UNIVERSITY OF CALIFORNIA, SAN DIEGO

Image and Video Transmission over Noisy Channels

A dissertation submitted in partial satisfaction of the

requirements for the degree Doctor of Philosophy

in

Electrical Engineering (Communications Theory and Systems)

by

Xiaodong Tian

Committee in charge:

Professor Kenneth Zeger, Chair
Profesoor Pamela Cosman
Professor William Hodgkiss
Professor Lawrence Milstein
Professor Geoffrey M. Voelker

2005

The dissertation of Xiaodong Tian is approved. And it is acceptable in quality and form for publication on microfilm:

_____

_____

_____

_____

_____

_____

Chair

University of California, San Diego

2005

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

First, I would like to express my appreciation to my advisor Professor Kenneth Zeger for giving me the opportunity to conduct Ph.D research at UCSD, and for his support and help over the years. I would also like to thank Professors Pamela Cosman, William Hodgkiss, Lawrence Milstein, and Geoffrey Voelker for serving on my committee.

I also want to thank my lab mates and friends for their help. An incomplete list follows: Greg Sherwood, Tamas Frajka, Qinghua Zhao, Dirck Schilling, Jon Rogers, Song Cen, Yan Ye, Hugo Tullberg, Jilei Hou, and Yan Zhang.

Last but not least, I am grateful to my parents, my brother, and my wife for their love, support and encouragement. This dissertation is dedicated to them.

The text of Chapter 2, in part, is a reprint of the material as it appears in: P.G. Sherwood, X. Tian, and K. Zeger. Channel code blocklength and rate optimization for progressive image transmission. *IEEE Wireless Communications and Networking Conference*, pp 978-982, Vol.2, Sept. 1999. The text of Chapter 3, in part, is a reprint of the material as it appears in: P.G. Sherwood, X. Tian, and K. Zeger. Efficient image and channel coding for wireless packet networks. *International Conference on Image Processing*, pp 132-135, Vol. 2, Sept. 2000. The text of Chapter 4, in full, has been submitted for publication as: X. Tian. Asymptotic Analysis of Multiple Description Codes Over Noisy Channels, *IEEE Transactions on Information Theory*, 2004. The text of Chapter 5, in part, has been submitted for publication as: X. Tian. Punctured trellis coded quantization and modulation. *IEEE Transactions on Communications*, 2004. The

text of Chapter 6, in part, is a reprint of the material as it appears in: X. Tian. Efficient transmission power allocation for wireless video communications. *Proc. of IEEE Wireless Communications and Networking Conference, Atlanta*, 2004, Vol. 4, pp. 2058–2063, March 2004; and in part, has been submitted for publication as: X. Tian. Efficient transmission power allocation for wireless video communications under average power constraints. *IEEE Transactions on Circuits and Systems for Video Technology*, 2004. With the exception of the first two publications I was the primary researcher and my advisor Kenneth Zeger supervised the research which forms the basis for this dissertation.

ABSTRACT OF THE DISSERTATION

Image and Video Transmission over Noisy Channels

by

Xiaodong Tian

Doctor of Philosophy in Electrical Engineering

(Communications Theory and Systems)

University of California, San Diego, 2005

Professor Kenneth Zeger, Chair

This thesis studies how to effectively transmit images, video or generic types of sources over noisy channels. Source-channel coding methods such as rate allocation between source and channel codes, unequal error protection, and error-resilient source coding, are explored.

First, we consider the problem of progressive image coding over noisy channels. Motivated by turbo coding methods where performance improves with blocklength, we investigate techniques of optimally selecting both the channel code blocklength and the code rate to balance the tradeoff among error protection, source coding rate, and delay.

Next, we investigate the problem of image coding over packet loss channels. In order to improve error resiliency, our proposed scheme combines multiple description quantization, entropy coding, and data partitioning.

Next, we consider a cascade model of a multiple description coder followed by

block channel for transmission over noisy channels. A high-rate analysis shows the optimized cascade multiple description scheme may outperform the optimized cascade single description scheme only under certain conditions such as time-varying channels with channel states unknown to the encoder.

Next, we generalize a combined Trellis Coded Quantization/Modulation (TCQ/TCM) scheme by puncturing the component TCQ and TCM coders, and then optimize the tradeoff between source coding and channel coding for the given trellis. Simulation results show that significant performance improvement can be obtained for the proposed scheme over the fixed-rate TCQ/TCM scheme.

Lastly, we consider the problem of transmission power allocation for video transmission over a wireless channel. Assuming fixed-length packetization and adaptive FEC, we propose simple optimization algorithms to allocate transmission power to video packets under average transmission power constraints, with or without the channel state information (CSI) available to the transmitter. Simulations results demonstrate the performance improvement of the proposed schemes over the conventional schemes.

Chapter 1

Introduction

In this dissertation, we study the transmission of images, video and generic types of sources over noisy channels, such as wireless communication channels. Channel coding is crucial for such systems to achieve adequate performance. Jointly designing the source coder and the channel coder/transmitter has been proved to be beneficial in many cases.

1.1 Overview of Source-Channel Coding

A typical communication system consists of a system encoder, a channel, and a system decoder. As shown in Figure 1.1, the system encoder can be broken down to a source encoder, a channel encoder, and a modulator. The source encoder reduces the redundancy in the source to save transmission bandwidth while converting the source into strings of codewords as input to the channel encoder, the channel encoder adds useful redundancy to combat channel errors, and the modulator converts the coded bits into waveforms suitable for transmission. The system decoder can be broken down similarly.

A well-known theorem due to Shannon states that with limitless source coding

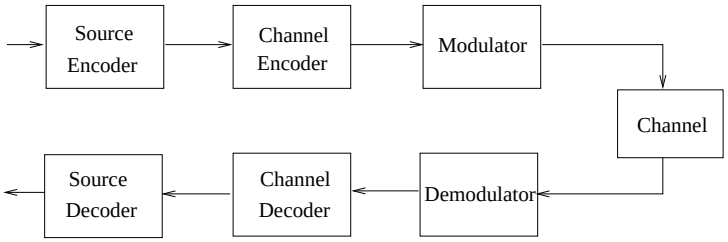


Figure 1.1: Diagram of a Communication System

dimension and channel coding blocklength, optimal source coding and channel coding can be considered separately without loss of optimality. The optimal channel coder is designed to achieve a practically error-free performance at the channel capacity, and the optimal source coder is designed to approach the source rate-distortion function at the given bit-rate.

However, with practical constraints on delay and complexity, the separation theorem no longer holds. In many practical communication systems such as wireless video communication systems, jointly designing the source coder and channel coder can improve the system performance significantly. This approach is termed "joint source-channel coding", or simply "source-channel coding".

Below we briefly discuss some source-channel coding methods and their applications in image and video transmission.

- Rate Allocation Between Source Coding and Channel Coding

  Consider a tandem system with a source coder and a channel coder, which may be designed separately. Given the transmission rate constraint, it is easy to see that there exists a tradeoff between source coding and channel coding in terms of the end-to-end distortion. If too much rate is allocated to channel coding, the error probability of channel decoding may be low, but because the available rate for source coding is small, the distortion due to source coding will be large. On

the other hand, if too much rate is allocated to source coding, the distortion due to source coding may be small, but because there is no adequate error protection, the distortion due to channel errors will be large.

This tradeoff between source and channel coding is analyzed theoretically in [38], where the asymptotic (high rate) upper and lower bounds on the channel code rate that minimizes the average distortion of a vector quantizer of fixed finite dimension are established. It is shown that given a total transmission rate constraint, the optimal channel code rate is less than the channel capacity.

As stated in [27], the optimal rate allocation between source codes and channel codes is one of the most important aspects of any source-channel coding designs. Examples of rate allocation between source codes and channels codes in image and video coding can be found in [56, 9]. Our work in Chapter 5 on a generalized Trellis Coded Quantization/Modulation scheme can also be classified into this category.

- Source-Optimized Channel Coding

  In source-optimized channel coding, the source code is designed for a noiseless channel. A channel code is then designed for this source code to minimize the end-to-end distortion for the given noisy channel. While there exist source codes such as fixed-length codes that are robust to channel errors, variable-length codes (VLCs) are extremely sensitive to channel errors. Careful packetization or insertion of synchronization codewords may improve the robustness of the source-channel coder.

  For example, the so-called Set-Partitioning-In-Hierarchical-Trees (SPIHT) coder is an efficient progressive image coder, i.e., the image is decodable at intermediate bit-rates and the decoding quality improves as more bits are received [66]. In [71],

SPIHT-encoded images are protected with Forward-Error-Control (FEC) codes over Binary Symmetric Channels (BSCs). The SPIHT code was partitioned into fixed-length blocks, followed by a channel coder consisting of a concatenation of a Cyclic Redundancy Check (CRC) coder for error detection and a Rate Compatible Punctured Convolutional (RCPC) coder for error correction. The channel code rate was selected empirically based on the channel condition such that practically no distortion due to channel noise could be detected.

Since in practice some source-coded bits are often more sensitive to channel errors than others, the error protection can be applied in an unequal fashion based on the importance of the information to provide improved efficiency and some graceful degradation. Examples of unequal error protection (UEP) for image coding can be found in [54, 61]. Our work in Chapter 2 on progressive image coding also falls into this category, where stronger FEC was used at earlier part of the coded image since the source bitstream is only decodable up to the first error.

For video coding, UEP can be applied to the layers of a scalable video bitstream. The breakup of a frame into subsets of varying quality lends itself naturally to employing a UEP scheme, in which the base layer is better protected than the enhanced layers. Examples of this approach can be found in [45, 46]. Our work in Chapter 6, where transmission power is allocated to a single layer of video packets according to their error sensitivity, can also be classified into this category.

- Channel-Optimized Source Coding

  In channel-optimized source coding, redundancy in source coding is utilized to make source coded symbols robust to channel errors. For example, vector quantizers (VQs) can be made robust to channel errors by carefully designing the index assignment that maps source codewords to binary indices [86]. Without modi-

fying the VQ codewords, proper index assignment makes sure that likely channel errors cause relatively small increase in distortion. Another example is the channel-optimized vector quantizer (COVQ), where the encoding regions (source encoder) and the codewords (source decoder) are iteratively optimized for a given channel crossover probability [21]. Since VQs are fixed-length codes, they are more robust to channel errors than VLCs.

Channel-optimized source codes have also been developed for VLCs. For example, in [63] an error-resilient entropy code is designed for transmitting videos over noisy channels. A class of reversible variable length codes is described in [81]. The basic idea here is that after a loss of synchronization due to a channel decoding error, the decoder will determine an appropriate codeword boundary and decode in the reverse direction to recover the bits in between.

Multiple Description Coding (MDC) can be considered as a special case of channel-optimized source coding that matches the source coder to a packet (description) loss channel. Multiple descriptions are generated for a source with correlation between the descriptions. The goal of MDC is to achieve some acceptable quality if only a subset of the description is received, and an improved quality if more descriptions are received. An example of MDC based image coding can be found in [68]. In Chapter 3, we will also employ a similar approach for robust image coding over packet loss channels.

- Jointly-Optimized Source-Channel Coding

  Jointly optimizing a source code and a channel code is often difficult. In practice, source-optimized channel coding can often be combined with channel-optimized source coding using an iterative design, which is a locally optimal technique. An example of this approach is [27], where COVQ and RCPC codes were itera-

tively optimized, with the addition of optimal bit allocation. An iterative decoding scheme of VLCs and channel codes was given in [31].

Our work in Chapter 4, where the channel code rate r and multiple description redundancy parameter a are jointly optimized in a cascade model of a multiple description coder followed by block channel coders, also falls into this category.

So far we have assumed that the encoder has accurate knowledge of the current channel state information (CSI) so that the source-channel coder can use this information to improve the end-to-end performance. For a fixed channel, the channel gain information can be considered as a special case of the CSI. However, in practical systems, the channel may be unknown to the encoder or the CSI known to the encoder may be inaccurate due to the time-varying nature of the channel or inaccurate channel estimation. Some discussions on the source-channel coding strategies in this scenario can be found in [19, 24, 39].

It is worthy of mentioning that methods that improve the channel performance, such as diversity reception, channel equalization, and space-time coding, can be directly used with source-channel coding; they change the channel transition probabilities used by the source-channel coder.

1.2 Outline of This Dissertation

This dissertation presents our work in several aspects of source-channel coding/transmission over noisy channels, including that of images, video and generic sources.

In Chapter 2, we study the problem of progressive image coding over noisy channels. If decoding at multiple rates is not important, then using long blocklength channel codes provides the best performance in many cases, as shown in the case of turbo coding where performance improves with blocklength. One penalty associated

with long blocklength codes is increased decoding complexity. For progressive coding, there is the additional penalty of decoding delay. In order to balance the tradeoff among error protection, source coding rate, and delay, we investigate techniques of optimally selecting both the channel code blocklength and code rate. We propose a general performance measure for evaluating progressive transmission and use dynamic programming to determine the channel code parameters based on the progressive performance. Performance results are provided for packet erasure channels and bit error channels, and are shown to outperform conventional methods where fixed number of information bits are used in each codeword.

In Chapter 3, we design a multiple description based image coding scheme over packet loss channels. In order to improve the error resiliency of images over varying packet loss channels, multiple description coding is combined with data partitioning and entropy coding in our scheme, and a multidimensional extension of the two-channel multiple description scalar quantizer (MDSQ) is proposed. In addition, the SPIHT image coder is employed for efficient entropy coding of multiple description images. The performance of our proposed scheme degrades gracefully as channel conditions worsen. Experimentally, with 4 descriptions, the new coder outperforms previous reports at any loss rate.

In Chapter 4, we investigate a cascade model of a multiple description coder followed by block channel coders for source-channel coding of Gaussian sources over noisy channels. Traditionally multiple description coding (MDC) is designed for packet loss channels. Our motivation here is to apply MDC to noisy channels such as wireless channels where FEC codes need to be used to achieve adequate performance. We conduct a high-resolution analysis to minimize a distortion upper bound over the channel code rate r and multiple description (MD) redundancy parameter a. Analytic results show that for time-invariant channels, in most channel conditions, the optimized single

description (SD) scheme outperforms the optimized MD scheme; when the channels are time-varying and the channel states are unknown to the encoder, the MD scheme may outperform the SD scheme. This result is useful for determining whether a multiple description coder as opposed to a single description coder should be used in conjunction with channel coding.

In Chapter 5, we study a joint source-channel coding/modulation method. In particular, we develop a generalized Trellis-Coded Quantization/Modulation scheme, which is a low-complexity source-channel coding scheme that takes modulation into account. One of our motivations is to enforce proper rate allocation in the scheme to match it with the given channel, which, as discussed earlier in this chapter, is very important for achieving low end-to-end distortion by source-channel coders. Another motivation is to make use of the duality between source coding and channel coding, which is an under-utilized concept in obtaining source coding methods. By combining these ideas and using the flexibility of rate allocation offered by a puncturing mechanism, we design a new source-channel coding method achieving large performance gains at modest increase of complexity.

In Chapter 6, we consider the problem of transmission power allocation for video transmission over a wireless channel. Reducing the transmission power will reduce the interference between users sharing a wireless channel, and may extend battery life in a mobile device. Previous work in this area assumed source coding only without the use of channel coding, or assumed power allocation on the frame level instead of the packet level, or required future CSI for power allocation. Under the assumptions of using fixed-length packetization and adaptive forward error control (FEC), we propose simple power allocation algorithms to allocating transmission power to video packets, with the goal of minimizing the amount of distortion in the reconstructed video sequence given certain channel bandwidth and average transmission power constraints. Depending on whether

the CSI is known at the transmitter or not, two scenarios are considered. In addition to numerical algorithms, closed-form expressions are also presented to characterize the roles of the source error sensitivity and the CSI in optimally allocating transmission power. Simulations are conducted for a time-correlated Rayleigh fading channel, and results demonstrate the performance improvement of the proposed algorithms over the conventional algorithms.

1.3 Contributions of This Dissertation

The contributions of this dissertation are:

1. A new general performance measure for evaluating progressive image transmission systems operating over noisy channels, and a dynamic programming optimization algorithm to optimize both block length and rate of channel codes for progressive image transmission.

2. A new image coding scheme that combines multiple description quantization, entropy coding, and data partitioning to improve the error resiliency of images over varying packet loss channels, and a new method of entropy coding of multiple description images, using the SPIHT algorithm.

3. An asymptotic analysis of a cascade model of a multiple description coder and channel coders showing that only under certain conditions, such as time-varying channels with channel states unknown to the encoder, may the optimized cascade multiple description scheme outperform the optimized cascade single description scheme.

4. A new source coding scheme (Punctured Trellis Coded Quantization) that achieves fractional bit-rates at low complexity, and a new source-channel coding and mod-

ulation scheme (Punctured Trellis Coded Quantization/Modulation) that achieves flexible rate allocation and significant performance gains at modest increase of complexity.

5. A new power allocation scheme for wireless video transmission that incorporates fixed-length packetization and adaptive channel coding and allocates the transmission power to video packets, and new closed-form formulas characterizing the roles of the source error sensitivity and the CSI in optimal transmission power allocation.

Chapter 2

Channel Code Blocklength and Rate Optimization for Progressive Image Transmission

2.1 Introduction

Systems composed of embedded wavelet-based image coders followed by channel coding result in some of the best systems currently known for transmitting images over certain noisy channels [71, 69, 11, 54]. Selecting the proper channel code rate is important so that bits are not wasted on unnecessary redundancy, while keeping the error probability sufficiently low. Adjusting the blocklength can further improve the performance of certain channel codes.

It is known from information theory that long blocklengths can achieve good performance at rates close to capacity, and many known codes exhibit improved performance as the blocklength is increased. Examples include turbo codes, low-density parity check (LDPC) codes, and Reed-Solomon codes on erasure channels. The results

presented in [54] for image transmission on a packet erasure channel demonstrate that excellent performance at a target transmission rate can be achieved by using the maximum possible channel codeword blocklength. For an erasure rate ϵ and channel code rate $r \leq 1 - \epsilon$, a useful upper bound on the block error probability is [26, p.531]

$$P_E(r, \epsilon, N) \leq 2^{-ND_2(1-r||\epsilon)}$$

where N is the blocklength and $D_2(a||b) = a\log_2(a/b) + (1-a)\log_2((1-a)/(1-b))$ is the information divergence. This upper bound shows the advantage of long blocklengths (i.e., large N) for rates below capacity.

If decoding at multiple rates is not important, then using long blocklength channel codes provides the best performance in many cases. One penalty associated with long blocklength codes is increased decoding complexity. For progressive coding, there is the additional penalty of decoding delay. Since the decoder typically must wait until the entire block has been received and decoded before reliable source bits are available, the distortion vs. rate curve flattens out over the duration of a blocklength and decreases only at the end of the block. The result is that the curve is staircase shaped, where the size of the steps depends on the channel code blocklength as illustrated in Figure 2.1. The tradeoff is that long blocklengths improve error protection but they also decrease progressivity.

Previous work [11, 49, 3] has considered methods of selecting the best channel code rate for image transmission on noisy channels. A gradient-based technique was used in [3] to determine the best rate allocation between source and channel codes for a given transmission rate. In [49], Lagrangian methods were used to select the channel code rate schedule for fixed length information blocks to minimize distortion at a final target rate. Although progressive source coders were used, progressive performance was

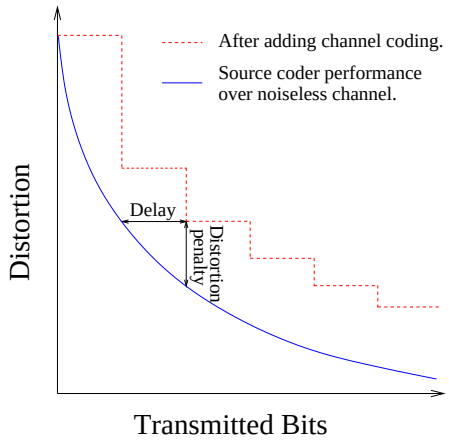
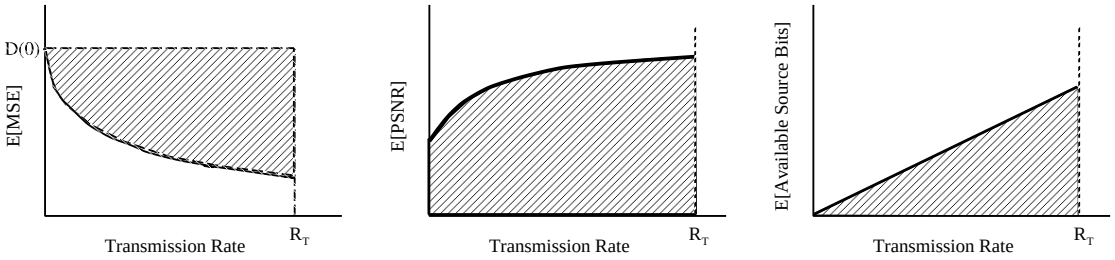


Figure 2.1: Distortion vs. Rate with and without Channel Coding.



(a) Performance measure based on MSE.

(b) Performance measure based on PSNR.

(c) Performance measure based source bits.

Figure 2.2: Three Possible Progressive Performance Measures Each Equal to the Area of the Shaded Region.

not considered in [49, 3].

In [11], optimization was also performed for a specific transmission rate, but the optimization criterion and rate-compatible properties of the channel codes allowed optimal transmission at many lower rates. A dynamic programming approach was presented for determining channel code rate schedules for fixed-length information blocks using optimization criteria based on MSE, PSNR, and number of available source bits. The optimization based on number of available source bits was suggested as the best approach since it reduces complexity, eliminates the need to transmit the rate schedule, and allows a single rate schedule to incorporate the optimized rate schedules for many

lower transmission rates as prefixes by using the rate-compatible properties of the codes considered.

We extend the work in [11] to consider the optimization of blocklength and channel code rate, and we use a more general performance function to characterize progressive performance. Specifically, this chapter includes the following contributions:

- A general performance measure is presented for evaluation of progressive performance;
- A dynamic programming solution is presented for optimizing blocklength and rate;
- Results for a number of error rates compare optimizations based on PSNR and optimizations based on available source bits with and without fixed information block sizes;
- Results for Reed-Solomon codes on erasure channels and for several codes including turbo codes and LDPC codes on bit error channels, demonstrate the importance of optimizing blocklengths.

2.2 A Performance Measure for Progressive Transmission

In order to optimize the channel code parameters, it is necessary to define a performance measure for progressive transmission. In image coding, typically the expected MSE or PSNR at a target transmission rate is used. For progressive coding, the desire is generally to maximize the expected performance at all rates up to the target transmission rate. We

propose a family of performance measures given by

$$\mu = \int_0^{R_T} W(x)F(x)dx \tag{2.1}$$

where R_T is the terminal transmission rate, $F(x)$ is the expected performance for a given channel at transmission rate x, and $W(x)$ is a non-negative, real-valued "weighting function" which allows unequal emphasis for different rates. The performance function $F(x)$ is large for good performances and small for poor performances, in contrast to the more typical use of a "distortion" function. The performance measure μ includes, as a special case, the conventional approach of measuring performance at a single rate (or possibly a small set of rates), by using impulse functions as part of the weighting function (e.g., $W(x) = \delta(x - r_1) + \delta(x - r_2) + \ldots$ where $\delta(x)$ is the Dirac delta function).

Possible performance measures include the expected MSE, expected PSNR, or number of available source bits as noted in [11]. The PSNR-based performance is closely tied to image quality but does not emphasize the large distortions at low rates as much as the MSE-based performance. Performance based on available source bits assumes all bits as equally important but it allows simplifications in the optimization and eliminates overhead information. In both of these cases, the goal (assuming $W(x) = 1$) is to maximize the area under $F(x)$. If we define $F(x) = D(0) - D(x)$, where $D(x)$ is the expected MSE at rate x, then the goal is also to maximize μ. In the context of dynamic programming algorithms discussed in the next section, μ is the reward. Figure 2.2 shows examples of these performance measures where the shaded region represents the area computed by the integral with $W(x) = 1$.

2.3 Optimization Using Dynamic Programming

The optimal schedule of channel code parameters can be determined using dynamic programming where the goal is to maximize the reward based on the progressive performance measure (i.e., μ in (2.1)). The notation is based on that in [8].

Given a transmission rate constraint R_T, the goal of the optimization problem is to determine the *code schedule*, $\pi = \{(N_1, K_1), (N_2, K_2), \ldots, (N_{M(\pi)}, K_{M(\pi)})\}$, subject to $\sum_{i=1}^{M(\pi)} N_i \leq R_T$ that maximizes μ. Each component of the code schedule contains a blocklength parameter, N_i, and an information size parameter, K_i. For a given system consisting of source coder, channel coder, and channel, each code schedule, π, determines an expected performance-rate function as in (2.1), denoted $F_\pi(x)$. Since the transmitted data consists of a sequence of codewords, the overall reward will be based on the incremental reward associated with each codeword. The incremental reward of a codeword with parameters (N_i, K_i) under code schedule π is given by

$$r_\pi(N_i, K_i) = \int_{\widetilde{R}_{i-1}}^{\widetilde{R}_i} W(x)(F_\pi(x) - F_\pi(\widetilde{R}_{i-1}))dx + (F_\pi(\widetilde{R}_i) - F_\pi(\widetilde{R}_{i-1})) \int_{\widetilde{R}_i}^{R_T} W(x)dx \tag{2.2}$$

where $\widetilde{R}_i = \sum_{j=1}^{i} N_j$ is the number of transmitted bits up to and including codeword i, and $\widetilde{R}_0 \equiv 0$. The first term of (2.2) is the incremental reward during transmission of the codeword while the second term is the accumulated incremental reward from the end of the codeword to the target rate. The integrals typically simplify to sums due to the discrete nature of the problem, and for certain performance measures such as the one based on available source bits, the integrand simplifies due to the constant reward for each bit.

The underlying assumption used in combining the incremental rewards from multiple codewords into the overall reward is the serial decoding requirement of the

source decoder. A source decoder is said to have a *serial decoding requirement* if a substantial portion of the bit stream must be decoded in a sequential and uninterrupted fashion for correct interpretation. Many embedded coders, especially those employing adaptive entropy coding, have this property. The effect is that decoding terminates at the first uncorrectable (and detected) error so no reward is accumulated from subsequent correctly received codewords. A "reward-to-go" function associated with codeword i is defined as

$$J_i(R_T, \pi) = r_\pi(N_i, K_i) + (1 - P_E(N_i, K_i))J_{i+1}(R_T, \pi) \tag{2.3}$$

where $P_E(N_i, K_i)$ is the probability of block decoding error for codeword i. The $(1 - P_E(N_i, K_i))$ term multiplying J_{i+1} is the result of the serial decoding requirement. The goal of the optimization problem is to find the optimal code schedule π^* that satisfies

$$\pi^* = \arg\max_{\pi} J_1(R_T, \pi) \tag{2.4}$$

with the condition that $J_i(R_T, \pi) = 0$ for some sufficiently large i under all admissible policies due to the transmission rate constraint R_T.

In the following sections, this general framework is applied to specific channel conditions and performance results are provided for the resulting rate schedules.

2.4 Progressive Transmission over Erasure Channels

This section considers the problem of transmitting an image over a packet erasure channel with erasure rate ϵ. The packets are assumed to be of fixed length, and this length is a parameter of the optimization. The information is protected using erasure correcting codes across the packets (similar to that found in [69, 54]), such as Reed-Solomon codes, which allow for error-free transmission of K information packets out of a block

of length N as long as any K packets are received.

The codes are systematic with the information packets transmitted first so these packets are available immediately to the source decoder until the first erasure. After the first erasure, further source decoding is delayed until K packets of the current codeword have been received at which point erasure correction is possible. Assuming the weighting function from (2.2) is unity (i.e., $W(x) = 1$) for simplicity, the incremental reward is derived below for this case.

For the first K_i packets of the codeword, the incremental reward is based on the expected number of packets before the first erasure. Let $f(x)$ be the performance of the source coder at transmission rate x on a noiseless channel with no channel coding (e.g., number of available source packets or bits (where $f(x) = x$), PSNR, etc). Note that the rate parameter is in units of packets rather than bits, for simplicity, and the conversion to bits involves a constant multiplicative factor due to the constant packet size. The expected performance over this range of transmitted packets under code schedule π is given by

$$F_\pi(M+\widetilde{R}_{i-1}) = \sum_{j=0}^{M-1} f(\widetilde{K}_i+j)(1-\epsilon)^j\epsilon + f(\widetilde{K}_i+M)(1-\epsilon)^M \qquad 1 \le M \le K_i \quad (2.5)$$

where $\widetilde{K}_i = \sum_{j=1}^{i} K_i$ is the number of cumulative information packets and ϵ is the packet erasure rate.

During the transmission of the $N_i - K_i$ parity packets, the performance only changes at the point where a total of K_i packets have been received for that codeword (at which point erasure correction is possible). The expected performance over this

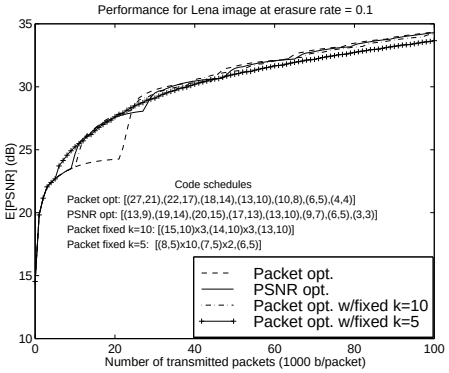
range is given by

$$F_\pi(M+\widetilde{R}_{i-1}) = \sum_{j=0}^{K_i-1} P_{me}(M-j-1, M-K_i) f(\widetilde{K}_i+j)(1-\epsilon)^j\epsilon + (1-P_E(M,K_i))f(\widetilde{K}_i+K_i) \quad K_i+1 \le M \le N_i \tag{2.6}$$

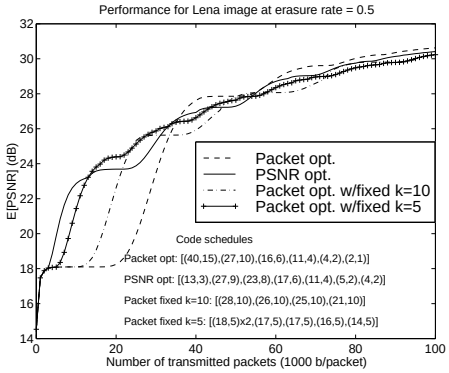
where $P_{me}(x,y) = \sum_{j=y}^{x} \binom{x}{j}(1-\epsilon)^{x-j}\epsilon^j$ is the multiple erasure probability (i.e., the probability of at least y erasures in x packets) and $P_E(M,K_i) = P_{me}(M, M-K_i+1)$ is the block decoding error probability of a code with parameters (M, K_i).

Optimization based on number of available source bits (packets) is attractive due to the reduced complexity and the elimination of the need to transmit the code schedule (i.e., the code schedule is image independent and can be computed by the receiver). However, performance measures based on MSE and PSNR are better indicators of image quality, so experiments were performed to compare the relative performance of the optimization methods. The overhead required for transmission of the code schedule in the case of the MSE or PSNR performance measures is assumed to be negligible so it is not included in the test results. This assumption is reasonable since the code schedule is typically fairly short (i.e., not many codewords) and can be compressed. Alternatively, the operational distortion-rate function is well modeled in practice by a function of the form $D(R) = \sum_{i=1}^{n} a_i e^{-\lambda_i R}$ with as few as two or three terms, as noted in [3]. This functional approximation could be used in determining the code schedule, and the function parameters could be transmitted instead of the actual schedule which would allow the receiver to compute the schedule.

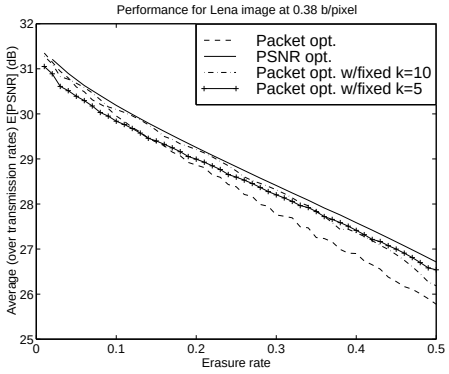
Figure 2.3 shows the results of experiments using the SPIHT [66] image coder with arithmetic coding along with code schedules computed from performance measures based on PSNR and available source bits or equivalently packets. The resulting code schedules are evaluated in terms of expected PSNR for two different erasure rates in



(a) E[PSNR] vs. rate for $\epsilon = 0.1$.



(b) E[PSNR] vs. rate for $\epsilon = 0.5$.



(c) Performance vs. erasure rate.

Figure 2.3: Performance Results for the 512x512 Lena Image Transmitted over an Erasure Channel.

Figures 2.3(a) and 2.3(b). In addition, a method similar to that in [11], referred to as packet optimization with "fixed k," is applied to Reed-Solomon codes. Results are provided for code schedules with $k = 5$ and $k = 10$ information packets per codeword.

In the "fixed k" method, the information size of each codeword is fixed and the blocklength (equivalently rate) is selected, the performance measure is available source bits (packets), and the optimization is for the final transmission rate so $W(x) = \delta(x - R_T)$ in (2.2). The resulting code schedules from the optimization are monotonic

non-increasing in blocklength (or non-decreasing in channel code rate) due to the serial decoding requirement. Initial information packets are protected more heavily than later packets, even though every packet contains one unit of reward, since a packet can only be used if all earlier information packets are correctly received. The order of transmission of the codeword packets can be altered so that optimal (in terms of available source bits) code schedules are achieved at certain lower transmission rates as suggested in [11]. This progressive transmission is achieved by transmitting punctured versions of the codewords starting with the last codeword parameters in the schedule. Additional parity packets are sent as needed during transmission to achieve all the intermediate code rates specified in the schedule between the last codeword and the codeword in question. Optimal schedules in terms of available source packets are achieved for transmission rates which equal the cumulative sum of the blocklengths starting from the last codeword and proceeding to the first codeword in the schedule. There are a few points to mention regarding the reordering for progressive transmission in [11]:

- The optimality is for the particular value of k and it corresponds to available source packets, not expected PSNR;
- The optimality is only guaranteed at certain transmission rates (equal to the cumulative sum of blocklengths starting from the last codeword), so other transmission orders may be better at other intermediate transmission rates;
- The best order for performance at all intermediate transmission rates is dependent on the erasure rate, and it is generally better to transmit according to the natural schedule order (i.e., all packets of each codeword in sequence) under high erasure conditions;
- The transmission rates of optimality are not selectable, but are instead determined by the code schedule computed for the final transmission rate.

The results in Figures 2.3(a) and 2.3(b) show that the optimization based on available source packets and that based on expected PSNR achieve nearly the same expected PSNR at the final transmission rate. The code schedules are also listed in the figures. Generally, the code schedules for the packet-based optimization consist of codewords with non-increasing blocklengths and nearly identical rates with the possible exception of the final codewords due to the transmission rate constraint. This is in contrast with the "fixed k" method, where the code rate increases monotonically. The PSNR-based optimization selects a shorter but lower-rate initial codeword to reduce the delay while still keeping the error probability low, and then the remainder of the schedule is similar to the packet-based schedule, though the code rate tends to be slowly increasing. As the first two figures in Figure 2.3 demonstrate, the relative performance of the different methods depends on the erasure rate. Figure 2.3(c) shows the performance in terms of the expected PSNR averaged over all the intermediate transmission rates (essentially a normalized version of the area under the curve) over a range of erasure rates. The PSNR-based optimization has the best performance at all rates as expected, while the other curves cross at different erasure rates. Also the spread in relative performance increases with erasure rate.

2.5 Progressive Transmission over Bit-Error Channels

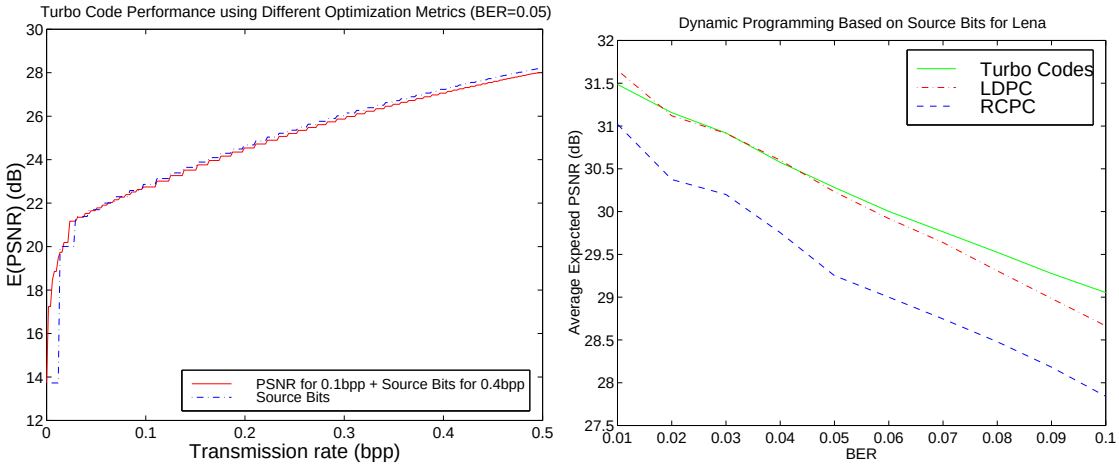
The same optimization techniques can be applied for turbo coding over bit error channels by simply changing the incremental reward function. For the erasure channel, information bits can be immediately used by the source decoder up to the first erasure since any data that arrives on this channel is known to be good. However, in this case, the information bits cannot be used before the entire codeword has been received since they are likely to contain errors which cannot be corrected until the remainder of the

codeword arrives. The result is that the incremental reward of a codeword only has the second term from (2.2), and the performance curves have more of a staircase shape. The incremental reward for a codeword with parameters (N_i, K_i) under code schedule π is given by

$$r_\pi(N_i, K_i) = (1 - P_E(N_i, K_i))(f(\widetilde{K}_i + K_i) - f(\widetilde{K}_i))(R_T - \widetilde{R}_i). \qquad (2.7)$$

It is important that decoding errors are detected, because the source decoder will lose synchronization and corrupt the image with high probability if its input contains errors. Therefore an outer cyclic redundancy code (CRC) is used for error detection if the error-correcting code itself does not have this capability. Three codes are considered, namely turbo codes, LDPC codes and rate-compatible punctured convolutional (RCPC) codes. Among them, both turbo codes and RCPC codes need to be concatenated with CRC while LDPC codes need not.

The turbo codes used in the experiments consist of a family of punctured codes produced from a rate $1/3$ mother code. Note that using punctured codes allows the possibility of optimizing only the rate (i.e., the "fixed-k" method mentioned in the previous section) and using the progressive transmission order, but recent results suggest that non-punctured codes perform significantly better for the same code parameters [6]. This is one of the motivations for optimizing both code rate and blocklength as the "fixed-k" method is not applicable to non-punctured codes. Also note that it is the block error probability that is important here and not the bit error probability, so choice of interleavers becomes important for large blocklengths. For this reason, s-random interleavers [17], which enforce a minimum spread between adjacent information bits after interleaving, were used because they result in better block error probabilities than random interleavers.

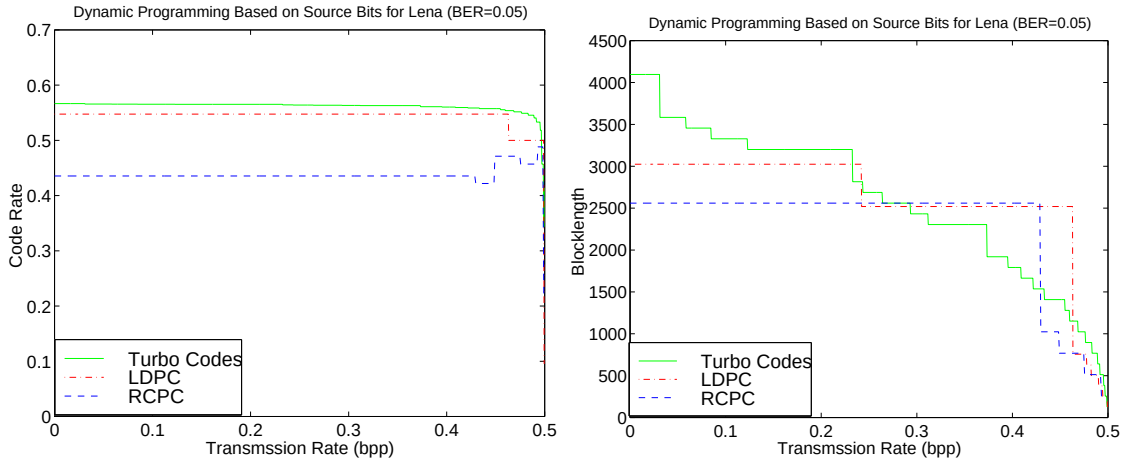


(a) Different Metrics (b) Different Channel Codes

Figure 2.4: Comparison of Different Metrics and Channel Codes

Before starting our experiments, we must decide which metric should be used. In Figure 2.4(a) a source-bits based metric and a hybrid metric are compared. The hybrid metric is used to reduce the complexity but achieve code schedules close to those of PSNR metric. From this graph, we see that at the early stage of transmission the source-bits-based optimization results in longer code blocklength thus longer delay than the PSNR-based optimization, but the average performance of the two metrics are very close. This contrasts with the results for packet erasure channels as shown in Figure 3, where with a packet size of 1000, a longer codeword would cause significantly longer delay. In the case of bit-error channels, because practical maximum code blocklength is on the order of a thousand bits, the longer code blocklengths selected by the source-bits-based optimization have little effect on the overall performance in terms of the average expected PSNR at practical rates for image transmission. Therefore for the complexity reason, the source-bits-based scheme is employed in the following experiments.

The performance results comparing Turbo Codes, LDPC and RCPC are shown in Figure 2.4(b). It's clear that Turbo Codes and LDPC outperform RCPC by about 0.5



(a) Code Rate (b) Blocklength

Figure 2.5: Comparison of Different Channel Codes

$\sim$ 1.0 dB.

Figure 2.5 shows the optimized code schedules for different channel codes. Even though the details of these curves depend on the available code parameters supplied to the optimization program, we can still find some structure in them: the code rate is essentially constant and the blocklength decreases monotonically. This structure is similar to those computed for the erasure channel using the packet-based metric. If the PSNR-based metric were used, the resulting code schedule would be very different for initial codewords: shorter but lower rate codewords would be selected, as in the case of erasure channels using PSNR-based metric. Note that for turbo codes and LDPC codes the block error probability has a strong dependence on blocklength, but for RCPC codes this dependence is much weaker. This partly explains why in Figure 2.5 (b) the blocklength varies more for turbo codes and LDPC codes than for RCPC codes.

2.6 Conclusion

We have proposed a general performance measure for evaluating progressive image transmission systems operating over noisy channels. Using this performance measure, a dynamic programming optimization algorithm was presented for determining the best code schedule where both block length and rate are optimized. Performance results were presented for erasure channels and for several codes over bit-error channels and demonstrated the importance of optimizing blocklengths. We found that the performance measure based on available source bits provides the best tradeoff in many cases. While the performance gains of the image dependent PSNR-based optimization may not warrant the additional complexity in every application, the method is useful for evaluating the relative performance of simpler coding schemes.

This chapter, in part, is a reprint of the material as it appears in: P.G. Sherwood, X. Tian, and K. Zeger. Channel code blocklength and rate optimization for progressive image transmission. *IEEE Wireless Communications and Networking Conference*, pp 978-982, Vol.2, Sept. 1999. The dissertation author was the second author of this paper.

Chapter 3

Efficient Multiple Description Coding of Images for Packet Loss Channels

3.1 Introduction

As multimedia communications are used for harsh environments such as low power wireless devices, error resilient source coding becomes more important. Wireless devices are especially challenging because they operate in difficult environments (e.g., multiply fading channels). The goal is to perform well under noisy conditions and degrade gracefully as channel conditions worsen.

Typically, multimedia applications operate at a level in the network protocol stack that is conceptually disjoint from the physical layer. Therefore the channel appears as a packet erasure channel to the application (i.e., packets arrive error-free or else are lost). This contrasts other scenarios, where the packet erasure model is handled by adding cyclic redundancy check (CRC) codes at the application layer. Therefore we focus on a packet erasure channel with possibly unknown or varying erasure rates. This channel is an n-channel generalization of the 2-channel model typically assumed in

multiple description (MD) coding (e.g. [77]). Due to the similarity, previous work on both multiple descriptions and coding for packet erasure channels is relevant.

Previous work on MD coding includes methods which allow a wide range of redundancy levels to be introduced across a small set of descriptions (usually two descriptions). The MD scalar quantizer in [77] uses an index assignment to map a regular quantizer index to a pair of indices with a selectable amount of redundant coding between them. Application of MD scalar quantizers to image coding was considered in [68], including discussion of extensions to more than two descriptions. Correlating transforms are used to introduce redundancy between pairs of values in [80, 30]. In [29] quantized overcomplete frame expansions are used to produce a small set of descriptions with some redundancy.

Data partitioning is another technique used to add error resiliency to a source coder (e.g. [65, 70, 40, 52]). High performance compression algorithms, such as the SPIHT image coder [66], often require the source coder bits to be decoded in a sequential and uninterrupted fashion for correct interpretation. Carefully partitioning the data into separate streams prior to coding can improve the error resilience. The idea is that the source decoder can use more of the correctly received information because independent pieces of the transmission can be correctly interpreted despite losing other portions. Data partitioning often also facilitates the application of unequal error protection channel coding. This method is an attractive technique for situations where the decoder complexity and graceful degradation under varying channel conditions are important. The cost is a loss of compression performance because the dependencies among the partition elements cannot be fully exploited.

If the number of packets is large, Forward Error Control (FEC) channel codes at the output of the source coder work well. When the codes' block lengths are long, a large number of rates are possible and thus the code rate can be accurately matched to

the channel. Also the FEC can be applied in an unequal fashion based on the importance of the information to provide improved efficiency and some graceful degradation (e.g. [54, 61]). The major costs of this method include decoding complexity, decoding delay (i.e., loss of progressivity), and sensitivity to channel mismatch.

The best choice among known methods depends on the number of packets (or descriptions) to be transmitted. Often the network imposes constraints by limiting the maximum packet size (i.e., the path maximum transmission unit (MTU)) or by distributing the information into a set of naturally separate descriptions as in a multichannel transmission system. In this chapter, we assume the number of packets is fairly small (i.e., fewer than 10), and investigate extensions of the two-channel multiple description mappings to create efficient and error resilient image coders.

3.2 Image Coding Structure

Many high performance image coding algorithms entropy code the scalar quantized coefficients of a linear transform (e.g. wavelet). We consider a similar structure where the scalar quantizer is replaced by an MDSQ, and each of the M descriptions is coded by an independent entropy coder. Each description is coded separately using the SPIHT algorithm with arithmetic coding (see Figure 3.1). This method (called MDSQ-SPIHT hereinafter) works well for MDSQ's with high redundancy, since each description image is roughly a coarsely quantized version of the original image.

As described in [77], an MDSQ is a cascade of a scalar quantizer followed by an index assignment which maps 1 index to M indices. We consider uniform quantization for the central quantizers with a fixed image-wide stepsize. If the amount of redundancy is fixed for all subbands, then the stepsizes for side quantizers are the same for all wavelet coefficients. The redundancy is tuned by selecting the number of central

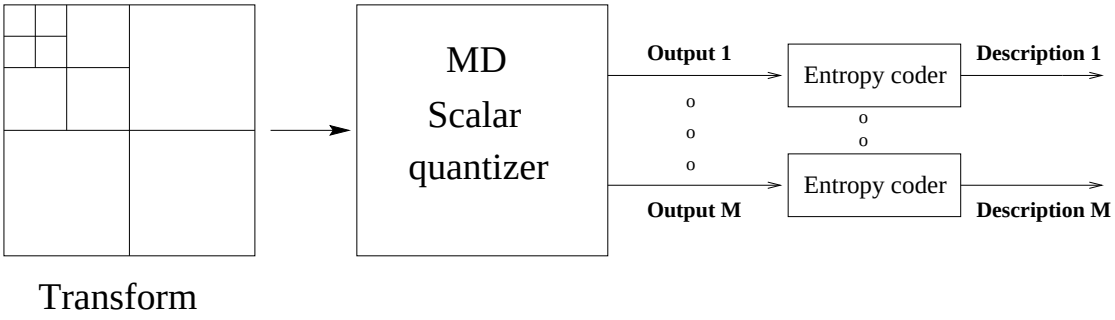


Figure 3.1: Block Diagram of an Image Coder for Packet Loss Channels.

quantizer cells relative to the number of bits in the description indices. An example of two index assignments is shown in Figure 3.2 for the two-description case.

1	2						
3	4	5					
	6	7	9				
		8	10	11			
			12	13	15		
				14	16	17	
					18	19	21
						20	22

(a) Index assignment with high redundancy

1	3	5	7				
2	8	10	12	14			
4	9	15	17	19	21		
6	11	16	22	23	25	27	
	13	18	24	29	30	32	34
		20	26	31	36	37	39
			28	33	38	41	43
				35	40	42	44

(b) Index assignment with less redundancy

Figure 3.2: Example MD Index Assignment Mappings for Two Descriptions.

Each of the numbered cells in the mappings represents a cell of the central quantizer, and each quantized coefficient is mapped into a pair of coordinates which specify its position in the grid. In these examples, the central quantizer indices are placed along diagonals of the grid, and the redundancy is controlled by selecting the number of diagonals. When only one of the two descriptions is received, all the central cells along

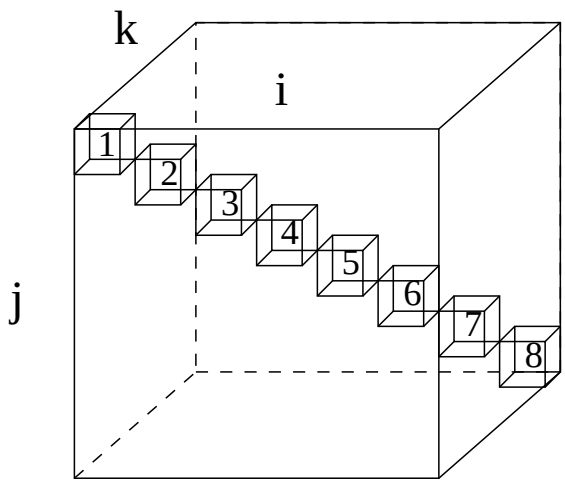


Figure 3.3: Example of a Three-Dimensional MD Mapping.

the given row or column are possible, so it is important to minimize the spread (i.e., the difference between the minimum and maximum indices in the row or column) for good error resilience.

The same concept can be extended to more than two descriptions by using higher dimensional mappings. Cells from the central quantizer are placed along diagonals of an M-dimensional hypercube. Figure 3.3 illustrates the concept with a three-dimensional mapping which includes only the main diagonal resulting in a (3,1,3) repetition code of the central quantizer indices. Adding other diagonals around the main diagonal reduces the redundancy and allows the channel coding rate to be adjusted. The algorithm used to generate the multi-dimensional mappings adds elements successively along each of the diagonals beginning at the upper corner and proceeding to the lower corner. At each step, the diagonals are sorted according to the minimum value along each dimension and the central quantizers cells are assigned in that order to reduce the spread.

The redundancy of the mapping is directly related to the the number of hypercube cells *not* assigned a central quantizer index. With Q central quantizer cells, S indices for each description, and M descriptions, the channel code rate of the mapping

is

$$r_c = \frac{\log_2(Q)}{M \log_2(S)}. \tag{3.1}$$

Thus it is difficult to create maps which provide high channel code rates with few central quantizer cells and many descriptions. These situations occur in the context of image coding when coding the coefficients in the higher frequency subbands. In such situations, data partitioning can be applied so that information about a specific high frequency coefficient is only sent in a subset of the packets. A similar type of data partitioning was used in the context of image coding for packet erasures in [40, 52].

The choice of mapping also has an impact on the individual entropy coders for each description. Low redundancy mappings result in high-entropy descriptions, reducing the effectiveness of the entropy coders. Also, the individual descriptions may not be matched very well in terms of rate and error resilience. It is important to match the descriptions so the loss of any description is approximately equivalent, thus improving error resilience.

Even though the MD mapping is designed to have approximately equal spread among all descriptions, it may not result in well-matched description rates since they will not be equally compressible. One method of balancing the descriptions is to use different permutations of the index M-tuple throughout the image. Therefore the MD mapping used for a particular coefficient is a composition of a multi-dimensional MD mapping and a coefficient-dependent permutation (e.g., the permutation may depend on the location or sign of the coefficient). For two descriptions, the sign of the coefficient can be used to select the permutation. For example, if the central quantizer index 1 is mapped into indices $(1, 0)$ for side quantizers, then the central quantizer index -1 is mapped into $(0, -1)$.

We propose two permutation schemes for an arbitrary number of descriptions

in order to balance the bit-rates among descriptions. The first one, a local permutation scheme applicable to any coders, rotates the M-tuple component order for a coefficient positioned at (x, y) according to the value of $(x + y)$ mod M. The second scheme is suitable for zero-tree type wavelet coders. It uses the first scheme only on the root coefficients (lowest subband), and all the descendants of a root coefficient will have the same index order as their ancestor. Both schemes can equalize the bit-rates among descriptions, but experiments show that the second scheme is slightly better than the first scheme when the SPIHT coder is used.

3.3 Experiment Results

The 9/7 Daubechies wavelet and a 5-level wavelet decomposition was used for the 512x512 Lena image, and a 4-level decomposition was used for the 256x256 Lena image. The total bit rate in the experiments ranges from 1.0 bpp to 1.25 bpp. An MDSQ is used to generate the M descriptions, and they are supplied to the SPIHT entropy coder. The stepsize for the central quantizer is determined by a bisection search algorithm to meet the total bit budget.

First we compare different index permutation schemes when SPIHT is used as the entropy coder for each description. For the 512x512 Lena image coded at a total rate of 1 bpp in 2 descriptions, the zero-tree based permutation scheme is about 0.2 dB better than the local permutation scheme in terms of reconstruction PSNR for both the central quantizer and the side quantizers, and is about 0.3 dB better than the sign-based permutation scheme. Therefore the zero-tree based permutation scheme is used for the remaining results in this section.

Figure 3.4 compares the proposed method and that of Miguel *et al.* [52] and Servetto *et al.* [68] for the case of 2 descriptions. The results are comparable to those of

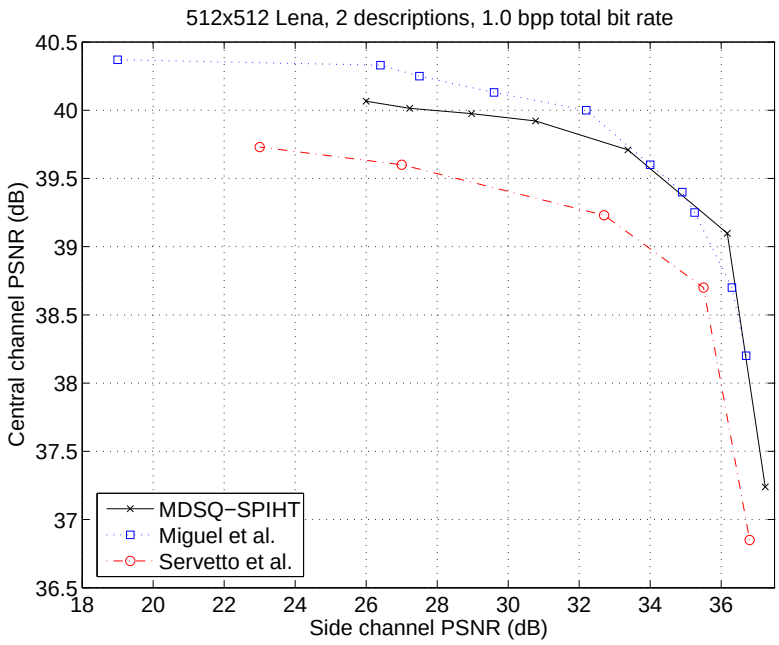


Figure 3.4: PSNR Results for MDSQ-SPIHT, the MD-SPIHT of Miguel *et al.* [52] and the MDSQ of Servetto *et al.* [68].

[52] in the high redundancy region (i.e., low central PSNR region) but are slightly worse in the low redundancy region where the data partitioning used in [52] is able to achieve very low redundancy. The improvement over [68] can be attributed to the improved entropy coder provided by the SPIHT algorithm as well as the better index permutation scheme based on zero-trees over that based on signs.

Figure 3.5 shows some examples of decoded images for a 4-description coder based on MDSQ-SPIHT with 3 diagonals at a total of 1.22 bpp with various numbers of received descriptions. Figure 3.6 shows the average PSNR performance results of the same 4-description coder for a number of different redundancy levels. As more diagonals are added to the mapping (reducing the redundancy), the performance with no description loss improves but the performance with heavy description loss degrades. Compared to the method of Mohr *et al.* based on unequal loss protection (ULP) [53], the proposed method with 5 diagonals outperforms ULP by $1.6 \sim 3.0$ dB depending the

number of descriptions received.

In Figure 3.7, the number of descriptions is 8, and we compare MDSQ-SPIHT with the scheme of Miguel *et al.* [52], and with the ULP scheme of Mohr *et al.* [54]. Using 5 diagonals, our coder performs close to their schemes when there is no description loss, but is about 3 to 5 dB worse than that of [54] when there is a loss of 1 to 2 descriptions; on the other hand, MDSQ-SPIHT performs better than [54] when at least 4 descriptions are lost, and the gain is as high as 10 dB when only 1 description is received.

In order to improve the performance at low loss rates, we apply the data partitioning technique to our coder. In this case, while the lower frequency bands are coded with 8 descriptions, high frequency bands are only coded with 4 descriptions, which are alternated into one of the two groups – each having 4 descriptions – in a predetermined order. This reduces the total bit rate given the stepsize, thus achieving finer quantization given the total bit budget. In Figure 3.7, the PSNR for the cases with no description loss or just 1 description loss improves by about 1 dB, but the PSNR with only 1 description received drops by about 4 dB.

3.4 Conclusion

In this chapter, a novel scheme for robust image coding over wireless packet networks is proposed. We investigate the use of a combination of techniques including MD scalar quantization, efficient entropy coding of MDSQ output, and data partitioning to improve the error resiliency of images when transmitted over packet erasure channels. A multi-dimensional extension of the two-channel multiple description scalar quantizer is proposed. A novel wavelet image coder based on a SPIHT lossless encoding and index alternation scheme is proposed for MDSQ, and shows better performance than a wavelet based image coder previously designed for two-description scalar quantizers.

Performance results obtained for 4 descriptions shows that for a small number of descriptions, the proposed method outperforms the state-of-the-art ULP method [54] by a significant margin. It is also demonstrated that with the number of diagonals in the index mapping as a parameter, the tradeoff between the performance of high loss rate and that of low loss rate can be achieved. The coder shows graceful degradation as descriptions are lost, and is very robust under high loss rate.

This chapter, in part, is a reprint of the material as it appears in: P.G. Sherwood, X. Tian, and K. Zeger. Efficient image and channel coding for wireless packet networks. *International Conference on Image Processing*, pp 132-135, Vol. 2, Sept. 2000. The dissertation author was one of the primary investigators and the second author of this paper.

(a) 1 description received

(b) 2 descriptions received

(c) 3 descriptions received

(d) 4 descriptions received

Figure 3.5: Decoded Images for 256x256 Lena and MDSQ-SPIHT

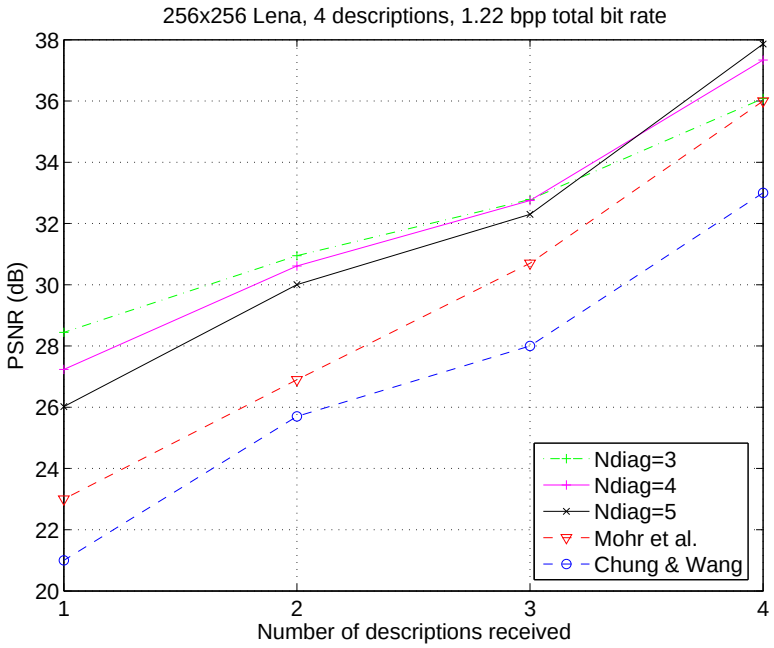


Figure 3.6: PSNR Results for MDSQ-SPIHT, the Unequal Loss Protection scheme of Mohr *et al.* [53] and the Lapped Orthogonal Transform Based Scheme of Chung and Wang [13].

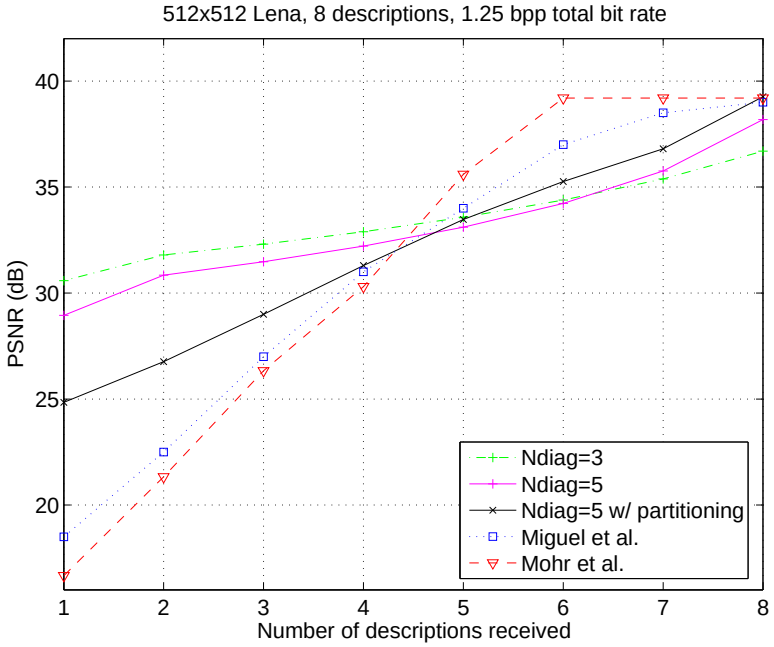


Figure 3.7: PSNRs for MDSQ-SPIHT, MD-SPIHT (Miguel *et al.* [52]) and Unequal Loss Protection (Mohr *et al.* [54]).

Chapter 4

Asymptotic Analysis of Multiple Description Codes Over Noisy Channels

4.1 Introduction

Multiple Description Coding (MDC) is a source-channel coding technique used to generate multiple descriptions (mostly two descriptions) of a source. Both descriptions are transmitted over independent channels to the receiver. The goal of MDC is to achieve some acceptable quality if only one description is received, and an improved quality if both descriptions are received.

Extensive work on MDC has been done on finding the achievable rate-distortion bounds for MDC, and on developing practical coders to approach those bounds. For example, Ozarow constructed the rate-distortion region for the case of a memoryless Gaussian source and a squared-error distortion criterion [58]. Vaishampayan developed the Multiple Description Scalar Quantizer (MDSQ) [75], using an index assignment to

map a regular quantizer index to a pair of indices with selectable amount of redundant coding between them. Its asymptotic performance was later studied in [76].

So far most work on MDC assumes ideal multiple description (MD) channels, i.e, descriptions arrive error-free or else are completely lost. Examples of such channels are found in lossy packet networks such as Internet, where congestion and delays lead to packet loss, and significant performance improvement can be achieved by using MDC [2]. However, in wireless communication systems, due to channel fading, interference and noise, bits within each description are subject to errors, so additional channel coding beyond MDC is necessary to maintain a low error probability. In [72], MDC was considered over noisy channels. An MDSQ encoder was used to generate two descriptions, each of which was then coded using a rate $1/2$ convolutional code for an AWGN channel. With an iterative decoding algorithm to decode multiple descriptions, it was shown that there is an optimal MD redundancy for a given channel condition. In [64], MDC was applied to noisy channels including Binary Symmetric Channels (BSCs). It was assumed that the MDC bounds for Gaussian sources are achieved by the source coder, and that some Rate-Compatible Punctured Convolutional (RCPC) codes and Cyclic-Redundancy-Check (CRC) codes are used for channel coding. It was shown through numerical optimization that compared with a two-layer source coder with unequal error protection, the MD coding scheme is more effective only at very high channel error probabilities. In [15], an MDC system with MDSQ followed by error correcting block codes was considered, and the performance of the MD system and that of its corresponding single description (SD) system were compared over BSCs. It was shown through numerical optimization that in most channel conditions an optimized SD system performs better than an optimized MD system.

In this paper, we consider a problem similar to that of [15], but through an analytical approach. In particular, we analyze the asymptotic performance of a cascade model

of a MD coder and channel coders over BSCs. Under this model, diversity is achieved through the use of two independent channels to transmit source information. These two channels can be in the form of frequency, time or code division multi-access. Error detection codes such as CRCs are used to allow descriptions to be discarded if they contain any residual errors after channel decoding. Given a per-channel transmission rate R, in order to minimize the average distortion, there is a tradeoff among source coding resolution, MD redundancy and channel coding redundancy. We follow the approach used in [38], where both exponential formula given by high-resolution quantization theory and exponential bounds to channel coding error probability were used to derive tight bounds on the tradeoff between source and channel coding. Assuming large R, we derive the MD redundancy parameter and the channel code rate that minimize a distortion upper bound. Then we compare the performance of the MD system and its corresponding SD system for both time-invariant and time-varying channels.

4.2 MD Rate-Distortion Bound

Before applying MDC to noisy channels, we first consider an MD quantizer for a noiseless channel [58] [76]. Assume that two descriptions are generated for the source. When both descriptions are received, the quantizer is called *central quantizer*, and the distortion is called *central distortion* (denoted d_0); when only one description is received, the quantizer is called *side quantizer*, and the distortion is called *side distortion* (denoted d_1 and d_2 for the two descriptions, respectively). In [58], it was shown that for a memoryless unit-variance Gaussian source, given a rate per source component R_1 and R_2, the

set of squared distortion (d_0, d_1, d_2) can be achieved if and only if

$$\begin{aligned} d_1 &\geq 2^{-2R_1} \\ d_2 &\geq 2^{-2R_2} \\ d_0 &\geq 2^{-2(R_1+R_2)} \frac{1}{1-(\sqrt{\Pi}-\sqrt{\Delta})^2} \end{aligned}$$

where $\Pi = (1-d_1)(1-d_2)$ and $\Delta = d_1 d_2 - 2^{-2(R_1+R_2)}$.

For simplicity, we consider the balanced case and large rates, i.e., $R_1 = R_2 = \tilde{R} \gg 1$ and $d_1 = d_2$. Furthermore, assume that $d_1 = \nu 2^{-2\tilde{R}}$ where $\nu \triangleq b2^{2\tilde{R}a}$, $0 \leq a \leq 1$ and $b > 0$. It will become clear that a is a parameter to balance the tradeoff between the side distortion and central distortion.

In [76], the asymptotic rate-distortion bounds for $a = 0$ and $0 < a < 1$ were obtained. By including the case of $a = 1$, one can show that

$$d_0 d_1 = 2^{-2(2\tilde{R})+O(1)+o(\tilde{R})}.$$

Similar results were also derived in [76] for MDSQ. The performance of the optimum level-constrained MDSQ is given by

$$d_0 d_1 = \frac{3\pi^2}{16} 2^{-2(2\tilde{R})+o(\tilde{R})}$$

and of the optimum entropy-constrained MDSQ by

$$d_0 d_1 = \frac{\pi^2 e^2}{144} 2^{-2(2\tilde{R})+o(\tilde{R})}.$$

For a good MD coder with blocklength k, its performance can not be worse than that of MDSQ, which has blocklength 1, but can not be better than the rate-distortion

bound, which can be achieved with large blocklength. Therefore we conclude that for memoryless Gaussian source $X \in \mathbf{R}^k$,

$$d_0 d_1 = 2^{-2(2\tilde{R})+O(1)+o(\tilde{R})}$$

where

$$d_1 = 2^{-2\tilde{R}(1-a)+O(1)+o(\tilde{R})} \tag{4.1}$$

$$d_0 = 2^{-2\tilde{R}(1+a)+O(1)+o(\tilde{R})}. \tag{4.2}$$

4.3 System Model

We consider a memoryless Gaussian source $X \in \mathbf{R}^k$. An MD quantizer is used to encode the source and generate two bitstreams (each at m bits/source vector). Then each of the two bitstreams is supplied to a channel encoder (with CRC for error detection) of blocklength n bits, and then passed to a BSC. The two BSCs are assumed to have the same crossover probability ϵ, unless specified otherwise.

From the block diagram shown in Figure 4.1, one can see that this model is similar to that of [64] and [15]. By using error detection, the channels here are equivalent to packet erasure channels, where the packet erasures are not pure erasures but instead are caused by the inability of FEC to correct bit errors in a block. It is assumed in this model that the channel errors in these two channels are independent.

For each channel, at a large channel transmission rate per source component $R = n/k$, the effect of error detection codes on the channel code rate $r = m/n$ is negligible. We address the question: For a given R, what MD source redundancy parameter a and channel code rate r minimize an upper bound of the average mean-squared distortion from sender to receiver?

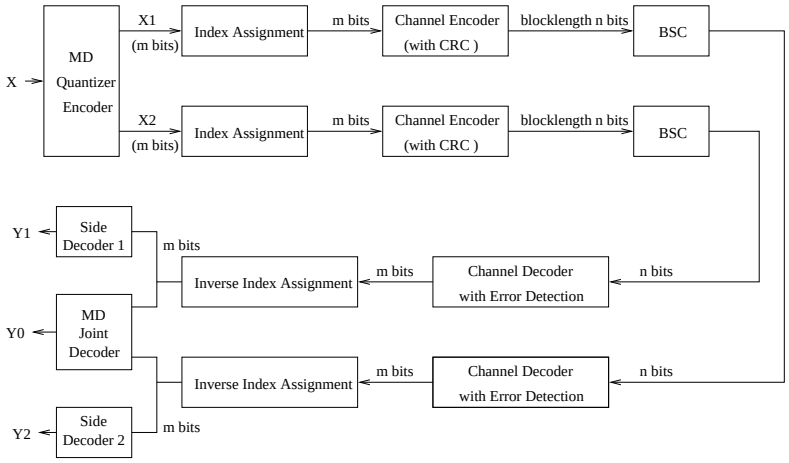


Figure 4.1: Cascaded MD Quantizer and Channel Coder System

For MD quantizer Q, we consider the balanced case as described in Section 4.2, with the source coding rate per source component $\tilde{R} = Rr$. We also assume that for each channel coder, the quantizer codevectors $y_1, \cdots, y_M$ are mapped to $M = 2^m = 2^{kRr}$ $n-$bit channel codewords through an index assignment π.

Now we are ready to analyze the above cascaded system, following the approach used in [38]. As in [38], we consider only r below the channel capacity $C = 1+\epsilon \log \epsilon + (1-\epsilon)\log(1-\epsilon)$. Assuming $\epsilon < \frac{1}{2}$, we have $C > 0$.

4.4 Analysis of Tradeoff Between MD Coding and Channel Coding

In this section, we derive the MD redundancy parameter a and channel code rate r that minimize a distortion upper bound.

Let P_e be the probability of incorrect decoding of either channel decoder for the

index assignment π. Then the average distortion $D_R(Q, \epsilon, \pi)$ satisfies

$$\begin{aligned} D_R(Q, \epsilon, \pi) &= (1 - P_e)^2 d_0 + 2P_e(1 - P_e)d_1 + P_e^2 O(1) \\ &\leq d_0 + 2P_e d_1 + P_e^2 O(1). \end{aligned} \tag{4.3}$$

These 3 terms correspond to the cases where there is no channel decoder error, only one channel decoder is in error, and both channel decoders are in error, respectively.

Shannon's channel coding theorem guarantees that, for channel code rates r below capacity, channel codes exist for which

$$P_e \leq 2^{-kRE_{ex}(r)+o(R)} \text{ as } R \to \infty \tag{4.4}$$

where $E_{ex}(r)$ is the "expurgated error exponent" for the BSC, given by

$$E_{ex}(r) = \sup_{\rho_1 \geq 1} g(\rho_1, r) \tag{4.5}$$

with

$$g(\rho_1, r) \triangleq \rho_1[1 - r - \log(1 + \delta^{1/\rho_1})] \tag{4.6}$$

and $\delta = 2\sqrt{\epsilon(1 - \epsilon)}$ [25] [38]. Here the logarithm is base two, which is also true for the rest of the paper.

Substituting (4.1), (4.2) and (4.4) into (4.3) yields

$$\begin{aligned} D_R(Q, \epsilon, \pi) \leq\ & 2^{-2Rr(1+a)+O(1)+o(R)} + 2 \cdot 2^{-2Rr(1-a)+O(1)+o(R)} \cdot 2^{-kRE_{ex}(r)+o(R)} \\ & + O(1) \cdot 2^{-2kRE_{ex}(r)+o(R)}. \end{aligned} \tag{4.7}$$

Let $D_{ex}(a, r)$ denote the right-hand side of (4.7), which is a distortion upper

bound to be minimized. Then we have the following lemma.

Lemma 4.1 *As $R \to \infty$, the pair (a, r) that minimizes $D_{ex}(a, r)$ obeys the following:*

$$(a, r) = \underset{a \in [\frac{1}{3}, 1], r \in (0,1]}{argsup} (1 + a)r \tag{4.8}$$

$$\textit{subject to} \qquad E_{ex}(r) = \frac{4a}{k} r. \tag{4.9}$$

Proof:: See Appendix A.

Define

$$B_{ex}(r) = \frac{kE_{ex}(r)}{4r}. \tag{4.10}$$

Then for (a, r) that satisfies (4.9), we have that given r, a can be determined by

$$a = B_{ex}(r). \tag{4.11}$$

Conversely, given a, r can be determined by

$$r = B_{ex}^{-1}(a) \tag{4.12}$$

where $B_{ex}^{-1}(\cdot)$ is the inverse function of $B_{ex}(\cdot)$.

Define

$$f_{ex}(a) = (1 + a) \cdot B_{ex}^{-1}(a). \tag{4.13}$$

Then it can be shown that as $R \to \infty$, $\min\limits_{r} D_{ex}(a, r) = 2^{-2R \cdot f_{ex}(a) + O(1)}$. The goal is to find $a \in [\frac{1}{3}, 1]$ so that $f_{ex}(a)$ is maximized.

Considering (4.5), we define

$$\rho(r) = \underset{\rho_1 \geq 1}{argsup}\, g(\rho_1, r). \tag{4.14}$$

For a given r, $\rho(r)$ is either greater than or equal to 1. We examine these two cases separately.

- $\rho(r) > 1$

  In this case, $\rho(r)$ must be a stationary point for $g(\rho, r)$. Letting $\frac{\partial g(\rho,r)}{\partial \rho} = 0$ yields

$$r(\rho) = 1 - \log(1 + \delta^{1/\rho}) - \frac{\delta^{1/\rho} \log(1/\delta)}{\rho(1 + \delta^{1/\rho})}. \tag{4.15}$$

  Substituting (4.15) into (4.5), we have

$$E_{ex}\left(r(\rho)\right) = \frac{\delta^{1/\rho}}{1 + \delta^{1/\rho}} \log(1/\delta). \tag{4.16}$$

  Lemma 4.2 *For (a, r) that satisfies (4.9) and $\rho(r) > 1$,*

  1. *a is an increasing function of ρ.*
  2. *The stationary point ρ^*_{ex} for $f_{ex}(\cdot)$ is given by*

$$\rho^*_{ex} = \frac{4}{k}. \tag{4.17}$$

  3. *The stationary point a^*_{ex} for $f_{ex}(\cdot)$ satisfies*

$$\begin{aligned} f'_{ex}(a) &< 0, \quad \text{if } a < a^*_{ex}; \\ f'_{ex}(a) &> 0, \quad \text{if } a > a^*_{ex}. \end{aligned} \tag{4.18}$$

  Proof:: See Appendix B.

  Combining (4.17) with (4.11), (4.10), (4.16) and (4.15), we can determine the

corresponding stationary point a^*_{ex} for $f_{ex}(\cdot)$ as follows:

$$a^*_{ex} = \frac{\delta^{k/4}\log(1/\delta)}{\frac{4}{k}(1+\delta^{k/4})[1-\log(1+\delta^{k/4})] - \delta^{k/4}\log(1/\delta)}. \tag{4.19}$$

- $\rho(r) = 1$

  In this case, (4.5) becomes

$$E_{ex}(r) = 1 - r - \log(1+\delta). \tag{4.20}$$

  Substituting (4.9) into (4.20) yields

$$r = \frac{1-\log(1+\delta)}{1+4a/k}. \tag{4.21}$$

  Replacing (4.12) and (4.21) into (4.13), and calculating the derivative of $f_{ex}(a)$, we have

$$f'_{ex}(a) = \frac{1-\log(1+\delta)}{(1+4a/k)^2}(1-\frac{4}{k}). \tag{4.22}$$

Having examined the two cases for $\rho(r)$, now we are ready to derive the parameter a that minimizes D_{ex}.

Theorem 4.3 *As $R \to \infty$, the parameter a that minimizes the distortion upper bound*

D_{ex} *is*

$$a_{ex} = \begin{cases} \frac{1}{3}, & \text{if } k < 4,\ a^*_{ex} \geq 1 \\ 1, & \text{if } k < 4,\ a^*_{ex} \leq \frac{1}{3} \\ \frac{1}{3}, & \text{if } k < 4,\ \frac{1}{3} < a^*_{ex} < 1 \text{ and } f_{ex}\left(\frac{1}{3}\right) > f_{ex}(1) \\ 1, & \text{if } k < 4,\ \frac{1}{3} < a^*_{ex} < 1 \text{ and } f_{ex}\left(\frac{1}{3}\right) \leq f_{ex}(1) \\ 1, & \text{if } k \geq 4. \end{cases} \tag{4.23}$$

Proof:: See Appendix C.

Intuitively, when the source dimension k is large, rate being spent on channel coding would be more efficient than being spent on MD redundancy, because given the total transmission rate, large block length would make the block decoding error probability P_e so small that large side distortion d_1 becomes less significant.

4.5 Performance Comparison Between MD and SD Based Schemes

In this section, we compare the performance of the MD-based scheme and SD-based scheme over BSCs. As shown in Figure 4.2, a channel error correcting code with an error detecting code is used in the SD scheme, as in the MD scheme. For fair comparison, we assume that the rate constraint for the SD scheme is the same as the total rate constraint for the MD scheme. Furthermore, we assume that the SD coder and the MD coder have the same source dimension k, unless noted otherwise. In other words, the SD rate constraint is $2R$ bits/source sample, and the channel code blocklength $n_{SD} = 2kR = 2n_{MD}$.

Consider a distortion upper bound based on the "expurgated error exponent" for

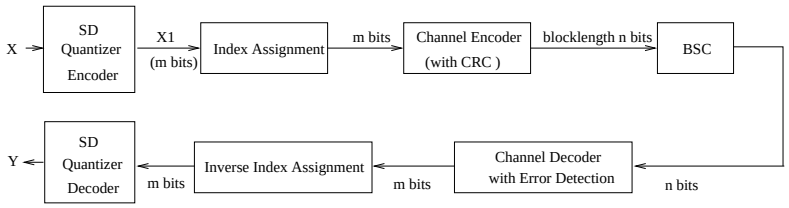


Figure 4.2: Cascaded SD Quantizer and Channel Coder System

the SD scheme [38]. As $R \to \infty$,

$$D_{SD}(r_{SD}) \leq 2^{-4R \cdot r_{SD} + O(1)} + 2^{-2kR \cdot E_{ex}(r_{SD}) + o(R)} \tag{4.24}$$

4.5.1 Time-Invariant Channels

Assume that each BSC in the MD and SD scheme is time-invariant, and has the same crossover probability ϵ. Also assume that this information is known to the encoder. Then we can minimize the distortion upper bounds of both schemes for the given ϵ.

For the SD scheme, in order to minimize the distortion upper bound of (4.24), r_{SD} is determined by solving $E_{ex}(r_{SD}) = \frac{2}{k} r_{SD}$, yielding

$$D_{SD} \leq 2^{-2R \cdot f_{SD} + O(1)}, \text{ as } R \to \infty \tag{4.25}$$

with the optimized distortion exponent

$$f_{SD} = 2 r_{SD}. \tag{4.26}$$

For the MD scheme, as shown in Section 4.4, for large R, the distortion upper bound D_{ex} can be written as

$$D_{ex} = 2^{-2R \cdot f_{MD} + O(1)} \tag{4.27}$$

where $f_{MD} = (1 + a_{ex}) \cdot r_{MD}$, with r_{MD} satisfying (4.9). Since from Theorem 4.3, the value of a_{ex} that minimizes D_{ex} is either 1 or $\frac{1}{3}$, we examine these two cases separately.

- $a_{ex} = 1$

  In this case, from (4.9) we have $E_{ex}(r_{MD}) = \frac{4}{k} r_{MD}$. Since $E_{ex}(r)$ is non-negative and a decreasing function of r (as illustrated in Figure 4.7 in Appendix A), it can be shown that $r_{SD} > r_{MD}$, yielding $f_{SD} > f_{MD}$.

- $a_{ex} = \frac{1}{3}$

  In this case, from (4.9) we have $E_{ex}(r_{MD}) = \frac{4/3}{k} r_{MD}$. Similarly as in the previous case, it can be shown that $E_{ex}(r_{SD}) > E_{ex}(r_{MD})$, yielding $2r_{SD} > (4/3)r_{MD}$, thus $f_{SD} > f_{MD}$.

These results indicate that under our assumptions, the optimized SD scheme may outperform the optimized MD scheme in terms of the average distortion. One assumption used is that when comparing the SD scheme with the MD scheme, the channel blocklength for SD is twice that for MD, so that the source blocklength (thus source delay) is the same for SD and MD coders. Since channel codes with longer blocklength tend to have smaller decoding error probability, this assumption enables the SD scheme to perform better than the MD scheme. However, if the channel blocklength must be kept the same due to constraints on complexity, then as will be illustrated in the next section, given the same total rate, the MD scheme may perform better than the SD scheme when k is small and ϵ is large.

4.5.2 Time-Varying Channels

Previously we assumed that the channels in the MD and SD schemes are time-invariant. Now we consider the case of time-varying channels. Assume that there exist a good

state (with the crossover probability ϵ^g) and a bad state (with the crossover probability ϵ^b, and $\epsilon^g < \epsilon^b$) for each channel, but the encoder has no knowledge of the instantaneous channel state information except that of its statistics. As a result, balanced rates for both descriptions ($R = R_1 = R_2$, and $r = r_1 = r_2$) should still be used for the MD scheme. Since in practice the probability of a channel being in the bad state is usually much smaller than that in the good state, the situation that both channels are in the bad state can be omitted. Therefore we consider a situation when one channel is in the good state and another is in the bad state, and minimize the distortion upper bound for the MD scheme.

Similarly as in (4.7), we can obtain the distortion upper bound as

$$\begin{aligned} D_{MD}^{gb}(a,r) \quad \leq \quad & 2^{-2Rr(1+a)+O(1)+o(R)} + 2^{-2Rr(1-a)+O(1)+o(R)} \cdot 2^{-kRE_{ex}^{b}(r)+o(R)} \\ & + O(1) \cdot 2^{-kRE_{ex}^{g}(r)-kRE_{ex}^{b}(r)+o(R)} \end{aligned}$$

where $E_{ex}^{b}(r)$ and $E_{ex}^{g}(r)$ are $E_{ex}(r)$ defined in (4.5) with $\epsilon = \epsilon^b$ and $\epsilon = \epsilon^g$, respectively.

Define

$$F_{ex}^{gb}(a,r) = \min\{r(1+a), [2r(1-a) + kE_{ex}^{b}(r)]/2, [kE_{ex}^{g}(r) + kE_{ex}^{b}(r)]/2\}. \quad (4.28)$$

Then it can be shown that as $R \to \infty$,

$$D_{MD}^{gb}(a,r) \leq 2^{-2R \cdot F_{ex}^{gb}(a,r)+O(1)}.$$

Define $B_{ex}^{b}(r) = \frac{kE_{ex}^{b}(r)}{4r}$, and its inverse function $r^b(a) = (B_{ex}^{b})^{-1}(a)$.

Assume that the 3 arguments of the $\min(\cdot)$ function in (4.28) are equal to each

other at $a = a_{123}^{gb}$, which can be determined by solving

$$E_{ex}^{g}(r^b(a_{123}^{gb})) = \frac{2r^b(a_{123}^{gb})}{k}(1 - a_{123}^{gb}). \tag{4.29}$$

Define $f_{MD}^{gb} = \sup_{(a,r)} F_{ex}^{gb}(a, r)$, and $f(a, r) = r(1 + a)$, then we have the following lemma:

Lemma 4.4 *The distortion exponent f_{MD}^{gb} can be determined by*

$$\begin{aligned} f_{MD}^{gb} &= \sup_{a \in [a_{123}^{gb}, 1]} f(a, r^b(a)) \\ &= \max\{f(a_{123}^{gb}, r^b(a_{123}^{gb})), f(1, r^b(1))\}. \end{aligned}$$

This lemma can be proved using a similar approach to that of Section 4.4.

Now we consider the SD scheme, and assume that the channel is also time-varying and subject to the good state and the bad state. Since the SD scheme optimized for the good state would perform poorly when the channel is in the bad state, we optimize the SD scheme for the bad state instead, whose performance in the good state cannot be worse. In other words, we minimize the SD distortion upper bound for the channel in the bad state, and determine f_{SD}^{b} similarly as in (4.26) with $\epsilon = \epsilon^b$. If the resultant f_{SD}^{b} is greater than f_{MD}^{gb}, then the SD scheme performs better in both good and bad states than the MD scheme. Otherwise, the MD scheme may outperform the SD scheme.

4.6 Results

In this section, we evaluate the analytic results in Section 4.4 and Section 4.5 with plots.

In Figure 4.3, function $f_{ex}(a)$ is obtained through numerical evaluations. For

each a considered, we evaluate r_{ex} as determined by (4.9) [1] , and evaluate $f_{ex}(a)$ according to $f_{ex}(a) = (1+a) \cdot r_{ex}$. Then $f_{ex}(a)$ is plotted for three cases: $k = 1$, $\epsilon = 0.1$; $k = 10$, $\epsilon = 0.1$; $k = 2$, $\epsilon = 10^{-5}$. Remember that the larger $f_{ex}(a)$, the smaller the distortion upper bound D_{ex}. One can see that in the first case, a_{ex}, i.e. the value of parameter a that minimizes D_{ex}, is clearly $1/3$, and in the second case the a_{ex} is clearly 1, while in the third case $a = 1$ is slightly better than $a = 1/3$.

In Figure 4.4, the parameter a_{ex} that minimizes the distortion upper bound D_{ex} as determined by Theorem 4.3 is plotted as a function of k and ϵ. Numerical evaluations of the best a (like those in Figure 4.3) that minimizes D_{ex} would give the exactly same results.

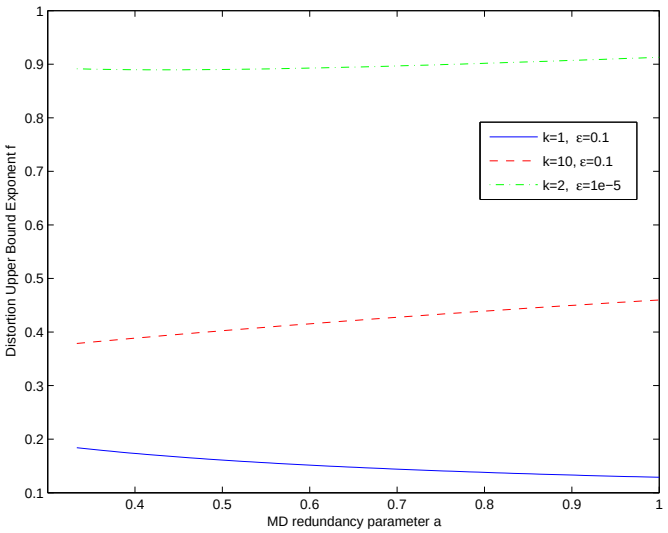


Figure 4.3: The Distortion Exponent f_{ex} as a Function of MD Redundancy Parameter a

In Figure 4.5, the distortion exponents f_{MD} and f_{SD} corresponding to the minimized distortion upper bounds for the MD and SD schemes are compared for $\epsilon = 0.1$

[1] The steps to numerically evaluate r_{ex} are: (i) Given ϵ and k, obtain function $r(\rho)$ from (4.15); (ii) Numerically solve for ρ that satisfies (4.9) where $E_{ex}(r)$ is given by (4.6); (iii) If the resultant ρ is greater than or equal to 1, determine r_{ex} from (4.15), otherwise substitute $\rho = 1$ into (4.6) and solve (4.6) for r.

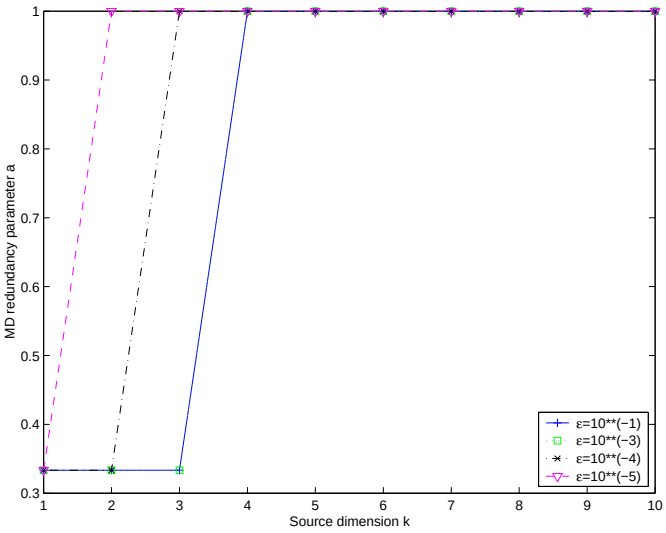
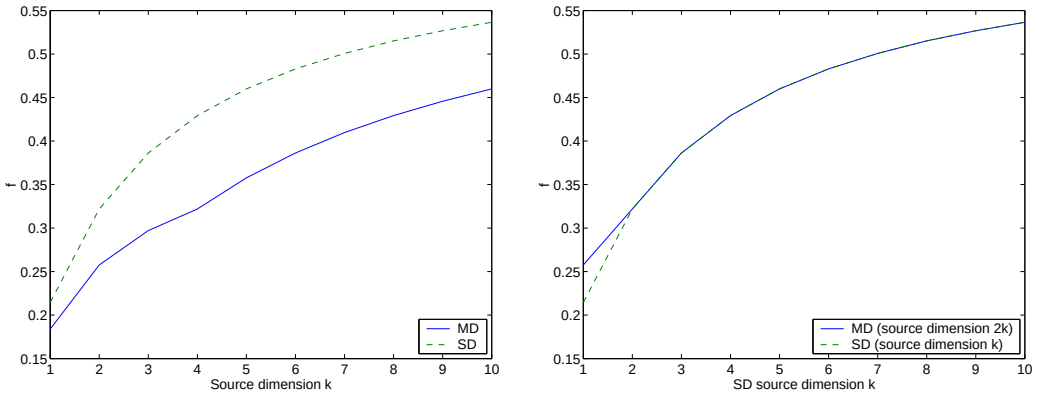


Figure 4.4: MD Redundancy Parameter a that Minimizes the Distortion Upper Bound D_{ex}

using numerical evaluations. In Figure 4.5 (a), it is assumed that the same source dimension is used for the MD coder and SD coder, making the channel blocklength in the SD scheme double that in the MD scheme. The numerical results demonstrate that under this model, the optimized MD scheme performs worse than the optimized SD scheme in terms of the distortion upper bound, as expected from the analytical results in Section 4.5.

However, as discussed in Section 4.5, if we allow the source dimension of the MD coder to be twice that of the SD coder so that the channel blocklength is kept the same for both schemes, then as shown in Figure 4.5(b), the optimized distortion exponent for the MD scheme is smaller than that for the SD scheme at very low source dimensions, while they are identical at higher source dimensions.

In Figure 4.6, the distortion upper bounds of the MD and SD schemes are compared for time-varying channels that are subject to the good state with $\epsilon = \epsilon^g$ and the bad state with $\epsilon = \epsilon^b$. The same source dimension k is used for both schemes. The MD distortion exponent f_{MD}^{gb} is determined based on Lemma 4.4. In Figure 4.6 (a), when the



(a) Same source dimension for MD and SD

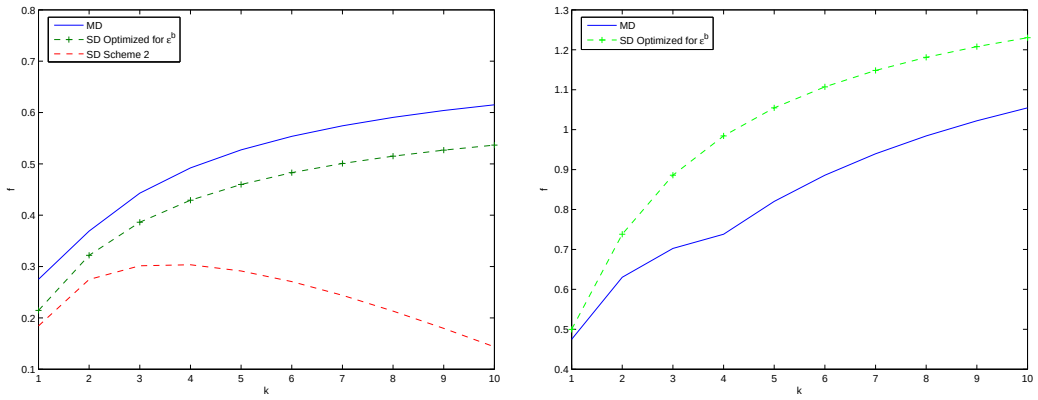
(b) Same channel blocklength for MD and SD

Figure 4.5: Comparison of Distortion Exponent f_{MD} and f_{SD} for Time-Invariant Channels at $\epsilon = 0.1$

SD distortion exponent is optimized for $\epsilon^b = 0.1$, it leads to smaller distortion exponent than that of the MD scheme optimized for $\epsilon^b = 0.1$ and $\epsilon^g = 0.01$, meaning that the MD scheme may have better performance in this case. This can also be shown by selecting the SD channel code rate r_{SD} satisfying $\min(2r_{SD}, kE^g_{ex}(r_{SD})) = f^{gb}_{MD}$ so that the SD scheme in the good state has the same distortion exponent with the MD scheme, and then evaluating the distortion exponent of the SD scheme in the bad state according to $\min(2r_{SD}, kE^b_{ex}(r_{SD}))$. From the resultant curve in Figure 4.6 (a), where it is labelled as "SD Scheme 2", one can see that in this case the distortion exponent f in the bad state for the SD scheme is much smaller than that of the MD scheme, suggesting that the MD scheme may outperform the SD scheme.

In Figure 4.6 (b), we consider $\epsilon^b = 0.01$ and $\epsilon^g = 0.001$. When the SD distortion exponent f is optimized for ϵ^b, it yields larger distortion exponent than that for the optimized MD scheme, suggesting that in this case the SD scheme may outperform the MD scheme.

In general, when ϵ^b is large and $\epsilon^g \ll \epsilon^b$, the MD scheme may have better distortion performance than the SD scheme in the time-varying channel.



(a) $\epsilon^b = 0.1, \epsilon^g = 0.01$ (b) $\epsilon^b = 0.01, \epsilon^g = 0.001$

Figure 4.6: Comparison of Distortion Exponent f_{MD}^{gb} and f_{SD} for Time-Varying Channels

4.7 Conclusion

In this paper we have investigated the asymptotic performance of cascaded MD source codes and block channel codes for BSCs. In this model there exists a tradeoff between source coding resolution, MD redundancy and channel coding rate. Analytic results show that when the channels are time-invariant, allocating redundancy to channel coding in the SD scheme may be more effective than allocating redundancy to MD coding; on the other hand, when the channels are time-varying and the channel state is unknown to the encoder and the channel bit error probability is large, the MD scheme may outperform the SD scheme.

Even though we have focused on the distortion upper bound for the MD scheme

in this paper, an analysis on a distortion lower bound based on the "sphere-packing exponent" of the BSC [38] can also be conducted for the MD scheme, and similar conclusions as above can be drawn.

Appendix A: Proof of Lemma 4.1

Define

$$F_{ex}(a,r) = \min\left(r(1+a), [2r(1-a) + kE_{ex}(r)]/2, kE_{ex}(r)\right). \tag{4.30}$$

Denote the above 3 arguments of the $\min(\cdot)$ function in (4.30) by $t_1(a,r)$, $t_2(a,r)$ and $t_3(r)$, respectively. If one term is the smallest term among these 3 terms, then as $R \to \infty$, the distortion due to this term dominates $D_{ex}(a,r)$, with

$$D_{ex}(a,r) = 2^{-2R\cdot F_{ex}(a,r)+O(1)}.$$

Assume that one term in (4.30) is smaller than the other two terms, then it can be shown that by adjusting a or r, $D_{ex}(a,r)$ can be decreased until this term equals to another term. For example, if $t_1(a,r)$ is smaller than the other two terms, then by fixing a and increasing r, a point can be reached such that $t_1(a,r) = t_3(r)$ (or $t_1(a,r) = t_2(a,r)$), because $t_1(a,r) = r(1+a)$ increases as r increases, $t_3(r) = kE_{ex}(r)$ decreases as r increases, and $t_3(C) = 0$. If $t_2(a,r)$ is smaller than the other two terms, then by fixing r and decreasing a, a point can be reached such that $t_2(a,r) = t_1(a,r)$ (or $t_2(a,r) = t_3(r)$), because $t_2(a,r) = [2r(1-a) + kE_{ex}(r)]/2$ increases as a decreases, and $t_2(0,r) \geq t_1(0,r)$.

Therefore, in order to minimize D_{ex} for large R, we only need to consider the following 4 cases, where at least two terms in (4.30) are equal to each other.

- Case 1: Assume that $t_1(a,r) = t_2(a,r) = t_3(r)$. Then

$$\begin{aligned} a &= \frac{1}{3}, \\ E_{ex}(r) &= \frac{4/3}{k} r, \\ D_{ex}(a,r) &= 2^{-2Rr \cdot 4/3 + O(1)} = 2^{-2kRE_{ex}(r) + O(1)}, \text{ as } R \to \infty. \end{aligned}$$

- Case 2: Assume $t_1(a,r) = t_2(a,r) < t_3(r)$. Then

$$\begin{aligned} a &> \frac{1}{3}, \\ E_{ex}(r) &= \frac{4a}{k} r, \\ D_{ex}(a,r) &= 2^{-2Rr \cdot (1+a) + O(1)}, \text{ as } R \to \infty. \end{aligned}$$

  Intuitively, in this case sufficient channel coding is used so that the probability that both channels fail at the same time is small.

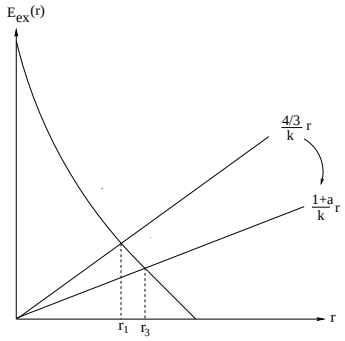


Figure 4.7: Function $E_{ex}(r)$ for Case 3

- Case 3: Assume that $t_1(a,r) = t_3(r) < t_2(a,r)$. Then

$$\begin{aligned} a &< \frac{1}{3}, \\ E_{ex}(r) &= \frac{1+a}{k} r, \\ D_{ex}(a,r) &= 2^{-2kRE_{ex}(r)+O(1)}, \text{ as } R \to \infty. \end{aligned}$$

  Let $a_1 = \frac{1}{3}$, $a_3 < a_1$, and let r_i and a_i satisfy (4.31) for $i = 1, 3$. As illustrated in Figure 4.7, since $E_{ex}(r)$ is a decreasing function of r, we have that $r_3 > r_1$ and $E_{ex}(r_3) < E_{ex}(r_1)$, meaning that D_{ex} in this case is always larger than that in Case 1. Intuitively, this is because in this case too much rate is allocated to MD redundancy to achieve small side distortion, but too little rate is allocated to source coding (for the central quantizer) and channel coding.

- Case 4: Assume that $t_2(a,r) = t_3(r) < t_1(a,r)$. Then

$$\begin{aligned} a &> \frac{1}{3}, \\ E_{ex}(r) &= \frac{2(1-a)}{k} r, \\ D_{ex}(a,r) &= 2^{-2Rr \cdot 2(1-a)+O(1)}, \text{ as } R \to \infty. \end{aligned}$$

  By an argument similar to that in Case 3, it can be shown that this case always has a larger D_{ex} than Case 1. Intuitively, this is because in this case, too much rate is allocated to source coding resolution (for the central quantizer), but too little rate is allocated to MD redundancy and channel coding.

Summarizing all cases, for large R, the pair (a,r) that minimizes $D_{ex}(a,r)$ satisfies (4.8) and (4.9).

■

Appendix B: Proof of Lemma 4.2

From (4.16), calculating the derivative of E_{ex} yields

$$\frac{dE_{ex}}{d\rho} = \frac{\delta^{1/\rho}\,(\log\delta)^2}{(1+\delta^{1/\rho})^2\rho^2\log e}.$$

From (4.15), we have

$$\frac{dr}{d\rho} = -\frac{\delta^{1/\rho}(\log\delta)^2}{(1+\delta^{1/\rho})^2\rho^3\log e} < 0. \tag{4.31}$$

Therefore

$$E'_{ex}(r) = \frac{dE_{ex}/d\rho}{dr/d\rho} = -\rho(r) \tag{4.32}$$

where $\rho(r)$ satisfies (4.15).

(4.13) can be rewritten as

$$f_{ex}(r) = [1 + B_{ex}(r)] \cdot r.$$

Thus

$$f'_{ex}(r) = B'_{ex}(r)r + 1 + B_{ex}(r) = \frac{kE'_{ex}(r)}{4} + 1. \tag{4.33}$$

Since

$$E''_{ex}(r) = \frac{dE'_{ex}(r)}{dr} = \frac{dE'_{ex}(r)/d\rho}{dr/d\rho} = \frac{(1+\delta^{1/\rho})^2\rho^3\log e}{\delta^{1/\rho}(\log\delta)^2} > 0,$$

we have

$$f''_{ex}(r) = \frac{kE''_{ex}(r)}{4} > 0. \tag{4.34}$$

From (4.11), (4.10), (4.32) and (4.5), we have

$$\frac{da}{dr} = -\frac{k\rho\left[1-\log(1+\delta^{1/\rho})\right]}{4r^2} < 0. \tag{4.35}$$

From (4.35) and (4.31), it follows that that a is an increasing function of ρ.

Assuming that there exists a stationary point a_{ex}^* such that $f_{ex}'(a_{ex}^*) = 0$, $r_{ex}^* \triangleq r(a_{ex}^*)$ must also be a stationary point such that $f_{ex}'(r_{ex}^*) = 0$. Since $f_{ex}''(r) > 0$, $a'(r) < 0$ and $f_{ex}'(a) = f_{ex}'(r)/a'(r)$, we conclude that for $a < a_{ex}^*$, $f'(a) < 0$; for $a > a_{ex}^*$, $f'(a) > 0$.

Considering (4.33) and (4.32), the stationary point ρ_{ex}^* for $f_{ex}(\cdot)$ is

$$\rho_{ex}^* = \frac{4}{k}.$$

■

Appendix C: Proof of Theorem 4.3

As defined by (4.14), ρ is larger than or equal to 1. From Lemma 4.2, a is an increasing function of ρ for $\rho > 1$. It follows that if $\rho(a)$ is ever equal to 1, then there must exist a_0, such that for all $a > a_0$, $\rho(a) > 1$, and for all $a \leq a_0$, $\rho(a) = 1$. Two special cases are: (i) $a_0 < 0$, which means that $\rho(a)$ is not equal to 1 for any $a \in [0, +\infty)$; (ii) $a_0 = +\infty$, which means that $\rho(a) = 1$ for all $a \in [0, +\infty)$.

Now we consider 3 cases separately: $k < 4$, $k = 4$, and $k > 4$.

- $k < 4$

  In this case, (4.17) indicates that the stationary point ρ_{ex}^* for $f_{ex}(\cdot)$ satisfies $\rho_{ex}^* > 1$. Furthermore, a_{ex}^* can be computed from (4.19). Based on the definition of a_0,

$a_0 < a^*_{ex}$. For $a \leq a_0$, $\rho(a) = 1$, thus from (4.22), we have $f'_{ex}(a) < 0$. For $a > a_0$, $\rho(a) > 1$, thus from (4.18), we have $f'_{ex}(a) < 0$ for $a_0 < a < a^*_{ex}$, and $f'_{ex}(a) > 0$ for $a > a^*_{ex}$. Therefore, $f_{ex}(a)$ is a decreasing function for $a < a^*_{ex}$, and an increasing function for $a > a^*_{ex}$.

Using this result on $f'_{ex}(a)$, since the permissible $a \in [\frac{1}{3}, 1]$, it follows that the parameter a that maximizes $f_{ex}(a)$ is either $\frac{1}{3}$ or 1. If the stationary point a^*_{ex} is outside the permissible interval of a, then the far end of the interval is the parameter a that maximizes f_{ex}. That is, $a_{ex} = \frac{1}{3}$, if $k < 4$ and $a^*_{ex} \geq 1$; $a_{ex} = 1$, if $k < 4$ and $a^*_{ex} \leq \frac{1}{3}$. However, if a^*_{ex} lies within the permissible interval, then one need compare the function values at the two ends of the interval and choose the value of a that has a larger $f_{ex}(a)$.

- $k = 4$

  In this case, (4.17) indicates that the stationary point $\rho^*_{ex} = 1$. Furthermore, a^*_{ex} can be computed from (4.19). Based on the definition of a_0, $a_0 = a^*_{ex}$. It follows that for $a \leq a^*_{ex}$, $\rho(a) = 1$, and thus from (4.22), we have $f'_{ex}(a) = 0$; for $a > a^*_{ex}$, $\rho(a) > 1$, and from (4.18), we have $f'_{ex}(a) > 0$.

  Using this result on $f'_{ex}(a)$, since the permissible $a \in [\frac{1}{3}, 1]$, it follows that if $a^*_{ex} \leq 1$, then the parameter a that maximizes $f_{ex}(a)$ is 1; if $a^*_{ex} > 1$, then $f_{ex}(a)$ is constant for any $a \in [\frac{1}{3}, 1]$. In the latter case, for simplicity, we may choose 1 for a. Therefore, the parameter a that maximizes f_{ex} is $a_{ex} = 1$, if $k = 4$.

- $k > 4$

  In this case, (4.17) indicates that the stationary point for $f_{ex}(\cdot)$ satisfies $\rho^*_{ex} < 1$. But this contradicts to the condition $\rho \geq 1$. Therefore we must have $a^*_{ex} < a_0$. For $a > a_0$, $\rho(a) > 1$ and $f'_{ex}(a) > 0$; on the other hand, for $a \leq a_0$, (4.22) indicates

that $f'_{ex}(a) > 0$. In both situations, $f'_{ex}(a) > 0$. Since the permissible $a \in [\frac{1}{3}, 1]$, the parameter a that maximizes f_{ex} is $a_{ex} = 1$, if $k > 4$.

■

This chpater, in full, has been submitted for publication as: X. Tian. Asymptotic Analysis of Multiple Description Codes Over Noisy Channels, *IEEE Transactions on Information Theory*, 2004. The dissertation author was the primary investigator and single author of this paper.

Chapter 5

Punctured Trellis Coded Quantization and Modulation

5.1 Introduction

Trellis coded modulation (TCM) is a combined coding and modulation technique for digital transmission over band-limited channels. It can achieve significant coding gain without bandwidth expansion or reduction of information rate [74]. Trellis coded quantization (TCQ) was introduced as a natural dual to TCM [51]. By combining a trellis structure with a quantizer with an expanded set of quantization levels, TCQ outperforms many other source coding techniques of comparable complexity in encoding both memoryless sources and sources with memory.

TCQ and TCM were combined by Fischer and Marcellin [22] to produce an effective joint source and channel coding/modulation system. By using the same trellis code for TCQ and TCM, and by using a consistent labeling between TCQ levels and TCM symbols, symbol errors at the TCM decoder are unlikely to cause large mean squared errors at the TCQ decoder. As a result, the combined TCQ/TCM scheme is

more error resilient than the scheme where TCQ and TCM are separately designed.

In [79], Wang and Fischer designed a combined TCQ/TCM scheme that uses a channel-optimized TCQ and that optimizes TCM constellation points with a quasi-Newton method. Compared with [22], the performance at low channel SNRs is improved, but the training algorithm is fairly complex. In [35] and [36], a symbol-by-symbol *Maximum a Posteriori* (MAP) algorithm was modified as a soft-decoder for the combined TCQ/TCM, and the TCQ encoder and the soft-decoder were designed iteratively to reduce the distortion. Significant improvement was obtained for very noisy channels, again at the cost of large complexity.

While it is known that jointly designing the source coder, channel coder, and modulator may improve the overall performance, a lower complexity effective method is one that optimally allocates rates between the source coder and channel coder. There is considerable amount of work in the literature that utilizes the tradeoff between source coding and channel coding. For example, in [5], a trellis-coded vector quantization (TCVQ) source coder and a convolutional channel coder were jointly designed; while in [12], a TCQ coder was followed by a Rate-Compatible Punctured Convolutional (RCPC) coder. In both cases the rate allocation between source codes and channel codes was considered.

In this chapter, we explore such tradeoffs between source coding and channel coding in the context of TCQ/TCM schemes for an AWGN channel, under the constraint that each source sample is coded into one channel symbol. Inspired by the success of TCQ derived from TCM using a duality between quantization and modulation, and that of the Punctured TCM scheme (PTCM) of Wolf and Zehavi [84], we propose a new source coding scheme PTCQ as the dual of PTCM, providing efficient source coding at fractional bit-rates with low complexity. Then we generalize the combined TCQ/TCM scheme to a PTCQ/PTCM scheme, providing the flexibility of adjusting the tradeoff be-

tween source coding and channel coding with a single underlying trellis, thus improving the performance with a noisy channel.

An example to illustrate the proposed PTCQ/PTCM scheme is shown in Figure 5.1, where every 2 source samples are coded into 5 information bits, which are then expanded to 6 bits, forming two 8-PSK symbols. In contrast, in the TCQ/TCM scheme of [22] with 8-PSK, each source sample is coded into 2 information bits, which are expanded to 3 bits, forming one 8-PSK symbol.

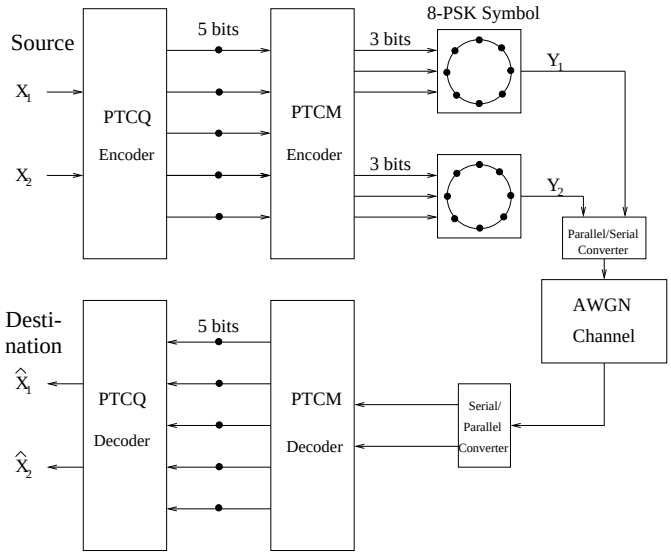


Figure 5.1: An Example of PTCQ/PTCM for 8-PSK

The rest of the paper is organized as follows. In Section 5.2, an overview of the PTCM scheme based on [84] is presented. In Section 5.3, the PTCQ/PTCM scheme is proposed. Simulation results are presented in Section 5.4. Section 5.5 summarizes this chapter.

5.2 Overview of PTCM

In [84], a puncturing technique widely used for binary convolutional codes was applied to TCM. By not transmitting certain output bits of the trellis codes, a class of higher-rate

trellis codes is formed, called PTCM.

5.2.1 Encoding

Consider the PTCM scheme for M-PSK with an underlying convolutional code of rate $1/2$. The convolutional coder in TCM is replaced by a punctured convolutional coder with rate $r \in [\frac{1}{2}, 1)$. For 8-PSK, this achieves an overall transmission rate of $R \in [2, 3)$ bits/symbol. At time k, PTCM produces 3 bits $(x_k^2\ x_k^1\ x_k^0)$, where x_k^1 and x_k^0 are coded bits, and x_k^2 is an uncoded bit. The following mapping (denoted $S_{PR}(\cdot)$) is assumed between the binary digits $(x_k^2\ x_k^1\ x_k^0)$ and PSK constellation points: $(000) = 0^o,\ (001) = 45^o,\ (011) = 90^o,\ (010) = 135^o, (100) = 180^o,\ (101) = 225^o,\ (111) = 270^o$, and $(110) = 315^o$, where the signal phases are used to specify the PSK constellation points. This scheme was termed *pragmatic codes* in [84] because it allows existing rate $1/2$ convolutional encoders and decoders to be used for 8-PSK TCM after slight modifications. For QPSK, PTCM achieves a transmission rate of $R \in [1, 2)$ bits/symbol. A Gray mapping $S_{GR}(\cdot)$ is used to map binary digits $(x_k^1\ x_k^0)$ to QPSK constellation points: $(00) = 0^o$, $(01) = 90^o$, $(11) = 180^o$, and $(10) = 270^o$.

For an underlying convolutional code of rate $1/2$, if the number of symbols that are transmitted per puncture period P is denoted by S, then for every P input (information) bits, the punctured encoder generates $2S$ output bits, yielding a code rate of $r = P/(2S)$. For QPSK, these $2S$ output bits are grouped into S QPSK symbols, so the overall bit-rate of the PTCM coder is P/S bits/symbol. For 8-PSK, an additional S uncoded bits can be combined with those $2S$ coded bits into S 8-PSK symbols, so the overall bit-rate of the PTCM is $(P + S)/S$ bits/symbol. Some examples of puncturing tables and their corresponding parameters are given in Table 5.1,

Punc. Table	$\begin{pmatrix} 1 \\ 1 \end{pmatrix}$	$\begin{pmatrix} 11x1 \\ 1x11 \end{pmatrix}$	$\begin{pmatrix} 11x \\ 1x1 \end{pmatrix}$	$\begin{pmatrix} 11x1x \\ 1x1x1 \end{pmatrix}$
P	1	4	3	5
S	1	3	2	3
Code-rate r	$1/2$	$2/3$	$3/4$	$5/6$
Bit-rate for 8-PSK (bit/sym.)	2	$2\frac{1}{3}$	$2\frac{1}{2}$	$2\frac{2}{3}$
Bit-rate for QPSK (bit/sym.)	1	$1\frac{1}{3}$	$1\frac{1}{2}$	$1\frac{2}{3}$

Table 5.1: PTCM parameters

5.2.2 Decoding

Due to puncturing, one TCM symbol may contain coded bits from two trellis transitions. In order to use a conventional Viterbi algorithm for decoding, sub-optimal branch metrics need to be used to approximate the optimal branch metrics. In this chapter, we adopt the sub-optimal branch metrics from [83].

Denote the Euclidean distance between the received signal y_k and the transmitted 8-PSK symbol corresponding to bits $(x_k^2\ x_k^1\ x_k^0)$ as

$$d_{x_k^2 x_k^1 x_k^0} = ||y_k - S_{PR}(x_k^2 x_k^1 x_k^0)||. \tag{5.1}$$

If there is no puncturing at time k, then define the following sub-optimal branch metric [83]:

$$\begin{aligned} \Delta M_{AP}(x_k^1 x_k^0) &= \min(d^2_{0x_k^1 0}, d^2_{0x_k^1 1}, d^2_{1x_k^1 0}, d^2_{1x_k^1 1}) \\ &+ \min(d^2_{00x_k^0}, d^2_{01x_k^0}, d^2_{10x_k^0}, d^2_{11x_k^0}) \end{aligned} \tag{5.2}$$

where "AP" refers to the approximate or sub-optimal nature of this metric.

If there is puncturing at time k so that coded bits x_k^1 and x_k^0 are generated by two trellis transitions, then define the sub-optimal metrics of the branchs corresponding to x_k^1 and x_k^0 as the first and second term of the right-hand side of (5.2), respectively.

By using the above branch metrics and the Viterbi algorithm, the PTCM decoder can determine the sequence of coded bits x_k^1 and x_k^0. The uncoded bit x_k^2 can be chosen such that $d^2_{x_k^2 x_k^1 x_k^0} = \min(d^2_{1 x_k^1 x_k^0}, d^2_{0 x_k^1 x_k^0})$.

For QPSK, the symbols are mapped from coded bits ($x_k^1 x_k^0$) and no uncoded bits x_k^2 are used, so the above definitions of Euclidean distance and branch metrics need to be adjusted accordingly. For example, (5.1) is replaced by $d_{x_k^1 x_k^0} = ||y_k - S_{GR}(x_k^1 x_k^0)||$, and (5.2) is replaced by $\Delta M_{AP}(x_k^1 x_k^0) = \min(d^2_{x_k^1 0}, d^2_{x_k^1 1}) + \min(d^2_{0 x_k^0}, d^2_{1 x_k^0})$.

5.3 Proposed PTCQ/PTCM Scheme

5.3.1 PTCQ

We propose a source-optimized PTCQ scheme as the dual of PTCM, using the mechanism by which TCQ is derived from TCM. In this scheme, the PTCQ encoder is similar to the PTCM decoder, and the PTCQ decoder is similar to the PTCM encoder. The encoding trellis of PTCQ is defined by a punctured convolutional encoder. The output of the PTCQ encoder is a sequence of binary codewords, which are converted to a sequence of quantization levels at the decoder. By combining a punctured trellis structure with a quantizer, PTCQ can achieve quantization at fractional bit-rates that are greater than 1 bit/source sample. In its simplest form, with an underlying convolutional code of rate $1/2$, PTCQ coded at R bits/source sample requires $2^{\lfloor R \rfloor + 1}$ quantization levels (codewords), partitioned into 4 subsets, each of $2^{\lfloor R \rfloor - 1}$ codewords. The sub-optimal branch metrics for PTCQ are similar to those described in Section 5.2 for PTCM, except that now $S_{PR}(\cdot)$ and $S_{GR}(\cdot)$ map binary digits to quantization levels instead of PSK symbols. The Viterbi algorithm is used to encode the source sequence. The PTCQ quantization levels can be trained using the Lloyd algorithm and a training sequence.

Assume that a doubled signal constellation is used for the underlying trellis. If M is the number of inputs to the convolutional coder, then the rate of the underlying convolutional code is $M/(M+1)$. For the rate $1/2$ code considered in this chapter, $M = 1$. The number of states of the trellis is denoted N_s. It can be shown that the encoding complexity of PTCQ consists of:

$2^{\lfloor R \rfloor + 1}$ multiplies,

$2^{\lfloor R \rfloor + 1} + 2^{M+1} + N_s 2^M P/S$ additions,

$[(2^{\lfloor R \rfloor} - 1)(M+1)2^{M+1} + N_s(2^M - 1)P/S$

$+(2^{\lfloor R \rfloor - M} - 1)]$ two-way comparisons

per source sample. The fractional bit-rate quantizaiton realized by PTCQ can also be realized by TCVQ [23]. The full search encoding complexity per (scalar) source sample of $L-$dimensional TCVQ is ((3) of [23]):

2^{RL+1} multiplies,

$2^{RL+1} + (N_s 2^M / L)$ additions,

$\{\left[2^{M+1}(2^{RL-M} - 1) + N_s(2^M - 1)\right]/L\}$

two-way comparisons.

It is easy to see that the encoding complexity of PTCQ is only slightly higher than that of TCQ (TCVQ with $L = 1$), but is much smaller than that of TCVQ at fractional bit-rates ($L > 1$), and the difference becomes greater for larger L.

5.3.2 PTCQ/PTCM

We propose PTCQ/PTCM as a combination of PTCQ and PTCM with proper rate allocation. This combination is done in the same way as how TCQ and TCM are combined into the TCQ/TCM scheme. In particular, a same punctured trellis structure is used for PTCQ and PTCM, and consistent labeling between PTCQ levels and PTCM symbols is employed.

A key component of this scheme is to exploit the flexibility of rate allocation offered by the puncturing mechanism. Assuming that the encoder has the knowledge of the channel SNR (via a feedback channel), it can select appropriate puncturing patterns based on this knowledge, and thus allocate rates between source coding and channel coding to match the given channel SNR.

The overall performance of the PTCQ/PTCM can also be improved by adaptive modulation. When the channel is favorable, a higher-order modulation scheme can be used to increase the system's bits per symbol capacity, and conversely, a more robust lower order modulation scheme can be used when the channel becomes inferior. For example, at high channel SNRs, 8-PSK may be used to transmit $R \in [2, 3)$ bits/symbol; while at low channel SNRs, QPSK may be used to transmit $R \in [1, 2)$ bits/symbol.

5.4 Results

In this section, we present numerical results of the proposed PTCQ and PTCQ/PTCM schemes.

The combined PTCQ/PTCM is designed for transmitting certain memoryless sources through an AWGN channel according to the algorithm described previously. A sequence of 100 blocks, each of 1000 samples, is used for training; a separate sequence of data of the same block length is used for testing. A 4-state rate 1/2 convolutional

	Gaussian Source		Uniform Source	
Encoding rate (bit/sample)	R=1.5	R=2.5	R=1.5	R=2.5
SQNR for PTCQ (dB)	7.48	12.79	9.15	15.49
SQNR for TCVQ (dB)	7.65	13.27	9.38	15.70

Table 5.2: Quantization Performance of PTCQ and TCVQ

encoder specified by $h_0 = 5_8$ and $h_1 = 7_8$ is used. The modulation constellation is 8-PSK or QPSK, depending on the bit-rate. The performance results are given in the form of SQNR vs. channel SNR. SQNR is defined as $\mathrm{SQNR} = \frac{E\{y^2\}}{E\{(y-\hat{y})^2\}}$, where y is the signal to be quantized, and $\hat{y}$ is the reconstructed signal. The channel SNR is defined as $\mathrm{SNR} = \frac{E_s}{2\sigma_n^2}$, where E_s is the energy of the PSK signal, and σ_n^2 is the variance of the Gaussian noise along each dimension.

Before presenting the simulation results of the combined PTCQ/PTCM scheme, let us examine the quantization performance of PTCQ. The simulation results at 2 fractional bit-rates for memoryless Gaussian and uniform sources are shown in Table 5.2. Also shown in this table are the coding performance results of TCVQ, which are obtained based on our implementation of the TCVQ coder (Structure 1) of [23]. One can see that the performance of PTCQ is close to that of TCVQ. But as noted in Section 5.3.1, PTCQ is less complex than TCVQ.

In Figure 5.2, the performance results of PTCQ/PTCM for a memoryless Gaussian source at 6 bit-rates ($R = 1, 1\frac{1}{3}, 1\frac{2}{3}, 2, 2\frac{1}{3}$, and $2\frac{2}{3}$ bits/symbol) are shown. Also shown in the graph are the performance results of the reference system – the source-optimized TCQ/TCM scheme at 2 bits/symbol [22]. These results show that performance improvement is possible at low and high SNRs if the bit-rate can be optimally selected for the given channel SNRs. Similar performance improvement can also obtained for memoryless uniform and Laplacian sources.

In Figure 5.3, the optimal rate allocation for PTCQ/PTCM is shown as a function

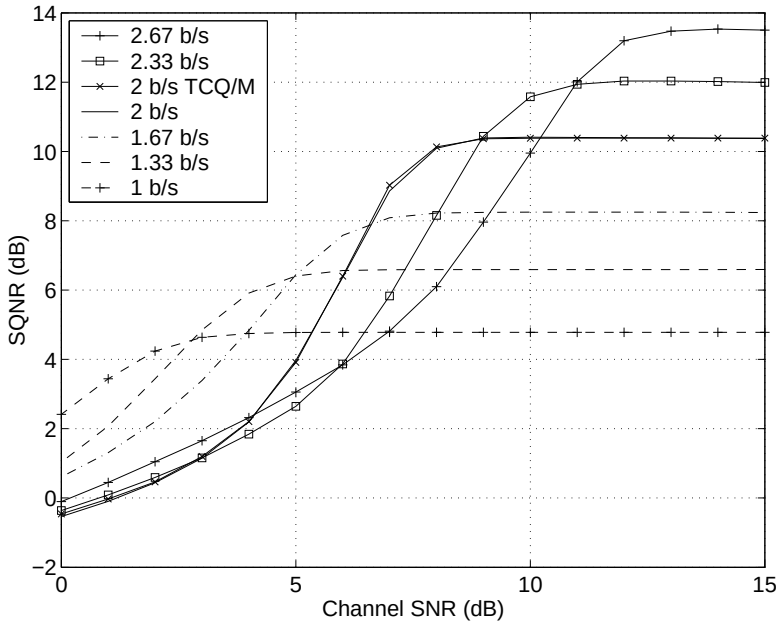


Figure 5.2: PTCQ/PTCM for Gaussian Source

of the channel SNR for Gaussian, uniform, and Laplacian sources. These results are obtained by measuring the end-to-end distortions via simulations and selecting, among the 6 rates, the rate that gives the least distortion. Similar results can also be obtained by analyzing the distortion due to quantization and that due to channel noise with the punctured trellis code, but the analysis is omitted here. A rule-of-thumb observed from these results is that among the 6 available rates, the best rate corresponds to the largest possible bit-rate with the symbol error rate of PTCM decoding being less than 1%.

In Figure 5.4, the performance results of the PTCQ/PTCM using the optimal rates as described above are compared with those of other methods for a memoryless Gaussian source. Compared with a reference system (TCQ/TCM at 2 bits/symbol) [22] in which the Viterbi algorithm (VA) was used, by choosing from 6 rates, PTCQ/PTCM can achieve about 3 dB gain in SQNR at both low channel SNRs and high channel SNRs, while the gain is smaller at intermediate SNRs. Figure 5.4 also shows the performance

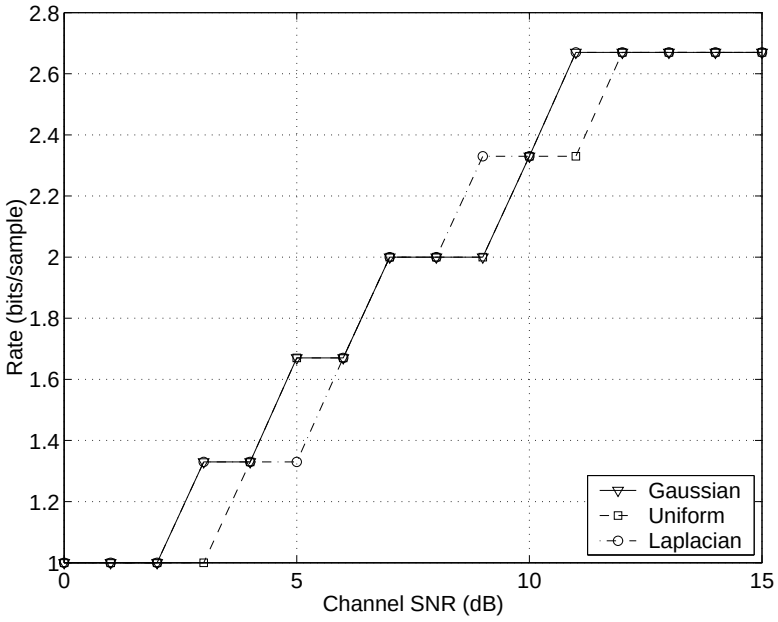


Figure 5.3: Rate Allocations Results for PTCQ/PTCM Based on Simulations

of the iterative MAP decoding algorithm of [35], where an 8-state trellis was used for a TCQ/TCM 8-PSK system. It can be seen that the performance of the algorithm of [35] is similar to that of the PTCQ/PTCM scheme at low and medium channel SNRs, but its performance is worse than PTCQ/PTCM at high channel SNRs. Its decoding complexity is much higher than that of PTCQ/PTCM [1].

So far we have assumed that bit-wise puncturing is used in the PTCQ/PTCM scheme; however, symbol-wise puncturing [82] can also be used to form a symbol-wise PTCQ/PTCM scheme. Bit-wise puncturing refers to the puncturing mechanism in which some or all transmitted symbols contain coded bits from more than one trellis

[1] It can be shown that for the $4-$state punctured trellis codes used in this section, the total decoding complexity of PTCQ/PTCM is no greater than 16 multiplies, 34 additions, and 33 two-way comparisons per source sample. For comparison, a conservative estimate of the complexity of the decoding algorithm in [35] [36] is $3BN_s^2$ multiplies and $2BN_s^2$ additions per source sample, where B is the number of parallel branches between states. Since $B = 2$ and $N_s = 8$ were used, this amounts to 384 multiplies and 256 additions per source sample.

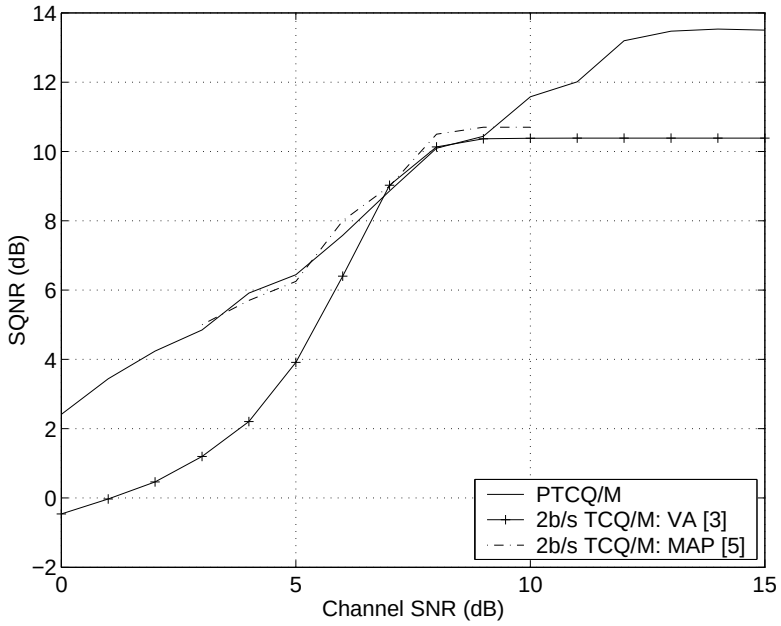


Figure 5.4: Comparison of PTCQ/PTCM with Other Methods

transition, while symbol-wise puncturing refers to the case when each transmitted symbol corresponds to one trellis transition. According to our experimental results (which are not shown here), bit-wise and symbol-wise PTCQ/PTCM schemes have similar performance results, but the symbol-wise scheme is simpler.

5.5 Conclusion

In this chapter, we have proposed a Punctured TCQ scheme as the dual of the Punctured TCM scheme, providing a variety of fractional bit-rates that are greater than 1 bit/sample. PTCQ is shown to be able to achieve performance close to TCVQ but at less complexity. We have also proposed a combined PTCQ/PTCM scheme so that by using proper modulation order and puncturing pattern, rates can be optimally allocated between source coding and channel coding to match the given channel SNR. Simulation

results show that significant performance improvement can be achieved by the proposed scheme over the reference TCQ/TCM scheme at both low and high channel SNRs, while its complexity is much lower than that of a scheme with competitive performance at low channel SNRs.

This chpater, in part, has been submitted for publication as: X. Tian. Punctured trellis coded quantization and modulation. *IEEE Transactions on Communications*, 2004. The dissertation author was the primary investigator and single author of this paper.

This chpater, in part, has been submitted for publication as: X. Tian. Punctured trellis coded quantization and modulation. *IEEE Transactions on Communications*, 2004. The dissertation author was the primary investigator and single author of this paper.

Chapter 6

Efficient Transmission Power Allocation for Wireless Video Communications Under Average Power Constraints

6.1 Introduction

Next-generation wireless communication systems must accommodate multimedia services. Transmitting videos over wireless systems is a challenging task due to the time-varying nature of wireless channels and the limited resources of wireless systems such as bandwidth and power. Thus resource allocation constitutes a key problem in wireless video communications.

A common type of resource allocation is the optimal allocation of rates between source coding and channel coding. Since most video codecs use motion compensated prediction, decoding errors of forward error control (FEC) will propagate into future

frames and may cause severe performance degradation. To make video decoders resilient to channel decoding errors, proper source synchronization and error concealment strategies must be employed. While video is inter-coded in most frames to maintain coding efficiency, intra-macroblock (MB) or intra-frame refreshing is often used to stop error propagation. All these factors make the theoretical analysis of rate allocation between source coding and channel coding extremely difficult.

In [88], a robust source coding algorithm is proposed for transmitting videos over packet networks. The expected decoder distortion is analyzed at pixel-level precision for a given packet loss rate, and the intra/inter mode is selected for each MB based on a rate-distortion model. Even though this approach of pixel-level estimation accurately accounts for both spatial and temporal error propagation, it does not allow global optimization of rate allocation between source coding and channel coding for the video sequence, as it estimates the distortion recursively frame by frame. Its computational complexity is also high, since both the first and second moments of the reconstructed values at the decoder need to be calculated. In [73], the error propagation effect is analyzed under the assumption that the channel-induced error variance is the same on average for each frame. An empirical model of the distortion due to quantization (source coding) and that due to channel errors for an H.263 video codec is proposed, so that the optimal intra-MB refreshing rate and FEC code rate can be chosen for the video sequence. However, this algorithm requires running computations over the video sequence a few times to generate measurement points before the optimal parameters can be selected, making it not suitable for real-time communications. In [34], an algorithm for joint source and channel coding is proposed for video transmission over a binary symmetric channel. In addition to developing a rate-distortion model for source coding, the relationship between the average channel distortion over the sequence and the source/channel coding parameters such as the intra-MB refreshing rate and FEC code

rate is also derived, allowing optimal selection of these parameters for the video sequence. However, the analysis is strictly based on the assumption that the decoder error concealment strategy is simply to copy the MBs at the same location from the previous decoded frame, which limits the performance of the codec.

Another type of resource allocation for wireless video communications involves the optimal allocation of the total consumed power or the transmission power.

Reducing the total consumed power for mobile devices improves their battery life. A common optimization approach for this type of problems is to allocate the total consumed power to different components of the device, including source/channel coding processing and transmission, because the power consumed for processing, which is proportional to complexity, is not negligible when evaluating the total consumed power for wireless video communications. In [87], the transmission power and the bit allocation between source coding and channel coding are jointly optimized, according to wireless channel conditions and video quality requirements, to minimize the total consumed power for a single user and a group of users in a cell. In [50], a similar optimization problem is considered for the H.263 source coder, the Reed-Solomon channel coder and a two-state Markov channel model, using the distortion model of video transmission proposed in [73]. This approach is extended to the context of multiple antennas and space-time coding in [85].

On the other hand, reducing the transmission power will reduce the interference between users sharing a wireless channel and thus enhance the system capacity. Transmission power allocation may also extend the battery life for hand-held devices, since the transmission power is a significant part of the total consumed power. In [20], a joint source coding and transmission power allocation scheme is proposed. In this scheme, no channel coding is used, and the transmission power is allocated adaptively to different video segments based on their relative importance. The optimization problem

is formulated such that the transmission energy required to transmit a video frame is minimized at some acceptable decoded video quality and transmission delay. However, the algorithm is based on the assumption that each MB is a separate packet for transmission, which, as acknowledged in [20], has a low coding efficiency because of the large overhead it would incur if implemented. In order to jointly select the source coding mode/stepsize and transmission power for packets in a video frame, the algorithm also involves building a source coding tree to take into account the mode selection for each packet, making it computationally complex. In [43], a received power management scheme is described for a two-class video stream in a base channel and a motion-vector (MV) channel. The analysis is based on intra-frame refreshing once every N frames, which requires larger buffer than that for intra-MB refreshing schemes. A different quality of service is provided to each frame within an intra refreshing period to minimize the average received power while satisfying a distortion constraint, with more power allocated to the earlier frames within the refreshing period. A fixed-rate FEC is used for various channel Signal-to-Noise Ratios (SNRs) in this scheme. In [44], a transmit power management scheme is proposed for video transmission in CDMA systems. A quasi-static flat-fading channel is assumed, and the bit error rate (BER) is adjusted to the changes of image characteristics on the Group of Blocks (GOB) level to meet the performance requirement. The solution presented there requires non-causal CSI for the whole video sequence known to the transmitter before the transmit power for any single GOB can be allocated, and no discussion on how to apply the approach to practical scenarios is given. In [10], the resource requirement for a CDMA system with a large number of users is approximated as the product of the source coding rate and the energy per information bit to noise ratio, and under this assumption, the source coding rate, channel coding rate and the transmission power level for a wireless video user are jointly selected to improve the overall performance of the video transmission. In a sense,

this approach allocates the source rate and transmission power across different users of a CDMA cell.

In this work, we will investigate the transmission power allocation approach for wireless video communications under certain bandwidth and average transmission power constraints. By focusing on methods that allocate transmission power to different source segments according to their source error sensitivity and channel conditions, our work is in spirit similar to that of [20], [43] and [44]. However, we consider a more realistic scenario, where a fixed-length packetization structure is used to facilitate error concealment without sacrificing too much coding efficiency, and the FEC code rate is adapted to the channel SNR to provide a reasonably low bit error rate after channel decoding, and no future CSI is required at the transmitter. While these assumptions may improve the video transmission performance of the conventional scheme, they could potentially reduce the achievable performance gain of a transmission power allocation scheme over the conventional scheme. It is the purpose of this chapter to explore the transmission power allocation approach under these realistic assumptions, and examine its usefulness accordingly. By adopting a frame-level analysis approach, we propose simple transmission power allocation algorithms to apply unequal error protection to video segments of different error sensitivity. Depending on whether information regarding the state of the channel is available at the transmitter, two cases are considered under this framework. We further present closed-form solutions that require knowledge of the video packet error sensitivity information and the CSI for the whole sequence, which can be interpreted as a bound to the performance of practical algorithms, and help to illustrate the roles of the source error sensitivity information and the CSI in allocating transmission power. Simulations are conducted for a time-correlated Rayleigh fading channel with multiple transmit antennas, and results show that the proposed power allocation algorithms outperform the conventional algorithms under both cases.

The rest of the chapter is organized as follows. In Section 2, the basic system model is described, including the source coder, channel coder and channel model. In Section 3, two power allocation algorithms are proposed. Simulation results are presented in Section 4. Section 5 summarizes the chapter.

6.2 System Model

The system considered in this chapter is shown in Figure 6.1, and its components will be discussed below in details.

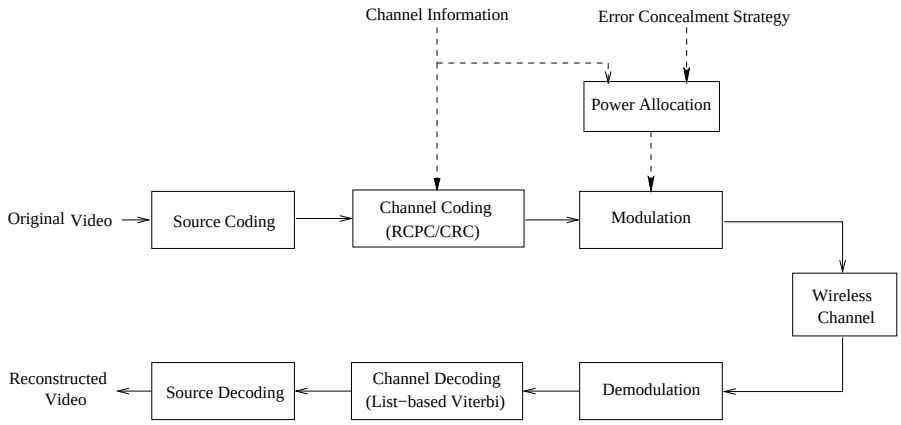


Figure 6.1: System Diagram for Wireless Video Transmission

6.2.1 Video Encoder

For source coding, we use a single-layer H.263 compliant video coder [14] . A frame-level rate control algorithm is used to determine the quantization parameters for each frame. No advanced options of H.263 are used, except that a 17-bit synchronization marker is inserted in each GOB. This frequency of synchronization markers balances the error resiliency and coding efficiency of the video codec.

Periodic intra refreshing is used to mitigate the effects of temporal error propagation. Refreshing only a portion of each picture on an MB basis instead of refreshing a whole frame produces a more constant bit-rate. Encoding each MB in the intra mode at least once every 2 seconds can limit the video quality degradation without a significant reduction in coding efficiency [18]. This can be implemented by initializing a counter to a random value for each MB and encoding the MB in the inter mode unless the counter reaches a given threshold; in that case, the MB is encoded in the intra mode and the counter is reset.

6.2.2 Video Decoder

We assume that the first frame in the sequence is intra-coded and enough transmission power is used so that it is correctly decoded. For all subsequent frames, if a packet is labeled by the channel decoder as lost due to decoding errors, all the MBs with which the packet overlaps will be error concealed. A temporal error concealment technique is employed, with the median of the MVs of the nearest 3 MBs in the GOB above used for the lost MB [88] . Any lost MB in the first row will be assumed to have a zero-MV.

Note that other error concealment strategies can also be used in conjunction with the proposed power allocation algorithms in this work.

6.2.3 Packetization

Assume that the length of each channel packet is N_c, and for a given channel code rate r_c, the number of source bits that can be accommodated in a packet is $N_s(r_c)$. Without additional header information, the loss of a packet would cause all the source bits before the next GOB to be lost and thus incur large distortion. In order to improve the error resiliency of the codec, a 16-bit header is added to the source part of each packet in the

format shown in Figure 6.2 . The first 7 bits indicate the index of the next MB whose starting bit is in the packet; the number 7 is chosen because for QCIF videos there are 99 ($< 2^7$) MBs in a frame. The following 9 bits indicate the position of the starting bit of the next MB within the source portion of the packet, allowing up to 512 source bits in a packet. If the current packet does not overlap any MB boundary, then all the 16 bits in the header are set to 1 as a flag. When the previous packet is lost but the current one is correctly received, the source decoder may resume normal decoding starting from the MB determined from the 16-bit header. This 16-bit overhead is normally smaller than that of inserting synchronization markers at the MB level.

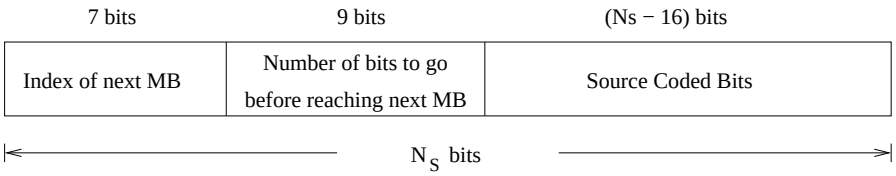


Figure 6.2: Format of the Source Portion of a Packet

Furthermore, for simplicity, we assume that no packet overlaps the frame boundary. This can be fulfilled by padding zeros to the last packet in each frame.

6.2.4 Channel Coding

We use a 16-bit Cyclic Redundancy Code (CRC) outer coder for error detection, and a Rate-Compatible-Punctured Convolutional (RCPC) coder for error correction [71]. The RCPC codes are obtained by puncturing a 16-state convolutional code with puncturing period 8, as given in Table I of [33]. A total of 13 available code rates with code index n_c from 0 to 12 are $8/9$, $4/5$, $2/3$, $4/7$, $1/2$, $4/9$, $4/10$, $4/11$, $1/3$, $4/13$, $2/7$, $4/15$ and $1/4$, respectively. The channel code rate actually used is selected from these rates to adapt to the channel SNR, and the details will be discussed later in this chapter.

A list-based Viterbi algorithm is used to find the best candidate in the RCPC

trellis, and the CRC is checked to see if it is a valid codeword. If there is an error detected, then find the next candidate, and check the CRC, and so on. If a valid codeword can not be found after checking the listed paths for N_L times, an error is declared for the source decoder to handle. In the simulations we adopt $N_L = 100$, as further increasing this number does not improve the error performance significantly [71].

6.2.5 Fading Model and Transmit Diversity

Power control can be used to combat channel fluctuations by modifying the transmission power so that a certain received signal strength or BER is maintained over the communication link. Based on how fast a radio signal fluctuates over time or travel distance, the mechanisms behind radio propagation can be classified into large-scale fading effects such as path loss and shadowing, and small-scale fading effects such as multipath fading [62]. Since the frequency bands of the uplink and downlink are separated by more than the channel coherence bandwidth, multipath fading on the uplink is uncorrelated with that on the downlink; on the other hand, path loss and shadowing are usually assumed to be identical on both links. As a result, open-loop power control can be used to estimate the channel state on the opposite link and compensate for path loss and shawdowing reasonably well without requiring feedback from the receiver; on the other hand, when the information on the channel state is available at the transmitter through feedback, and if the round-trip delay is smaller than the correlation time of the channel, closed-loop power control can be used to compensate for rapid multipath fading [55]. In this work, we assume that large-scale fading has been compensated by open-loop power control, so only small-scale fading needs to be considered. A commonly used small-scale channel model is that of Rayleigh fading. In this work we use the Jakes' model to generate time-correlated Rayleigh fading parameters [16] . The maximum Doppler shift

of channel fading can be computed by

$$f_D = f_c \frac{v}{c} \tag{6.1}$$

where f_c is the carrier frequency, v is the mobile speed, and c is the speed of light. Additive White Gaussian Noise (AWGN) is added to the signal at the output of the channel.

To combat channel fading, time diversity (interleaving), combined with FEC, is often used to provide significant performance improvement. However, time interleaving may result in large delay when the channel varies slowly. Antenna diversity, on the other hand, can be effective in reducing the effects of fading without introducing large delays. In this work, we employ a transmit diversity technique introduced in [1], with two transmit antennas and one receive antenna. The signal sequence is encoded in the two transmit antennas, and a signal combiner is employed at the receiver. By obtaining the equivalent channel gains and additive noise, and then integrating them into the channel decoder, antenna diversity can be achieved. The details of the transmit diversity design can be found in [1] .

6.3 Transmission Power Allocation

We assume that for the given channel type and channel SNR, the channel code rate is constant in the whole video sequence. This assumption is made because adjusting channel code rate at the packet level would require a reliable side channel, or incur considerable overhead. In order to make efficient resource allocation, we adaptively allocate transmission power at the packet level according to the relative importance of the packets. This adaptivity can be achieved without incurring additional overhead in

bit rates.

In this chapter, video quality is measured as the mean squared error (MSE) between the original pixel values and the reconstructed pixel values at the decoder averaged over all frames in the sequence, which is also called the end-to-end distortion. Let D^q denote the quantization-induced distortion, i.e., the overall MSE between the original pixel values and the quantized pixel values at the encoder, and let D^c denote the channel-induced distortion, i.e., the overall MSE between the quantized pixel values at the encoder and the reconstructed pixel values at the decoder. Under the assumption that the quantization-induced distortion and the channel-induced distortion are statistically independent, the end-to-end distortion at the decoder can be approximated by [73] [34]

$$D = D^q + D^c. \tag{6.2}$$

As discussed earlier, due to the effect of error propagation in inter-frame coding, the theoretical analysis of the channel-induced distortion is rather difficult, and its pixel-level estimation is computation-intensive. Therefore we consider a frame-level approximation of the analysis similar to that used in [73] and [42].

Let $d^c_{MB}(i,j)$ denote the MSE between the quantized pixel values at the encoder and the reconstructed pixel values at the decoder for the pixels of the j-th MB in the i-th frame when this MB is lost and error-concealed. Then the expected channel-induced distortion in the i-th frame that is caused by channel errors in the same frame can be given by

$$D_i^{c,0} = \frac{1}{N_{MB}} \sum_{j=1}^{N_{MB}} \rho_{MB}(i,j) d^c_{MB}(i,j) \tag{6.3}$$

where N_{MB} is the number of MBs per frame, and $\rho_{MB}(i,j)$ is the loss probability of the j-th MB in the i-th frame.

For an inter-coded video sequence, channel errors in a frame will propagate into

subsequent frames. Assume that $D_i^{c,l}$ denotes the expected distortion in the $(i+l)$-th frame (where $l > 0$) that is caused by channel errors in the i-th frame. Then define $\mu_i^{(l)}$ as follows:

$$\mu_i^{(l)} = \frac{D_i^{c,l}}{D_i^{c,l-1}},\ l = 1, 2, \cdots . \tag{6.4}$$

Let D_i^c denote the total distortion caused by channel errors in the i-th frame, including the expected error propagation effect on the subsequent frames. Define the error propagation gain as

$$\mu_i = \frac{D_i^c}{D_i^{c,0}}. \tag{6.5}$$

Assuming that channel errors do not propagate beyond T frames, μ_i is given by [42]

$$\mu_i = 1 + \sum_{m=1}^{T} \prod_{l=1}^{m} \mu_i^{(l)}. \tag{6.6}$$

At low channel error rates, the channel-induced distortion for the sequence D^c can be approximated as follows by averaging the distortion caused by channel errors in all frames of the sequence:

$$D^c = \frac{1}{N_F} \sum_{i=1}^{N_F} D_i^c \tag{6.7}$$

where N_F is the number of frames in the sequence.

The essence of a frame-level distortion analysis such as that in [73] and [42] is to assume that $\mu_i^{(l)}$ is determined by the statistics at the frame-level (instead of that at the pixel-level), such as the ratio of the number of inter-coded blocks in a frame over the total blocks per frame. For the intra-MB refreshing scheme employed in this work, the number of inter-coded blocks in a frame is about the same for each frame. Therefore, the error propagation gain μ_i can be assumed to be a constant for all frames. This assumption allows us to consider each individual frame and to look for a near-optimal power allocation on a frame-by-frame basis.

Next we formulate the optimization problem in the context of our packetization scheme. For simplicity, we assume that the effect of the case when two neighboring packets are lost and affect the same MB at their boundary can be ignored. Furthermore, assume that the error concealment strategy described in Section 6.2.2 is used, and that the three nearest MVs in the GOB above the MB in consideration are correctly decoded. These approximations are reasonably good when the average packet loss probability is low; they allow independent optimization for each packet and thus greatly simplify the optimization problem. Under these assumptions, (6.3) can be approximated by

$$D_i^{c,0} \approx \frac{1}{N_{MB}} \sum_{k=1}^{N_i} \rho(i,k) d^c(i,k). \tag{6.8}$$

where N_i is the number of packets in the i-th frame, $\rho(i,k)$ is the loss probability of the k-th packet in the i-th frame, and $d^c(i,k)$ is the sum of $d_{MB}^c(i,j)$ for all MBs in the i-th frame that are affected by the loss of the k-th packet in the same frame.

Let $P_t(i,k)$ denote the transmission power for the k-th packet in the i-th frame. Define the power profile for the i-th frame

$$\mathbf{P_i} = \{P_t(i,1), P_t(i,2), \cdots, P_t(i,N_i)\}.$$

Note that for a given channel code rate r_c, $\rho(i,k)$ is a decreasing function of $P_t(i,k)$. However, this function also depends on what channel information is available at the transmitter. Under one scenario, the transmitter knows only the average channel SNR but not the CSI, therefore no closed-loop power control can be conducted to compensate for small-scale fading. Under another scenario, an estimate of the CSI (e.g., the estimate of the average fading power $|\hat{h}(i,k)|^2$ for the k-th packet in the i-th frame, which may be obtained from the delayed value for the previous packet) is available at the transmitter

through feedback, so closed-loop power control can be conducted to compensate for small-scale fading, and ρ is actually a function of $(|\hat{h}|^2 \cdot P_t)$. Under both scenarios, either the function ρ of P_t or that of $(|\hat{h}|^2 \cdot P_t)$ can be estimated via simulations.

Now we are ready to describe our power allocation algorithms.

6.3.1 Proposed Power Allocation Algorithm 1

We propose a power allocation algorithm that allocates the transmission power in a frame so that D_i^c is minimized while satisfying the total average power constraint in the frame. Considering (6.5) and (6.8) and the assumption that μ_i is a constant for all frames, the objective becomes

$$\underset{\mathbf{P_i}}{\text{minimize}} \qquad \frac{1}{N_{MB}} \sum_{k=1}^{N_i} \rho(i,k) d^c(i,k) \tag{6.9}$$

$$\text{subject to} \qquad \frac{1}{N_i} \sum_{k=1}^{N_i} P_t(i,k) \leq P_{TH} \tag{6.10}$$

where P_{TH} is the constraint for the average power in the sequence. As mentioned earlier, $\rho(i,k)$ is a function of $P_t(i,k)$.

This problem can be converted to an unconstrained optimization problem using the Lagrangian relaxation method. Given a Lagrangian multiplier λ, we have the following cost function:

$$J_i = \sum_{k=1}^{N_i} \{\rho(i,k) d^c(i,k) + \lambda P_t(i,k)\}. \tag{6.11}$$

Minimizing this cost function can be achieved by minimizing each of its components

and summing them up:

$$\begin{aligned} \min_{\mathbf{P_i}} \quad & \{\sum_{k=1}^{N_i}[\rho(i,k)d^c(i,k) + \lambda P_t(i,k)]\} \\ = \quad & \sum_{k=1}^{N_i} \min_{P_t(i,k)} \{\rho(i,k)d^c(i,k) + \lambda P_t(i,k)\}. \end{aligned} \tag{6.12}$$

Depending on whether the CSI estimate is available at the transmitter, λ can be determined as follows.

- **Case 1: Only the average channel SNR is available**

  As discussed earlier, in this case, $\rho(i,k)$ is a function of $P_t(i,k)$. Assuming a delay of one video frame in encoding is tolerable, an optimal λ can be found based on all the source and channel information for the frame that is available at the transmitter. The procedure is as follows: For a given λ, the power profile $\mathbf{P_i}$ can be determined from (6.12). By varying λ and finding $\mathbf{P_i}$ that satisfies the average power constraint (6.10), the problem of (6.9) and (6.10) can be solved numerically. Further discussions on the Lagrangian method can be found in [57].

- **Case 2: The CSI estimate is available**

  Assume that $|\hat{h}(i,k)|^2$ is available at the transmitter, then as discussed earlier, $\rho(i,k)$ becomes a function of $|\hat{h}(i,k)|^2 \cdot P_t(i,k)$. Unlike in the previous case, where the solution to the optimization problem does not depend on any instantaneous channel information so that an optimal λ can be chosen to meet the constraint of (6.10) for the $i-$th frame, here a sub-optimal λ needs to be chosen for each packet based on information available up to that packet. A simple algorithm is used to update λ at each packet in the $i-$ frame as follows:

$$\lambda_{i,k+1} = \lambda_{i,k} \left(1 + \alpha \frac{\sum_{j=1}^{k} (P_t(i,j) - P_{TH})}{k P_{TH}} \right) \tag{6.13}$$

where α is a small positive number (e.g. 0.02 for the simulations described in Section 6.4 at 70 mph) which should be proportional to the mobile speed.

6.3.2 Proposed Power Allocation Algorithm 2

In the previous algorithm, the average transmission power is constant in all frames, regardless of the error sensitivity of each frame. Now we consider an algorithm that allows the average transmission power to vary from frame to frame, and that allocates more power to the more error-sensitive frames where high motion or frequent scene change occurs.

We propose a power allocation algorithm that minimizes the average transmission power in a video frame so that D_i^c is no greater than a given threshold. Considering (6.5) and (6.8) and the assumption that μ_i is constant for all frames, the objective becomes

$$\underset{\mathbf{P_i}}{\text{minimize}} \qquad \frac{1}{N_i} \sum_{k=1}^{N_i} P_t(i,k) \tag{6.14}$$

$$\text{subject to} \qquad \frac{1}{N_{MB}} \sum_{k=1}^{N_i} \rho(i,k) d^c(i,k) \leq D_{TH} \tag{6.15}$$

where D_{TH} is the threshold for the distortion induced by channel errors in each frame.

Note that there exists certain symmetry between {(6.9), (6.10)} and {(6.14), (6.15)}. Exchanging $P_t(i,k)/N_i$ and $\rho(i,k)d^c(i,k)/N_{MB}$ in the first set of equations leads to the second set of equations, except for a difference of constants between P_{TH} and D_{TH}. As a result, this problem of (6.14) and (6.15) under the two scenarios of the CSI availability can also be solved numerically in a similar way to that in Section 6.3.1.

The algorithm's details are omitted here due to space limit.

In order to compare this algorithm with the previous algorithm, D_{TH} is selected for the sequence such that $\frac{1}{N_F N_i}\sum_{i=1}^{N_F}\sum_{k=1}^{N_i} P_t(i,k) = P_{TH}$.

In order to compare the above joint coding and power allocation algorithms under the second scenario of the CSI availability with conventional schemes where power control and source/channel coding are separately designed, we describe a conventional power control algorithm here. In this algorithm, the transmitter tries to adjust the transmission power in order to compensate for channel fading using the CSI estimate at each packet, while maintaining a fixed average transmission power. That is, the transmission power for the $k-$th packet in the $i-$th frame is specified according to

$$P_t(i,k) = \frac{P_{r0}(i)}{|\hat{h}(i,k)|^2} \tag{6.16}$$

where $P_{r0}(i)$ is the target received power for the frame. $P_{r0}(\cdot)$ can be adjusted at the end of every frame by a fixed stepsize (e.g. ± 0.02 dB) to satisfy the average power constraint P_{TH}.

6.3.3 Analysis of the Power-Distortion Tradeoff

Assume that the CSI estimate is available at the transmitter. Furthermore, assume that

$$\rho(i,k) = b[P_t(i,k)|\hat{h}(i,k)|^2]^{-a} \tag{6.17}$$

where a and b are positive constants to be determined through simulations and curve-fitting, and functions of channel code rate r_c, mobile speed, and CSI availability.

Consider the optimization of the power-distortion tradeoff for the whole se-

quence under the average power constraint:

$$J = \sum_{i=1}^{N_F}\sum_{k=1}^{N_i}[\mu\rho(i,k)d^c(i,k) + \lambda P_t(i,k)] \tag{6.18}$$

$$\text{subject to} \qquad \frac{1}{\sum_{i=1}^{N_F} N_i}\sum_{i=1}^{N_F}\sum_{k=1}^{N_i} P_t(i,k) = P_{TH}. \tag{6.19}$$

Letting $\frac{\partial J}{\partial P_t(i,k)} = 0$ yields

$$P_t(i,k) = [\frac{\mu ab|\hat{h}(i,k)|^{-2a}d^c(i,k)}{\lambda}]^{1/(a+1)}. \tag{6.20}$$

Substituting (6.20) into (6.19), we have

$$\lambda^{-1/(a+1)} = \frac{P_{TH}}{\frac{1}{\sum_{i=1}^{N_F} N_i}\sum_{i=1}^{N_F}\sum_{k=1}^{N_i}[\mu ab|\hat{h}(i,k)|^{-2a}d^c(i,k)]^{1/(a+1)}}. \tag{6.21}$$

Combining (6.21) and (6.20) yields

$$P_t(i,k) = \frac{[|\hat{h}(i,k)|^{-2a}d^c(i,k)]^{1/(a+1)}}{\frac{1}{\sum_{i=1}^{N_F} N_i}\sum_{i=1}^{N_F}\sum_{k=1}^{N_i}[|\hat{h}(i,k)|^{-2a}d^c(i,k)]^{1/(a+1)}} \cdot P_{TH}. \tag{6.22}$$

It can be easily verified that (6.22) satisfies the average power constraint (6.19).

If the CSI is not available at the transmitter, we may set $|\hat{h}(i,k)|$ to 1 in (6.17). Note that in this case the constants a and b may have different values from those in the case of CSI being available at the transmitter. Then we have

$$P_t(i,k) = \frac{d^c(i,k)^{1/(a+1)}}{\frac{1}{\sum_{i=1}^{N_F} N_i}\sum_{i=1}^{N_F}\sum_{k=1}^{N_i} d^c(i,k)^{1/(a+1)}} \cdot P_{TH}. \tag{6.23}$$

If only the CSI but no source error sensitivity information $d^c(i,k)$ is available at

(a) *Mother and Daughter*: Frame 15 (b) *News*: Frame 10

Figure 6.3: Video Sequences Used in Simulations

the transmitter, we may set $d^c(i,k) = 1$ in (6.22) and obtain

$$P_t(i,k) = \frac{|\hat{h}(i,k)|^{-2a/(a+1)}}{\frac{1}{\sum_{i=1}^{N_F} N_i} \sum_{i=1}^{N_F} \sum_{k=1}^{N_i} |\hat{h}(i,k)|^{-2a/(a+1)}} \cdot P_{TH} \tag{6.24}$$

which can be used to evaluate the performance of power allocation based on the CSI alone.

6.4 Experimental Results and Discussions

We use two well-known QCIF video sequences in the simulations: one low motion sequence *Mother and Daughter*, and one high motion sequence *News*. A frame of each sequence is shown in Figure 6.3. Both sequences are coded at 10 fps for 100 frames, and the data rate is 200 Kbits/sec. The length of a channel packet N_c is chosen to be about 400 bits, and it may vary slightly corresponding to different RCPC code rates.

As pointed in [73], intra coding and FEC can be exchanged to some extent with-

out causing significant performance loss. This is especially true when the FEC rate can be optimally selected for a given intra coding rate. For simplicity, we use a fixed refreshing rate $\beta = \frac{1}{20}$ in the simulations, unless noted otherwise. Each frame is encoded with a fixed quantization step size, which is adapted frame by frame to meet a given target bit budget.

In the simulations, a normalized flat Rayleigh fading process is generated using the Jakes' Model. Assuming $f_c = 900$ MHz, $v = 70$ mph, and the data rate $B = 200$ Kbits/sec, the normalized Doppler value $f_D/B = 4.67 \times 10^{-4}$. BPSK modulation is employed. Considering the use of two transmit antennas with each subject to independent fading, the SNR is defined as $\gamma = \frac{2E_s}{2\sigma_n^2} = \frac{E_s}{\sigma_n^2}$, where E_s is the energy of the transmitted PSK signal averaged over the video sequence, and σ_n^2 is the variance of the Gaussian noise along each dimension. This parameter can be used to indicate the level of the average transmission power.

The Y-component $PSNR$ value is given by

$$PSNR_Y = 10 \log_{10} \frac{255^2}{D_Y} \tag{6.25}$$

where D_Y is the mean squared difference of the luminance values between the original frame (or sequence) and the corresponding reconstructed frame (or sequence). Unless specified otherwise, the simulation results are averaged over 50 different channel realizations.

From Figure 6.9 to Figure 6.12, it is assumed that the CSI estimate for the current packet $|\hat{h}(i,k)|^2$ is available at the transmitter, so closed-loop power control is employed to combat Rayleigh fading. For all other figures, it is assumed that only the average SNR is known at the transmitter, and the CSI estimate is available at the receiver but not at the transmitter, so no closed-loop power control is conducted to combat Rayleigh

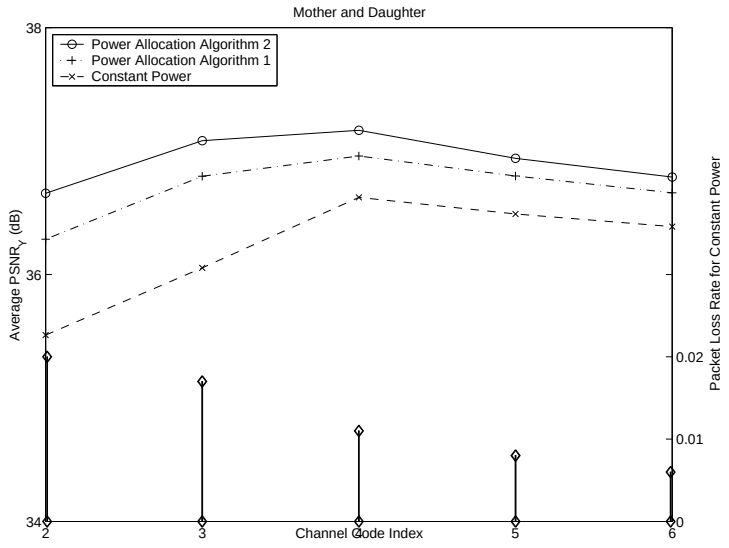


Figure 6.4: Performance Results vs. Channel Code Index

fading. Under both scenarios, the probability of packet loss for a given channel code rate is determined through simulations and modeled through curve-fitting.

In Figure 6.4, the performance results for several channel code rates at the same SNR ($\gamma = 15$ dB) are shown for *Mother and Daughter*. We can see that the best results occur when the channel code rate is chosen such that the average packet loss rate in the sequence $P_e \approx 1\%$ for the constant power algorithm. A too low P_e may consume too much rate for channel coding and also leave power allocation little room for improvement; a too high P_e may cause large channel-induced distortion, and in this case power allocation given certain average power constraints will not be of much help either. As confirmed by simulations to be usually a good choice, and for simplicity, we will choose the channel code rate that gives $P_e \approx 1\%$ for the constant power algorithm, unless noted otherwise. Then the source coding parameters are determined accordingly by a rate-control algorithm to meet the bit budget.

In Figure 6.5 and Figure 6.6, the PSNR results vs. the frame indices are com-

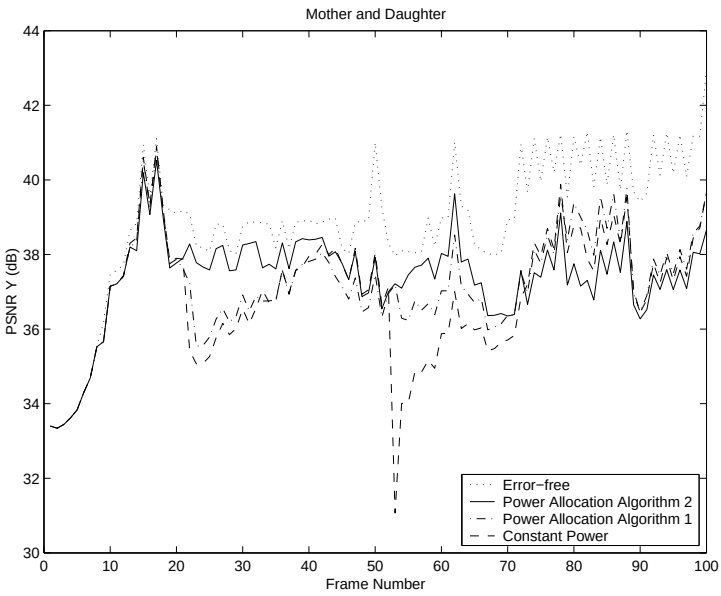


Figure 6.5: PSNR vs. Frame Number for *Mother and Daughter*

pared between the constant power algorithm and proposed power allocation algorithms at $\gamma = 15$ dB and $n_c = 4$ ($r_c = 1/2$) for a channel realization. It can be seen that the proposed algorithms greatly reduce the performance degradation caused by channel fading. In Figure 6.7 and Figure 6.8, the average PSNR results vs. the SNRs are shown. At the given SNRs, on average the first proposed power allocation algorithm is 0.37 dB and 0.68 dB better in PSNR than the constant power algorithm for *Mother and Daughter* and *News*, respectively; and for the second proposed power allocation algorithm, the improvement is 0.57 dB and 0.85 dB on average, respectively. More importantly, the proposed algorithms reduce the performance variation from frame to frame and enhance the subjective quality significantly. Also shown in the figures are the performance results of power allocation based on (6.23), which requires the knowledge of the video packet error sensitivity information for the whole video sequence and can serve as an upper bound (in terms of PSNR) to practical power allocation algorithms. On average this bound is 0.59 dB and 0.91 dB better than the constant power algorithm for *Mother and Daughter* and *News*, respectively.

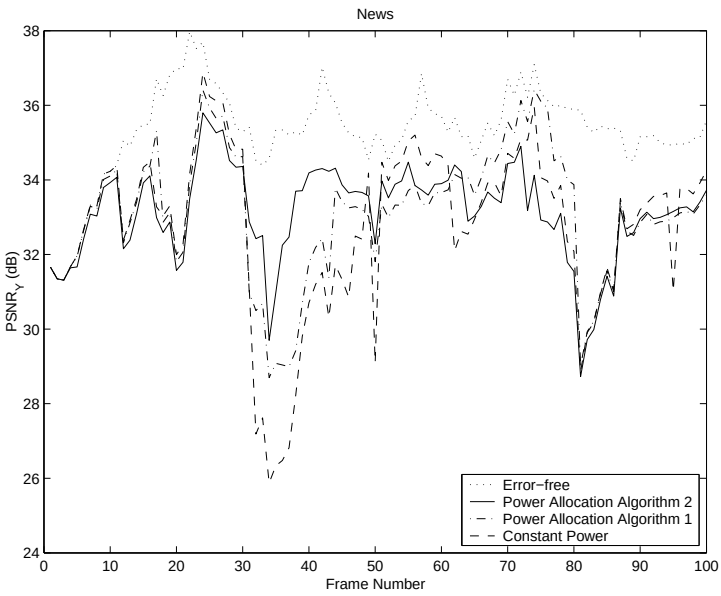


Figure 6.6: PSNR vs. Frame Number for *News*

In Figure 6.9 and Figure 6.10, the average PSNR results vs. the SNRs are shown where the proposed algorithms take advantage of the estimates of the average channel fading power available at the transmitter. For simplicity, an estimate of the average channel fading power for a packet is obtained by using the actual value of the average channel fading power for the previous packet. At the given SNRs, on average the first proposed power allocation algorithm is 0.54 dB and 1.16 dB better in PSNR than the conventional closed-loop power control algorithm for *Mother and Daughter* and *News*, respectively; and for the second proposed power allocation algorithm, the improvement is 0.72 dB and 1.25 dB on average, respectively. Also shown in the figures are the performance results of power allocation based on (6.22), which requires the knowledge of the source error sensitivity information and the CSI for the whole video sequence and can serve as an upper bound (in terms of PSNR) to practical power allocation algorithms. On average this bound is 0.76 dB and 1.38 dB better than the conventional power control algorithm for *Mother and Daughter* and *News*, respectively.

In Figure 6.11 and Figure 6.12, we assume that the CSI is available at the trans-

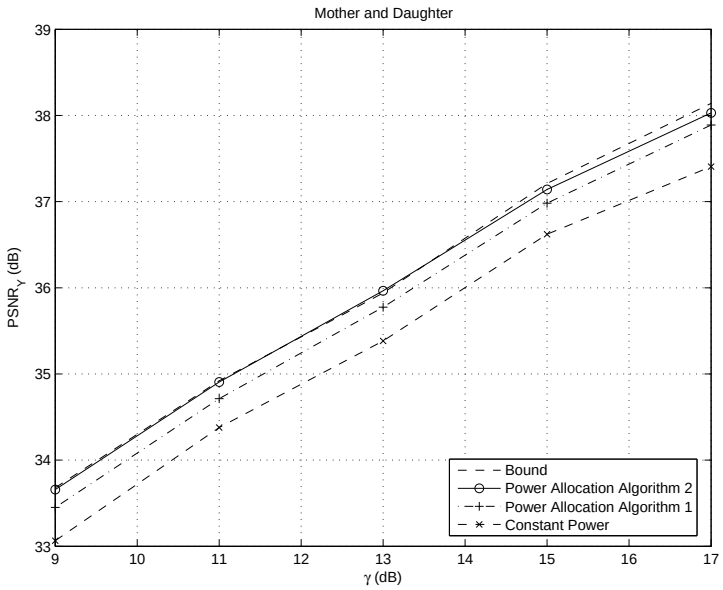


Figure 6.7: Performance Results for *Mother and Daughter*

mitter and the channel code rate is fixed with $r_c = 1/2$ for transmission of *Mother and Daughter* and *News*, respectively. As the average transmission power reduces, the channel error rate increases and the overall performance drops sharply. In addition to the "Bound" curves determined by (6.22), the bound curves determined by (6.24) and labelled as "Bound w/o Using $d^c(i,k)$", are also shown in the figures to indicate the amount of improvement obtainable using optimal power control alone without taking advantage of the source error sensitivity information. It can be seen that at high channel block error rate, properly allocating power using the CSI alone can improve the performance significantly over the conventional power control algorithm. This observation is similar to that in [48]. However, when the channel block error rate is low, allocating power using CSI alone does not help much, and most of the overall gains from allocating power can be attributed to the use of the source error sensitivity information.

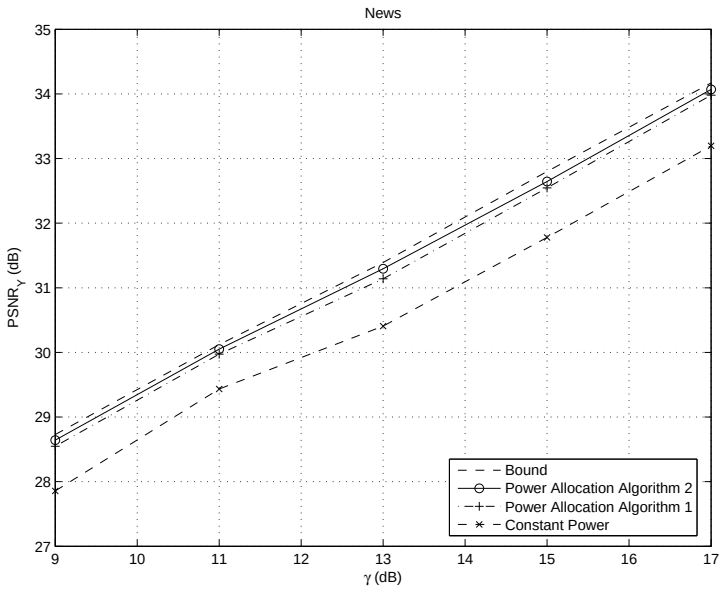


Figure 6.8: Performance Results for *News*

6.4.1 Impact of Individual Components of the Proposed Scheme

The proposed power allocation scheme consists of several essential components, such as adaptive channel coding, fixed-length packet structure, and power allocation on packet level, etc. We will study the impact of each individual component on the system performance by comparing it with alternative designs. We assume in the following experiments that only the average SNR but no CSI estimate is known at the transmitter.

In Figure 6.13, a fixed channel code rate scheme is considered, and its performance results are compared with those of the proposed scheme, where the channel code rate is selected based on the channel SNR. For each scheme, two curves for the constant power algorithm and the second proposed power allocation algorithm are included in the plot; the curve for the first proposed power allocation algorithm, which is not shown in the plot, would lie in between. It can be observed that when using a fixed $r_c = 1/2$, the performance deteriorates much faster than the adaptive r_c scheme as the SNR decreases. Also shown in the figure are the fixed-r_c performance curves for intra-refreshing

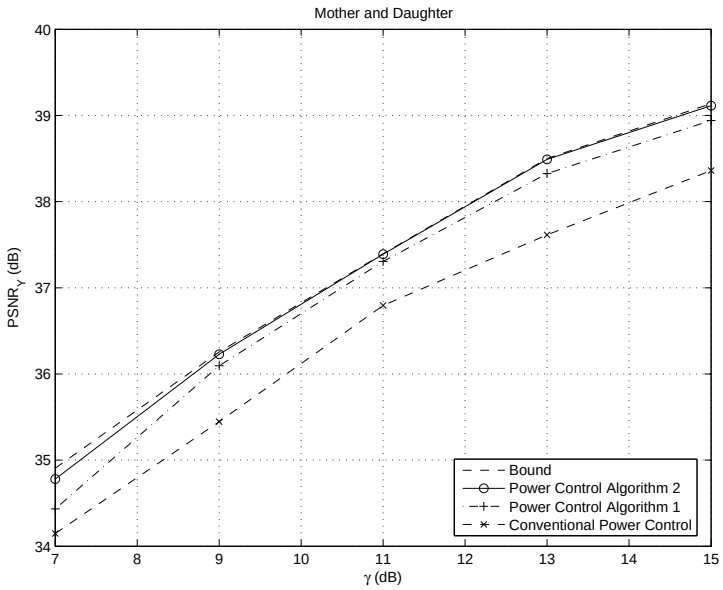


Figure 6.9: Performance Results for *Mother and Daughter* with CSI

rate $\beta = \frac{1}{10}$ instead of $\beta = \frac{1}{20}$. It can observed that a higher intra-refreshing rate would decrease the performance loss of the fixed r_c scheme at low SNRs. However, even under that case, power allocation cannot compensate for all of the performance loss due to the mismatch between the fixed channel code rate and the optimal one for the given channel SNR. This indicates that while transmission power allocation is useful, it is not a replacement for adapting channel coding to the channel SNR.

In the proposed scheme, a fixed-length packetization structure with a header as shown in Figure 6.2 is used to facilitate the decoder to recover from channel decoding errors. In order to evaluate the efficiency of this scheme, we consider an alternative scheme where a slice structure header is added to each coded MB (i.e. any MB that is not skipped) in the H.263 stream using Annex K [32] but no header for fixed-length packetization is used. This incurs additional overhead in the source coded stream but saves the overhead for the fixed-length packet header. In Figure 6.14, the results of the two schemes are compared for *Mother and Daughter*. When transmission power is allocated to packets of the same length, the performance of the proposed scheme is on

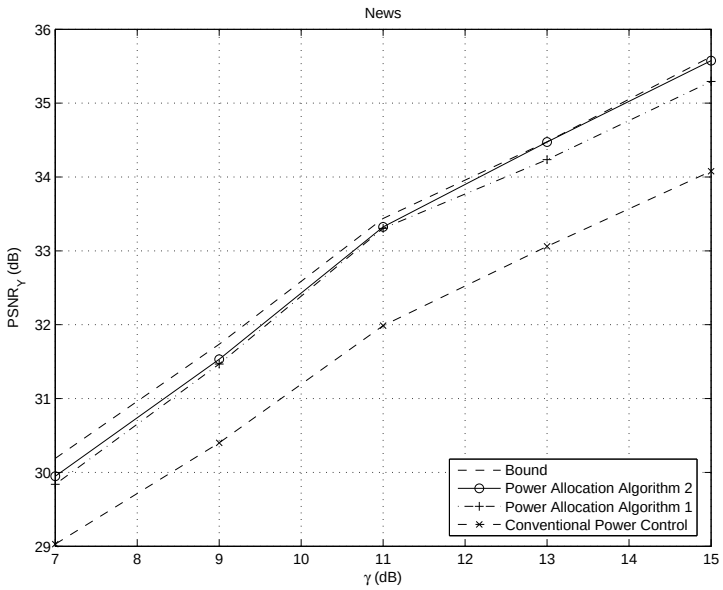


Figure 6.10: Performance Results for *News* with CSI

average 0.60 dB better than the scheme with the slice structure.

In the proposed scheme, transmission power is allocated to packets after FEC. For comparison purposes, we consider another scheme in which power is allocated on the MB level without FEC, since allocating power on the MB level can be readily achieved when FEC is not used. In order for the decoder to recover from channel errors, a slice structure header is inserted to every coded MB. Since the length of MBs is variable, we take it into account when calculating the average power in a frame for power allocation. Comparison results for the two schemes are shown in Figure 6.15 for *Mother and Daughter*. Even though the 1.38 dB performance improvement obtained from power allocation for the MB scheme is greater than the 0.49 dB improvement obtained from power allocation for the fixed-length packetization scheme with FEC, the absolute performance of the latter is still much better than that of the former, especially when the channel SNR is low. This justifies the use of FEC in the power allocation schemes. Power allocation on the MB level with the use of FEC may require side-channel information on the MB lengths and is not considered here.

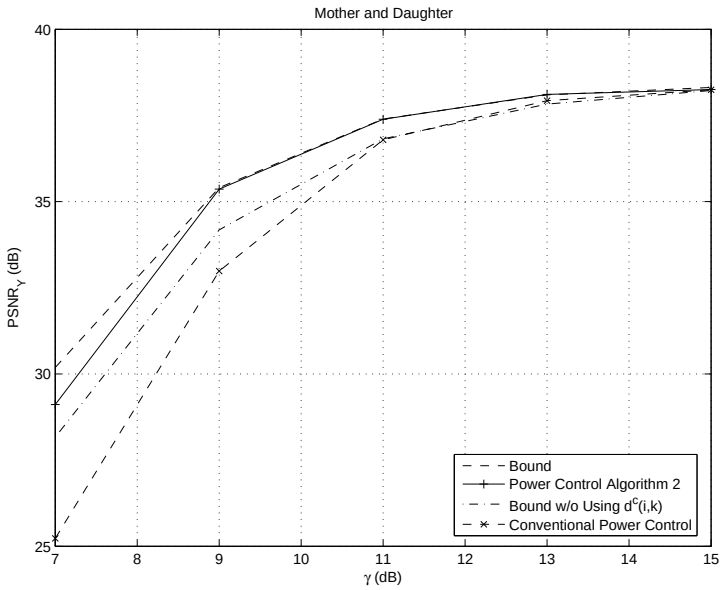


Figure 6.11: Performance Results for *Mother and Daughter* with CSI and Fixed r_c

Lastly, we consider allocating power on the frame level, in which case the transmission power is constant for each video frame but can vary from frame to frame. The frame-level allocation is performed by choosing the minimum transmission power for the $i-$th frame subject to the distortion constraint (6.15) for the frame. In Figure 6.16, the frame-level allocation results are compared with those of the packet-level allocation using the second proposed algorithm for *Mother and Daughter*. While the average improvement is 0.57 dB for the packet-level power allocation algorithm, it is 0.30 dB for the frame-level algorithm. This shows that packet-level power allocation can provide better performance than frame-level power allocation.

6.5 Conclusion

In this chapter, we have investigated the transmission power allocation approach for wireless video transmission under more realistic assumptions than those used in previous work. In particular, we consider a fixed-length packetization scheme and select a

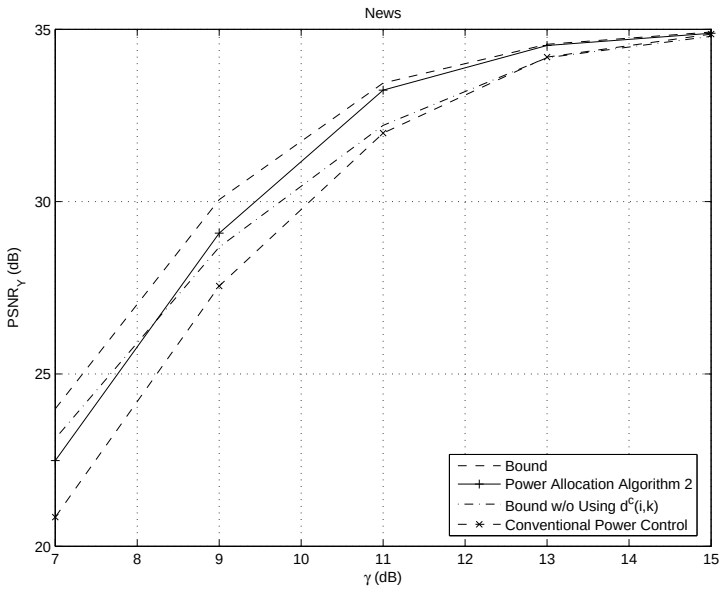


Figure 6.12: Performance Results for *News* with CSI and Fixed r_c

RCPC code rate to match the channel SNR reasonably well. Under these assumptions and adopting a frame-level analysis approach, we propose simple transmission power allocation algorithms to take advantage of the different error-sensitivity of different video segments, with or without the CSI available to the transmitter. In addition to numerical algorithms, closed form expressions are also presented to characterize the roles of the source error sensitivity and the CSI in optimally allocating transmission power. Extensive simulation results demonstrate the performance improvement of the proposed schemes over the conventional schemes. Experiments are also conducted to study the impact of the various components of the proposed scheme.

This chpater, in part, is a reprint of the material as it appears in: X. Tian. Efficient transmission power allocation for wireless video communications. *Proc. of IEEE Wireless Communications and Networking Conference, Atlanta*, 2004, Vol. 4, pp. 2058–2063, March 2004; and in part, has been submitted for publication as: X. Tian. Efficient

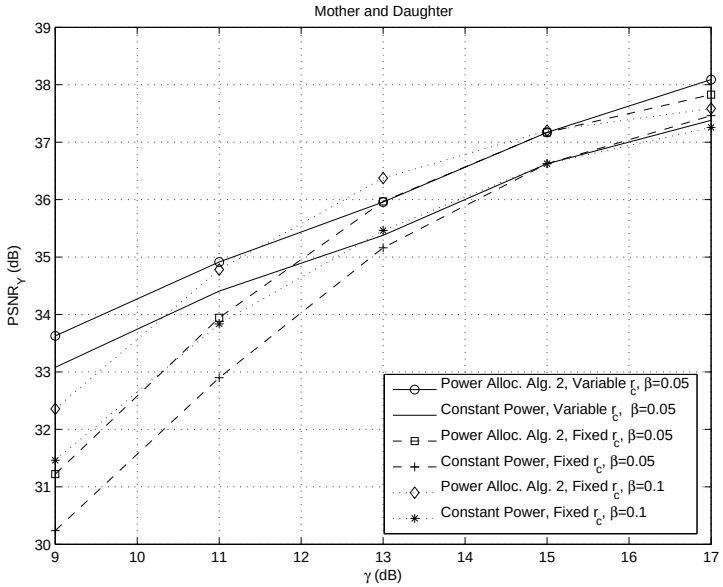


Figure 6.13: Performance Comparison of Schemes with Different Channel Code Rates and Intra-Refreshing Rates

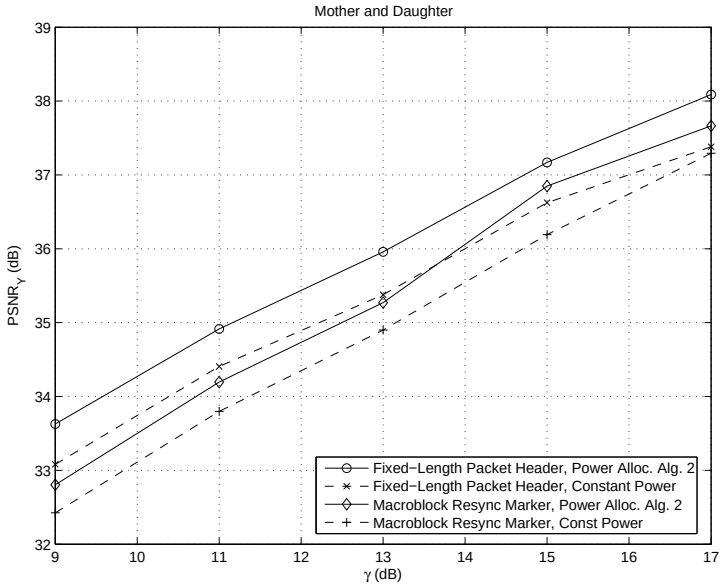


Figure 6.14: Performance Comparison of Schemes with and without Fixed-Length Packetization

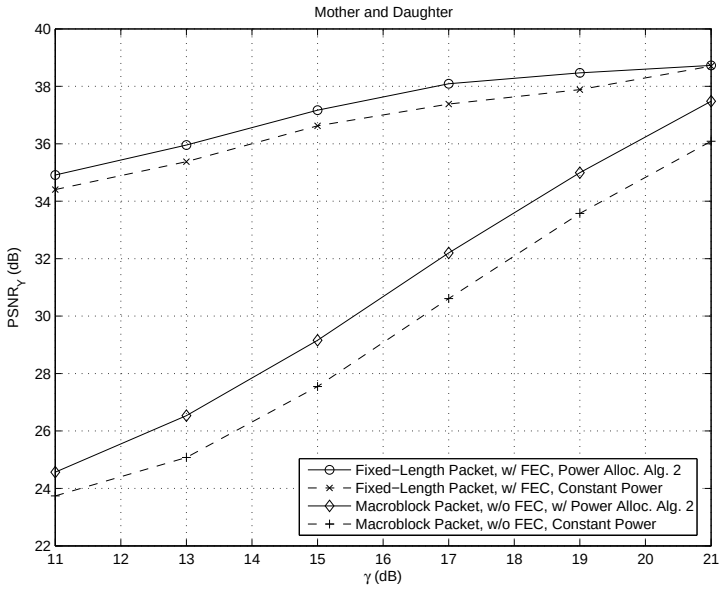


Figure 6.15: Performance Comparison of Packet-Level and Macroblock-Level Power Allocation

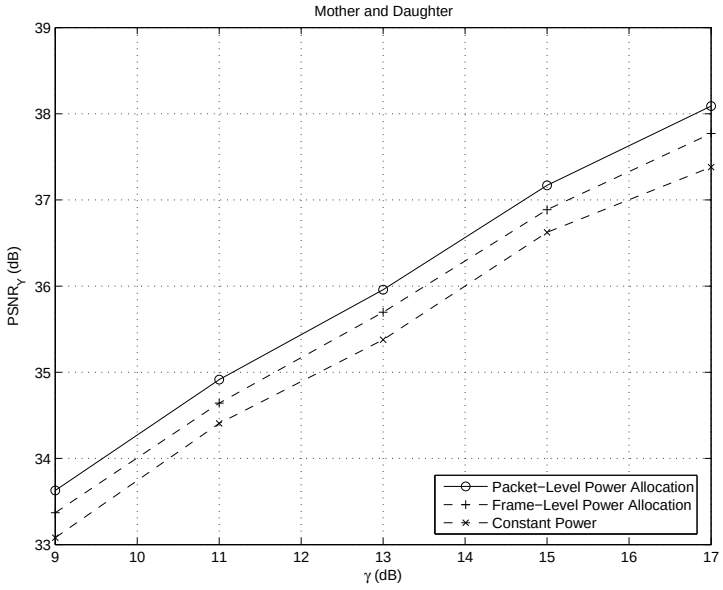


Figure 6.16: Performance Comparison of Frame-Level and Packet-Level Power Allocation

transmission power allocation for wireless video communications under average power constraints. *IEEE Transactions on Circuits and Systems for Video Technology*, 2004. The dissertation author was the primary investigator and single author for these two papers.

Chapter 7

Future Work

In this chapter, we briefly discuss some possible future research projects in the area of source-channel coding, with applications to image and video transmission.

7.1 M-Channel Multiple Description Scalar Quantizer

A major advantage of MDSQ [75] is that it is independent of the correlation between signals and thus can be very robust when applied to time-varying sources or channels. For M-channel MDSQ, when M is small, this approach may have certain performance advantage over other MDC methods such as polyphase selective quantization [40] or FEC codes with UEP [54].

In Chapter 3, we conducted some preliminary study of the M-channel MDSQ by devising some hand-crafted index assignments for image coding applications. In future work, further investigation may be conducted for the M-channel MDSQ problem. The goal is to design $M-$channel index assignment schemes of good performance and low search complexity.

7.2 Joint Source and Channel Coding for Sources with Residual Redundancy and Channels with Memory

Residual redundancy at the source coder output refers to the bit-rate of the source coder in excess of the source entropy. It includes redundancy due to the non-uniform distribution of the source and redundancy due to the memory of the source. These redundancies can be exploited by a joint source-channel coder to improve the overall performance over noisy channels [78] [60] [67].

Turbo codes have been shown to be able to obtain high coding gains even at low SNRs [7]. They have been successfully applied to joint source-channel coding systems [37] [59] [4]. The relatively large interleaver in a turbo coding system, while crucial to its performance gains, may potentially reduce the effectiveness of a joint source-channel decoder. This type of phenomenon was observed in [60] for Markov channels with no explicit channel coding used, where it was shown that a coding scheme without interleavers may outperform the scheme with interleavers by a large margin. In [41], the turbo decoder was modified for a Gilbert-Elliot burst channel model to improve the bit error performance, but no source coding was considered.

As future work, the idea of joint source and turbo coding for channels with memory and sources with residual redundancy should be explored, with the results being applied to applications such as image transmission over wireless channels.

7.3 Transmission Power Allocation for Video Communications Under Peak Power Constraints

In Chapter 6, we investigated transmission power allocation schemes for video coding subject to average transmission power constraints. By varying transmission power over

video packets according to their different degrees of importance, the end-to-end distortion can be reduced.

When the CSI is known at the transmitter, power control is often used to reverse the impact of the channel fading. However, in practice, in addition to the average power constraints, there also exist peak transmission power constraints [47, 48]. When the peak power exceeds that dynamic range of the amplifier, the transmitted signal will be clipped, causing performance degradation. If the dynamic range of the amplifier has to be increased, it will directly translate into significantly more expensive devices, limiting widespread industrial applications.

As future work, source-channel coding strategies for video transmission should be investigated under both average and peak power constraints.

7.4 Adaptive Transmission Power and Modulation for Wireless Video Communications

In Chapter 6, we investigated the transmission power allocation approach for video coding for a fixed modulation scheme. As shown in [28], a variable-rate and variable-power modulation scheme over fading channels may exhibit a large ($5 \sim 10$ dB) power savings relative to variable-power fixed-rate transmission. This motivates us to extend this adaptive power and rate concept to video coding in future work, and examine how large the improvement the joint source-channel coding approach may bring.

Bibliography

[1] S.M. Alamouti. A simple transmit diveristy technique for wireless communications. *IEEE Trans. on CSVT*, 16:1451–1458, October 1998.

[2] M. Alasti, K. Sayrafian-Pour, A. Ephremides, and N. Farvardin. Multiple description coding in networks with congestion problem. *IEEE Transactions on Information Theory*, pages 891–902, March 2001.

[3] S. Appadwedula, D.L. Jones, K. Ramchandran, and I. Konzintsev. Joint source channel matching for a wireless communication link. In *Proc. ICC-98*, pages 482–486, 1998.

[4] B.A. Banister, B. Belzer, and T.R. Fischer. Robust image transmission using jpeg2000 and turbo-codes. *IEEE Signal Processing Letters*, 9:117–119, April 2002.

[5] B. Belzer, J.D. Villasenor, and B. Girod. Joint source channel coding of images with trellis coded quantization and convolutional codes. In *Proc. of International Conference on Image Processing*, pages 85–88, 1995.

[6] S. Bennedetto, R. Garello, and G. Montorsi. A search for good convolutional codes to be used in the construction of turbo codes. *IEEE Transactions on Communications*, 46(9):1101–1105, September 1998.

[7] C. Berrou and A. Glavieux. Near optimum error correcting coding and decoding: Turbo-codes. *IEEE Transactions on Communications*, 44:1261–1271, October 1996.

[8] D.P. Bertsekas. *Dynamic Programming: Deterministic and Stochastic Models*. Prentice Hall, 1987.

[9] M. Bystrom and T. Stockhammer. Depedent source and channel rate allocation for video transmission. *IEEE Transactions on Wireless Communications*, 3:258–268, January 2004.

[10] Y.S. Chan and J.W. Modestino. A joint source coding-power control approach combined with adaptive channel coding for video transmission over CDMA cellular networks. In *Proc. of IEEE 58th Vehicular Technology Conference*, volume 5, pages 3415–3419, 2003.

[11] V. Chande and N. Farvardin. Joint source-channel coding for progressive transmission of embedded source coders. In *Proc. DCC '99*, 1999.

[12] S.S. Channappayya, G.P. Abousleman, and L.J. Karam. Joint source-channel coding of images using punctured convolutional codes and trellis-coded quantization. In *Proc. of IEEE International Symposium on Circuits and Systems*, pages 133–136, 2001.

[13] D.-M. Chuang and Y. Wang. Multiple description image coding using signal decomposition and reconstruction based on lapped orthogonal transforms. *IEEE Transactions on Circuits and Systems for Video Technology*, 9(6):895–908, September 1999.

[14] G. Cote, B. Erol, M. Gallant, and F. Kossentini. H.263+: video coding at low bit rates. *IEEE Trans. on CSVT*, 8:849–866, November 1998.

[15] H. Coward, R. Knopp, and S. D. Servetto. On the performance of multiple description codes over bit error channels. In *Proc. of the IEEE International Symposium on Information Theory*, page 240, 2001.

[16] P. Dent, G.E. Bottomley, and T. Croft. Jakes' model revisted. *Electronics Letters*, 29(1162-1163), June 1993.

[17] D. Divsalar and F. Pollara. Turbo codes for PCS applications. In *Proc. ICC-95*, pages 54–59, 1995.

[18] C. Dubuc, D. Boudreau, and F. Patenaude. The design and simulated performance of a mobile video telephony application for satellite third-generation wireless systems. *IEEE Transactions on Multimedia*, 3:424–431, December 2001.

[19] M. Effros. Robustness to channel variation in source coding for transmission across noisy channels. In *Proc. of International Conference on Acoustics, Speech, and Signal Processing*, pages 2961–2964, 1997.

[20] Y. Eisenberg, C.E Juna, T.N. Pappas, and A.K. Katsaggelos. Joint source coding and transmission power management for energy efficient wireless video communictions. *IEEE Trans. on CSVT*, 12:411–424, June 2002.

[21] N. Farvardin and V. Vaishampayan. On the performance and complexity of channel-optimized vector quantization. *IEEE Transactions on Information Theory*, 37:155–160, January 1991.

[22] T.R. Fischer and M.W. Marcellin. Joint trellis coded quantization/modulation. *IEEE Transactions on Communications*, 39:172–176, February 1991.

[23] T.R. Fischer, M.W. Marcellin, and M. Wang. Trellis-coded vector quantization. *IEEE Transactions on Information Theory*, 37:1551–1566, November 1991.

[24] S. Gadkari and K. Rose. Vector quantization with transmission energy allocation for time-varying channels. *IEEE Transactions on Communications*, pages 149–157, January 1999.

[25] R. Gallager. *Information Theory and Reliable Communication*. New York: Wiley, 1968.

[26] Robert G. Gallager. *Information Theory and Reliable Communication*. John Wiley and Sons, Inc., 1968.

[27] A. Goldsmith and M. Effros. Joint design of fixed-rate source codes and multiresolution channel codes. *IEEE Transactions on Communications*, pages 1301–1312, October 1998.

[28] A.J. Goldsmith and S.G. Chua. Variable-rate variable-power MQAM for fading channels. *IEEE Transactions on Communications*, 46:1218–1230, October 1997.

[29] V.K. Goyal, J. Kovačević, and M. Vetterli. Quantized frame expansions as source-channel codes for erasure channels. In *Proceedings DCC '99. Data Compression Conference*, pages 326–335, 1999.

[30] V.K. Goyal, J. Kovačević, R. Arean, and M. Vetterli. Multiple description transform coding of images. In *Proc. ICIP 98*, 1998.

[31] A. Guyader, E. Fabrr, C. Guillemot, and M. Robert. Joint source-channel turbo decoding of entropy-coded sources. *IEEE Journal on Selected Areas in Communications*, 9:1680–1696, September 2001.

[32] ITU-T Recommendation H.263. Video coding for low bit rate communication. February 1998.

[33] J. Hagenauer. Rate compatible convolutional codes and their applications. *IEEE Transactions on Communications*, 36:389–400, April 1988.

[34] Z. He, J. Cai, and C.W. Chen. Joint source channel rate-distortion analysis for adaptive mode selection and rate control in wireless video coding. *IEEE Trans. on CSVT*, 12:511–523, June 2002.

[35] K-P. Ho and K.H. Chei. Soft-decoding combined trellis-coded quantization/modulation. In *Proc. of IEEE Global Telecommunications Conference*, pages 2451–2455, 1999.

[36] K-P. Ho and K.H. Chei. Optimal soft-decoding for combined trellis-coded quantization/modulation. *IEEE Transactions on Communications*, 48:901–904, June 2000.

[37] K.P. Ho. Soft-decoding vector quantizer using reliability information from turbo-codes. *IEEE Communications Letters*, 3:308–310, July 1999.

[38] B. Hochwald and K. Zeger. Tradeoff between source and channel coding. *IEEE Transactions on Information Theory*, pages 1412–1424, September 1997.

[39] H. Jafarkhani and N. Farvardin. Design of channel-optimized vector quantizers in the presence of channel mismatch. *IEEE Transactions on Communications*, pages 118–124, January 2000.

[40] W. Jiang and A. Ortega. Multiple description coding via polyphase transform and selective quantization. In *Proceedings of the SPIE*, number 3653, pages 998–1008, 1999.

[41] J.H. Kang, W.E. Stark, and A.O. Hero. Turbo codes for fading and burst channels. In *Proc. of IEEE Global Telecommunications Conference*, pages 40–45, 1998.

[42] I.M. Kim and H.M. Kim. A new resource allocation scheme based on a PSNR criterion for wireless video transmission to stationary receivers over Gaussian channels. *IEEE Transactions on Wireless Communications*, 1:393–401, June 2002.

[43] I.M. Kim and H.M. Kim. An optimum power management scheme for wireless video service in CDMA systems. *IEEE Transactions on Wireless Communications*, 2:81–91, January 2003.

[44] I.M. Kim and H.M. Kim. Transmit power optimization for video transmission over slowly-varying rayleigh-fading channels in CDMA systems. *IEEE Transactions on Wireless Communications*, 3:1411–1415, September 2004.

[45] L.P. Kondi, F. Ishtiqa, and A.K. Katsaggelos. Joint source-channel coding for motion-compensated dct-based snr scalable video. *IEEE Transactions on Image Processing*, 9:1042–1052, September 2002.

[46] C. Lan, T. Chu, K.R. Narayanan, and Z. Xiong. Scalable image and video transmission using irregular repeat-accumulate codes with fast algorithm for optimal unequal error protection. *IEEE Transactions on Communications*, 52:1092–1101, July 2004.

[47] J.S. Lee and L.E. Miller. Anaylysis of peak-to-average power ratio for IS-95 and third generation CDMA forward link waveforms. *IEEE Trans. on Vehicular Technology*, 50:1004–1013, July 2001.

[48] Y.H. Lee and Y. Bar-ness. Power adaptation for BPSK signaling with average and peak power constraints in Rayleigh fading channels. *IEEE Transactions on Communications*, 51:1871–1876, November 2003.

[49] J. Lu, A. Nosratinia, and B. Aazhang. Progressive source-channel coding of images over bursty error channels. In *Proc. ICIP 98*, pages 127–131, 1998.

[50] X. Lu, Y. Wang, and E. Erkip. Power efficient H.263 video transmission over wireless channels. In *Proc. of IEEE ICIP*, volume 1, pages 533–536, 2002.

[51] M.W. Marcellin and T.R. Fischer. Trellis coded quantization of memoryless and gauss-markov sources. *IEEE Transactions on Communications*, 38:82–93, January 1990.

[52] A.C. Miguel, A.E Mohr, and E.A. Riskin. SPIHT for generalized multiple description coding. In *Proc. ICIP 99*, 1999.

[53] A.E Mohr, E.A. Riskin, and R.E. Ladner. Generalized multiple description coding through unequal loss protection. *submitted to IEEE Transactions on Circuits and Systems for Video Technology*, 1999.

[54] A.E Mohr, E.A. Riskin, and R.E. Ladner. Graceful degradation over packet erasure channels through forward error correction. In *Proceedings DCC '99. Data Compression Conference*, pages 462–475, 1999.

[55] A.M. Monk and L.B. Milstein. Open-loop power control error in a land mobile satellite system. *IEEE Journal on Selected Areas in Communications*, 13:205–212, February 1995.

[56] A. Nosratinia, L. Lu, and B. Aazhang. Source-channel rate allocation for progressive transmission of images. *IEEE Transactions on Communications*, 51:186–196, February 2003.

[57] A. Ortega and K. Ramchandran. Rate-distortion methods for image and video compression: an overview. *IEEE Signal Processing Magazine*, pages 23–50, November 1998.

[58] L. Ozarow. On a source-coding problem with two channels and three receivers. *Bell Syst. Tech. J.*, pages 1909–1921, December 1980.

[59] Z. Peng, Y.-F. Huang, and D.J. Costello Jr. Turbo codes for image transmission-a joint channel and source decoding approach. *IEEE Journal on Selected Areas in Communications*, 18:868–879, June 2000.

[60] N. Phamdo, F. Alajaji, and N. Farvardin. Quantization of memoryless and gauss-markov sources over binary markov channels. *IEEE Transactions on Communications*, 45:668–675, June 1997.

[61] R. Puri and K. Ramchandran. Multiple description souce coding using forward error correction codes. In *Thirty-Third Asilomar Conference on Signals, Systems and Computers*, 1999.

[62] T.S. Rappaport. Wireless communications: principles and practice. *Prentice Hall PTR*, 1996.

[63] D.W. Redmill and N.G. Kingsbury. The EREC: an error-resilient technique for coding variable-length blocks of data. *IEEE Transactions on Image Processing*, pages 565–574, April 1996.

[64] A. Reibman, H. Jafarkhani, M. Orchard, and Y. Wang. Performance of multiple description coders on a real channel. In *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing*, pages 2415–2418, 1999.

[65] J. Rogers and P. Cosman. Wavelet zerotree image compression with packetization. *IEEE Signal Processing Letters*, 5:105–107, May 1998.

[66] A. Said and W. A. Pearlman. A new, fast, and efficient image codec based on set partitioning in hierarchical trees. *IEEE Trans. on CSVT*, 6(3):243–250, June 1996.

[67] K. Sayood, F. Liu, and J.D. Gibson. A constrained joint source/channel coder design. *IEEE Journal on Selected Areas in Communications*, 12:1584–1593, December 1994.

[68] S. Servetto, K. Ramchandran, V. Vaishampayan, and K. Nahrstedt. Multiple description wavelet based image coding. *IEEE Transactions on Image Processing*, 9:813-826, May 2000.

[69] P.G. Sherwood and K. Zeger. Error protection for progessive image transmission over memoryless and fading channels. *IEEE Transactions on Communications*, 46(12):1555–1559, December 1998.

[70] P.G. Sherwood and K. Zeger. Macroscopic multistage image compression for robust transmission over noisy channels. In *Proceedings VCIP '99. Visual Communication and Image Processing*, 1999.

[71] P.G. Sherwood and K. Zeger. Progressive image coding over noisy channels. *IEEE Signal Processing Letter*, 4:189–191, July 1997.

[72] M. Srinivasan. Iterative decoding of multiple descriptions. In *Proc. Data Compression Conference*, pages 463–472, 1999.

[73] K. Stuhlmuller, N. Farber, M. Link, and B. Girod. Analysis of video transmission over lossy channels. *IEEE Journal on Selected Areas in Communications*, 18:1012–1032, June 2000.

[74] G. Ungerböck. Channel coding with multilevel/phase signals. *IEEE Transactions on Information Theory*, 28:55–67, January 1982.

[75] V. Vaishampayan. Design of multiple description scalar quantizers. *IEEE Transactions on Information Theory*, pages 821–834, May 1993.

[76] V. Vaishampayan and J.-C. Batllo. Asymptotic analysis of multiple description quantizers. *IEEE Transactions on Information Theory*, pages 278–284, January 1998.

[77] V. Vaishampayan. Design of multiple description scalar quantizers. *IEEE Transactions on Information Theory*, pages 821–834, May 1993.

[78] R.E. van Dyck and D.J. Miller. Transport of wireless video using separate, concatenated, and joint source-channel coding. *Proceedings of the IEEE*, 87:1734–1350, October 1999.

[79] M. Wang and T.R. Fischer. Joint trellis coded quantization designed for noisy channels. *IEEE Transactions on Information Theory*, 40:1792–1802, December 1994.

[80] Y. Wang, M. T. Orchard, and A. Reibman. Multiple description image coding for noisy channels by pairing transform coefficients. In *Proc. IEEE Workshop on Multimedia Signal Processing*, pages 419–424, 1997.

[81] J. Wen and J.D. Villasenor. Reversible variable-length codes for robust images and video transmission. In *Proc. of 31st Asilomar Conference–Sig.,Sys., and Comp.*, pages 973–979, 1997.

[82] R.D. Wesel, X. Liu, and W. Shi. Punctured symbol puncturing of trellis codes. In *Proc. of 31st Asilomar Conference Signals, Systems and Computers*, pages 172–176, 1997.

[83] T. Woerz and R. Schweikert. Performance of punctured pragmatic codes. In *Proc. of IEEE Global Telecommunications Conference*, pages 664–669, 1995.

[84] J.K. Wolf and E. Zehavi. P^2 codes: pragmatic trellis codes utilizing punctured convolutional codes. *IEEE Communications Magazine*, 33:94–99, February 1995.

[85] H. Yousefi'zadeh, H. Jafarkhani, and M. Moshfeghi. Power optimization of wireless media systems with space-time block codes. *IEEE Transactions on Image Processing*, 13:873–884, July 2004.

[86] K. Zeger and A. Gersho. Pseudo-gray coding. *IEEE Transactions on Communications*, 38:2147–2158, December 1990.

[87] Q. Zhang, Z. Ji, W. Zhu, and Y.Q. Zhang. Power-minimized bit allocation for video communication over wireless channels. *IEEE Trans. on CSVT*, 12:398–410, June 2002.

[88] R. Zhang, S.L. Regunathan, and K. Rose. Video coding with optimal inter/intra-mode switching for packet loss resilience. *IEEE Trans. on CSVT*, 18:966–976, June 2000.