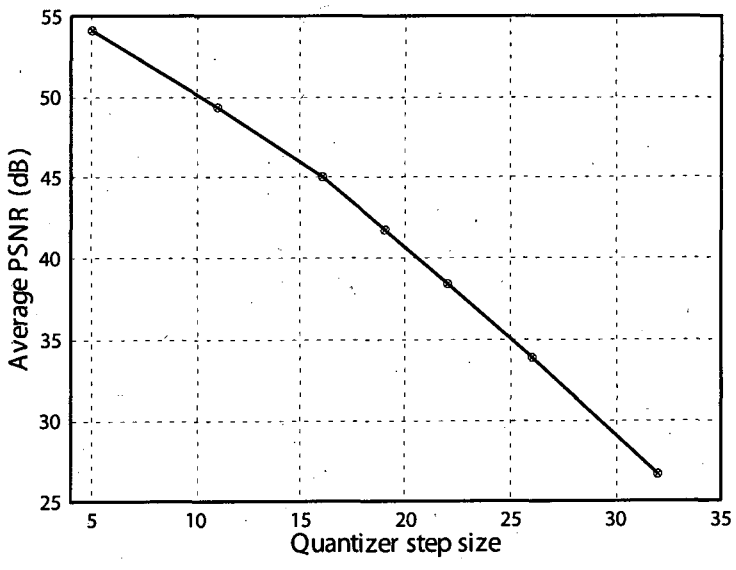
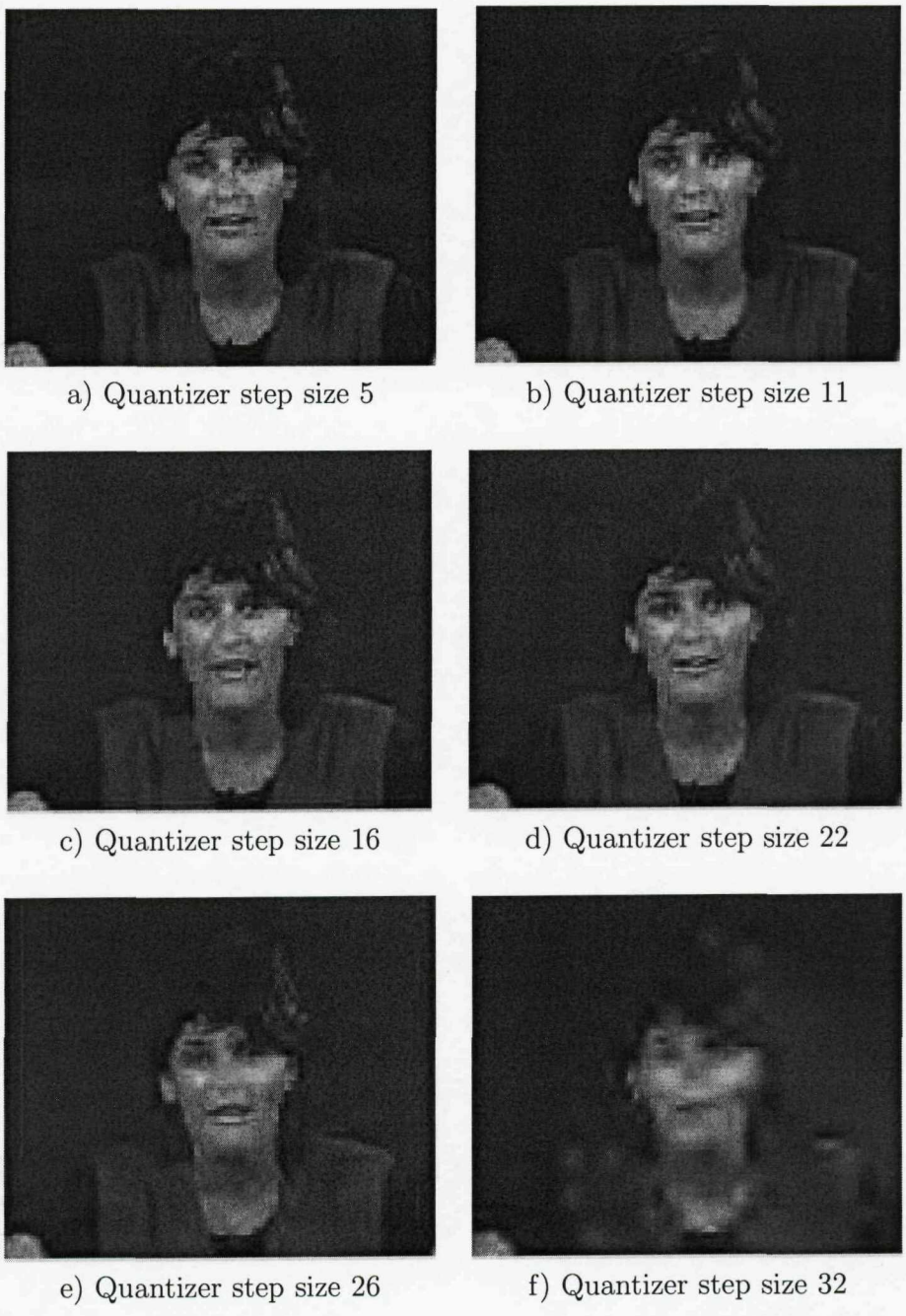


Figure 2.15: The average video quality (PSNR) versus quantizer step size ranging from 5 to 31 using the (144x176)-pixel QCIF *"Miss_America"* sequence encoded by the Dirac video codec.

Let us now briefly consider the effect of the quantizer step size on video bit-rate. The results of Figure 2.15 show that if the step size is high, typically a low bitrate is generated. However, the reconstructed values provide a crude representation of the original video signal. If the step size is small, the reconstructed values match the original signal more closely at the cost of a reduced compression efficiency. In other words, the larger the step size, the lower the video quality. The Dirac video codec invokes Rate-Distortion Optimization (RDO) [100, 101] by selecting the most appropriate one from 60 different quantizer step sizes for quantizing each subband in a video frame.



a) Quantizer step size 5

b) Quantizer step size 11

c) Quantizer step size 16

d) Quantizer step size 22

e) Quantizer step size 26

f) Quantizer step size 32

Figure 2.16: The effect of quantization step size on the video quality in the 1^{st} I-frame of the (144x176)-pixel QCIF *"Miss_America"* sequence.

2.5 Chapter Conclusions

In this chapter we investigated an efficient wavelet-based video codec, namely the Dirac codec, proposed by the BBC's Research and Development Department. This source codec is capable of efficiently compressing video and yet achieving a remarkably high reconstructed video quality at a low bit-rate. Hence, the Dirac video codec can be employed for the sake of providing high quality video services as in current television broadcast.

Commencing with the brief overview of the wavelet decomposition, the overview of the Dirac video codec, shown in Figure 2.2, was provided in Section 2.2. Moreover, a range of efficient techniques invoked in the Dirac video codec, such as for example the spatial prediction within the lowest LL subband, dead-zone based rate-distortion optimization aided quantizer [53], motion compensation using overlapped block, context-based adaptive binary arithmetic coding [54], is presented in Section 2.3.

In Section 2.4, the subjective video quality performance of the Dirac video codec was investigated. The experiment results of Figures 2.13 and 2.12 illustrated the effect of GOP structure on compression efficiency: the more frame in the GOP, the more compression efficiency the codec gets, and vice versa. Moreover, the relative frequency of occurrence for the various Dirac video codec parameters shown in Figure 2.13 illustrated that in inter-frame coded mode the bitrate contribution of the motion vectors becomes quite dominant, up to half of the total frame size of 2000 bits. Additionally, the lowest LL subband data is the most significant frequency parameter in the intra-frame coded mode.

Furthermore, the effective of different quantizer step size on video bit-rate depicted in Figure 2.15 demonstrated that the larger the step size invoked (corresponding to the lower bitrate), the lower video quality is. This explains the employment of 60 different quantizer step sizes for quantizing each subband in video frame. In the light of these results, the various video transmission scheme for the Dirac's bitstream will be deeply studied in the next chapters.

Chapter 3

Unequal Protection Aided Wavelet Compressed Video Telephony

3.1 Introduction

As mentioned in Chapter 2, the Dirac video codec is capable of achieving a high reconstructed video quality at a low bit-rate. It was developed by the BBC as an open-source design alternative to the Motion Picture's Expert Group's MPEG4 codec [2], which uses quite different DCT-based design principles. As most video codecs, the Dirac video codec invokes motion compensation for removing the temporal redundancy in the successive video frames. Furthermore, in order to achieve a high compression efficiency, Context-based Adaptive Binary Arithmetic Coding (CABAC) [54] is employed for compressing both the motion vectors as well as the subbands' coefficients. When aiming for low-bitrate operation, the bitrate contribution of the header information becomes quite dominant, for example about 30% of the total frame size of 2000 bits. Naturally, this is the most error-sensitive segment of the encoded video frame, since no useful information can be decoded in its absence. Additionally, the employment of CABAC and predictive coding renders the compressed video sequence sensitive to channel errors, since CABAC potentially propagates errors to the succesive motion vectors/subband coefficients, while the predictive coding inflicts error propagation upon the future frames.

As a novel contribution, the aim of our bit-sensitivity study is to quantify the average PSNR degradation caused by each erroneously decoded Dirac-video codec parameter

in the sequence, so that appropriate channel coding can be assigned to each parameter. In [7], the authors showed that for a video-coded sequence having non-uniform error sensitivity, the effect of a channel error may significantly vary from one bit to another. Furthermore, as a main contribution of this chapter we proposed an Unequal Error Protection (UEP) scheme designed for the Dirac video codec by using Irregular Convolutional Codes (IRCCs) which are constituted by a family of different-rate subcodes. Our UEP scheme benefits from the high design flexibility of IRCCs, hence excellent convergence properties can be achieved, while having unequal error protection capabilities matched to the specific requirements of the Dirac video codec. More specifically, the constituent codes of the IRCC are decoded jointly and iteratively by exchanging extrinsic information with the inner code, and the IRCC assigns a lower coding rate to highly sensitive bits and a higher coding rate to less sensitive bits. Our simulation results demonstrate that irregular convolutional codes exhibit excellent convergence properties in the context of iterative decoding, whilst having an unequal error protection capability, which is exploited in this chapter to protect the different-sensitivity video bits of the Dirac video codec. As a benefit, the UEP using IRCCs exhibits an iteration gain of about 1dB over the Equal Error Protection (EEP) benchmarker employing regular convolutional codes, when communicating over a 5-path dispersive AWGN channel.

The remainder of this chapter is organized as follows. Section 3.2 investigates the Dirac video-coded sequence, highlighting its hierarchical encoded sequence structure. Based on this structure, the relative frequency of occurrence for the various Dirac parameters is also investigated as the basis of our error resilience study in Section 3.4.

Section 3.3 provides the Dirac video codec's bit-sensitivity study. It is worth noting that there are a number of different techniques used in the literature for quantifying the bit error sensitivity of video codecs, and the outcome of the investigations to be conducted will highly depend on a number of factors, such as the video source material used and the target bitrate of the video codec [7]. Here, we consider a simplified objective video quality measure based bit-sensitivity evaluation, which attempts to examine the effects of corrupting the video parameters on the video quality degradation inflicted. Moreover, the associated error propagation effects are also investigated for further understanding the bit error sensitivity of the Dirac video codec.

In Section 3.4, a UEP scheme using turbo-equalized Irregular Convolutional Codes (IRCCs) is designed for the transmission of the Dirac video-coded sequence. Rather than

being decoded separately, the constituent codes of the IRCC are decoded jointly and iteratively by exchanging extrinsic information with the rate-1 pre-coder as our inner code. The IRCC is optimized to match the characteristics of both the Dirac codec and those of the channel, so that the UEP scheme employed maximizes the attainable iteration gain and hence the overall system performance.

3.2 Dirac Video-Coded Sequence

3.2.1 Structure of the Dirac Video-Coded Sequence

The Dirac video encoded sequence is constituted by the concatenation of the sequence header and the frame data. The sequence header conveys the default decoding parameters of the entire video sequence and it has to be fixed during the entire video session. By contrast, the frame data, including the picture header, are constructed depending on a number of parameters, such as for example whether the intra-frame or inter-frame coding mode is used and on other diverse video properties. Since the Dirac video-encoded sequence has a complex structure, it is rather difficult to visualize the corresponding sequence in a well-presented, highly organized manner. In order to simplify this task, we partition the Dirac video-encoded sequence into several categories, as seen in Figure 3.1. Let us now investigate each category seen in Figure 3.1.

◊ *SEQUENCE HEADER*

The sequence header of Figure 3.1 consists of the default decoding parameters of a video-encoded sequence. For example, the video format is chosen from the following set defined in Section 2.2: YUV411, YUV420, YUV422 or YUV444; the video resolution is the 176x144 pixels Quarter Common Intermediate Format (QCIF), the 352x288 pixels Common Intermediate Format (CIF) or the 1920x1080 pixels High-Definition Television (HDTV) format; the video frame structure is based on macroblocks of either 8x8-pixel blocks non-overlapped or 12x12-pixel overlapped blocks. The error-free reception of these parameters enables the Dirac decoder to ensure that it uses the same reconstructed video buffer contents as the encoder's local decoder during the decoding process. Additionally, the so-callled *codec identifier* and the *video frame rate* are also

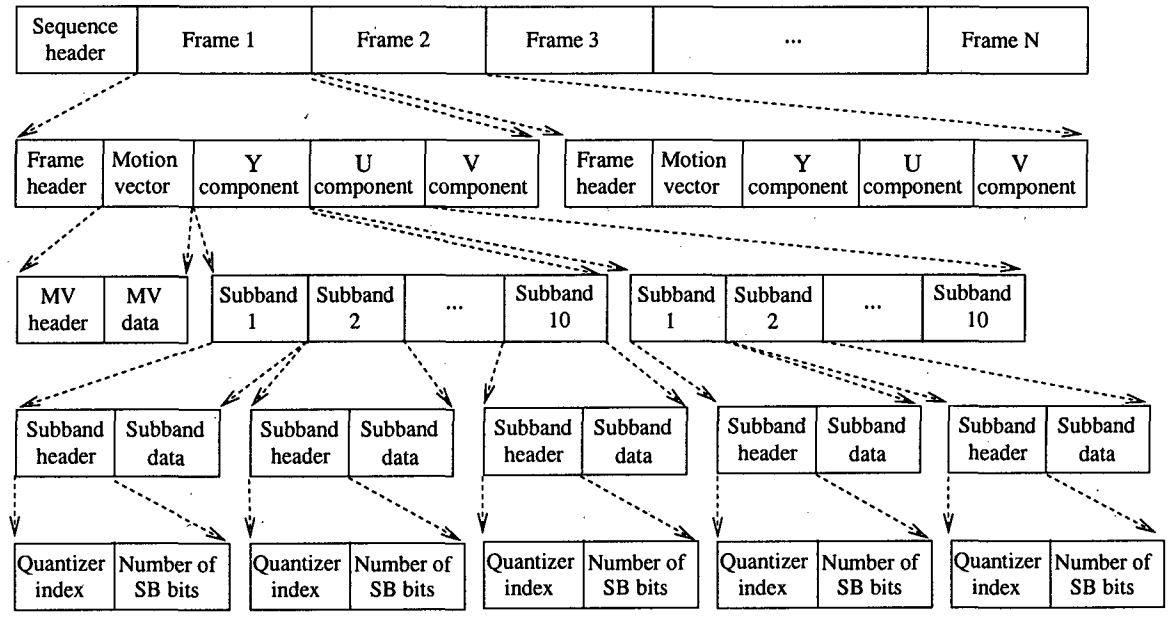


Figure 3.1: Hierarchical structure of the Dirac video sequence.

contained in the sequence header for the sake of signaling to the video decoder, how the video sequence was encoded and how to display the decoded video sequence.

◊ *FRAME DATA*

As mentioned above, the structure of the so-callled frame data units seen in Figure 3.1 varies, conveying information, such as the intra-frame or inter-frame coding mode and a range of other picture properties. The frame data unit contains the decoding parameters of a certain encoded frame, such as the *frame header*, the *motion vectors* and 30 subbands' DWT coefficients for all the *picture's* Y, U, V components.

More specifically, starting with the 32 bits describing the *current frame index*, the frame data informs the decoder of the starting point of new frame and the temporal order of the specific frames considered in a video sequence. Following this information, a number of other coding control parameters, such as a specific flag indicating *whether* or *not* to *skip* decompressing a particular frame, the *expiry time* of this frame in the reconstruction buffer, the *encoded mode*[1], the number of reference frames and the corresponding reference frame indices are presented.

[1]Intra, L1 or L2

◊ *MOTION VECTOR*

As seen in Figure 3.1, the motion vector parameters are divided into two parts: *motion vector header* and *motion vector data*. Here the motion vector header represents the number of bits used to encode all the motion vectors of the frame considered, which is employed for signaling the end of an encoded stream to the CABAC decoder described in Section 2.3.3. By contrast, the motion vector data provides the input data for the CABAC decoder. Noting that length of the motion-compensated data varies on a frame-by-frame basis, depending both on its encoding mode and on the temporal correlation between the current video frame and the reference frames.

◊ *SUBBAND DATA*

The frame's Y, U and V component data is given by compressing the DWT coefficients of the original video frame in the intra-frame mode or those of the motion compensated prediction error in the inter-frame mode, as discussed in Section 2.2. The DWT coefficients of the 10 subbands are generated by wavelet transforming the Y, U or V component, which are then separately quantized and encoded using the CABAC encoder described in Section 2.3.3. The quantizer index and the number of bits used to encode all the DWT coefficients of the subband considered is signalled to the decoder in the subband header. The Dirac video decoder uses the header information to reset the adaptive arithmetic decoding state as and when there is a VLC termination problem. On the other hand, the actual subband's data is used for decoding the quantized DWT coefficients. Observe in Figure 3.1 that the index of the specific quantizer out of the 60 legitimate schemes plus the total number of bits representing a particular subband is explicitly signalled to the receiver.

3.2.2 Simulated Coding Statistics

In this section, video-coding simulations were conducted in order to quantify the relative frequency of occurrence for the various coding parameters in the Dirac video sequence. More specifically, Figure 3.2 shows the probability of occurrence for the various Dirac parameters in the intra-frame coded mode, while Figure 3.3 represents that in the inter-frame coded mode. These figures were derived using the QCIF-resolution *"Miss_America"*

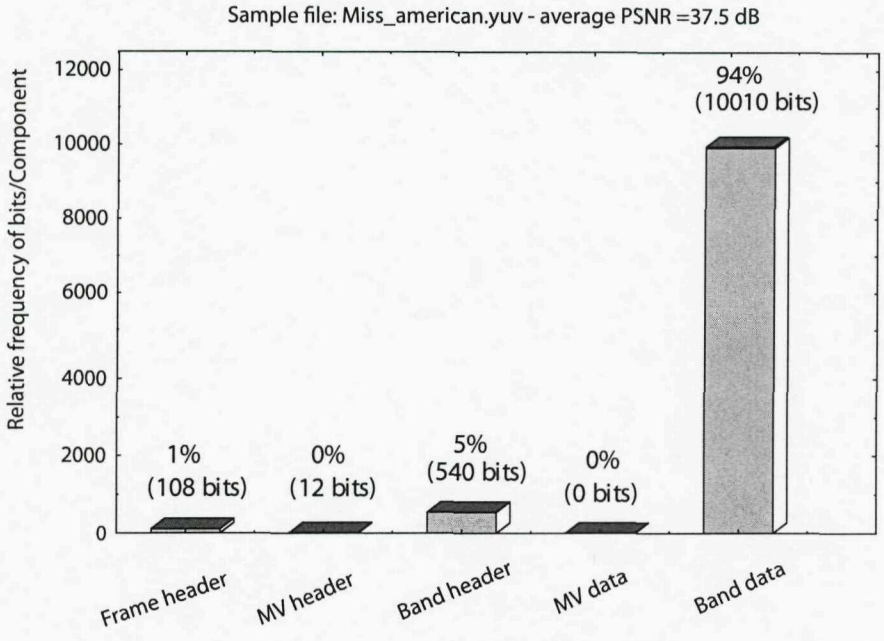


Figure 3.2: The relative frequency of occurrence for the various Dirac parameters in the intra-frame coded mode for the *"Miss_America"* QCIF video sequence encoded at 30 frame/s at a bitrate of 60kbit/s.

video clip.

Let us now discuss the results seen in Figure 3.2, which presents the relative frequency of occurrence for the various Dirac parameters in the intra-frame coded mode. It can be seen that the majority of - namely 94% - the intra-coded frame's bits represent the subband data, while all the other parameters are represented by the remaining 6% of the bits. Naturally, the actual specific proportion of these fractions becomes different in the different frames of a video stream. In our experiment characterized in Figure 3.2, the average PSNR was 37.5 dB for the "Miss_America" QCIF video sequence, which requires about 10,000 bits for encoding all the subband data of the (176x144)-pixel QCIF frame.

In the inter-frame coded mode, the motion compensation removes much of the temporal redundancy in the video frame, hence the bitrate of the inter-frame coded MCER was significantly reduced. Observe from Figures 3.3 and 2.13 of Chapter 2 that the average bitrate contribution of an inter-frame coded picture is 1801 bits, which is significantly lower than that of the intra-frame coded pictures, which was found to be 10670 bits. Figure 2.13 also shows that the bitrate contribution of the bi-directionally predicted frames is lower than that of the forward predicted frames. More details on these parameters can be found in Section 2.4.

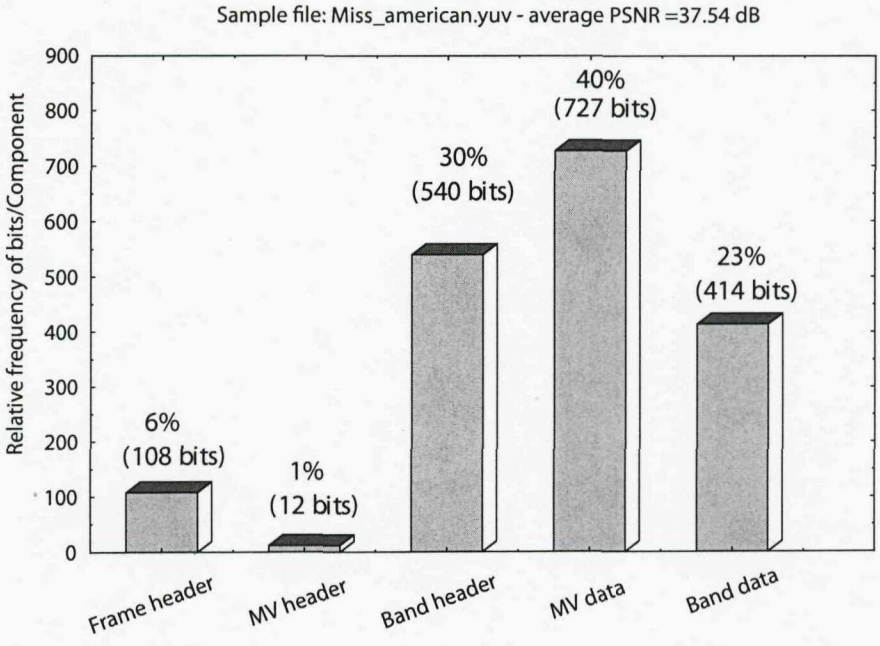


Figure 3.3: The relative frequency of occurrence for the various Dirac parameters in the inter-frame coded mode for the *"Miss_America"* QCIF video sequence encoded at 30 frame/s at a bitrate of 60kbit/s.

3.3 Bit Sensitivity Study

In order to provide robust source-matched error protection for the wavelet coding-based video stream, it was subjected to bit sensitivity investigations by systematically corrupting all of its bits in particular parameter of all consecutive video frames in the encoded video frame while keeping the rest of the parameter un-contaminated and then evaluating the average PSNR degradation inflicted. The simulations were carried out using a 150-frame duration (144x176)-pixel QCIF-resolution *"Miss_America"* video clip. The PSNR degradations were recorded for each of the parameters under the following conditions:

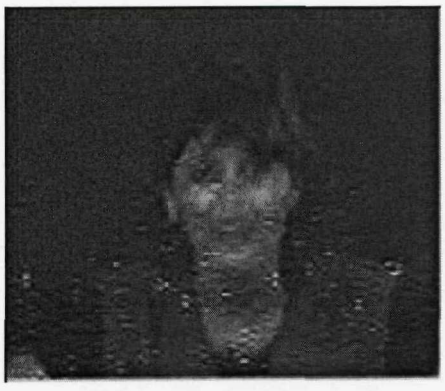
1. The effects of bit corruption on the Dirac video header parameters were somewhat difficult to estimate, because these parameters contain vital frame-synchronization information related to both the entire video sequence and to a particular video frame. The decoder cannot even commence its decoding operations, when the frame-synchronization parameters are corrupted.

2. When for example the sensitivity of bit 1 of an intra-frame coded picture was investigated, this bit was consistently corrupted in each occurrence of an intra-frame coded picture, while keeping all other bits of the same frame intact.

a) Corrupted first bit of the LL subband

b) Corrupted 1500^{th} bit of the LL subband

c) Corrupted first bit of the 9^{th} subband



d) Corrupted first bit of the 8^{th} subband

e) Corrupted first bit of the 7^{th} subband

f) Original frame

Figure 3.4: Subjective image quality degradation due to the corruption of the Y component in the 1^{st} frame of the *"Miss_America"* QCIF video sequence encoded at 30 frame/s using the intra-frame mode.

More details on the bit allocation used within a frame are provided in Table 3.1, where Y, U, V represent the luminance and the two colour difference components of the video frame. An overview of the objective Peak Signal to Noise Ratio (PSNR) degradations

	The number of subband bits		
	Y component	U component	V component
DC subband	1888	116	502
Subband 9	466	28	109
Subband 8	1049	145	230
Subband 7	1208	110	273
Subband 6	490	0	65
Subband 5	1038	60	104
Subband 4	1504	0	334
Subband 3	0	0	0
Subband 2	215	0	0
Subband 1	302	0	0

Table 3.1: The number of subband bits in the first frame of the *"Miss_America"* QCIF video sequence using the intra-frame coding mode at a bitrate of 60kbit/s and 30 frame/s.

inflicted by the systematic bit corruption events upon the subband DWT coefficients encoded using the intra-frame coding mode is given in Figure 3.5, while Figure 3.4 depicts the subjective video quality degradation due to the corruption of the Y component for the *"Miss_America"* QCIF video sequence. Observe in this figure that

1. The bit sensitivity is typically higher at the beginning of a subband and then reduces towards the end of the subband, which is because an error at the beginning of a CABAC segment corrupts more of the remaining bits of the subband than one near the end of it.

2. The subband's bit sensitivity also depends on the particular subband's spatial frequency. More specifically, at the same bit index of a low-frequency subband, the bit sensitivity is typically higher than that of a high-frequency subband and vice versa.

3. The number of DWT coefficients in a subband is an other factor influencing the sensitivity of a subband, since the so-called context-based CABAC continuously updates its parameters based on the past received symbols for the sake of decoding the remaining part of the subband concerned. Consequently, the rest of the entire subband data is affected by a transmission error inflicted upon one of the first bits of that subband.

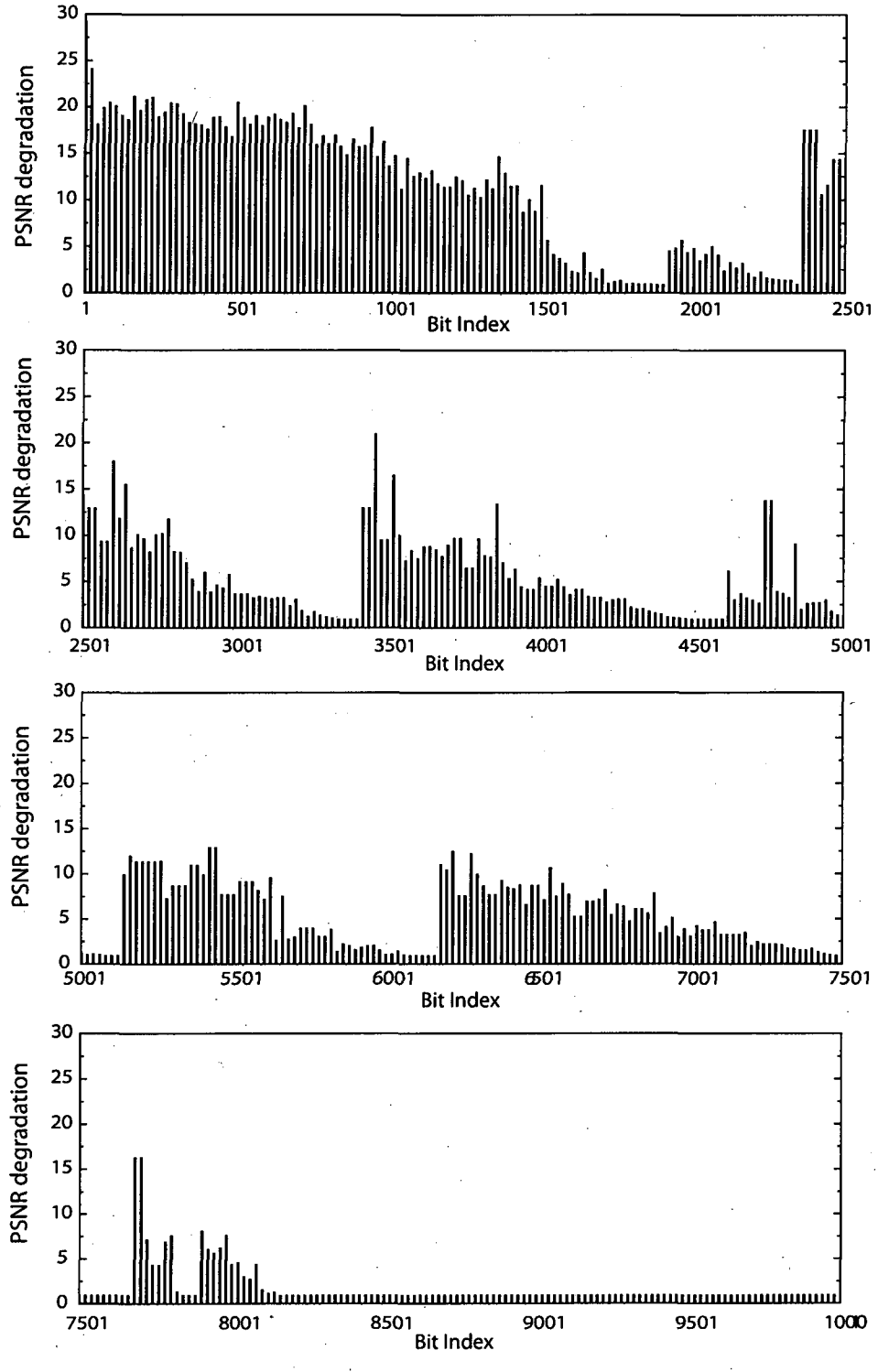


Figure 3.5: Average PSNR degradation due to the bit corruption events for the *"Miss_America"* QCIF video sequence encoded at 30 frame/s using intra-frame mode, where the frame-structure of Figure 3.1 and the bit allocation used within a frame provided in Table 3.1 were followed.

a) Corrupted first bit of motion vector

b) Corrupted 700^{th} bit of motion vector

c) Corrupted 1400^{th} bit of motion vector

d) Corrupted first bit of the LL subband

Figure 3.6: Subjective image quality degradation due to various bit corruption events in the 2^{nd} frame of the *"Miss_America"* QCIF video sequence encoded at 30 frame/s and 60kbit/s using the inter-frame mode.

For the sake of investigating the effects of bit-corruption in the inter-coded frames, our simulations were conducted for the following situations:

- The first intra-coded frame (frame 1/150) is an uncorrupted frame.
- The number of bits required for the different parameters in the second frame of the *"Miss_America"* QCIF video sequence using the inter-frame coding mode at 30 frame/s and 60kbit/s is shown in Table 3.2.
- Errors are inflicted upon the first inter-coded frame (frame 2/150) as detailed below.
- A corrupted motion vector parameter is encountered in the first inter-coded frame (frame 2/150), more specifically on the 1^{st}, 700^{th} and 1400^{th} bit in the motion vector

	The number of bits		
	Y component	U component	V component
Subband 10 (DC subband)	73	38	0
Subband 9	0	45	0
Subband 8	0	0	0
Subband 7	0	0	23
Subband 6	87	34	0
Subband 5	63	40	0
Subband 4	0	40	0
Subband 3	0	0	0
Subband 2	77	164	0
Subband 1	0	0	0
Motion vector	1497		

Table 3.2: The number of bits required for different parameters in the second frame of the *"Miss_America"* QCIF video sequence using the inter-frame coded mode at 30 frame/s and 60kbit/s.

data. This may lead to displacing a block differently from what was used by the local decoder of the encoder and hence propagating errors to all remaining motion vectors.

- Corrupting the 1^{st} bit of the lowest LL subband data in the first inter-coded frame (frame 2/150), leading to error propagation for the rest of the lowest LL subband's DWT coefficients.

- Corrupting the 1^{st} bit of the 5^{th} subband in the first inter-coded frame (frame 2/150), leading to error propagation for the rest of the 5^{th} subband's DWT coefficients.

As seen in Figure 3.6, the subjective quality degradation imposed by the corruption of the motion data parameters is typically high, while that of the subband DWT coefficients is less dramatic. In order to provide further insights, the bit allocation used within the first inter-frame coded picture is given in Table 3.1. The corresponding objective PSNR degradations inflicted by the systematic bit corruption events in inter-coded frames is depicted in Figures 3.7 and 3.8 at 30 frame/s and 60kbit/s. More specifically, the bit indices spanning from the 1^{st} to 1497^{th} represent the case when the encoded sequence of motion vectors in the inter-coded frame was corrupted, while the bit indices ranging from the 1498^{st} to 1570^{th} represent the case, when the encoded sequence of the lowest LL subband in the inter-coded frame was corrupted. These bit indices can be located with the aid of Table 3.2 for the second inter-coded frame. Again, as seen in these figures, the

bit sensitivity is typically higher at the beginning and then reduces towards the end of each parameter. Moreover, the most sensitive parameter in the inter-frame coded pictures is the motion vector, whose bitrate contribution is significantly higher than that of the 30 subbands.

Observe from Figures 3.5, 3.7 and 3.8 that the sensitivity of approximately the last 500 bits in the motion vector field remains moderate (the PSNR degradation is usually confined to the range below 8dB). This characteristic behavior can also be observed for the subband data in both the intra-frame and inter-frame coded mode.

During the inter-frame mode's bit error sensitivity study, investigations were conducted using the second frame of the *"Miss_America"* video sequence. The first frame, corrupted at the first bit position of the Y component in the lowest LL subband is depicted on the left of Figure 3.9, while the corresponding frame forward predicted by the motion-compensation scheme is shown at the right of Figure 3.9.

The results seen in Figures 3.10 and 3.11 are also interesting to analyze. Since all video frames in a Group Of Picture (GOP) directly or indirectly use the most recent intra-coded frame as a reference frame, the corruption of the intra-frame coded picture results in error propagation to subsequent inter-coded frames. The PSNR degradation seen in Figure 3.10 for the subsequent frames is almost as high as that in the originally corrupted intra-frame coded picture itself. Figure 3.11 shows that similar trends are also valid for the corruption of a P-frame in the GOP.

In summary, as Figures 3.10 and 3.11 reveal, the I-frame errors or P-frame errors result in image quality degradations that diminish fairly slowly with time. Additionally, the errors will propagate to the forthcoming inter-frame coded pictures, until the next intra-frame coded picture arrives. In contrast, the errors in a B-frame do not affect the image quality of subsequent frames as drastically as the I-frame or P-frame errors.

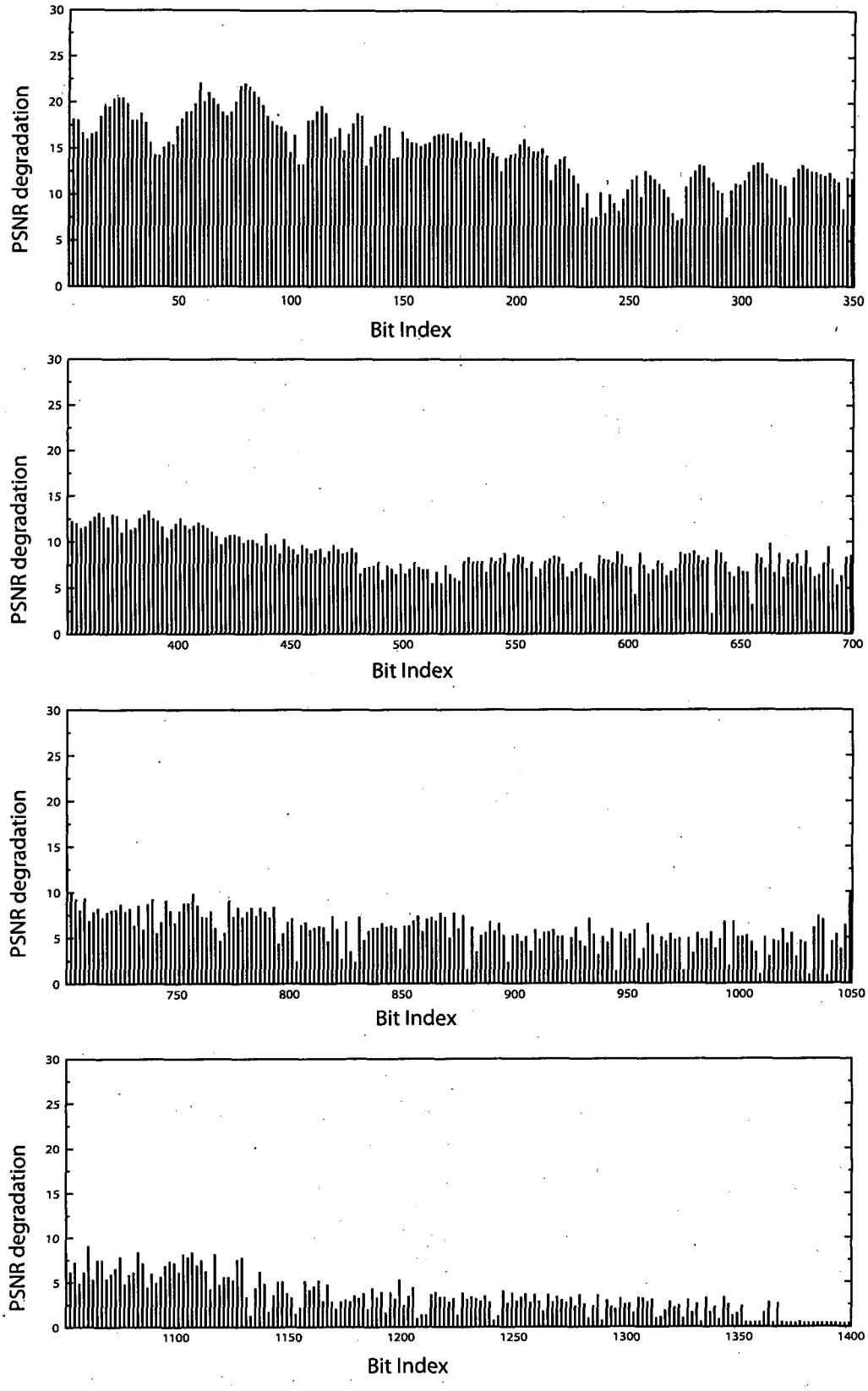


Figure 3.7: Average PSNR degradation due to the bit corruption events for the *"Miss_America"* QCIF video sequence encoded at 30 frame/s and 60kbit/s using the inter-frame mode.

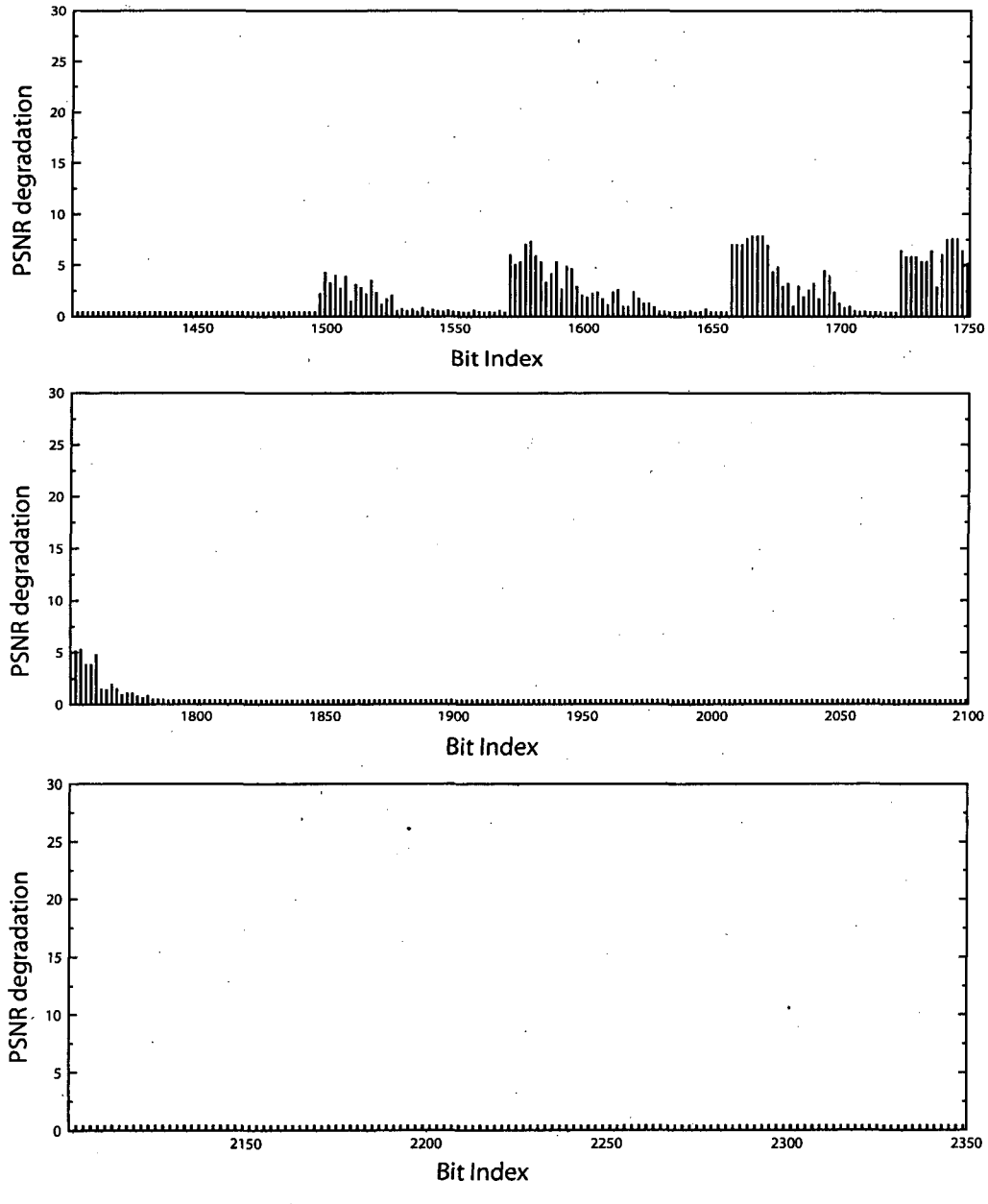
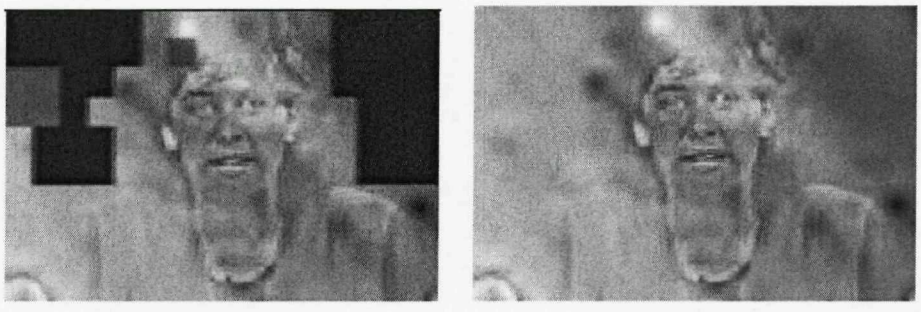


Figure 3.8: (Continued from Figure 3.7) Average PSNR degradation due to the bit corruption events for the *"Miss_America"* QCIF video sequence encoded at 30 frame/s and 60kbit/s using the inter-frame mode.



(a) (b)

Figure 3.9: Image quality degradation due to the corruption of the Y component's first bit in the lowest LL subband (a) and the resultant motion-predicted second frame (b) for the *"Miss_America"* QCIF video sequence encoded at 30 frame/s and 60kbit/s using 2181 bits for this frame.

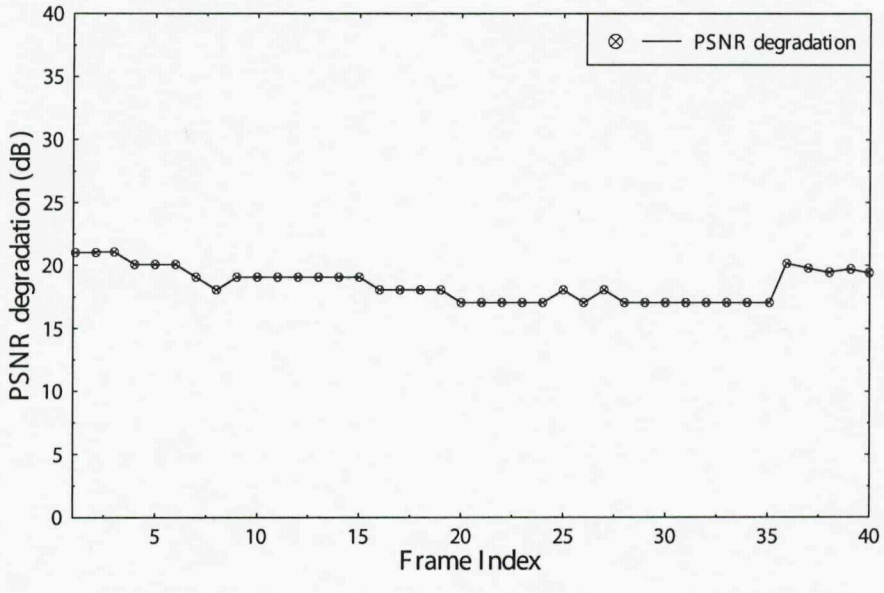


Figure 3.10: PSNR degradation due to the corruption of the first bit of the lowest LL subband in intra-frame mode and its propagation to the subsequent frames for the *"Miss_America"* QCIF video sequence encoded at 30 frame/s and 60kbit/s.

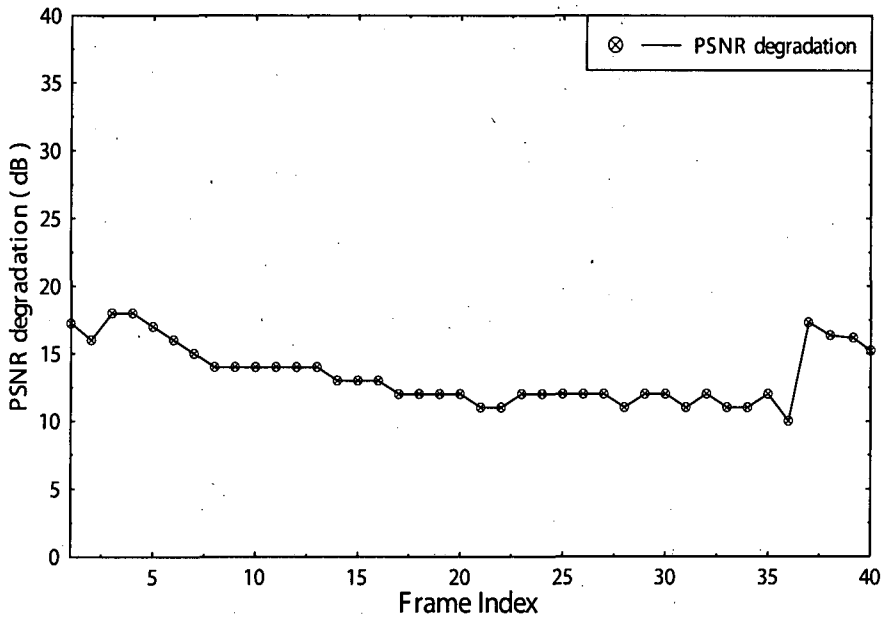


Figure 3.11: PSNR degradation due to the corruption of the first bit of the motion vectors in the first forward predicted frame and its propagation to the subsequent frames for the *"Miss_America"* QCIF video sequence encoded at 30 frame/s and 60kbit/s.

3.4 Unequal Error Protection Scheme for Dirac Video Transmission

3.4.1 Introduction

As quantified in Section 3.3, the Dirac video-encoded sequence exhibits different bit error sensitivities for the different bit positions. If the channel encoder has the ability to exploit the knowledge of the source bit error sensitivity, Unequal Error Protection (UEP) has the potential to outperform Equal Error Protection (EEP). This is achieved by applying a lower channel coding rate to the more sensitive portions of the data sequence, and a higher channel coding rate to the less sensitive portions, while keeping the overall channel coding rate the same as that of the EEP scheme.

At the source encoder the implementation of UEP often involves data partitioning, which partitions an encoded bitstream into groups of bits with different sensitivities. At the channel encoder specific classes of error control codes, such as rate compatible punctured convolutional (RCPC) codes [102] and rate compatible punctured turbo (RCPT) codes [103], can be employed to provide different levels of error protection to the groups of bits with different sensitivities. These codes use the same mother code to provide different

channel coding rates with the aid of puncturing.

Recently, in [104,105] Tüchler studied the construction of irregular convolutional codes (IRCCs) and proposed several design criteria. These IRCCs were constructed from a family of convolutional codes having different code rates and were designed with the aid of extrinsic information transfer (EXIT) charts [106]. In general, EXIT chart analysis only provides accurate performance estimates, which have a long interleaver block length. However, it was shown in [104] that by using an appropriate optimization criterion, the concatenated system is capable of performing well even for short interleaver block lengths. Since the constituent codes have different coding rates, the resultant IRCC is capable of providing UEP.

The novel contribution of this section is that turbo-equalized IRCCs are successfully applied to robust Dirac-encoded video transmission. An iterative detection aided unequal error protection Dirac videophone scheme using an IRCC as the outer code is designed for the transmission of the Dirac-encoded video sequence. Rather than being decoded separately, the constituent codes of the IRCC are decoded jointly and iteratively by exchanging extrinsic information with the inner code. Furthermore, the characteristics of the Dirac codec and those of the channel are employed to optimize IRCCs, so that UEP is achieved while maximizing the iteration gain of the system. Simulation results are provided to demonstrate that the UEP scheme outperforms the EEP scheme in terms of both the achievable bit error ratio (BER) and the PSNR.

3.4.2 Source-Coding Assisted Design of IRCCs

In this section, an UEP scheme is designed by selecting the best-matched IRCCs for the Dirac video sequence. In our work, we first partition the unequal-sensitivity information of a video-encoded sequence into a number of protection classes and then employ IRCCs to protect them differently. As mentioned above, IRCCs allow us to control the protection of the different-sensitivity classes. Based on the investigations of Section 3.3, our partitioning algorithm splits the Dirac video frame into three classes, namely the Most Significant Bits (MSB), Very Significant Bits (VSB) and the Least Significant Bits (LBSs). The class a certain bit belongs to is dynamically decided as follows:

1. The video header bits, including the frame header, motion vector header and all 30-subbands' headers constitute MSBs. An error corrupting the MSBs will result in the

decoder's loss of synchronization with the encoder and hence results in catastrophic decoding failure of the subsequent process.

2. All bits of the motion vectors are assigned to the LSB class, if the number of motion vector bits is less than 500. This is, because it was documented in [82] that despite their inter-frame propagation an error in any of them will inflict a relatively low distortion in the decoded frame, which results in a PSNR degradation of less than 8dB.

3. All bits of a subband's DWT coefficients are also deemed to be LSBs, when the number of subband bits is less than 500. The video PSNR degradation imposed by the corruption of any of them is typically confined to the range below 8dB.

4. By contrast, the motion vector and subband parameter encoding bits are split into two different classes, if the specific subband generates more than 500 bits. More specifically, the least sensitive 500 bits are considered as LSBs, while the remaining bits will be deemed to be VSBs.

Naturally, this bit-allocation scheme would be different for higher video resolutions, such as for example 4CIF and HDTV formats, which predominantly portray different video content. Although the three protection classes are interleaved during the encoding and decoding process, their classifier information (namely whether they are MSBs, VSBs or LSBs) is recovered by the receiver based on the strongly-protected side-information provided by the MSBs.

Furthermore, the packetization algorithm of Figure 3.13 concatenates these classes in order to form fixed-length packets to transmit. It is worth noting that due to the different size of the classes and fixed-length subcodes of the IRCC scheme, the concatenation of these classes may lead to different code-rates for protecting bits potentially belonging to the same protection class.

Table 3.3: The average ratio of the three protection classes in the Dirac video sequence.

	The relative frequency of occurrence		
	MSB class	VSB class	LSB class
Probability (%)	30.41	19.77	49.82

An interesting question arising for the design of the IRCC scheme is how to decide the ratio of the number of bits in each subcode of the IRCC. The basic criterion is to try to keep the same ratio as in the original Dirac video-encoded sequence. Unfortunately, this ratio is time varying. Table 3.3 shows the average ratio computed from the Dirac-encoded *"Miss_America"* QCIF video clip scanned at 30 frame/s. More details are depicted in Figure 3.12. We can see from Figure 3.12 that although the ratio of the three protection classes varies from frame to frame, their range is limited. Hence we chose a ratio that falls in the middle of each of the ranges. Additionally, three short IRCC subcodes, which encode [1.3% 4.8% 6.5%] of the IRCC packet, are also used for the sake of dealing with the fluctuation of the associated ratio.

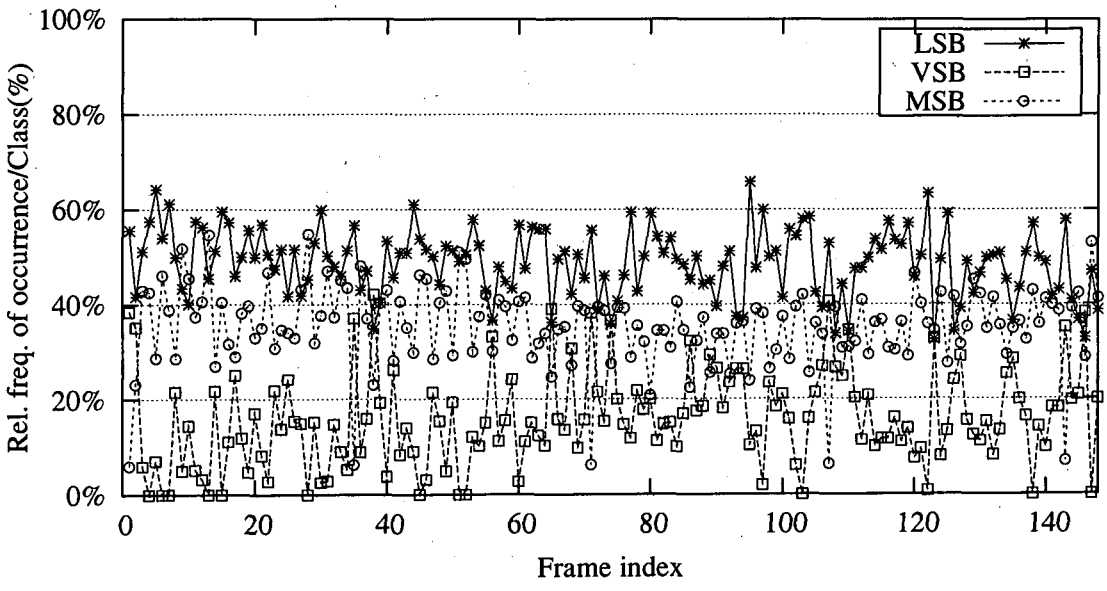


Figure 3.12: The relative frequency of occurrence for the various Dirac video bit portions computed from the first 40 frames of *"Miss_America"* QCIF video sequence encoded at 30 frame/s and 60kbit/s.

Another factor in deciding the ratio of MSB, VSB, and LSB bits is the fairness of comparison. To compare the result of the EEP and UEP scheme, we have to fix the overall channel rates to be the same. More specifically, the EEP and UEP schemes are designed as follows:

1. For the EEP scheme, a classic half-rate maximum free-distance Non-Systematic Convolutional (NSC) code was employed as the outer channel code.

2. For the UEP scheme, we employ the IRCC consisting of a set of subcodes having coding rates of [0.35 0.45 0.5 0.55 0.6 0.8] and a constraint length of 5, which encode [34% 1.3% 4.8% 22% 6.5% 31%] of the input data bits, respectively. This results in the IRCC having an average coding rate of 0.5 and a constraint length of 5 for approximately $K = 2000$ video bits per packet.

3.4.3 System Model

In this section, the achievable performance of the Dirac-encoded video telephone scheme was investigated, when communicating over dispersive AWGN channels. In order to reduce the channel induced Inter-Symbol Interference (ISI), powerful turbo equalization schemes were employed [64]. The associated system model is depicted in Fig. 3.13.

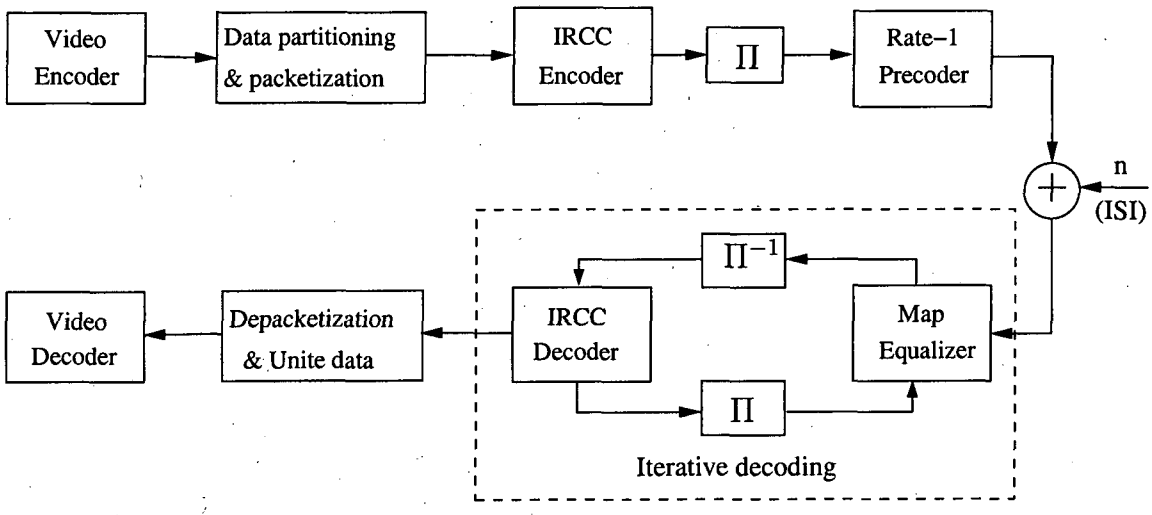


Figure 3.13: System Model

At the transmitter, the video-compressed frames output by the Dirac encoder are first partitioned into three different bit error sensitivity classes, as detailed in Section 3.4.2. These classes are then combined into fixed-length packets of $K = 2000$ bits before being fed into the channel encoder. More specifically, in order to guarantee that the MSBs is protected by the lowest-rate subcodes and that the VSBs and LSBs are protected by higher-rate subcodes, the packetization scheme has to decide the ratio of the packet's bits assigned to each protection class. If the video-compressed frame generates more than 1 packet, the ratio of [34% 28% 38%] of the first $K = 2000$-bit packet was assigned to the

MSBs, VSBs, LSBs, respectively, and the remaining bits of each class were stored in the class' buffer. When a higher sensitivity class' buffer is not fully filled, for example the MSB's buffer, some of the VSBs will be assigned to it, and so on.

The channel encoder consists of an outer encoder and a rate-1 precoder [107], which are separated by the interleaver Π of Figure 3.13. The outer code adds redundancy for the sake of providing protection for the video data, while the precoder renders the channel to appear recursive[2], which assists in reducing the associated BER [107], when employing iterative detection.

On the other hand, at the receiver of Figure 3.13 a turbo equalizer is invoked, which consists of an inner Maximum A Posteriori (MAP) channel equalizer [108] and an outer channel decoder. Both of them are capable of processing soft inputs and generating soft outputs, which are exchanged between the two components in a number of consecutive iterations for the sake of improving the attainable decoding/detection performance. At the last iteration hard decisions are carried out, and the resultant bit stream is depacketized by using the side-information provided by the MSBs for reassembling the video-compressed frames, which are then fed into the Dirac decoder of Figure 3.13 for the sake of reconstructing the video sequence.

To be specific, the proposed scheme employs an IRCC scheme as the outer channel code, which was designed to match the channel's characteristics as well as those of the Dirac video sequence. An IRCC is constructed from a set of subcodes [109], which have different coding rates and are generated using puncturing from the same mother convolutional code. Each subcode encodes a fraction of the input video bits, while maintaining the overall coding rate at a specified value. For more details on the design of IRCCs, please refer to [104, 109], where the same IRCC was used, except that the coding rate ratios of the subcodes were different. The design method proposed in [109] was used here for optimizing the IRCCs for the specific dispersive channel considered and for the wavelet video codec invoked. As our benchmarker scheme, a classic maximum free-distance Non-Systematic Convolutional (NSC) code was employed as the outer channel code, which provides EEP for the video bits.

[2]The term 'recursive' implies here that the system has an infinite impulse response.

3.4.4 Convergence Analysis of the Turbo Equalizer Using EXIT Charts

It is worth noting that for a non-dispersive channel, there is no need to use an equalizer. In this case, the performance of the outer channel code represents a bound on the achievable receiver performance. In order to achieve an improved performance, for example the outer code can be replaced by a near-capacity turbo code, as suggested in [110]. Alternatively, in a dispersive channel context the non-iterative FEC decoder can be combined with an equalizer and may exchange *extrinsic* information with the equalizer [108]. When the channel is non-dispersive, it can be made to appear recursive to the receiver, thus providing a potentially increased iterative detection performance. To this effect, recursive rate-one precoders may be invoked, which are capable of effectively rendering the channel to appear recursive and hence they have been shown to yield significant performance gains in iterative equalization aided systems [107]. A further advantage of the precoder is that the receiver's complexity owing to its inclusion does not increase significantly.

Following the design of the IRCCs outlined in Section 3.4.2, the design of our UEP system may be viewed as a joint source/channel codec optimization task, where our aim is to ensure that the extrinsic information transfer function of the outer IRCC matches that of the inner rate-1 precoder, because it was shown in [111] that this leads to near-capacity operation.

For the sake of investigating the convergence behaviour of the turbo equalizer aided receiver of Figure 3.13, the rate-1 precoder having a generator polynomial of $g_0 = 1 + D^2$ was invoked. A five-path dispersive channel having an impulse response of

$$\begin{aligned} h[n] &= 0.227\delta[n] + 0.46\delta[n-1] + 0.688\delta[n-2] \\ &\quad +0.46\delta[n-3] + 0.227\delta[n-4] \end{aligned} \tag{3.1}$$

and hence imposing severe ISI was selected from [112] for our investigations. The EXIT functions of the six IRCC subcodes having coding rates of [0.35 0.45 0.5 0.55 0.6 0.8], respectively, and that of the amalgamated IRCC having an overall coding rates of 1/2 is depicted in Figure 3.14. The EXIT chart of the system of Figure 3.13 using the above IRCC and the MAP turbo channel equalizer for communicating over the 5-path channel of Equation 3.1 is shown in Figure 3.15.

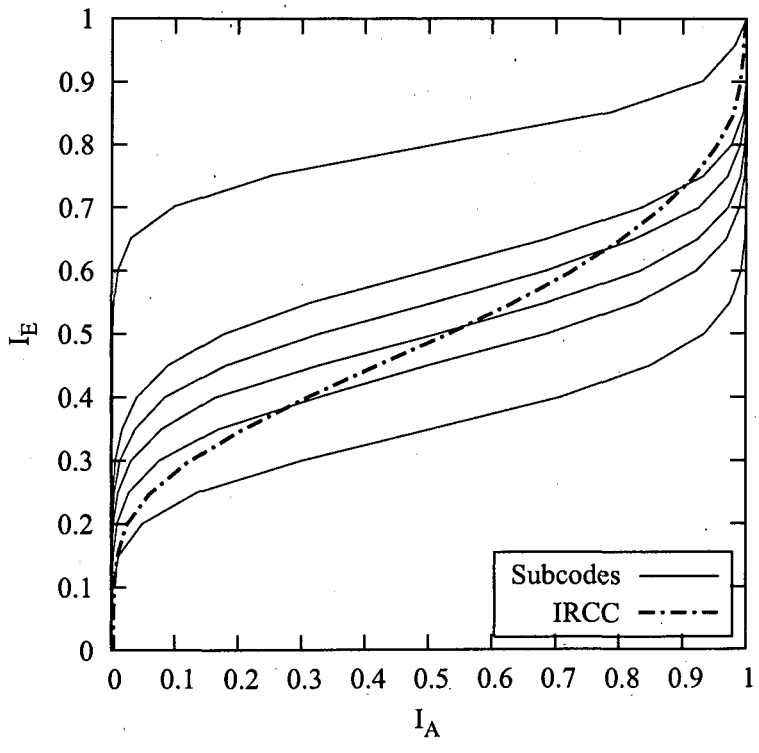


Figure 3.14: The EXIT function of a rate-1/2 IRCC and that of its subcodes

As seen from the EXIT charts of Figure 3.15, the convergence SNR threshold for the system using the IRCC is slightly above 1dB, while that for the EEP system using the rate-1/2 NSC code is close to 3dB. More quantitatively, we found that the former required an SNR of 1.3dB, while the latter necessitated 2.4dB for maintaining an open EXIT chart, which facilitates convergence to a low BER.

3.4.5 Simulation Results

The achievable system performance was evaluated for $K = 2000$ bits per video frames facilitating the employment of an interleaver length of $L = 4000$ bits at a total average IRCC code-rate of 0.5 across a single video frame, as summarized in Table 3.4. The five-path dispersive channel of Equation 3.1 imposing severe ISI was selected for our investigations. Again, the outer IRCC has an average coding rate of 0.5 and a constraint length of 5. The UEP IRCC scheme consists of a set of subcodes having coding rates of

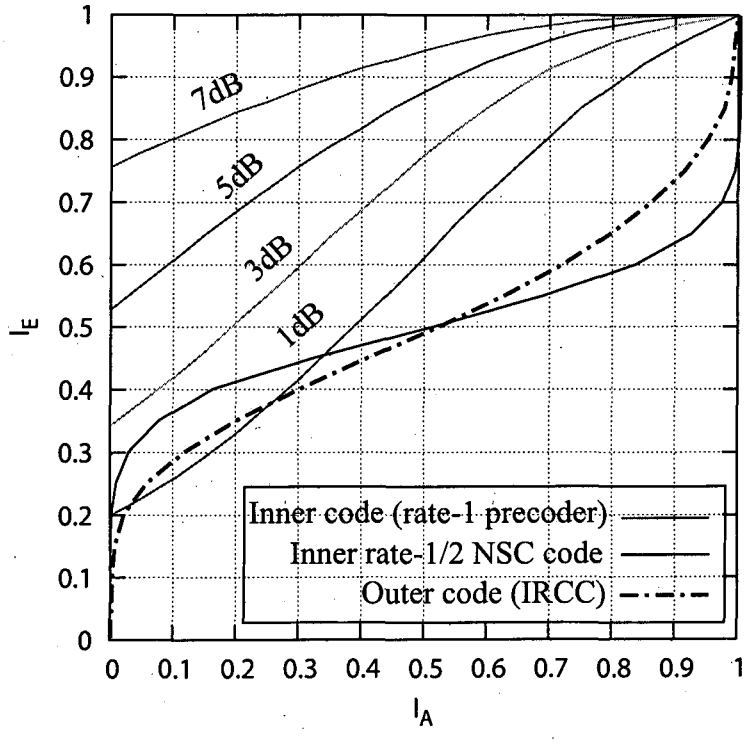


Figure 3.15: EXIT charts of the system of Figure 3.13 using the above IRCC and the MAP turbo channel equalizer for communicating over the 5-path channel of Equation 3.1.

[0.35 0.45 0.5 0.55 0.6 0.8], which encode [34% 1.3% 4.8% 22% 6.5% 31%] of the input data bits, respectively. For our outer NSC benchmarker scheme the generator polynomials of $g_0 = 1 + D + D^2 + D^4$, and $g_1 = 1 + D^3 + D^4$ were used. For both schemes, the rate-1 precoder having a generator polynomial of $g_0 = 1 + D$ was invoked, and again both the channel equalizer and the channel decoder employed the Maximum A *Posteriori* (MAP) algorithm [113]. In all simulations BPSK modulation and a dispersive AWGN channel were assumed.

The BER performances of the considered UEP scheme recorded during the iterative decoding process are depicted in Figure 3.16. The result was achieved by repeating the transmission of the 150-frame *"Miss_America"* video sequence 50 times, which was about $1.5 \cdot 10^7$ transmitted bits, and evaluating the video source sample reconstruction $PSNR$ following the completion of each decoding iteration. Clearly, as more and more decoding iterations are completed, the resultant source sample reconstruction SNR can be expected to increase, until iterative decoding convergence is achieved. However, the associated

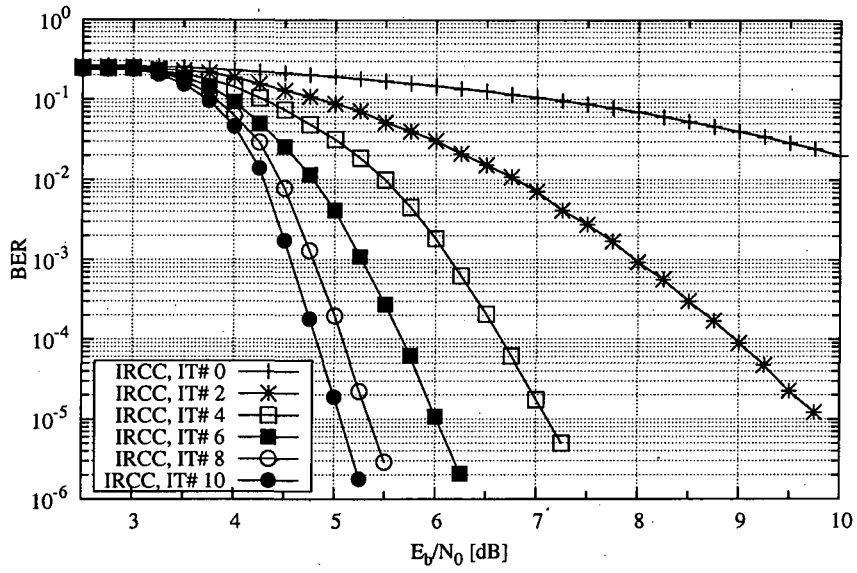


Figure 3.16: BER performance of the UEP scheme designed for Dirac video transmission over the five-path dispersive channel of Equation 3.1 imposing severe ISI. The system parameters are summarized in Table 3.4

total computational complexity will also increase as more and more decoding iterations are completed. Clearly, there is a tradeoff between the achievable reconstruction quality and the computational complexity imposed. Hence we opted for using $I = 10$ iterations in the proposed UEP scheme.

Table 3.4: The parameters of the videophone scheme.

The system parameters	
Average video frame size	2000 bits
Code rate	1/2
Interleaver length	4000
Number of decoder iterations	$I = 10$
Outer NSC polynomials	$g_0 = 1 + D + D^2 + D^4$
	$g_1 = 1 + D^3 + D^4$
Constraint length	5
Precoder polynomial	$g_0 = 1 + D$

The BER profiles of the UEP scheme are plotted in Figure 3.17. As intended, the different protection classes of the Dirac video-encoded frame benefited from different degrees

of IRCC-aided protection. As a result, the MSB class is capable of achieving near-error-free detection for channel E_b/N_0 values in excess of 4.25dB, while the LSB class requires an approximately 5.25dB E_b/N_0 for achieving a near error-free BER performance.

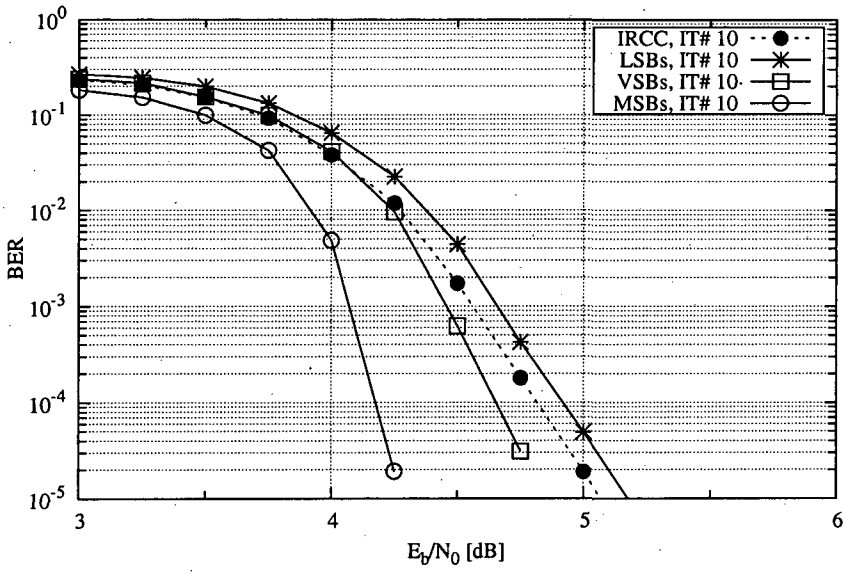


Figure 3.17: BER performance of the three protection classes of the UEP scheme designed for Dirac video transmission over the five-path dispersive channel of Equation 3.1 imposing severe ISI.

The video PSNR performances of both the UEP and EEP system are depicted in Figure 3.18. It can be seen that the UEP system attains a near-unimpaired PSNR of 33dB for channel E_b/N_0 values in excess of 4.5dB. By contrast, the EEP scheme requires an approximately 5.3dB E_b/N_0 for achieving the same PSNR. This result matches our performance prediction inferred from the EXIT chart analysis of Section 3.4.4

3.5 Chapter Conclusions

The main contributions of this chapter are constituted by our bit sensitivity study and the proposed UEP scheme using turbo-equalized IRCCs for the transmission of the Dirac video-coded sequence.

We commenced our discourse in Section 3.2 by visualizing the structure of the Dirac video-encoded sequence. The Dirac video-encoded sequence is then partitioned into several parameter categories, as seen in Figure 3.1. Hence, the video-coding simulations of

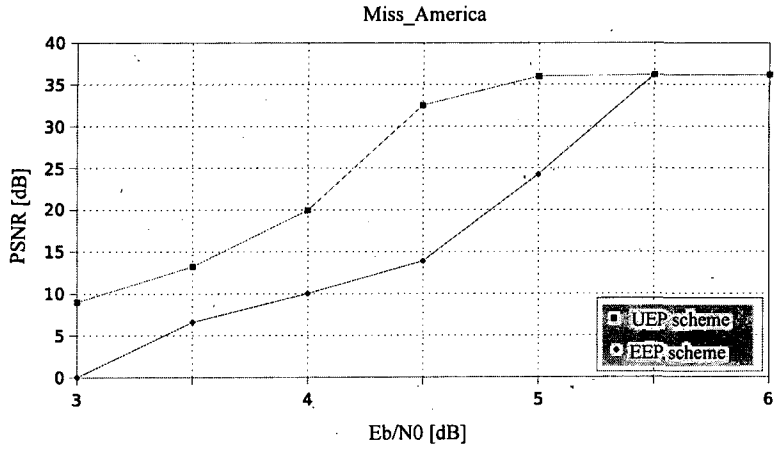


Figure 3.18: Comparison of the achievable video PSNR using both EEP and UEP. All system parameters are summarized in Table 3.4

Section 3.2.2 were conducted in order to quantify the relative frequency of occurrence for the various coding parameters in the Dirac video sequence. The simulation results shown in Figures 3.2 demonstrated that in the intra-frame coded mode the majority of the intra-coded frame's bits - namely 94% - represent the subband data, while all the other parameters are represented by the remaining 6% of the bits. By contrast, in the inter-frame coded mode, the bitrate contribution of motion vectors becomes quite dominant, for example about 40% of the total frame size of 1800 bits, while that of the inter-frame coded MCER was significantly reduced, namely to 23% of the total frame size, as seen in Figure 3.3.

Furthermore, in order to provide robust source-matched error protection for the wavelet coding-based video stream, the Dirac video codec's bit-sensitivity investigations of Section 3.3 were carried out by systematically corrupting all of its bits in a particular parameter of all consecutive encoded video frames, while keeping the rest of the parameters un-contaminated and then evaluating the average PSNR degradation inflicted. The objective Peak Signal to Noise Ratio (PSNR) degradations inflicted by the systematic bit corruption events depicted in Figures 3.5, 3.7 and 3.8 showed that

1. The Dirac video-encoded sequence exhibits different bit error sensitivities for the different bit positions.

2. In the inter-frame coding mode, the subjective quality degradation imposed by the corruption of the motion data parameters is typically high, while that of the subband

DWT coefficients is less dramatic.

3. The bit sensitivity is typically higher at the beginning of a subband and then reduces towards the end of the subband.

4. The subband's bit sensitivity also depends on the particular subband's spatial frequency. More specifically, at the same bit index of a low-frequency subband, the bit sensitivity is typically higher than that of a high-frequency subband and vice versa.

5. The number of DWT coefficients in a subband is an other factor influencing the sensitivity of a subband, since the so-called CABAC technique of Section 2.3.3 continuously updates its parameters based on the past received symbols for the sake of decoding the remaining part of the subband concerned. Consequently, the rest of the entire subband data is affected by a transmission error inflicted upon one of the first bits of that subband.

6. The sensitivity of approximately the last 500 bits in the subband data remains moderate (the PSNR degradation is usually confined to the range below 8dB).

Additionally, the results shown Figures 3.10 and 3.11 revealed that the corruption of the I-frame or P-frame will propagate to the forthcoming inter-frame coded pictures until the next I-frame arrived, and the image quality degradations diminish only fairly slowly with time. In contrast, the corruption of a B-frame do not affect the video quality as drastically as the I-frame or P-frame errors.

As a basis of our error resilience study, Section 3.3 proposed a partitioning scheme, which partitions the unequal-sensitivity information of a video-encoded sequence into three protection classes, namely the MSBs, VSBs and LBSs. The class a certain bit belongs to is dynamically decided as follows:

1. The video header bits, including the frame header, motion vector header and all 30-subbands' headers constitute MSBs. An error corrupting the MSBs will result in the decoder's loss of synchronization with the encoder and hence results in catastrophic decoding failure of the subsequent process.

2. All bits of the motion vectors are assigned to the LSB class, if the number of motion vector bits is less than 500. This is, because it was documented in [82] that despite their inter-frame propagation, an error in any of them will inflict a relatively low

distortion in the decoded frame, which results in a PSNR degradation of less than 8dB.

3. All bits of a subband's DWT coefficients are also deemed to be LSBs, when the number of subband bits is less than 500. The video PSNR degradation imposed by the corruption of any of them is typically confined to the PSNR range below 8dB.

4. By contrast, the motion vector and subband parameter encoding bits are split into two different classes, if the specific subband generates more than 500 bits. More specifically, the least sensitive 500 bits are considered as LSBs, while the remaining bits will be deemed to be VSBs.

The average ratio of these protection classes in the Dirac video sequence was provided in Table 3.3. As a result, Section 3.4 proposed a UEP scheme using turbo-equalized Irregular Convolutional Codes (IRCCs) for the transmission of the Dirac video-coded sequence, as seen in Figure 3.13. In the design of IRCCs, we aimed for matching the extrinsic information transfer function of the IRCC to that of the recursive rate-1 precoder, while maintaining unequal error protection capabilities matched to the requirements of the wavelet encoded video sequence. Furthermore, the UEP scheme was compared to a classic maximum free-distance NSC code considered as our EEP benchmarker scheme in terms of both the achievable BER and PSNR. The simulation results shown in Figures 3.14, 3.15, 3.16, 3.17 and 3.18 revealed that

1. The convergence SNR threshold for the system using the IRCC is lower than that for the EEP system using the rate-1/2 NSC code. More quantitatively, we found that the former required an SNR of 1.3dB, while the latter necessitated 2.4dB for maintaining an open EXIT chart, which facilitates convergence to a low BER.

2. The employment of a sufficiently low coding rate for the MSB class allows us to preserve the integrity of the corresponding bits.

3. The proposed UEP scheme is capable of attaining a near unimpaired video quality for channel SNRs in excess of about 4.5dB over a five-path dispersive AWGN channel and it outperforms the EEP scheme by about 0.8dB in terms of the required E_b/N_0.

Future work could improve the Dirac video transmission scheme's performance in three different aspects. First, the question of optimum coding rate allocation for UEP remains

unanswered. Amongst several reasons, this is partly owing to the fact that the difference in ratio of the three protection classes fluctuated from frame to frame. Hence the IRCCs using a fixed proportion of subcodes may lead to some limitations. For example, if the encoded frame contains many MSBs, they may have to be protected by relatively weak high-rate subcodes when aiming for a fixed total channel-coded rate and hence their channel decoding process may fail in the presence of transmission errors.

Secondly, the potential problem of losing synchronization has to be addressed when the VLCs are corrupted. Significant image quality degradations could be avoided by transmitting strongly protected side information bits for the synchronization of the video decoder.

Finally, the natural residual redundancy remaining in the Dirac video-encoded sequence can be exploited for supporting the iterative exchange of *extrinsic* information between source and channel decoding. This has the potential of further improving the error correcting capability of the entire communication system. Combining this with an UEP scheme is expected to lead to a better video transmission system.

Chapter 4

Iterative Source-Channel Decoding for Video Telephony

4.1 Introduction

The intentionally imposed or inherent unintentional residual redundancy found in source encoded bitstreams can be exploited for supporting joint SoftBit-Source Decoding and Channel Decoding, which has the potential of improving both the error correcting capability as well as the subjective audio or video quality of communication systems. Numerous iterative source-channel decoding approaches have been investigated in the literature [44, 114, 115], leading to a variety of guidelines. For example, in [44] and [114] the source decoder exploited the inherent residual redundancy in the encoded bitstream for the sake of providing an increased error protection. By contrast in [115] Buttigieg deliberately increased the inherent residual redundancy in the encoded bitstream, creating a class of Variable-Length Error-Correcting (VLEC) codes, which combines the benefits of variable-length coding with improved distance properties. In a manner similar to that of [115], a class of Over-Complete Mapping (OCM) scheme was created, which intentionally increased redundancy in the encoded bitstream for the sake of improving the error correcting capability of the softbit-based source decoding algorithm detailed in [116].

More explicitly, this chapter investigate Iterative Source-Channel Decoding (ISCD) designed for robust Dirac-encoded video transmission over fading channels. Our first contribution is that of exploiting the residual redundancy inherent in the Dirac-encoded

bitstream for the sake of providing an increased error protection. This can be achieved in an iterative 'turbo-detection' fashion, which exchanges *extrinsic* information between the channel decoder and the softbit source decoder. Moreover, the performance of the Dirac-encoded video transmission scheme was investigated by analyzing the convergence behaviour of an ISCD scheme using EXIT charts. Simulation results provided in this chapter demonstrate that direct iterative source-channel decoding of Dirac-encoded video transmissions does not perform well, because the amount of inherent redundancy in the Dirac-encoded sequence is limited.

As a counter-measure, in this chapter a powerful ISCD scheme is proposed, where the softbit source decoder benefits both from any residual redundancy inherent in the source encoded bitstream as well as from the intentional redundancy imposed by the novel OCM schemes to be introduced in Section 4.3.2.

The design of ISCD is based on formulating the necessary and sufficient condition for ensuring the iterative decoder's convergence to an infinitesimally low Bit Error Ratio (BER). We then proceed by proposing an algorithm for designing a suitable mapping for iterative source-channel decoding. More explicitly, we demonstrate that the necessary and sufficient condition for ensuring successful ISCD is the employment of the over-complete bit-to-symbol mapping having a minimum Hamming distance of $d_H = 2$. Furthermore, the overall system's robustness against transmission errors as well as the implementational complexity of the OCM-assisted Softbit Source Decoding (OCM-SSD) scheme was characterized. We continue by quantifying the effects of interleaver delay on the OCM-SSD's performance by analyzing its EXIT [106] characteristics in conjunction with a Unity Rate Code (URC) [117] employed as the inner codec and the former as the outer codec. This allows us to design an attractive video transmission scheme having a high error resilience at a reasonable complexity and a low delay. Our results outlined in Section 4.3.3 demonstrate that imposing redundancy on the source-coded bitstream with the aid of the OCM scheme, which is effectively as EXIT-chart Optimized Block Code (ECOBC) has the potential of outperforming even the powerful benchmarker of [116] at the same overall transmission rate.

The rest of this chapter is organized as follows. In Section 4.2, an overview of ISCD is presented and the dependence of its performance on the potential redundancy of the source data is investigated using EXIT charts. Section 4.3 demonstrates how this scheme exploits the intentionally introduced redundancy of the OCM, accompanied by our de-

sign guidelines. Furthermore, the benefits of our design guidelines is demonstrated by investigating the attainable performance of three identical-rate Dirac video transmission schemes. The OCM-SSD is characterized in Section 4.4 by investigating its implementational complexity, the trade-off between the achievable decoding performance or coding gain and the mapping rate as well as the effects of interleaver delay on the OCM-SSD's performance. Finally, our conclusions are provided in Section 4.5.

4.2 Iterative Source-Channel Decoding Using the Dirac Video Codec

According to Shannon's source and channel coding separation theorem [1], these operations may be carried out separately, following separate optimization of each component. However, in the context of limited-complexity, limited-delay, lossy source codecs, which exploit the psycho-visual and psycho-acoustic properties of human perception, Shannon's lessons have limited applicability, as detailed in the preface of [7]. This is particularly so, when the real-time, interactive video system experiences bursty errors inflicted by dispersive fading channels, rather than independent random errors. As a result, the joint design of source and channel coding modules can improve the achievable performance of practical communications systems.

In general, joint source-channel decoding (JSCD) schemes can be divided into iterative JSCD and non-iterative JSCD, according to whether there is iterative decoding between the source and channel decoders. While non-iterative JSCD provides excellent results for moderate channels, these methods fail to perform well for low channel SNRs or for channels infested by dispersive fading. For these scenarios, iterative source-channel decoding (ISCD) may be used for mitigating the effects of transmission errors. Bauer and Hagenauer proposed a novel iterative source-channel decoding approach for variable-length coded (VLC) memoryless sources [118] by applying the "*turbo*" principle [119]. The proposed scheme modified the conventional hard VLC decoder by invoking an a *posteriori* probability (APP) soft decoder and then by facilitating soft *extrinsic* information exchange between the APP channel decoder and the APP VLC decoder. Their work was first applied to symbol by symbol VLC decoding [120] using a symbol-based VLC-trellis. This philosophy was later extended to bit by bit VLC decoding using a bit-based VLC-

trellis [118]. The approach of [120] outperforms separate source and channel decoding by about 2.3dB.

More recently, Adrat and Vary have designed a softbit source decoding technique [116], which enables iterative source-channel decoding in order to exploit the residual redundancy inherent in the source-encoded bitstream for the sake of achieving an additional error protection. In the following section, we first briefly review this technique and then investigate its performance for the transmission of Dirac video-encoded sequences over an uncorrelated Rayleigh fading channel with the aid of EXIT charts. We will also investigate the effects of different residual redundancies inherent in the source-encoded sequence by employing the novel overcomplete mapping scheme of Section 4.3.

4.2.1 Definitions and Notations

In this treatise, we use capital letters, such as X, to denote sets of finite-length binary strings, while lower-case letters, such as x, are reserved for denoting the finite-length binary strings themselves. We also use capital letters for superscripts as in X^K to denote a set constituted by K-bit binary strings, while lower-case letters for the subscripts as in x_i to indicate the i-*th* bit of the binary string x. Furthermore, the superscript [extr] is used in conjunction with the subscript j as in $x_j^{[extr]}$ to denote all bits of the binary string x, except for the j-th bit, i.e. we have $x_j^{[extr]} = [x_1, ..., x_{j-1}, x_{j+1}, ..., x_N]$. Moreover,we use $\|X\|$ to denote the cardinality of a set X.

Definition 1 *The Hamming distance between two binary codewords of equal length is the number of positions for which the corresponding bits are different.*

Definition 2 *The minimum Hamming distance of the set X is defined as the minimum of the Hamming distances between all possible codeword pairs of the set X: $d_h(X) = min\Big\{ d_h(c_i, c_j) | c_i, c_j \in X \Big\}$..*

Definition 3 *The mapping φ from set X to set Y φ: $X \rightarrow Y$ is referred to as an Over-Complete Mapping (OCM), if X is a strict subset of the set Y, i.e. we have $(X \subset Y)$.*

Definition 4 *The Hamming distance of the over-complete mapping φ: $X \rightarrow Y$ is defined as the Hamming distance of the set $\varphi(X)$.*

Definition 5 *The over-complete mapping $f(X)$ is termed as an EXIT-optimized mapping, if the EXIT curve of the outer softbit source decoder may reach the point* $\left(I_{SBSD}^{L^{apri}}, I_{SBSD}^{L^{extr}}\right) = (1,1)$ *in the context of the ISCD scheme.*

4.2.2 Soft-bit Source Decoding

The conventional SBSD scheme introduced in [116] determines the *extrinsic* information from the natural residual redundancy, that inherently manifests itself in terms of the correlation of the source encoded bits. This residual redundancy manifests itself in terms of the non-uniform probability of occurrence or in term of the auto-correlation of the resultant bit patterns. In an attempt to simplify the receiver, we characterize this redundancy with the aid of the non-uniform symbol probability distribution of the natural binary K-bit symbol representation of the compressed binary video stream. More explicitly, the source bit statistics were evaluated by dividing the wavelet video encoded bitstream into a K-bit/symbol sequence. Let us now portray the softbit-based decoding technique in more detail.

Let $U = \{u = [u_1, u_2, ..., u_K]\}$ be the natural binary representation of the K-bit symbol having a symbol probability distribution of $\{p(u) = p([u_1, u_2, ..., u_K])\}$. If we denote the reliability information of the K-bit symbol corresponding to the transmitted symbol $u = [u_1, u_2, ..., u_K]$ by $\hat{u} = [\hat{u}_1, \hat{u}_2, ..., \hat{u}_K]$, then the probability of the K-bit symbol is given by the product of its single-bit probabilities as follows:

$$p(\hat{u}|u) = \prod_{j=1}^{K} p(\hat{u}_j|u_j). \tag{4.1}$$

The extrinsic channel-output information $p(\hat{u}_i^{[extr]}|u_i^{[extr]})$ generated for each desired bit u_i is obtained as

$$p(\hat{u}_i^{[extr]}|u_i^{[extr]}) = \prod_{\substack{j=1 \\ j\neq i}}^{K} p(\hat{u}_j|u_j). \tag{4.2}$$

Finally, the extrinsic value of each bit expressed in terms of the corresponding Logarithmic Likelihood Ratio (LLR) can be obtained by combining the channel's output information

and our *a priori* knowledge of the corresponding N-bit symbol as follows:

$$\begin{aligned} LLR(u_i) &= log\frac{\sum\limits_{u_i^{[extr]}} p\big(u_i^{[extr]}|u_i=+1\big).p\big(\hat{u}_i^{[extr]}|u_i^{[extr]}\big)}{\sum\limits_{u_i^{[extr]}} p\big(u_i^{[extr]}|u_i=-1\big).p\big(\hat{u}_i^{[extr]}|u_i^{[extr]}\big)} \\ &= log\frac{\sum\limits_{u_i^{[extr]}} p\big(u_i^{[extr]}|u_i=+1\big)\prod\limits_{\substack{j=1\\ j\neq i}}^{K} p\big(\hat{u}_j|u_j\big)}{\sum\limits_{u_i^{[extr]}} p\big(u_i^{[extr]}|u_i=-1\big)\prod\limits_{\substack{j=1\\ j\neq i}}^{K} p\big(\hat{u}_j|u_j\big)}. \end{aligned} \tag{4.3}$$

Note that $p\big(\hat{u}_i^{[extr]}|u_i^{[extr]}\big)$ can also be expressed in terms of an LLR value as [116]:

$$p\big(\hat{u}_i^{[extr]}|u_i^{[extr]}\big) = \Psi \cdot exp\Big(\sum_{u_l \in u_i^{[extr]}} \frac{u_l}{2} \cdot L_{CD}^{[extr]}(u_l)\Big), \tag{4.4}$$

where $L_{CD}^{[extr]}$ represents the extrinsic LLR values generated by the soft-output inner decoder, while the factor Ψ cancels out in Equation 4.3. Hence, the numerator of Equation 4.3 can be rewritten as follows

$$\begin{aligned} A &= \sum_{u_i^{[extr]}} p\big(u_i^{[extr]}|u_i=+1\big) \cdot exp\Big(\sum_{u_l \in u_i^{[extr]}} \frac{u_l}{2} \cdot L_{CD}^{[extr]}(u_l)\Big), \\ &= \sum_{u_i^{[extr]}} exp\Big[log\Big(p\big[u_i^{[extr]}|u_i=+1\big]\Big)\Big] \cdot exp\Big(\sum_{u_l \in u_i^{[extr]}} \frac{u_l}{2} \cdot L_{CD}^{[extr]}(u_l)\Big), \\ &= \sum_{u_i^{[extr]}} exp\Big(\Big[log\Big(p\big[u_i^{[extr]}|u_i=+1\big]\Big)\Big] + \sum_{u_l \in u_i^{[extr]}} \frac{u_l}{2} \cdot L_{CD}^{[extr]}(u_l)\Big). \end{aligned} \tag{4.5}$$

Similarly, the denominator of Equation 4.3 may be formulated as:

$$B = \sum_{u_i^{[extr]}} exp\Big(\Big[log\Big(p\big[u_i^{[extr]}|u_i=-1\big]\Big)\Big] + \sum_{u_l \in u_i^{[extr]}} \frac{u_l}{2} \cdot L_{CD}^{[extr]}(u_l)\Big). \tag{4.6}$$

Hence we have

$$LLR(u_i) = log\frac{\sum_{u_i^{[extr]}} exp\left(\left[log\left(p\left[u_i^{[extr]}|u_i=+1\right]\right)\right] + \sum_{u_l \in u_i^{[extr]}} \frac{u_l}{2} \cdot L_{CD}^{[extr]}(u_l)\right)}{\sum_{u_i^{[extr]}} exp\left(\left[log\left(p\left[u_i^{[extr]}|u_i=-1\right]\right)\right] + \sum_{u_l \in u_i^{[extr]}} \frac{u_l}{2} \cdot L_{CD}^{[extr]}(u_l)\right)}. \tag{4.7}$$

4.2.3 Iterative Source-Channel Decoding Using the Dirac Video Codec

4.2.3.1 System Overview

The block diagram of the ISCD scheme considered is shown in Figure 4.1. At the transmitter, the Dirac wavelet video codec is employed for compressing the video frames. The output binary sequence u_k is interleaved using the random bit-interleaver Π of Figure 4.1 in order to obtain the interleaved sequence $\overline{u}_k$. Subsequently, the interleaved sequence $\overline{u}_k$ is encoded by the Generalized Low-Density Parity-Check (GLDPC) Code, which was detailed in [121] and will be briefly reviewed in Section 4.2.3.2, before its transmission over an uncorrelated Rayleigh fading channel.

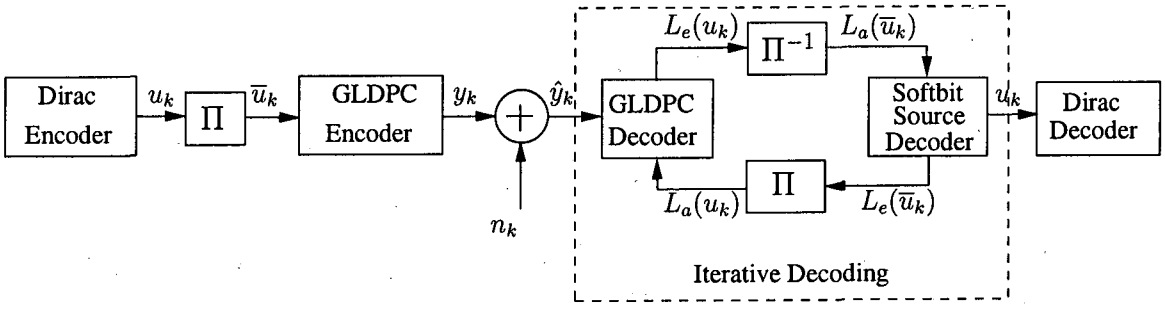


Figure 4.1: Iterative source and channel decoding model

At the receiver, the reliability information is evaluated in an iterative '*turbo-detection*' fashion, which exchanges *extrinsic* information between the GLDPC decoder and the softbit source decoder. This reliability information can be evaluated either in terms of *bit probabilities* or as their LLRs.

More specifically, during the first stage of iteration the GLDPC decoder computes the *a posteriori* values $L(u_k)$ both from the channel's output $\hat{y}_k$ and from the *a priori* information $L_a(u_k)$, which is fed back from the softbit source decoder. Note that for the first iteration $L_a(u_k) = 0\ \forall k$. At the output of the GLDPC decoder, the *a priori* values $L_a(u_k)$ are subtracted from the *a posteriori* values $L(u_k)$ for the sake of generating the *extrinsic* LLR values $L_e(u_k)$. The *extrinsic* LLR values $L_e(u_k)$ of the GLDPC decoder are deinterleaved by the soft-bit deinterleaver Π^{-1} of Figure 4.1 and then fed to the softbit source decoder as the *a priori* information $L_a(\overline{u}_k)$.

In the second stage of iterations, the softbit source decoder generates the *extrinsic* LLR information $L_e(\overline{u}_k)$ from the *a priori* information $L_a(\overline{u}_k)$ provided by the GLDPC decoder. Similarly, the *extrinsic* LLR values $L_e(\overline{u}_k)$ are interleaved by the soft-bit interleaver Π of Figure 4.1 and then fed back to the GLDPC decoder. After the last iteration, hard-decision estimation is invoked in oder to obtain the Dirac-encoded video frame and the resultant decoded video frame is decoded by the Dirac video decoder.

4.2.3.2 Generalized Low-Density Parity-Check Codes

Low Density Parity Check (LDPC) codes were devised by Gallager [122] and generalized by Tanner [123]. However, these codes remained relatively unexplored until the Hamming-code based GLDPC codes were proposed by Boutros [124] as well as by Lentmaier and Zigangirov [121]. GLDPC decoding is based on a combination of low-complexity soft-input soft-output (SISO) decoding of short, linear block codecs, such as for example Hamming-codes, binary or non-binary Bose-Chaudhuri-Hocquenghem (BCH) codes. To construct a GLDPC code, one has to replace each single parity check equation of the classic LDPC codes by the Parity-Check Matrix (PCM) of a simple linear block code referred to as a constituent code [121, 124, 125].

In our simulations, the GLDPC code proposed by Boutros [124] was invoked. The PCM of this GLDPC code is a sparse matrix H constructed by concatenating J sub-matrices, namely $H_1, H_2, ..., H_W$. The submatrices $H_1, H_2, ..., H_W$ seen in Figure 4.2 are the PCMs of the super-codes $C_1, C_2, ..., C_W$, respectively, that are constructed from the

constituent codes $C(n, k)$.

$$\mathbf{H_1} = \begin{bmatrix} H_0 & 0 & \dots & 0 \\ 0 & H_0 & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & H_0 \end{bmatrix}$$

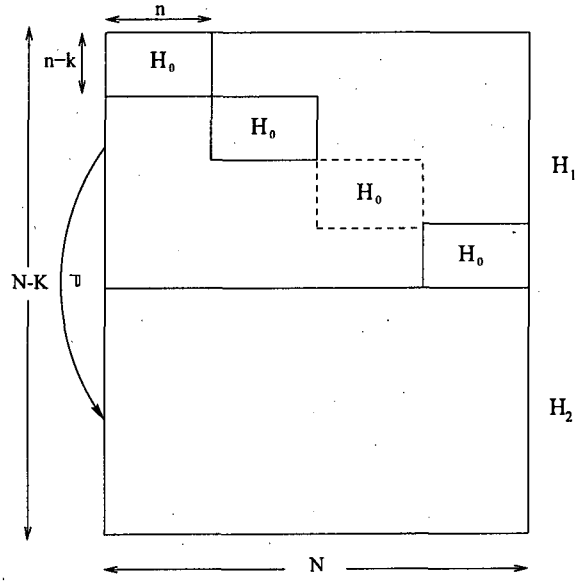


Figure 4.2: Parity-check matrix H of a GLDPC(N, K) code.

More explicitly, H_1 is a *block diagonal matrix* constructed by replacing the unity values on the main diagonal of an identity matrix with the parity-check matrix H_0 of the constituent code $C(n, k)$. All the other submatrices $H_2, ..., H_W$ are formed by the pseudo-random permutation of all the columns of the submatrix H_1 without interleaving the elements of the columns. This operation, formulated as $H_j = \pi_j(H_1)$ for $j = 2, ..., W$, indicates that the super-codes $C_2, ..., C_W$ are derived by randomly interleaving the super-code C_1. The codeword C of the GLDPC(N, K) code is the combined symbol of all super-codes $C_1, C_2, ..., C_W$. More specifically, the codeword C of a GLDPC code should be checked by the PCM H, constructed by concatenating the W number of PCMs $H_1, H_2, ..., H_W$ of the super-codes $C_1, C_2, ..., C_W$, therefore we have $C * H_1 = C * H_2 = ... = C * H_W = 0$.

If the rate of the constituent code $C(n, k)$ is $r = k/n$, the designed rate R of the

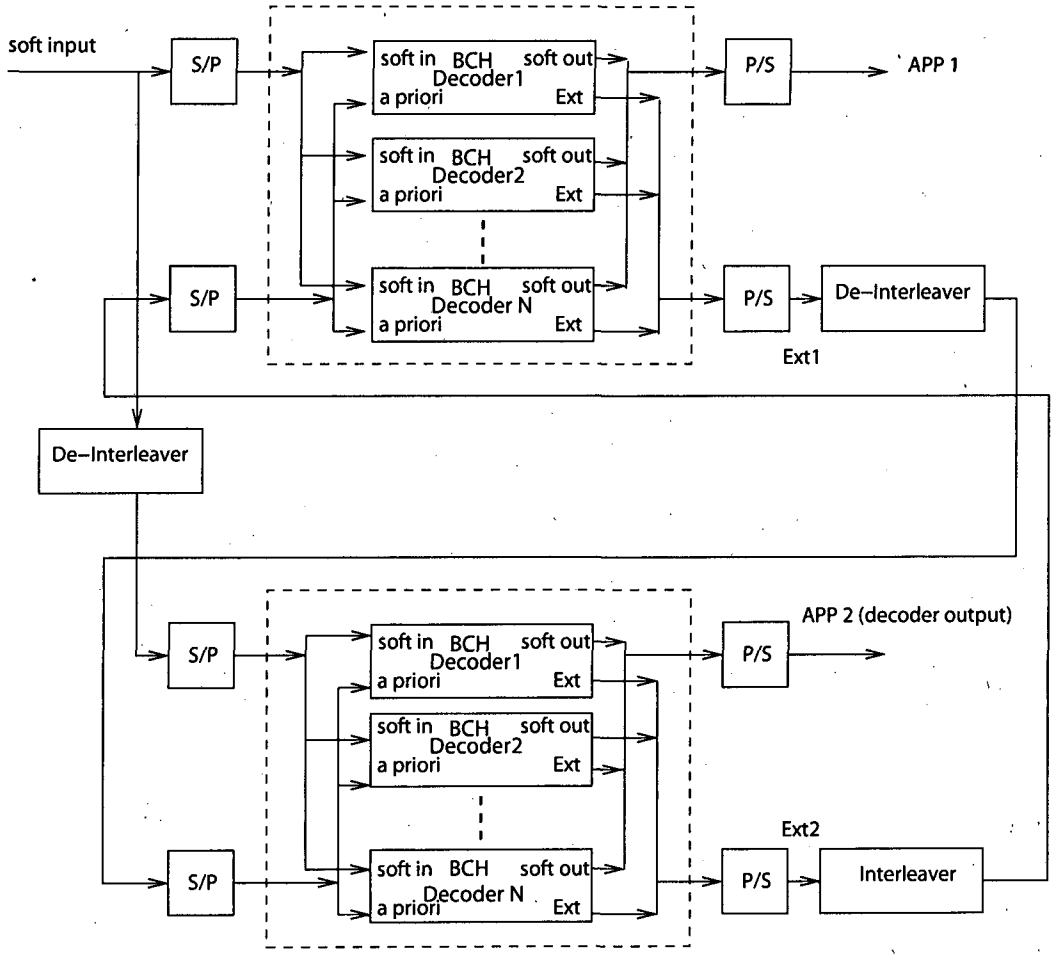


Figure 4.3: The schematic of the GLDPC decoder.

resultant GLDPC code is given by:

$$R =\geq 1 - W(1-r). \tag{4.8}$$

It has been shown in [124, 125] that GLDPC codes having $W = 2$ super-codes have the benefit of exhibiting the highest possible code rate of $R = 1 - W(1-r)$, a low-complexity decoder structure, as well as an excellent Bit Error Ratio (BER) performance, which are desirable properties. Thus, in our simulations we only consider GLDPC codes having $W = 2$ super-codes. The schematic of the GLDPC decoder having $W = 2$ is shown in Figure 4.3. More details can be found in [121], [124], [125].

4.2.4 EXIT Functions of Soft-bit Source Decoder

In our simulations, the wavelet video-encoded bitstreams of the 150-frame *"Miss_America"* video sequence, the 300-frame *Akiyo* sequence and the 300-frame *Mother-Daughter* video sequence were used as training sequences. The source statistics were evaluated by dividing the wavelet video encoded bitstream into a 3 bits/symbol sequence, since this 3-bit/source symbol representation was found to maximize the potential data rate reduction in terms of the associated entropy of the source symbols. The relative frequency of the 3-bit source symbol modelling the Probability Density Function (PDF) of the Dirac video-encoded bitstream, which was extracted from our simulations, is shown in Table 4.1.

Symbol	Probability	Symbol	Probability
000	0.2097020	100	0.1268470
001	0.0960067	101	0.0903356
010	0.0854893	110	0.0857518
011	0.1193480	111	0.1865200

Table 4.1: The relative frequency of the 3-bit source symbol modelling the Probability Density Function (PDF) of the Dirac video-encoded bitstream extracted from our simulations.

The parameters of the ISCD scheme investigated are summarized in Table 4.2. In our simulations, the 150-frame *"Miss_America"* video sequence was divided into fixed-length packets of 2280 bits for transmission over the uncorrelated Rayleigh fading channel. For the iterative receiver, four decoding iterations were executed. The results were evaluated by repeating the transmission of the 150-frame *"Miss_America"* video sequence 50 times, which was about $1.5 \cdot 10^7$ transmitted bits. The *extrinsic* information transfer functions of the outer softbit source decoder and inner GLDPC decoder are plotted in the form of EXIT charts in order to predict the achievable performance of the iterative receivers, as seen in Figure 4.4.

The EXIT charts recorded for the ISCD scheme and the corresponding BER performance are depicted in Figures 4.4 and 4.5, respectively. Observe from Figure 4.4 that since the outer softbit source decoder exhibits a slowly increasing EXIT function, it can provide only modest *extrinsic* information to the inner channel decoder, hence the overall performance can hardly be improved by the iterative process. This is confirmed by the

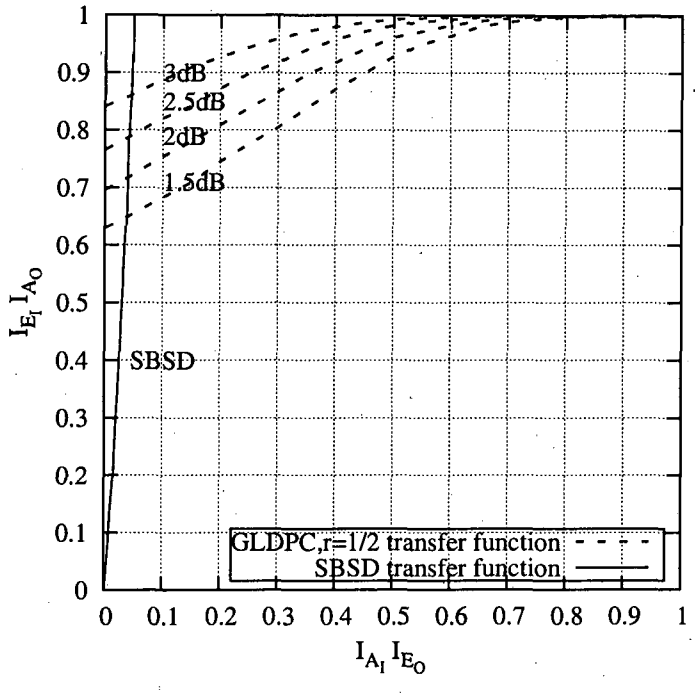


Figure 4.4: EXIT characteristics of the ISCD scheme designed for Dirac video transmission over the uncorrelated Rayleigh fading channel. The distribution of the source-encoded bitstream is presented in Table 4.1 using M=3bits/symbol.

Table 4.2: The parameters of the ISCD videophone scheme

Video codec	Dirac
Bit-Interleaver Π_1	2280 bits
GLDPC component codes	BCH(6,8,1)
Component code rate of GLDPC	$r = \frac{6}{8}$
Code rate of GLDPC	$\frac{1}{2}$
Number of inner GLDPC iterations	3
Number of outer iterations	2
Modulation	BPSK

BER performance seen in Figure 4.5, where only a modest iteration gain is attained.

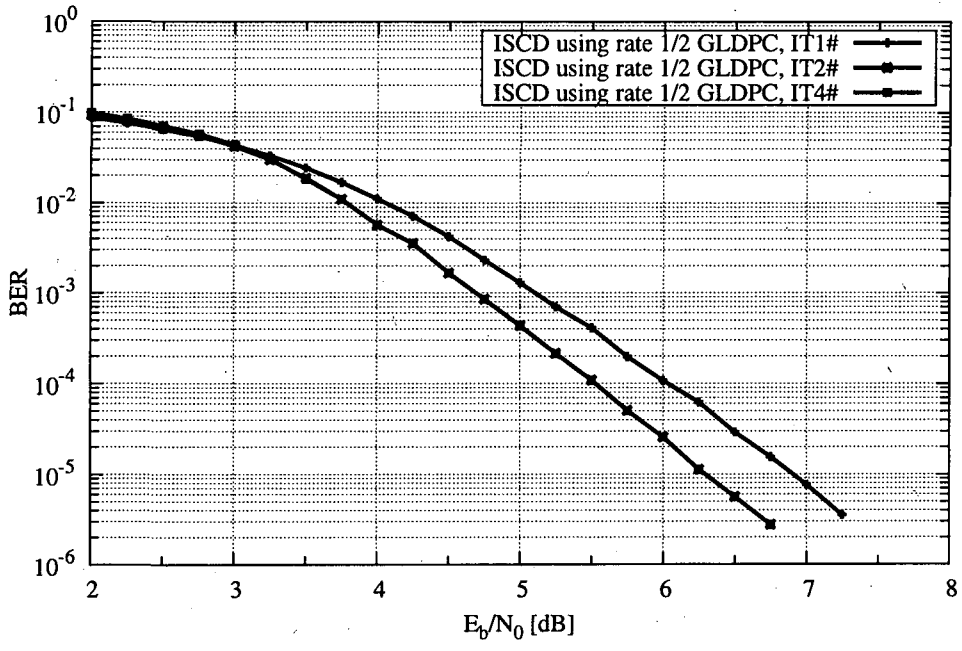


Figure 4.5: BER performance of the ISCD scheme designed for Dirac video transmission over the uncorrelated Rayleigh fading channel. The distribution of the source-encoded bitstream is presented in Table 4.1 for M=3bits/symbol.

4.2.5 Discussions

The aim of ISCD schemes is to extract as much *extrinsic* information L^{extr}_{SBSD}, L^{extr}_{CD} from both constituent decoders as possible for the sake of assisting each other. This leads to that the intersection of the component decoders' EXIT curves must be avoided for the sake of maintaining an open EXIT tunnel. If this cannot be achieved, their intersection should be located at the highest possible (L^{extr}_{SBSD}, L^{extr}_{CD}) point[1] in the EXIT chart.

It is worth noting that when using the Dirac video encoder there is little residual redundancy left in the source-encoded bitstream. Hence it is difficult to achieve further performance improvements with the aid of ISCD. The question arises then, whether there is any merit in intentionally imposing redundancy on the source-coded bitstream for the sake of constructing an attractive ISCD scheme. A further dilemma is, how to partition the limited total bit-rate budget between the source and the channel encoders, so that

[1]The particular ISCD's EXIT curves is considered to provide the "highest possible" performance if the sum of squares of the mutual information, i.e. $(I^{L^{extr}}_{SBSD})^2 + (I^{L^{extr}}_{CD})^2$ is maximized.

the channel-induced impairment of the recovered image is minimized. This problem will be resolved in following section.

4.3 Overcomplete Mapping-Assisted Iterative Source-Channel Decoding

As mentioned above, in state-of-the-art video encoders there is little residual redundancy left in the source-encoded bitstream and thus the achievable performance benefits of ISCD schemes are limited. In order to improve the attainable ISCD gains, we propose to find high performance over-complete mapping schemes, which artificially introduce redundancy into the source-encoded bitstreams. A specific manifestation of the artificial redundancy introduced by the over-complete mapping is that certain bit patterns become illegitimate and hence do not occur at all after mapping. For example, *Mapping 1* of Table 4.1 introduces the bit patterns {0000, 0001, 0010, 0011, 1100, 1101, 1110, 1111} with a probability of zero, i.e they are illegitimate. In other words, only eight out of the 16 possible symbols are actively used.

In order to investigate the dependence of the SBSD's performance on the potential redundancy inherent in the source encoded data, both the rate-1 mapping and the over-complete mapping were employed. For fair comparison, different over-complete mappings having the same rate of 3/4, but different Hamming distances were invoked in our simulations. Furthermore, in the following simulations we also use the relative frequency based modelling of the K-bit source-symbols of the wavelet video-encoded bitstream, as shown in Table 4.1.

4.3.1 Overcomplete Mapping Philosophy

In order to analyze the dependence of the SBSD's performance on the potential redundancy inherent in the source encoded data, the over-complete *Mappings 1* and *2* of rate 3/4 detailed in Table 4.3, were employed. The unsophisticated *Mapping 1* scheme of Table 4.3 simply concatenates a logical 0 or 1 to the source-encoded symbol for the sake of illustration. The design of the more sophisticated *Mapping 2* scheme will be detailed in Section 4.3.2.

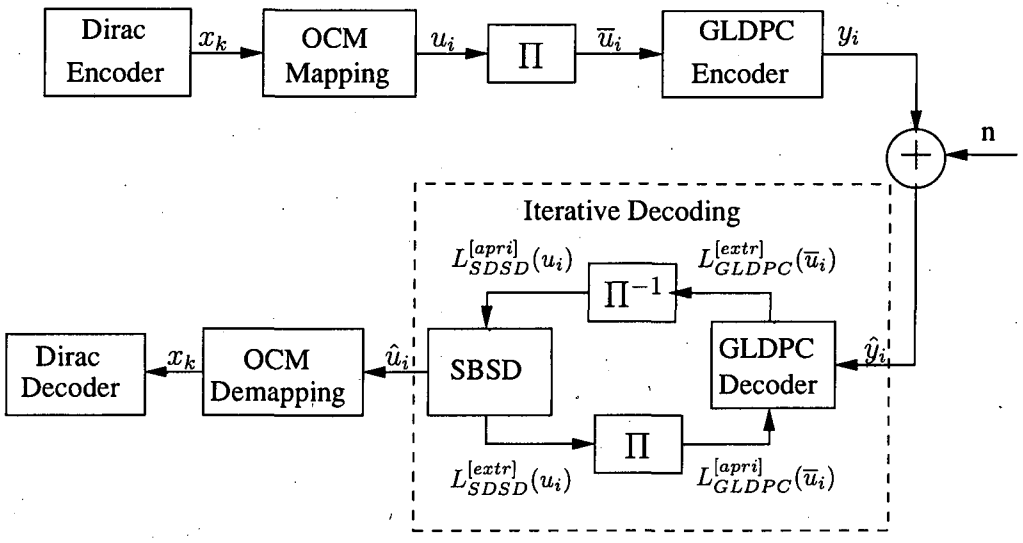


Figure 4.6: Overcomplete mapping-assisted iterative source-channel decoding

Table 4.3: source symbol mappings used in the simulations.

Symbol	Mapping 0	Mapping 1	Mapping 2
000	000	1000	0000
001	001	1001	1001
010	010	1010	1010
011	011	1011	0011
100	100	0100	1100
101	101	0101	0101
110	110	0110	0110
111	111	0111	1111

The block diagram of the overcomplete mapping-assisted iterative source-channel decoding scheme is shown in Figure 4.6. The *extrinsic* information transfer function characteristics of the mapping-assisted softbit source decoder are depicted in Figure 4.7. Observe from Figure 4.7 that the softbit source decoder using the *Mapping 0* of Table 4.3 has a negligible *extrinsic* information contribution. This is because the Mapping 0 does not increase the inherent redundancy in the encoded bitstream. Moreover, the *extrinsic* information contribution of the softbit source decoder is significantly increased, when using overcomplete mappings, i.e the *Mapping 1* and the *Mapping 2* of Table 4.3. In other words, the performance of SBSD strongly depends on the presence or absence of potential redundancy, which manifests itself in terms of the more uniform and hence uncorrelated

or less uniform distribution of the source-encoded K-bit data symbols.

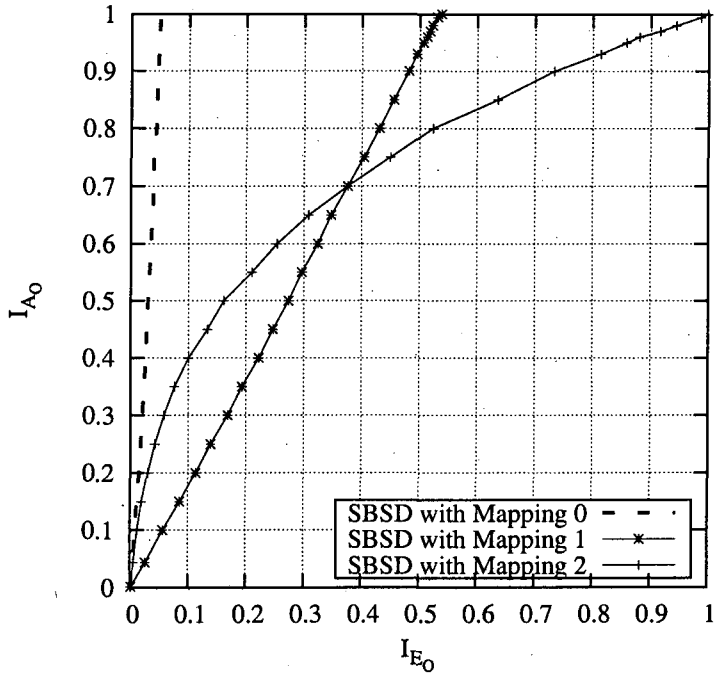


Figure 4.7: The *extrinsic* information transfer characteristics of the overcomplete mapping-assisted softbit source decoder. The mappings are detailed in Table 4.3

However, as evidenced by Figure 4.7, the artificially introduced redundancy imposed by these over-complete mappings cannot ensure that the EXIT curve of the SBSD reaches the point $(1, 1)$ at the upper-right corner of the EXIT chart, which would be necessary for achieving convergence to an infinitesimally low BER. In order to create over-complete mapping schemes for successful SBSD, in following section we will formulate *Lemma* 1. Furthermore, *Algorithm* 1 will outline how to construct near-capacity ISCD schemes, and as an additional result, the cardinality of EXIT-optimized mapping sets is also investigated.

4.3.2 Designing OCMs for Iterative Source-Channel Decoding

4.3.2.1 Design Criteria

Lemma 1 *Assume that the source data has a K-bit symbol probability distribution of $P = \{p_1, p_2, ..., p_{2^K}\}$, $p_i > 0 \; \forall i = 1, 2, ..., 2^K$. In the context of the ISCD scheme depicted in Figure 4.6, the necessary and sufficient condition for the EXIT curve of SBSD decoder*

to reach the point $(I_{SBSD}^{L^{apri}}, I_{SBSD}^{L^{extr}}) = (1,1)$ *is that the ISCD employs the over-complete mapping having a Hamming distance of at least* $d_H = 2$.

Proof: Let $OCM[N,K]$ denote the rate K/N OCM employed by the ISCD scheme of Figure 4.6. We first prove that if the EXIT curve of a SBSD decoder of Figure 4.6 reaches the point $(I_{SBSD}^{L^{apri}}, I_{SBSD}^{L^{extr}}) = (1,1)$, then the $OCM[N,K]$ must have a Hamming distance of at least $d_H = 2$.

Assuming that the $OCM[N,K]$ has a Hamming distance of $d_H = 1$, we have at least one $OCM[N,K]$-mapped symbol pair (u^m, u^n), $m \neq n$ having a Hamming distance of $d_H = 1$. This means that there exists a value $j \in \{1,...,N\}$, for which we have

$$\begin{aligned} u^m &= (c_1,...,c_{j-1},1,c_{j+1},...,c_N), \\ u^n &= (c_1,...,c_{j-1},0,c_{j+1},...,c_N). \end{aligned} \tag{4.9}$$

Let $\hat{u} = [\hat{u}_1, \hat{u}_2, ..., \hat{u}_N]$ be the *a priori* information sequence corresponding to the transmitted bit sequence $u^m = (c_1,...,c_{j-1},1,c_j,c_{j+1},...,c_N)$. Since $I_{SBSD}^{L^{apri}} = 1$, $u_j^m = 1$ and $I_{SBSD}^{L^{extr}} = 1$, we have

$$\begin{cases} \hat{u}_j = +\infty; \\ LLR(u_j) = +\infty. \end{cases}$$

Moreover, the assumption of $I_{SBSD}^{L^{apri}} = 1$ leads to

$$p\left(\hat{u}_\lambda | u_\lambda\right) = \begin{cases} 1 & \text{if } u_\lambda = c_\lambda; \\ 0 & \text{otherwise.} \end{cases} \tag{4.10}$$

Hence we have:

$$p\left(\hat{u}_\lambda^{[extr]} | u_\lambda^{[extr]}\right) = \prod_{\substack{i=1 \\ i \neq \lambda}}^{N} p\left(\hat{u}_i | u_i\right) = \begin{cases} 1 & \text{if } u_\lambda^{[extr]} = u_\lambda^{m,[extr]} \\ 0 & \text{otherwise.} \end{cases} \tag{4.11}$$

Combining (4.11) with (4.3) yields

$$LLR(u_j) = log \frac{p\left(u_j^{m,[extr]} | u_j^m = +1\right)}{p\left(u_j^{m,[extr]} | u_j^m = -1\right)}. \tag{4.12}$$

Note that from (4.9) we have $u_j^{m,[extr]} = u_j^{n,[extr]}$. Therefore, (4.12) can be written as

$$LLR(u_j) = log\frac{p\left(u_j^{m,[extr]}|u_j^m = +1\right)}{p\left(u_j^{n,[extr]}|u_j^m = -1\right)} = log\frac{p\left(u_j^{m,[extr]}|u_j^m = +1\right)}{p\left(u_j^{n,[extr]}|u_j^n = +1\right)} = log\frac{p_m}{p_m}. \quad (4.13)$$

Because we have $0 < p_m, p_n < 1$, from (4.13) we infer that $LLR(u_j) \neq +\infty$. This contradicts to Equation (4.10), and hence we have no other option but to conclude that the $OCM[N, K]$ has a Hamming distance of at least $d_H = 2$ □.

Let us now show that if the $OCM[N, K]$ has a Hamming distance of at least $d_H = 2$, then the EXIT curve of the SBSD decoder of Figure 4.6 reaches the point $(I_{SBSD}^{L^{apri}}, I_{SBSD}^{L^{extr}}) = (1, 1)$. Let $\hat{u} = [\hat{u}_1, \hat{u}_2, ..., \hat{u}_N]$ be the *a priori information* sequence at the input of the SBSD decoder and $u^0 = [u_1^0, u_1^0, ..., u_N^0]$ be the corresponding transmitted bit sequence. Using the same argument as above, we have

$$LLR(u_\lambda) = log\frac{p\left(u_\lambda^{0,[extr]}|u_\lambda = +1\right)}{p\left(u_\lambda^{0,[extr]}|u_\lambda = -1\right)}. \quad (4.14)$$

The lemma is proven if we can show that $LLR(u_\lambda)$ tends to $+\infty$, as $\hat{u}_\lambda \rightarrow +\infty$, and $LLR(u_\lambda)$ tends to $-\infty$, as $\hat{u}_\lambda \rightarrow -\infty$. Firstly, we consider the case, when $LLR(u_\lambda)$ approaches $+\infty$. Since $I_{SBSD}^{L^{apri}} = 1$, we have $I(\hat{u}, u^0) = 1$, which implies that u_λ^0 tends to the probability of unity, corresponding to a high confidence concerning its logical value, regardless whether it is a logical *one* or *zero*, if we have $\hat{u}_\lambda \rightarrow +\infty$. Equation (4.3) therefore becomes

$$LLR(u_\lambda) = log\frac{p\left(u_\lambda^{0,[extr]}|u_\lambda = u_\lambda^0\right)}{p\left(u_\lambda^{0,[extr]}|u_\lambda = \overline{u_\lambda^0}\right)} = log\frac{p\left(u^0\right)}{p\left(u_\lambda^{0,[extr]}|u_\lambda = \overline{u_\lambda^0}\right)}, \quad (4.15)$$

where $\overline{u_\lambda^0}$ is the inverted version of a binary value u_λ^0. Since the mapping defined in *Lemma 1* has a Hamming distance of at least $d_H = 2$, for a certain value of $u_\lambda^{[extr]}$, there is only one $u(\lambda)$ value, which results in $p\left(u_\lambda^{[extr]}|u_\lambda\right) \neq 0$. Therefore the denominator of Equation 4.15 must be 0, and the right-hand side of Equation 4.15 tends to $+\infty$, corresponding to a near-unity probability of $p(u^0) \approx 1$. A similar argument is valid, when $LLR(u_\lambda)$ tends to $-\infty$, which completes the proof of *Lemma 1* □

4.3.2.2 EXIT-Optimized Mapping Design

Let f: $X^K \rightarrow Y$ ($N > K$ and $Y \subseteq X^N$) be the rate K/N EXIT-Optimized mapping. Note that the overcomplete mapping f casting the set X^K to the set Y, namely

$$f(x_i) = y_i \qquad i = 1, 2, ..., 2^K,\ x_i \in X^K$$

has a Hamming distance of $d_H = 2$. Hence, the mapping φ: $X^K \rightarrow Y$ established by

$$\begin{cases} \varphi(x_1) = y_2, \\ \varphi(x_2) = y_1, \\ \varphi(x_k) = y_k \qquad \forall k \in \{3, ..., 2^P\} \end{cases}$$

is also a rate K/N EXIT-Optimized mapping. Therefore, in order to design a rate K/N EXIT-Optimized mapping $OCM[N, K]$ we can select a subset $Y \subseteq X^N$, which has a Hamming distance of $d_H = 2$ and a cardinality of $\| Y \|$, obeying $\| Y \| \geq \| X^K \| = 2^K$. Let us now demonstrate how to construct the above subset Y.

Let $X^{N-1} = \left\{ x = [x_1, x_2, ..., x_{N-1}] | x_i \in \{0, 1\} \right\} \subset X^N$ denote the set of legitimate natural binary representations of a $(N-1)$-bit symbol and $I(x) = \sum_{i=1}^{N} x_i \cdot 2^{N-i}$ be the index of symbol x. Since we have $K \leq (N-1)$, hence $\| X^{N-1} \| = 2^{N-1} \geq 2^K = \| X^K \|$ holds. For a binary string $U = [u_1, u_2, ..., u_{2^{N-1}}]$, we define the set X_A^N as follows:

$$X_A^N = \left\{ [x, u_{I(x)}] \mid x \in X^{N-1} \right\} = \left\{ [x_1, x_2, ..., x_{N-1}, u_{I(x_1, x_2, ..., x_{N-1})}] | \ x_i \in 0, 1 \right\}. \quad (4.16)$$

Observe that the set X_A^N has a cardinality of 2^{N-1}. Therefore, the above-mentioned set Y and the corresponding over-complete mapping scheme would be unambiguously determined, if we could find a suitable binary sequence $U = [u_1, ..., u_{2^{N-1}}]$ for which the set X_A^N has a Hamming distance of $d_H = 2$.

For $P = 2^{N-1}$, we have $\frac{P}{2} = 2^{N-2}$ and for $\forall k = 1, 2, ..., \left(\frac{P}{2} - 1\right)$, $[b_1, b_2, ..., b_{N-1}]$ is assumed to be the $(N-1)$-bit binary representation of k, yielding $k = \sum_{i=1}^{N-1} b_i \cdot 2^{N-1-i}$. Since $k < \frac{P}{2} = 2^{N-2}$, we have $b_1 = 0$, and therefore $k = \sum_{i=2}^{N-1} b_i \cdot 2^{N-1-i}$. Hence, $k + \frac{P}{2} =$

$k + 2^{N-2} = \sum_{i=2}^{N-1} b_i \cdot 2^{N-1-i} + 1 \cdot 2^{N-2}$. This means that $[1, b_2, ..., b_{N-1}]$ or $[\overline{b_1}, b_2, ..., b_{N-1}]$ is an $(N-1)$-bit binary representation of $\left(k + \frac{P}{2}\right)$, where $\overline{b_1}$ denotes the inversion of a binary value b_1. Therefore, if a binary sequence $[v_1, v_2, ..., v_{N-2}]$ ensures that the set

$$Z_A^{N-1} = \{[b_1, ..., b_{N-2}, v_{I(b_1,...,b_{N-2})}]\}$$

has a Hamming distance of two, then the sets

$$Z_B^{N-1} = \{[\overline{b_1}, ..., b_{N-2}, v_{I(b_1,...,b_{N-2})}]\} \quad \text{and} \quad Z_C^{N-1} = \{[\overline{b_1}, ..., b_{N-2}, \overline{v_{I(b_1,...,b_{N-2})}}]\}$$

also have a Hamming distance of two. However, the union of Z_A^{N-1} and Z_B^{N-1} denoted by $\left(Z_A^{N-1} \cup Z_B^{N-1}\right)$ has a Hamming distance of one, while the union of Z_A^{N-1} and Z_C^{N-1} denoted by $\left(Z_A^{N-1} \cup Z_C^{N-1}\right)$ has a Hamming distance of two.

Hence the binary sequence $U = [u_1, u_2, ..., u_{2^{N-1}}] = \left\{[u_{I(x)}], x \in X^{N-1}\right\}$ used for the set of N-bit symbols X_A^N is given by constructing the subsequence $U^1 = [u_1, u_2, ..., u_{2^{N-2}}]$ used for the $(N-1)$-bit symbol subset X_A^{N-1}, which is then concatenated it with its complement $\overline{U^1} = [\overline{u_1}, ..., \overline{u_{2^{N-2}}}]$. This operation is repeated until the subset X_A^N has only two elements. Finally, we concatenate the resultant bit for constructing the binary subsequence, which is used for the subset X_A^1.

In fact, our algorithm simply divides each segment $[0 \ldots N-1]$ into two identical-length sub-segments of $[0 \ldots \frac{N-1}{2}]$ and $[\frac{N-1}{2} + 1 \ldots N-1]$, and then sets the *start value* of the second sub-segment $u_{\frac{N-1}{2}+1}$ equal to the inverted version of the first sub-segment's *start value*. More specifically, in our algorithm, the above-mentioned operations have been carried out with the aid of three variables: the start position χ_{Start} and the end position χ_{Stop} of the current subsequence represented by the binary sequence $U = [u_0, ..., u_{2^{N-1}}]$, as well as the bit u_{Start}, indicating that the two-bit string "$u_{Start}\overline{u_{Start}}$" is chosen at the final step of the algorithm.

A. Algorithm

Algorithm 1 ***the prefix-bit selection algorithm;***

Input: N, u_{Start};

Output: *A binary sequence* $u = \{u_0, ..., u_{2^{N-1}}\}$;

Step 1: $\chi_{Start} = 0$, $\chi_{Stop} = 2^N - 1$,

Step 2: *For the set* $\{[\chi_{Start} \cdots \chi_{Stop}], u_{Start}\}$ *do*

{

Step 3: *Set* $\chi_{Mid} = \dfrac{\chi_{Stop} - \chi_{Start} + 1}{2}$, $\quad u_{Mid} = \overline{u_{Start}}$

where $\overline{u_{Start}}$ *is the inverse of the binary value* u_{Start}

Step 4: *if* $(\chi_{Start} = \chi_{Stop})$ *then goto* ***Step 7;***

Step 5: *Repeat* ***Step 2*** *for the set* $\{[\chi_{Start} \cdots \chi_{Mid} - 1], u_{Start}\}$;

Step 6: *Repeat* ***Step 2*** *for the set* $\{[\chi_{Mid} \cdots \chi_{Stop}], u_{Mid}\}$;

}

Step 7: *Stop.*

As an example, Table 4.4 shows the results of applying Algorithm 1 to $u_{Start} = 0$ for various N values. For the sake of illustrating the dependence of the results of Algorithm 1 on the difference $K = (\chi_{Stop} - \chi_{Start})$ as well as on the value of u_{Start}, the resultant binary sequences of Table 4.4 are partitioned into two identical-length subsequences. Observe that the second subsequence u_{Mid} printed in *italic font* can be obtained by complementing the first subsequence u_{Start} printed in *standard font*. Furthermore, the sequences u_{Start} and $u_{Start+K}$ calculated by applying Algorithm 1 to the set $\{[\chi_{Start} \cdots \chi_{Stop}], u_{Start}\}$ and to $\{[\chi_{Start+K}, \chi_{Stop+K}], u_{Start}\}$ are identical. In other words, the result of Algorithm 1 only depends on the values of $K = (\chi_{Stop} - \chi_{Start})$ and u_{Start}.

Table 4.4: Prefix bit sequence u as a function of N.

Symbol size N	Bit sequence s_N
1	01
2	01 *10*
3	0110 *1001*
4	01101001 *10010110*
5	0110100110010110 *1001011001101001*

B. Example of constructing an EXIT-optimized mapping scheme

- *Example 1:* constructing an EXIT-optimized over-complete rate-3/4 mapping scheme;

A rate-3/4 over-complete mapping scheme f: $X^3 \rightarrow X^4$, namely Mapping 2 of Table 4.3, can be constructed from the set $X^N = X^{4-1} = X^3$ as follows:

Stage 1 Set $N = 2^3 = 8$, $u_1 = 0$.

Stage 2 Using *Step 2, 3, 4, 5, 6* of **Algorithm 1** to divide the segment $[1 \ldots 8]$ into $[1 \ldots 4]$ and $[5 \ldots 8]$, and sets $u_5 = \overline{u_1} = 1$;

Stage 3 Using *Step 2, 3, 4, 5, 6* of **Algorithm 1** to divide the segment $[1 \ldots 4]$ into $[1, 2]$ and $[3, 4]$, and sets $u_3 = \overline{u_1} = 1$;

Stage 4 Using *Step 2, 3, 4, 5, 6* of **Algorithm 1** to divide the segment $[5 \ldots 8]$ into $[5, 6]$ and $[7, 8]$, and sets $u_7 = \overline{u_5} = 0$;

Stage 5 for sub-segment $[1, 2]$ set $u_2 = \overline{u_1} = 1$;

Stage 6 for sub-segment $[3, 4]$ set $u_4 = \overline{u_3} = 0$;

Stage 7 for sub-segment $[5, 6]$ set $u_6 = \overline{u_5} = 0$;

Stage 8 for sub-segment $[7, 8]$ set $u_8 = \overline{u_7} = 1$;

Stage 9 Concatenates the resultant bits, yielding $s^3 = \{u_1, u_2, u_3, u_4, u_5, u_6, u_7, u_8\}$ = "01101001". This result is presented in the fourth row, second column of Table 4.4;

Stage 10 Generates the natural binary representations of the $(N - 1) = 3$-bit-symbol set X^3, as seen in the second column of Table 4.3;

Stage 11 Inserts the first binary symbol u_1 = "0" of the sequence s_3 at the beginning of the first entry, namely x_1 = "000" of X^3 in Table 4.3 to form the first output symbol z_1 = "0000" of Mapping 2, the second binary symbol u_2 = "1" of the sequence s_3 at the beginning of the second entry x_2 = "001" of X^3 to form the second output symbol z_2 = "1001" of Mapping 2, etc. The result is shown in the fourth column of Table 4.3;

Stage 12 The rate-3/4 EXIT-optimized over-complete mapping f is given by Table 4.3 as

$$f(x_i) = z_i \qquad i = 0, ..., 7.$$

- *Example 2:* constructing an EXIT-optimized over-complete rate-3/5 mapping scheme. The subset Y of a rate-3/5 EXIT-optimized mapping scheme can be constructed from the set $X^N = X^{5-1} = X^4$ by following the eight first stages of **Example 1**.

More specifically, the binary symbols of the resultant sequence s_4 = "0110100110010110" of Table 4.4 are inserted at the beginning of the corresponding entries of $X^4 = \{0000, 0001, 0010, 0011, 0100, 0101, 0110, 0111, 1000, 1001, 1010, 1011, 1100, 1101, 1110, 1111\}$ not shown in the Tables to form the subset $Y = \{00000, 10001, 10010, 00011, 10100, 00101, 00110, 10111, 11000, 01001, 01010, 11011, 01100, 11101, 11110, 01111\} = \{y_0, y_1, ..., y_{15}\}$. A subset of the rate-3/5 mapping scheme $F = \{f_j | f_j : X^3 \rightarrow X^5\}$ is given by

$$f_j(x_i) = y_{i+j} \qquad i = 0, ..., 7,\ j \in \{1, ..., 7\}.$$

C. Lemma

Lemma 2 *The set X_A^N constructed above has a minimum Hamming distance of $d_H = 2$.*

Proof: Assume that X_A^N has a Hamming distance of $d_H = 1$. This implies that there exists an index pair $i, j \in I(X^N)$, $i = \sum_{k=0}^{N-1} x_{i,k} \cdot 2^{N-1-k}$, $j = \sum_{k=0}^{N-1} x_{j,k} \cdot 2^{N-1-k}$, $i \neq j$ for which

$$d_h[(u(i), x_{i,1}, ..., x_{i,N}), (u(j), x_{j,1}, ..., x_{j,N})] = 1.$$

If we have $i = j$, then $u(i) = u(j)$, and hence

$$d_h[(u(i), x_{i,1}, ..., x_{i,N}), (u(j), x_{j,1}, ..., x_{j,N})] = 0.$$

Therefore we have $i \neq j$, which means that $d_h[(x_{i,1}, ..., x_{i,N}), (x_{j,1}, ..., x_{j,N})] \geqslant 1$. Combining this with the assumption of $d_h[(u(i), x_{i,1}, ..., x_{i,N}), (u(j), x_{j,1}, ..., x_{j,N})] = 1$, we have

$$u(i) = u(j). \tag{4.17}$$

Thus $d_h[(x_{i,1}, ..., x_{i,N}), (x_{j,1}, ..., x_{j,N})] = 1$, which means that there exists a value $k \in \{0, ..., N-1\}$, for which we have

$$\begin{aligned} i &= (c_0, ..., c_{k-1}, c_k, c_{k+1}, ..., c_N), \\ j &= (c_0, ..., c_{k-1}, \overline{c_k}, c_{k+1}, ..., c_N). \end{aligned} \tag{4.18}$$

Observe that i and j belong to the same set $[\chi_{(Start,k)}, \chi_{(Stop,k)}, u_{Start_k}]$ after carrying out $(k-1)$ partitioning steps in Algorithm 1. At the time instant k we have $i \in [\chi_{(Start,k)}, \chi_{(Mid,k)}, u_{Start_k}]$, and $j \in [\chi_{(Mid,k)}, \chi_{(Stop,k)}, u_{Mid_k}]$. As mentioned above, if we

have the prefix sequences $u_{Start_k} = [u_{(Start_k,0)}, ...u_{(Start_k,K)}]$ and $u_{Mid_k} = [u_{(Mid_k,0)}, ...u_{(Mid_k,K)}]$ generated by Algorithm 1 from the set $[\chi_{(Start,k)}, \chi_{(Mid,k)}, u_{Start_k}]$ and from the set $[\chi_{(Mid,k)}, \chi_{(Stop,k)}, u_{Mid_k}]$, respectively, then $u_{(Start_k,p)} = \overline{u_{(Mid_k,p)}}$, $\forall p \in \{0, ..., K\}$. It may be readily seen from Equation 4.18 that the index of i in u_{Start_k} and the index of j in u_{Mid_k} are the same, namely $\sum_{p=k+1}^{N-1} c_p \cdot 2^{N-1-p}$ and hence we have $u(i) = \overline{u(j)}$. This contradicts to Equation 4.17, and hence we have no other option but to conclude that X_A^N has a Hamming distance of at least $d_H = 2$ □

C. Cardinality of an EXIT-Optimized Mapping

Lemma 3 *The cardinality*[2] *of the EXIT-optimized mapping set* f: $X^K \rightarrow X^N$ $(K < N)$ *is equal to* $\left(2 \cdot \frac{2^{N-1}!}{(2^{N-1}-2^K)!}\right)$.

Proof: Observe that the set X_A^{N-1} constructed by using Algorithm 1 for $u_{Start} = 0$ and exhibiting a Hamming distance of at least $d_H = 2$, has the cardinality of $\|X_A^{N-1}\| = 2^{N-1}$. Hence, all the possible injective mappings f: $X^K \rightarrow X_A^{N-1}$ are EXIT-optimized mappings. It is clear that the number of possible injective mapping[3] schemes f: $X^K \rightarrow X_A^{N-1}$ is equal to the number of possible ways of obtaining an ordered subset of 2^K elements from a set of 2^{N-1} elements. In other words, the cardinality of the injective mapping set $F = \{f : X^K \rightarrow X_A^{N-1}\}$ is equal to the number of possible permutations $P_{2^{N-1}}^{2^K} = \frac{2^{N-1}!}{(2^{N-1}-2^K)!}$, where "!" denotes the factorial operator. Note that X_B^{N-1} constructed by using Algorithm 1 for $u_{Start} = 1$ also has a Hamming distance of at least $d_H = 2$ and X_A^{N-1}, X_B^{N-1} are non-overlapping sets obeying $X^N = X_A^{N-1} \cup X_B^{N-1}$. Using the same argument allows us to quantify the number of possible injective mapping schemes $\psi : X^K \rightarrow X_B^{N-1}$, which is also equal to the number of the possible permutations $P_{2^{N-1}}^{2^K} = \frac{2^{N-1}!}{(2^{N-1}-2^K)!}$, hence the number of over-complete mapping schemes f: $X^K \rightarrow X^N$ having a minimum Hamming distance of $d_H = 2$ is $2 \cdot \frac{2^{N-1}!}{(2^{N-1}-2^K)!}$, which concludes the proof of the lemma □

[2]The cardinality of a set is the number of elements in the set.

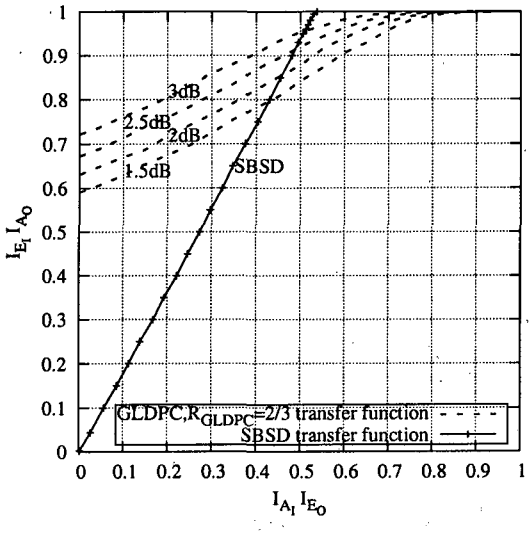
[3]The injective mapping is a one-to-one mapping.

4.3.3 Simulation Results

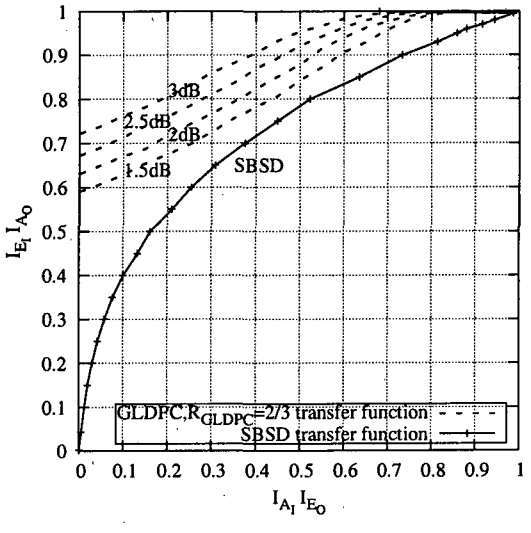
In this section, the attainable performance of the ISCD schemes of Figure 4.1 and Figure 4.6 are investigated for transmission over an uncorrelated Rayleigh fading channel. The various system parameters are summarized in Table 4.5.

Table 4.5: The simulation parameters

Parameters	System 1	System 2	System 3
Overall code-rate	0.5	0.5	0.5
Video codec	Dirac	Dirac	Dirac
Mapping rate	1	3/4	3/4
		(*Mapping 1*)	(*Mapping 2*)
bit-Interleaver Π_1	2280 bits	3648 bits	3648 bits
GLDPC code-rate	1/2	2/3	2/3
Constituent codes	BCH(8,6,1)	BCH(20,18,1)	BCH(20,18,1)
Modulation	BPSK	BPSK	BPSK



a) System 2 of Table 4.5



b) System 3 of Table 4.5

Figure 4.8: EXIT characteristics of systems specified in Table 4.5 and designed for transmission over the uncorrelated Rayleigh fading channel.

The convergence behaviour of these systems is characterized in Figure 4.4 and Figure 4.8. Observe by comparing Figure 4.8.a and 4.8.b that both the amount of redun-

dancy imposed by the over-complete mapping as well as the Hamming distance of the over-complete mapping play an important role. More specifically, the intersection of the EXIT-curves of System 2 using *Mapping 1* of Table 4.3 is significantly closer to the $(1, 1)$ point of the EXIT-chart and consequently to the E_b/N_0 value of infinitesimally low BER than that of System 1. Furthermore, at the same amount of redundancy, i.e. mapping rate, the EXIT-curve of System 3 indeed reaches the top-right corner of the EXIT chart, whereas that of System 2 fails to do so, as evidenced by Figure 4.8. Hence, although not shown here, the iterative decoding gain of System 1 erodes beyond two iterations, while that of System 2 and System 3 improves further upon iterating four times.

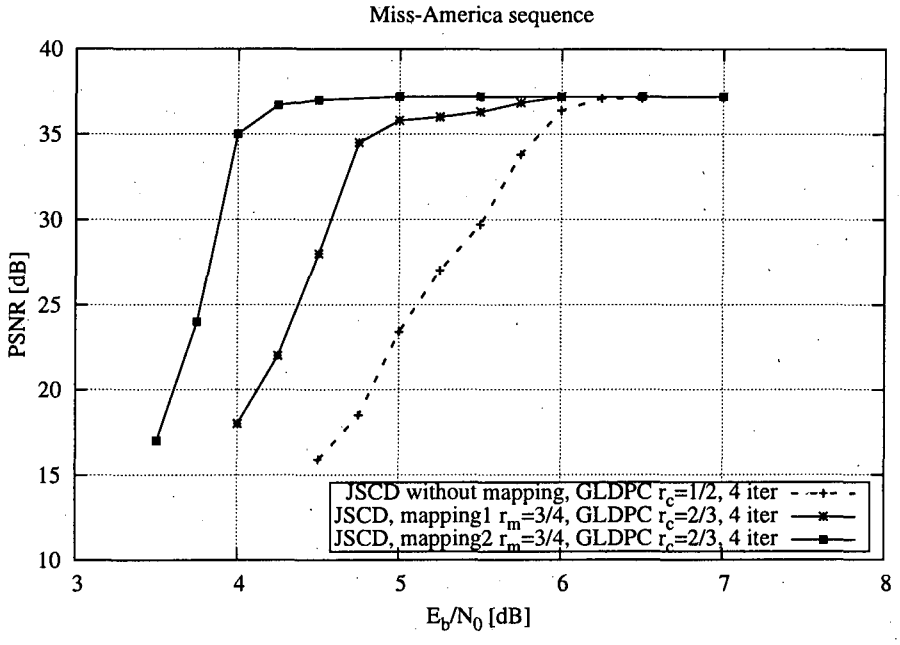
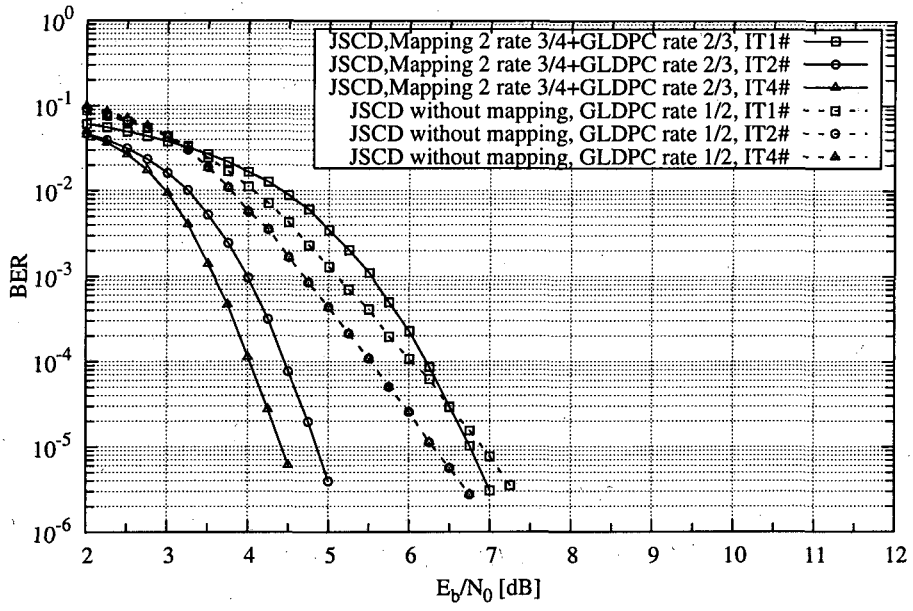
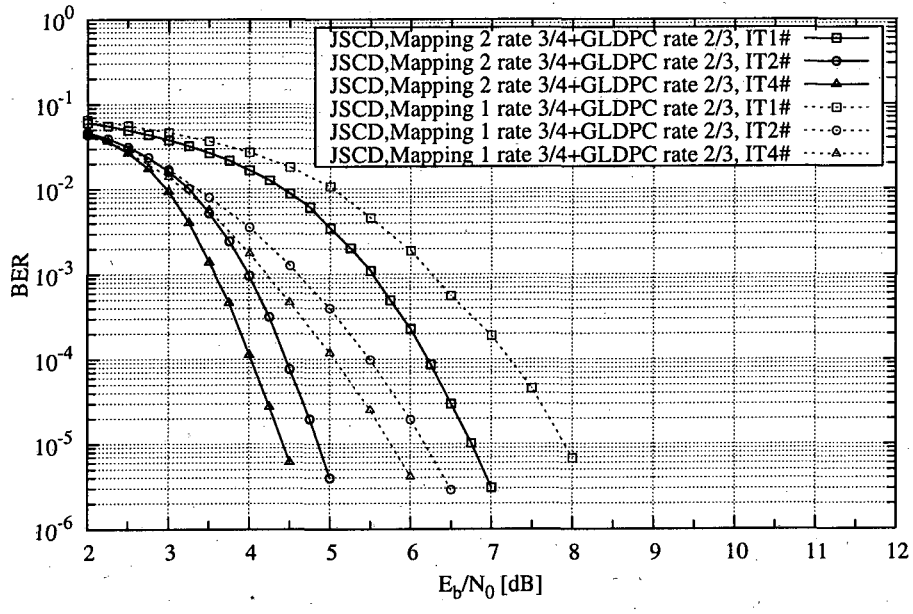


Figure 4.9: Comparison of the achievable video PSNR using system summarized in Table 4.5.

The associated BER performances are shown in Figure 4.10 for an overall code-rate of 0.5. More specifically, System 3 outperformed System 2 by about $1.75dB$, and it has an approximately $2dB$ gain also in comparison to System 1 for $I = 4$ iterations. The video PSNR performances of the systems summarized in Table 4.5 are depicted in Figure 4.9, while the achievable subjective video qualities at different channel E_b/N_0 values can be seen in Figure 4.11. Observe from Figures 4.10,4.9 and Figure 4.11 that System 3 attains a near-unimpaired video reconstruction quality for a channel E_b/N_0 value of 4dB, although at this point the overall BER is about $\times 10^{-4}$. Furthermore, at the same channel E_b/N_0



(a) System 1 and System 3 of Table 4.5.



(b) System 2 and System 3 of Table 4.5.

Figure 4.10: BER performance of the systems summarized in Table 4.5 and designed for transmission over the uncorrelated Rayleigh fading channel.

value of 4dB the video reconstruction quality of System 2 is significantly lower than that of System 3, although at this point the overall BER is about 2×10^{-3}.

Figure 4.11: Subjective video quality of the systems summarized in Table 4.5 for transmission over the uncorrelated Rayleigh fading channel at $E_b/N_0 = 4.5dB$: the second video frame of the *"Miss_America"* sequence extracted from System 1 (left); System 2 (middle); System 3 (right).

4.4 Characterization of OCM-SSD

4.4.1 Computational Complexity Analysis

The complexity of the proposed system may be evaluated for example by evaluating the number of multiply-add operations required to compute the *extrinsic* information of a single source symbol. In our computational complexity evaluation of the SBSD, the number of multiply-add operations required to compute the *extrinsic* information of a single K-bit source symbol is used as the complexity metric.

We first consider the complexity of the SBSD technique. Since every symbol has K bits, $u_j^{[extr]}$ has $(K-1)$ bits and hence we have a total of $2^{(K-1)}$ different $u_\tau^{[extr]}$ values. Observe that $(K-1)$ multiplications, $(K-1)$ additions and an exponentiation is required to compute $p\left(\hat{u}_i^{[extr]}|u_i^{[extr]}\right)$ of Equation 4.4. This leads to a total of $(K-1)\cdot 2^{(K-1)} + 2^{(K-1)} = K \cdot 2^{(K-1)}$ multiplications, $(K-1)\cdot 2^{(K-1)} + 2^{(K-1)} = K \cdot 2^{(K-1)}$ additions and $2^{(K-1)}$ exponentiations necessitated for calculating the numerator of Equation 4.3. Note that the same number of operations is required for calculating the denominator of Equation 4.3. Therefore, the total number of operations necessitated by calculating the *extrinsic* information can be summarized, as seen in Table 4.6. Let us now consider the complexity of the rate-K/N over-complete mapping-aided SBSD technique. Note that

Table 4.6: Total number of operations required to compute the *extrinsic* information using the K-bit source symbol SBSD technique.

Operator	For a single bit	For a transmitted symbol
Multiplications	$2 \cdot K \cdot 2^{K-1} = K \cdot 2^K$	$K \cdot K \cdot 2^K = K^2 \cdot 2^K$
Additions	$2 \cdot K \cdot 2^{K-1} = K \cdot 2^K$	$K \cdot K \cdot 2^K = K^2 \cdot 2^K$
Exponentiations	$2 \cdot 2^{K-1} = 2^K$	$K \cdot 2^K$
Logarithm	1	K

the mapped symbol having a length of N and the rate-K/N over-complete mapping-aided softbit source decoding can be carried out by modifying the N-bit source symbol based SBSD algorithm. More specifically, since $u_j^{[extr]}$ has $(N-1)$ bits, the calculation of the expression $p(\hat{u}_i^{[extr]}|u_i^{[extr]})$ in Equation 4.4 requires $(N-1)$ multiplications, $(N-1)$ additions and an exponentiation. Moreover, only 2^K encoded-symbols have non-zero symbol probabilities, hence 2^{K-1} different $u_j^{[extr]}$ values are necessitated for calculating the expression $\sum_{u_i^{[extr]}} p(u_i^{[extr]}|u_i = +1).p(\hat{u}_i^{[extr]}|u_i^{[extr]})$. This leads to the total number of operations required for calculating the numerator of Equation 4.3 calculated as

$$\begin{aligned} &\text{Multiplications:} && (N-1) \cdot 2^{K-1} + 2^{K-1} = N \cdot 2^{K-1}; \\ &\text{Additions:} && (N-1) \cdot 2^{K-1} + 2^{K-1} = N \cdot 2^{K-1}; \\ &\text{Exponentiations:} && 2^{K-1}. \end{aligned}$$

Recall that the same number of operations is required for calculating the denominator of Equation 4.3. Therefore, the total number of operations necessitated by calculating the *extrinsic* information can be summarized as seen in Table 4.7.

Table 4.7: Total number of operations required to compute the *extrinsic* information using the mapping[N,K].

Operator	For a single bit	For a transmitted symbol
Multiplications	$2 \cdot N \cdot 2^{K-1} = N \cdot 2^K$	$N \cdot N \cdot 2^K = N^2 \cdot 2^K$
Additions	$2 \cdot N \cdot 2^{K-1} = N \cdot 2^K$	$N \cdot N \cdot 2^K = N^2 \cdot 2^K$
Exponentiations	$2 \cdot 2^{K-1} = 2^K$	$N \cdot 2^K$
Logarithm	1	N

The results of Tables 4.6 and 4.7 reveal that the computational complexity of the

rate-K/N over-complete mapping-aided SBSD is slightly higher than that of the K-bit source symbol SBSD and at the rate of K/N the computational complexity of the over-complete mapping-aided SBSD employing rate K/N mapping is significantly lower than that employing rate $2K/2N$ mapping. This property can be exploited to design wireless video systems, trading off the achievable system performance against the implementational complexity.

4.4.2 Performance versus the Code-Rate

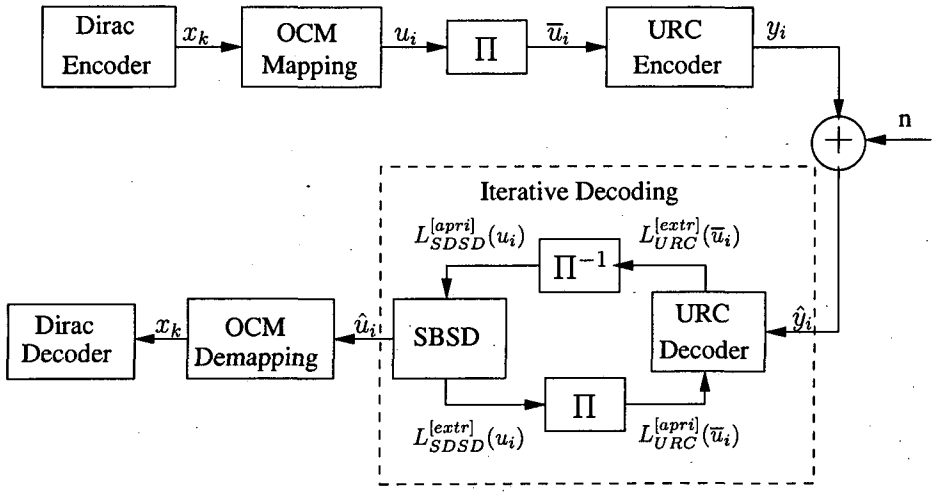


Figure 4.12: System model of the video telephony scheme employing OCM-SSD, which is referred to as ECOBC Encoder/Decoder, as outer code and unity-rate code as inner code.

For the sake of investigating the performance of the OCM-SSD, this section considers a turbo detection scheme employing OCMs having a variety of coding rates. More explicitly, in our simulations we invoke rate K/N OCMs having (N,K) parameters of $[(10,2),(10,3),(7,3),(8,4),(7,4),(10,6),(9,6),(8,6),(10,7),(10,9)]$, which correspond to the coding rates of $[0.2, 0.3, 0.429, 0.5, 0.545, 0.6, 0.66, 0.75, 0.7, 0.9]$. The OCM-SSD, employed as our outer codec, is serially concatenated and iteratively decoded by exchanging extrinsic information with the URC [117] of Figure 4.13 used as the inner codec. The block diagram of the proposed scheme is shown in Figure 4.12. In order to make a fair comparison, the transmission of a random K-bit-symbol sequence having a uniform probability distribution is considered in all of our simulations.

To elaborate a little further, Figure 4.13 portrays the URC employed in our turbo detection scheme, where D is a shift register stage and $\oplus$ represents the modulo-2 operation.

Table 4.8: The first five OCMs[N,K] used in the simulations of Section 4.4.2.

4-bit source symbol	The mapped symbols of	
	OCM[8,4]	OCM[7,4]
0000	00000000	0000000
0001	00001111	0000111
0010	00110011	0011001
0011	00111100	0011110
0100	01010101	0101010
0101	01011010	0101101
0110	01100110	0110011
0111	01101001	0110100
1000	10010110	1001011
1001	10011001	1001100
1010	10100101	1010010
1011	10101010	1010101
1100	11000011	1100001
1101	11001100	1100110
1110	11110000	1111000
1111	11111111	1111111

3-bit source symbol	The mapped symbols of	
	OCM[7,3]	OCM[10,3]
000	0000000	0000000000
001	0001111	0000011111
010	0110011	0011100011
011	0111100	0011111100
100	1010101	1100100101
101	1011010	1100111010
110	1100110	1111000110
111	1101001	1111011001

2-bit source symbol	The mapped symbols of
	OCM[10,2]
00	0000000000
01	0000111111
10	0111000111
11	0111111000

Figure 4.13(b) shows the trellis diagram of the URC considered. The trellis transitions are denoted by c_i/x_i, where c_i denotes the input bit of the URC at time instant i, while x_i indicates the corresponding URC's output bit.

Let us first investigate the EXIT chart [106] of the proposed scheme. Our aim is to maintain an open EXIT chart tunnel, which results in iterative decoding convergence to an infinitesimally low probability of error. All EXIT functions seen in Figure 4.14 were obtained for the uncorrelated Rayleigh fading channel by using the histogram-based LLR measurements of [106].

As an additional result, Figure 4.15 and Figure 4.16 provide a plot of the attainable capacity of the scheme considered in comparison to the channel capacity of the BPSK-modulated Rayleigh fading channel. As seen in Figure 4.16, the OCM-SSD gain peaks near the coding rate of 1/2.

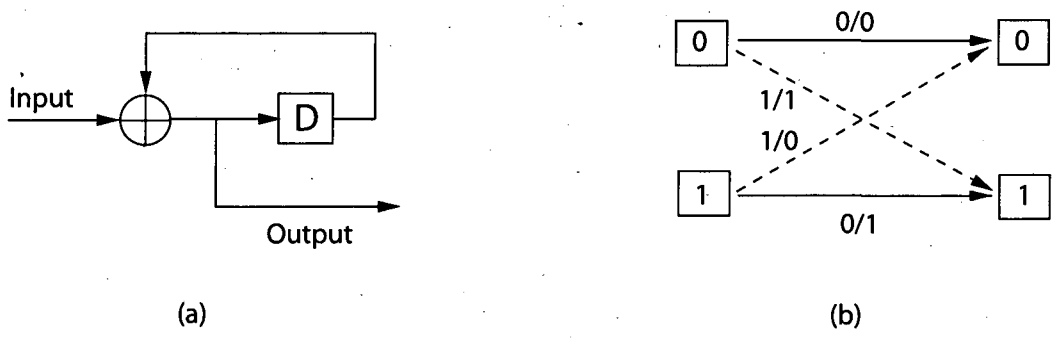


Figure 4.13: (a) Shift register representation of the unity rate memory-1 code having a generator polynomial of $1 + D$ and (b) its trellis transitions.

4.4.3 Performance versus Hamming Distance of the OCM.

In this section, the effects of the OCM's Hamming distance on the OCM-SSD's performance are investigated. This is carried out by investigating the various EXIT functions of the OCM-SSD decoders having the same mapping-rate, but different Hamming distances *and/or* source symbol lengths. More explicitly, in our simulations we invoke rate 1/2 OCMs having input/output symbol lengths (N,K) of $[(6,3),(8,4),(12,6),(16,8)]$, which have the corresponding Hamming distances of $[3,4,4,5]$. Note that in order to make a fair comparison, a random K-bit-symbol sequence having a uniform probability distribution is transmitted in all of our simulations.

Let us now investigate the EXIT characteristics of the OCM-SSD decoders having the same mapping-rate, but different Hamming distances and/or source symbol lengths. The four lower curves of Figure 4.17 show the inverted EXIT functions obtained for the OCM-SSD decoders employing rate-1/2 mappings, but different Hamming distances and/or source symbol lengths. These EXIT functions were obtained by simulating the OCM-SSD encoding and decoding of 400 frames of 10 000 randomly generated bits with a uniform K-bit symbol probability distribution. As seen in Figure 4.17, the area property of EXIT charts[4] [126] is valid for all the inverted EXIT functions of the OCM-SSD. Moreover, the extrinsic information contribution of the OCM-SSD decoder employing OCM having a higher Hamming distance becomes more effective than their lower Hamming distance counterpart for I_A values in excess of 0.5. This is achieved at the cost of increasing the computational complexity of the OCM-SSD decoders, as mentioned in Section 4.4.1.

[4]The area property of EXIT charts implies that the area under the outer code's EXIT curve is equal to the code-rate. Hence, identical-rate codes exhibit an identical area, despite their different shapes.

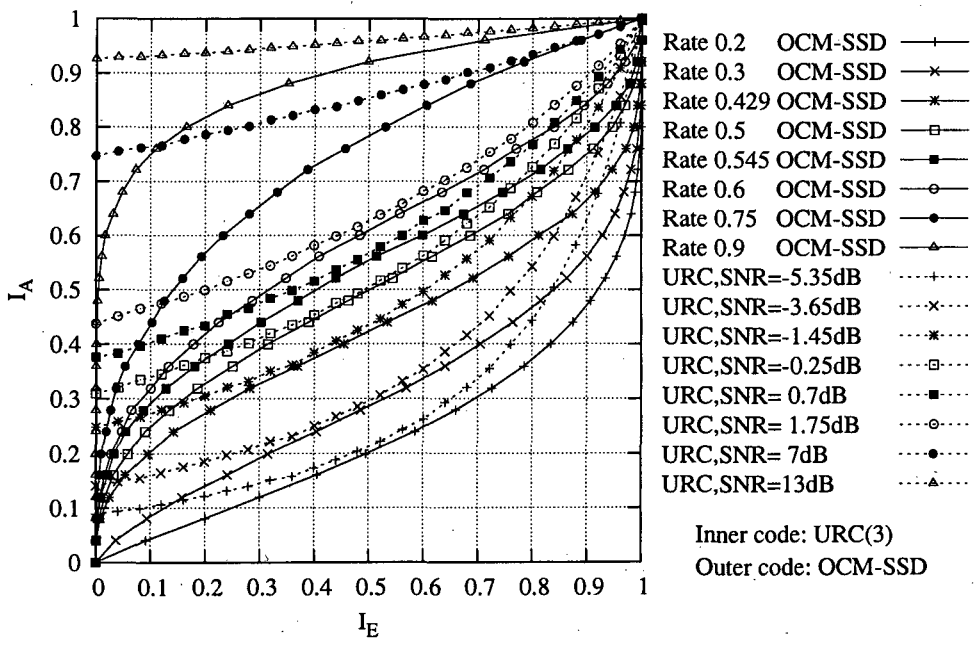


Figure 4.14: The EXIT characteristics of the transmission scheme employing OCM-SSD decoder as outer decoder and the URC having octal generator polynomial of 3 as inner decoder.

4.4.4 Performance versus System Delay

Note that in order to mitigate the effects of bursty errors, interleavers have been utilized in communication systems. However, the affordable length of the interleaver is typically determined by the end-to-end delay constraint. In this section, the effect of interleaver length on the OCM-SSD decoder's attainable performance is studied by investigating the convergence behaviour of the ISCD scheme employing OCM-SSD as the outer codec and the URCs of Section 4.4.2 as the inner codec.

As seen in Figures 4.18 and 4.19, the actual bit-by-bit decoding trajectory of the system having an interleaver length of 100 000 bits matches the EXIT characteristic of the OCM-SSD fairly accurately. This behaviour remains similar for a reduced interleaver length of 10 000 bits. However, the system having a short interleaver length of 1000 bits fails to reach the upper right corner of the EXIT chart. These findings suggest that an infinitesimally low probability of error may be achieved provided that the system employs an interleaver size in excess of 10 000 bits. At 30 frames/sec video scanning rate and 2000 bits/video frame for example this corresponds to a total delay of five video frames

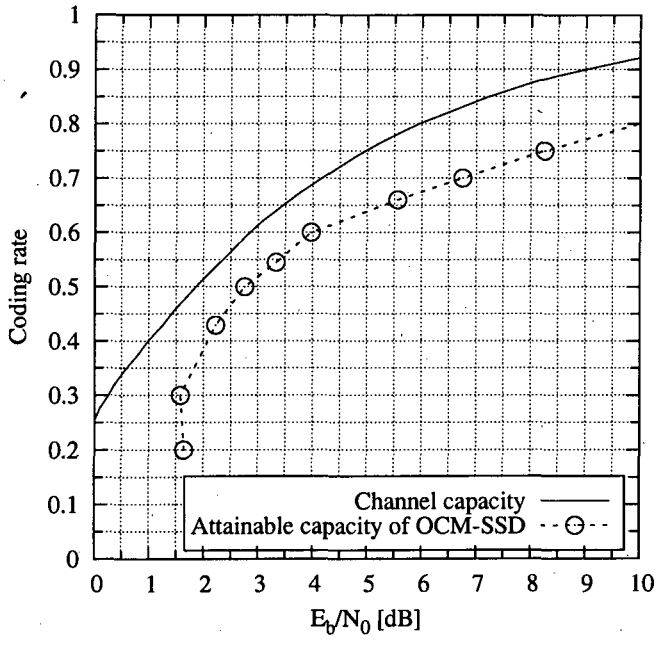


Figure 4.15: The maximum effective throughput for which open EXIT chart tunnels could be achieved versus BPSK-modulated Rayleigh fading channel E_b/N_0 for the scheme depicted in Figure 4.12 using different rate mappings

or to about 5:30= 167ms delay.

4.5 Chapter Conclusions

In this chapter, we have investigated an over-complete mapping assisted Iterative Source-Channel Decoding (ISCD) scheme designed for robust Dirac-encoded video transmission over fading channels. Commencing with a brief review of the softbit source decoding technique proposed by Adrat, Vary and Spittka [116, 127, 128], Section 4.2 investigated the performance of the ISCD scheme of Figure 4.1 designed for the transmission of Dirac video-encoded sequences over an uncorrelated Rayleigh fading channel by analyzing its convergence behaviour using EXIT charts. The EXIT charts recorded for the ISCD scheme depicted in Figure 4.4 revealed that when using the Dirac video encoder, the outer softbit source decoder exhibits a slowly increasing EXIT function and hence it can provide only modest *extrinsic* information for the inner channel decoder. As a result, the overall performance can hardly be improved by the iterative detection process. This was confirmed by the BER performance shown in Figure 4.5, where only a modest iteration

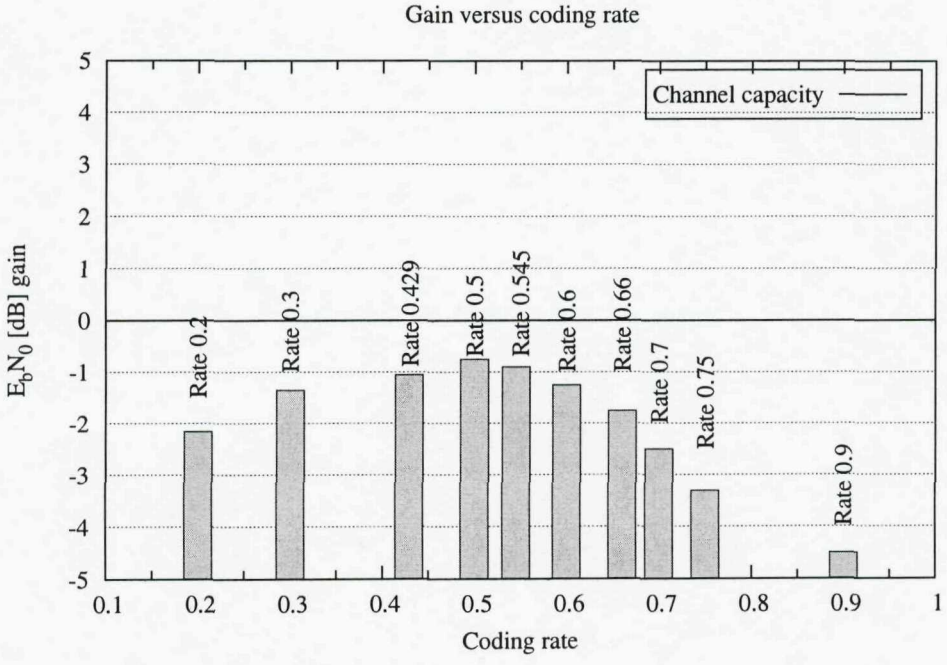


Figure 4.16: The attainable capacity of the OCM-SSD-URC scheme versus mapping-rate in comparison with channel capacity of the BPSK-modulated Rayleigh fading channel.

gain is attained. This result demonstrates that direct iterative source-channel decoding of the Dirac-encoded video stream does not perform well, because the amount of inherent redundancy in the Dirac-encoded sequence is limited.

In our novel approach, we partitioned the total available bit-rate budget between the source and channel codecs, creating a novel class of over-complete source-mapping schemes. As a counter-measure, in Section 4.3 the novel ISCD scheme portrayed in Figure 4.6 was proposed, where the softbit source decoder benefits both from the modest residual redundancy inherent in the source encoded bitstream as well as from the intentional redundancy imposed by the novel OCM schemes. The EXIT charts recorded for the OCM-assisted softbit source decoder shown in Figure 4.7 demonstrated that the performance of the OCM-assisted softbit source decoder strongly depends on the the redundancy imposed by the OCMs. However, the artificially introduced redundancy imposed by the OCMs cannot ensure that the EXIT curve of the SBSD reaches the point $(1, 1)$ at the upper-right corner of the EXIT chart, which would be necessary for achieving convergence to an infinitesimally low BER.

In Section 4.3 the design of OCM schemes was addressed by formulating the necessary

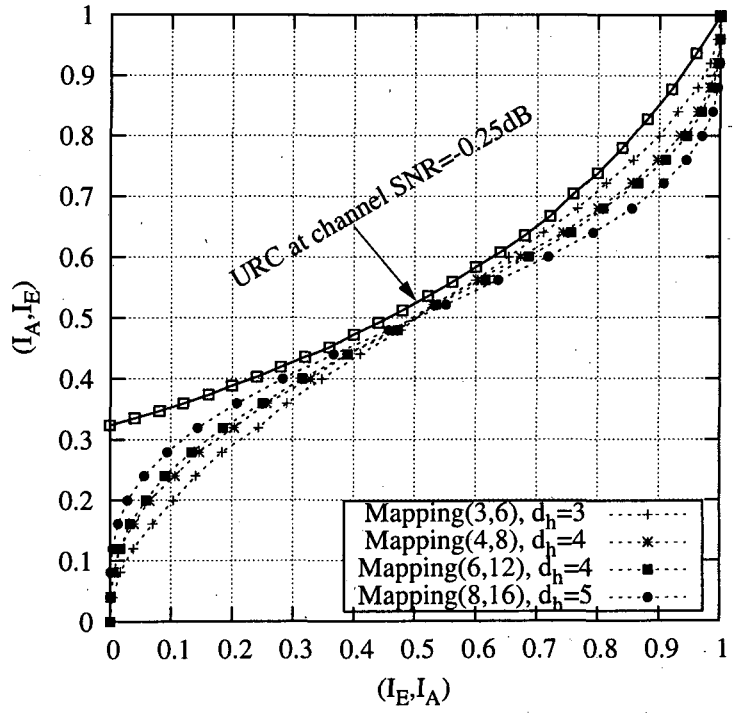


Figure 4.17: The inverted EXIT functions (dash-lines) of the OCM-SSD employing rate-1/2 OCMs, but different Hamming distances and/or source symbol lengths.

and sufficient condition for ensuring the iterative decoder's convergence to an infinitesimally low BER and then proceeded by proposing an algorithm for designing a suitable OCM scheme for iterative source-channel decoding. As a result, the attainable performance of the various video telephony schemes summarized in Table 4.9 for transmission over an uncorrelated Rayleigh fading channel and investigated in [85] illustrates the benefits of our ISCD design.

More explicitly, the relative frequency of the $K = 3$-bit source symbols modelling the probability distribution of the Dirac video-encoded bitstream are shown in Table 4.1, which was used for designing System 1 and System 2 of Table 4.9. By contrast, in System 3 and System 4 of Table 4.9, a uniform probability distribution of the $K = 3$-bit source symbols was assumed, which corresponds to the assumption that the high-compression H.264 video codec removed all the source redundancy. The results of [85] are portrayed in Figures 4.21 and 4.20, illustrating that at a fixed total system bandwidth, the turbo detection scheme using the H.264 video codec and disabling the error concealment tools but amalgamated both with the proposed OCM-SSD and the additional Generalized Low-

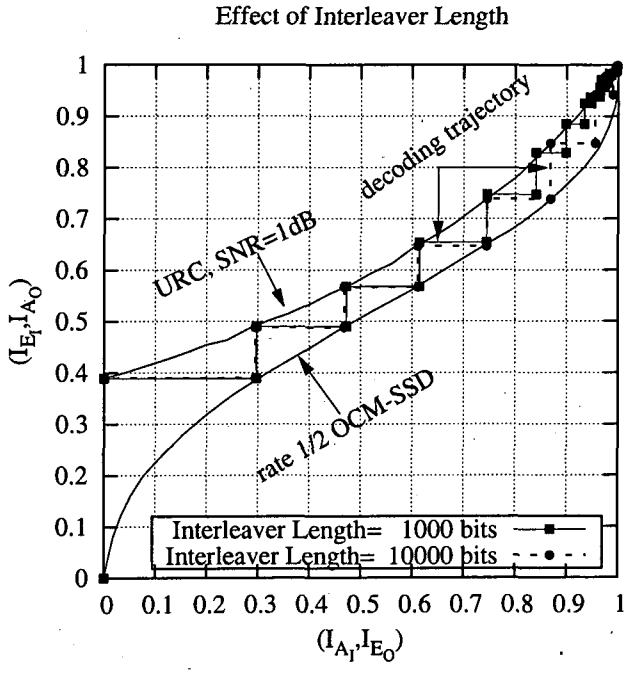


Figure 4.18: The effect of different interleaver delays on the ISCD's performance: EXIT-chart and Decoding Trajectory at channel SNR=1dB.

Density Parity-Check (GLDPC) code of [125] may lead to an approximately $2.5dB$ channel E_b/N_0 gain in comparison to the benchmarker having the same total bit-rate. More explicitly, the benchmarker was the H.264 scheme with the error concealment tools [129] enabled, hence resulting in a higher video rate, which was then amalgamated with the same-rate GLDPC code, as our proposed scheme but dispensed with our novel OCM-SSD. Moreover, the results of [85] and [130, 131] indicate that our OCM-SSD technique is widely applicable, regardless of the choice of source codec - even for speech and audio codecs.

The OCM-SSD is characterized in Section 4.4 by investigating its implementational complexity, the trade-offs between the achievable decoding performance and the mapping rate as well as the effects of interleaver delay on the OCM-SSD's performance. More specifically, the computational complexity imposed by OCM-assisted softbit source decoder was investigated in Section 4.4.1 by evaluating the number of multiply-add operations required to compute the *extrinsic* information of a single source symbol. The results of Tables 4.6 and 4.7 demonstrated that the computational complexity of the rate-K/N OCM-assisted SBSD is slightly higher than that of the K-bit source symbol SBSD and at the rate of

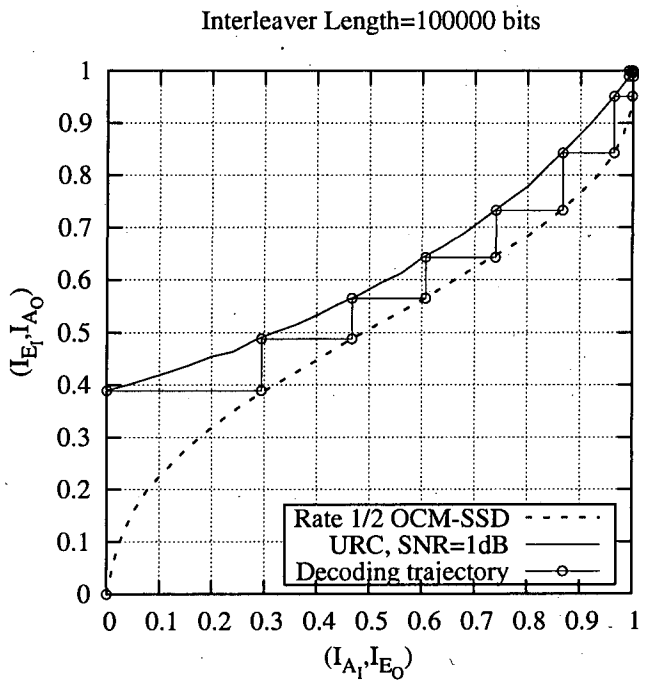


Figure 4.19: The effect of interleaver delay on the ISCD's performance: EXIT-chart and Decoding Trajectory at channel SNR=1dB.

K/N the computational complexity of the over-complete mapping-aided SBSD employing rate K/N mapping is significantly lower than that employing rate $2K/2N$ mapping. This property can be exploited to design wireless video systems, trading off the achievable system performance against the implementational complexity.

We also inspected the system's overall robustness against transmission errors in Section 4.4.2 by analyzing the EXIT characteristics of the OCM-SSD decoder, as seen in Figure 4.14. The plot of the attainable capacity in comparison to the channel capacity of the BPSK-modulated Rayleigh fading channel of Figures 4.15 and 4.16 showed that the attainable OCM-SSD gain is the highest near the coding rate of 1/2.

Furthermore, in Section 4.4.3 the effects of the OCM's Hamming distance on the OCM-SSD's performance were investigated by examining the various EXIT functions of the OCM-SSD decoders having the same mapping-rate, but different Hamming distances *and/or* source symbol lengths. The results shown in Figure 4.17 illustrated that the extrinsic information contribution of the OCM-SSD decoder employing an OCM scheme having a higher Hamming distance becomes more effective than their lower Hamming

Table 4.9: the simulation parameters of the various video telephony schemes

Parameters	System 1	System 2	System 3	System 4
Video codec	Dirac	Dirac	H264 without error concealment	H264 with error concealment
Mapping-rate (summarized in Table 4.3)	3/4 (Mapping-1)	3/4 (Mapping-2)	3/4 (Mapping-2)	1 (Mapping-0)
Average video frame size	2160 bits	2160 bits	2148 bits	2872 bits
Bit-Interleaver Π_1	2880 bits	2880 bits	2864 bits	2872 bits
GLDPC code-rate	2/3	2/3	2/3	2/3
Constituent codes	BCH(20,18,1)	BCH(20,18,1)	BCH(20,18,1)	BCH(20,18,1)
Number of iterations	8	8	8	8
Transmission packet size	4320	4320	4296	4308
Modulation	BPSK	BPSK	BPSK	BPSK

distance counterparts, especially for I_A values in excess of 0.5. This is achieved at the cost of increasing the computational complexity of the OCM-SSD decoders, as mentioned in Section 4.4.1.

The effect of interleaver length on the OCM-SSD decoder's attainable performance was studied in Section 4.4.4 by investigating the convergence behaviour of the ISCD scheme employing OCM-SSD as the outer codec and the URCs of Section 4.4.2 as the inner codec. As shown in Figures 4.18 and 4.19, the actual bit-by-bit decoding trajectory of the system having an interleaver length of 100 000 bits matches the EXIT characteristic of the OCM-SSD fairly accurately. This behaviour remains similar for a reduced interleaver length of 10 000 bits. However, the system having a short interleaver length of 1000 bits fails to reach the upper right corner of the EXIT chart. These findings suggest that an infinitesimally low probability of error may be achieved, provided that the system employs an interleaver size in excess of 10 000 bits. At 30 frames/sec video scanning rate and 2000 bits/video frame for example this corresponds to a total delay of five video frames or to about 167ms delay. The above results allow us to design an attractive video transmission scheme having a high error resilience at a reasonable complexity and a low delay.

Our future work may improve the OCM-SSD's performance by exploiting the residual redundancy which manifests itself in terms of the auto-correlation of source symbols.

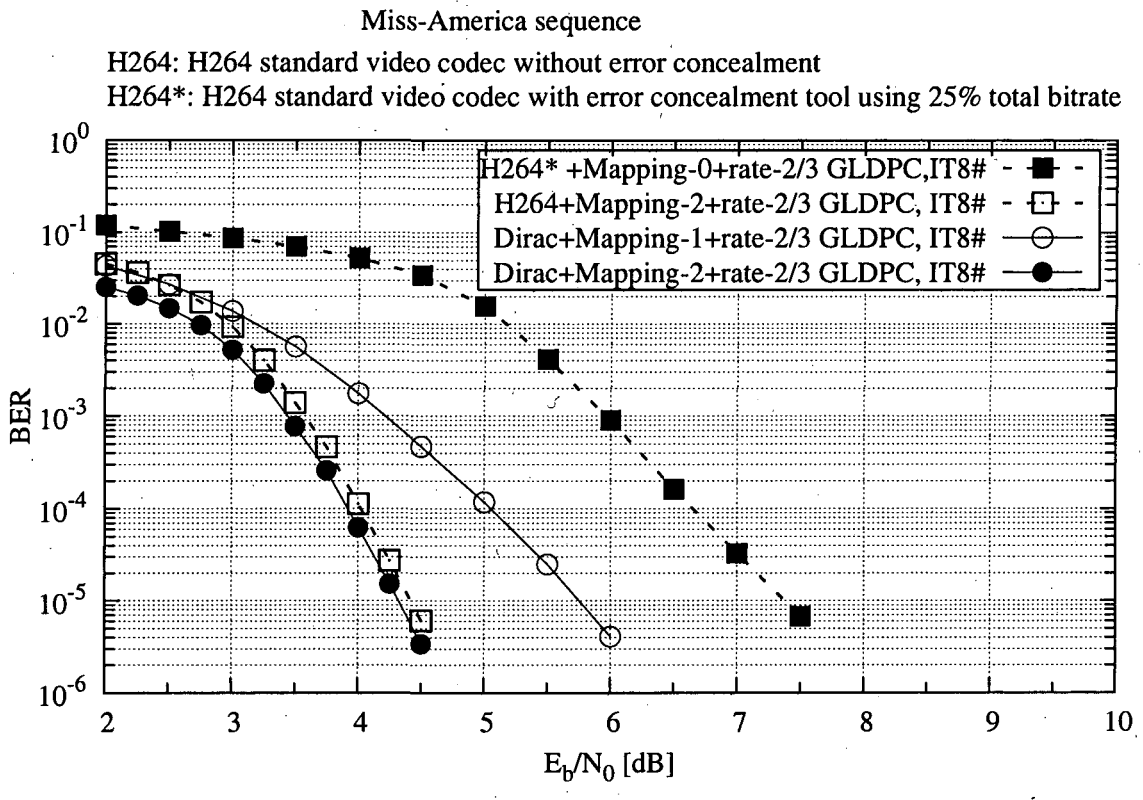


Figure 4.20: BER performance of the video telephony schemes characterized in Table 4.9 designed for transmission over the uncorrelated non-dispersive Rayleigh fading channel.

Furthermore, the Dirac video transmission performance may be improved by designing a suitable irregular over-complete mapping scheme, which allocates a lower mapping rate to highly sensitive bits and a higher mapping rate to less sensitive bits, while keeping the overall mapping-rate of the irregular over-complete mapping scheme the same as that of a single over-complete mapper. These irregular OCM schemes may also be conveniently assigned directly to the individual video codec parameters, which would prevent the propagation of errors across the various video parameters. These video parameters have inherently different lengths.

Another interesting design alternative is to consider a purely non-binary scheme, which transmits the N-bit OCM-SSD symbols using an N-bit/symbol modem and N-bit/symbol non-binary LDPC codes, as well as non-binary URCs. Finally, all the solutions of this chapter are also applicable to audio/speech systems.

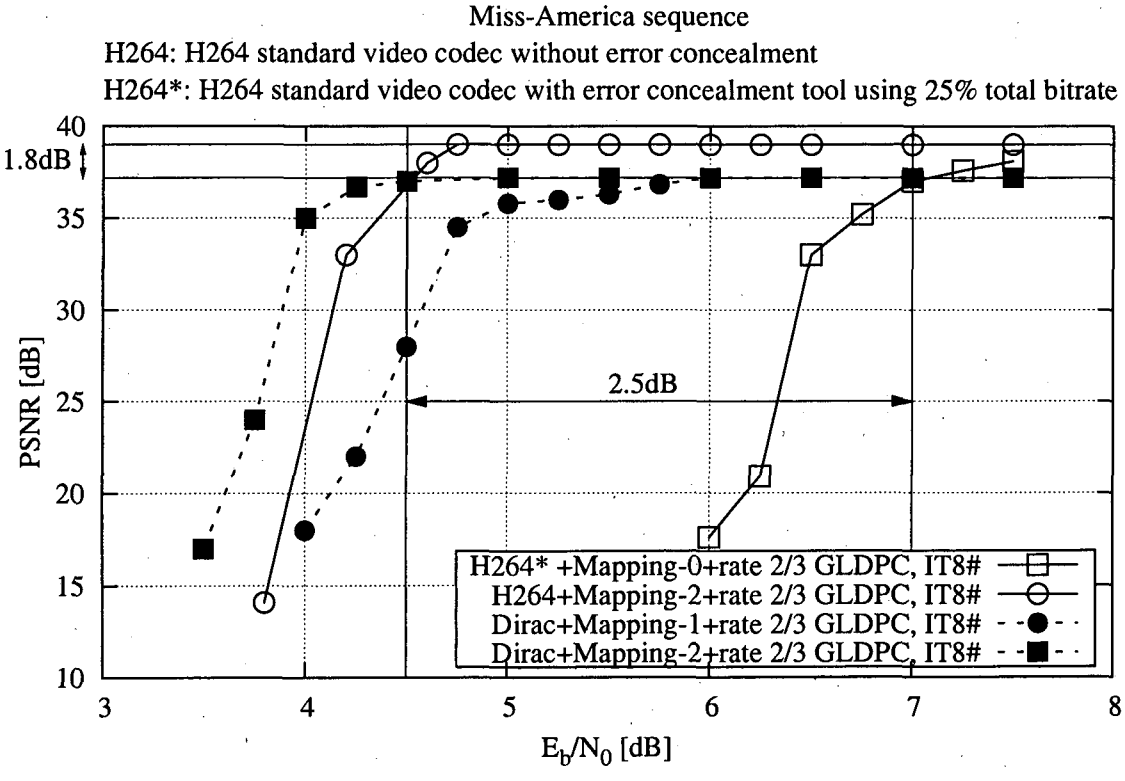


Figure 4.21: Comparison of the achievable video PSNR using the system characterized in Table 4.9.

Chapter 5

Irregular Iterative Source-Channel Decoding for Video Telephony

5.1 Introduction

Following a period of intensive research in wireless video telephony [7], [132], the resultant attractive services offered at various bit rates and video qualities have become a commercial reality, as often seen on television. Motivated by these events, the goal of this chapter is to design the wavelet-coded wireless videophone schemes using the Dirac video codec [8] of Chapter 2 as our source codec.

As quantified in Section 3.3, the Dirac video-encoded sequence exhibits different bit error sensitivities for the different bit positions. Provided that the channel decoder has the ability to exploit the knowledge of the source bit's error sensitivity, UEP has the potential to outperform EEP. This is achieved by applying a lower forward error correction (FEC) coding-rate for protecting the more sensitive portions of the video-encoded sequence and a higher FEC coding-rate for the less sensitive portions, while keeping the overall channel coding rate the same as that of the EEP scheme.

Furthermore, although Dirac codec attempts to remove all the redundancy from the source, there is often residual redundancy left in the encoded bitstreams. This residual redundancy can be utilized for the sake of providing an increased error protection in an iterative 'turbo-detection' fashion, which exchanges *extrinsic* information between the channel decoder and the softbit source decoder. Moreover, one can deliberately in-

crease the residual redundancy with the aid of of appropriately designed over-complete mapping-assisted softbit source decoding (OCM-SSD), as proposed in Chapter 4 in order to dramatically improve the attainable system performance, when iterative source-channel decoding is adopted.

As a further counter-measure, this chapter propose unequal error protection irregular over-complete mapping for wavelet video telephony using iterative source and channel decoding. The philosophy of the proposed video telephony scheme is that we attempt to exploit as much redundancy inherent in the Dirac-encoded bitstream as possible for the sake of improving the system's BER performance, while protecting the more sensitive bits of the Dirac video-encoded sequence with a lower OCM rate. Moreover, a near-instantaneously adaptive transceiver is proposed for transmitting the Dirac video-encoded bitstream. This allows us to design an attractive video transmission scheme, which attains iterative detection gains without increasing the system's delay, regardless of the specific choice of the source codec.

The rest of this chapter is organized as follows. In Section 5.2, unequal error protection based Irregular Over-Complete Mapping (IrOCM) is designed for wavelet-encoded video telephony using iterative source and channel decoding. Rather than being designed separately, the constituent over-complete mapping schemes of the IrOCM are designed jointly and then iteratively exchange extrinsic information with the rate-2/3 Generalized Low-Density Parity-Check (GLDPC) code of [125] employed as our inner code in Figure 5.1. The IrOCM is optimized in order to match the characteristics of both the Dirac codec and those of the channel, so that the UEP scheme employed maximizes the attainable iteration gain and hence the overall system performance. As a result, the benefits of our design are illustrated by investigating the attainable performance of two identical-rate Dirac video transmission schemes. Section 5.3 proposed a near-instantaneously adaptive transceiver for wireless video telephony, which is capable of estimating and adaptively exploiting the probability density function (PDF) of the video bitstream. This allows our wireless transceiver to be applied, regardless of the specific choice of source codec - even for speech and audio codecs. Finally, our conclusions are provided in Section 4.5.

5.2 Unequal Error Protection Irregular Over-Complete Mapping for Wavelet Video Telephony

5.2.1 Introduction

Since the different bits of the compressed video sequence produced by the BBC's wavelet-based video codec exhibit different error sensitivity, efficient UEP schemes have to be used to ensure that the perceptually more important bits benefit from more powerful protection. Furthermore, in the context of turbo detection the OCM aided soft videobit based source decoder should also match the characteristics of the channel codes for the sake of attaining a good iterative decoding performance. In this section, we address this design dilemma by proposing a novel IrOCM scheme constituted by appropriately weighted different-rate over-complete mapping schemes, protecting the appropriate fractions of the different stream by different-rate OCMs. Our turbo detection scheme has attractive iterative decoding convergence properties, while maintaining UEP capabilities matched to the requirements of the wavelet encoded video sequence. An EXIT chart based design procedure is presented and used in the context of protecting the different-sensitivity video bits of the BBC's wavelet-based Dirac video codec. As a benefit, the UEP scheme using rate 2/3 irregular over-complete mapping exhibits a transmit power advantage of about 0.7dB at an error probability of 0.01% over the equal-protection benchmarker scheme employing rate 2/3 over-complete mapping, when communicating over the uncorrelated Rayleigh-fading channel. We will demonstrate that this leads to an approximately 1dB channel SNR gain, when considering the video PSNR performance of our EEP iterative source-channel decoding scheme using an OCM having the same rate and $I = 6$ iterations.

The rest of this section is organized as follows. In Section 5.2.2, the schematic of the iterative detection aided UEP Dirac videophone scheme using IrOCM is presented. Section 5.2.3 investigates the design of the IrOCM in the context of Dirac-encoded video transmission. The convergence analysis of our IrOCM-assisted turbo detection scheme using EXIT charts is presented in Section 5.2.4. Finally, our overall system performance results are provided in Section 5.2.5, which will demonstrate that the UEP scheme outperforms the EEP scheme in terms of the achievable PSNR.

5.2.2 System Model

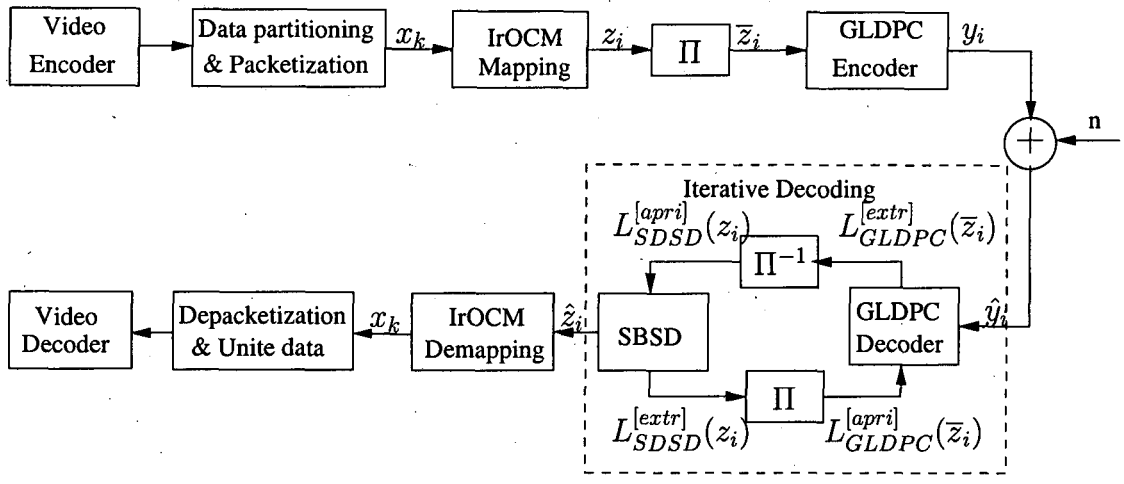


Figure 5.1: System model of an iterative detection aided UEP Dirac videophone scheme using IrOCM.

The block diagram of the iterative detection aided UEP Dirac videophone scheme using IrOCM is shown in Figure 5.1. At the transmitter, the video-compressed frames output by the Dirac encoder are first partitioned into three different bit error sensitivity classes, namely the Most Significant Bits (MSBs), Very Significant Bits (VSBs) and the Least Significant Bits (LSBs), as detailed in Section 3.4.2. These classes are then combined into fixed-length packets of $K = 2034$ bits x_k, where the 'Packet type' bits of Figure 5.2 explicitly indicate whether the packet contains MSBs. A one-bit flag is sufficient for signalling this classifier, which is repeated twice in order to allow Majority Logic Decision (MLD) at the receiver. Since one in three bits are conectable by the MLD scheme, this regime is capable of tolerating a 33% channel error rate. Note that all the MSBs of a certain video-encoded frame are contained in the same transmitted packet, and each transmitted packet contains no more than one frame's MSBs, as seen in Figure 5.2.

The $K = 2034$-bit packet x_k is mapped to the binary sequence z_i, $i = 1, 2, ..., 3060$ using the irregular over-complete mapping to be designed in Section 5.2.3, and resultant sequence z_i is then interleaved using the random 3060-bit interleaver Π of Figure 5.1. Finally, the interleaver's output sequence $\overline{z}_i$, $i = 1, 2, ..., 3060$ is encoded by the rate-2/3 Generalized Low-Density Parity-Check (GLDPC) Code described in Section 4.2.3.2, before its transmission over an uncorrelated Rayleigh fading channel.

At the receiver, the soft-decision based reliability information $\hat{y}_j$, $j = 1, 2, ..., 4590$ is evaluated in an iterative 'turbo-detection' fashion, which exchanges extrinsic information

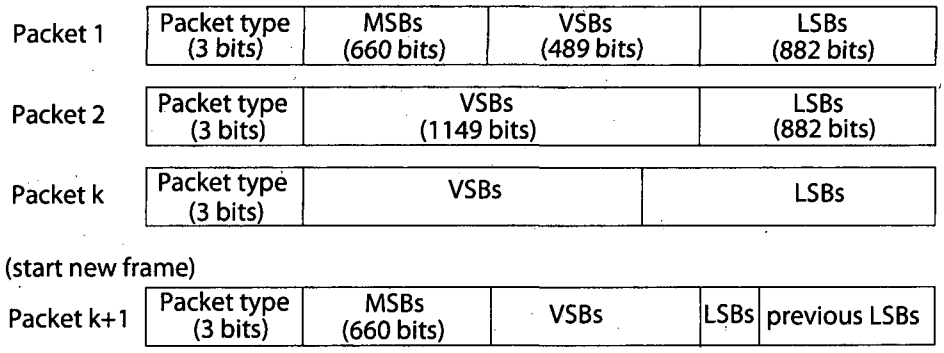


Figure 5.2: Packetization scheme for the Dirac video-encoded sequence

between the GLDPC decoder and the overcomplete mapping-assisted softbit source decoder of Section 4.3. This reliability information can be evaluated either in terms of *bit probabilities* or as their *log-likelihood ratios* (LLRs). At the last iteration softbit-estimation is carried out, the resultant bit sequence $\hat{z}_j$, $j = 1, 2, ..., 3060$ is then demapped for the sake of reconstructing the video bits of $K = 2034$-bit packet x_k. Based on the side-information provided by the MSB portion, the depacketizer rearranges the video bits of $K = 2040$-bit packet, and the resultant bit sequence is united to form reconstructed video frame, which is then fed into the Dirac video decoder for the sake of reconstructing the video sequence.

5.2.3 Design of Irregular Over-Complete Mapping

In this section, we exploit our knowledge of the Dirac video codec's bit error sensitivity, which was investigated in Section 3.3 for the sake of designing an iterative detection aided UEP Dirac videophone scheme. More specifically, as seen in Figure 5.1 at the transmitter side we invoke data partitioning, which partitions the Dirac video encoded sequence into three classes of bits having different error-sensitivities and then apply a lower-rate OCM scheme for protecting the more sensitive classes, and a higher-rate one for the less sensitive classes, while keeping the overall rate of the IrOCM arrangement the same as that of the EEP benchmarker scheme.

Noting that an IrOCM arrangement may be constructed by employing a family of OCM schemes having rates of $[r_1, r_2, ..., r_P]$, let L denote the total number of mapped bits generated by the IrOCM from a group of K input information bits. Each constituent OCM maps a fraction of $\alpha_k \cdot r_k \cdot L$, $k = 1, ..., P$ information bits and generates $\alpha_k \cdot L$

mapped bits. Given the target code rate of $R \in [0,1]$, the weighting coefficient α_k has to satisfy:

$$\sum_{k=1}^{P} \alpha_k = 1, \quad \sum_{k=1}^{P} \alpha_k \cdot r_k = R, \quad \text{and } \alpha_k \in [0,1] \; \forall k \; .$$

As seen in Figure 5.1, the mapped symbols of the constituent OCMs of the IrOCM are decoded jointly and iteratively by exchanging extrinsic information with the inner GLDPC code. Furthermore, the characteristics of the Dirac codec and those of the GLDPC code must be taken into account when optimizing the IrOCM scheme, so that UEP is achieved while maximizing the iteration gain of the system.

It is worth noting that due to the various sizes of the protection classes and owing to the employment of fixed-length component OCMs by the IrOCM scheme, the concatenation of these classes may lead to using different mapping rates within a specific protection class.

Table 5.1: The average ratio of the three protection classes in the Dirac video sequence.

	The relative frequency of occurrence		
	MSB class	VSB class	LSB class
Probability (%)	30.41	19.77	49.82

An interesting question arising for the design of the IrOCM is how to decide the ratio of the bits assigned to each component OCM of the IrOCM. The basic criterion is to aim for keeping the same ratio as in the original Dirac video encoded sequence. Unfortunately, this ratio is time-varying. Table 5.1 shows the average ratio of the three classes computed for the Dirac-encoded QCIF *"Miss_America"* video sequence scanned at 30 frames/s. The ratio of these classes is further characterized in Figure 5.3. We can see from Figure 5.3 that although the ratio varies from frame to frame, their range is limited. Hence, we opted for choosing a ratio that guarantees that the MSB class is protected by the lowest-rate OCM, while the VSB and LSB class should be protected by a higher-rate one.

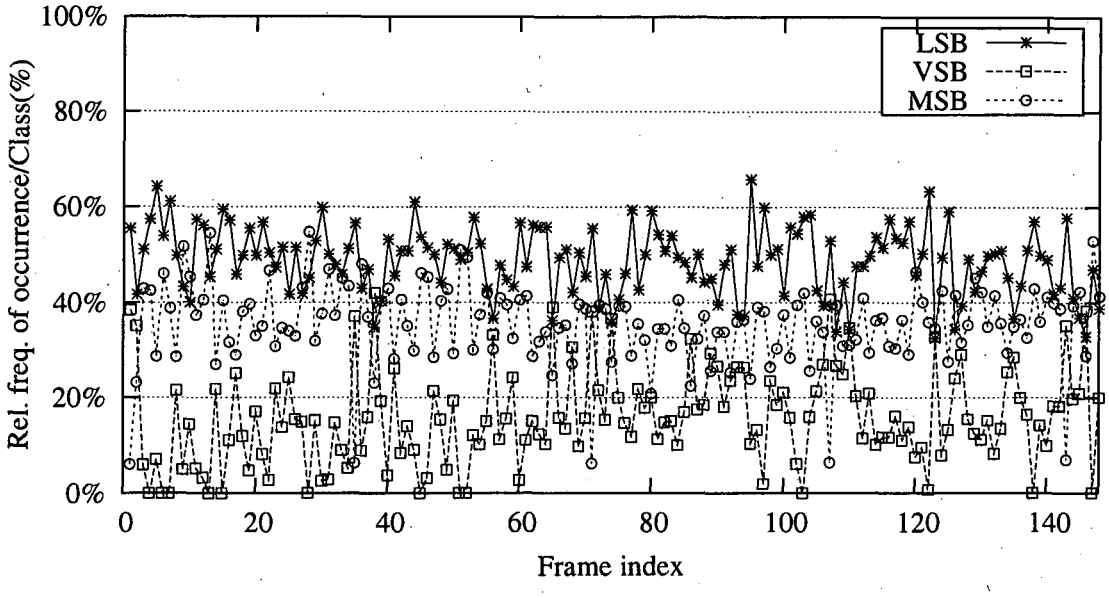


Figure 5.3: The relative frequency of occurrence for the MSB, VSB, LSB Dirac video-encoded bits computed from the *"Miss_America"* QCIF video sequence scanned at 30 frames/s.

Table 5.2: The 6-bit source symbol OCMs used in the simulations of Section 5.2.3.

Source symbol	The mapped symbols of			
	OCM[7,6]	OCM[9,6]	OCM[10,6]	OCM[11,6]
000000	0000000	000000010	0000000000	00000000000
000001	0000011	000001101	0000000111	00000001111
000010	0000101	000010100	0000011001	00000110011
000011	0000110	000011011	0000011110	00000111100
000100	0001001	000100100	0000101010	00001010101
000101	0001010	000101011	0000101101	00001011010
000110	0001100	000110010	0000110011	00001100110
000111	0001111	000111101	0000110100	00001101001
001000	0010001	001000100	0001001011	00010010110
001001	0010010	001001011	0001001100	00010011001
001010	0010100	001010010	0001010010	00010100101

Continued on next page

Table 5.2 – continued from previous page

Source symbol	The mapped symbols of			
	OCM[7,6]	OCM[9,6]	OCM[10,6]	OCM[11,6]
001011	0010111	001011101	0001010101	00010101010
001100	0011000	001100010	0001100001	00011000011
001101	0011011	001101101	0001100110	00011001100
001110	0011101	001110100	0001111000	00011110000
001111	0011110	001111011	0001111111	00011111111
010000	0100001	010000100	0110000001	01100000011
010001	0100010	010001011	0110000110	01100001100
010010	0100100	010010010	0110011000	01100110000
010011	0100111	010011101	0110011111	01100111111
010100	0101000	010100010	0110101011	01101010110
010101	0101011	010101101	0110101100	01101011001
010110	0101101	010110100	0110110010	01101100101
010111	0101110	010111011	0110110101	01101101010
011000	0110000	011000000	0111001010	01110010101
011001	0110011	011001111	0111001101	01110011010
011010	0110101	011010110	0111010011	01110100110
011011	0110110	011011001	0111010100	01110101001
011100	0111001	011100100	0111100000	01111000000
011101	0111010	011101011	0111100111	01111001111
011110	0111100	011110010	0111111001	01111110011
011111	0111111	011111101	0111111110	01111111100
100000	1000001	100000100	1010000010	10100000101
100001	1000010	100001011	1010000101	10100001010
100010	1000100	100010010	1010011011	10100110110
100011	1000111	100011101	1010011100	10100111001
100100	1001000	100100010	1010101000	10101010000
100101	1001011	100101101	1010101111	10101011111
100110	1001101	100110100	1010110001	10101100011
100111	1001110	100111011	1010110110	10101101100

Continued on next page

Table 5.2 – continued from previous page

Source symbol	The mapped symbols of			
	OCM[7,6]	OCM[9,6]	OCM[10,6]	OCM[11,6]
101000	1010000	101000010	1011001001	10110010011
101001	1010011	101001101	1011001110	10110011100
101010	1010101	101010100	1011010000	10110100000
101011	1010110	101011011	1011010111	10110101111
101100	1011001	101100100	1011100011	10111000110
101101	1011010	101101011	1011100100	10111001001
101110	1011100	101110010	1011111010	10111110101
101111	1011111	101111101	1011111101	10111111010
110000	1100000	110000010	1100000011	11000000110
110001	1100011	110001101	1100000100	11000001001
110010	1100101	110010100	1100011010	11000110101
110011	1100110	110011011	1100011101	11000111010
110100	1101001	110100100	1100101001	11001010011
110101	1101010	110101011	1100101110	11001011100
110110	1101100	110110010	1100110000	11001100000
110111	1101111	110111101	1100110111	11001101111
111000	1110001	111000100	1101001000	11010010000
111001	1110010	111001011	1101001111	11010011111
111010	1110100	111010010	1101010001	11010100011
111011	1110111	111011101	1101010110	11010101100
111100	1111000	111100010	1101100010	11011000101
111101	1111011	111101101	1101100101	11011001010
111110	1111101	111110100	1101111011	11011110110
111111	1111110	111111011	1101111100	11011111001

Another factor in deciding this ratio is the fairness of our comparisons. More specifically, in order to compare the performance of the EEP and UEP schemes, we have to let the rate of IrOCM and that of the single OCM arrangement to be the same. More explicitly, the EEP and UEP schemes are chosen as follows:

1. For the EEP scheme, a rate-6/9 over-complete mapping OCM[9,6] of Table 5.2 was employed as the outer code, which provides equal protection for the $K = 2040$ video bits per transmitted packet.

2. For the UEP scheme, we employ an IrOCM that consists of a family of component OCMs having rates of $\left[\frac{6}{7}, \frac{6}{10}, \frac{6}{11}\right]$ summarized in Table 5.2, which encode [882, 486, 666] video bits, respectively. This results in the IrOCM having an appropriately chosen coding rate of 2/3 for the $K = 2034$ video bits per transmitted packet.

5.2.4 Convergence Analysis Using EXIT Charts

Let us first investigate the EXIT characteristics [106] of the various OCM components of the IrOCM scheme. In all cases, the EXIT curves seen in Figure 5.4 were generated using uncorrelated Gaussian distributed *a priori* LLRs with the assumption that each of the 6-bit source symbols[1] has the same probability of occurrence. All EXIT functions were obtained by using the histogram-based LLR measurements of [106].

For the IrOCM-assisted SBSD decoder (outer decoder), we denote the mutual information between the *a priori* input A and the mapped bits C as $I_A^o = I(C;A)$, while the mutual information between the *extrinsic* output E and the mapped bits C is denoted as $I_E^o = I(C;E)$. Then the transfer function of outer decoder can be defined as:

$$I_E^o = T_o(I_A^o) \tag{5.1}$$

which maps the input variable I_A^o to the output variable I_E^o. Similarly, for the GLDPC decoder (inner decoder), we denote the mutual information between the *a priori* input A and the transmitted information bits X as $I_A^i = I(X;A)$. Furthermore, we denote the mutual information between the *extrinsic* output E and the transmitted information bits X as $I_E^i = I(X;E)$. Note that the *extrinsic* output of the inner code also depends on the channel SNR or E_b/N_0. Hence the transfer function of the inner code is defined as:

$$I_E^i = T_i(I_A^i, E_b/N_0) \tag{5.2}$$

[1]All bits of the MSB, VSB and LSB classes were assigned to 6-bit source-symbols, which were then protected by the [6/7, 6/10, 6/11]-rate OCMs

We also denote the EXIT transfer function of the k^{th} component softbit source decoder $k = 1, 2, 3$ of the amalgamated IrOCM decoder as $I_E^{o,k} = T_{o,k}(I_A^{o,k})$. The inverted transfer functions $I_A^{o,k} = T_{o,k}^{-1}(I_E^{o,k})$ of the IrOCM's component softbit source decoders and that of the IrOCM decoder, namely $I_A^o = T_o^{-1}(I_E^o)$ are again depicted in Figure 5.4.

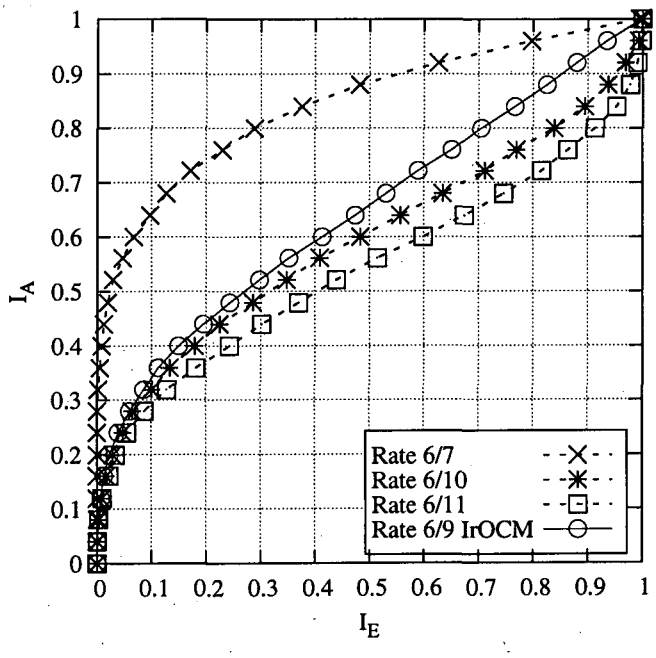


Figure 5.4: The EXIT transfer function of the rate-2/3 IrOCM and that of its components

Note that the EXIT transfer function of the IrOCM decoder can also be obtained as the weighted superposition of the $P = 3$ component softbit source decoders' transfer functions as follows

$$T_o(I_A^o) = \sum_{k=1}^{P} \alpha_k \cdot T_{o,k}(I_A^o), \tag{5.3}$$

where α_k is the weighting coefficient mentioned above. The EXIT charts [106] of the UEP scheme using IrOCM as well as the EEP scheme using rate-6/9 over-complete mapping are shown in Figure 5.5.

As seen from the EXIT chart, there is a wider EXIT chart tunnel between the inner decoder's curve and the outer decoder's curve for the proposed UEP scheme in comparison to the EEP scheme using the rate-6/9 OCM. Hence, given a limited number of iterations, we would predict that the UEP system using the IrOCM may be expected to perform better than the EEP scheme employing the rate-6/9 OCM in the range of

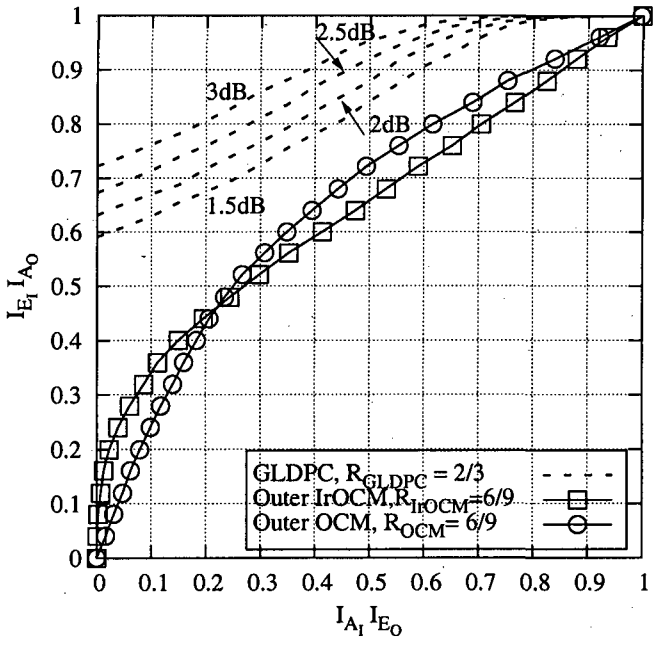


Figure 5.5: The EXIT transfer functions of the rate-6/9 over-complete mapping-assisted outer decoder and the designed IrOCM-assisted outer decoder, as well as that of the inner decoder at E_b/N_0=1.5, 2.0, 2.5 and 3dB.

E_b/N_0=1.5...3dB.

5.2.5 Simulation Results

In this section, the attainable performance of the proposed UEP scheme of Figure 5.1 is investigated for transmission over an uncorrelated Rayleigh fading channel. The various system parameters are summarized in Table 5.3.

The associated BER performances are shown in Figure 5.6 for an overall code-rate of 0.44. More specifically, the proposed UEP system outperformed the EEP system by about $0.7dB$ for $I = 6$ iterations.

The video PSNR performances of the UEP scheme and EEP scheme summarized in Table 5.3 are depicted in Figure 5.7. Observe from Figs. 5.6 and 5.7 that our UEP system attains a near-unimpaired video reconstruction quality for a channel E_b/N_0 value of 2.75dB, although at this point the overall BER is about $1.6 \cdot 10^{-4}$. Furthermore, at the same channel E_b/N_0 value of 3.25dB the video reconstruction quality of the EEP system is significantly lower than that of the UEP system, although at this point the overall BER

Table 5.3: The simulation parameters

Parameters	EEP scheme	UEP scheme
Overall code-rate	4/9	4/9
Video codec	Dirac	Dirac
Mapping rate	rate-6/9 OCM	rate-6/9 IrOCM
Bit-Interleaver Π_1	3060 bits	3060 bits
GLDPC code-rate	2/3	2/3
Constituent codes	BCH(12,10,1)	BCH(12,10,1)
Number of iterations	6	6
Transmission packet size	4590	4590
Modulation	BPSK	BPSK

of the EEP system is about $5 \cdot 10^{-4}$.

5.3 Near-Instantaneously Adaptive Transceivers for Wireless Video Telephony

5.3.1 Introduction

In order to accommodate the variations of the relative frequency of the K-bit source symbol modelling the PDF of the different video-encoded bitstreams, the employment of near-instantaneously adaptive video transceivers is desirable. This allows to make the video communication system suitable for different video sources, i.e. different video codecs and different parameters K.

Furthermore, the K-bit source symbol PDF of a certain video codec depends on the characteristics of the video source, such as for example the motion-activity and the video format. Hence, there is an inevitable difference between the specific symbol PDF of the video sequence to be transmitted and the PDF generated using off-line training sequences.

The above observation has led to the design of near-instantaneously adaptive transceivers for the Dirac-based wireless video transceiver, which motivated the research to be outlined in the following sections. Focussing my attention on the OCM-assisted softbit decoding aspects, i.e. on the probabilities of the 64 legitimate 6-bit source symbols of the MSBs,

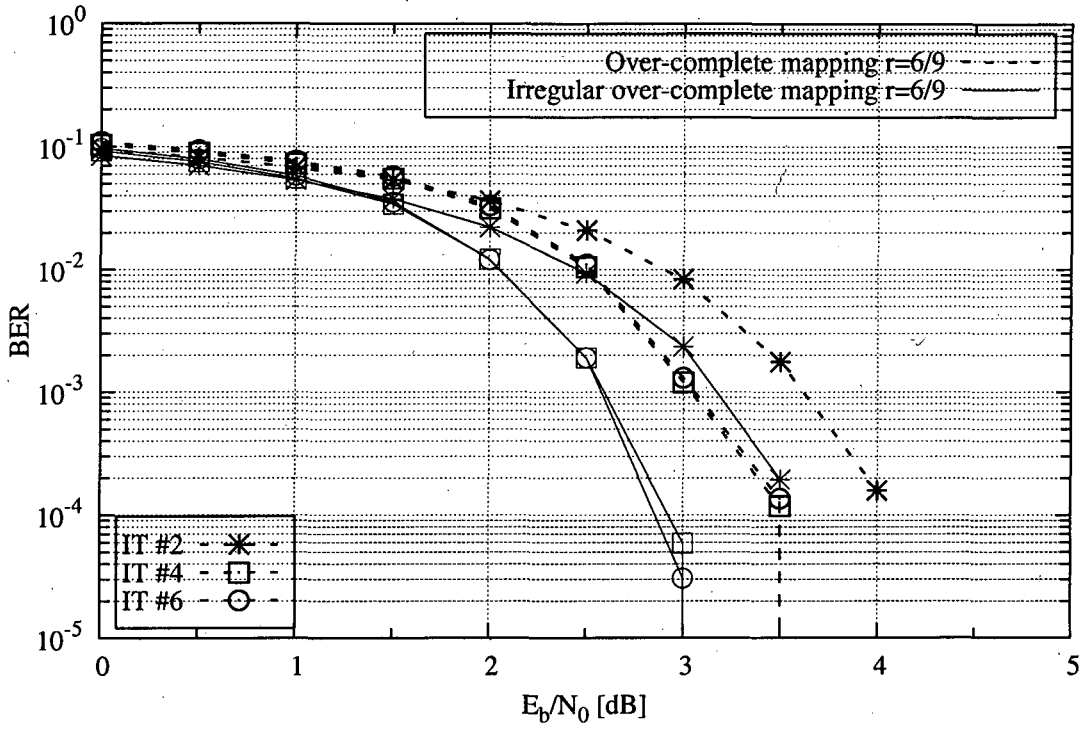


Figure 5.6: BER performance of the EEP system and the UEP system of Table 5.3 designed for transmission over the uncorrelated Rayleigh fading channel.

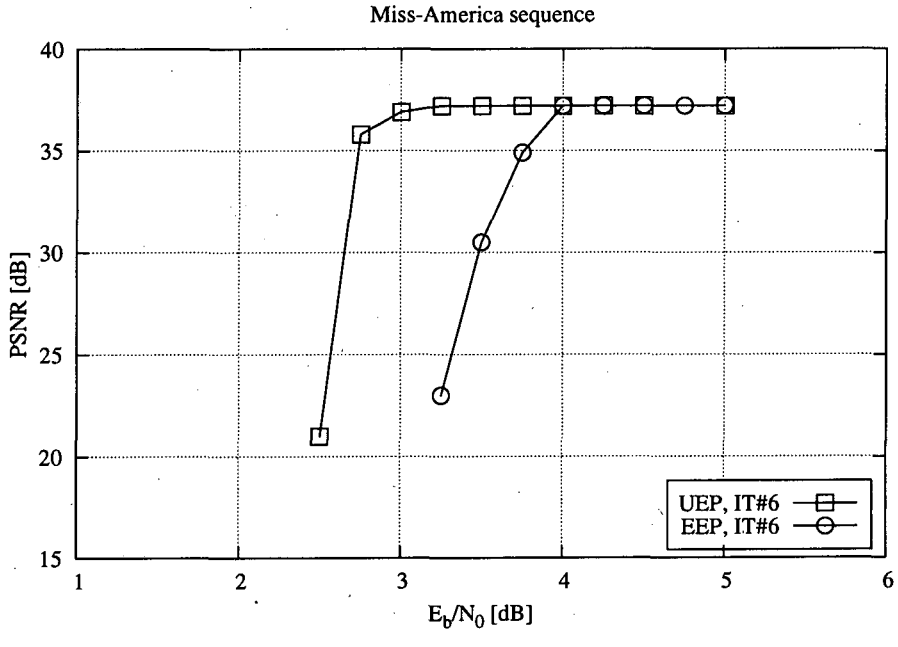


Figure 5.7: Comparison of the achievable video PSNR using the systems characterized in Table 5.3.

VSBs and LSBs, an adaptive transceiver was designed by instantaneously updating the OCM-SSD decoder's PDF. More specifically, the reconstructed video-encoded sequences at the output of a depacketizer of Figure 5.1 and Figure 5.8 is used for quantifying the relative frequency of occurrence for all the 64 legitimate 6-bit source symbols of the three Dirac video bit protection classes. The resultant relative frequencies of the 64 legitimate 6-bit symbols are then fed back to the OCM-SSD decoder for updating its symbol PDF. Figure 5.8 shows the schematic of the near-instantaneously adaptive transceiver in the context of the Dirac-based wireless video telephony system.

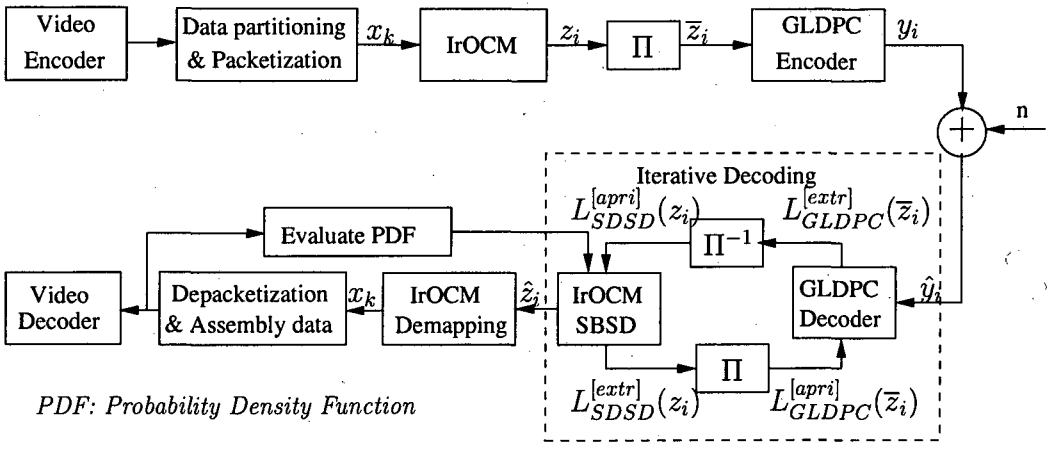


Figure 5.8: Near-instantaneously adaptive wireless video transceivers.

Note that the OCM-SSD decoder assumes having an identical 6-bit source symbol probability for the decoding of the first received packet. Once the iteration process was completed, the MSBs, VSBs and LSBs of the transmitted packet are reconstructed and the relative frequency of occurrence for all the 64 legitimate 6-bit source symbols of these classes have to be updated for assisting in the decoding of the forthcoming packet. These operations are continued until the entire Dirac video-encoded sequence has been decoded. Before, we have to note that the PDF-update process is naturally prone to error propagation, when the symbol error probability is high, but it has the benefit of requiring no extra side-information transmission in support of the decoder's operation.

5.3.2 Simulation Results

In this section, we investigate the attainable performance of the various video telephony schemes considered for transmission over an uncorrelated Rayleigh fading channel for the sake of illustrating the benefits of our adaptive transceiver benefiting from the on-

line symbol probability update process. The simulation parameters are summarized in Table 5.4.

Table 5.4: the simulation parameters of the various video telephony schemes

Parameters	System 1	System 2	System 3
System model	UEP scheme of Fig. 5.1	UEP scheme of Fig. 5.1	UEP scheme of Fig. 5.8
IrOCM's source symbol PDF	6-bit/symbol identical symbol prob.	6-bit/symbol PDF of Table 5.5	6-bit/symbol adaptive PDF
Video codec	Dirac	Dirac	Dirac
Mapping	rate 2/3 IrOCM	rate 2/3 IrOCM	rate 2/3 IrOCM
Bit-Interleaver Π_1	3060 bits	3060 bits	3060 bits
GLDPC code-rate	2/3	2/3	2/3
Constituent codes	BCH(20,18,1)	BCH(20,18,1)	BCH(20,18,1)
Transmission packet size	4590	4590	4590
Number of iterations	8	8	8
Modulation	BPSK	BPSK	BPSK

More explicitly, we assume having the same time-invariant 6-bit source symbol probability of 1/64 for all component OCMs of the IrOCM of System 1, while in System 3 the time-invariant adaptively updated symbol probabilities of the adaptive transceiver of Figure 5.8 were applied. Furthermore, the 64 symbol probabilities extracted from our simulations using the training sequences of the 150-frame *"Miss_America"* video-encoded sequence, the 300-frame *Akiyo* video-encoded sequence, the 300-frame *Mother-Daughter* video-encoded sequence are summarized in Table 5.5, which were utilized in System 2 of Table 5.4.

Table 5.5: The probability of all the 64 legitimate 6-bit source symbols of the three classes in the Dirac video sequence.

Symbol	The probability density function of		
	MSB class	VSB class	LSB class
000000	0.18312	0.01884	0.01620
000001	0.02393	0.01760	0.01580
000010	0.01060	0.01726	0.01601
000011	0.05044	0.01786	0.01605
000100	0.06686	0.01545	0.01560
000101	0.01120	0.01661	0.01578
000110	0.00677	0.01640	0.01568
000111	0.07501	0.01724	0.01614
001000	0.02393	0.01760	0.01580
001001	0.00313	0.01643	0.01542
001010	0.00139	0.01612	0.01562
001011	0.00659	0.01668	0.01566
001100	0.00874	0.01443	0.01522
001101	0.00146	0.01551	0.01540
001110	0.00088	0.01531	0.01530
001111	0.00980	0.01610	0.01575
010000	0.01060	0.01726	0.01601
010001	0.00139	0.01612	0.01562
010010	0.00061	0.01582	0.01582
010011	0.00292	0.01637	0.01586
010100	0.00387	0.01416	0.01542
010101	0.00065	0.01522	0.01560
010110	0.00039	0.01502	0.01549
010111	0.00434	0.01580	0.01595
011000	0.05044	0.01786	0.01605
011001	0.00659	0.01668	0.01566

Continued on next page

Table 5.5 – continued from previous page

Symbol	The probability density function of		
	MSB class	VSB class	LSB class
011010	0.00292	0.01637	0.01586
011011	0.01389	0.01693	0.01591
011100	0.01841	0.01465	0.01546
011101	0.00309	0.01575	0.01564
011110	0.00186	0.01554	0.01554
011111	0.02066	0.01634	0.01600
100000	0.06686	0.01545	0.01560
100001	0.00874	0.01443	0.01522
100010	0.00387	0.01416	0.01542
100011	0.01841	0.01465	0.01546
100100	0.02441	0.01267	0.01503
100101	0.00409	0.01363	0.01520
100110	0.00247	0.01345	0.01510
100111	0.02738	0.01414	0.01555
101000	0.01120	0.01661	0.01578
101001	0.00146	0.01551	0.01540
101010	0.00065	0.01522	0.01560
101011	0.00309	0.01575	0.01564
101100	0.00409	0.01363	0.01520
101101	0.00069	0.01465	0.01538
101110	0.00041	0.01446	0.01528
101111	0.00459	0.01520	0.01573
110000	0.00677	0.01640	0.01568
110001	0.00088	0.01531	0.01530
110010	0.00039	0.01502	0.01549
110011	0.00186	0.01554	0.01554
110100	0.00247	0.01345	0.01510
110101	0.00041	0.01446	0.01528
110110	0.00025	0.01427	0.01518

Continued on next page

Table 5.5 – continued from previous page

Symbol	The probability density function of		
	MSB class	VSB class	LSB class
110111	0.00277	0.01500	0.01563
111000	0.07501	0.01724	0.01614
111001	0.00980	0.01610	0.01575
111010	0.00434	0.01580	0.01595
111011	0.02066	0.01634	0.01600
111100	0.02738	0.01414	0.01555
111101	0.00459	0.01520	0.01573
111110	0.00277	0.01500	0.01563
111111	0.03072	0.01577	0.01609

The associated BER performances are shown in Figure 5.9. More specifically, at the same overall code-rate of 4/9, System 2 and System 3 attain slightly better BER performances than System 1, when using $I = 8$ iterations. This is because the OCM-SSD decoders of System 2 and System 3 have exploited both the residual redundancy, which manifests itself in terms of the non-uniform probability distribution of the 64 6-bit symbols inherent in the Dirac video encoded bit-stream, as well as the intentional redundancy imposed by IrOCM for mitigating the effects of transmission errors. By contrast, the less sophisticated System 1 has exploited only the intentional redundancy imposed by IrOCM for doing so, but assumed having a 6-bit symbol probability of 1/64. Furthermore, since there is only a marginal difference between the symbol-PDF of Table 5.5 extracted from our simulations using the training sequences and the near-instantaneously updated PDFs, the achievable performance of System 2 is similar to that of System 3, which instantaneously updates the OCM-SSD decoder's PDF.

5.4 Chapter Conclusions

This chapter exploited our knowledge of the Dirac video codec's bit error sensitivity detailed in Section 3.3 and the OCM-assisted Softbit Source Decoding technique detailed

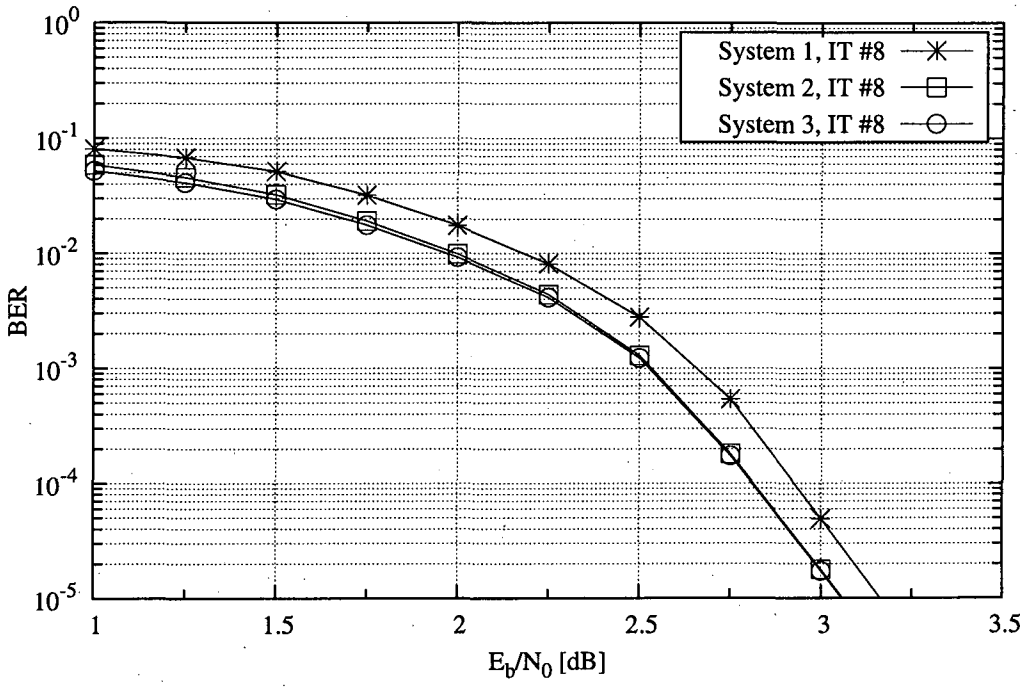


Figure 5.9: Comparison of the achievable BER performance of the systems characterized in Table 5.4 for transmission over the uncorrelated non-dispersive Rayleigh fading channel.

in Chapter4 for the sake of designing an attractive wavelet-based wireless video telephone scheme.

More specifically, in Section 5.2, an unequal error protection based IrOCM scheme was designed for wavelet-encoded video telephony using iterative source and channel decoding. The design of IrOCM-based UEP was carried out in the context of Dirac-encoded video transmission in Section 5.2.3 and by analyzing the convergence behaviour of the IrOCM-based UEP scheme using EXIT charts in Section 5.2.4. More explicitly, rather than being designed separately, the constituent over-complete mapping schemes of the IrOCM were designed jointly and they iteratively exchanged extrinsic information with the rate-2/3 Generalized Low-Density Parity-Check (GLDPC) code of [125] employed as our inner code in Figure 5.1.

The IrOCM scheme was optimized in order to match the characteristics of both the Dirac codec and those of the channel, so that the IrOCM-assisted turbo detection scheme maximized the attainable iteration gain and hence the overall achievable system performance. A wider EXIT chart tunnel was created between the inner decoder's and the outer decoder's EXIT curve for the proposed IrOCM-based UEP scheme in comparison

to the EEP scheme using the rate-6/9 OCM scheme, as seen in Figure 5.5. Given a limited number of iterations, the IrOCM-based UEP scheme using the IrOCM may be expected to perform better than the EEP scheme employing the rate-6/9 OCM in the range of E_b/N_0=1.5...3dB. This was confirmed by the results of Figure 5.6, where the proposed IrOCM-based UEP scheme attained a transmit power advantage of about 0.7dB at an error probability of 0.01% over the equal-protection benchmarker scheme employing rate-2/3 over-complete mapping, when communicating over the uncorrelated Rayleigh-fading channel. This leads to an approximately 1dB channel SNR gain, when considering the video PSNR performance of the EEP iterative source-channel decoding benchmarker scheme using OCM having the same rate and $I = 6$ iterations, as seen in Figure 5.7.

It was noted in Section 5.3.1 that the relative frequency of the K-bit source symbols used for modelling the PDF of a certain Dirac-encoded video bitstream depends on the specific value of K as well as on the characteristics of the video source, such as for example the motion-activity and the video format. Hence, there is an inevitable difference between the specific symbol PDF of the video sequence to be transmitted and the PDF generated using off-line training sequences. This motivated the employment of near-instantaneously adaptive video transceivers, which benefitted from updating the PDF of the video-encoded sequences, in an on-line fashion as detailed in Section 5.3.

More explicitly, the reconstructed video-encoded sequences at the output of the depacketizer of Figure 5.8 was used for quantifying the relative frequency of occurrence for all the 64 legitimate 6-bit source symbols of the three Dirac video bit protection classes. The resultant relative frequencies of the 64 legitimate 6-bit symbols were then fed back to the OCM-based softbit source decoder for updating its symbol PDF. The associated BER performances shown in Figure 5.9 revealed that at the same overall code-rate of 4/9, System 2 and System 3 of Table 5.4 attain slightly better BER performances than System 1, when using $I = 8$ iterations. This is because the OCM-based softbit source decoders of System 2 and System 3 have exploited both the residual redundancy, which manifests itself in terms of the non-uniform probability distribution of the 64 6-bit symbols inherent in the Dirac video encoded bit-stream, as well as capitalizing on the intentional redundancy imposed by IrOCM for mitigating the effects of transmission errors. By contrast, the less sophisticated System 1 has exploited only the intentional redundancy imposed by IrOCM for doing so, but assumed having a 6-bit symbol probability of 1/64. Furthermore, since there is only a marginal difference between the symbol-PDF of Table 5.5 extracted

from our simulations using the training sequences and the near-instantaneously updated PDFs, the achievable performance of System 2 is similar to that of System 3, which instantaneously updates the OCM-SSD decoder's PDF. We note that the employment of the proposed adaptive transceiver allows our wireless video telephony scheme to attain iterative detection gains without increasing the system's delay, regardless of the specific choice of the source codec, which make it widely applicable, even for audio transmissions.

Future work could improve the Dirac-based video transceiver's performance by exploiting the residual redundancy, which manifests itself in terms of the auto-correlation of the source symbols. However, this naturally leads to increasing the computational complexity of the system considered.

Chapter 6

Conclusions and Future Research

Joint Source-Channel Coding and decoding (JSCC) is a state-of-the-art technique used for designing robust video transmission systems. Significant progress was made throughout the last decades in optimizing each individual component of generations of wireless communication systems [44–46,116,118,133–136]. Accordingly, superior performance was achieved for JSCC schemes in comparison to separate source and channel coding. This observation motivated the research presented in this thesis. Three particular JSCC schemes were investigated in this thesis, namely *iterative detection aided unequal error protection using irregular convolutional codes* [82], *overcomplete mapping-assisted iterative source-channel decoding* [84] and *near-instantaneously adaptive unequal error protection irregular over-complete mapping* [86]. A key philosophy underlying these schemes is the well-known "turbo principle", which originates from the renowned turbo codes [80,137–145]. We note that the research ideas proposed in this thesis are restricted to the Dirac video codec. In fact, these ideas can be extended to communication systems adopting other source codecs. For instance, the iterative source-channel decoding scheme proposed in Chapter 4 and the near-instantaneously adaptive wireless video transceiver proposed in Chapter 5 are suitable for arbitrary source codecs [85], including audio transmissions [130]. In the following sections, these schemes are detailed, together with a discussion of our future work.

6.1 Iterative Detection Aided Irregular Convolutional Coded Video Scheme

As quantified in Section 3.3, the Dirac video-encoded sequence exhibits different bit error sensitivities for the different bit positions. If the channel encoder has the ability to exploit the knowledge of the source bit error sensitivity, UEP has the potential to outperform EEP. This is achieved by applying a lower channel coding rate to protect the more sensitive portions of the data sequence, and a higher channel coding rate to the less sensitive portions, while keeping the overall channel coding rate the same as that of the EEP scheme.

As a novel contribution, in Chapter 3 an UEP scheme was designed for the Dirac video codec by using IRCC. Our IRCC was constituted by a family of convolutional codes having different code rates, which were designed with the aid of EXIT charts [106] for achieving convergence to an infinitesimally low BER, while having unequal error protection capabilities matched to the specific requirements of the Dirac video codec. In general, EXIT chart analysis assumes having a high interleaver length. However, it was shown in [104] that by using an appropriate optimization criterion, the concatenated system becomes capable of performing well even for relatively short interleaver block lengths. Since the constituent codes have different coding rates, the resultant IRCC is inherently capable of providing UEP.

More specifically, channel encoder consists of an outer IRCC encoder and a rate-1 precoder [107], which are separated by the interleaver Π of Figure 3.13. The outer IRCC code adds redundancy for the sake of providing protection for the video data, while the inner precoder renders the channel to appear recursive[1], which assists in reducing the associated BER [107], when employing iterative detection. Furthermore, the IRCC is capable of assigning a lower coding rate to highly sensitive bits and a higher coding rate to less sensitive bits.

To elaborate a little further, based on the Dirac video codec's bit-sensitivity investigations provided in Section 3.3 our partitioning scheme divides the unequal-sensitivity information of a video-encoded sequence into three protection classes, namely the MSBs, VSBs and LBSs. The average ratio of these protection classes in the Dirac video se-

[1]The term 'recursive' implies here that the system has an infinite impulse response.

quence was provided in Table 3.3. The various system parameters are summarized in Table 3.4. More explicitly, the five-path dispersive channel of Equation 3.1 imposing severe ISI and the outer IRCC having an average coding rate of 0.5 were selected for our investigations. The UEP IRCC scheme consists of a set of subcodes having coding rates of [0.35 0.45 0.5 0.55 0.6 0.8], which encode [34% 1.3% 4.8% 22% 6.5% 31%] of the input data bits, respectively. For our outer NSC benchmarker scheme the generator polynomials of $g_0 = 1 + D + D^2 + D^4$ and $g_1 = 1 + D^3 + D^4$ were used. For both schemes, the rate-1 precoder having a generator polynomial of $g_0 = 1 + D^2$ was invoked and both the channel equalizer as well as the channel decoder employed the Maximum A *Posteriori* (MAP) algorithm [113]. In all simulations BPSK modulation and a dispersive AWGN channel were assumed.

The achievable system performance was evaluated for $K = 2000$ bits per video frames and half-rate coding, which facilitated the employment of an interleaver length of $L = 4000$ bits across a single video frame, i.e without extra video delay. The results shown in Figures 3.16, 3.17 and 3.18 revealed that

1. For very noisy channels, where E_b/N_0 is relatively low, UEP schemes are better than EEP in terms of both the achievable BER and PSNR.

2. The convergence SNR threshold for the system using the IRCC is lower than that for the EEP system using the rate-1/2 NSC code. More quantitatively, we found that the former required an SNR of 1.3dB, while the latter necessitated 2.4dB for maintaining an open EXIT chart, which facilitates convergence to a low BER.

3. The employment of a sufficiently low coding rate for the MSB class allows us to preserve the integrity of the corresponding bits.

4. The proposed UEP scheme is capable of attaining a near unimpaired video quality for channel SNRs in excess of about 4.5dB over a five-path dispersive AWGN channel and it outperforms the EEP scheme by about 0.8dB in terms of the required E_b/N_0.

6.2 Overcomplete Mapping-Assisted Iterative Source-Channel Decoding

Joint source-channel decoding schemes can be divided into iterative JSCD and non-iterative JSCD, according to whether there is iterative decoding between the source and channel decoders. While non-iterative JSCD provides excellent results for moderate-SNR channels, these methods fail to perform well for low channel SNRs or for error-infested dispersive fading channels. For these scenarios source-decoding techniques can be designed to exploit the residual redundancy inherent in the source-encoded bitstream for the sake of achieving additional error protection [44,116,146–149]. Based on these source-decoding techniques, several iterative source-channel decoding approaches have been investigated, leading to a variety of guidelines. For example, in [116] and [44] the source decoder exploited the inherent residual redundancy in the encoded bitstream for the sake of providing an increased error protection. By contrast, in [115] Buttigieg and Farrell deliberately increased the inherent residual redundancy in the encoded bitstream, creating a class of VLEC codes, which combine the benefits of variable-length coding with improved distance properties.

More recently, Adrat and Vary [128, 150] have introduced an iterative source-channel decoding scheme, which improves the SBSD performance using a powerful index assignment. The remarkable ISCD scheme of [114] used EXIT charts for demonstrating that the performance of the ISCD scheme employing the SBSD technique of [116] was highly dependent on the amount of redundancy inherent in both the source and channel codes. It was also demonstrated that the redundancy of the source-encoded bitstream plays a role, which is as important as the intentionally imposed redundancy of the channel code. Furthermore, in [114] the authors demonstrated that a necessary, but insufficient condition of ensuring successful ISCD is that the minimum residual redundancy quantified in terms of the auto-correlation of source codec parameters should be $\rho_{min} = 0.77$, which corresponds to a potential data rate reduction by $\Delta N = 0.659$ bit/symbol at $N = 3bits/symbol$. It is worth noting that in state-of-the-art video encoders, such as the MPEG4, JPEG2000 [5] and Dirac [82] schemes, there is little residual redundancy left in the source-encoded bitstream. Hence it is difficult to achieve further performance improvements with the aid of ISCD. In a manner similar to that of [115], a class of OCMs was created, which intentionally increased the redundancy in the encoded bitstream for the sake of improving the error

correcting capability of the softbit-based source decoding algorithm detailed in [116].

As a counter-measure, in Section 4.3 a novel ISCD scheme portrayed in Figure 4.6 was proposed, where the softbit-source decoder benefits both from the modest residual redundancy inherent in the source encoded bitstream as well as from the intentional redundancy imposed by the novel OCM schemes. The EXIT charts recorded for the OCM-assisted softbit-source decoder shown in Figure 4.7 demonstrated that the performance of the OCM-assisted softbit-source decoder strongly depends on the the redundancy imposed by the OCMs. However, the artificially introduced redundancy imposed by the OCMs cannot ensure that the EXIT curve of the SBSD reaches the point $(1, 1)$ at the upper-right corner of the EXIT chart, which would be necessary for achieving convergence to an infinitesimally low BER.

In Section 4.3 the design of OCM was addressed by formulating the necessary and sufficient condition for ensuring the iterative decoder's convergence to an infinitesimally low BER and then proceeded by proposing an algorithm for designing a suitable OCM scheme for iterative source-channel decoding. As a result, the attainable performance of the various video telephony schemes summarized in Table 6.1 for transmission over an uncorrelated Rayleigh fading channel investigated in [85] shown in illustrate the benefits of our ISCD design.

Table 6.1: the simulation parameters of the various video telephony schemes

Parameters	System 1	System 2	System 3	System 4
Video codec	Dirac	Dirac	H264 without error concealment	H264 with error concealment
Mapping-rate (summarized in Table 4.3)	3/4 (Mapping-1)	3/4 (Mapping-2)	3/4 (Mapping-2)	1 (Mapping-0)
Average video frame size	2160 bits	2160 bits	2148 bits	2872 bits
Bit-Interleaver Π_1	2880 bits	2880 bits	2864 bits	2872 bits
GLDPC code-rate	2/3	2/3	2/3	2/3
Constituent codes	BCH(20,18,1)	BCH(20,18,1)	BCH(20,18,1)	BCH(20,18,1)
Number of iterations	8	8	8	8
Transmission packet size	4320	4320	4296	4308
Modulation	BPSK	BPSK	BPSK	BPSK

More explicitly, the relative frequency of the $K = 3$-bit source symbols modelling

the probability distribution of the Dirac video-encoded bitstream are shown in Table 4.1, which was used for designing System 1 and System 2. By contrast, in System 3 and System 4, a uniform probability distribution of the $K = 3$-bit source symbols was assumed, which corresponds to the assumption that the high-compression video codec H.264 removed all the source redundancy. The results of [85], as seen in Figures 6.2 and 6.1, illustrate that at a fixed total system bandwidth, the turbo detection scheme using the H.264 video codec disabling the error concealment tools but amalgamated both with the proposed OCM-SSD and the additional Generalized Low-Density Parity-Check (GLDPC) code of [125] may lead to an approximately $2.5dB$ channel E_b/N_0 gain in comparison to the benchmarker having the same total bit-rate, which operated the H.264 scheme with the error concealment tools enabled, hence resulting in a higher video rate and amalgamated with the same-rate GLDPC code, but dispensing with our novel OCM-SSD. Moreover, the results of [85] and [130] indicate that our OCM-SSD technique is widely applicable, regardless of the choice of source codec - even for speech and audio codecs.

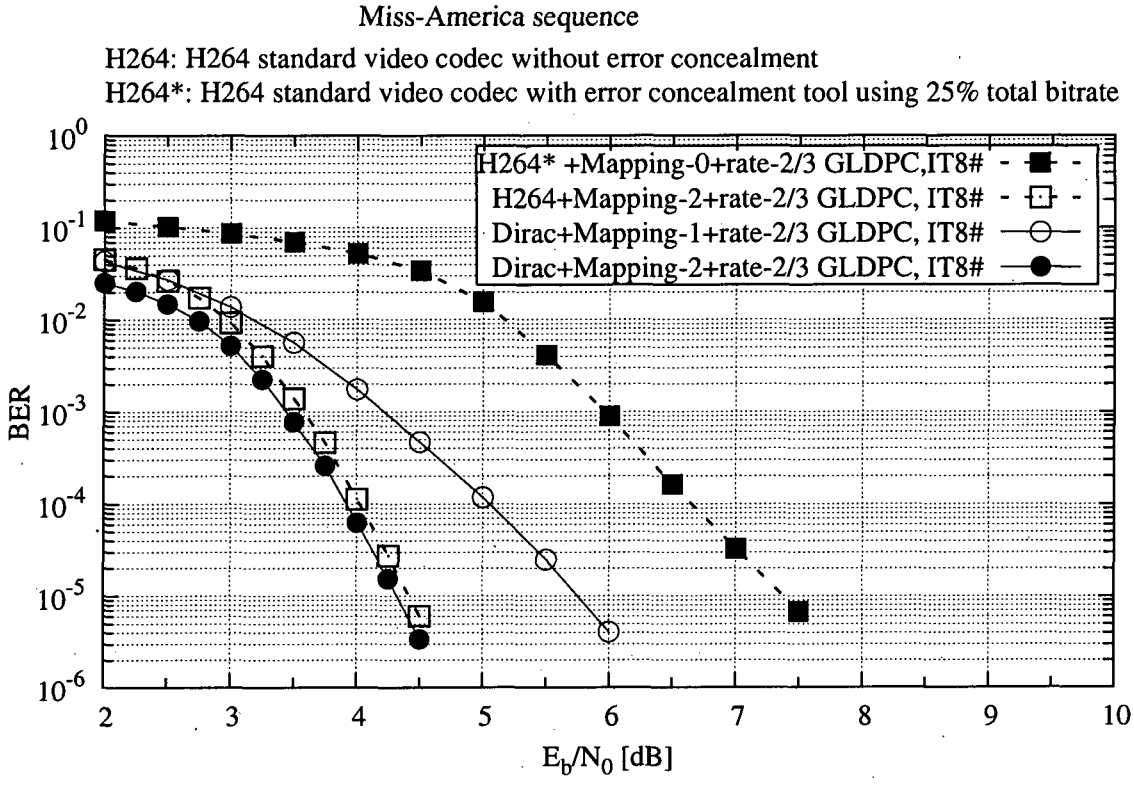


Figure 6.1: BER performance of the video telephony schemes characterized in Table 4.9 designed for transmission over the uncorrelated non-dispersive Rayleigh fading channel.

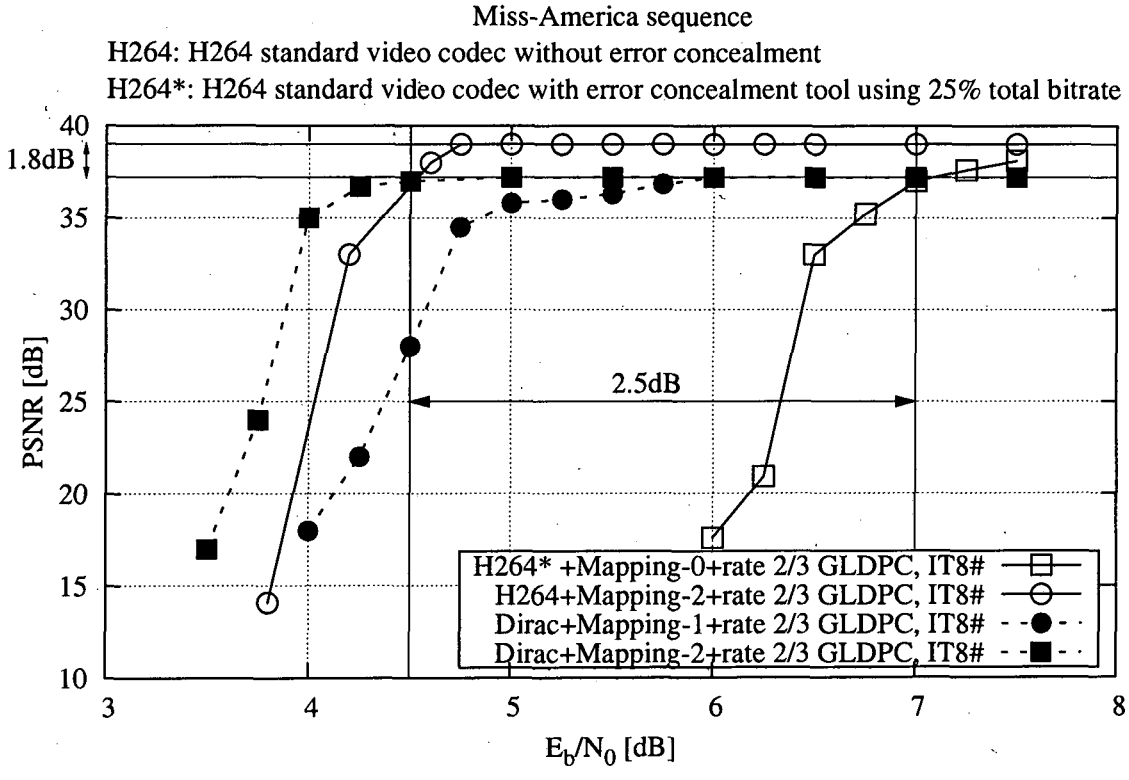


Figure 6.2: Comparison of the achievable video PSNR using the system characterized in Table 4.9.

Furthermore, Section 4.4 characterized the OCM-SSD by investigating its implementational complexity, the trade-off between the achievable decoding performance and the mapping rate as well as the effects of interleaver delay on the OCM-SSD's performance. More specifically, the computational complexity imposed by OCM-assisted softbit-source decoder was investigated in Section 4.4.1 by evaluating the number of multiply-add operations required to compute the *extrinsic* information of a single source symbol. The results of Tables 4.6 and 4.7 exhibited that the computational complexity of the rate-K/N OCM-assisted SBSD is slightly higher than that of the K-bit source symbol SBSD and at the rate of K/N the computational complexity of the over-complete mapping-aided SBSD employing rate K/N mapping is significantly lower than that employing rate $2K/2N$ mapping. This property can be exploited to design wireless video systems, trading off the achievable system performance against the implementational complexity.

We also inspected the system's overall robustness against transmission errors in Section 4.4.2 by analyzing the EXIT characteristics of the OCM-SSD decoder, as seen in

Figure 4.14. The plot of the attainable capacity in comparison to the channel capacity of the BPSK-modulated Rayleigh fading channel of Figures4.15 and 4.16 showed that the OCM-SSD gain peaks near the coding rate of 1/2.

Additionally, in Section 4.4.3 the effects of the OCM's Hamming distance on the OCM-SSD's performance were investigated by examining the various EXIT functions of the OCM-SSD decoders having the same mapping-rate, but different Hamming distances *and/or* source symbol lengths. The results shown in Figure4.17 illustrated that the extrinsic information contribution of the OCM-SSD decoder employing OCM having a higher Hamming distance becomes more effective than their lower Hamming distance counterpart for I_A values in excess of 0.5. This is achieved at the cost of increasing the computational complexity of the OCM-SSD decoders, as mentioned in Section 4.4.1.

The effect of interleaver length on the OCM-SSD decoder's attainable performance was studied in Section 4.4.4 by investigating the convergence behaviour of the ISCD scheme employing OCM-SSD as the outer codec and the URCs of Section 4.4.2 as the inner codec. As shown in Figures4.18 and 4.19, the actual bit-by-bit decoding trajectory of the system having an interleaver length of 100 000 bits matches the EXIT characteristic of the OCM-SSD fairly accurately. This behaviour remains similar for a reduced interleaver length of 10 000 bits. However, the system having a short interleaver length of 1000 bits fails to reach the upper right corner of the EXIT chart. These findings suggest that an infinitesimally low probability of error may be achieved provided that the system employs an interleaver size in excess of 10 000 bits. At 30 frames/sec video scanning rate and 2000 bits/video frame for example this corresponds to a total delay of five video frames or to about 5:30= 167ms delay.

6.3 Near-Instantaneously Adaptive Unequal Error Protection Irregular Over-Complete Mapping

In Chapter 5 we exploited our knowledge of the Dirac video codec's bit error sensitivity of Section 3.3 and the OCM-assisted Softbit Source Decoding technique detailed in Chapter 4 for the sake of designing an attractive wavelet-based wireless video telephony. The philosophy of the proposed video telephony is that we try to exploit as much redundancy inherent in the Dirac-encoded bitstream as possible for improving the system's

BER performance, while protecting the more sensitive portions of the Dirac video-encoded sequence with a lower mapping rate.

More specifically, in Section 5.2, IrOCM-based UEP scheme is designed for wavelet-encoded video telephony using iterative source and channel decoding, as shown in Figure 5.1. The design of IrOCM-based UEP scheme was carried out by investigating the design of the IrOCM in the context of Dirac-encoded video transmission in Section 5.2.3 and analyzing convergence behaviour of the IrOCM-based UEP scheme using EXIT charts in Section 5.2.4. More explicitly, rather than being designed separately, the constituent OCM schemes of the IrOCM are designed jointly and then iteratively exchange extrinsic information with the rate-2/3 Generalized Low-Density Parity-Check (GLDPC) code of [125] employed as our inner code.

The IrOCM is optimized in order to match the characteristics of both the Dirac codec and those of the channel, so that the IrOCM-assisted turbo detection scheme employed maximizes the attainable iteration gain and hence the overall system performance. A wider EXIT chart tunnel between the inner decoder's curve and the outer decoder's curve for the proposed IrOCM-based UEP scheme in comparison to the EEP scheme using the rate-6/9 OCM, as seen in Figure 5.5 provided that given a limited number of iterations, the IrOCM-based UEP scheme using the IrOCM may be expected to perform better than the EEP scheme employing the rate-6/9 OCM in the range of E_b/N_0=1.5...3dB. This was confirmed by the results of Figure 5.6, where the proposed IrOCM-based UEP scheme attains a transmit power advantage of about 0.7dB at an error probability of 0.01% over the equal-protection benchmarker scheme employing rate-2/3 over-complete mapping, when communicating over the uncorrelated Rayleigh-fading channel. This leads to an approximately 1dB channel SNR gain, when considering the video PSNR performance of the EEP iterative source-channel decoding benchmarker scheme using OCM having the same rate and $I = 6$ iterations, as seen in Figure 5.7.

It was noted that the relative frequency of the K-bit source symbol modelling the PDF of a certain Dirac-encoded video bitstreams depends on different parameters K as well as the characteristics of the video source, such as for example the motion-activity and the video format. Hence, there is an inevitable difference between the specific symbol PDF of the video sequence to be transmitted and the PDF generated using off-line training sequences. This motivated the employment of near-instantaneously adaptive video transceivers, which benefitted from updating the PDF of the video-encoded sequences as

provided in Section 5.3.

More explicitly, the reconstructed video-encoded sequences at the output of a depacketizer Figure 5.8 is used for quantifying the relative frequency of occurrence for all the 64 legitimate 6-bit source symbols of the three Dirac video bit protection classes. The resultant relative frequencies of the 64 legitimate 6-bit symbols are then fedback to the OCM-based softbit-source decoder for updating its symbol PDF. The associated BER performances shown in Figure 5.9 revealed at the same overall code-rate of 4/9, System 2 and System 3 of Table 5.4 attain slightly better BER performances than System 1, when using $I = 8$ iterations. This is because the OCM-based softbit-source decoders of System 2 and System 3 have exploited both the residual redundancy, which manifests itself in terms of the non-uniform probability distribution of the 64 6-bit symbols inherent in the Dirac video encoded bit-stream, as well as the intentional redundancy imposed by IrOCM for mitigating the effects of transmission errors. By contrast, the less sophisticated System 1 has exploited only the intentional redundancy imposed by IrOCM for doing so, but assumed having a 6-bit symbol probability of 1/64. Furthermore, since there is only a marginal difference between the symbol-PDF of Table 5.5 extracted from our simulations using the training sequences and the near-instantaneously updated PDFs, the achievable performance of System 2 is similar to that of System 3, which instantaneously updates the OCM-SSD decoder's PDF.

We note that the employment of the proposed adaptive transceiver allows our wireless video telephony scheme attain iterative detection gains without increasing the system's delay, regardless of the specific choice of the source codec, which make it widely applicable, even for audio transmissions.

6.4 Future Research Directions

Future work may improve the Dirac-encoded video transmission performance in five aspects. First, the optimum coding rate allocation for UEP remains unanswered. This is because the difference in ratio of three classes is varied from frame to frame. The UEP scheme using fixed proportion of the subcodes may lead to protecting the most error-sensitive segment of the encoded video frame with high-rate subcodes if the encoded frame containing many MSBs and VSBs, and vice versa.

Second, It would be interesting to expand the application of the ISCD technique employing over-complete mapping-aided softbit source decoding, particularly to other video and audio codecs, such as for example the MPEG4 audio/video coding standards, in the light of exemplifying a wide range of applications of the over-complete mapping scheme. Furthermore, a list codebooks which are optimized for different rate over-complete mappings should be provided.

Third, the loss of synchronization issue needs to be addressed. This is of vital importance for the design of video transmission over binary erasure channel (BEC), such as for example wireless internet communication. Significant image quality degradation could be avoided by transmitting side information bits and/or retransmitting I-frames for video sequence synchronization. These synchronization bits should be best protected since the loss of synchronization will ruin the reconstructed image.

Fourth, most of the techniques studied in this treatise were proposed for low-rate applications. However, they are also benecial in high-quality, high-bitrate wireless video broadcast systems communicating over highly dispersive channels at bitrates between 1 and 10 Mbps using for example 4CIF resolution video sequences.

Finally, future work could improve the Dirac-based video telephony's performance by exploiting the residual redundancy which manifests itself in terms of the auto-correlation of source symbols. The exploitation of pixel-domain video redundancy potentially allows the realisation of a signicant error resilience capability. However, this leads to the cost of increasing the computational complexity of the considered system.

List of Symbols

C: GLDPC supercode

$c(i,j)$: Wavelet subband coefficient

$d_h(x,y)$: the Hamming distance between two binary sequences of the same length x and y.

$d_h(X)$: the Hamming distance of a set X.

E_b: Bit energy.

E_b/N_0: Ratio of bit energy to noise power spectral density.

g_i: The synthesis filter coefficients of the DWT

H: GLDPC's parity-check matrix

H_i: The Block diagonal matrix used in GLDPC

$H(X)$: The entropy of the random variable X

$H(X,Y)$: The joint entropy of the random variables X and Y.

$H(X|Y)$: The conditional entropy of the random variable X given Y

h_i: The analysis filter coefficients of the DWT

$I(X;Y)$: The mutual information between the random variables X and Y

$I(X;Y|Z)$: The conditional mutual information between the random variables X and Y given the random variable Z

L: The interleaver block length

$LLR\Big(u_\tau(\lambda)\Big)$: The LLR *extrinsic* value of $u_\tau(\lambda)$

M: The input symbol size of over-complete mapping

$P(u_k)$: The symbol probability.

Q: Quantization factor.

R: Coding rate.

r: The rate of IRCC's subcode.

u_τ: The τ-th M-ary symbol of the input sequence at the soft-bit source decoder

$\hat{x}$: The estimation of x.

x_k: The transmitted bit at instance k.

y_k: The received bit at instance k.

π: Interleaver.

π^{-1}: Deinterleaver.

$\varphi : X \rightarrow Y$: The over-complete mapping from set X to set Y

Glossary

ARQ	Automatic Request, Automatic request for retransmission of corrupted data
AWGN	Additive White Gaussian Noise
BCH	Bose-Chaudhuri-Hocquenghem. A class of forward error correcting codes (FEC)
CABAC	Context-based Adaptive Binary Arithmetic Coding, used for video compression in Dirac, H.264.
CAVLC	Context-adaptive variable-length coding
CIF	Common Intermediate Format Frames containing 352 pixels vertically and 288 pixels horizontally
DCT	A discrete cosine transform, transforms data into the frequency domain. Commonly used for video compression by removing high frequency components in the video frames
DPCM	Differential Pulse-Code Modulation, the basic scheme for predictive compression technique
DWT	Discrete Wavelet Transform, transforms data into the frequency domain, used in Dirac video codec, JPEG2000
FEC	Forward Error Correction
GLDPC	The Generalized Low-Density Parity-Check Code
GOP	Group Of Pictures, a particular structure for the encoded sequence of picture in video compression

GPRS	General Packet Radio Service
HDTV	High-Definition Television
IRCCs	Irregular Convolutional Codes
ISCD	Iterative Source-Channel Decoding
ISI	Intersymbol Interference
ISO	The International Organization for Standardization
ITU-T	The ITU Telecommunication Standardization Sector (ITU-T) coordinates standards for telecommunications on behalf of the International Telecommunication Union (ITU)
JPEG	Joint Photographic Experts Group, also a video coding standard designed by this group that is widely used
JSCD	Joint Source-Channel Decoding
LDPC	The Low-Density Parity-Check Code
MB	An abbreviation for Macroblock
MCER	the Motion-Compensated Error Residual is formed by subtracting the current block from its predicted block. Commonly used for video compression by removing temporal redundancy between adjacent video frames
ME	Motion Estimation, estimations of the motions in a frame
MPEG	Motion Picture Expert Group, also a video coding standard designed by this group that is widely used
MSE	Mean Square Error.
NSC	Non-Systematic Convolutional Codes
PCM	The Parity-Check Matrix (PCM) of GLDPC code

PSNR Peak Signal to Noise Ratio, noise energy compared to the maximum possible signal energy. Commonly used to measure video image quality

QCIF Quarter Common Intermediate Format Frames containing 176 pixels vertically and 144 pixels horizontally

RCPC Rate Compatible Punctured Convolutional codes

RCPT Rate Compatible Punctured Turbo codes entryEXIT EXtrinsic Information Transfer

SISO Soft-Input Soft-Output decoding

UEP Unequal Error Protection, a transmission scheme allocates a lower coding rate to highly sensitive bits and a higher coding rate to less sensitive bits

VBSMC Variable block-size motion compensation, used for video compression in H.264/AVC/MPEG-4 Part 10

VCEG The ITU-T Video Coding Experts Group

VQ vector quantization. Commonly used for video compression by removing spatial redundancy

Bibliography

[1] C. E. Shannon, "Coding theorems for a discrete source with a fidelity criterion," *Institute of Radio Engineers, International Convention Record*, vol. 7, Part 4, pp. 142–163, 1959.

[2] R. Koenen, "MPEG4: Multimedia for our time," *In IEEE Spectrum*, vol. 36, Number 2, pp. 26–33, February 1999.

[3] T. Sikora, "The MPEG-4 video standard verification model," *IEEE Transaction On Circuits And Systems For Video Technology*, vol. 7, pp. 19–31, February 1997.

[4] Joint Video Team (JVT) of ISO/IEC MPEG and ITU-T VCEG, *Joint Final Committee Draft (JFCD) of Joint Video Specification (ITU-T Rec. H.264 ISO/IEC 14496-10 AVC)*, August 2002.

[5] The ISO/IEC JPEG Committee, *JPEG2000 Verification Model 3.0 (ISO/IEC JTC1/SC29/WG11 JPEG WG1N 1137)*, November 1998.

[6] A. Luthra, G. J. Sullivan, and T. Wiegand, "Introduction to the special issue on the h.264/avc video coding standard," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 13, pp. 557–559, July 2003.

[7] L. Hanzo, P. J. Cherriman, and J. Streit, *Video Compression and Communications: From Basics to H.261, H.263, H.264, MPEG2, MPEG4 for DVB and HSDPA-Style Adaptive Turbo-Transceivers*. UK: John Wiley IEEE Press, 2007 2nd Edition.

[8] Sourceforge website, *http://sourceforge.net/projects/dirac*.

[9] ISO/IEC 13818-2 MPEG-2 Video Coding Standard, "Information technology - Generic coding of moving pictures and associated audio information: Video," March 1995.

[10] B. G. Haskell, A. Puri, and A. N. Netravali, *Digital Video: An Introduction to MPEG-2.* New York: Chapman & Hall, 1997.

[11] S. Vembu, S. Verdu, and Y. Steinberg, "The source-channel separation theorem revisited," *IEEE Transactions on Information Theory,*, vol. 41, pp. 44–54, January 1995.

[12] K. Immink, P. Siegel, and J. Wolf, "Codes for digital recorders," *IEEE Transaction Information Theory*, vol. 44, pp. 2260–2299, 1998.

[13] J. Bucklew and N. Gallagher, "Some properties of uniform step size quantizers," *IEEE Transactions on Information Theory*, vol. Volume 26, Issue 5, pp. 610–613, September 1980.

[14] M. G. Strintzis and D. Tzovaras, "Optimal construction of subband coders using Lloyd-Max quantizers," *IEEE Transactions on Image Processing*, vol. Volume 7, Issue 5, pp. 649–667, May 1998.

[15] R. Gray, "Vector quantization," *IEEE Acoustics, Speech and Signal Processing Magazine*, pp. 4–29, April 1984.

[16] A. Gersho and R. Gray, *Vector Quantization and Signal Compression.* Dordrecht: Kluwer Academic Publishers, 1992.

[17] L. Torres and J. Huguet, "An improvement on codebook search for vector quantisation," *IEEE Transactions on Communications*, vol. 42, pp. 208–210, February 1994.

[18] V. Sitaram, C. Huang, and P. Israelsen, "Efficient codebooks for vector quantisation image compression with an adaptive tree search algorithm," *IEEE Transactions on Communications*, vol. 42, pp. 3027–3033, November 1994.

[19] A. Gersho and B. Ramamurthi, "Image coding using vector quantization," *Proceding of IEEE international Conference on Acoustics, Speech, Signal Processing*, pp. 428–431, 1982.

[20] P. Mathieu, M. Antonini, M. Barlaud, and I. Daubechies, "Image coding using wavelet transform," *IEEE transaction on image processing*, vol. 1, pp. 205–220, 1992.

[21] J. Johnston, "Transform coding of audio signals using perceptual noise criteria," *IEEE Journal on Selected Areas of Communications*, vol. 6, no. 2, pp. 314–323, 1988.

[22] A. E. Jacquin, "Image coding based on a fractal theory of iterated contractive image transformation," *IEEE Transactions on Image Processing*, vol. 1, pp. 18–30, January 1992.

[23] J. O'Neal and T. Natarajan, "Coding isotropic images," *IEEE Transactions on Information Theory*, vol. 23, no. 6, pp. 697–707, 1977.

[24] D. Blasiak and W. Y. Chan, "Efficient wavelet coding of motion compensated prediction residuals," *In International Conference on Image Processing*, vol. 2, pp. 287–290, Chicago 1998.

[25] R. Young and N. Kingsbury, "Frequency-domain motion estimation using a complex lapped transform," *IEEE Transactions on Image Processing*, vol. 2, pp. 2–17, January 1993.

[26] R. Young and N. Kingsbury, "Video compression using lapped transforms for motion estimation/compensation and coding," in *Proceedings of the SPIE Communication and Image Processing Conference*, (Boston, MA), pp. 1451–1463, SPIE, November 1992.

[27] H. Yamaguchi, "Adaptive DCT coding of video signals," *IEEE Transactions on Communications*, vol. 41, pp. 1534–1543, October 1993.

[28] Y. J. Chang, C. W. Lin, and Y. C. Chen, "Hierarchical motion estimation algorithm based on pyramidal successive elimination," *Proceedings of International Computer Symposium*, pp. 41–44., October 1998.

[29] K. Rao and P. Yip, *Discrete Cosine Transform: Algorithms, Advantages and Applications.* New York: Academic Press Ltd., 1990.

[30] G. K. Wallace, "The JPEG still picture compression standard," *Communications of the Association for Computing Machinery*, vol. 34, no. 4, pp. 30–44, 1991.

[31] International Standards Organization, *ISO/IEC 11172 MPEG 1 International Standard, 'Coding of moving pictures and associated audio for digital storage media up to about 1.5 Mbit/s, Parts 1–3.*

[32] ITU-T, *Recommendation H.261: Video codec for audiovisual services at px64 Kbit/s*, March 1993.

[33] Y. Mayer, "Wavelet: Algorithm and application," *Society for Industrial and Applied Mathematics*, pp. 13–31, 101–105, 1993.

[34] G. Strang and T. Nguyen, *Wavelets and filter banks.* MA: Wellesley-Cambridge, 1996.

[35] I. Daubechies and W. Sweldens, "Factoring wavelet transform into lifting steps," *Journal of Fourier Analysis and Applications*, vol. 4, no. num. 3, pp. 247–269, 1998.

[36] Z. Li, F. Wu, S. Li, , and E. Delp, "wavelet video coding via a spatially adaptive lifting structure," *IEEE Conference on Acoustics, Speed Signal Processing*, vol. 3, pp. 93–96, 2003.

[37] ISO/IEC JTC1/SC29/WG11, "Information technology - coding of moving pictures and associated audio for digital storage media at up to about 1.5 Mbits/s. ," in *Part 3: Audio. ISO/IEC 11172-3:1993*, 1993.

[38] ISO/IEC and ITU-T, "Information technology - generic coding of moving pictures and associated audio," in *Part 2: Video. Draft ISO/IEC 13818-2 (MPEG-2) and ITU-T Recommendation H.262*, ISO/IEC, (Geneva), 1994.

[39] T. Wiegand, X. Zhang and B. Girod, "Long-term memory motion-compensated prediction," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 9, pp. 70–84, February 1999.

[40] B. Liu and A. Zaccarin, "New fast algorithms for the estimation of block motion vectors," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 3, pp. 148–157, April 1993.

[41] RealVideo website, *http://www.realnetworks.com/products/codecs/realvideo.html.*

[42] x264 VideoLAN website, *http://www.videolan.org/developers/x264.html.*

[43] R. E. V. Dyck and D. J. Miller, "Transport of wireless video using separate, concatenated, and joint source-channel coding," *Proceedings of IEEE*, vol. 87, pp. 1734–1750, Oct. 1999.

[44] M. Park and D. J. Miller, "Joint source-channel decoding for variable-length encoded data by exact and approximate map sequence estimation," *IEEE Transaction on Communications*, pp. 1–6, Jan.

[45] J. Kliewer and R. Thobaben, "Iterative Joint Source-Channel Decoding of Variable-Length Codes Using Residual Source Redundancy," pp. 919–929.

[46] Q. Chen and K. P. Subbalakshmi, "Joint source-channel decoding for MPEG-4 video transmission over wireless channels," *IEEE Journal on Selected Areas in Communications*, vol. 21, no. 10, pp. 1780–1789, 19-22 April 2003.

[47] F.-M. Wang and S. Liu, "Hybrid video coding for low bit-rate applications," vol. v, pp. 481–484, 1994.

[48] ITU, *Joint Photographic Experts Group ISO/IEC, JTC/SC/WG8, CCITT SGVIII. JPEG technical specifications, revision 5. Report JPEG-8-R5*, January 1990.

[49] ITU-T, "Video coding for low bitrate communication," *ITU-T Recommendation H.263; version 1*, November 1995.

[50] L. Chiariglione, "The development of an integrated audiovisual coding standard: Mpeg," *Proceedings of the IEEE*, vol. 83, pp. 151–157, February 1995.

[51] R. Schafer and T. Sikora, "Digital video coding standards and their role in video communications," *Proceedings of the IEEE*, vol. 83, pp. 907–924, June 1995.

[52] A. Skodras, C. Christopoulos, and T. Ebrahimi, "The jpeg 2000 still image compression standard," *IEEE Signal Processing Magazine*, vol. 18, pp. 36–58, September 2001.

[53] J. Yu, "Advantages of uniform scalar dead-zone quantization in image coding system," *International Conference on Communications, Circuits and Systems, ICCCAS 2004*, vol. 2, pp. 805–808, 27-29 June 2004.

[54] D. Marpe, H. Schwarz, and T. Wiegand, "Context-based adaptive binary arithmetic coding in the h.264/avc video compression standard," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 13, Issue 7, pp. 620–636, July 2003.

[55] D. Marpe, G. Blättermann, G. Heising, and T. Wiegand, "Adaptive codes for h.26l," *ITU-T SG16/Q.6 VCEG-L-13*, January 2001.

[56] ITU-T Rec. H.26L/ISO/IEC 11496-10, "Advanced video coding," *Final Committee Draft, Document JVT-E022*, September 2002.

[57] ITU-T/SG 16/VCEG(formerly Q.15 now Q.6), "H.26L test model long term number 7 (TML-7), Doc. VCEG-M81," April 2001.

[58] G. J. Sullivan, T. Wiegand and T. Stockhammer, "Draft H.26L video coding standard for mobile applications," in *Proceedings of IEEE International Conference on Image Processing*, pp. 573–576, October 2001.

[59] T. Wiegand, G. Sullivan, G. Bjntegaard, and A. Luthra, "Overview of the H.264/AVC video coding standard," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 13, Issue: 7, pp. 560–576, July 2003.

[60] T. Wedi, "Motion compensation in h.264/avc,," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 13, pp. 577–586, July 2003.

[61] H. Y. Lin, Y. H. Lu, B. D. Liu, and J. F. Yang, "A Highly Efficient VLSI Architecture for H.264/AVC CAVLC Decoder," *IEEE Transactions on Multimedia*, vol. Volume 10, Issue 1, pp. 31–42, January 2008.

[62] T. C. Chen, Y. W. Huang, C. Y. Tsai, B. Y. Hsieh, and L. G. Chen, "Architecture Design of Context-Based Adaptive Variable-Length Coding for H.264/AVC," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. Volume 53, Issue 9, pp. 832–836, September 2006.

[63] ISO/IEC/JTC 1/SC 29, *ITU-T Recommendation H.264 (2005) — ISO/IEC 14496-10:2005 'Information technology - Coding of audio-visual objects- Part 10: Advanced Video Coding'*, 2005.

[64] L. Hanzo, T. Liew, and B. Yeap, *Turbo Coding, Turbo Equalisation and Space Time Coding for Transmission over Wireless channels.* New York, USA: John Willy IEEE Press, 2002.

[65] S. Lin, D. C. Jr., and M. Miller, "Automatic-repeat-request error-control schemes," *IEEE Communications Magazine*, vol. 22, pp. 5–17, December 1984.

[66] G. Djuknic and D. Schilling, "Performance analysis of an ARQ transmission scheme for meteor burst communications," *IEEE Transactions on Communications*, vol. 42, pp. 268–271, February/March/April 1994.

[67] S. Kallel, "Analysis of a type ii hybrid arq scheme with code combining," *IEEE Transactions on Communications*, vol. 38, pp. 1133–1137, August 1990.

[68] J. Hamorsky and L. Hanzo, "Performance of the turbo hybrid automatic repeat request system type ii," *IEEE Information Theory Workshop (ITW'99)*, p. 51, 27 June-1 July, Metsovo, Greece 1999.

[69] W. Tan and A. Zakhor, "Video multicast using layered fec and scalable compression," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 11, pp. 373–387, March 2001.

[70] P. C. Cosman, J. K. Rogers, P. G. Sherwood, and K. Zeger, "Image transmission over channels with bit errors and packet erasures," *Proceedings Thirty-Fourth Asilomar Conference on Signals, Systems and Computers, (Pacific Grove, CA)*, pp. 1621–1625, 1998.

[71] B. Girod, K. Stuhlmuller, M. Link, and U. Horn, "Packet loss resilient internet video streaming," *in Proceedings of Visual Communications and Image Processing VCIP-99, (San Jose, CA)*, pp. 833–844, January 1999.

[72] J. G. Apostolopoulos, "Error-resilient video compression via multiple state streams," *International Workshop on Very Low Bitrate Video Coding (VLBV99), Kyoto, Japan*, pp. 168–172, October 1999.

[73] L. Hanzo, J. S. Blogh, and S. Ni, "3G, HSDPA, HSUPA and FDD Versus TDD Networking: Smart Antennas and Adaptive Modulation." (http://www-mobile.ecs.soton.ac.uk.), 2008.

[74] Harri Holma and Anti Toskala, *HSDPA/HSUPA for UMTS.* Chichester, UK: John Wiley and Sons, 2006.

[75] R. Aravind, M. R. Civanlar, and A. R. Reibman, "Packet loss resilience of mpeg-2 scalable video coding algorithms," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 6, pp. 426–435, October 1996.

[76] I. J. WG11, "Information technology - generic coding of audio-visual objects,," in *in Part 2: Visual. Draft ISO/IEC 14496-2 (MPEG-4), version 1*, ISO/IEC, (Geneva), 1998.

[77] U. Horn, K. Stuhlmuller, M. Link, and B. Girod, "Robust internet video transmission based on scalable coding and unequal error protection," *Signal Processing: Image Communication*, vol. 15, pp. 77–94, September 1999.

[78] H. Yang, R. Zhang, and K. Rose, "Drift management and adaptive bit rate allocation in scalable video coding," *In Proceedings of the IEEE International Conference on Image Processing (ICIP)*, vol. 2, pp. 49–52, September 2002.

[79] R. Thobaben and J. Kliewer, "Robust decoding of variable-length encoded markov sources using a three-dimensional trellis," *IEEE communications letters*, vol. 7, pp. 320–322, July 2003.

[80] C. Berrou, A. Glavieux, and P. Thitimajshima, "Near shannon limit error-correcting coding and decoding: Turbo-codes," *IEEE International Conference on Communications (ICC) 93, Geneve, Switzerland*, vol. 2, pp. 1064–1070, 23-26 May 1993.

[81] L. Hanzo, F. Somerville, and J. Woodard, "Voice and audio compression for wireless communications: Principles and applications for fixed and wireless channels." (http://www-mobile.ecs.soton.ac.uk.), 2007.

[82] A. Q. Pham, J. Wang, L. L. Yang, and L. Hanzo, "An iterative detection aided unequal error protection wavelet video scheme using irregular convolutional codes," *IEEE 63rd Vehicular Technology Conference, Melbourne, Australia,*, pp. 2484 – 2488, May 2006.

[83] Sourceforge website, *http://sourceforge.net/projects/ffmpeg.*

[84] A. Q. Pham, L. L. Yang, and L. Hanzo, "Joint optimization of iterative source and channel decoding using over-complete source-mapping," *in Proceedings of IEEE 66th Vehicular Technology Conference VTC2007-Fall Baltimore*, pp. 1072–1076, 30 September - 3 October 2007.

[85] A. Q. Pham, L. L. Yang, N. S. Othman, and L. Hanzo, "Exit-chart optimized block codes for wireless video telephony." Accepted for publication in the IEEE Transactions on Circuits and Systems for Video Technology, 2008.

[86] A. Q. Pham, J. Wang, L. L. Yang, and L. Hanzo, "Unequal error protection irregular over-complete mapping for wavelet coded wireless video telephony using iterative source and channel decoding," *in Proceedings of the IEEE Wireless Communications and Networking Conference (WCNC-2008), Las Vegas, Nevada, USA*, pp. 1251 – 1255, 31 March-3 April 2008.

[87] M. G. Albanesi, I. de Lotto, and L. Carrioli, "Image compression by the wavelet decomposition," *European Transactions on Telecommunications and Related Technologies(ISSN 1120-3862)*, vol. 3, no. 3, pp. 265–274, May-June 1992.

[88] D. Marpe and H. L. Cycon, "Very low bit-rate video coding using wavelet-based techniques," *IEEE Transactions on Circuits and Systems for Video Technology,*

vol. Volume 9, Issue 1, pp. 85–94, Feb. 1999.

[89] S. G. Mallat, "A theory for multiresolution signal decomposition: the wavelet representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. Volume 11, Issue 7, pp. 674–693, July 1989.

[90] A. Grapes, "An introduction to wavelets," *IEEE Computational Science and Engineering*, vol. 2, no. num. 2, pp. 50–61, Summer, 1995. (published by the IEEE Computer Society).

[91] M. Vetterli and C. Herley, "Wavelet and filter bank: Theory and design," *IEEE transactions on signal processing*, vol. 40, pp. 2207–2232, 1992.

[92] Y. Q. Shi and H. Sun, *Image and Video Compression for Multimedia Engineering.* CRC Press, 2000.

[93] I. H. Witten, R. M. Neal, and J. G. Cleary, "Arithmetic coding for data compression," *Communications of the ACM*, vol. 30, No. 6, pp. 520–540, 1987.

[94] D. Marpe and H. L. Cycon, "Efficient pre-coding techniques for wavelet-based image compression," *Proceedings of International Picture Coding Symposium (PCS'97)*, pp. 45–50, 1997.

[95] Z. Ya-Qin and Z.Sohail, "Motion-compensated wavelet transform coding for color video compression," *IEEE Transactions on Circuits and Systems for Video technology*, vol. 2, No. 3, pp. 285–296, September 1992.

[96] K. M. Nam, J. S. Kim, R. H. Park, and Y. S. Shim, "A fast hierarchical motion vector estimation algorithm using mean pyramid," *IEEE Transactions on Circuits and Systems for Video technology*, vol. 5, pp. 344–351, August 1995.

[97] J. Vanne, E. Aho, T. D. Hamalainen, and K. Kuusilinna, "A high-performance sum of absolute difference implementation for motion estimation," *IEEE Transactions on Circuits and Systems for Video technology*, vol. 16, No. 7, pp. 876–883, July 2006.

[98] D. Agrafiotis, D. R. Bull, and C. N. Canagarajah, "Enhanced error concealment with mode selection," *IEEE Transactions on Circuits and Systems for Video technology*, vol. 16, No. 8, pp. 960–973, August 2006.

[99] B. Yan and N. K. Wing, "A novel selective motion vector matching algorithm for error concealment in mpeg-4 video transmission over error-prone channels," *IEEE Transactions on Consumer Electronics*, vol. 49, No. 4, pp. 1416–1423, November 2003.

[100] T. Berger, *Rate Distortion Theory.* Englewood Cliffs, NJ: Prentice-Hall, 1971.

[101] G. J. Sullivan and T. Wiegand, "Rate-distortion optimization for video compression," *IEEE Signal Processing Magazine*, vol. 15, pp. 74–90, November 1998.

[102] J. Hagenauer, "Rate-compatible punctured convolutional codes (rcpc codes) and their applications," *IEEE Transactions on Communications*, vol. 36, pp. 389–400, April 1988.

[103] F. Babich, G. Montorsi, and F. Vatta, "Design of rate-compatible punctured turbo (rcpt) codes," *IEEE International Conference on Communications,*, vol. 3, pp. 1701–1705, 28 April-2 May 2002.

[104] M. Tüchler, "Design of serially concatenated systems depending on the block length," *IEEE Transactions on Communications*, vol. 52, pp. 209–218, February 2004.

[105] M. Tüchler and J. Hagenauer, "Exit charts of irregular codes," *in Proceedings of Conference on Information Science and Systems, Princeton University [CDROM]*, March 2002.

[106] S. ten Brink, "Convergence behaviour of iteratively decoded parallel concatenated codes," *IEEE Transactions on Communications*, pp. 1727–1737, October 2001.

[107] I. Lee, "The effect of precoder on serially concatenated coding systems with an isi channel," *IEEE Transactions on Communications*, vol. 49, pp. 1168–1175, July 2001.

[108] A. G. Lillie, A. R. Nix, and J. P. McGeehan, "Performance and design of a reduced complexity iterative equalizer for precoded isi channels," *in IEEE Vehicular Technology Conference (Orlando, USA)*, pp. 299–303, October 2003.

[109] J. Wang, N. S. Othman, L. L. Yang, and L. Hanzo, "Turbo detected unequal error protection general configuration irregular convolutional codes designed for the wideband advanced multirate speech codec," *In Proceedings of the IEEE 62nd Semiannual Vehicular Technology Conference*, 25-28 September 2005.

[110] D. Raphaeli, "Combined turbo equalization and turbo decoding," *IEEE Communications Letters*, vol. 2, pp. 107–109, April 1998.

[111] I. Land, S. Huettinger, P. A. Hoeher, and J. Huber, "Bounds on information combining," *in Procceding of International Symposium on Turbo Codes & Related Topics, Brest, France*, pp. 39–42, September 2003.

[112] J. Proakis, *Digital Communications.* 4th edition: New York: McGraw-Hill, 2001.

[113] L. R. Bahl, J. Cocke, F. Jelinek, and J. Raviv, "Optimal decoding of linear codes for minimal symbol error rate," *IEEE Transactions on Information Theory*, vol. 20, pp. 284–287, March 1974.

[114] M. Adrat, T. Clevorn, J. Brauers, and P. Vary, "Minimum terms of residual redundancy for successful iterative source-channel decoding," *Communications Letters, IEEE,*, vol. 10, Issue 11, pp. 778 – 780, November 2006.

[115] V. Buttigieg and P. G. Farrell, "Variable-length error-correcting codes," *IEE Proceedings in Communications*, vol. 147, no. 4, pp. 211–215, August 2000.

[116] M. Adrat, P. Vary, and J. Spittka, "Iterative source-channel decoder using extrinsic information from softbit-source decoding," *IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 2653–2656, 7-11 May 2001.

[117] R. G. Maunder and L. Hanzo, "Concatenated irregular variable length coding and irregular unity rate coding." Submitted to IEEE Transactions on Wireless Communications, 2007.

[118] R. Bauer and J. Hagenauer, "On variable length codes for iterative source/channel decoding," *in Procceding of IEEE Data Compression Conference(DCC'01), Snowbird, UT, USA*, pp. 272–282, March 2001.

[119] J. Hagenauer, "The turbo principle - tutorial introduction and state of the art," *in Procceding of International Symposium on Turbo Codes & Related Topics, Brest, France*, pp. 1–11, September 1997.

[120] R. Bauer and J. Hagenauer, "Symbol-by-symbol map decoding of variable length codes," *in Procceding of the 3rd ITG-Conference on Source and Channel Coding*, pp. 111–116, January 2000.

[121] M. Lentmaier and K. Zigangirov, "On generalized low-density parity-check codes based on hamming component codes," *IEEE Communications Letters*, vol. 3, No 8, pp. 248–250, August 1999.

[122] R. G. Gallager, "Low density parity-check codes," *MIT Press, Cambridge, MA*, 1963.

[123] R. M. Tanner, "A recursive approach to low complexity codes," *IEEE Transactions on Information Theory*, pp. 533–547, September 1981.

[124] J. Boutros, O. Pothier, and G. Zemor, "Generalized low density (tanner) codes," *IEEE Internationl Conference on Communications (ICC 1999)*, vol. 1, pp. 441–445, June 1999.

[125] R. Y. S. Tee, F. Kuo, and L. Hanzo, "Multilevel generalised low-density parity-check codes," *Electronics Letters*, vol. 42 (3), pp. 167–168, 2006.

[126] A. Ashikhmin, G. Kramer, and S. ten Brink, "Code rate and the area under extrinsic information transfer curves,," *in IEEE International Symposium on Information Theory*, p. 115, 2002.

[127] M. Adrat, "Iterative source-channel decoding for digital mobile communications," *RWTH Aachen University: Ph.D Thesis*, 2003.

[128] M. Adrat and P. Vary, "Iterative source-channel decoding: improved system design using exit charts," *EURASIP Journal on Applied Signal Processing*, pp. 928–941, 2005.

[129] M. Tesanovic, D. R. Bull, D. Agrafiotis, and A. Doufexi, "Comparison of standard-based H.264 error-resilience techniques and multiple-description coding for robust MIMO-enabled video transmission," *Proceedings of the SPIE Visual Communications and Image Processing (VCIP), San Jose.*, vol. 6508, pp. 6508–05, 2007.

[130] N. S. Othman, M. El-Hajjar, A. Q. Pham, O. Alamri, S. X. Ng, and L. Hanzo, "Over-complete source-mapping aided AMR-WB MIMO Transceiver using three-stage iterative detection," *in Proceedings of the 2008 IEEE International Conference on Communications (ICC 2008), CD-ROM, Beijing, China*, 19-23 May 2008.

[131] Nasruminallah, M. El-Hajjar, N. S. Othman, A. Q. Pham, and L. Hanzo, "Over-complete mapping aided, soft-bit assisted iterative unequal error protection H.264 joint source and channel decoding," *accepted for the 68th IEEE Vehicular Technology Conference, Calgary, Canada*, 21 - 24 September, 2008.

[132] S. X. Ng, J. Y. Chung, and L. Hanzo, "Turbo-detected unequal protection mpeg-4 wireless video telephony using trellis coded modulation and space-time trellis coding," *IEE Proceedings on Communications*, vol. 152, Issue 6, pp. 1116–1124, December 2005.

[133] Y. Wang and S. Yu, "Joint Source-Channel Decoding for H.264 Coded Video Stream," *IEEE Transactions on Consumer Electronics*, vol. 51, No.4, pp. 1273–1276, November 2005.

[134] S. X. Ng, F. Guo, J. Wang, L.-L. Yang, and L. Hanzo, "Joint source-coding, channel-coding and modulation schemes for AWGN and Rayleigh fading channels," *Electronics Letters*, vol. 39, pp. 1259–1261, Aug. 2003.

[135] C. Bergeron and C. Lamy-Bergot, "Soft-input decoding of variable-length codes applied to the H.264 standard," in *IEEE Workshop on Multimedia Signal Processing*, pp. 87–90, 2004.

[136] R. Thobaben and J. Kliewer, "Low-complexity iterative joint source-channel decoding for variable-length encoded Markov sources," *IEEE Transactions on Communications*, vol. 53, pp. 2054–2064, December 2005.

[137] B. Vucetic and J. Yuan, *Turbo Codes Principles and Applications*. The Netherlands: Kluwer Academic Publishers, 2000.

[138] P. Jung and M. Nasshan, "Performance evaluation of turbo codes for short frame transmission systems," *IEE Electronic Letters*, pp. 111–112, January 1994.

[139] P. Jung, "Novel low complexity decoder for turbo-codes," *IEE Electronic Letters*, pp. 86–87, 19 January 1995.

[140] S. Benedetto and G. Montorsi, "Unveiling turbo codes: Some results on parallel concatenated coding schemes," *IEEE Transactions on Information Theory*, vol. 42, pp. 409–428, March 1996.

[141] C. Berrou and A. Glavieux, "Near optimum error correcting coding and decoding: turbo codes," *IEEE Transactions on Communications*, vol. 44, pp. 1261–1271, October 1996.

[142] S. Benedetto and G. Montorsi, "Design of parallel concatenated convolutional codes," *IEEE Transactions on Communications*, vol. 44, pp. 591–600, May 1996.

[143] G. Caire and G. Lechner, "Turbo codes with unequal error protection," *IEE Electronic Letters*, vol. 32, pp. 629–631, 28 March 1996.

[144] P. Hoeher, "New Iterative (Ṫurbo̱) Decoding Algorithms," in *International Symposium on Turbo Codes and related topics*, (Brest, France), pp. 63–70, September 1997.

[145] G. Battail, "A conceptual framework for understanding turbo-codes," in *International Symposium on Turbo Codes and related topics*, (Brest, France), pp. 55–62, September 1997.

[146] M. W. Y. Takishima and H. Murakami, "Reversible variable length codes," *IEEE Transactions on Communications*, vol. 43, pp. 158–162, Feb/Mar/Apr 1995.

[147] J. D. V. J. Wen, "A class of reversible variable length codes for robust image and video coding," *Proceedings 1997 IEEE International Conference on Image Processing*, vol. 2, pp. 65–68, October 1997.

[148] Takishima, Y. and Wada, M. and Murakami, H., "Reversible variable length codes," *IEEE Transactions on Communications*, vol. 43, no. 234, pp. 158–162, 1995.

[149] J. Wang and L.-L. Yang and L. Hanzo, "Iterative construction of reversible variable-length codes and variable-length error-correcting codes," *IEEE Communications Letters*, vol. 8, pp. 671–673, November 2004.

[150] T. Clevorn, M. Adrat, and P. Vary, "Turbo decodulation using highly redundant index assignments and multi-dimensional mappings," *4th International Symposium on Turbo Codes & Related Topics in connection with 6th Int. ITG-Conference on Source and Channel Coding, Munich, Germany [CDROM]*, April 2006.

Index

P

Q

R

S

U

V