VIDEO TRANSMISSION OVER WIRELESS NETWORKS

A Dissertation

by

SHENGJIE ZHAO

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

May 2004

Major Subject: Electrical Engineering

VIDEO TRANSMISSION OVER WIRELESS NETWORKS

A Dissertation

by

SHENGJIE ZHAO

Submitted to Texas A&M University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Approved as to style and content by:

Zixiang Xiong (Co-Chair of Committee)	Xiaodong Wang (Co-Chair of Committee)
Andrew K. Chan (Member)	Jyh Charn Liu (Member)
Chanan Singh (Head of Department)	

May 2004

Major Subject: Electrical Engineering

ABSTRACT

Video Transmission over Wireless Networks. (May 2004)

Shengjie Zhao, B.S., University of Science and Technology of China;

M.S., China Aerospace Institutes

Co–Chairs of Advisory Committee: Dr. Zixiang Xiong
Dr. Xiaodong Wang

Compressed video bitstream transmissions over wireless networks are addressed in this work. We first consider error control and power allocation for transmitting wireless video over CDMA networks in conjunction with multiuser detection. We map a layered video bitstream to several CDMA fading channels and inject multiple source/parity layers into each of these channels at the transmitter. We formulate a combined optimization problem and give the optimal joint rate and power allocation for each linear minimum mean-square error (MMSE) multiuser detector in the uplink and two types of blind linear MMSE detectors, i.e., the direct-matrix-inversion (DMI) blind detector and the subspace blind detector, in the downlink. We then present a multiple-channel video transmission scheme in wireless CDMA networks over multipath fading channels. For a given budget on the available bandwidth and total transmit power, the transmitter determines the optimal power allocations and the optimal transmission rates among multiple CDMA channels, as well as the optimal product channel code rate allocation. We also make use of results on the large-system CDMA performance for various multiuser receivers in multipath fading channels. We employ a fast joint source-channel coding algorithm to obtain the optimal product channel code structure. Finally, we propose an end-to-end architecture for multi-layer progres-

sive video delivery over space-time differentially coded orthogonal frequency division multiplexing (STDC-OFDM) systems. We propose to use progressive joint source-channel coding to generate operational transmission distortion-power-rate (TD-PR) surfaces. By extending the rate-distortion function in source coding to the TD-PR surface in joint source-channel coding, our work can use the 'equal slope' argument to effectively solve the transmission rate allocation problem as well as the transmission power allocation problem for multi-layer video transmission.

It is demonstrated through simulations that as the wireless channel conditions change, these proposed schemes can scale the video streams and transport the scaled video streams to receivers with a smooth change of perceptual quality.

To my parents, my wife Huamei, and my son Brian

ACKNOWLEDGMENTS

I would like to thank my advisors, Dr. Zixiang Xiong and Dr. Xiaodong Wang of Columbia University, for their inspirational guidance and encouragement during my Ph.D. program. I would also like to thank my committee members, Dr. Andrew K. Chan and Dr. Jyh Charn Liu, for their valuable comments and time. Special thanks go to Dr. Edward Dougherty for his support for my first year of study.

I wish to thank Nortel Networks for providing me with financial support from 2000 until 2002, the Dwight Look College of Engineering for awarding me the Fred J. Benson Graduate Fellowship from 1999 until 2000, the Department of Electrical Engineering for providing me with a Teaching Assistant position from the summer of 2002 until the spring of 2003, and Dr. Jae Choong Han for providing me with a graduate assistant position from 2003 until 2004.

I want to take this opportunity to recognize all my colleagues in the Wireless Communication Lab and Multimedia Lab at Texas A&M University, especially Dr. Xiaobo Zhou of Harvard University, Jianping Hua, Yong Sun, Guosen Yue, and Xuechao Du for the constructive and insightful discussions.

I am very grateful to my parents, Baogui Zhao and Yaqin Jin, my twin brother Shengmin Zhao, and my sister Shengli Zhao, for their continuous support and tremendous care.

Finally, but most importantly, I am especially indebted to my wife, Huamei Tang, for her never-ending love, great understanding, patience, support and for boosting my morale in hard times during my study in United States, while she chose to stay in China to take care of our son Brian Nlong Zhao with my parents-in-law, who I am also indebted to.

TABLE OF CONTENTS

CHAPTER Page

LIST OF TABLES

LIST OF FIGURES

FIGURE Page

FIGURE Page

CHAPTER I

INTRODUCTION

A. Background

The emerging third generation (3G) and fourth generation (4G) wireless technologies [1] are making high-fidelity video over wireless channels a reality. In addition, Internet protocol (IP) based architecture for 3G wireless systems promises to provide next-generation wireless services such as voice, high-speed data, Internet access, audio and video streaming on an all IP network [2]. However, wireless video [3, 4, 5] has bandwidth, delay, and loss requirements, many existing mobile networks cannot provide a guaranteed quality of service because temporally high bit error rates are unavoidable during fading periods.

Video sequence requires huge amount of data to process whereas the network bandwidth and the capacity of storage media are limited, thus video compression plays a central role in modern multimedia communications. Although network bandwidth and digital devices' storage capacity are increasing rapidly, video compression continues to play an essential role due to the exponential growth in size of multimedia content. Today compressed video data represent the dominant source of internet traffic. During the past decade, wavelets-based transform coding compression techniques have made a big advance in the field of video compression [6, 7, 8, 9, 10, 11, 12, 13]. The evolving next-generation video compression standard certainly will be based on wavelets. Therefore the research on networked wavelets-based compressed video data becomes very important [3, 4, 5].

The journal model is *IEEE Transactions on Circuits and Systems for Video Technology*.

Signal fading due to multipath propagation is a dominant source of impairment in wireless communication systems, causing high bit error rate (BER) and thus packets loss in networks [14]. Transmission errors typically occur during short fade intervals. Since the transmission of the information bits is packet-based, a high BER results in high packet loss ratio and hence the use of error control techniques is necessary. The purpose of error control is to use the available transmission rate, as determined by the congestion control mechanism, to mitigate the effects of packet loss. Error control is generally achieved by transmitting some amount of redundant information to compensate for the loss of potentially important packets. Real-time services such as video delivery transmission require that packet be received within a bounded delay. Therefore, error control techniques, such as forward error correction (FEC), are often used for mobile channels [15, 16]. Unequal error protection (UEP) using rate-compatible codes was popularized by Hagenauer [17] and Albanese *et al.* [18]. It can be achieved by fixing the source block length K and varying the channel block length $N_{\max}$ across the different source layers. Using an iterative descent algorithm, optimal error control in the form of UEP for receiver-driven layered multicast was considered in [19].

Furthermore the joint source-channel coding approach [20] has been motivated primarily by many communications applications involving speech, image, and video transmission. The trend in the evolution of joint source-channel coding systems has been to use more accurate source and channel models, as well as more sophisticated state-of-the-art source and channel coding techniques. Given a fixed BER, one can design an optimal system which minimizes the distortion of transmitted video bitstream subject to a total transmission rate [19]. In heterogeneous packet networks such as the Internet and wireless WAN, switches in the network are typically unaware of the structure or content of the packets they process. They provide only

a simple first-in-first-out queuing/scheduling policy and indiscriminately discard incoming packets when the output queues are full or the time is out. On the other hand, standards such as MPEG-4 [21] fine-granularity scalable (FGS) [22] produce layered bitstreams which are characterized by a natural hierarchy of layers. Such standard techniques produce coded bitstreams with unequal importance. This facilitates the use of different types of digital devices with different computational, display and memory capabilities. However, layered coding also makes the bitstreams more susceptible to bit errors or lost packets. The recently proposed product channel code framework renders all transmitted packets equally important and matches well with networks that have no means of prioritizing packets [23].

Truly portable communications drive much of the development of wireless technology. The end users themselves require a lightweight and compact interface to the network in the form of a pocket-sized, battery-powered transceiver or terminal. As such, power control [24, 25] is a critical issue in the design of mobile systems and gives rise to important practical constraint. Power control issues include assigning transmit power levels to channels subject to acceptable signal quality, providing varying levels of service to different priority classes, and maintaining connections in the presence of user movements. Given a set of channels to be connected, one must assign the transmit power level to each of them [26]. A great deal of work has been done on power allocation. Algorithms have been developed and shown to minimize the number of channels required to accommodate every user [26] and to minimize the total transmitted power [24, 25]. These approaches are very suitable for handling voice traffic. For some multimedia applications, however, other objectives may be more appropriate. For instance, a scheme based on maximizing the minimum SNR with a constraint on the total transmit power was proposed in [27]. The objectives is to guarantee the minimum required voice quality and reserve the highest possible

system capacity to receivers.

Moreover the combination of power control and joint source-channel coding for wireless video has an increasing importance. Several joint source-channel coding schemes that take into account transmission power and bandwidth have been studied. Schemes for AWGN channels have been studied in [28], and extensions to Rayleigh channels were considered in [29] with an objective of optimizing joint source-modulation performance subject to an average transmission power constraint. A framework for jointly considering error resilience and concealment techniques at the source coding level, as well as transmission power management at the physical layer has been proposed in [30], with the goal of limiting the amount of distortion in received video sequence while minimizing the transmission energy. A multi-channel wireless video transmission framework in conjunction with multiuser detection techniques, which minimizes the expected distortion of transmitted video bitstream subject to a total transmitted power and a total transmission rate, has been proposed by us in [31]. In addition, we proposed in [32] an optimal resource allocation method for multi-layer wireless video transmission by using the large-system performance analysis results for various multiuser receivers in multipath fading channels.

Recent advances in computing technology, data compression, high-bandwidth storage devices, high-speed networks, and the third generation (3G) wireless technology have made it feasible to provide the delivery video over wireless channels. Future wireless communication systems promise to offer a variety of multimedia services which require reliable transmission at high data rates. In order to achieve such high data rates, transmission over OFDM channels is of great interest.

OFDM can largely eliminate the effects of inter-symbol interference for high-speed transmission in very dispersive environments, and it readily supports interference suppression and space-time coding for enhanced efficiency [33]. One impor-

tant problem in the design of OFDM system is to optimize the system transmission bandwidth and power level. Each sub-carrier in OFDM system has two variables: transmission power and bit rate. The bit rate is related to modulation rate which is defined as the number of bits per transmission, e.g., 4 for 16-PSK. Meanwhile space-time coding (STC) employs diversity techniques, which integrate multiple antenna with coding techniques to achieve higher capacity and reduce co-channel interference in multiple access [34, 35, 36]. Furthermore, space-time differential detection (DD)is an attractive technique in flat-fading environments since it is very robust and does not require carrier phase tracking [37, 38]. However, performance of the conventional DD method in flat-fading channels is limited by an irreducible error floor if the fading bandwidth is larger than zero. Decision-feedback differential detection (DF-DD) is found to be a very effective method to reduce such an error floor with very low computational complexity. Recently Liu *et al.* developed the DF-DD receiver for space-time coded OFDM systems [39]. Therefore we adopt the DF-DD receiver in our work of transporting video over space-time coded OFDM systems.

Video transmission with OFDM or the integration of STC with OFDM has been studied recently [40, 41, 42]. In [40] the authors discussed scalable video transmission with a precoded OFDM system which is different from conventional OFDM systems with antenna diversity, where there is only one transmitter and one receiver antenna. In [41] the authors proposed multi-layer video transmission over space-time coded OFDM systems, which used a fixed transmission rate per OFDM frame and a fixed power level per sub-carrier. In [42] the authors dealt with MPEG-4 video transmission over OFDM systems using adaptive bit loading techniques. However [40] and [41] only considered a transmission framework with a fixed transmission rate. Furthermore, neither of these three works considered the power allocation problem for multi-layer video transmission, which is critical to wireless multimedia data transmission.

Multimedia delivery over wireless networks [43, 44] is an important topic that requires high transmission reliability and stringent end-to-end delay. Because wireless links are usually error-prone, bandlimited and time-varying, error control schemes are necessary to obtain high transmission reliability. In general, forward error correction (FEC) codes have been used for delivery transmission because they maintain a constant throughput and a bounded delay. In delivery transmission applications, a good performance is desirable as soon as possible. Suppose that an acceptable reconstruction quality is already reached at a low intermediate rate, then the transceiver can stop the transmission at this early stage and thus save a lot of time. Sherwood *et al.* [45] therefore proposed a scheme of progressive image transmission, which minimized the average of the expected distortion over a set of intermediate transmission rates given a target transmission rate. They also presented a dynamic programming algorithm to obtain an optimal solution. However the algorithm itself is not appropriate for delivery applications because its time complexity is quadratic with respect to the target transmission rate. Recently Stankovic *et al.*[46] proposed a fast unequal error protection strategy that allowed efficient progressive transmissions of images. In contrast to [45], the algorithm of [46] can be implemented in real-time. Therefore we proposed to adopt this algorithm in our proposed system.

B. Dissertation Outline

The dissertation presents some approaches to providing robust and efficient transmission of video over personal communications networks employing wireless access such as CDMA networks with multiuser detection, CDMA networks with large-system performance results, and space-time differentially coded OFDM systems. Our purpose is to provide good video quality most of the time while limiting the degradation

incurred under heavy fading conditions.

In Chapter II, We consider error control and power allocation for transmitting wireless video over CDMA networks in conjunction with multiuser detection. We map a layered video bitstream to several CDMA fading channels and inject multiple source/parity layers into each of these channels at the transmitter. At the receiver, we employ linear minimum mean-square error (MMSE) multiuser detector in the uplink and two types of blind linear MMSE detectors, i.e., the direct-matrix-inversion (DMI) blind detector and the subspace blind detector, in the downlink, for demodulating the received data. For given constraints on the available bandwidth and transmit power, the transmitter determines the optimal power allocation among different CDMA fading channels and the optimal number of source and parity packets to send that offer the best video quality. We formulate a combined optimization problem and give the optimal joint rate and power allocation for each of these three receivers. Simulation results show a performance gain of up to 3.5 dB with joint optimization over with rate optimization only.

In Chapter III, We present a multiple-channel video transmission scheme in wireless CDMA networks over multipath fading channels. We map an embedded video bitstream, which is encoded into multiple independently decodable layers by 3D-ESCOT video coding technique, to multiple CDMA channels. One video source layer is transmitted over one CDMA channel. Each video source layer is protected by a product channel code structure. A product channel code is obtained by the combination of a row code based on rate-compatible punctured convolutional code(RCPC) with cyclic redundancy check(CRC) error detection, and a source-channel column code, i.e., systematic rate-compatible Reed-Solomon(RS) style erasure code. For a given budget on the available bandwidth and total transmit power, the transmitter determines the optimal power allocations and the optimal transmission rates among

multiple CDMA channels, as well as the optimal product channel code rate allocation, i.e. the optimal unequal Reed-Solomon code source/parity rate allocations and the optimal RCPC rate protection for each channel. In formulating such an optimization problem, we make use of results on the large-system CDMA performance for various multiuser receivers in multipath fading channels. The channel is modelled as the concatenation of wireless BER channel and a wireline packet erasure channel with a fixed packet loss probability. By solving the optimization problem we obtain the optimal power level allocation and the optimal transmission rate allocation over multiple CDMA channels. For each CDMA channel we also employ a fast joint source-channel coding algorithm to obtain the optimal product channel code structure. Simulation results show that the proposed framework allows the video quality to degrade gracefully as the fading worsens or the bandwidth decreases, and it offers improved video quality at the receiver.

In Chapter IV, We propose an end-to-end architecture for multi-layer progressive video delivery over space-time differentially coded orthogonal frequency division multiplexing (STDC-OFDM) systems. An input video sequence is compressed by 3D-ESCOT into a layered bitstream. We input the bitstream in parallel to multiple STDC-OFDM channels. Each layer is transmitted over one STDC-OFDM channel. Different video source layers are protected by different error protection schemes in order to achieve unequal error protection. In progressive transmission, the reconstruction quality is important not only at the target transmission rate but also at the intermediate rates. So the error protection strategy needs to optimize the average performance over the set of intermediate rates. We propose to use progressive joint source-channel coding to generate operational transmission distortion-rate (TD-R) functions and operational transmission distortion-power(TD-P) functions for multiple wireless channels before forming the operational transmission distortion-power-

rate (TD-PR) surfaces. Lagrange multipliers are then employed on the fly to obtain the optimal power allocation and optimal rate allocation among multiple channels, subject to constraints on the total transmission rate and the total power level. Progressive joint source-channel coding offers the scalability feature to handle bandwidth variations and changes in channel conditions. By extending the rate-distortion function in source coding to the TD-PR surface in joint source-channel coding, our work can use the 'equal slope' argument to effectively solve the transmission rate allocation problem as well as the transmission power allocation problem for multi-layer video transmission. Experiments show that our scheme achieves significant PSNR improvement over a non-optimal system with the same total power level and total transmission rate.

Finally, Chapter V contains the conclusions.

Our contributions, presented in our publications [31, 32, 47, 48, 49, 50, 51], are briefly described as follows.

- We propose multi-layer video robust transmission frameworks over
  - CDMA networks employing receivers with multi-user detection capability by using the 3-D SPIHT video coder.
  - CDMA networks employing receivers with large-system performance results by using the 3-D ESCOT video coder.
  - wideband wireless channels using space-time differentially coded OFDM systems.
- We formulate combined problems for optimal resource allocation, such as power levels and transmission bandwidth, for various wireless networks.
- We propose to use multiple product code structures to packetize the layered

video bit-stream.

- We propose to use progressive joint source-channel coding to generate operational transmission distortion-power-rate (TD-PR) surfaces for multi-layer video streams.
- We propose a new practical algorithm for video delivery applications over space-time coded differentially over OFDM systems.
- The proposed systems have better performance for wireless video transmission.
- Our systems provide better video quality most of time while limiting the degradation incurred heavy fading conditions.

CHAPTER II

JOINT ERROR CONTROL AND POWER ALLOCATION FOR VIDEO TRANSMISSION OVER CDMA NETWORKS WITH MULTIUSER DETECTION*

A. Introduction

In this chapter, we simultaneously address rate control and power allocation in an integrated framework, by using layered source and channel coding in conjunction with exact linear minimum mean-square error (MMSE) multiuser detector and two types of blind MMSE detectors: direct-matrix-inversion (DMI) detector and subspace detector. Layered source coding is achieved with three-dimensional set partitioning in hierarchical trees (3-D SPIHT) [6], while layered channel coding is accomplished using a systematic rate-compatible Reed-Solomon (RS) style erasure code [52]. Both source layers and parity packets are generated at the transmitter for easy adaptation to changing channel bandwidth and packet loss ratios, which are determined by power allocation. After estimating channel parameters such as the SNR or SINR, the transmitter assigns the optimal power level to each channel, computes the optimal allocation of the available transmission rate between the source and channel codes, and transmits the packet data for the optimal collection of source and channel packets. Conceptually our work can be viewed as an extension of [19], where the packet loss ratio (or power level) is the same for all channels. Of course, power control does not come into play in the Internet multicast scenario considered in [19].

Using exact linear MMSE multiuser detector and blind linear MMSE detectors, we can receive good video quality even at transmission rate as low as 25 kilobits per second (Kbps) in each channel. We essentially transmit *one* layered video bitstream using *multiple* wireless channels. This is done by mapping multiple source/parity layers to different channels according to the importance of these source layers in the bitstreams. Depending on the video source coder, the importance level can be assigned according to significance or resolution. For example, with MPEG-4 fine-granularity scalable (FGS) coding [22], one could assign the base layer and the enhancement layer bitstreams to different channels. In 3-D embedded subband coding with optimal truncation (3-D ESCOT) [7], one can layer the bitstream by resolution and hence map bitstream layers corresponding to different resolutions to different wireless channels. In our work, we transmit 2, 6, 17 and 25 source layers and their associated parity packets over four wireless channels, respectively, when a color QCIF (176×144) sequence is coded with 3-D SPIHT at 50 layers (one thousand bytes each) per group of frames (GOF), with the GOF size being 32.

In addressing a topic as broad as wireless video [3, 4, 5], which could encompass source coding, channel coding, modulation, channel equalization, and multiuser detection, ideally an end-to-end approach should be taken that considers all components together. In practice, however, depending on their background, researchers typically focus on one or two areas (e.g., error-resilient video coding [53], channel coding [54], joint source-channel coding (JSCC) [20].) Bringing our expertise in source coding (signal processing) and wireless communications together, this work in this chapter combines JSCC with multiuser detection. The issue of power allocation naturally arises in the scenario we consider. Choosing power allocation as a means of adjusting the SINRs (hence packet loss ratios) of different CDMA fading channels, in the form of unequal power level assignment, provides an additional degree of freedom with

respect to JSCC via error control alone, therefore can achieve higher overall peak signal to noise ratio (PSNR) values. As an example, for transmission of the QCIF "Foreman" sequence at 50 Kbps with roughly the same average power level, a PSNR gain of about 3.5 dB is achieved with joint power allocation and error control over with rate optimization only.

The remainder of this chapter is organized as follows: Section B introduces source coding and our FEC coding model while Section C describes the CDMA channel, linear MMSE and blind multiuser detection. Section D is devoted to joint error control and power allocation. Section E presents both analytical and simulation results. Section F concludes the chapter.

B. Source Coding and FEC Coding Model

In this section we describe the schemes for source coding, packetization and FEC coding adopted in our proposed system.

1. Source Coding

The widespread availability and acceptance of wireless service make it the natural next step to support video transmission over wireless channels. With a broadband wireless network in place, the key bottleneck of wireless visual phone is video compression because full motion video requires at least 8 Mbps bandwidth. A compression ratio of over 200:1 is needed for transmission of video over a 32 Kbps wireless link! International standards like MPEG-2 [55] and H.263+ [56] for video compression have been developed during the past decade for a number of important commercial applications (e.g., satellite TV and DVD).

However, these standard algorithms cannot meet general needs of wireless video,

because they are not designed or optimized with wireless applications in mind. Real-time wireless multimedia applications must be interoperable cross platforms and adaptive to available bandwidth as determined by the antenna size. Scalable coding, also known as layered, embedded, or progressive coding, is very desirable in these scenarios because it encodes a video source in layers, like an onion, that facilitate easy bandwidth adaptation. But it is extremely difficult to write a compression algorithm that can layer the data properly, without a performance penalty. That is: a scalable compression algorithm inherently delivers lower quality than an algorithm that can optimally encode the source monolithically, like a solid ball. So the difficulty is minimizing the effect of this structural constraint on the efficiency of the compression algorithm, both in terms of computational complexity and quality delivered at a given bandwidth.

Standard algorithms do not do well in this regard. Experiments with MPEG-2 [55] and H.263+ [56] in scalable mode show that, compared with monolithic (non-layered) coding [55, 57], the average PSNR loses roughly one dB with each layer. Furthermore, it is difficult for these coding schemes to achieve scalability because there is always a potential drifting problem [58] associated with predictive coding. The main focus of the MPEG-4 standard [21] is object-based coding and scalability in it is very limited. MPEG-4's streaming video profile on FGS coding [22] only provides flexible rate scalability and the coding performance is still about 1-1.5 dB lower than that of a monolithic coding scheme [59]. In addition, error propagation [60] due to packet loss is particularly severe if the video coding scheme exploits temporal redundancy of the video sequence, like it is done in H.263+ and MPEG-4.

Wireless video will play a major role in shaping how new compression algorithms are defined and computer or network resources are used in the 21st century. As such, researchers have been looking into new scalable video coding techniques (e.g., 3-D

wavelet video coding [6, 61, 62, 63, 64, 65]) over and beyond the H.263+ and MPEG-4 standards. 3-D wavelet video coding deviates from the standard motion compensated DCT approach in H.263+ or MPEG-4. Instead, it seeks for alternative means of video coding by exploiting spatio-temporal redundancies via 3-D wavelet transformation. Promising results have been reported. For example, Choi and Woods [65] presented better results than MPEG-1 using a 3-D subband approach together with hierarchical variable size block-based motion compensation. In particular, the 3-D SPIHT video coder, which is a 3-D extension of the celebrated SPIHT image coder [66], was chosen by Microsoft as the basis of its next-generation streaming video technology [19]. The latest embedded video coder [7] showed for the first time that 3-D wavelet video coding outperforms MPEG-4 coding by as much as two dB for most low motion and high motion sequences.

In our work, we choose to use the 3-D SPIHT coder because: 1) it is 100% embedded, i.e., there is no performance penalty due to scalability; 2) it achieves comparable performance to MPEG-2 and H.263 while being embedded; and 3) it is insensitive to error propagation, like most other 3-D wavelet video coders. In the following we briefly review the 2-D and 3-D SPIHT algorithms.

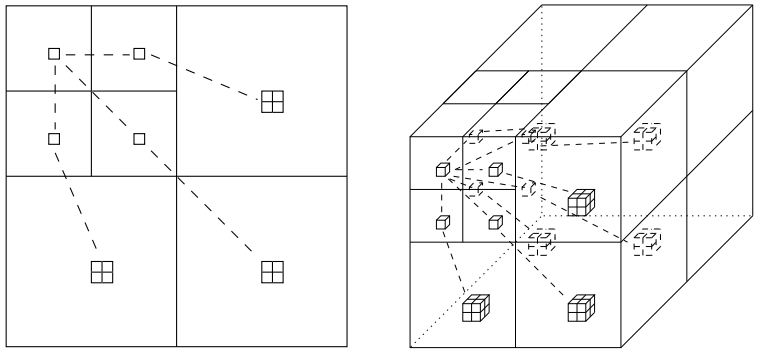
The 2-D SPIHT algorithm [66], like the embedded zerotree wavelet (EZW) coding algorithm [67], views wavelet coefficients as a collection of spatial orientation trees, with each tree consisting of coefficients from all subbands that correspond to the same spatial location in an image (see Fig. 1 (a)). It uses multipass "zerotree" coding to transmit the largest wavelet coefficients (in magnitude) first. A set of tree coefficients is significant if the largest coefficient magnitude in the set is greater than or equal to a certain threshold (e.g., a power of two); otherwise, it is insignificant. Similarly, a coefficient is significant if its magnitude is greater than or equal to the threshold; otherwise, it is insignificant. In each pass the significance of a larger set in the tree

is tested first: if the set is insignificant, a binary "zerotree" bit is used to set all coefficients in the set to zero; otherwise, the set is partitioned into subsets (or child sets) for further significance tests. After all coefficients are tested in one pass, the threshold is halved before the next pass.

The underlying assumption of SPIHT coding is that most images can be modeled as having decaying power spectral densities. That is: if a parent node in the wavelet coefficient tree is insignificant, it is very likely that its descendants are also insignificant. The zerotree symbol is used very efficiently in this case to signify a spatial subtree of zeros.

When the thresholds are powers of two, SPIHT coding can be thought of as a bit-plane coding scheme. It encodes one bit-plane at a time, starting from the most significant bit. With the sign bits and refinement bits (for coefficients that become significant earlier) being coded on the fly, SPIHT achieves *embedded* coding in the wavelet domain using three lists: the list of significant pixels (LSP); the list of insignificant pixels (LIP); and the list of insignificant sets (LIS). The 2-D SPIHT coder performs competitively with most other coders published in the literature [68], while possessing desirable features such as relatively low complexity and rate embeddedness. It represents the current state-of-the-art of wavelet image coding.

The 2-D SPIHT algorithm [66] was extended to 3-D embedded SPIHT video coding in [6]. Besides motion compensation, the 3-D SPIHT algorithm is in principle the same as 2-D SPIHT, except that 3-D wavelet coefficients are treated as a collection of 3-D spatio-temporal orientation trees (see Fig. 1 (b)) and that context modeling in arithmetic coding is more involved. A block-based motion estimation scheme is implemented in the 3-D SPIHT coder in [6], and an option for not using motion estimation is also allowed to reduce the encoding complexity. Global affine



(a) (b)

Fig. 1. EZW-style 2-D spatial orientation tree and 3-D spatio-temporal orientation tree.

motion compensation was combined with the 3-D SPIHT algorithm in [69]. Every 32 frames form a GOF and the Daubechies 9/7 biorthogonal filters of [70] are used in all three dimensions to perform a separable wavelet decomposition. The temporal transform and 2-D spatial transform are done separately by first performing three levels of a dyadic wavelet decomposition in the temporal direction, and then within each of the resulting temporal bands, performing three levels of a 2-D spatial dyadic decomposition. transform Spatio-temporal orientation trees coupled with powerful SPIHT sorting and refinement turns out to be very efficient. Even without motion compensation, the 3-D SPIHT coder provides comparable performance to H.263 objectively and subjectively when operating at bit rates of 30 to 60 Kbps. It outperforms MPEG-2 at the same bit rate (1.5 to 4 Mbps). In addition to being rate scalable, the 3-D SPIHT video coder allows multiresolutional scalability in encoding and decoding in both time and space. This added functionality along with many desirable features, such as full embeddedness for progressive transmission, precise rate control for constant bit rate traffic, and low complexity for possible software only video applications, makes the video coder an attractive candidate for applications like wireless video.

2. Packetization

We assume that the video source is encoded by the 3-D SPIHT [6] coder. An embedded bitstream is first generated for each GOF, each bitstream then partitioned into a sequence of packets, and the l-th packet assigned to the l-th source layer, $l = 1, \cdots, L$, as shown in Fig. 2. Typical parameters for such a construction might be the following for a 25 frame-per-second QCIF (176×144) video sequence: 32 frames per GOF, 50 packets per GOF, and 1000 bytes per packet payload. This implies that a GOF has a duration of 1.28s; that there are 50 source layers; that each source layer has a bit rate of 6.25 Kbps; and that a receiver that receives all 50 source layers can decode the video to about 0.5 bit per pixel (312.5 Kbps).

Because each GOF is encoded into an embedded bitstream that is partitioned into a sequence of packets, any nested subsets of these packets are decodable to a level of quality commensurate with the total bit rate of the subset. Thus there is a sequential dependency between these packets. Let $\Delta R_i \geq 0$ and $\Delta D_i \geq 0$ denote the expected increase in rate (per GOF) if the i-th packet is transmitted and the expected decrease in distortion (per GOF) if the i-th packet is decoded, respectively, the total transmission rate from the first packet through the l-th packet is simply

$$R(l) = \sum_{i=1}^{l} \Delta R_i. \tag{2.1}$$

Similarly, the expected distortion experienced by the receiver is

$$D(l) = D_0 - \sum_{i=1}^{l} P_{\preceq i} \Delta D_i, \tag{2.2}$$

where D_0 is the expected distortion when the rate is zero (i.e., when the transmitter does not send any source layers) and $P_{\preceq i}$ is the probability that the i-th packet and all its preceding packets are received. Note that $P_{\preceq i}$ depends on the channel model

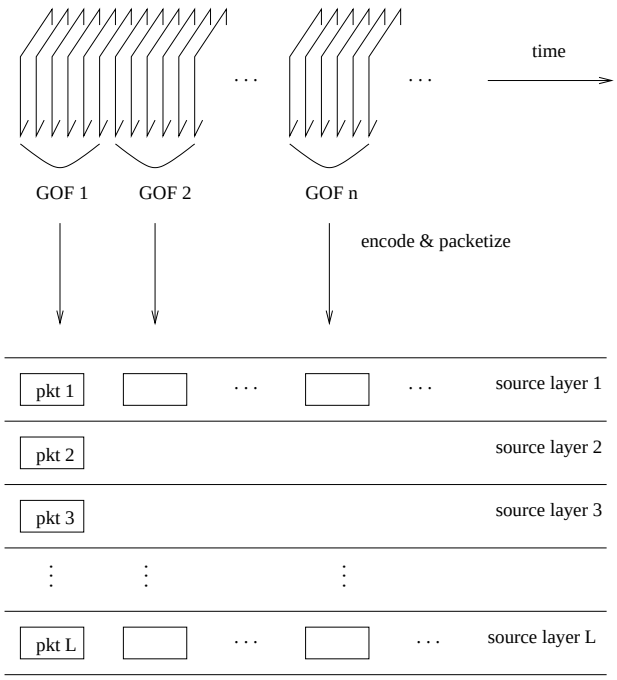


Fig. 2. Source packetized into layers.

and possibly on packet transmission sequence.

3. FEC Coding

Now, we outline the FEC coding scheme. Each source layer is partitioned into coding blocks having K source packets per coding block. The block size K is constant across all source layers. For each block of K source packets in a source layer, we assume that $N_{\max} - K$ parity packets are produced using a systematic $(N_{\max}, K)$ RS style erasure correction code [52]. $N_{\max} - K$ is the maximum amount of redundancy that will be needed by the transmitter to protect the source layer. The $N_{\max} - K$ parity packets are generated byte-wise from the K source packets, using for each byte the generator matrix from the RS style code over the finite Galois field $GF(2^8)$. As long as the total number of correctly received packets in an RS coded block is greater than or equal to K, all K source packets can be recovered.

Each of the parity packets so generated is placed together with its source packets as a layer and will be assigned to the same channel. A group of source/parity layers are transmitted through a CDMA channel. In this work, we assume that a total of 4 CDMA channels are used to transmit the 50 layers of source/parity data packets, with channel 1 transmitting the first 2 layers, channel 2 transmitting the next 6 layers, channel 3 transmitting the next 17 layers, and channel 4 transmitting the last 25 layers, as shown in Fig. 3. Note that for traditional voice communication in CDMA systems, each user occupies one channel. However, in multimedia communications, each user may occupy more than one channel to transmit data (e.g., here the video user occupies 4 channels). Nevertheless, we use the terms *user* and *channel* interchangeably from now on. The CDMA channel model and the multiuser receivers will be discussed in Section C.

The transmitter now has many layers to send. It can transmit any collection of source layers and any collection of parity packets associated with those source layers. The transmitter buffers frames as they arrive. When a GOF is accumulated, it encodes the GOF and packetizes the resulting layered bitstream. After K such GOFs, the transmitter computes the $N_{\max} - K$ parity packets for each coding block of K source packets. For a fixed transmission rate and a fixed power level, the transmitter chooses to transmit the optimal number of source and parity packets highlighted in Fig. 3, based on the optimization procedure given in Section D.

The receiver instantly recovers as many source packets as possible from the received source and parity packets, and decodes them. Playback begins after exactly K GOFs of coding delay.

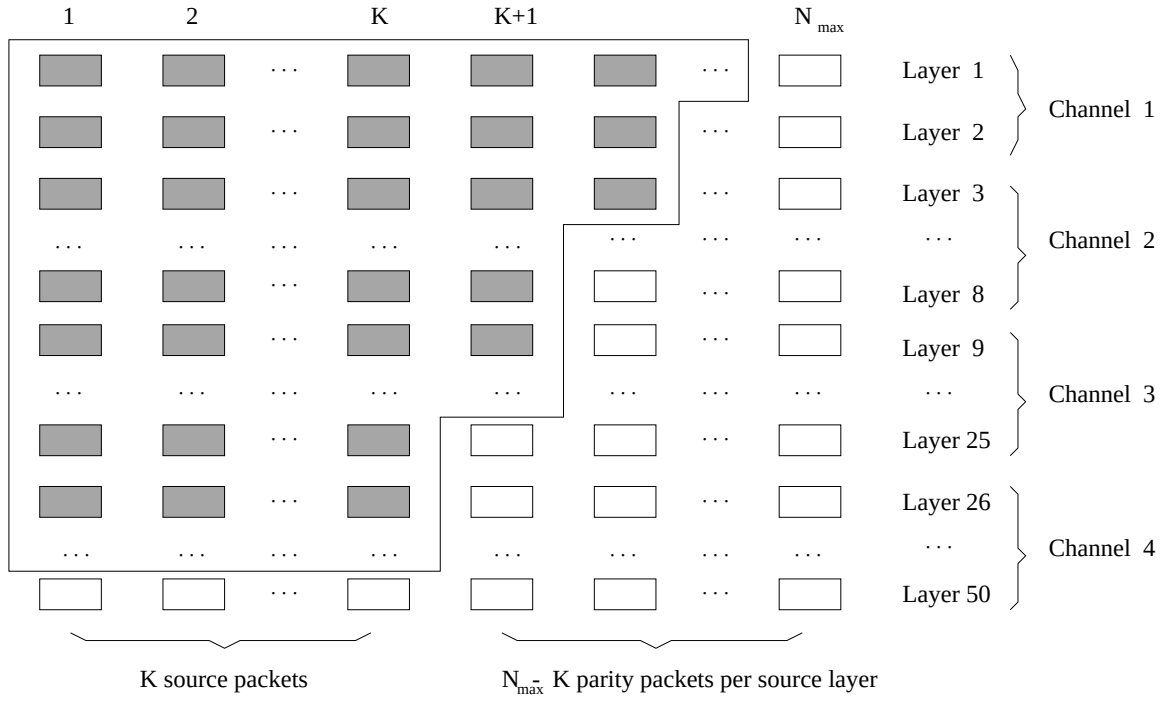


Fig. 3. Relationship between source/parity layers and channels. The transmitted packets are highlighted.

C. CDMA Channel and Linear MMSE Multiuser Detection

In this work, the video data packets are transmitted through CDMA channels. Here we propose employing linear multiuser receiver for demodulating the received data. Specifically, we consider both the exact linear MMSE multiuser detector when the channel conditions are known to the receiver (i.e., uplink); and the blind linear MMSE receiver when the channel conditions are unknown to the receiver (i.e., downlink).

1. Synchronous CDMA Signal Model

We start by introducing the most basic multiple-access signal model, namely, a baseband, G-user, time-invariant, synchronous, additive white Gaussian noise (AWGN) system, employing periodic (short) spreading sequences, and operating with a coher-

ent BPSK modulation format. The continuous-time waveform received by a given user in such a system can be modeled as follows

$$r(t) = \sum_{k=1}^{G} \sqrt{P_k} \sum_{i=0}^{M-1} b_k[i] s_k(t - iT) + n(t), \tag{2.3}$$

where M is the number of data symbols per user in the data frame of interest; T is the symbol interval; P_k, $\{b_k[i]\}_{i=0}^{M-1}$ and $s_k(t)$ denote respectively the power level, the transmitted symbol stream, and the normalized signaling waveform of the k-th user; and $n(t)$ is the baseband white Gaussian ambient channel noise with power spectral density σ^2. It is assumed that for each user k, $\{b_k[i]\}_{i=0}^{M-1}$ is a collection of independent equiprobable ± 1 random variables, and the symbol streams of different users are independent. The user signaling waveform is of the form

$$s_k(t) \;=\; \frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} c_{j,k} \psi(t - jT_c), \quad 0 \le t < T, \tag{2.4}$$

where N is the processing gain; $\{c_{j,k}\}_{j=0}^{N-1}$ is a signature sequence of ± 1's assigned to the k-th user; and $\psi(\cdot)$ is a chip waveform of duration $T_c = \frac{T}{N}$ and with unit energy, i.e., $\int_0^{T_c} \psi(t)^2 \mathrm{d}t = 1$.

At the receiver, the received signal $r(t)$ is filtered by a chip-matched filter and then sampled at the chip rate. The sample corresponds to the j-th chip of the i-th symbol is given by

$$r_j[i] \triangleq \int_{iT+jT_c}^{iT+(j+1)T_c} r(t)\psi(t - iT - jT_c)\mathrm{d}t, \quad j = 0, \cdots, N-1; \quad i = 0, \cdots, M-1. \tag{2.5}$$

The resulting discrete-time signal corresponding to the i-th symbol is then given by

$$\boldsymbol{r}[i] = \sum_{k=1}^{G} \sqrt{P_k} b_k[i] \boldsymbol{s}_k + \boldsymbol{n}[i] = \boldsymbol{S}\sqrt{\boldsymbol{P}}\boldsymbol{b}[i] + \boldsymbol{n}[i], \tag{2.6}$$

with

$$\boldsymbol{r}[i] \triangleq \begin{bmatrix} r_0[i] \\ r_1[i] \\ \vdots \\ r_{N-1}[i] \end{bmatrix}, \quad \boldsymbol{s}_k \triangleq \frac{1}{\sqrt{N}} \begin{bmatrix} c_{0,k} \\ c_{1,k} \\ \vdots \\ c_{N-1,k} \end{bmatrix}, \quad \boldsymbol{n}[i] \triangleq \begin{bmatrix} n_0[i] \\ n_1[i] \\ \vdots \\ n_{N-1}[i] \end{bmatrix}, \tag{2.7}$$

where $n_j[i] = \int_{jT_c}^{(j+1)T_c} n(t)\psi(t - iT - jT_c)\mathrm{d}t \sim \mathcal{N}(0, \sigma^2)$ is a Gaussian random variable; $\boldsymbol{n}[i] \sim \mathcal{N}(\boldsymbol{0}, \sigma^2 \boldsymbol{I}_N)$; $\boldsymbol{S} \triangleq [\boldsymbol{s}_1 \ \cdots \ \boldsymbol{s}_G]$; $\boldsymbol{P} \triangleq \mathsf{diag}(P_1, \cdots, P_G)$; and $\boldsymbol{b}[i] \triangleq [b_1[i] \ \cdots \ b_G[i]]^T$.

2. Linear MMSE Detector

Suppose that we are interested in demodulating the data bits of a particular user, say user 1, $\{b_1[i]\}_{i=0}^{M-1}$, based on the received waveforms $\{\boldsymbol{r}[i]\}_{i=0}^{M-1}$. A *linear* receiver for this purpose is a vector $\boldsymbol{w}_1 \in \mathbb{R}^N$, such that the desired user's data bits are demodulated according to

$$z_1[i] = \boldsymbol{w}_1^T \boldsymbol{r}[i], \tag{2.8}$$

$$\hat{b}_1[i] = \mathrm{sign}\left\{z_1[i]\right\}. \tag{2.9}$$

Substituting (2.6) into (2.8), the output of the linear receiver $\boldsymbol{w}_1$ can be written as

$$z_1[i] = \sqrt{P_1}\left(\boldsymbol{w}_1^T \boldsymbol{s}_1\right) b_1[i] + \sum_{k=2}^{G} \sqrt{P_k}\left(\boldsymbol{w}_1^T \boldsymbol{s}_k\right) b_k[i] + \boldsymbol{w}_1^T \boldsymbol{n}[i]. \tag{2.10}$$

In (2.10), the first term contains the useful signal of the desired user; the second term contains the signals from other undesired users – the so-called *multiple-access interference* (MAI); and the last term contains the ambient Gaussian noise. The simplest linear receiver is the conventional matched-filter, where $\boldsymbol{w}_1 = \boldsymbol{s}_1$. It is well

known that such a matched-filter receiver is optimal only in a single-user channel (i.e., $G = 1$). In a multiuser channel (i.e., $G > 1$), this receiver may perform poorly since it makes no attempt to ameliorate the MAI, a limiting source of interference in multiple-access channels.

The linear minimum mean-square error (MMSE) detector is designed to minimize the total effect of the MAI and the ambient noise at the detector output. Specifically, it is given by the solution to the following optimization problem

$$\boldsymbol{w}_1 = \arg \min_{\boldsymbol{w} \in \mathbb{R}^N} E\left\{ \left(b_1[i] - \boldsymbol{w}^T \boldsymbol{r}[i] \right)^2 \right\} = \boldsymbol{C}_r^{-1} \boldsymbol{s}_1, \tag{2.11}$$

with

$$\boldsymbol{C}_r \triangleq E\left\{ \boldsymbol{r}[i] \boldsymbol{r}[i]^T \right\} = \sum_{k=1}^{G} P_k \boldsymbol{s}_k \boldsymbol{s}_k^T + \sigma^2 \boldsymbol{I}_N = \boldsymbol{S}\boldsymbol{P}\boldsymbol{S}^T + \sigma^2 \boldsymbol{I}_N. \tag{2.12}$$

Denote the normalized cross-correlation matrix of the signal set $\boldsymbol{s}_1, \cdots, \boldsymbol{s}_G$ as

$$\boldsymbol{R} \triangleq \boldsymbol{S}^T \boldsymbol{S} = \begin{pmatrix} \rho_{11} & \cdots & \rho_{1G} \\ \vdots & \vdots & \vdots \\ \rho_{G1} & \cdots & \rho_{GG} \end{pmatrix}, \tag{2.13}$$

where $\rho_{ij} \triangleq \boldsymbol{s}_i^T \boldsymbol{s}_j$.

Since it is assumed that the user bit streams are independent, and the noise is independent of the user bits, following [71, 72, 73, 74], the signal-to-interference-plus-noise ratio (SINR) at the output of the linear detector $\boldsymbol{w}_1$ is given by

$$\mathrm{SINR}_1 \triangleq \frac{E\left\{ z_1[i] \mid b_1[i] \right\}^2}{E\{Var\{ z_1[i] \mid b_1[i] \}\}} = \frac{P_1 \left(\boldsymbol{w}_1^T \boldsymbol{s}_1 \right)^2}{\sum_{k=2}^{G} P_k \left(\boldsymbol{w}_1^T \boldsymbol{s}_k \right)^2 + \sigma^2 \|\boldsymbol{w}_1\|^2}, \tag{2.14}$$

where

$$\boldsymbol{w}_l^T \boldsymbol{s}_k = \frac{1}{P_l}\left[\boldsymbol{R}(\boldsymbol{R}+\sigma^2\boldsymbol{P}^{-1})^{-1}\right]_{k,l}, \quad k,l=1,\cdots,G, \tag{2.15}$$

$$\|\boldsymbol{w}_1\|^2 = \frac{1}{P_1^2}\left[(\boldsymbol{R}+\sigma^2\boldsymbol{P}^{-1})^{-1}\boldsymbol{R}(\boldsymbol{R}+\sigma^2\boldsymbol{P}^{-1})^{-1}\right]_{1,1}. \tag{2.16}$$

It is shown in [75] that the output of a linear MMSE detector is well approximated by a Gaussian distribution. Thus the bit error rate can be expressed as

$$P_{b,1} = Q(\sqrt{\mathrm{SINR}_1}), \tag{2.17}$$

where $Q(x) \triangleq \frac{1}{\sqrt{2\pi}}\int_x^\infty \exp\left(-\frac{t^2}{2}\right)\mathrm{d}t$.

3. Blind Detector

It is seen from (2.12) that the linear MMSE detector $\boldsymbol{w}_1$ is a function of the signature sequences $\boldsymbol{S}$ of all G users. Recall that for the matched-filter receiver, the only prior knowledge required is the desired user's signature sequence $\boldsymbol{s}_1$. In the downlink of a CDMA system, the mobile receiver typically only has the knowledge of its own signature sequence, but not of those of the other users. Hence it is of interest to consider the problem of *blind* implementation of the linear detector, i.e., without the requirement of knowing the signature sequences of the interfering users.

Let the eigendecomposition of $\boldsymbol{C}_r$ in (2.12) be

$$\boldsymbol{C}_r \;=\; \boldsymbol{U}_s\boldsymbol{\Lambda}_s\boldsymbol{U}_s^T + \sigma^2\boldsymbol{U}_n\boldsymbol{U}_n^T, \tag{2.18}$$

where $\boldsymbol{\Lambda}_s = \mathsf{diag}(\lambda_1,\cdots,\lambda_G)$ contains the largest G eigenvalues of $\boldsymbol{C}_r$; $\boldsymbol{U}_s = [\boldsymbol{u}_1,\cdots,\boldsymbol{u}_G]$ contains the eigenvectors corresponding to the largest G eigenvalues in $\boldsymbol{\Lambda}_s$; $\boldsymbol{U}_n = [\boldsymbol{u}_{G+1},\cdots,\boldsymbol{u}_N]$ contains the $(N-G)$ eigenvectors corresponding to the smallest eigenvalue σ^2 of $\boldsymbol{C}_r$. It is known that $\mathsf{range}(\boldsymbol{U}_s) = \mathsf{range}(\boldsymbol{S})$ is the signal subspace; and

$\mathsf{range}(\boldsymbol{U}_n) \perp \mathsf{range}(\boldsymbol{S})$ is the noise subspace. The linear MMSE detector $\boldsymbol{w}_1$ in (2.11) can also be written in terms of the signal subspace components as [76]

$$\boldsymbol{w}_1 = \boldsymbol{U}_s \boldsymbol{\Lambda}_s^{-1} \boldsymbol{U}_s^T \boldsymbol{s}_1. \tag{2.19}$$

Corresponding to the two forms of the linear MMSE detector (2.11) and (2.19), there are two approaches to its blind implementation. In the direct-matrix-inversion (DMI) method, the autocorrelation matrix $\boldsymbol{C}_r$ in (2.11) is replaced by the corresponding sample estimate. That is

$$\hat{\boldsymbol{C}}_r = \frac{1}{M} \sum_{i=1}^{M} \boldsymbol{r}[i] \boldsymbol{r}[i]^T, \tag{2.20}$$

$$\hat{\boldsymbol{w}}_1 = \hat{\boldsymbol{C}}_r^{-1} \boldsymbol{s}_1. \quad \text{[DMI blind linear MMSE detector]} \tag{2.21}$$

In the subspace method, the eigencomponents $\boldsymbol{\Lambda}_s$ and $\boldsymbol{U}_s$ in (2.19) are replaced by the corresponding eigenvalues and eigenvectors of the sample autocorrelation matrix $\hat{\boldsymbol{C}}_r$. That is

$$\begin{aligned} \hat{\boldsymbol{C}}_r &= \frac{1}{M} \sum_{i=1}^{M} \boldsymbol{r}[i] \boldsymbol{r}[i]^T \\ &= \hat{\boldsymbol{U}}_s \hat{\boldsymbol{\Lambda}}_s \hat{\boldsymbol{U}}_s^T + \hat{\boldsymbol{U}}_n \hat{\boldsymbol{\Lambda}}_n \hat{\boldsymbol{U}}_n^T, \end{aligned} \tag{2.22}$$

$$\hat{\boldsymbol{w}}_1 = \hat{\boldsymbol{U}}_s \hat{\boldsymbol{\Lambda}}_s \hat{\boldsymbol{U}}_s^T \boldsymbol{s}_1, \quad \text{[Subspace blind linear MMSE detector]} \tag{2.23}$$

where $\hat{\boldsymbol{\Lambda}}_s$ and $\hat{\boldsymbol{U}}_s$ contain respectively the largest G eigenvalues and the corresponding eigenvectors of $\hat{\boldsymbol{C}}_r$; $\hat{\boldsymbol{\Lambda}}_n$ and $\hat{\boldsymbol{U}}_n$ contain respectively the remaining eigenvalues and eigenvectors of $\hat{\boldsymbol{C}}_r$.

According to [71, 72, 73, 74], the output SINR_1 of the blind detector can be expressed as

$$\frac{P_1\left(\boldsymbol{w}_1^T\boldsymbol{s}_1\right)^2}{\sum_{k=2}^{G}P_k\left(\boldsymbol{w}_1^T\boldsymbol{s}_k\right)^2+\sigma^2\|\boldsymbol{w}_1\|^2+\frac{1}{M}\left[(G+1)\boldsymbol{w}_1^T\boldsymbol{s}_1-2\sum_{k=1}^{G}P_k^2\left(\boldsymbol{w}_1^T\boldsymbol{s}_k\right)\boldsymbol{w}_k^T\boldsymbol{s}_k+(N-G)\tau\sigma^2\right]}, \tag{2.24}$$

where $\boldsymbol{w}_l^T\boldsymbol{s}_k$ and $\|\boldsymbol{w}_1\|^2$ are given respectively by (2.15) and (2.16), and

$$\tau\sigma^2=\begin{cases}\boldsymbol{w}_1^T\boldsymbol{s}_1, & \text{[DMI blind detector]}\\ \frac{\sigma^4}{P_1^2}\left[\left(\boldsymbol{R}+\sigma^2\boldsymbol{P}^{-1}\right)^{-1}\boldsymbol{P}^{-1}\boldsymbol{R}^{-1}\right]_{1,1}. & \text{[Subspace blind detector]}\end{cases} \tag{2.25}$$

Note that the last term in the denominator of (2.24) represents the noise power due to the estimation error. Also note that the performance difference between the two forms of blind detectors is due to the term τ given by (2.25). It is shown in [71, 72, 73, 74] that in realistic channels the subspace blind detector outperforms the DMI blind detector and that the output of the blind detector is approximately Gaussian. Therefore the BER can be expressed as

$$P_{b,1}=Q\left(\sqrt{\text{SINR}_1}\right). \tag{2.26}$$

Remark 1: In our proposed system for video transmission over CDMA networks, the video stream occupies up to 4 CDMA channels. The other users in the same network act as interference. In the CDMA uplink, the base station has the knowledge of the signature sequences of all users, and therefore can use the exact linear MMSE receiver. For the downlink, on the other hand, the mobile station typically knows only the signature sequence of its own, and hence must employ the blind detector.

Remark 2: Although here we consider a simple synchronous CDMA channel model, as shown in [77, 78, 79], the asynchronous multipath fading CDMA channel model

bears a similar form as (2.6), and therefore the techniques developed in this work can be extended to multipath fading channels.

D. Optimal Rate Allocation and Power Allocation

In this section, we consider the problem of finding the collection of source and parity packets which the transmitter should send to the channels, and the power level distribution across the CDMA channels, given a maximum transmission rate and a maximum power level in order to minimize the average distortion of the received video data. In general, this is equivalent to the problem of combined optimal allocation of transmission rate between source and channel codes with a constraint on the total transmission rate, and optimal allocation of power level with a constraint on the total power level. Note that optimal allocation of transmission rate between source and channel codes for a given transmission rate is addressed in [19].

1. Problem Formulation

Let $\overline{G}$ be the number of CDMA channels used for transmitting a video stream (in this work, $\overline{G} = 4$). Denote $\boldsymbol{p} = [P_1, P_2, \cdots, P_{\overline{G}}]$ as the *power level allocation vector* for the transmitted powers of the $\overline{G}$ channels used for transmitting the video data packets. Here we assume that the powers of other users in the network (i.e., $P_{\overline{G}+1}, \cdots, P_G$) are fixed. One of our objective is to optimize the power allocation $\boldsymbol{p}$ vector subject to a total power constraint such that the distortion is minimized. Note that from Sections C.2 and C.3, the bit error rate, and therefore the packet loss probability in a channel, depends on the power levels of all channels, not just the power of that particular channel.

Let L_j be the number of source layers transmitted over the j-th channel (in this

work, $L_1 = 2, L_2 = 6, L_3 = 17$ and $L_4 = 25$). Then the total number of source layers L is

$$L = \sum_{j=1}^{\overline{G}} L_j. \tag{2.27}$$

Let K be the number of source packets per code block per source layer; $N_{\max}$ be the maximum number of source packets plus parity packets per code block; and N_i ($0 \leq N_i \leq N_{\max}$) be the number of source packets plus parity packets in the code block for the i-th source layer transmitted (see highlighted packets in Fig. 3.). Note that we have either $N_i = 0$ or $N_i \geq K$, because the transmitter either does not transmit the i-th layer at all, or at least transmits the K source packets at the i-th layer. When $N_i = 0$, it means that no packet corresponding to the i-th source layer is transmitted; when $N_i = K$, only the K source packets are transmitted for the i-th source layer; and when $N_i > K$, denote $r_i = N_i/K$ as the redundancy per GOF transmitted to the receiver for the i-th layer. We call $\boldsymbol{r} = (r_1, \cdots, r_l)$ the *rate allocation vector* [19]. The rate allocation vector specifies how many source and parity packets to transmit for each source layer when the transmitter transmits the first l source layers. In this way, the rate allocation vector specifies the allocation of the total transmission rate between source packets and parity packets. Any given rate allocation vector $\boldsymbol{r}$ induces a total transmission rate (in terms of transmitted packets per GOF)

$$R(\boldsymbol{r}) = \sum_{i=1}^{l} r_i = \frac{1}{K} \sum_{i=1}^{l} N_i. \tag{2.28}$$

Given a specified data rate, we can determine the number of source layers, l, to be transmitted. Assume these layers are transmitted using up to $\overline{G}$ CDMA channels.

By (2.2), the total distortion at the receiver is given by

$$D(\boldsymbol{r}, \boldsymbol{p}) = D(l) = D_0 - \sum_{i=1}^{l} P_{\preceq i} \Delta D_i, \tag{2.29}$$

where

$$\begin{aligned} P_{\preceq i} &= P(\text{the first } i \text{ layers are decoded correctly}) \\ &= \prod_{j=1}^{i} \theta_j(r_j, \boldsymbol{p}), \end{aligned} \tag{2.30}$$

where $\theta_j(r_j, \boldsymbol{p})$ is the probability that the j-th layer source packet is recovered correctly after channel decoding. Let the j-th layer packets be transmitted through the $c(j)$-th CDMA channel and $\epsilon_j(\boldsymbol{p})$ be the packet loss probability of the j-th layer. Assume the bit error occurs independently, then the packet loss probability for the j-th layer can be written as

$$\epsilon_j(\boldsymbol{p}) = 1 - (1 - P_{b,c(j)})^{n_b}, \tag{2.31}$$

where $P_{b,c(j)}$ is given by (2.17) for the exact linear MMSE detector or given by (2.26) for the blind detectors; n_b is the packet size in *bits* (in this work, $n_b = 8000$ bits). When $N_j \geq K$ and assume that a (N_j, K) RS style erasure code is used. Then

$$\theta_j(r_j, \boldsymbol{p}) = \frac{\kappa\Big(r_j, K, \epsilon_j(\boldsymbol{p})\Big)}{K}, \tag{2.32}$$

where $\kappa\Big(r_j, K, \epsilon_j(\boldsymbol{p})\Big)$ is the expected number of source packets that can be recovered after channel decoding with a (N_j, K) RS erasure code at source layer j, which is given by

$$\kappa\Big(r_j, K, \epsilon_j(\boldsymbol{p})\Big) = \sum_{m=1}^{K-1} \binom{r_j K}{m} \epsilon_j(\boldsymbol{p})^{r_j K - m} \Big(1 - \epsilon_j(\boldsymbol{p})\Big)^m \left(\frac{m}{r_j}\right)$$

$$+\sum_{m=K}^{r_jK}\binom{r_jK}{m}\epsilon_j(\boldsymbol{p})^{r_jK-m}\Big(1-\epsilon_j(\boldsymbol{p})\Big)^m K. \tag{2.33}$$

Hence the expected distortion (per GOF) for the given rate allocation vector $\boldsymbol{r}$ and power allocation vector $\boldsymbol{p}$ is

$$D(\boldsymbol{r},\boldsymbol{p}) = D_0 - \sum_{i=1}^{l}\left(\prod_{j=1}^{i}\theta_j(r_j,\boldsymbol{p})\right)\Delta D_i. \tag{2.34}$$

With the expected distortion expression (2.34) for any rate allocation vector $\boldsymbol{r}$ and any power level allocation vector $\boldsymbol{p}$ now in hand, we can optimize these two vectors to minimize the expected distortion subject to a transmission rate constraint as well as a power level constraint. That is, we consider the following constraint optimization problem

$$\min_{\boldsymbol{r},\boldsymbol{p}} D(\boldsymbol{r},\boldsymbol{p}) \;\text{ subject to }\; \sum_{i=1}^{l} r_i \le R \;\text{ and }\; \sum_{k=1}^{\overline{G}} P_k \le P, \tag{2.35}$$

where R and P are a total transmission rate and a total power level, respectively.

2. The Optimization Algorithm

One way to solve the above constraint optimization is by finding the rate allocation vector $\boldsymbol{r}$ and power allocation vector $\boldsymbol{p}$ that minimizes the Lagrangian

$$\begin{aligned} J(\boldsymbol{r},\boldsymbol{p},\lambda_1,\lambda_2) &= D(\boldsymbol{r},\boldsymbol{p}) + \lambda_1\sum_{i=1}^{l} r_i + \lambda_2\sum_{k=1}^{\overline{G}} P_k \\ &= D_0 + \sum_{i=1}^{l}\left[\left(-\prod_{j=1}^{i}\theta_j(r_j,\boldsymbol{p})\right)\Delta D_i + \lambda_1 r_i\right] + \lambda_2\sum_{k=1}^{\overline{G}} P_k, \end{aligned} \tag{2.36}$$

The solution to this problem is completely characterized by the set of distortion increments ΔD_i, which are determined by the source code and packetization, and the probability $\theta_j(r_j,\boldsymbol{p})$ with which the j-th layer source packet is recovered correctly,

which is in turn determined by the user spreading codes, power allocation, type of receiver, and channel code. There are many methods to find this solution in the case of one constraint on the transmission rate only [80, 81, 82].

We solve the problem by using an iterative approach that is based on the method of alternating variables for multivariable minimization [83]. The objective function $J(r_1, \cdots, r_l, P_1, \cdots, P_{\overline{G}})$ in (2.36) is minimized one variable at a time, keeping the other variables constant, until convergence. To be specific, let $\boldsymbol{r}^{(0)}$ and $\boldsymbol{p}^{(0)}$ be any initial rate allocation vector and power allocation vector, respectively. Let $\boldsymbol{r}^{(t)} = (r_1^{(t)}, \cdots, r_l^{(t)})$ and $\boldsymbol{p}^{(t)} = (P_1^{(t)}, \cdots, P_{\overline{G}}^{(t)})$ be determined for $t = 1, 2, \cdots$, as follows. Select one component $x \in \{r_1, \cdots, r_l, P_1, \cdots, P_{\overline{G}}\}$ to optimize at step t. This can be done in a round-robin style. Then for $r_i \neq x$, let $r_i^{(t)} = r_i^{(t-1)}$, or for $P_k \neq x$, let $P_k^{(t)} = P_k^{(t-1)}$. If $x = r_i$, then we perform the following rate optimization

$$\begin{aligned} r_i^{(t)} &= \arg\min_{r_i} J\left(r_1^{(t)}, \cdots, r_{i-1}^{(t)}, r_i, r_{i+1}^{(t)}, \cdots, r_l^{(t)}, P_1^{(t)}, \cdots, P_{\overline{G}}^{(t)}\right) \\ &= \arg\min_{r_i} \sum_{m=i}^{l} \left(-\prod_{j=1}^{m} \theta_j(r_j, \boldsymbol{p})\right) \Delta D_m + \lambda_1 r_i \end{aligned} \tag{2.37}$$

and if $x = P_k$, then we perform the following power optimization

$$\begin{aligned} P_k^{(t)} &= \arg\min_{P_k} J\left(r_1^{(t)}, \cdots, r_l^{(t)}, P_1^{(t)}, \cdots, P_{k-1}^{(t)}, P_k, P_{k+1}^{(t)}, \cdots, P_{\overline{G}}^{(t)}\right) \\ &= \arg\min_{P_k} \sum_{i=1}^{l} \left(-\prod_{j=1}^{i} \theta_j(r_j, \boldsymbol{p})\right) \Delta D_i + \lambda_2 P_k. \end{aligned} \tag{2.38}$$

For fixed λ_1 and λ_2, the one-dimensional minimization problems (2.37) and (2.38) can be solved using standard non-linear optimization procedures, such as gradient-descent-type algorithm [83]. Now in order to minimize the Lagrangian $J(\boldsymbol{r}, \boldsymbol{p}, \lambda_1, \lambda_2)$ given by (2.36), we proceed as follows: first for fixed (λ_2, P), we minimize $J(\boldsymbol{r}, \boldsymbol{p}, \lambda_1, \lambda_2)$ over (λ_1, r), then for fixed (λ_1, r), we minimize $J(\boldsymbol{r}, \boldsymbol{p}, \lambda_1, \lambda_2)$ over (λ_2, P). This pro-

cedure is repeated until convergence.

In our experiments, we always start with the initial rate allocation vector $\boldsymbol{r} = (1, 1, \cdots, 1)$ and the initial power allocation vector $\boldsymbol{p} = (\overline{P}, \overline{P}, \cdots, \overline{P})$ where $\overline{P}$ is the average power level over all channels for a given total power level. We cycle through the components, beginning with the component associated with the first layer and ending with the components associated with the last layer. The resulting power allocation $P_1^*, \cdots, P_{\overline{G}}^*$ specify the optimal packet loss probability in each channel for the resulting rate allocation $r_1^*, \cdots, r_l^*$, which in general will be in the form of UEP.

E. Numerical Results

1. CDMA System Setup

a. Uplink - Linear MMSE detector

For linear MMSE detection in CDMA uplink, the total number of channels is $G = 4$. We set $\rho_{ij} = 0.4$, when $i \neq j$; $\rho_{ij} = 1$, when $i = j$ for the normalized cross-correlation matrix $\boldsymbol{R}$ in (2.13). The power spectral density of baseband white Gaussian ambient channel noise is $\sigma^2 = 1$. The block size per coding block $K = 8$. The packet size $n_b = 8000$ bits. When the transmission rate is $R = 50, 100, 150$, and 200 Kbps, the corresponding number of CDMA channels to which the optimization algorithm is applied to, is 4, based on the layer-channel mapping structure shown in Fig. 3. The average power level of 4 channels, which results in the packet loss ratio 10%, is 23.85; The average power level 21.97 results in the packet loss ratio 20%. These are calculated using (2.14), (2.17) and (2.31) assuming $P_1 = P_2 = P_3 = P_4$.

b. Downlink - Blind detectors

For blind detection in CDMA downlink, the total number of channels is $G = 10$. When the transmission rate is $R = 50, 100, 150,$ and 200 Kbps, the corresponding number of CDMA channels to which the optimization algorithm is applied to, is 4, based on the layer-channel mapping structure shown in Fig. 3; the other 6 channels are considered as interference. Again $\rho_{ij} = 0.4$, when $i \neq j$; $\rho_{ij} = 1$, when $i = j$ for the normalized cross-correlation matrix $\boldsymbol{R}$ in (2.13). The processing gain $N = 15$ and the sample size $M = 8000$ bits, which is the packet length in *bits*. The power spectral density of baseband white Gaussian ambient channel noise is $\sigma^2 = 1$. The block size per coding block $K = 8$. The packet size $n_b = 8000$ bits. In addition, for the DMI blind MMSE detection, the average power levels, which result in the packet loss ratio 10% and 20%, are 27.65 and 25.39, respectively. For the subspace blind MMSE detection, the average power levels, which result in the packet loss ratio 10% and 20%, are 27.37 and 25.22, respectively. These power levels are calculated from (2.24), (2.26) and (2.31) assuming $P_1 = P_2 = \cdots = P_{10}$.

2. Analysis

In this subsection, we model the source as having an operational distortion-rate function $D(R) = A2^{-2R/B}$ for an arbitrary constant A and a scaling factor B. No actual source coding or channel coding is performed in this part.

We compare the performance of several systems. The first is a system with no error protection. The second is the FEC system with only rate optimization: UEP with redundancy up to $r = 2.0$ determined optimally for each layer and uniform power allocation across channels. The last is our hybrid FEC/power allocation system. They are all based on the linear MMSE, blind DMI and blind subspace detectors.

In this work, however, because we are only interested in the performance difference of the systems with FEC/power allocation, pure FEC and no error protection, we normalized the average powers for three detectors such that the packet loss probability for each detector are all equal to 10%, thus the average powers for three detectors are different, so comparing the performance among these detectors is unfair. We should compare the performance of systems with FEC/power allocation, pure FEC and no error protection. In addition, due to the same packet loss probability for three detectors in the systems with pure FEC and no error protection, we select the linear MMSE detector/receiver for these two systems as benchmark to compare. But for our FEC/power allocation system, we gave the performance curves for three detectors separately.

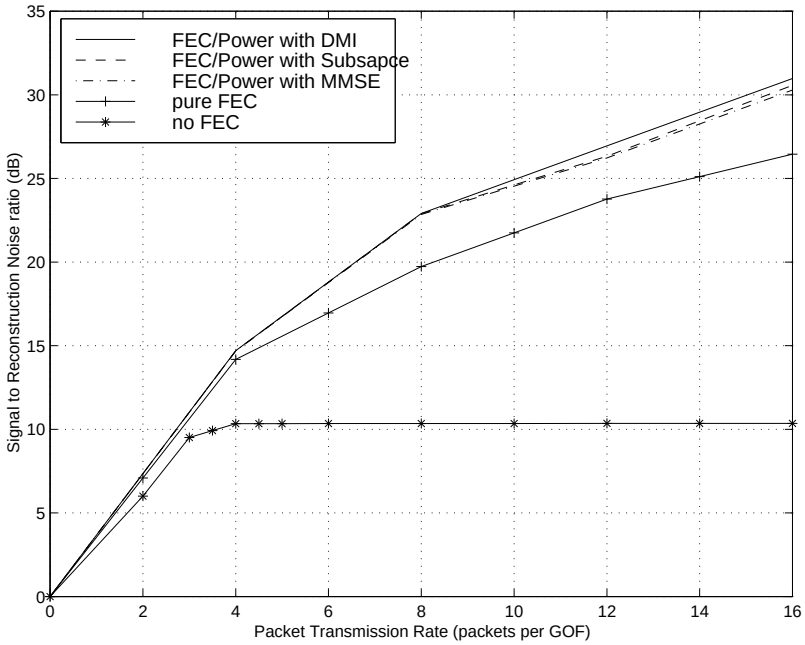


Fig. 4. Analytical results: Signal-to-reconstruction-noise ratio vs. packet transmission rate (packets per GOF).

Fig. 4 shows the peak signal to reconstruction noise ratio as a function of the packet transmission rate in packets per GOF. With no error protection, i.e.,

$(N_{\max}, K) = (8, 8)$, as the transmission rate increases, the system performance saturates, because with high probability there is a loss within the first few layers which renders subsequent layers useless, regardless of the number of transmitted layers. With UEP and uniform power allocation, the system performance improves significantly (as much as 16 dB). An additional performance improvement is achieved when joint optimization of rate and power allocation is performed, i.e., up to 5 dB at transmitting 16 packets per GOF. This demonstrates the superiority of the proposed joint optimization approach.

3. Simulations

In order to test the validity of our analytical results, we run simulations of the proposed system on real data transmitted over a simulated wireless environment. The color QCIF (144×176) sequence *foreman* is used in our experiments with a frame rate of 25 fps. We first partition the 256-frame sequence into eight GOFs and independently code each GOF using 3-D-SPIHT [6] to obtain an embedded bitstream; we then partition each embedded bit string into sequentially dependent packets containing 1000 bytes per packet. We assign the l-th packet from each embedded bit string to the l-th source layer. Thus, the duration of a GOF is $32/25 = 1.28$ s, while the transmission rate of each source layer is $8000/1.28 = 6.25$ Kbps. We produce up to 32 source layers for a total of up to 200Kbps for the source.

We compare the same systems as in the analysis part: one with no error protection, one with optimal UEP and last with our hybrid FEC/power allocation. They are based on the linear MMSE, blind DMI and blind subspace detectors. We evaluate each of these systems at four transmission rates: 50, 100, 150 and 200 Kbps. The end-to-end distortion is computed by averaging the distortion over each of the 256 frames of the test sequence, and averaging again over 10 independent transmission

trials.

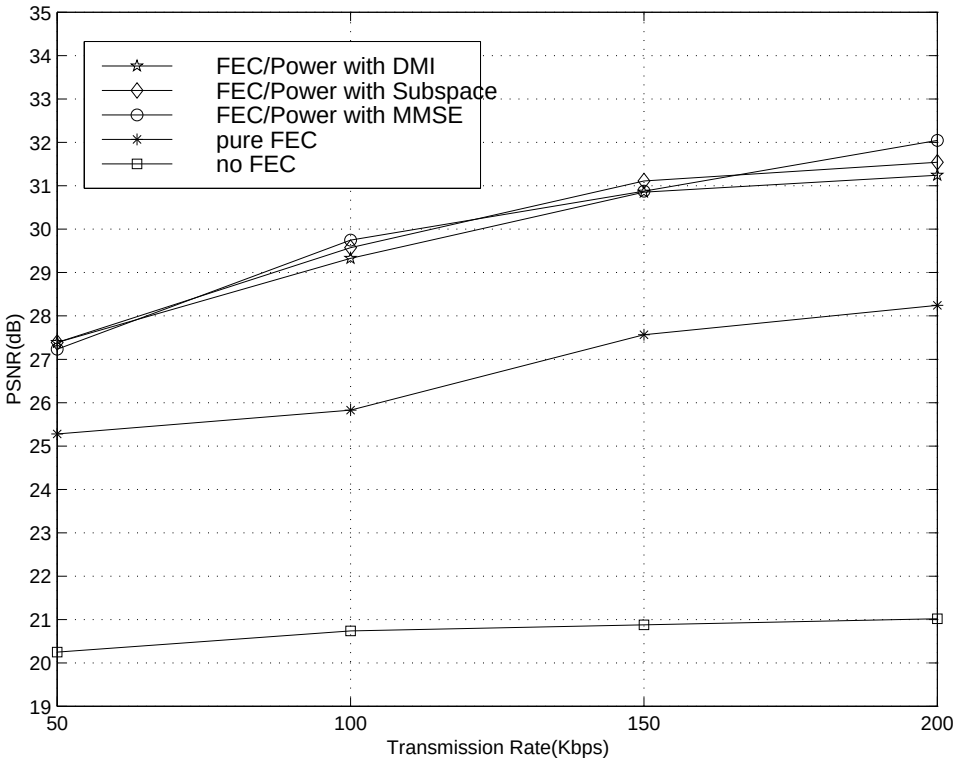


Fig. 5. PSNR vs. transmission rate for MMSE, DMI and subspace receivers.

Fig. 5 shows the peak signal to reconstruction noise ratio of each system at the above-mentioned four rates, respectively, for the Y component. Here, we normalized the average powers for three detectors such that the packet loss probability for each detector are all equal to 20%. The line at the bottom representing the system without error protection is $5-7$ dB worse than any of the other lines, those of which use protection of some kind. The middle line representing pure FEC with only rate optimization performs substantially better. The top three lines representing our hybrid FEC/power allocation systems with three receivers exceed pure FEC by up to 3.5 dB, confirming that power allocation indeed offers significant quality improvements in wireless video transmission. Our simulation results differ quantitatively from the analytical results in that the operational distortion-rate function of real encoded video

is not precisely $D(R) = A2^{-2R/B}$. As a matter of fact, the operational distortion-rate function varies from GOF to GOF.

In our system, the number of transmitted source and parity packets primarily depends on the packet loss ratios of different channels, which are determined by power allocation. Optimal power allocation at four transmission rate: 50, 100,150 and 200 Kbps for the three type of receivers are given in Figs. 6-8. Note that, at these rates, the optimization algorithm tell us that the number of CDMA channels the transmitter employs is 2, 3, 3 and 3, respectively. These results indicate that our joint optimization allows parsimonious use of wireless links at different bit rates. In addition, at each fixed bit rate, optimal power allocation results in unequal power level assignment.

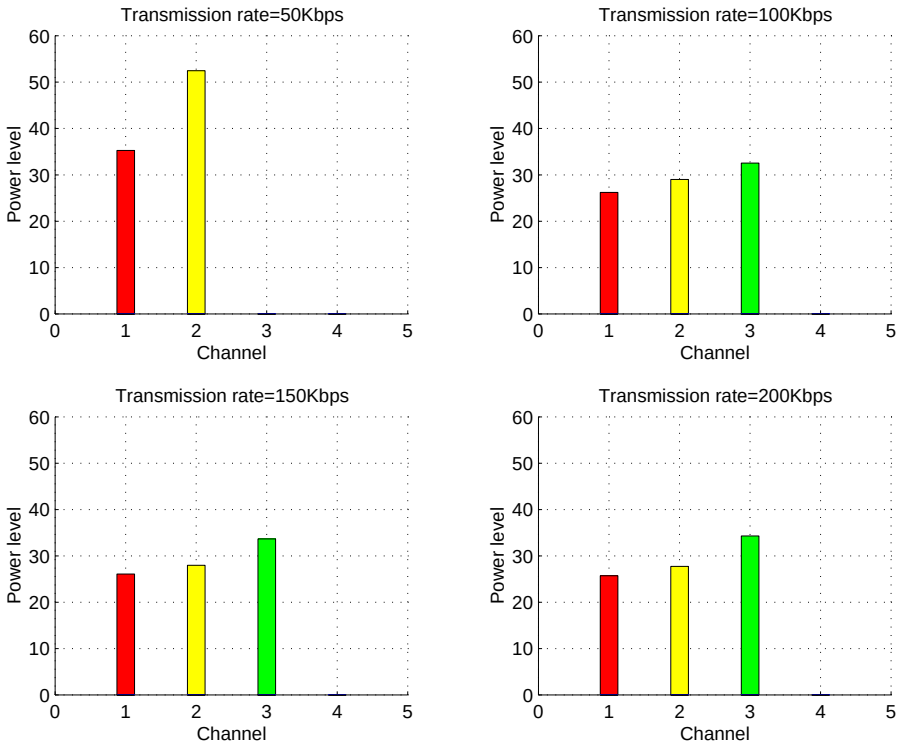


Fig. 6. Optimal power allocation for the linear MMSE receiver.

In Fig. 9, we depict the optimal source/parity packets distribution correspond-

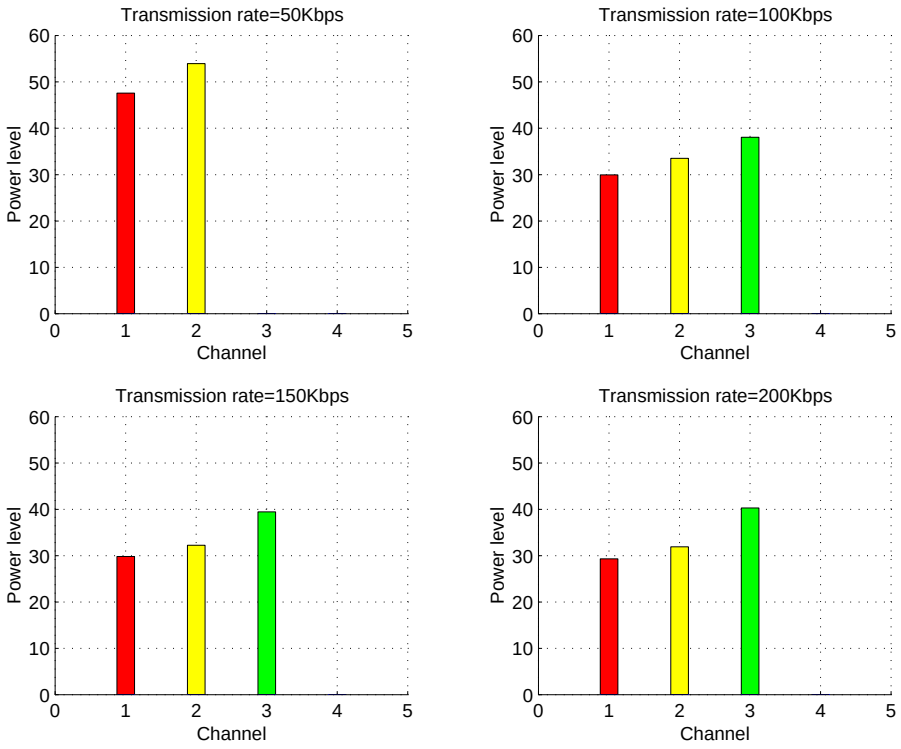


Fig. 7. Optimal power allocation for the blind DMI receiver.

ing to the optimal rate distribution in our exact linear MMSE simulation when the transmission rate is 200 Kbps (or 32 packets per GOF). A transmitter chooses to send a total of 25 source layers and 256 source/parity packets for $K = 8$ GOFs in this case. We notice from Fig. 9 that N_i is non-increasing within each channel. This is very reasonable given the sequential dependency that exists among source layers in the embedded 3-D SPIHT bitstream.

F. Conclusions

We have considered wireless video transmission over CDMA networks employing receivers with multiuser detection capability, by using the 3-D SPIHT video coder. Three types of multiuser receivers are considered, namely, the exact linear MMSE detector for the uplink, and the blind DMI detector as well as the blind subspace de-

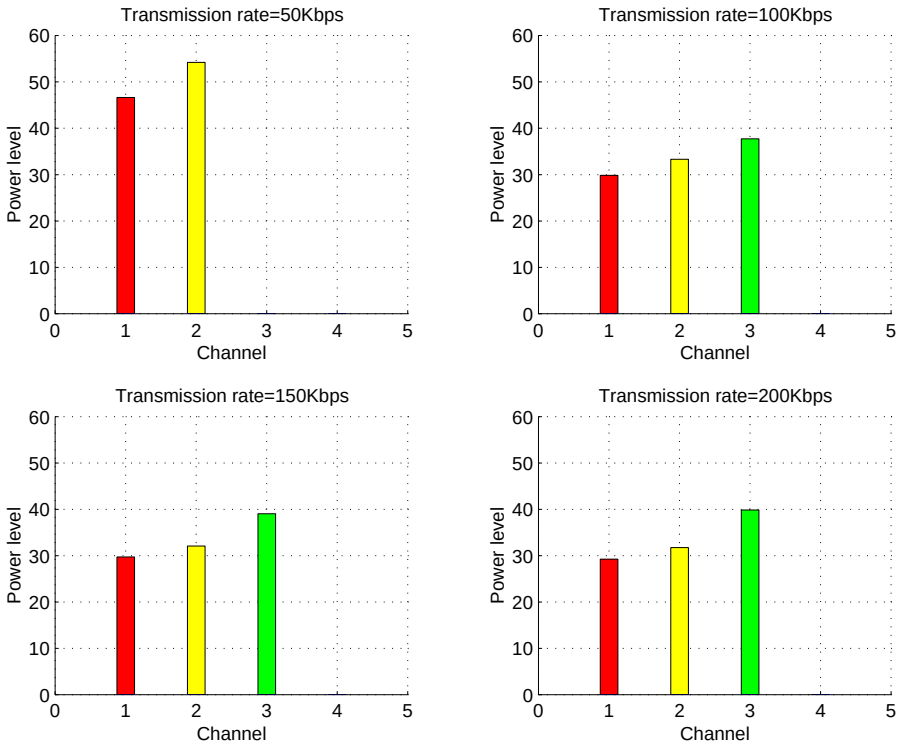


Fig. 8. Optimal power allocation for the blind subspace receiver.

tector for the downlink. We have proposed a scheme for joint optimization of channel coding rate allocation at different source layers, and the power allocation at different CDMA channels, to minimize the distortion on the received video data. Simulations show that the proposed joint optimal FEC/power allocation scheme offers a performance gain of up to 3.5 dB over the scheme with optimal FEC and equal power levels.

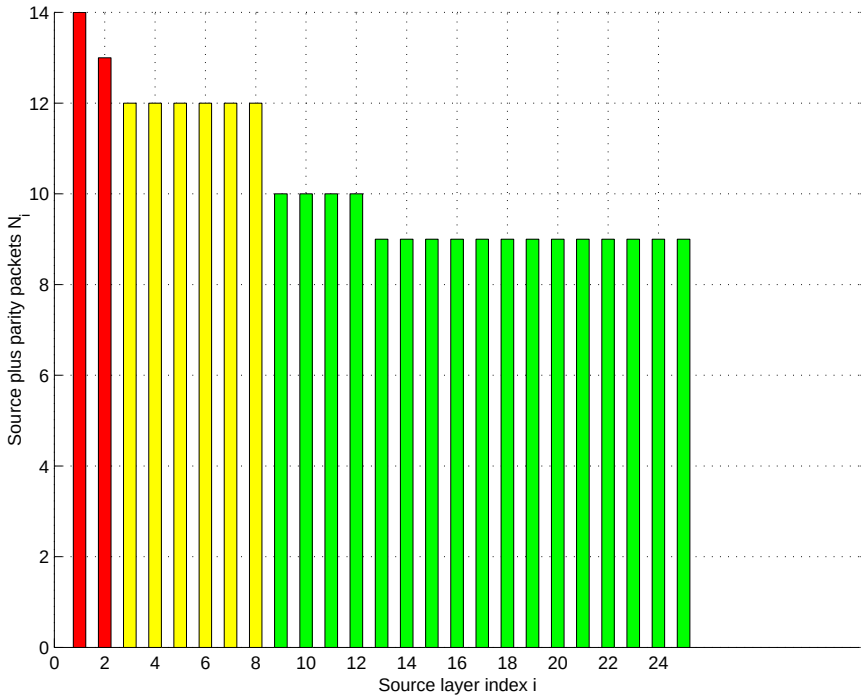


Fig. 9. Optimal rate allocation for the exact linear MMSE receiver when the transmission rate is 200 Kbps and the average power level is 21.97.

CHAPTER III

OPTIMAL RESOURCE ALLOCATION FOR WIRELESS VIDEO OVER CDMA NETWORKS*

A. Introduction

In this chapter, we consider transmitting layered and embedded source bitstream in terms of product channel code structure over wireless CDMA networks with multipath fading channels. Layered and embedded source coding is achieved using the three-dimensional embedded subband coding with optimal truncation (3-D ESCOT) [7] which has bit rate scalability and resolution scalability. A product channel code is obtained by the combination of a row code based on rate-compatible punctured convolutional code(RCPC) with cyclic redundancy check(CRC) error detection [17, 84], and a source-channel column code, i.e., systematic rate-compatible Reed-Solomon(RS) style erasure code [52]. The recently developed large-system analysis results for CDMA networks are used to characterize the performance of various receivers in multipath fading environment, which facilitates the formulation of our optimization problem [85]. We essentially map *one* layered video bitstream using *multiple* CDMA channels. Using 3-D ESCOT we layer the bitstream by resolution into multiple independently decodable layers corresponding to different resolutions. Next we packetize each layer of video bitstream in one product channel code structure, and then map one product channel code to one CDMA channel individually. In this work, given certain transmission rate and total transmission power over multiple CDMA channels, we transmit groups of product channel code packets over these multiple CDMA

*Parts of this chapter are ©2004 IEEE. Reprinted, with permission, from S. Zhao, Z. Xiong and X. Wang, "Optimal resource allocation for wireless video over CDMA networks," *IEEE Trans. on Mobile Computing*, vol. 3, no. 3 July-Sept. 2004.

channels. We address the problem of minimizing the total distortion of layered video bitstream subject to a total transmission rate and a total transmitted power level over multiple CDMA channels. By using the method of Lagrange multiplier, we get the optimal transmission rate allocation and power allocation vector over multiple CDMA channels. Given certain transmission rate for a particular channel and power allocation vector for multiple channels, we use a fast joint Reed-Solomon and RCPC rate algorithm to obtain the optimal Reed-Solomon source/parity distribution and RCPC code rate for one CDMA channel, i.e. one product channel code structure.

This work can be viewed as an extension of our previous work in Chapter II. The new contributions of this chapter include the following:

1. We use the 3-D ESCOT video bitstream which can be encoded into independently decodable layers. Such an independently decodable capability provides more system error resilience even during the deep fading period in wireless channel. At the same time, we may employ the optimal rate allocation algorithm separately for each layers (hence for each CDMA channel).

2. We use the product channel code framework in which the Reed-Solomon code is across different packets instead of within one packet, while the error detection ability of CRC and the rate-compatible property of RCPC make it easy for getting various video quality services. This new packetization and FEC coding method fully considers the characteristics of packet-switched networks which treat packets as equally important.

3. We treat the multipath fading channels and take into account effects of channel estimation error and various receiver structures.

4. We make use of the recently developed large-system performance results for

CDMA networks, which accurately characterize the average system performance under random fading and spreading.

All these new features make the framework proposed here suitable for transmitting video bitstreams robustly and for providing end-users with different QoS.

The remainder of this chapter is organized as follows: Section B introduces the performance of 3D ESCOT codec. Section C describes the large-system CDMA for multiuser receivers. Section D is devoted to channel coding. Section E describes the optimal rate and power allocation. Section F presents simulation results. Section G concludes the chapter.

B. Performance of 3D ESCOT Codec

Borrowing ideas from the EBCOT image coding algorithm [86], the original 3-D ESCOT algorithm was designed to achieve better coding performance for scalable wavelet video compression. Rate scalability is accomplished by independently coding wavelet coefficients in different subbands using bit plane coding. The main contribution in extending the 2-D EBCOT algorithm to 3-D ESCOT lies in 3-D context formation and modelling for arithmetic coding. In the following, we provide a high-level summary of the 3-D ESCOT algorithm and highlight its difference from the 3-D SPIHT algorithm [6]. Readers interested in the algorithmic details are referred to [7].

While the basic components of bit plane coding, namely, significance coding, sign coding and magnitude refinement, are the same in 3-D ESCOT and 3-D SPIHT, major differences between the two are as follows.

1) For significance coding, 3-D ESCOT forsakes the zerotree structure made so famous by the success of EZW [67] and SPIHT [66]. Instead, it completely relies on conditional arithmetic coding to exploit redundancies among significance bits. The

immediate advantage is that context formation for arithmetic coding in 3-D ESCOT does not have to be restricted to the rigid cubic structure imposed by zerotrees. The practice of not using zerotrees is supported by recent findings that show the coding gain due to parent-child dependencies in zerotree-based coders is much smaller than what was previously believed [87].

2) Compared to 3-D SPIHT, which applies adaptive arithmetic coding [88] on the significant bits only, 3-D ESCOT additionally uses high-order conditional arithmetic coding on the sign bits and refinement bits. Due to the use of L_2 metric in wavelet approximation, samples in all subbands except the lowest frequency one can be thought of being drawn from zero-mean random processes. The coefficient sign is equally probable to be positive or negative. The self entropy of the coefficient sign is thus at its maximum of one bit. Similar argument was made about the randomness of refinement bits in [67, 66]. Consequently, neither the sign bits nor the refinement bits are entropy coded in 3-D SPIHT. But the self entropy of the coefficient sign being one does not necessarily mean that the sign bits are uncompressible. In fact the high-order conditional entropy of the coefficient sign can be significantly lower than one bit as the sign patterns of wavelet coefficients often reflect the waveform structure of the input signal (e.g., less sign changes in smooth areas and more around image edges.) Applying high-order conditional arithmetic coding on the sign bits and refinement bits enables 3-D ESCOT to outperform 3-D SPIHT.

3) 3-D ESCOT introduces the concept of *fractional* bit planes. Using three passes – significance propagation, magnitude refinement and normalization – to process each bit plane in the natural raster scan order, 3-D ESCOT effectively partitions it into three non-overlapping fractional bit planes. The significance propagation pass, which involves significance coding and sign coding, processes coefficients that are highly likely to become significant in the current bit plane. It typically achieves a higher R-D

ratio than the magnitude refinement pass. The normalization pass finishes coding bits left over by the first two passes. This processing order of the three fractional bit planes thus follows the decreasing order of their perceived R-D significance levels. It ensures an R-D optimized embedded bitstream for each subband. Note that with context models built differently for significance coding, sign coding and magnitude refinement, processing significance coding and magnitude refinement in different fractional bit planes comes naturally in 3-D ESCOT.

4) Fractional bit plane coding in 3-D ESCOT ensures that the final bitstream is scalable with fine granularity at the fractional bit plane level. Coding each subband independently also offers flexibilities in the final bitstream formation to provide functionalities such as rate and resolution scalability. For instance, to achieve rate scalability, the classic equal slope criterion [89] is invoked for R-D optimal bit allocation among different subbands. The end result is that bitstreams generated from different fractional bit planes in all subbands are multiplexed together into one final layered bitstream according to the decreasing order of their R-D significance levels.

C. Large-System CDMA for Multiuser Receivers

Suppose that an RCPC code rate r_i is used for a packet protection and suppose that the channel is modelled as the concatenation of a multipath fading channel plus a wireline packet erasure channel with packet loss probability q, then the probability that a packet is considered to be lost is $\tilde{q} = q + (1-q)P_{r_i}$, where P_{r_i} is the probability that a packet protected with rate r_i cannot be correctly decoded with the RCPC decoder. Thus if N packets are protected by the same RCPC code rate r_i, the probability that n out of N packets are lost, denoted by $P_{n/N}$, is expressed as

$$P_{n/N} = \binom{N}{n} \tilde{q}^n (1-\tilde{q})^{N-n}. \tag{3.1}$$

Obviously $P_{n/N}$ is determined completely by P_{r_i} and q. Where the packet loss probability q for a wireline packet erasure channel is modelled as a constant. However P_{r_i} is related to the BER over wireless multipath fading channels, furthermore the transmit power level of wireless channels has important effects to the BER. Now let us consider a large-system CDMA channel in the rest of this section.

Consider a K-user baseband CDMA system with spreading gain N_c signaling over multipath fading channels. The received signal at time m can be written as

$$\boldsymbol{y}[m] = \sum_{k=1}^{K} b_k[m] \sum_{l=1}^{L} a_{kl}[m]\boldsymbol{s}_{kl}[m] + \boldsymbol{n}[m], \tag{3.2}$$

where $k \in \{1, \cdots, K\}$ indexes the multiple users, and $l \in \{1, \cdots, L\}$ indexes the paths of each user, $a_{kl}[m]$ is the channel gain for path l of user k over symbol period m, $b_k[m]$ is the transmitted symbol of user k over symbol period m, $\boldsymbol{s}_{kl}[m]$ is the signature sequence for path l of user k over symbol period m and is assumed to be an N_c-dimensional column vector with independent and identically distributed (i.i.d.) elements each with zero mean and variance $1/N_c$, and $\boldsymbol{n}[m]$ is a circularly symmetric complex white Gaussian noise with $\boldsymbol{E}[\boldsymbol{n}[m]] = \boldsymbol{0}$ and $\boldsymbol{E}[\boldsymbol{n}[m]\boldsymbol{n}^H[m]] = \sigma^2\boldsymbol{I}$. We assume throughout that the delay spread of the channel is small compared to the symbol time so that inter-symbol interference can be neglected. Note that the assumption that we know $\boldsymbol{s}_{kl}$ means that we implicitly assume knowledge of the timing of resolvable path l of user k.

Let $\boldsymbol{s}_k = [\boldsymbol{s}_{k1}, \cdots, \boldsymbol{s}_{kL}]$ and $\boldsymbol{a}_k = [a_{k1}, \cdots, a_{kL}]^T$ denote the sequence and channel gains corresponding to user k, and therefore $\boldsymbol{S} = [\boldsymbol{s}_1, \cdots, \boldsymbol{s}_K]$, $\boldsymbol{A} = \mathrm{diag}(\boldsymbol{a}_1, \cdots, \boldsymbol{a}_K)$, and $\boldsymbol{b} = [b_1, \cdots, b_K]^T$. The received signal over the symbol of interest then expressed as

$$\boldsymbol{y} = \boldsymbol{S}\boldsymbol{A}\boldsymbol{b} + \boldsymbol{n}.$$

The multipath multiuser receivers and corresponding analysis can be described as the data estimator and the channel estimator. The data estimator incorporates information from other symbol intervals only through the coupling with the channel estimator. The design and analysis of the data estimator is based on the assumption that the channel is statistically characterized by the mean and covariance structure supplied by the channel estimator. Assume that the error variances of all paths are equal and that the path estimates are uncorrelated and consider the limiting situation where $N_c \to \infty$ and $K \to \infty$ with $\alpha \triangleq K/N_c$ and L fixed, then the output signal-to–interference ratio of several receivers are obtained in [85] and are summarized as follows.

- **LMMSE Receiver:** letting $\boldsymbol{D} = \boldsymbol{E}_d[\boldsymbol{A}\boldsymbol{A}^H] = \text{diag}(\boldsymbol{E}_d[\boldsymbol{a}_1\boldsymbol{a}_1^H], \cdots, \boldsymbol{E}_d[\boldsymbol{a}_K\boldsymbol{a}_K^H])$, thus the LMMSE receiver for user 1 can be expressed as

$$\boldsymbol{c} = (\boldsymbol{S}\boldsymbol{D}\boldsymbol{S}^H + \sigma^2\boldsymbol{I})^{-1}\boldsymbol{s}_1\bar{\boldsymbol{a}}_1,$$

  then for LMMSE data estimation, we will have the approximate SIR results for the estimate $z = \boldsymbol{c}^H\boldsymbol{y}$ of b_1. In a multiple-path fading channel the gain for path l of user k is characterized by the estimate $\bar{a}_{kl}$ and error variance ξ_k^2. The SIR for the LMMSE receiver of user 1 can be expressed approximately as

$$\text{SIR}_1 = \frac{\sum_{l=1}^{L} |\bar{a}_{1l}|^2\beta_d}{1 + \xi_1^2\beta_d}, \tag{3.3}$$

  where β_d is the solution to the following fixed point equation

$$\beta_d = \Big[\sigma^2 + \frac{1}{N_c}\sum_{k=2}^{K}\Big((L-1)I(\xi_k^2, \beta_d) + I\Big(\sum_{l=1}^{L}|\bar{a}_{kl}|^2 + \xi_k^2, \beta_d\Big)\Big)\Big]^{-1}$$

  and $I(p, \beta) = \frac{p}{1+p\beta}$. Two special cases emerge

- when the channel is perfectly known, $\bar{a}_{kl} = a_{kl}$ and $\xi_k^2 = 0$; thus $\mathrm{SIR}_1 = |a_1|^2\beta_d$ and $\beta_d = \left[\sigma^2 + \frac{1}{N_c}\sum_{k=2}^{K} I(|a_k|^2, \beta_d)\right]^{-1}$.
- when in a single-path fading channel, $L = 1$; thus $\mathrm{SIR}_1 = \frac{|\bar{a}_1|^2\beta_d}{1+\xi_1^2\beta_d}$ and $\beta_d = \left[\sigma^2 + \frac{1}{N_c}\sum_{k=2}^{K} I(|\bar{a}_k|^2 + \xi_k^2, \beta_d)\right]^{-1}$.

Consider joint estimation of the channel parameters of all users over an *estimation window* of τ symbols, where the channel coherence time, over which the channel is essentially constant, is greater than τ symbol intervals. Based on periodically sending training data, τ would typically be chosen to be small fraction of the channel coherence time so that the training overhead is not too large. Thus the MSE for any path of any user converges almost surely as $N_c \to \infty$ to the nonrandom [85]

$$\xi^2 = \frac{\frac{\bar{p}}{L}}{1 + \frac{\bar{p}}{L}\beta_c}, \tag{3.4}$$

where β_c satisfies the equation $\beta_c = \left[\frac{\sigma^2}{\tau} + \frac{\alpha L}{\tau}\frac{\frac{\bar{p}}{L}}{1+\frac{\bar{p}}{L}\beta_c}\right]^{-1}$ and $\bar{p}$ is the average power of user 1 of interest.

- **Decorrelating Receiver:** For a decorrelating multipath-combining receiver, the first stage of the receiver corresponds to a decorrelator which treats every path of every user as a unique interferer. The output is given by

$$\overline{\boldsymbol{z}} = \boldsymbol{A}\boldsymbol{b} + \overline{\boldsymbol{n}},$$

where $\overline{\boldsymbol{n}}$ has covariance matrix $\sigma^2(\boldsymbol{S}^H\boldsymbol{S})^{-1}$. The second stage (for user 1) takes the first L components of $\overline{\boldsymbol{z}}$ and combines them taking into account the covariance structure of the noise and of the channel estimate, then the vector

$\overline{\boldsymbol{z}}_1$ consisting of the L decorrelating outputs of user 1 can be expressed

$$\overline{\boldsymbol{z}}_1 = \boldsymbol{a}_1 \boldsymbol{b} + \overline{\boldsymbol{n}}_1,$$

where $\overline{\boldsymbol{n}}_1$ has covariance $\boldsymbol{R}$ equal to the first $L \times L$ sub-block of $\sigma^2(\boldsymbol{S}^H\boldsymbol{S})^{-1}$. Then an LMMSE combiner for estimating b_1 from $\overline{\boldsymbol{z}}_1$ forms the decision statistic

$$z = \frac{\overline{\boldsymbol{a}}_1^H(\Xi_1 + \boldsymbol{R})^{-1}\overline{\boldsymbol{z}}}{1 + \overline{\boldsymbol{a}}_1^H(\Xi_1 + \boldsymbol{R})^{-1}\overline{\boldsymbol{a}}_1},$$

where $\Xi_1 = \boldsymbol{E}_d[(\boldsymbol{a}_1 - \overline{\boldsymbol{a}}_1)(\boldsymbol{a}_1 - \overline{\boldsymbol{a}}_1)^H]$, thus the receiver for user 1 does not require any information about the channels of the other users apart from the timing of the various resolvable multipath components. If $\alpha L = LK/N_c < 1$ in a multiple-path fading channel, then the approximate SIR for the decorrelating receiver of user 1 is also given by (3.3) with $\beta_d = \frac{1-\alpha L}{\sigma^2} = \left[\sigma^2 + \alpha L \frac{\sigma^2}{1-\alpha L}\right]^{-1}$. Two special cases emerge

- when the channel is known perfectly, $\bar{a}_{kl} = a_{kl}$ and $\xi_k^2 = 0$; thus $\mathrm{SIR}_1 = |a_1|^2\beta_d$ and $\beta_d = 1/\sigma^2$.
- when in a single-path fading channel, $L = 1$; thus $\mathrm{SIR}_1 = \frac{|\bar{a}_1|^2\beta_d}{1+\xi_1^2\beta_d}$ and $\beta_d = \frac{1-\alpha}{\sigma^2}$.

- **Matched Filter Receiver:** The matched filter receiver is based only on information about the desired user. The matched filter we consider here is simply $\boldsymbol{c} = \sum_{l=1}^{L} \bar{a}_{1l}\boldsymbol{s}_{1l}$ and leads to the following result for large systems. In a multiple-path fading model, the approximate SIR of user 1 is also given by (3.3) with $\beta_d = \left[\sigma^2 + \frac{1}{N_c}\sum_{k=2}^{K}\left((L-1)I(\xi_k^2) + \sum_{l=1}^{L}|\bar{a}_{kl}|^2 + \xi_k^2\right)\right]^{-1}$, where $I(p) = p$. Two special cases emerge

  - when the channel is perfectly known, $\bar{a}_{kl} = a_{kl}$ and $\xi_k^2 = 0$; thus $\mathrm{SIR}_1 =$

$|a_1|^2\beta_d$ and $\beta_d = \left[\sigma^2 + \frac{1}{N_c}\sum_{k=2}^{K}|a_k|^2\right]^{-1}$.

– when in a single-path fading channel, $L = 1$; thus $\mathrm{SIR}_1 = \frac{|\bar{a}_1|^2\beta_d}{1+\xi_1^2\beta_d}$ and $\beta_d = \left[\sigma^2 + \frac{1}{N_c}\Big(\sum_{k=2}^{K}(|\bar{a}_k|^2 + \xi_k^2)\Big)\right]^{-1}$.

Now let us compare the three receivers intuitively. Notice that for LMMSE receiver, the overall effect of interferer k is the same as the interference that would result in the single-path fading case from $L-1$ users with power ξ_k^2 and one user with power $\sum_{l=1}^{L}|\bar{a}_{kl}|^2 + \xi_k^2$. As the uncertainty increases, an interferer moves from looking like a single-path power interferer, to looking like L separate interferers with power reduced by a factor of L. Also notice that in decorrelating receiver, the SIR is independent of the powers of the interferers. The effective interference is $\frac{L}{\beta_d}$ for each interferer, and does not depend on its power. Furthermore notice that for matched filter receiver, $I(p) = p$ is the effective interference of user k, so the effective interference is linear in the interferer power which should be contrasted to the LMMSE receiver, for which $I(p,\beta) = \frac{p}{1+\beta p}$ at normalized β. The performance of the matched filter will become arbitrarily bad as the power of an interferer is increased, while for fixed β, the LMMSE effective interference for a high-power user approaches $1/\beta$. For the matched filter, L low-power interferers have exactly the same impact as one interferer with the same total power.

With the SIR expressions and the estimated channel values $\bar{a}_{kl}$ as well as its variance ξ_k^2 in hand for the three receivers, it is easy to get the BER expression for user 1, indeed, it is shown in [75] that the output of a linear multiuser detector is well approximated by a Gaussian distribution. Thus the bit error rate can be expressed as $P_{b,1} = Q(\sqrt{\mathrm{SIR}_1})$, where $Q(x) \triangleq \frac{1}{\sqrt{2\pi}}\int_x^\infty \exp\left(-\frac{t^2}{2}\right)\mathrm{d}t$.

D. Channel Coding

1. Product Code

In this section, we discuss the packetization of the embedded source code by using a product code. The row code of the product code is a concatenation of an outer CRC code and an inner RCPC code, while its column code is a systematic RS code. We use equal error protection(EEP) along the rows and unequal error protection(UEP) along the columns. We also put the earliest symbols of the embedded bitstream in the first columns as shown in Figure 10, which offers a better reconstruction quality for video sequence. We assume that the resulting joint source-channel product code is sent as N packets of L_p symbols each, which actually gives the transmission rate constraint of NL_p.

The embedded source code is first protected by the RS codes, then the CRC symbols are added to each row, finally each row is encoded with the same RCPC code. Let $\mathcal{R} = \{r_1, r_2, \ldots, r_M\}$, with $r_1 < \cdots < r_M$, be the finite set of available RCPC code rates. For certain RCPC code with rate $r_i \in \mathcal{R}$, we denote by L_{r_i} the sum of the number of source symbols and RS parity symbols used in a packet(i.e., a row) protected by RCPC code rate r_i. Thus, we have L_{r_i} source segments $S_1, \ldots, S_{L_{r_i}}$, where segment S_j, $1 \leq j \leq L_{r_i}$, consists of $m_j \in \{1, \ldots, N\}$ source symbols that are protected by $f_j = N - m_j$ RS symbols as shown in Fig. 10.

The N packets are composed of a product code structure and it can be sent over a CDMA channel. At the receiver, if the CRC detects an error, then the packet is considered lost (we suppose that all errors can be detected); otherwise each received packet is decoded with the RCPC decoder first. After all packets of a product code arrive, the RS decoding can be implemented. Suppose now that n out of N packets are lost, then the RS codes ensure that all segments that contain at most $N - n$

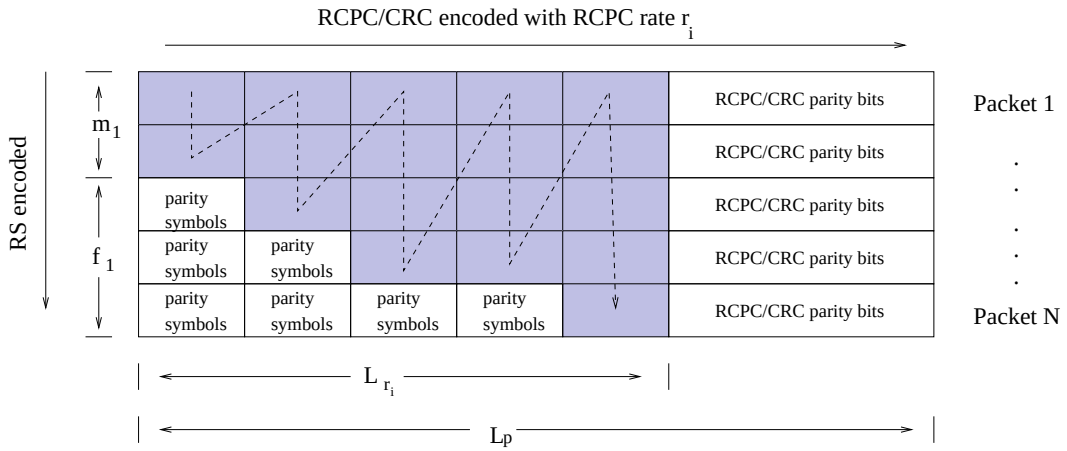


Fig. 10. Product code structure. There are $N = 5$ packets. Every shaded cell is a source symbol. The embedded source bitstream flows along the direction of the dashed line.

source symbols can be recovered. By adding the constraint $f_1 \geq f_2 \geq \cdots \geq f_{L_{r_i}}$, we guarantee that the receiver can decode at least the first j segments whenever at most f_j packets are lost. We denote by $\boldsymbol{F}_i$ the set of L_{r_i}-tuples $(f_1, \ldots, f_{L_{r_i}})$ such that $f_1 \geq f_2 \geq \cdots \geq f_{L_{r_i}}$ and $f_j \in \{0, \ldots, N-1\}$ for $j = 1, \ldots, L_{r_i}$. Let $\boldsymbol{f} = (f_1, \ldots, f_{L_{r_i}})$. Thus we can apply the two packet loss channel models in Section C to the product code structure.

Let ϕ denote the operational distortion-rate function of the source coder and let X be is the number of packets lost in such product code scheme. The expected distortion of L_{r_i}-segment product code protection is

$$D(\boldsymbol{f}, r_i) = \sum_{l=0}^{L_{r_i}} P_l(\boldsymbol{f})\phi(t_l), \tag{3.5}$$

where $t_l = \sum_{j=1}^{l} m_j = lN - \sum_{j=1}^{l} f_j$ for $l = 1, \ldots, L_{r_i}$, denotes the sum of source symbols when there are l source segments received correctly. $P_0(\boldsymbol{f}) = P(X > f_1)$ is the probability that none of the L_{r_i} source segments $S_1, \ldots, S_{L_{r_i}}$ can be recovered correctly by Reed-Solomon decoding; $P_l(\boldsymbol{f}) = P(f_{l+1} < X \leq f_l)$ for $l = 1, \ldots, L_{r_i} - 1$

is the probability that at most l out of the L_{r_i} source segments can be recovered correctly by Reed-Solomon decoding; and $P_{L_{r_i}}(\boldsymbol{f}) = P(X \leq f_{L_{r_i}})$ is the probability that all L_{r_i} source segments can be recovered correctly by Reed-Solomon decoding. The case of $t_0 = 0$ means that no source symbol is received correctly. Notice that for $l = 1, \ldots, L_{r_i} - 1$, we have $P_l(\boldsymbol{f}) = 0$ if $f_l = f_{l+1}$; and $P_l(\boldsymbol{f}) = \sum_{n=f_{l+1}+1}^{f_l} P_{n/N}$ if $f_l \neq f_{l+1}$, where $P_{n/N}$ is the probability that n out of N packets are lost, given by (3.1) in Section C.

An optimal product code $(r^*, \boldsymbol{f}^*)$ is given by an RCPC code rate $r^* \in \mathcal{R}$ and an L_{r^*}-segment RS protection that is the solution to the following minimization problem

$$(r^*, \boldsymbol{f}^*) = \arg\min_{r_i \in \mathcal{R}} \min_{\boldsymbol{f} \in \boldsymbol{F}_i} D(\boldsymbol{f}, r_i) = \arg\min_{r_i \in \mathcal{R}} \min_{\boldsymbol{f} \in \boldsymbol{F}_i} \sum_{l=0}^{L_{r_i}} P_l(\boldsymbol{f}) \phi(t_l). \tag{3.6}$$

Solving problem (3.6) by brute-force is impractical because the number of possible product codes is $\sum_{i=1}^{M} \binom{L_{r_i}+N-1}{L_{r_i}}$. We employ a fast method that finds an approximately optimal solution to problem (3.6). We do not try to minimize (3.5) for each RCPC code rate. We observe that the total number of parity protection symbols (RS and RCPC) for the product code corresponding to a distortion-optimal RS protection with the smallest expected distortion(among all RCPC code rates) is greater than that corresponding to a rate-optimal RS protection with the largest expected rate. Therefore we start with the rate-optimal RS protection that gives the largest expected rate and try to improve the associated product code by progressively increasing the number of protection symbols. This is done by alternatively applying the local search algorithm in [90] and decreasing the RCPC code rate. In the worst case, the algorithm computes $(N-1)L_{r_i} + 1$ times the cost function (3.5) for each $r_i \in \mathcal{R}$. The fast product code algorithm provides a near-optimal product code for wireless channels and is much faster than the original approach of [23], independent

of the packet length. Its very low memory requirements and a linear worst case time complexity make it suitable for real-time applications. The details of the fast product code algorithm are found in [91].

2. Multi-channel Product Code Framework

In this section, we outline the multi-channel product code FEC coding scheme. At first we form several independent decodable video source layers from a 3-D ESCOT embedded video bitstream, these layers are encoded from different subbands, for example, Fig.11(a) illustrates four subbands encoding for 2-D image; then using the joint rate and power allocation optimization algorithm, we get the optimal power allocation vector and optimal transmission rate allocation for multiple CDMA channels, as well as the optimal Reed-Solomon source/parity rate allocation and optimal RCPC code protection for each channel. One independent decodable layer corresponds to a product code. All packets of one product code are transmitted over one CDMA channel. After multiple product codes are formed for one embedded bitstream, we transmit all packets of these product codes over multiple CDMA channels.

In this chapter, we assume that a total of four CDMA channels are used to transmit the four groups of product code data packets for one embedded video bitstream. Fig.11(b) illustrates the framework we present. For example, suppose we use the 3-D ESCOT coding bitstream with the GOF size 16 frames and a fixed frame rate 30 fps, and want to transmit $N = 200$ packets with packet size 48 bytes during one GOF period and thus the transmission rate is 144 Kbps. By using the joint rate and power allocation optimization algorithm, we may send video bitstream with channel 1 transmitting the first 60 packets, channel 2 transmitting the next 25 packets, channel 3 transmitting the next 50 packets, and channel 4 transmitting the last 65 packets. Note that for traditional voice communication in CDMA systems, each user occupies

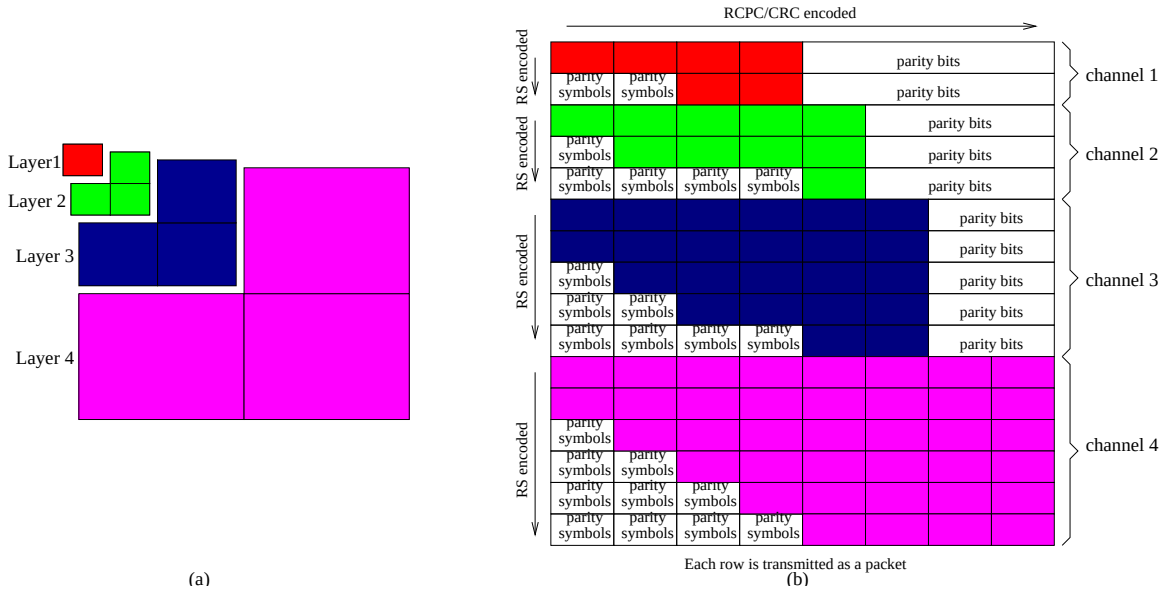


Fig. 11. (a) Layered image coding by resolution. (b) The proposed multi-channel product code structure multimedia transmission over CDMA networks.

one channel. However, in multimedia communications, each user may occupy more than one channel to transmit data (e.g., here the video user occupies four channels). Nevertheless, we use the terms *user* and *channel* interchangeably from now on.

The transmitter now has several groups of product code packets to send. The transmitter buffers frames as they arrive. When a block of groups of frames(GOF) is accumulated, it encodes the GOFs and packetizes the resulting independently decodable embedded layers. For a total transmission rate and a total power level over multiple CDMA channels, the transmitter chooses to get the optimal power level allocation vector and transmission rate allocation vector for multiple CDMA channels, and RS source/parity allocation and RCPC rate for each product code, as illustrated in Fig. 11(b). The receiver instantly recovers as many source symbols as possible from the received packets and decodes them for a block of GOFs. Playback begins after one block of GOFs of coding delay.

E. Optimal Rate and Power Allocation

1. Problem Formulation

Let $\overline{K}$ be the number of CDMA channels used for transmitting a video bitstream (in this work, $\overline{K} = 4$). Let $\boldsymbol{N} = (N_1, \ldots, N_{\overline{K}})$ be the vector for the number of packets for $\overline{K}$ product code structures, thus the transmission rate vector will be $\boldsymbol{N}L_p$ for $\overline{K}$ product code structures. Denote $\boldsymbol{p} = [p_1, p_2, \cdots, p_{\overline{K}}]$ as the *power level allocation vector* for the transmitted powers of the $\overline{K}$ channels used for transmitting the video bitstream. Here we assume that the powers of other users in the network (i.e., $p_{\overline{K}+1}, \cdots, p_K$) are fixed, each of those channels among $k = \overline{K}+1, \ldots, K$ is assigned the average power level over all K channels given a total power level. One of our objective is to optimize the power allocation $\boldsymbol{p}$ vector subject to a total power constraint such that the distortion is minimized. Note that from Section C, the bit error rate, and therefore the packets lost probability $P_{n/N}$ in a channel, depends on the power levels of all channels, not just the power of that particular channel. Thus by introducing the power allocation vector $\boldsymbol{p}$ over $\overline{K}$ CDMA channels, we can get the expected distortion expressions for $\overline{K}$ product code structure based on (3.5)

$$D(\boldsymbol{f}_k, r_{i_k}, \boldsymbol{p}) = \sum_{l=0}^{L_{r_{i_k}}} P_l(\boldsymbol{f}_k, \boldsymbol{p})\phi(t_{lk}), \tag{3.7}$$

where $k = 1, \ldots, \overline{K}$ index CDMA channels. Because of the independence of $\overline{K}$ source layers, the total expected distortion for $\overline{K}$ channels can be expressed

$$\sum_{k=1}^{\overline{K}} D(\boldsymbol{f}_k, r_{i_k}, \boldsymbol{p}) = \sum_{k=1}^{\overline{K}} \sum_{l=0}^{L_{r_{i_k}}} P_l(\boldsymbol{f}_k, \boldsymbol{p})\phi(t_{lk}). \tag{3.8}$$

We can now formulate the combined optimization problem of power allocation and transmission rate allocation for $\overline{K}$ channels, as well as product code rate allocations

for each of $\overline{K}$ channels. Given $\overline{K}$ channels for transmitting one video bitstream, we consider the following constraint optimization problem:

$$\min_{\boldsymbol{N},\boldsymbol{p}} \Big(\sum_{k=1}^{\overline{K}} \min_{r_{i_k} \in \mathcal{R}} \min_{\boldsymbol{f}_k \in \mathcal{F}_{i_k}} D(r_{i_k}, \boldsymbol{f}_k, \boldsymbol{p}) \Big) \text{ subject to } \sum_{k=1}^{\overline{K}} N_k L_p \leq \mathbb{R} \text{ subject to } \sum_{k=1}^{\overline{K}} p_k \leq \mathbb{P}, \quad (3.9)$$

where $\mathbb{P}$ and $\mathbb{R}$ are a total transmitted power and a total transmission rate over $\overline{K}$ CDMA channels , respectively.

2. Joint Rate and Power Allocation Algorithm

One way to solve the above constraint optimization (3.9) to get the optimal transmission rates $\boldsymbol{N}L_p$ and power allocation vector $\boldsymbol{p}$ for $\overline{K}$ CDMA channels is by finding them that minimizes the Lagrangian

$$J(\boldsymbol{N}; \boldsymbol{p}; \lambda_1, \lambda_2) = \sum_{k=1}^{\overline{K}} D(r_{i_k}, \boldsymbol{f}_k, \boldsymbol{p}) + \lambda_1 \sum_{k=1}^{\overline{K}} N_k L_p + \lambda_2 \sum_{k=1}^{\overline{K}} p_k. \quad (3.10)$$

The solution to this problem is completely characterized by the $\overline{K}$ optimal product codes. Obviously, we can solve (3.10) by using an iterative approach that is based on the method of alternating variables for multivariable minimization, which we used in our previous work [31]. The objective function $J(N_1, \ldots, N_{\overline{K}}; p_1, \ldots, p_{\overline{K}})$ in (3.10) is minimized one variable at a time, keeping the other variables constant, until convergence. To be specific, let $\boldsymbol{N}^{(0)}$ be any initial number of packets for $\overline{K}$ product code structures. Notice that because of the length of packet is constant L_p, the initial transmission rate vector will be $\boldsymbol{N}^{(0)}L_p$ for $\overline{K}$ product code structures. Let $\boldsymbol{p}^{(0)}$ be any initial power allocation vector. Let $\boldsymbol{N}^{(t)} = (N_1^{(t)}, \ldots, N_{\overline{K}}^{(t)})$ and $\boldsymbol{p}^{(t)} = (p_1^{(t)}, \cdots, p_{\overline{K}}^{(t)})$ be determined for $t = 1, 2, \cdots$, as follows. Select one component $x \in \{N_1, \cdots, N_{\overline{K}}, p_1, \cdots, p_{\overline{K}}\}$ to optimize at step t. This can be done in a round-robin style. Then for $N_k \neq x$, let $N_k^{(t)} = N_k^{(t-1)}$, or for $p_k \neq x$, let $p_k^{(t)} = p_k^{(t-1)}$. If $x = N_k$,

then we perform the following rate optimization

$$N_k^{(t)} = \arg\min_{N_k} J\left(N_1^{(t)}, \cdots, N_{k-1}^{(t)}, N_k, N_{k+1}^{(t)}, \cdots, N_{\overline{K}}^{(t)}, p_1^{(t)}, \cdots, p_{\overline{K}}^{(t)}\right), \quad (3.11)$$

and if $x = p_k$, then we perform the following power optimization

$$p_k^{(t)} = \arg\min_{p_k} J\left(N_1^{(t)}, \cdots, N_{\overline{K}}^{(t)}, p_1^{(t)}, \cdots, p_{k-1}^{(t)}, p_k, p_{k+1}^{(t)}, \cdots, p_{\overline{K}}^{(t)}\right). \quad (3.12)$$

During each iterative step, i.e. given a fixed power allocation vector $\boldsymbol{p}$ and a fixed transmission rate allocation $\boldsymbol{N}L_p$, we resort to use the fast product code optimization algorithm[90] to get the expected distortions over $\overline{K}$ channels and then the total expected distortion for one video bitstream. The summary of the overall algorithm is as follows.

1. Start with an initial values of $N_1^{(0)}, \ldots, N_{\overline{K}}^{(0)}$, $p_1^{(0)}, \ldots, p_{\overline{K}}^{(0)}$.

2. For each channel $k = 1, \ldots, \overline{K}$, employ the following procedures to get the optimal Reed-Solomon protection $\boldsymbol{f}_k^*$ and optimal RCPC protection r_k^*, and then get the distortion-optimal solution for each channel.

   (a) For each code rate $r_i \in \mathcal{R} = (r_1, \ldots, r_M)$, compute a rate-optimal solution to get the Reed-Solomon protection $\boldsymbol{f}_{k_i}$.

   (b) Determine the RCPC code rate $r_{k_m} \in (r_1, \ldots, r_M)$ which has the largest expected received source rate for this product code with Reed-Solomon protection $\boldsymbol{f}_{k_i}$.

   (c) Find a Reed-Solomon distribution $\boldsymbol{f}_k^*$ which has stronger protection (i.e. more parity symbols) than the rate-optimal solution $\boldsymbol{f}_{k_i}$.

   (d) Calculate the distortions for the product codes with Reed-Solomon protection $\boldsymbol{f}_k^*$ and those $r_i \in \mathcal{R}$ such that $i < r_{k_m}$, find the minimum distortion

and its corresponding r_k^*.

(e) Let $\boldsymbol{f}_{k_i} = \boldsymbol{f}_k^*$, go to Step (b) and repeat this procedure until r_{k_m} in Step (d) is the smallest RCPC code rate r_1.

3. Using the iterative procedure as above in (3.11) and (3.12) to get the optimal $J(N_1^{(t)}, \ldots, N_{\overline{K}}^{(t)}; p_1^{(t)}, \ldots, p_{\overline{K}}^{(t)})$ at step t.

4. Repeat Step 2 and 3 until $J(N_1, \ldots, N_{\overline{K}}; p_1, \ldots, p_{\overline{K}})$ converges.

Convergence is guaranteed, because $J(N_1, \ldots, N_{\overline{K}}; p_1, \ldots, p_{\overline{K}}; \lambda_1, \lambda_2)$ is nonincreasing and bounded below. The specified transmission rate $\mathbb{R}$ and the specified total power level $\mathbb{P}$ can be achieved by adjusting the Lagrange multiplier λ_1 and λ_2, respectively. Higher λ_1 will result in a lower transmission rate; lower λ_1 will result in a higher transmission rate. Similarly, higher λ_2 will result in a lower total power level; lower λ_2 will result in a higher total power level. Therefore, by adjusting λ_1, the overall rate constraint $\mathbb{R}$ can be met, and by adjusting λ_2, the overall power constraint $\mathbb{P}$ can also be met. Thus by using both Lagarange algorithm and fast product code algorithm we finally get the optimal transmission rates $\boldsymbol{N}^* L_p = (N_1^* L_p, \ldots, N_{\overline{K}}^* L_p)$, power level allocation vector $\boldsymbol{p}^* = (p_1^*, \ldots, p_{\overline{K}}^*)$, RCPC code rate $r_{i_1}^*, \ldots, r_{i_{\overline{K}}}^*$, and RS source/parity distribution $\boldsymbol{f}_1^*, \ldots, \boldsymbol{f}_{\overline{K}}^*$ over $\overline{K}$ CDMA channels.

F. Numerical Results

1. CDMA System Setup

- **LMMSE Receiver** In order to get the BER value for each CDMA channel, we selected processing gain $N_c = 256$ and system loading $\alpha = K/N_c = 0.5$, and so the number of users $K = 128$. For multiple-path fading channel, the resolvable paths $L = 5$ for each user; for non-perfectly known channel, the window size of

estimator $\tau = 10$ symbols. The BER is calculated according to the equation in Section C, where SIR is from (3.3). The estimated error variance ξ_k^2 is given by equation (3.4). The power level is digitized between its dynamic range and its stepsize is 0.1.

- **Decorrelating Receiver** In order to get the BER value for each CDMA channel, we selected processing gain $N_c = 256$ and because of the assumption of $LK/N_c < 1$ we chose the system loading $\alpha = K/N_c = 0.19$, and so the number of users $K = 49$. For multiple-path fading channel, the resolvable paths $L = 5$ for each user; for non-perfectly known channel, the window size of estimator $\tau = 10$ symbols. The BER is calculated according to the equation in Section C, where SIR is also from (3.3). The estimated error variance ξ_k^2 is given by equation (3.4). The power level is digitized between its dynamic range and its stepsize is 1.

- **Matched Filter Receiver** In order to get the BER value for each CDMA channel, we selected processing gain $N_c = 256$ and system loading $\alpha = K/N_c = 0.5$, and so the number of users $K = 128$. For multiple-path fading channel, the resolvable paths $L = 5$ for each user; for non-perfectly known channel, the window size of estimator $\tau = 10$ symbols. The BER is calculated according to the equation in Section C, where SIR is from (3.3). The estimated error variance ξ_k^2 is given by equation (3.4). The power level is digitized between its dynamic range and its stepsize is 0.5.

2. Simulations

We experiment with standard color QCIF (144×176) *akiyo* video sequences. Each sequence is in YUV format and has 288 pictures. We used 3-D ESCOT coding

bitstream with the GOP size 16 frames and a fixed frame rate 30 fps in all experiments. We used a 16-CRC code with generator polynomial 0x15935. The generator polynomials of the RCPC code were $(0117, 0127, 0155, 0171)$, the mother code rate was 1/4, and the puncturing rate was 8. Thus, the set of RCPC code rates is $\mathcal{R} = \{8/32, 8/31, \ldots, 8/10, 8/9\}$. The decoding of the RCPC code is done with a list Viterbi algorithm where the maximum number of candidate paths is 100. Under a given BER, we implement 1000 experiments for each RCPC code rate r_i to get the P_{r_i}, which is the probability that a packet protected with rate r_i cannot be correctly decoded with the RCPC decoder, as shown in Fig. 12 where for each RCPC code, the range of BER we considered is between 1% and 20%. Thus given P_{r_i}, it is easy to get $P_{n/N}$, the probability that n out of N packets are lost, based on the equation (3.1). With the $P_{n/N}$ in hand, we can apply the joint rate and power allocation algorithm as well as joint product code algorithm as described in previous section to get the expected distortion.

Fig. 13 shows the peak signal to reconstruction noise ratio for LMMSE receiver in three channel states such as perfectly known channel, single-path fading channel and multiple-fading channel for the optimal transmission rate over $\overline{K}$ channels and the optimal product code for each channel, with or without the optimal power level allocation. Here we only consider the Y component and calculate PSNR at four total transmission rates over $\overline{K}$ channels: 72, 144, 201.6 and 270 Kbps. The average power level of each one out of 128 channels is $p_1/\sigma^2 = 3.76$ dB which makes the BER in perfectly known channel to be 10%, thus the total power level of $\overline{K}$ channels which we used to transmit one video bitstream is $p/\sigma^2 = 9.78$ dB. We limited the dynamic range of the power level such that the BER for each CDMA channel is approximately between 1% and 20%. The important observation is that at any one of three channel states, we obtained good PSNR performance and especially, the

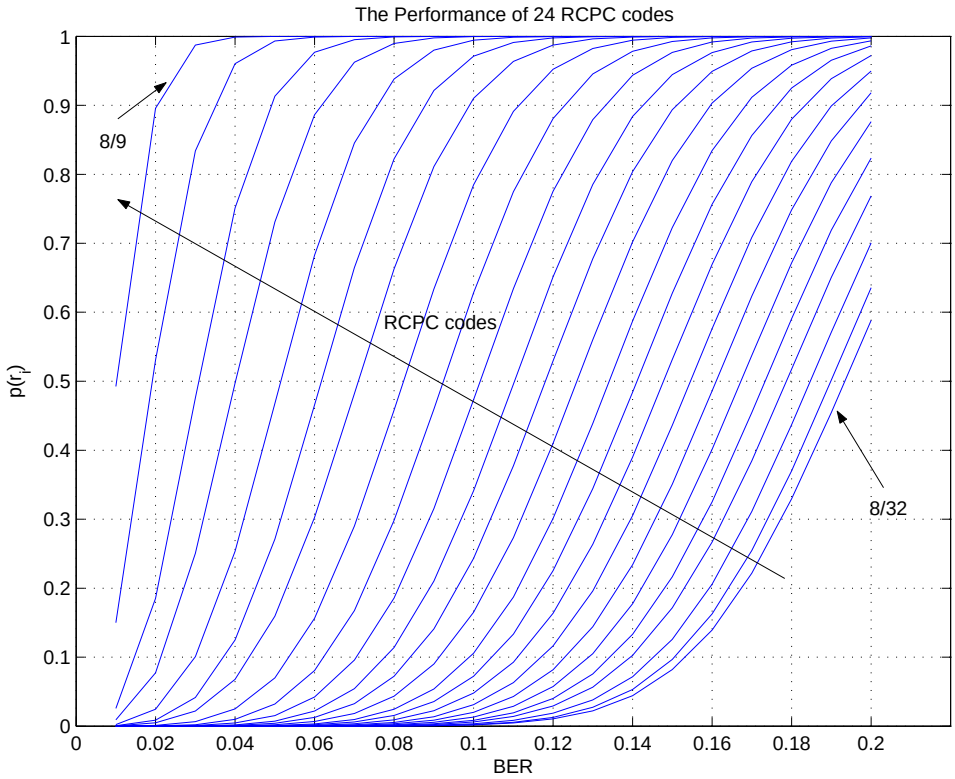


Fig. 12. The probability P_{r_i} that a packet protected with rate r_i cannot be correctly decoded with the RCPC decoder.

combined optimization of power allocation and transmission rate allocation achieved up to 1.5 dB performance gain than that only transmission rate optimization under the same channel state.

Fig. 14 shows the peak signal to reconstruction noise ratio for matched filter receiver in three channel states such as perfectly known channel, single-path fading channel and multiple-fading channel for the optimal transmission rate over $\overline{K}$ channels and the optimal product code for each channel, with or without the optimal power level allocation. Here we only consider the Y component and calculate PSNR at four total transmission rates over $\overline{K}$ channels: 72, 144, 201.6 and 270 Kbps. The average power level of each one out of 128 channels is $p_1/\sigma^2 = 9.47$ dB which makes the BER in perfectly known channel to be 10%, thus the total power level of $\overline{K}$ channels which

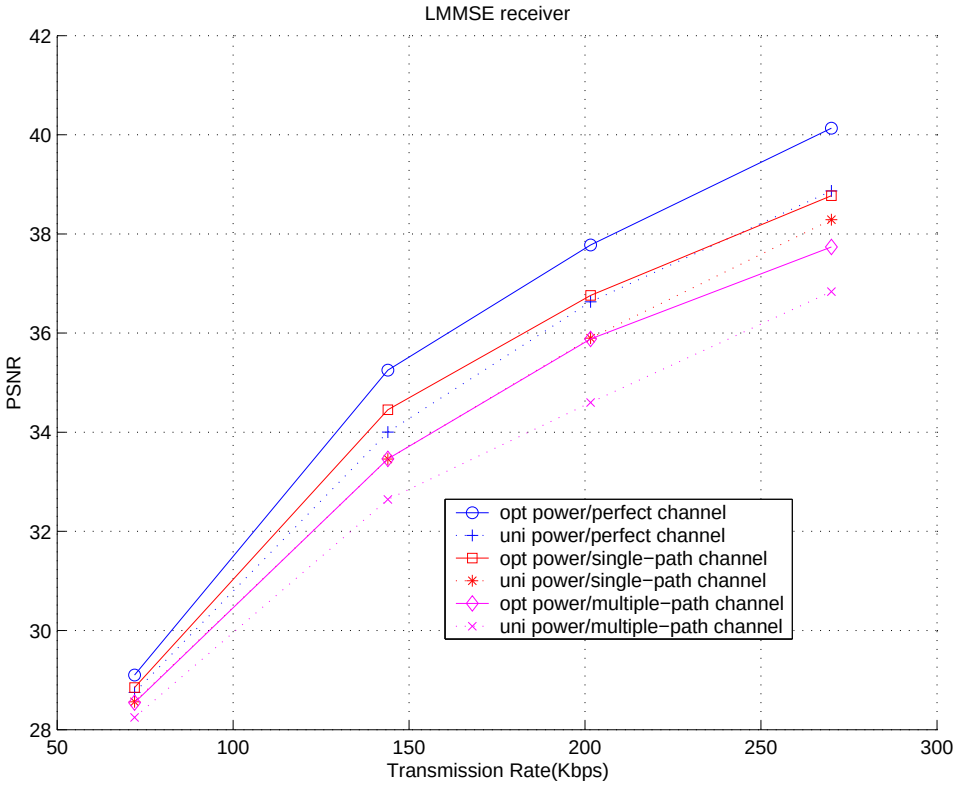


Fig. 13. PSNR vs. transmission rate for LMMSE receiver.

we used to transmit one video bitstream is $p/\sigma^2 = 15.49$ dB. We limited the dynamic range of the power level such that the BER for each CDMA channel is approximately between 1% and 20%. Similarly, the important observation is that at any one of three channel states, we obtained good PSNR performance and especially, the combined optimization of power allocation and transmission rate allocation achieved up to 1.5 dB performance gain than that only transmission optimization under the same channel state.

Fig. 15 shows the peak signal to reconstruction noise ratio for decorrelating receiver in three channel states such as perfectly known channel, single-path fading channel and multiple-fading channel for the optimal transmission rate over $\overline{K}$ channels and the optimal product code for each channel, with or without the optimal power level allocation. Here we only consider the Y component and calculate PSNR at

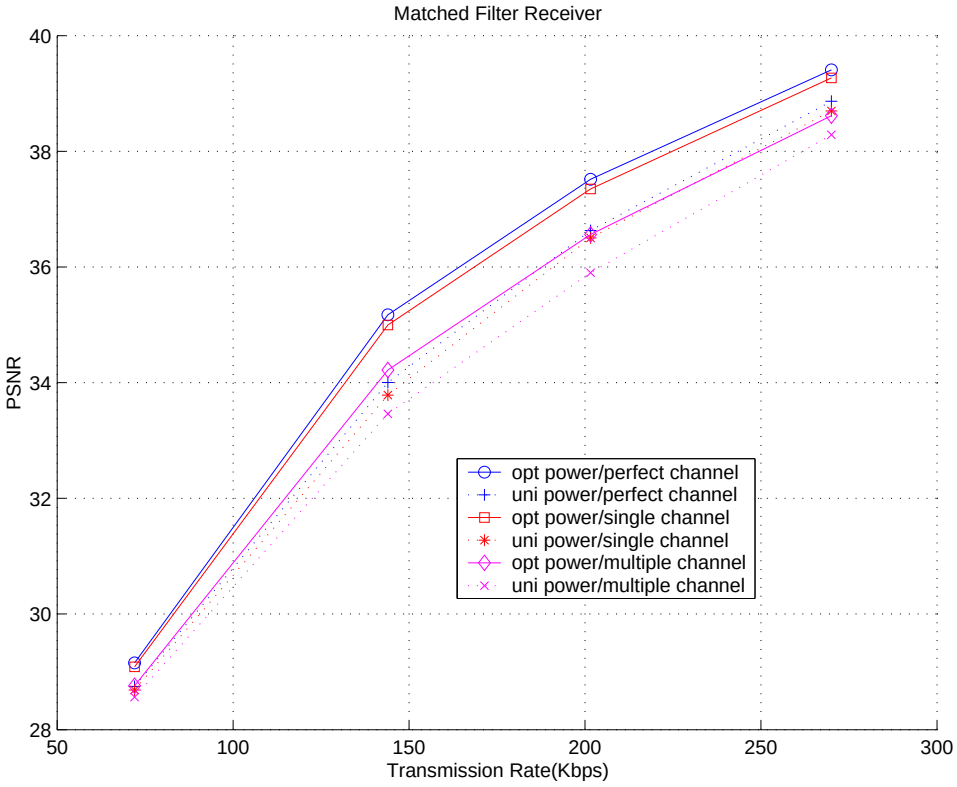


Fig. 14. PSNR vs. transmission rate for matched filter receiver.

four total transmission rates over $\overline{K}$ channels: 72, 144, 201.6 and 270 Kbps. The average power level of each one out of 49 channels is $p_1/\sigma^2 = 2.16, 3.59$ and 15.26 dB for perfectly known channel, single-path fading channel and multiple-path fading channel, respectively, which makes the BER in the three channel states to be 10% respectively. Thus the total power level of $\overline{K}$ channels which we used to transmit one video bitstream is $p/\sigma^2 = 8.18, 9.61$ and 21.28 dB for three channel states respectively. We limited the dynamic range of the power level such that the BER for each CDMA channel is approximately between 1% and 20%. Similarly, the important observation is that at any one of three channel states, we obtained good PSNR performance and especially, the combined optimization of power allocation and transmission rate allocation achieved up to 1.5 dB performance gain than that only transmission rate optimization.

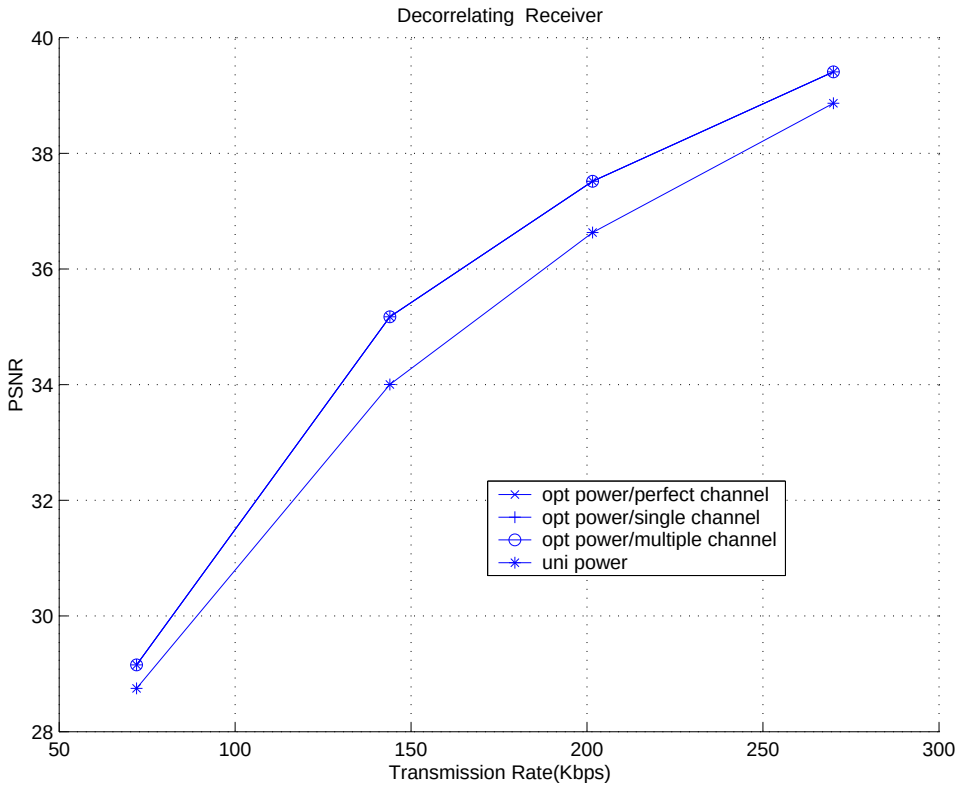


Fig. 15. PSNR vs. transmission rate for decorrelating receiver.

Fig. 16, Fig. 17 and Fig. 18 show the total power levels required for $\overline{K}$ CDMA channels, based on the optimal transmission rate over $\overline{K}$ channels and the optimal product code for each channel, with or without the optimal power level allocation, for LMMSE, matched filter and decorrelating receivers at perfectly known channel, single-path fading channel and multiple-fading channel, respectively, given a fixed PSNR performance curve. The important observation is that given a fixed PSNR performance curve, the required total power for the combined optimization of power level allocation and transmission rate allocation is less than that for only transmission rate optimization under the same channel state and moreover for certain PSNR performance, LMMSE receiver requires the least total power and matched filter requires the most total power.

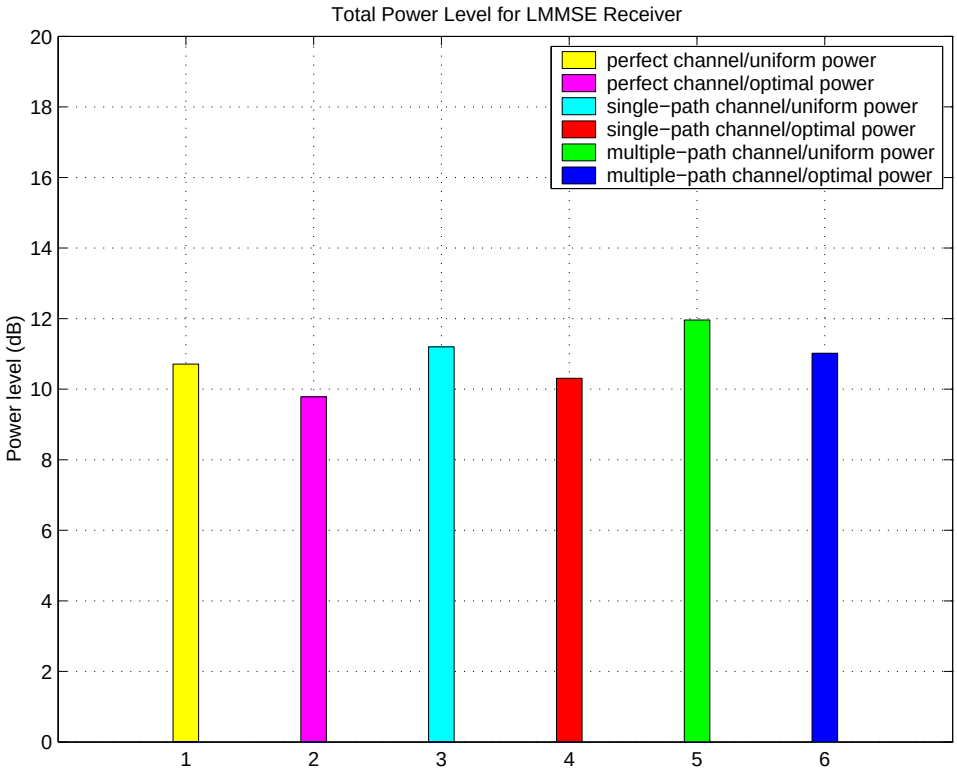


Fig. 16. The required total power for LMMSE receiver.

G. Conclusions

We have considered a multiple-channel video transmission scheme to achieve robust communication in wireless CDMA networks over fading multipath channels. The independently decodable layers from 3-D ESCOT and product code structure provide more system error resilience even during deep fading period in wireless channel and match well with networks such as the Internet that have no means of prioritizing packets. The recently developed large-system analysis results for CDMA networks are employed to characterize the performance of various receivers in multipath fading environment. We have proposed a framework for joint optimization of power allocation and transmission rate allocation over multiple CDMA channels, as well as joint optimization of Reed-Solomon source/parity rate allocation and RCPC code

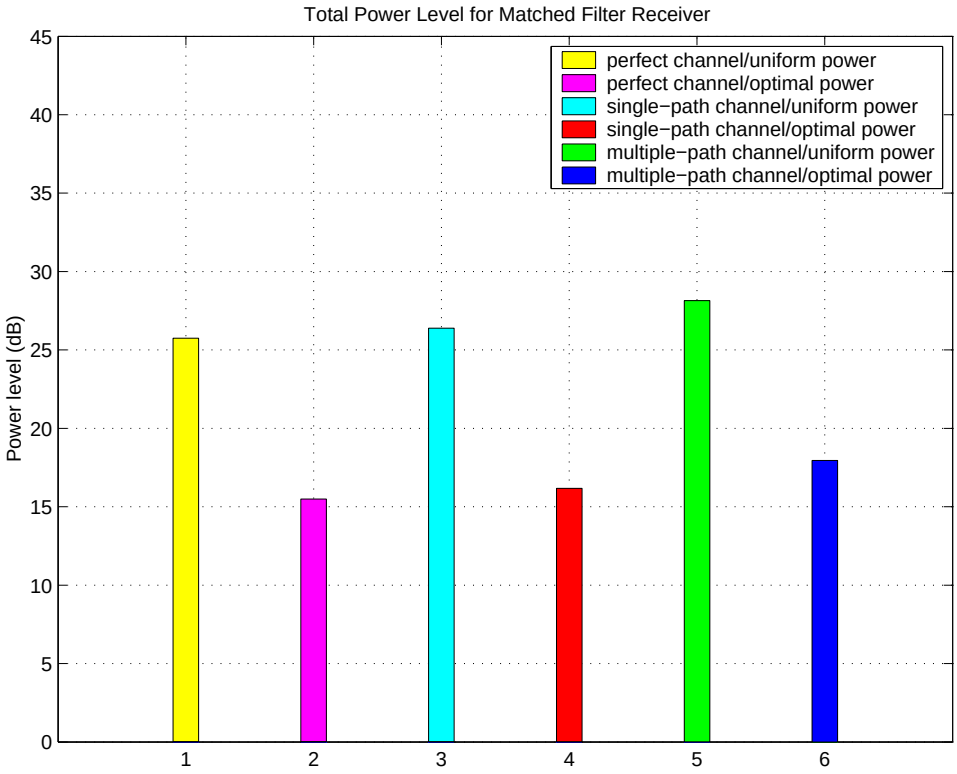


Fig. 17. The required total power for Matched Filter receiver.

rate protection for each channel , to minimize the distortion on the received video data. Simulation results show that the proposed framework allows the video quality to degrade gracefully as the fading worsens or the bandwidth decreases, and it offers improved video quality at the receiver. Moreover the proposed joint optimal transmission rate/power allocation scheme offers a performance gain of up to 1.5 dB over the scheme with only optimal transmission rate with uniform power levels.

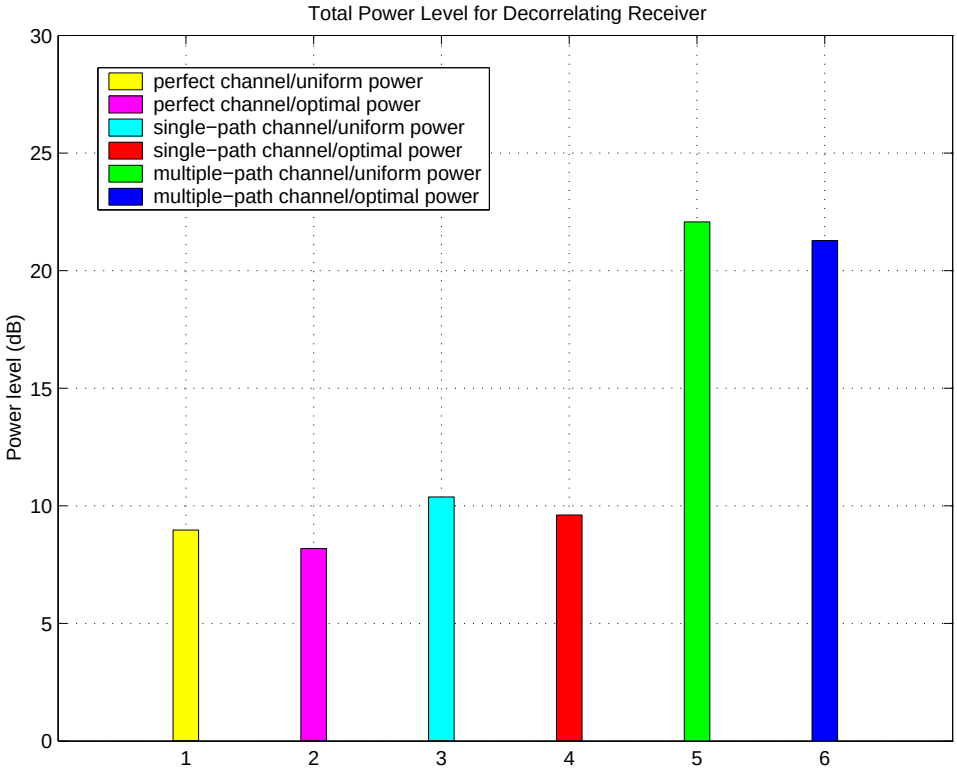


Fig. 18. The required total power for Decorrelating receiver.

CHAPTER IV

PROGRESSIVE VIDEO DELIVERY OVER WIDEBAND WIRELESS CHANNELS USING SPACE-TIME DIFFERENTIALLY CODED OFDM SYSTEMS

A. Introduction

In this chapter we consider a new system that integrates layered video coding, channel coding, signal modulation and antenna diversity for broadband wireless video transmission. Specifically we consider integrating space-time differentially coded orthogonal frequency division multiplexing (STDC-OFDM) with unequal error protection (UEP) and power control schemes for progressive video transmission. We adjust the transmission power and bit rate of each sub-carrier dynamically according to variations in channel selectivity. A multi-layer video stream is transmitted over the STDC-OFDM system, where several source layers are sent in parallel with different performance requirements (UEP). The available OFDM subcarriers are divided into several sets, each of which is associated with a different source layer. Each layer requires a particular performance (e.g. distortion of source bitstream), a specific bit rate and transmission power.

Layered (or embedded) source coding is achieved using the 3-D ESCOT coder [7], which provides bit rate scalability as well as resolution scalability. The UEP scheme for each source layer is obtained by the combination of a rate-compatible punctured convolutional(RCPC) code with cyclic redundancy check(CRC) error detection [17, 84].

We essentially input multiple source layers of *one* layered video bitstream into *multiple* STDC-OFDM channels. Using the 3-D ESCOT we first layer the bitstream

by resolution, generating multiple independently decodable layers corresponding to different resolutions. Next we packetize each layer of the video bitstream in one UEP structure before transmitting it with an individual STDC-OFDM channel. In this work, given certain total transmission rate and total transmission power over multiple STDC-OFDM channels, we transmit groups of UEP structures over these multiple STDC-OFDM channels. We address the problem of minimizing the total expected distortion of a layered video bitstream subject to a total transmission rate and a total transmission power level over multiple STDC-OFDM channels.

We propose a fast algorithm to obtain the optimal transmission rate allocation and the optimal power level allocation over multiple STDC-OFDM channels on the fly. This is based on using progressive joint source-channel codes [46] to generate operational transmission distortion-rate (TD-R) functions [92] and operational transmission distortion-power(TD-P) functions for multiple wireless channels, before forming the operational transmission distortion-power-rate (TD-PR) surfaces. We employ the Lagrange multipliers to obtain the optimal power allocation and optimal rate allocation among multiple channels with the constraints on the total transmission rate and the total power level. Assuming convexity of TD-PR surfaces, our work can use the 'equal slope' argument to effectively solve the transmission rate allocation problem as well as the transmission power allocation problem in joint source-channel coding for multiple-layer video transmissions. Because progressive joint source-channel coding offers the scalability feature, the re-optimizing and re-assembling of channel codewords are not needed in case the transmission rate changes.

To our best knowledge, there has been no prior published work on operational transmission distortion-power-rate surfaces (TD-PR). We recently came across He *et al.*' work [93] on power-rate-distortion (P-R-D) function for wireless video based on dynamic voltage scaling CMOS design technology. The authors of [93] do not

associate power levels with wireless channels, channel coding techniques, and video transmission as we do here. They only introduce the power level as an extra parameter to adjust video source coding parameters in CMOS circuit design.

The remainder of this chapter is organized as follows: Section B introduces the space-time differentially coded OFDM system for multi-layer video delivery. Section C describes the problem formulation. Section D is devoted to the fast optimal allocation algorithm for transmission rate and transmission power. Section E presents simulation results. Section F concludes the chapter.

B. Space-Time Differentially Coded OFDM System for Multi-layer Video Delivery

1. System Description

In this section, we briefly describe the space-time differentially coded OFDM system used for video transmission. Figures 19-20 show the structure of the transmitter and receiver, respectively. At the transmitter a video sequence is first compressed and partitioned into K layers according to the importance of the compressed data. Next each layer is encoded by an UEP structure (to be described as in Section C). Then the bitstream of each layer is coded with space-time differential coding. Finally after OFDM modulation each layer is transmitted by two antennas. Hence there are totally $2K$ transmission antennas in this system. At the receiver we employ K antennas, one for a different layer. In our system, the total number of sub-carriers is divided into K sets, each set is associated with a different layer. In other word, for each layer, the two OFDM frames at the transmitter and one OFDM at the receiver use the same OFDM set. Different layers use different OFDM sets. Because we use different sub-carriers for different layers, for each layer we only need to consider the transmission from the two transmitter antennas to the receiver antenna at the same layer. Assume the total

number of sub-carriers is N_c, thus the number of sub-carriers for transmitting each layer is N_c/K. Thus all OFDM blocks (i.e., IFFT blocks or FFT blocks) in Figures 19-20 take the same size N_c/K. Notice that we use different sets of sub-carriers for different layers (although with the same OFDM frame size). We also assume M-PSK modulation.

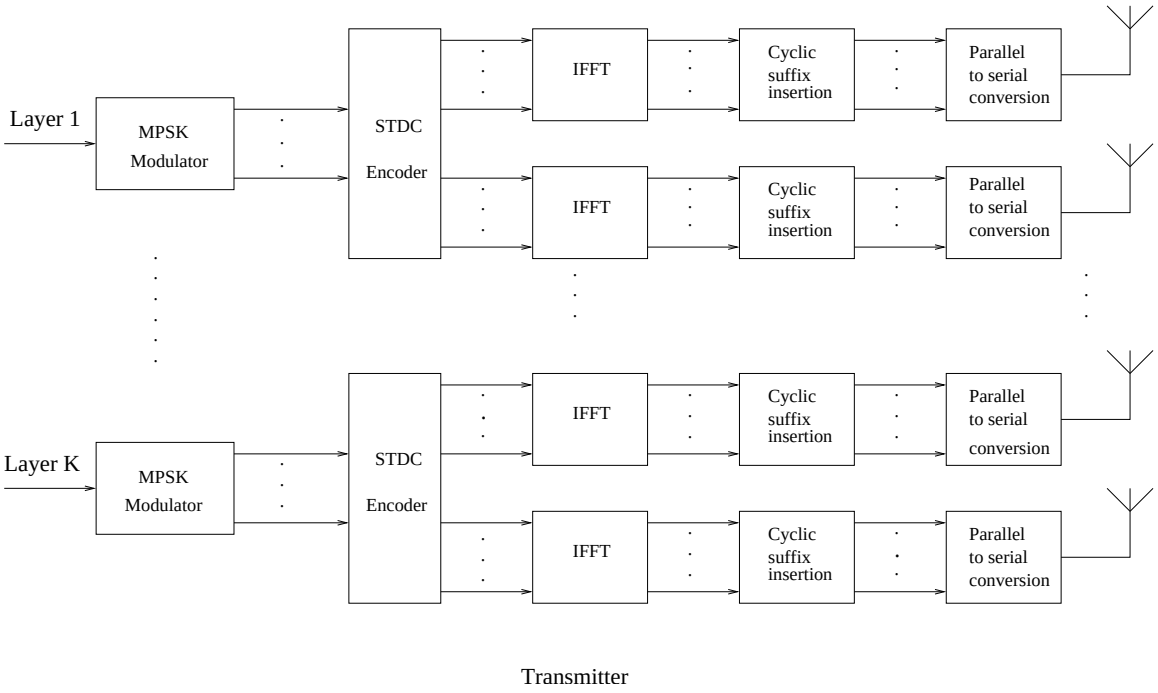


Fig. 19. The block diagram of the proposed multi-layer video space-time differentially coded OFDM transmitter system.

Several independently decodable video layers are first formed in the 3-D ESCOT embedded video bitstream, these layers are encoded from different subbands. For example, Figure 21(A) illustrates four-layer encoding for an 2-D image. Then using the fast optimal allocation algorithm for transmission rate and transmission power in Section D, we obtain the optimal power allocation and optimal transmission rate allocation for multiple STDC-OFDM channels, as well as the optimal UEP structure for each channel. Each independently decodable layer corresponds to one UEP structure.

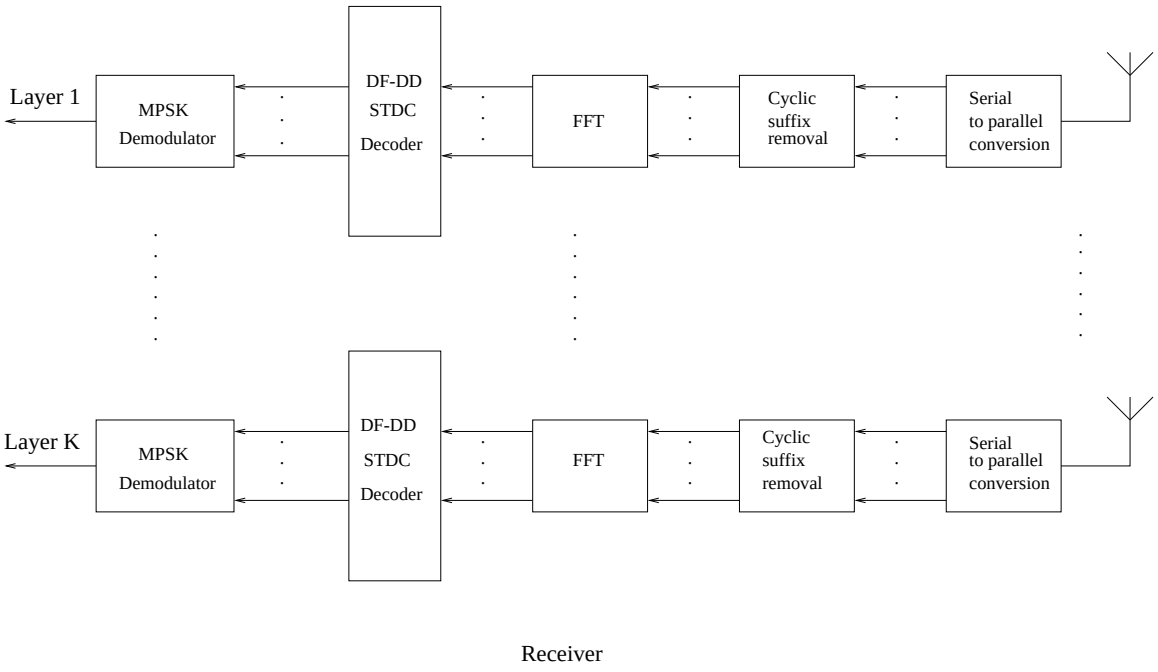


Fig. 20. The block diagram of the proposed multi-layer video space-time differentially coded OFDM receiver system.

All codewords of one UEP structure are transmitted over one STDC-OFDM channel. After multiple UEP structures are formed for one embedded video bitstream, we transmit all codewords in these UEP structures over multiple STDC-OFDM channels in parallel.

A total of K STDC-OFDM channels are used to transmit K groups of UEP structured codewords for one embedded video bitstream. Figure 21(B) illustrates the framework we present (here $K = 4$). The transmitter now has several groups of UEP structured codewords to send. The transmitter buffers frames as they arrive. When a block of groups of frames(GOF) is accumulated, it encodes the GOFs and packetizes the resulting independently decodable embedded layers. For a total transmission rate and a total power level over multiple STDC-OFDM channels, the transmitter chooses to get the optimal power level allocation vector and transmission rate allocation vector

for multiple STDC-OFDM channels, and the optimal RCPC codes for each UEP structure, as illustrated in Figure 21(B). The receiver instantly recovers as many source symbols as possible from the received codewords and decodes them for a block of GOFs. Playback begins after one block of GOFs of coding delay.

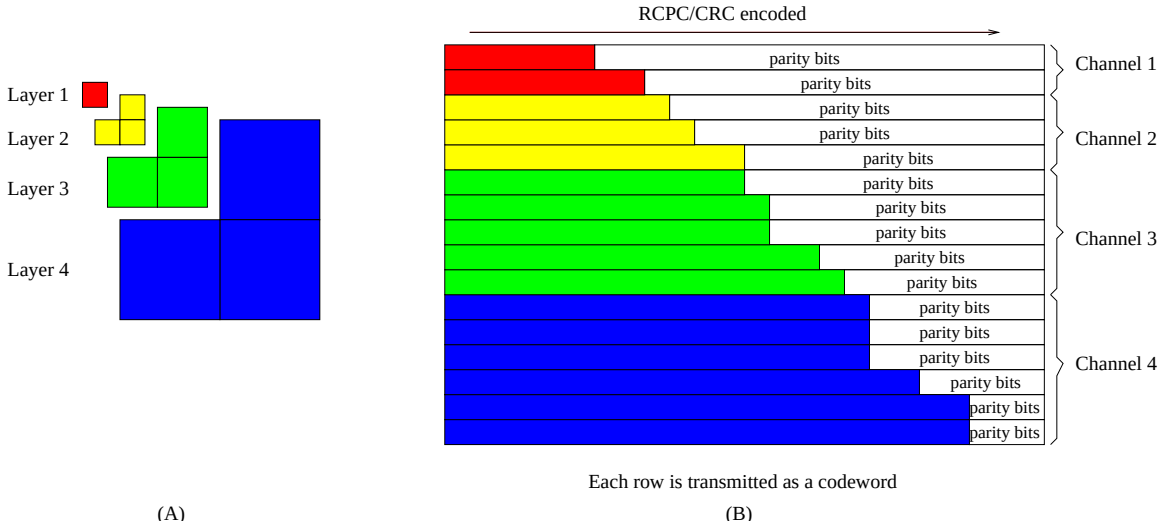


Fig. 21. (A) Layered image coding by resolution. (B) The proposed multi-layer video UEP structures for multimedia transmissions over multiple STDC-OFDM channels.

2. Performance Analysis of Multiple-Symbol Decision-Feedback Space-Time Differentially Coded OFDM System

In this section, we briefly review the performance analysis of multiple-symbol decision-feedback space-time differentially coded OFDM system [39]. Consider a communication system with two transmit antennas and one receive antenna. Let the PSK information symbols at time n be $a_n \in \left\{ \frac{1}{\sqrt{2}} e^{j\frac{2\pi k}{M}}, k = 0, 1, \cdots, M-1 \right\}$. Define the following matrices:

$$\underline{A}_n \triangleq \begin{bmatrix} a_{2n} & a_{2n+1} \\ -a_{2n+1}^* & a_{2n}^* \end{bmatrix}, \quad \underline{G}_n \triangleq \underline{A}_n \underline{A}_0^H, \quad \text{for } n \geq 0, \tag{4.1}$$

The space-time differential block code is recursively defined as follows:

$$\underline{X}_0 = \underline{A}_0, \quad \underline{X}_n = \underline{G}_n \underline{X}_{n-1}, \quad n = 1, 2, \cdots, \tag{4.2}$$

by a simple induction, it is easy to show that the matrix $\underline{X}_n$ has the following form

$$\underline{X}_n \triangleq \begin{bmatrix} x_{2n} & x_{2n+1} \\ -x_{2n+1}^* & x_{2n}^* \end{bmatrix}. \tag{4.3}$$

We now consider decoding of the space-time differential block code in flat-fading channels. At the transmitter, at time slot $2n$, the symbols on the first row of $\underline{X}_n$, x_{2n} and x_{2n+1} are transmitted simultaneously from antenna 1 and antenna 2, respectively. At time slot $2n+1$, the symbols on the second row of $\underline{X}_n$, $-x_{2n+1}^*$ and x_{2n}^* are transmitted simultaneously from the two antennas. In order to simplify the receiver structure, we make the assumption that the channels remain constant over two consecutive symbol intervals. Thus the received signals can be written as

$$\underbrace{\begin{bmatrix} y_{2n} \\ y_{2n+1} \end{bmatrix}}_{\underline{y}_n} = \underbrace{\begin{bmatrix} x_{2n} & x_{2n+1} \\ -x_{2n+1}^* & x_{2n}^* \end{bmatrix}}_{\underline{X}_n} \underbrace{\begin{bmatrix} \alpha_{2n}^{(1)} \\ \alpha_{2n}^{(2)} \end{bmatrix}}_{\underline{\alpha}_n} + \underbrace{\begin{bmatrix} v_{2n} \\ v_{2n+1} \end{bmatrix}}_{\underline{v}_n}. \tag{4.4}$$

where $\{\alpha_n^{(1)}\}_n$ and $\{\alpha_n^{(2)}\}_n$ are the fading processes associated with the channels between the two transmit antennas, and the receive antenna are modelled as mutually independent complex Gaussian variables with Jakes' correlation structure [14]. Each of them has normalized autocorrelation function $R(n) \triangleq E\{\alpha_{m+n}\alpha_m^*\} = E_s J_0(2\pi B_d n T)$, where E_s is the average symbol energy, $J_0(\cdot)$ is the zeroth order Bessel function of the first kind, T is the symbol interval, and B_d is the maximum Doppler shift, which is proportional to the vehicle speed carrier frequency.

Denote

$$\begin{aligned}
\boldsymbol{y}_n &= [\underline{y}_n^T\ \underline{y}_{n-1}^T\ \cdots\ \underline{y}_{n-P+1}^T]^T, \\
\boldsymbol{\alpha}_n &= [\underline{\alpha}_n^T\ \underline{\alpha}_{n-1}^T\ \cdots\ \underline{\alpha}_{n-P+1}^T]^T, \\
\boldsymbol{v}_n &= [\underline{v}_n^T\ \underline{v}_{n-1}^T\ \cdots\ \underline{v}_{n-P+1}^T]^T, \\
\boldsymbol{X}_n &= \mathrm{diag}\{\underline{X}_n, \underline{X}_{n-1}, \cdots, \underline{X}_{n-P+1}\}, \\
\boldsymbol{G}_n &= [\underline{G}_n\ \underline{G}_{n-1}\ \cdots\ \underline{G}_{n-P+2}].
\end{aligned}$$

Then we have

$$\boldsymbol{y}_n = \boldsymbol{X}_n\boldsymbol{\alpha}_n + \boldsymbol{v}_n. \tag{4.5}$$

The conditional log-likelihood function is given by

$$\log p(\boldsymbol{y}_n|\boldsymbol{G}_n) = -\boldsymbol{y}_n^H\boldsymbol{Q}_G^{-1}\boldsymbol{y}_n - \log\det(\boldsymbol{Q}_G) - 2P\log\pi, \tag{4.6}$$

where $\boldsymbol{Q}_G \triangleq E\Big\{\boldsymbol{y}_n\boldsymbol{y}_n^H\Big\} = E_s\boldsymbol{X}_n\Big(\boldsymbol{\Sigma}_P \otimes \boldsymbol{I}_2\Big)\boldsymbol{X}_n^H + \sigma^2\boldsymbol{I}_{2P}$, where $\otimes$ denotes the Kronecker matrix product, and the normalized $P\times P$ autocorrelation matrix has elements given by $\boldsymbol{\Sigma}_P[i,j] = J_0(2\pi B_d T(i-j))$. Denote $\boldsymbol{T} \triangleq \Big(\boldsymbol{\Sigma}_P + \frac{\sigma^2}{E_s}\boldsymbol{I}_P\Big)^{-1} = [t_{ij}]$, then we have the following decision-feedback decoding rule for $\underline{G}_n$:

$$\hat{\underline{G}}_n = \arg\min_{\underline{G}_n} \mathrm{Re}\Big\{\underline{y}_n^H\underline{G}_n\sum_{j=1}^{P-1} t_{0,j}\Big(\prod_{l=1}^{j-1}\hat{\underline{G}}_{n-l}\Big)\underline{y}_{n-j}\Big\} \tag{4.7}$$

The decision-feedback space-time differential decoding algorithm is summarized as follows [39].

Algorithm of MS-DF Space-Time Differential Decoding: Given the initial information symbol matrix $\underline{A}_0$, let $\hat{\underline{A}}_0 = \underline{A}_0$.

1. Based on the decision memory order P, the fading statistic $\boldsymbol{\Sigma}_n$, and the signal-to-noise ration $\frac{E_s}{\sigma^2}$, compute the feedback metric coefficients from $\boldsymbol{T} = \Big(\boldsymbol{\Sigma}_P +$

$\frac{\sigma^2}{E_s}\boldsymbol{I}_P\Big)^{-1}$.

2. Estimate the initial symbol matrices: for $n = 1, 2, \cdots, P-1$,

   (a) Estimate $\underline{\hat{G}}_n$ by simply quantizing $\underline{Y}_n^H \underline{Y}_{n-1}$.

   (b) Perform differential decoding according to $\underline{\hat{A}}_n = \underline{\hat{G}}_n \underline{\hat{A}}_0$.

3. For $n = P, P+1, \cdots$,

   (a) Estimate $\underline{\hat{G}}_n$ according to (4.7).

   (b) Perform differential decoding according to $\underline{\hat{A}}_n = \underline{\hat{G}}_n \underline{\hat{A}}_0$.

We next discuss the application of decision-feedback space-time differential decoding in OFDM systems. We consider a STDC-OFDM channel with Q subcarriers for one video layer, two transmit antennas and one receive antenna, signaling through a frequency- and time-selective fading channel. As illustrated in Figures 19-20, the information bits first go through a serial to parallel converter, then are encoded by a STDC encoder using the encoding algorithm introduced above. For each subcarrier $m, 0 \le m \le Q$, let $\{X_m[n]\}_n$ be the output of a space-time differential encoder. Then $\{X_m[n]\}_{m=0}^{Q-1}$ constitutes the OFDM symbol at time n. This OFDM symbol goes through an IFFT block to perform the inverse Fourier transform. Denote

$$\{x_m[2n]\}_{m=0}^{Q-1} = \text{IFFT}\Big[\{X_m[2n]\}_{m=0}^{Q-1}\Big], \{x_m[2n+1]\}_{m=0}^{Q-1} = \text{IFFT}\Big[\{X_m[2n+1]\}_{m=0}^{Q-1}\Big]. \quad (4.8)$$

After adding proper cyclic prefix, during time slot $2n$, the signals $\{x_m[2n]\}_{m=0}^{Q-1}$ are transmitted through antenna 1 serially, and $\{x_m[2n+1]\}_{m=0}^{Q-1}$ are transmitted through antenna 2 serially; during time slot $2n+1$, the signals $\{-x_m[2n+1]^*\}_{m=0}^{Q-1}$ and $\{x_m[2n]^*\}_{m=0}^{Q-1}$ are transmitted serially through antenna 1 and antenna 2, respectively.

For OFDM systems with proper cyclic extension and sample timing, with tolerable leakage, the channel frequency response between the i-th transmit antenna and the receive antenna at the n-th time slot and at the m-th subcarrier can be expressed as

$$H_m^{(i)}[n] = \sum_{l=0}^{L-1} \alpha^{(i)}[l;n] e^{-j2\pi ml/Q}, i = 1,2; m = 0,1,\cdots,Q-1. \tag{4.9}$$

Where $L = \lceil \tau_m \Delta_f + 1 \rceil$ denotes the maximum number of resolvable taps, with τ_m being the maximum multipath spread and Δ_f being the tone spacing of the OFDM systems; $\alpha^{(i)}[l;n]$ is the complex amplitude of the l-th tap, whose delay is lT. Each of them is independent with each other and has the same normalized correlation function and different average power $\sigma_l^2, l = 0,1,\cdots,L-1$. Assume Jakes fading channel model, $R_{\alpha^{(i)}}(l;n) = \sigma_l^2 J_0(2\pi B_d nT)$. At the receiver, the received symbols first go through the serial to parallel converter. Assuming proper cyclic prefix is employed at each OFDM symbol. Then by applying FFT to the time-domain received signals and assuming the channels remain static during two OFDM symbols, we obtain the frequency-domain received signals

$$\begin{bmatrix} Y_m[2n] \\ Y_m[2n+1] \end{bmatrix} = \begin{bmatrix} X_m[2n] & X_m[2n+1] \\ -X_m[2n+1]^* & X_m[2n]^* \end{bmatrix} \begin{bmatrix} H_m^{(1)}[2n] \\ H_m^{(2)}[2n] \end{bmatrix} + \begin{bmatrix} V_m[2n] \\ V_m[2n+1] \end{bmatrix}. \tag{4.10}$$

Where $H_m^{(1)}[n]$ and $H_m^{(2)}[n]$ are the complex channel frequency responses between the two transmit antennas and the receive antenna at time slot n and the m-th subcarrier. $V_m[2n]$ is the complex white Gaussian noise. The auto-correlation function of $H_m^{(i)}$ is

$$R_{H_m^{(i)}}(n) = J_0(2\pi B_d nT) \cdot \sum_{l=0}^{L-1} \sigma_l^2, \quad m = 0,1,\cdots,Q-1. \tag{4.11}$$

Comparing equations (4.10) with (4.4), we know the multiple-symbol decision-feedback space-time differential decoding method discussed above can be applied in parallel at

each subcarrier m to decode the information symbols.

Figure 22 shows the BER performance for one STDC-OFDM channel as suggested above in this Section. We chose the number of sub-carriers $Q = 128$ and a 3-tap frequency-selective channel between each transmit antenna and the receive antenna with normalized Doppler frequency $B_dT = 0.0075$. In the simulation, the fading processes vary from one OFDM symbol to another.

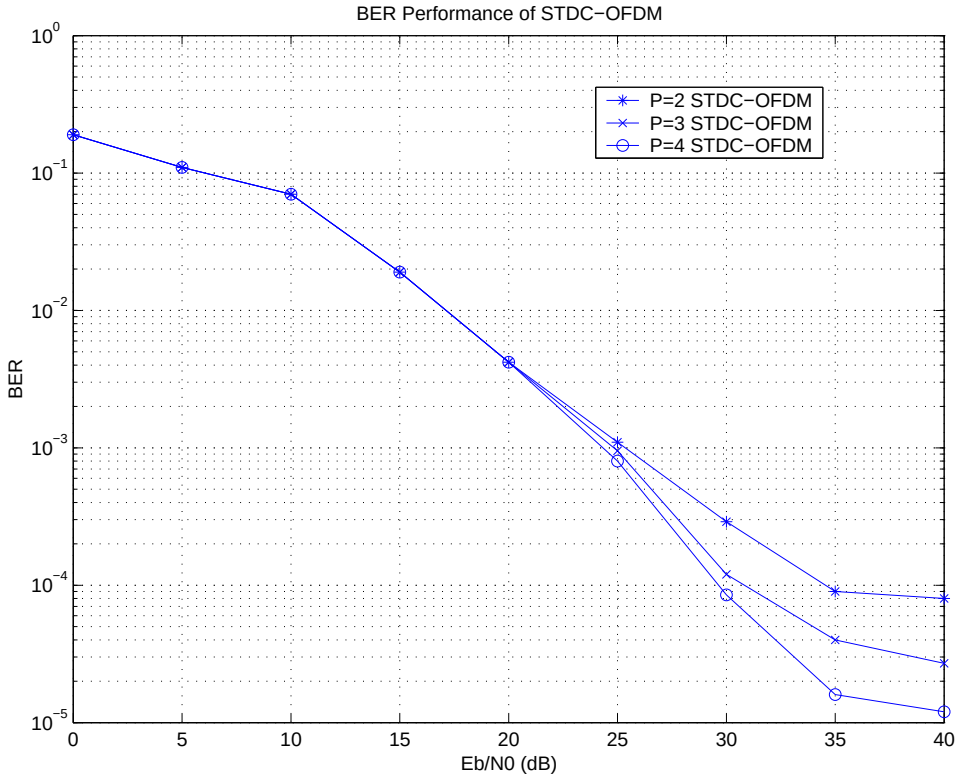


Fig. 22. Performance for space-time differentially coded OFDM system with QDPSK modulation in frequency-selective fading channels with normalized Doppler frequency $B_dT = 0.0075$ and decision memory order $P = 2, 3, 4$.

In Section E, based on the BER performance here, we implement 1000 experiments for each RCPC code rate r_i to get the $p(r_i)$, which is the probability that a codeword protected with rate r_i cannot be correctly decoded with the RCPC decoder. Furthermore through the equation (4.12), we can get the probability $P_i(\boldsymbol{r})$, which we

use in our optimal algorithm.

3. Channel Coding in the Form of UEP

We consider a joint source-channel coding system where source coder is an embedded coder, and the channel coder uses a finite number of channel codes with error detection and error correction capability [46]. In this work we use 3-D ESCOT as the source coder and the combination of a RCPC code with CRC error detection as the channel coder. After source coding, the channel encoder takes successive blocks of the compressed bitstream and transforms them into a sequence of channel codewords of fixed length L_p, which are sent over a wireless channel, as shown in Figure 23. If a received codeword can be correctly decoded, then the next codeword is considered by the decoder. Otherwise the decoding is stopped, and the video is reconstructed from the correctly decoded codewords. We assume that all decoding errors can be detected.

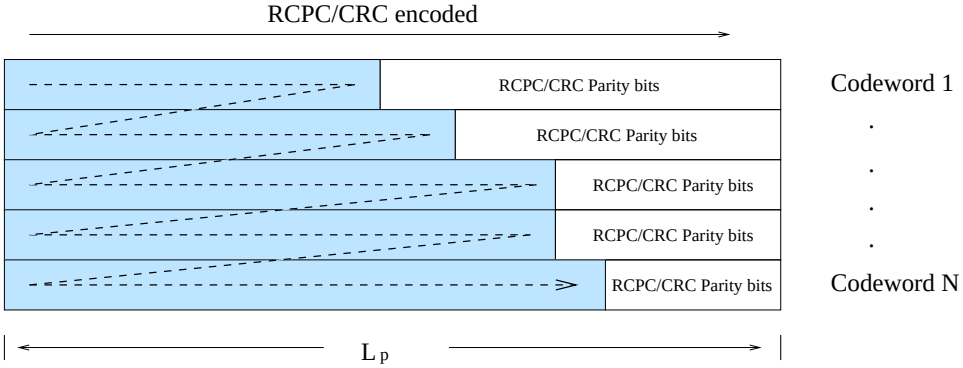


Fig. 23. Unequal error protection (UEP) structure. There are $N = 5$ codewords. The shaded parts are source symbols. The embedded source bitstream flows along the direction of the dashed line.

Let $\mathcal{R} = \{r_1, r_2, \ldots, r_m\}$ denote the set of channel rates with $r_1 < \cdots < r_m$ and $p(r_1) < \cdots < p(r_m)$, where $p(r_i)$ is the probability of a decoding error in a codeword of length L_p protected with code rate $r_i, i = 1, \cdots, m$. Given a transmission bit

budget B corresponding to a total number of $N = \lfloor B/L_p \rfloor$ transmitted codewords, an N-codeword *unequal error protection* (UEP) $\boldsymbol{r} = (r_{k_1}, \cdots, r_{k_N})$ protects the i-th source block, $i = 1, \cdots, N$, with channel code rate $r_{k_i} \in \mathcal{R}$. Then

$$P_i(\boldsymbol{r}) = \begin{cases} p(r_{k_1}) & i = 0, \\ p(r_{k_{i+1}}) \prod_{j=1}^{i} (1 - p(r_{k_j})) & 1 \le i \le N-1, \\ \prod_{j=1}^{N} (1 - p(r_{k_j})) & i = N, \end{cases} \tag{4.12}$$

is the probability that no decoding errors occur in the first i codewords with an error in the next one.

The end-to-end performance of an N-codeword UEP $\boldsymbol{r}$ can be measured by the expected distortion

$$E_N[d](\boldsymbol{r}) = \sum_{i=0}^{N} P_i(\boldsymbol{r}) d_i(\boldsymbol{r}), \tag{4.13}$$

where $d_0(\boldsymbol{r}) = d_0$ is a constant corresponding to the distortion if no source bits are received, and for $i > 1$, $d_i(\boldsymbol{r})$ is the reconstruction error using the first i source blocks. An N-codeword UEP $\boldsymbol{r}$ that minimizes (4.13) is called *distortion optimal*. To measure the progressive performance of an N-codeword UEP $\boldsymbol{r} = (r_{k_1}, \cdots, r_{k_N})$, the average of the expected performance over all N intermediate transmission rates can be computed

$$L_N[d](\boldsymbol{r}) = \frac{1}{N} \sum_{n=1}^{N} E_n[d](r_{k_1}, \cdots, r_{k_n}) = \frac{1}{N} \sum_{n=1}^{N} E_n[d](\boldsymbol{r}^{(1:n)}), \tag{4.14}$$

where $\boldsymbol{r}^{(1:n)}$ denotes the first n components of vector $\boldsymbol{r}$ with length N. An UEP $\boldsymbol{r}$ that minimizes the average expected distortion $L_N[d](\boldsymbol{r})$ is called *progressive distortion optimal*.

Thus we have the following algorithm to obtain a progressive distortion optimal UEP, which we employ to produce the operational transmission distortion-rate (TD-

R) functions given a fixed power level.

Algorithm of Producing Transmission Distortion-Rate Functions:

1. Set $k = 1, l = 1$, and $n = 0$. Use the linear-time rate-optimal algorithm [46] to compute a progressive rate optimal N-codeword UEP $\boldsymbol{r}_n$.

2. Let r be the k-th largest code rate used by $\boldsymbol{r}_n$. Let j be the index of the first codeword that $\boldsymbol{r}_n$ protects with r. If $r = r_1$, stop. Otherwise let $r_c \in \mathcal{R}$ be the l-th largest code rate smaller than r and define $\boldsymbol{r}_c$ to be the UEP obtained from $\boldsymbol{r}_n$ by protecting codeword j with r_c.

3. If $L_N[d](\boldsymbol{r}_c) < L_N[d](\boldsymbol{r}_n)$, set $\boldsymbol{r}_{n+1} = \boldsymbol{r}_c$, $n = n + 1$, $l = 1$, and go to Step 2.

4. If $j \neq 1$ and r_c is greater than the rate of codeword $j - 1$, set $l = l + 1$. If $j \neq 1$ and r_c is equal to the rate of codeword $j - 1$, set $l = 1$ and $k = k + 1$. If $j = 1$ and $r_c \neq r_1$, set $l = l + 1$. If $j = 1$ and $r_c = r_1$, stop.

5. Go to Step 2.

Readers who are interested in the complexity and convergence of this algorithm are referred to [46].

C. Problem Formulation

Now let us consider an optimal problem on the proposed space-time differentially coded OFDM system. Let $\boldsymbol{N} = (N_1, \ldots, N_K)$ be the vector for the number of codewords for K UEP structures, thus the transmission rate vector will be $\boldsymbol{N}L_p$ for K UEP structures. Denote $\boldsymbol{p} = [p_1, \cdots, p_K]$ as the power level allocation vector for the transmission powers of the K STDC-OFDM channels used for transmitting the video bitstream. Thus the OFDM frame power at the transmitter is $p_k/2$ for the k-th

layer. One of our objective is to optimize the power allocation $\boldsymbol{p}$ vector subject to a total power constraint such that the expected distortion is minimized. Note that from Sections B, the bit error rate, and therefore the probability $P_i(\boldsymbol{r})$, $i = 1, \cdots, N$ for one STDC-OFDM channel, depends on only the power of that particular channel. Thus by introducing the power allocation vector $\boldsymbol{p}$ and the transmission rate vector $\boldsymbol{N}L_p$ over K STDC-OFDM channels, we can get the expected distortion expression for the k-th UEP structure

$$D(\boldsymbol{r}_k, p_k) = \sum_{i=0}^{N} P_i(\boldsymbol{r}_k, p_k) d_i(\boldsymbol{r}_k), \tag{4.15}$$

obviously the expected distortion expression $D(\boldsymbol{r}_k, p_k)$ is the same as the one represented by $E_N[d](\boldsymbol{r})$ as in equation (4.13). However here we induce the power variable apparently into our notation. Where $\boldsymbol{r}_k = (r_{k_1}, \cdots, r_{k_N})$ is the N-codeword UEP structure for the k-th channel, and $k = 1, \ldots, K$ index STDC-OFDM channels. Because the independence of K source layers, the total expected distortion for K channels can be expressed as

$$\sum_{k=1}^{K} D(\boldsymbol{r}_k, p_k) = \sum_{k=1}^{K} \sum_{i=0}^{N} P_i(\boldsymbol{r}_k, p_k) d_i(\boldsymbol{r}_k). \tag{4.16}$$

We can now formulate a combined optimization problem of power allocation and transmission rate allocation over K STDC-OFDM channels, as well as UEP structure for each of K STDC-OFDM channels. Given K channels for transmitting one video bitstream, we consider the following constraint optimization problem:

$$\min_{\boldsymbol{N},\boldsymbol{p}} \Big(\sum_{k=1}^{K} \min_{\boldsymbol{r}_k} D(\boldsymbol{r}_k, p_k) \Big) \text{ subject to } \begin{cases} \sum_{k=1}^{K} N_k L_p \leq \mathbb{R}, \\ \sum_{k=1}^{K} p_k = \sum_{k=1}^{K} E_k N_k L_p \leq \mathbb{P}, \end{cases} \tag{4.17}$$

where $\mathbb{P}$ and $\mathbb{R}$ are a total transmission power and a total transmission rate over K STDC-OFDM channels , respectively. E_k is the energy per bit for the k-th layer.

D. Fast Optimal Allocation Algorithm for Transmission Rate and Transmission Power

One way to solve the above constraint optimization (4.17) is by obtaining the optimal transmission rates $\boldsymbol{N}L_p$ and power allocation vector $\boldsymbol{p}$ for K STDC-OFDM channels, that minimizes the Lagrangian cost

$$J(\boldsymbol{N};\boldsymbol{p};\lambda_1,\lambda_2)=\sum_{k=1}^{K}D(\boldsymbol{r}_k,p_k)+\lambda_1\sum_{k=1}^{K}N_kL_p+\lambda_2\sum_{k=1}^{K}p_k. \tag{4.18}$$

We can solve (4.18) by using an iterative approach that is based on the method of alternating variables for multivariable minimization [31]. However this method is time-consuming and not suitable for delay constrained video applications. Here we devise a new and practical algorithm to solve this problem.

We propose to use progressive joint source-channel codes to generate operational transmission distortion-rate (TD-R) functions and operational transmission distortion-power (TD-P) functions for multiple wireless channels, before forming the operational transmission distortion-power-rate (TD-PR) surfaces. The generation of operating TD-R function from the progressive UEP is straightforward. With the UEP structure obtained, one can generate the TD-R points through equation (4.15) for each of K channels. According to the BER performance of STDC-OFDM system in Section B and (4.15), the different power levels for each of K channels correspond to the different TD-R functions. Thus given a finite set of power levels for each channel, we can obtain a group of TD-R functions. Therefore for a fixed transmission rate, we can find the corresponding operational transmission distortion-power (TD-P)

function. Thus we can get the whole operational transmission distortion-power-rate (TD-PR) surface for each of the K channels. Figures 24 - 27, for instance, gives TD-PR surfaces for $K = 4$ STDC-OFDM channels with the 3-D ESCOT source coding and the target transmission rate 512 codewords by using the method we proposed here. The parameters for obtaining Figures 24- 27 will be given in next section.

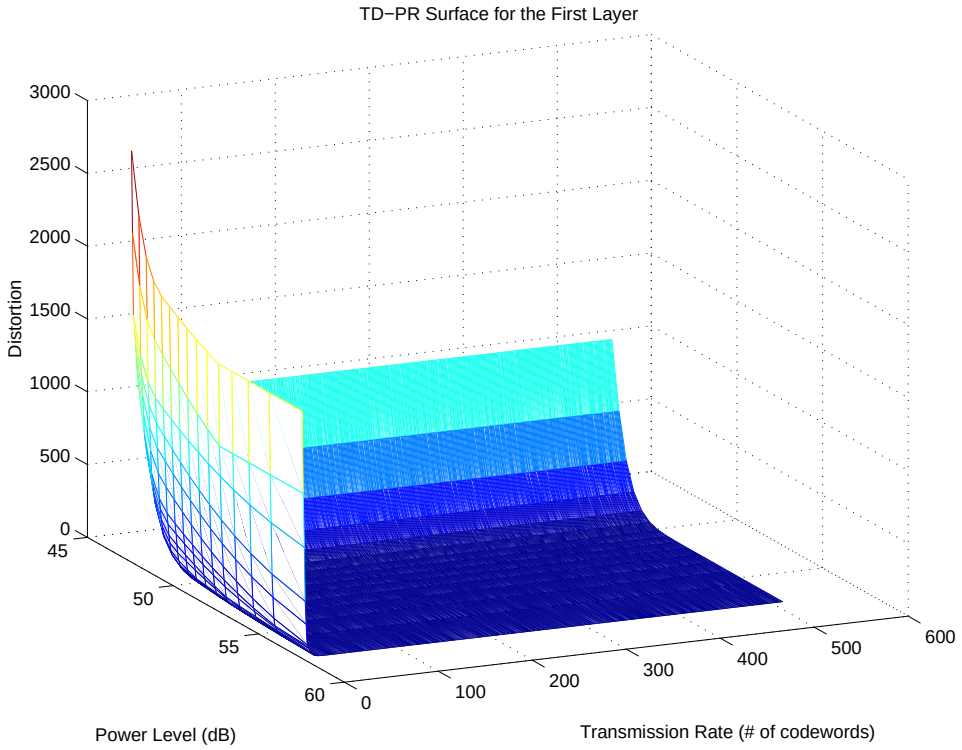


Fig. 24. The operational transmission distortion-power-rate surfaces for the first channel.

The advantages of using the progressive UEP algorithm are as follows. One is that it has low complexity and can work in wireless video applications. Indeed using the progressive UEP algorithm can guarantee the optimal performance not only at one target rate but also at all intermediate rates. Thus we can generate the TD-PR surfaces at one time for all available transmission rates. The re-optimization and re-assembly of UEP structures are not necessary when the transmission rate

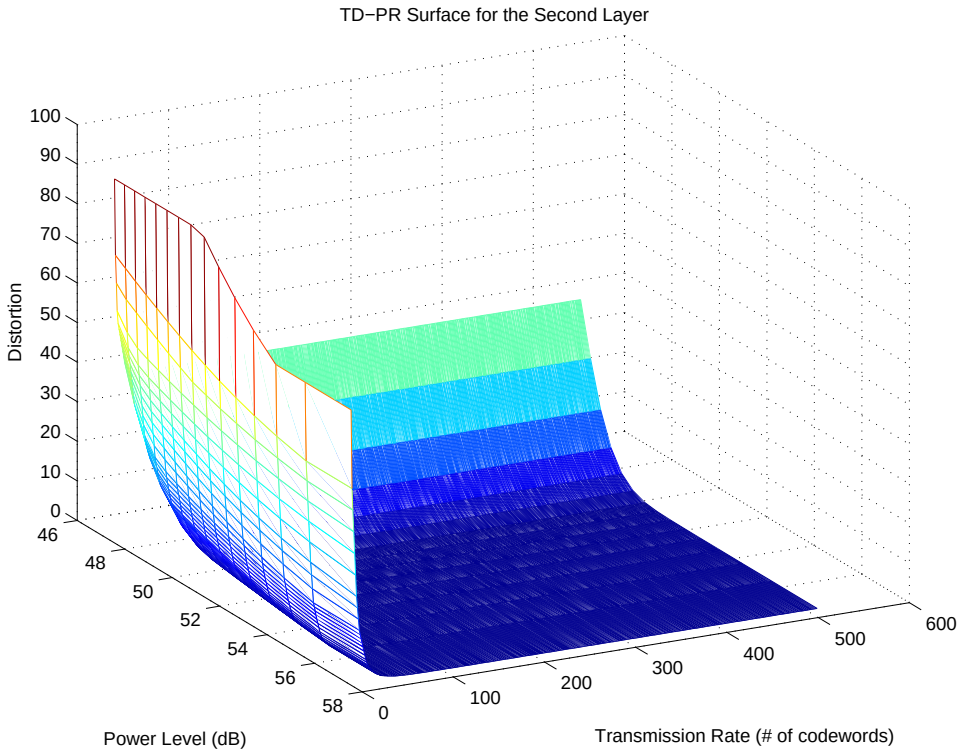


Fig. 25. The operational transmission distortion-power-rate surfaces for the second channel.

changes. The other is about the convexity of the resulting operational transmission distortion-power-rate surfaces. Now let us take a look at the convexity of the operational transmission distortion-power-rate surface generated by the progressive UEP structure. Along the axis of the transmission rate, we consider the convexity of the operational TD-R function. Since all codewords have the same length L_p, the TD-R slope can be represented by the difference of expected distortion between the consecutive rates for a fixed power level p_k, here for convenience we omit the notation of power level. The TD-R slope between rates (codewords) n and $n+1$ for the k-th channel is

$$\frac{\Delta D(\boldsymbol{r}_k^{(1:n)})}{L_p} = \frac{D(\boldsymbol{r}_k^{(1:n)}) - D(\boldsymbol{r}_k^{(1:n+1)})}{L_p}$$

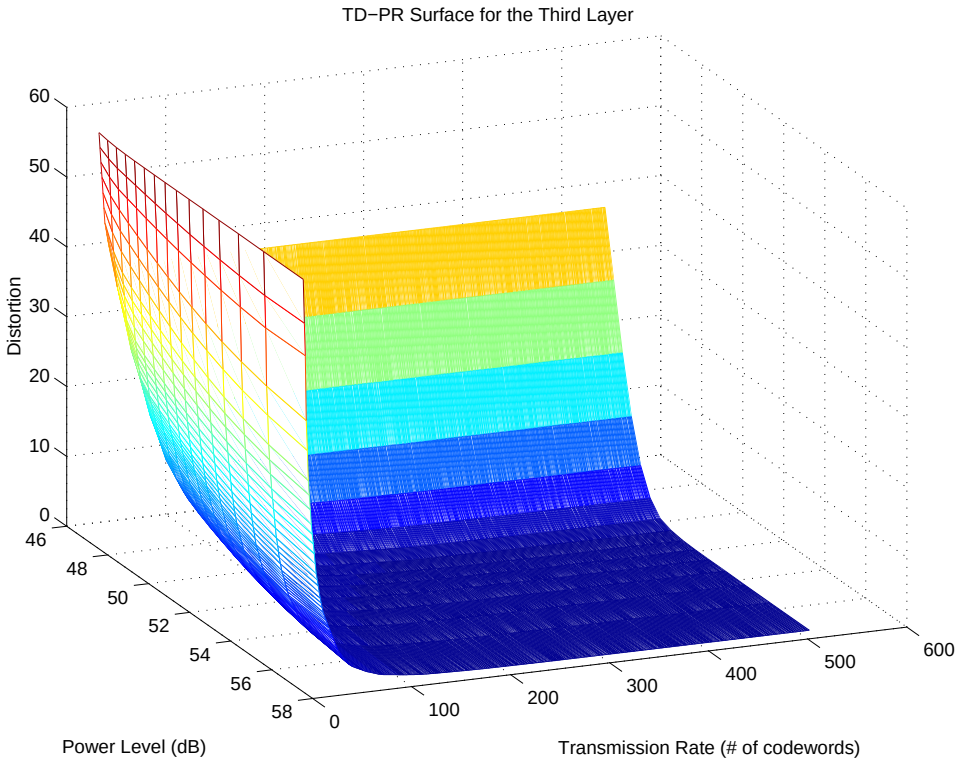


Fig. 26. The operational transmission distortion-power-rate surfaces for the third channel.

$$= \frac{\sum_{i=0}^{n} P_i(\boldsymbol{r}_k^{(1:n)}) - \sum_{i=0}^{n+1} P_i(\boldsymbol{r}_k^{(1:n+1)})}{L_p}$$
$$= \frac{(d_n(\boldsymbol{r}_k^{(1:n)}) - d_{n+1}(\boldsymbol{r}_k^{(1:n+1)})) \prod_{j=1}^{n+1} (1 - p(r_{k_j}))}{L_p}, \qquad (4.19)$$

where we use the property that $d_n(\boldsymbol{r}_k^{(1:n)}) = d_n(\boldsymbol{r}_k^{(1:m)})$ when $m \geq n$, which can be derived from the definition of progressive UEP structure. In equation (4.19), $\prod_{j=1}^{n+1}(1 - p(r_{k_j})$ is strictly decreasing with increasing n , whereas $d_n(\boldsymbol{r}_k^{(1:n)}) - d_{n+1}(\boldsymbol{r}_k^{(1:n+1)})$ corresponds to the MSE reduction caused by the source bits in $(n+1)$-th codeword. If the source D-R function is strictly convex, $d_n(\boldsymbol{r}_k^{(1:n)}) - d_{n+1}(\boldsymbol{r}_k^{(1:n+1)})$ is strictly decreasing with increasing n . The only possible exception is when $r_{k_{n+1}} > r_{k_n}$. Thus there might be points not lying on the convex hull around the rates where channel code rate changes. However if $r_{k_{n+1}} > r_{k_n}$, then $p(r_{kn+1}) > p(r_{k_n})$. Thus the factor

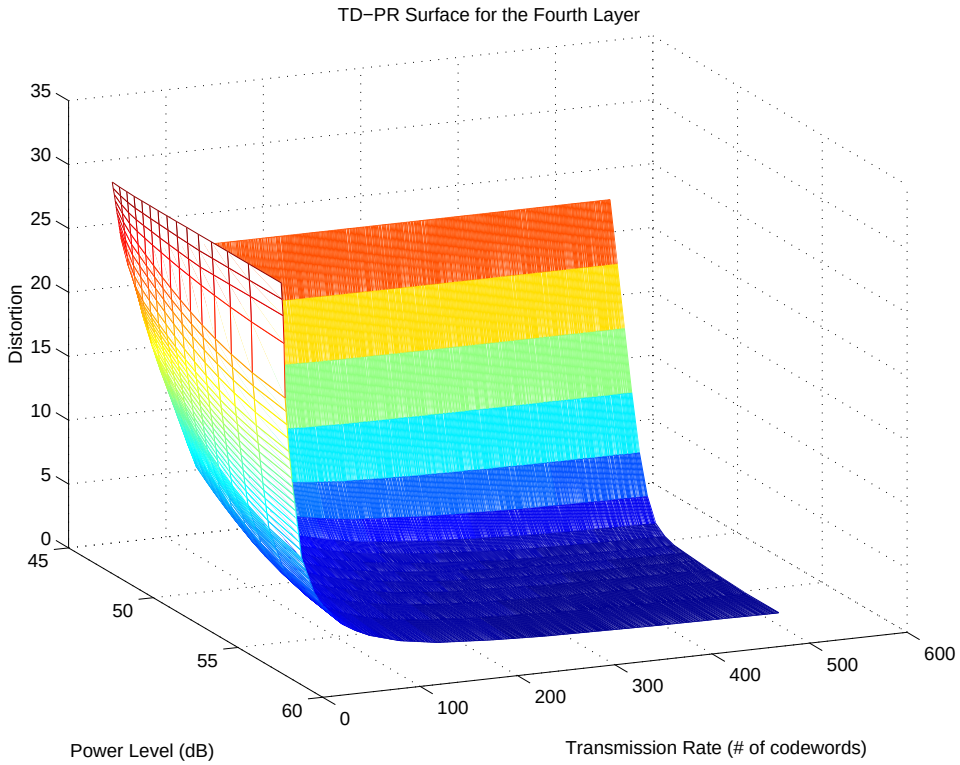


Fig. 27. The operational transmission distortion-power-rate surfaces for the fourth channel.

$1 - p(r_{k_{n+1}})$ will become smaller and give more reduction on the slope to compensate the gain obtained from the rate change, which suggests that most distortion-rate pairs are lying on the convex hull. The simulations show that the valid rate candidates are enough to meet practical requirements [92]. Similarly along the axis of power level, we consider the convexity of the operational TD-P function. The TD-P slope between power levels $p_k^{(n)}$ and $p_k^{(n+1)}$ given a fixed transmission rate for the k-th channel is

$$\begin{aligned} \frac{\Delta D(\boldsymbol{r}_k, p_k^{(n)})}{\Delta p_k^{(n)}} &= \frac{D(\boldsymbol{r}_k^{(n)}, p_k^{(n)}) - D(\boldsymbol{r}_k^{(n+1)}, p_k^{(n+1)})}{p_k^{(n)} - p_k^{(n+1)}} \\ &= \frac{\sum_{i=0}^{N} P_i(\boldsymbol{r}_k^{(n)}, p_k^{(n)}) d_i(\boldsymbol{r}_k^{(n)}) - \sum_{i=0}^{N} P_i(\boldsymbol{r}_k^{(n+1)}, p_k^{(n+1)}) d_i(\boldsymbol{r}_k^{(n+1)})}{p_k^{(n)} - p_k^{(n+1)}} \end{aligned} \tag{4.20}$$

Obviously, with increasing power level, the UEP structure may change according to

the progressive distortion optimal algorithm, i.e., $\boldsymbol{r}_k^{(n)} \neq \boldsymbol{r}_k^{(n+1)}$. Thus it is difficult to analyze the convexity of TD-P function mathematically. However intuitively we know that the source bits would increase with the increase of power levels. This is because higher power levels will lead to higher SNRs, and therefore lower BERs. Thus at a fixed transmission rate, the UEP structure will choose less error protection with more source bits. Indeed from the TD-PR surfaces we observe that the TD-P functions have enough valid rate candidates for the convex hull which we need.

Now let us get back to the equation (4.18) to see how to determine the optimal Lagranges λ_1 and λ_2 in order to satisfy the total transmission rate and the total power level. Computing partial derivatives with respect to the K transmission rates and K power levels, respectively,

$$\begin{cases} \frac{\partial J(\boldsymbol{N},\boldsymbol{p};\lambda_1,\lambda_2)}{N_k} = 0, \\ \frac{\partial J(\boldsymbol{N},\boldsymbol{p};\lambda_1,\lambda_2)}{p_k} = 0, \quad k = 1, \cdots, K \end{cases} \tag{4.21}$$

we obtain the following equations that the optimal transmission rate allocation and optimal power level allocation must satisfy

$$\begin{cases} -\frac{\partial D(\boldsymbol{r}_k,p_k)}{L_p \partial N_k} = \lambda_1, \\ -\frac{\partial D(\boldsymbol{r}_k,p_k)}{\partial p_k} = \lambda_2, \quad k = 1, \cdots, K. \end{cases} \tag{4.22}$$

The equations in (4.22) imply that the optimal transmission rate allocation over K channels should be lying on the equal slope points along the transmission rate axes over K TD-PR surfaces. Similarly the optimal power level allocation over K channels should be lying on the equal slope points along the power level axes over

K TD-PR surfaces. This suggests that we can find the optimal transmission rate allocation (or the optimal power level allocation) by only searching the equal slope points along the transmission rate axes (or along the power level axes). Furthermore, considering the convexity of the TD-PR surfaces over K channels, we can obtain the optimal transmission rate allocation and the optimal power level allocation by using a fast bisectional search.

The complete algorithm is given as follows.

1. Encode the input video source by using an embedded source coder like 3-D ESCOT to produce K source D-R functions.

2. Employ the algorithm in Section C to generate K operational distortion-power-rate (TD-PR) surfaces for K STDC-OFDM channels.

3. For a given total transmission rate and a total power level, using bisectional search over K TD-PR surfaces to get the optimal transmission rate allocation as well as the optimal power level allocation. The search can be started either along the transmission rate axes or along the power level axes. Without losing the generality, let us begin along the transmission rate axes.

   (a) Start with an initial transmission rate vector $\boldsymbol{N}^{(0)}$, whose components are lying on the equal slope points with the slope of $\lambda_1^{(0)}$ over K TD-PR surfaces along the transmission rate axes.

   (b) Use the bisectional search on the equal slope points along the transmission rate axes to adjust the λ_1 to meet the constraint requirement of the total transmission rate. Thus we get a new transmission rate allocation vector and power level allocation vector $\boldsymbol{N}^{(t)}$ and $\boldsymbol{p}^{(t)}$, as well as the lagrangian $J(\boldsymbol{N}^{(t)}; \boldsymbol{p}^{(t)}; \lambda_1^{(t)})$. Notice that at this moment, the components of $\boldsymbol{p}^{(t)}$ are

not necessarily lying on the equal slope points along the power level axes.

(c) Next start with $\boldsymbol{p}^{(t)}$, using the bisectional search on the equal slope points along the power level axes to adjust the λ_2 to meet the constraint requirement of the total power level. Thus we get a new transmission rate allocation vector and power level allocation vector $\boldsymbol{N}^{(t+1)}$ and $\boldsymbol{p}^{(t+1)}$, as well as the lagrangian $J(\boldsymbol{N}^{(t+1)}; \boldsymbol{p}^{(t+1)}; \lambda_2^{(t+1)})$. Notice that at this moment, the components of $\boldsymbol{N}^{(t+1)}$ are not necessarily lying on the equal slope points along the transmission rate axes.

(d) Check the convergence of $\Delta J(\boldsymbol{N}^{(t)}; \boldsymbol{p}^{(t)};)$. If it does not converge, let $\boldsymbol{N}^{(t)} = \boldsymbol{N}^{(t+1)}$ and $\boldsymbol{p}^{(t)} = \boldsymbol{p}^{(t+1)}$, go to Step (b). Otherwise stop.

By now we get the optimal transmission rate allocation and the optimal power level allocation , and finally the optimal UEP structures over K channels. We can use the obtained UEP structures to packetize the source bits and send them into wireless channels. The bisectional search, we employed in Step 3, is very fast compared with the generation of operational TD-PR surfaces. It only takes two or three iterations to converge to a stable solution. Of course the generation of operational TD-PR surfaces can also be suitable for wireless video applications because we employed the fast progressive algorithm in Section C for a finite set of available power levels. In addition, the K TD-PR surfaces only need to be produced once regardless of the change of the target total transmission rate and the target total power level.

The complete algorithm above consists of three steps. Under stringent delay requirements, the first two steps, used to generate K TD-PR surfaces, could be performed off-line. In general the resulting TD-PR surfaces will vary from one video sequence to another for a given embedded source encoder. However if we use the average result of several well-known benchmark video sequences like *foreman, akiyo,*

miss American, and so on, the first two steps could be computed off-line. Only the third step needs to be performed on the fly. Therefore the time delay of this algorithm above will be minimal at this situation.

E. Numerical Results

For OFDM channels, we assume a total of $N_c = 128$ sub-carriers. The video sequence is compressed and partitioned into $K = 4$ layers. Thus the number of sub-carriers of one OFDM frame each layer is $N_c/K = 32$. We also assume the time period of the total OFDM sub-carriers is $T_s = 200\mu s$ for one video sequence.

As we do in Chapter III, we experiment with standard color QCIF (144×176) *akiyo* video sequences. Each sequence is in YUV format and has 288 pictures. We used 3-D ESCOT coding bitstream with the GOP size 16 frames and a fixed frame rate 30 fps in all experiments. We used a 16-CRC code with generator polynomial 0x15935. The generator polynomials of the RCPC code were $(0117, 0127, 0155, 0171)$, the mother code rate was 1/4, and the puncturing rate was 8. Thus, the set of RCPC code rates is $\mathcal{R} = \{8/32, 8/31, \ldots, 8/10, 8/9\}$. The decoding of the RCPC code is done with a list Viterbi algorithm where the maximum number of candidate paths is 100. Under a given BER as shown in Section B, we implement 1000 experiments for each RCPC code rate r_i to get the $p(r_i)$, which is the probability that a codeword protected with rate r_i cannot be correctly decoded with the RCPC decoder. Figure 12 in Chapter III demonstrates all $p(r_i)$s we used for these 24 RCPC codes given a finite set of BERs. For each RCPC code, the range of BER we considered is approximately between 1% and 20%. Thus through the equation (4.12), we can get the probability $P_i(\boldsymbol{r})$, which we use in our optimal algorithm.

Figure 24 - 27 in Section D displays the operational transmission distortion-

power-rate surfaces for the four STDC-OFDM channels. We generate these surfaces by using the algorithm in Section C for the range of power level $[47.528\text{dB}, 57.494\text{dB}]$ while the transmission rate (number of codewords) is from 1 to 512. The modulation scheme is QDPSK and the decision memory order $P = 4$.

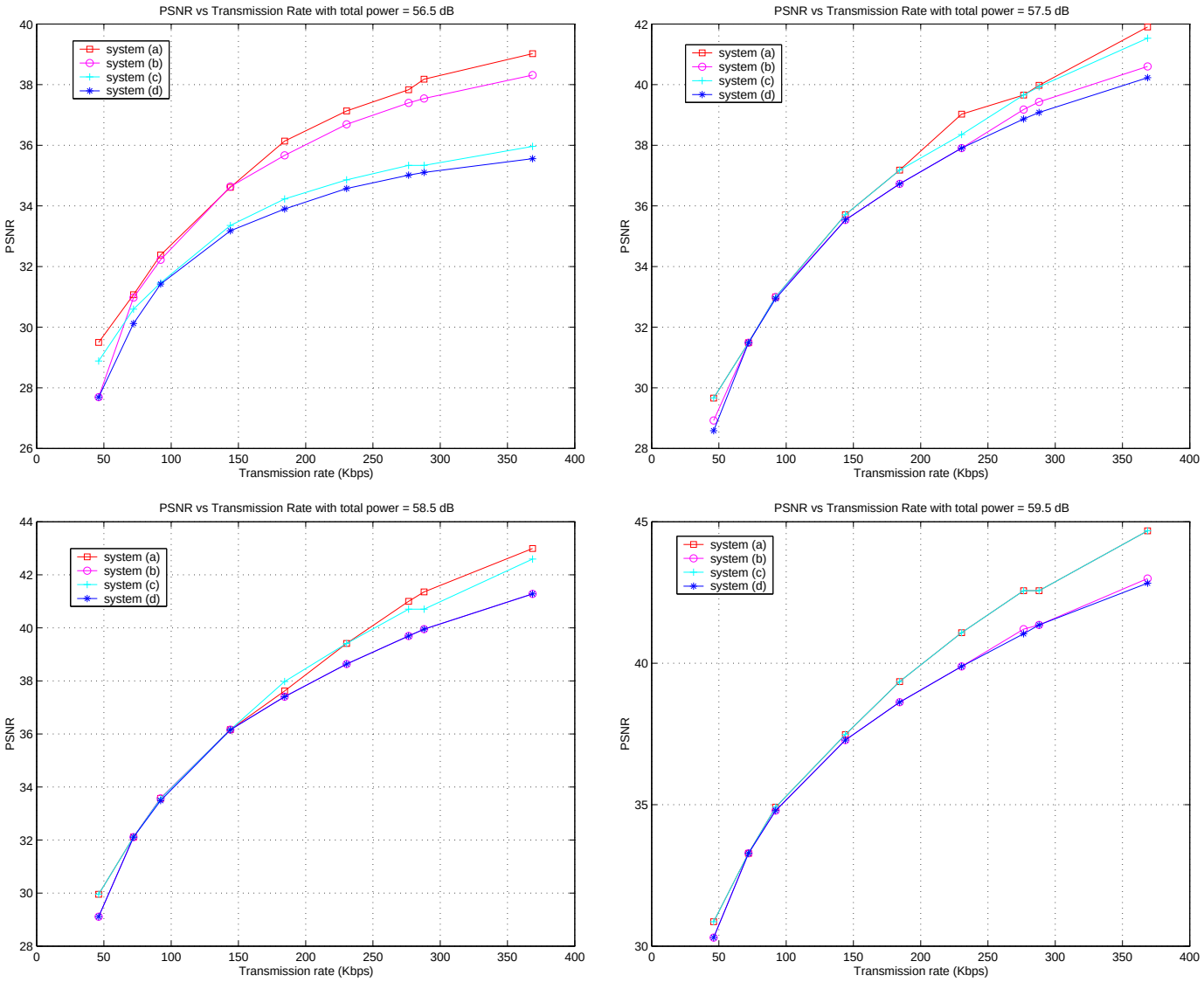


Fig. 28. PSNR vs. transmission rate with total power $\mathbb{P} = 56.5, 57.5, 58.5$, and 59.5 dB.

Figure 28 shows the peak signal to reconstruction noise ratio versus the total transmission rate of $K = 4$ STDC-OFDM channels in four systems: (a) joint opti-

mal transmission rate allocation/opitmal power level allocation, (b) uniform transmission rate allocation/optimal power level allocation, (c) optimal transmission rate allocation/uniform power level allocation, and (d) uniform transmission rate allocation/uniform power level allocation over $K = 4$ STDC-OFDM channels. For each total power level $\mathbb{P}$, we only consider the Y component and calculate PSNR at 9 total transmission rates over $K = 4$ channels: 46.08, 72.0, 92.16, 144.0, 184.32, 230.4, 276.48, 288.0, and 368.64 Kbps. We consider various total power levels $\mathbb{P} =$ 56.5, 57.5, 58.5, and 59.5 dB. We limit the dynamic range of the power level during power allocation such that the BER for each STDC-OFDM channel is approximately between 1% and 15%.

The important observation is that at moderate total power levels such as $\mathbb{P} = 56.5$ dB , systems (a), (b), and (c) achieve better PSNR performance than system (d). At these moderate total power levels, system (b) gains the better PSNR improvements than system (c). The reason is that when the total power level is not high enough, the reasonable power level allocation has dominant effects to the correctly receiving codewords. However with the total power level increasing, these PSNR improvements get small. Furthermore at some high total power levels such as $\mathbb{P} \geq 57.5$ dB, system (c) achieves better PSNR improvements than system (b). This is because in this situation, the high power level always guarantees the correct decoding, thus power level allocation will become useless. Obviously at all these situations system (a) we proposed always achieves the best PSNR performance.

Figure 29 shows the effect of the change of total power level $\mathbb{P}$ to the system performance. We can see that with the increase of total power level $\mathbb{P}$, all four systems basically have increasing PSNR performance. In particular with the increase of the transmission rate and power level, the PSNR performance is also increasing stably. This demonstrates the advantage of the systems we proposed.

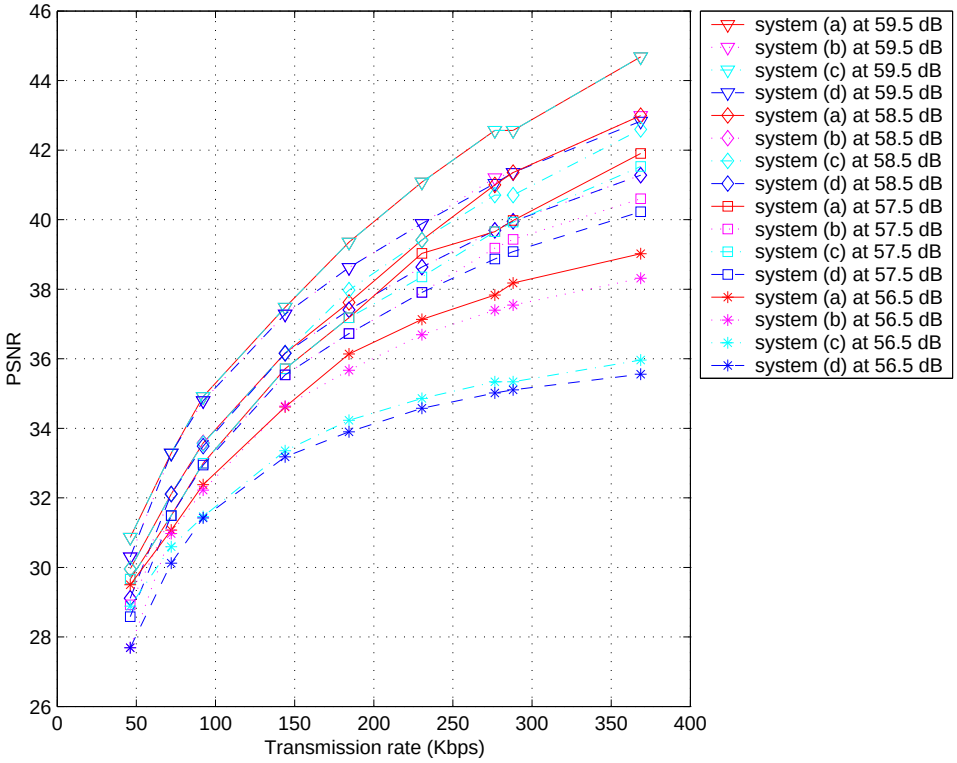


Fig. 29. Peak-signal-to-noise ratio vs. transmission rate with various total power levels $\mathbb{P}$.

Meanwhile, Figure 30 illustrates the received PSNR versus the total power level $\mathbb{P}$ with the total transmission rate $\mathbb{R} = 46.08$, 72.0, 184.32, and 288.0 Kbps. We can see that when the total power $\mathbb{P} \leq 57.5$ dB with the transmission rate between 72 and 288 Kbps, system (b) gains better PSNR performance than system (c). However when the total power level $\mathbb{P} \geq 57.5$ dB with the total transmission rate between 184.32 and 288 Kbps, system (c) gets better PSNR performance system (b). Moreover system (a) we proposed achieves the best PSNR performance in all these situations. However these curves imply that system (a) is not suitable for any total target transmission rate and any total power level. In practice the total power level should have a reasonable dynamic range in order to let the joint optimal power level and transmission rate allocation play best.

Figure 31 displays the effect of the change of the total transmission rate to the system performance. One important observation is that with the increase of the total transmission rate, all four systems have steady PSNR performance improvement. This demonstrates again the advantage of the systems we proposed, which can implement the robust and efficient transmission of video sequences over wireless channels. Figure 32 gives the optimal power level and transmission rate allocation (i.e., system (a)) by using our joint optimal power level and optimal transmission rate allocation algorithms at total power level $\mathbb{P} = 56.5$ dB and total transmission rate $\mathbb{R} = 184.32$ Kbps. At the optimal work point, the received PSNR = 36.1377 dB.

Finally we discuss the time complexity of the algorithm we proposed. As we claimed in Section D, the search of Step 3 has been proved to be fast compared with the generation of operational TD-PR surfaces. It only takes two or three iterations to converge to a stable solution, which takes a time of less than one second. Table I demonstrates the time complexity of generating the operational TD-PR surfaces over 4 STDC-OFDM channels. We run our codes on a DELL 1.8 GHz computer, which is in C language. The target transmission rate (number of codewords) is 512, which corresponds to the transmission rate of 368.64 Kbps for the video sequence *Akiyo*. The intermediate rates (number of codewords) are from 1 to 511. The samples of power level are $47.528, 47.946, 48.373, 48.811, 49.263, 49.732, 50.224, 50.745, 51.305, 51.915, 52.595, 53.378, 54.322, 55.558$ and 57.494 dB, which correspond to the BER between 15% and 1%,respectively. Obviously the algorithm we proposed is a good choice for video delivery applications. Moreover if we use the average result of several typical video sequences for TD-PR surfaces, we could generate these surfaces off-line. Thus we could realize faster implementation.

Table I. The time complexity of generating the operational TD-PR surfaces. (Unit: second)

TD-PR 1	TD-PR 2	TD-PR 3	TD-PR 4
13.472	13.74	14.341	14.843

F. Conclusions

A progressive video delivery transmission framework over wireless channels by combining layered source coding and space-time differentially coded OFDM system has been proposed. Unlike the existing scheme, we propose to transmit all the layers of a video sequence in parallel through different STDC-OFDM channels. The independently decodable layers from 3-D ESCOT and UEP structure provide more system error resilience even during deep fading period in wireless channels. We propose to use progressive joint source-channel codes to generate the operational transmission distortion-power-rate (TD-PR) surfaces for multiple wireless channels. A new fast algorithm is proposed to obtain the optimal power allocation and optimal transmission rate allocation among multiple channels, subject to constraints on the total transmission rate and the total power level. Based on the characteristics of the convex of TD-PR surfaces, the new algorithm can use the 'equal slope' argument to quickly solve the optimal problem. Experiments show that our scheme can run for wireless video delivery applications. We achieve significant PSNR improvement over a non-optimal system for multi-layer video space-time differentially coded OFDM system with the same total power level and total transmission rate.

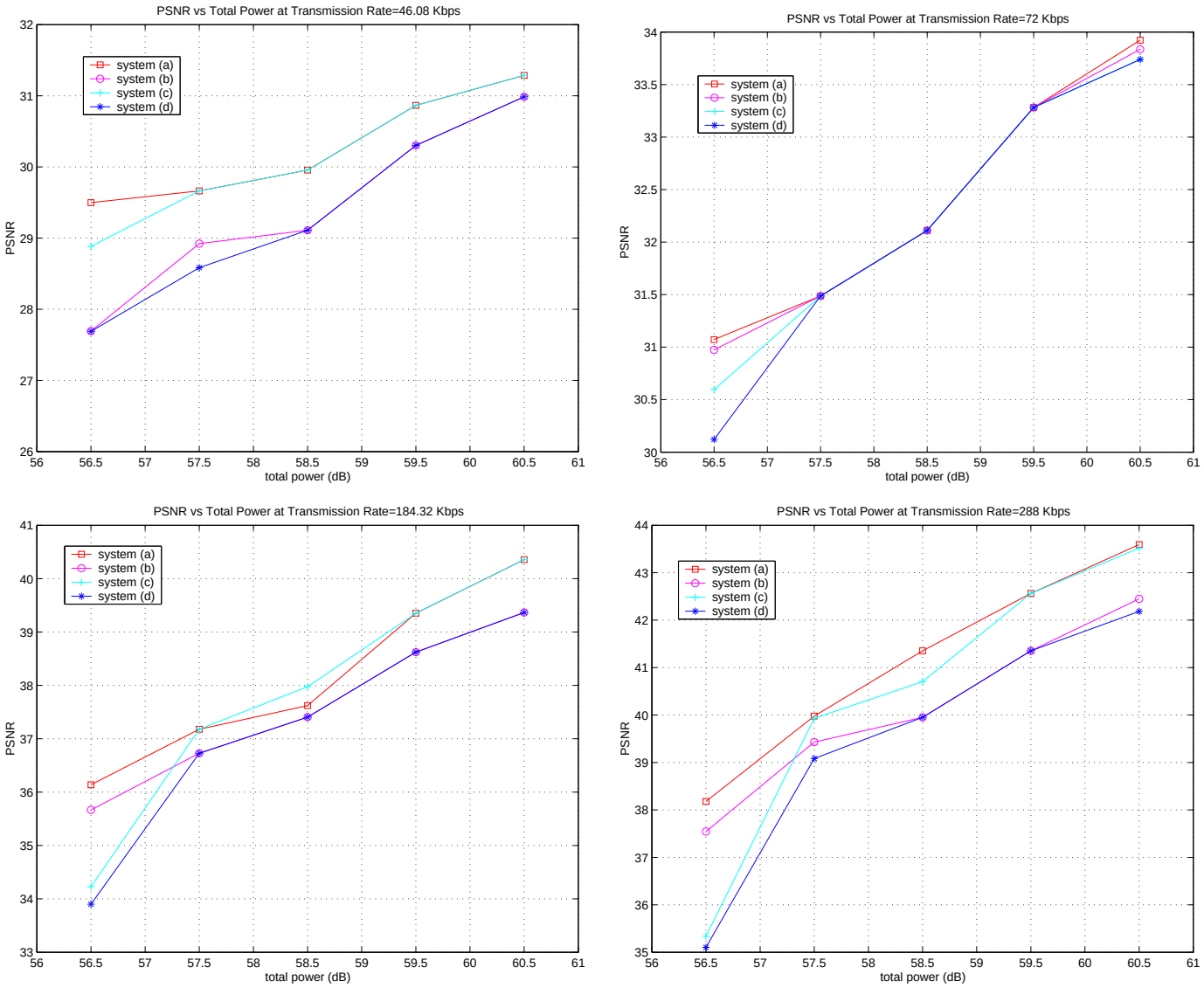


Fig. 30. PSNR vs. the total power level $\mathbb{P}$ with transmission rate 46.08, 72, 184.32 and 288 Kbps.

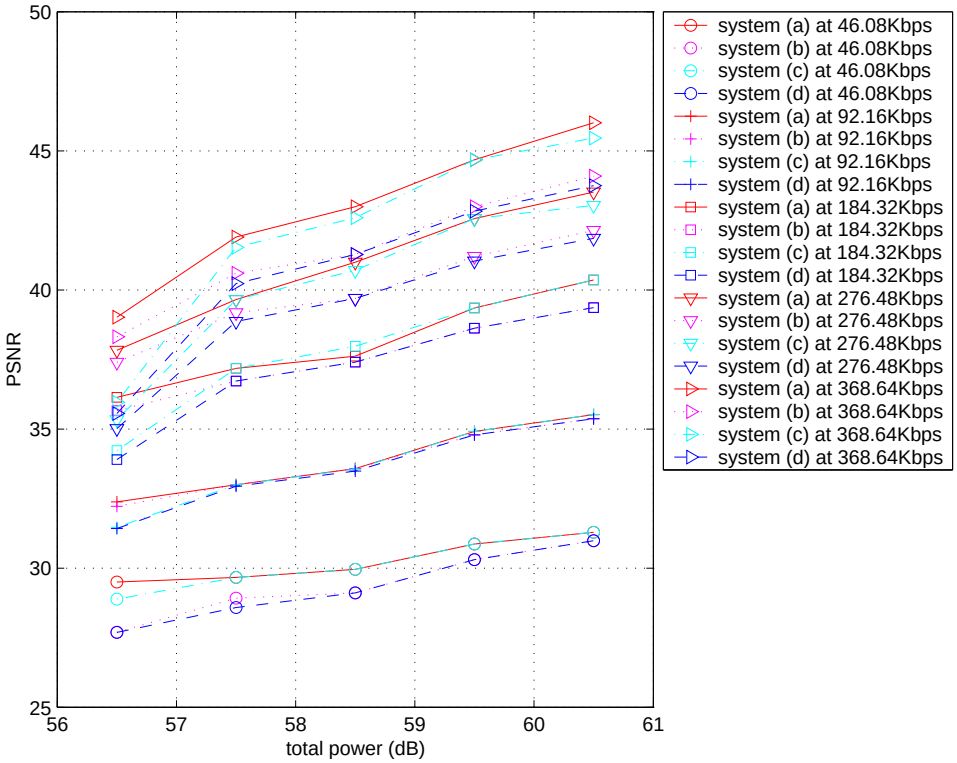


Fig. 31. PSNR vs. the total power level $\mathbb{P}$ with various transmission rates.

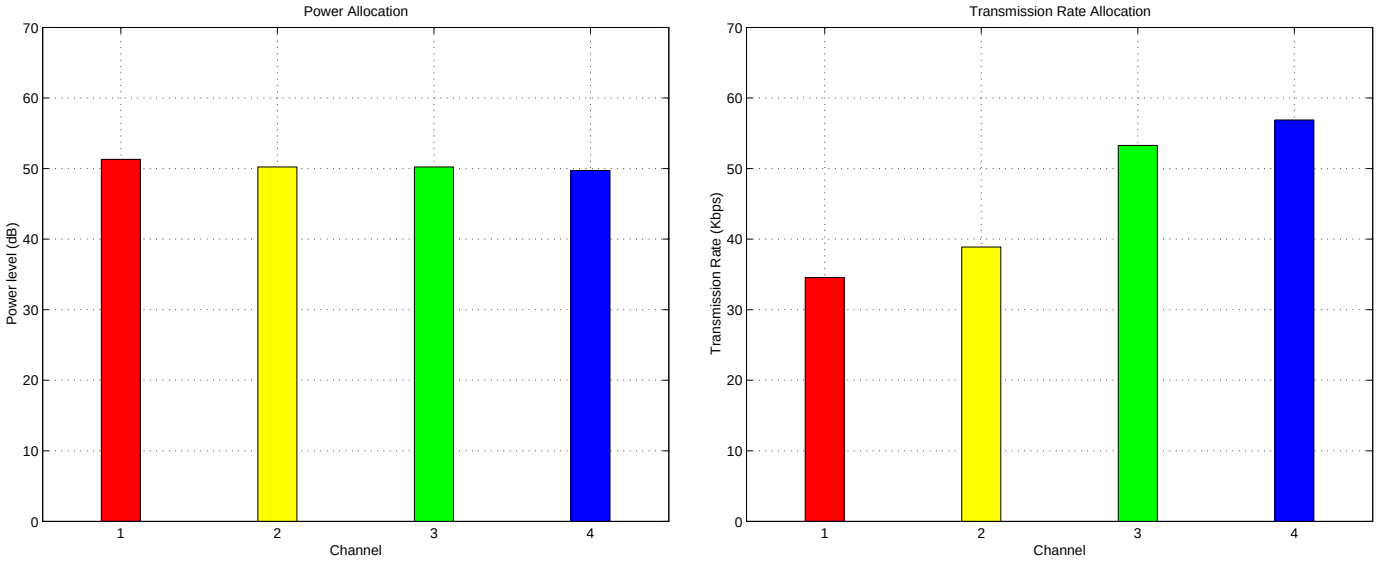


Fig. 32. Optimal power level and transmission rate allocation with a total power $\mathbb{P} = 56.5$dB and transmission rate $\mathbb{R} = 184.32$Kbps.

CHAPTER V

CONCLUSIONS

In this dissertation, we have presented several schemes for compressed video transmissions over wireless networks, namely, joint error control and power allocation for video transmission over CDMA networks with multiuser detection, optimal resource allocation for wireless video over CDMA networks with large-system performances, and progressive video delivery over wideband wireless channels using space-time differentially coded OFDM systems. These proposed schemes provide optimal resource allocation methods in various wireless channels. These frameworks of joint transmission rate control/power control essentially creates different qualities of service for different bandwidths and different transmitted power levels. The proposed schemes make optimal use of the available bandwidth and link budgets to transmit the video signals over wireless media, in order to reconstruct the signals at the receiver with the highest quality. Generally, as the wireless channel conditions change, these schemes can scale the video streams and transport the scaled video streams to receivers with a smooth change of perceptual quality. We believe these frameworks we proposed are good candidates for efficient and robust wireless video transmission applications.

REFERENCES

[1] P. Chaudhury, W. Mohr, and S. Onoe, "The 3GPP proposal for IMT-2000," *IEEE Communications Magazine*, pp. 72-81, December 1999.

[2] *Requirements for a 3G Network Based on Internet Protocol (All-IP) with Support for TIA/EIA-41 Interoperability*, 3GPP2 specification, October 2000.

[3] L. Hanzo, P. Cherriman, and E. Kuan, "Interactive cellular and cordless video telephony: state-of-the-art system design principles and expected performance," *Proc. of the IEEE*, vol. 88, pp. 1388-1413, September 2000.

[4] B. Girod and N. Farber, "Wireless Video," *Compressed Video over Networks*, A. Reibman, M.-T. Sun, eds., New York: Marcel Dekker, 2000.

[5] D. Wu, Y. Hou, and Y.-Q. Zhang, "Scalable video coding and transport over broad-band wireless networks," *Proc. of the IEEE*, vol. 89, pp. 6-20, January 2001.

[6] B.-J. Kim, Z. Xiong, and W. A. Pearlman, "Very low bit-rate embedded video coding with 3-D set partitioning in hierarchical trees (3-D SPIHT)," *IEEE Trans. on Circuits and Systems for Video Tech.*, vol. 10, pp. 1365-1374, December 2000.

[7] J. Xu, Z. Xiong, S. Li, and Y.-Q. Zhang, "3-D embedded subband coding with optimal truncation (3-D ESCOT)," *Applied and Computational Harmonic Analysis*, vol. 10, pp. 290-315, May 2001.

[8] P. Xiong, F. Wu, S. Li, Z. Xiong, and Y.-Q. Zhang, "Exploiting temporal correlation with flexible block-size motion alignment for 3D wavelet coding," in *Proc. VCIP'04*, San Jose, CA, January 2004.

[9] P. Chen and J. W. Woods, "Improved MC-EZBC with quarter-pixel motion vectors, MPEG document, *ISO/IEC JTC1/SC29/WG11, MPEG2002/M8366*, Fairfax, VA, May 2002.

[10] P. Chen and J. Woods, "Bidirectional MC-EZBC with lifting implementation, *IEEE Trans. Circuit and Systems for Video Tech.*, to appear.

[11] A. Secker and D. Taubman, "Lifting-based invertible motion adaptive transform (LI-MAT) framework for highly scalable video compression, *IEEE Trans. on Image Processing*, vol. 12, no. 12, pp. 1530-1542, December 2003.

[12] M. Flierl and B. Girod, "Video coding with motion-compensated lifted wavelet transforms," submitted to *Signal Processing: Image Communication*, June 2003.

[13] J.-R. Ohm, "Motion-compensated wavelet lifting filters with flexible adaptation, in *Proc. Intl. Workshop on Digital Communications*, Capri, 2002.

[14] W. Jakes, *Microwave Mobile Communications*, New York, John Wiley and Sons, 1974.

[15] R. Stedman, H. Gharavi, L. Hanzo, and R. Steele, "Transmission of subband-coded images via mobile channels," *IEEE Trans. Circuits and Systems for Video Tech.*, vol. 3, no. 1, pp. 15-26, February 1993.

[16] Y. Liu and Y. Q. Zhang, "A mobile data code division multiple access (CDMA) system with power control and its application to low-bit-rate image transmission," in *Proc. VTC'93*, pp. 770-773, 1993.

[17] J. Hagenauer, "Rate-compatible punctured convolutional codes (RCPC Codes) and their applications," *IEEE Trans. Commun.*, vol. 36, pp. 389-400, April 1988.

[18] A. Albanese, J. Blomer, J. Edmonds, M. Luby, and M. Sudan, "Priority encoding transmission," *IEEE Trans. Information Theory*, vol. 42, no. 6, pp. 1737-1744, November 1996.

[19] P. A. Chou, A. Mohr, A. Wang, and S. Mehrota, "Error control for receiver-driven layered multicast of audio and video," *IEEE Trans. Multimedia*, vol. 3, no. 1, pp. 108-122, March 2001.

[20] R. Van Dyck and D. J. Miller, "Transport of wireless video using separate, concatenated, and joint source-channel coding," *Proc. of the IEEE*, vol. 87, pp. 1734-1750, October 1999.

[21] ISO/IEC JTC 1/SC29/WG11 N2687, *MPEG-4 Video Verification Model Version 13.0,* March, 1999.

[22] H. Radha, M. van der Schaar, and Y. Chen, "The MPEG-4 fine-grained scalable video coding method for multimedia streaming over IP," *IEEE Trans. Multimedia*, vol. 3, no. 1, pp. 53-68, March 2001.

[23] Sachs, D.G.; Anand, R.; Ramchandran, K., "Wireless image transmission using multiple-description based concatenated codes," in *Proceedings of the Data Compression Conference (DCC 2000)*, pp. 569.

[24] J. Rulnick and N. Bambos, "Mobile power management for wireless communication networks," *Wireless Networks*, vol. 3, no. 1, pp. 3-14, 1997.

[25] R. Yates and C. Huang, "Integrated power control and base station assignment," *IEEE Trans. Vehicular Tech.*, vol. 44, no. 3, pp. 638-644, August 1995.

[26] S. Papavassiliou and L. Tassiulas, "Improving the capacity in wireless networks through integrated channel base station and power assignment," *IEEE Trans.*

Vehicular Tech., vol. 47, no. 2, pp. 417-427, June 1998.

[27] A. Grandhi, J. Zander, and R. Yates, "Constrained power control," *Intl. J. Wireless Personal Communications*, vol. 1, no. 1, pp. 1257-1270, 1995.

[28] M. Polley; S. Wee; W. Schreiber, "Hybrid channel coding for multiresolution HDTV terrestrial broadcasting," in *Proc. Int. Conf. Image Processing*, vol. 1, pp.243-247, November 1994.

[29] F. H. Liu; P. Ho; V.Cuperman, "Joint source and channel coding using a non-linear receiver," in *Proc. ICC*, vol. 3, pp.1502-1507, June 1993.

[30] Y. Eisenberg, C. E. Luna, T. N. Pappas, R. Berry, and A. K. Katsaggelos, "Joint source coding and transmission power management for energy efficient wireless video commninications," *IEEE Trans. Circuits and Systems for Video Tech.*, vol. 12, no. 6, pp. 411-424, June 2002.

[31] S. Zhao, Z. Xiong and X. Wang, "Joint error control and power allocation for video transmission over CDMA networks with multiuser detection," *IEEE Trans. Circuits and Systems for Video Tech.*, vol. 12, no.6, pp. 425-437, June 2002.

[32] S. Zhao, Z. Xiong and X. Wang, "Optimal resource allocation for wireless video over CDMA networks," *IEEE Trans. on Mobile Computing*, vol. 3, no. 3 July-Sept. 2004.

[33] B. Lu, X. Wang, and Y. Li, "Iterative receivers for space-time block coded OFDM systems in dispersive fading channels," in *IEEE Global Telecommunications Conference*, vol. 1, pp. 514 - 518, Nov. 25-29, 2001.

[34] S.M.Alamouti, "A simple transmit diversity technique for wireless communication," *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 8,

pp.1451-1458, October 1998.

[35] Z. Yang, B. Lu, and X. Wang, "Bayesian Monte Carlo multiuser receiver for space-time coded multicarrier CDMA systems," *IEEE Journal on Selected Areas in Communications*, vol. 19, no. 8, pp. 1625 - 1637, Aug. 2001.

[36] B. Lu and X. Wang, "Iterative receivers for multiuser space-time coding systems," *IEEE Journal on Selected Areas in Communications*, vol. 18, no. 11, pp. 2322 - 2335, Nov. 2000.

[37] B.L. Hughes, "Differential space-time modulation," in *Proc. 1999 IEEE Wireless Communications and Networking Conference (WCNC'99)*, New Orleans, LA, Mar. 1999.

[38] V. Tarokh and H. Jafarkhani, "A differential detection scheme for transmit diversity," *IEEE Journal on Selected Areas in Communications*, vol. 18, no. 7, pp. 1169-1174, July 2000.

[39] Y. Liu and X. Wang, "Multiple-symbol decision-feedback differential space-time decoding in fading channels," *EURASIP Journal on Applied Signal Processing, Special Issue on Space-Time and Its Applications*, vol. 2002, no. 3, pp. 297-304, March 2002.

[40] H. Zhang, X-G. Xia, Q. Zhang, and W. Zhu, "Precoded OFDM with adaptive vector channel allocation for scalable video transmission over frequency-selective fading channels," *IEEE Trans. on Mobile Computing*, vol. 1, no. 2, pp. 132-142, April-June 2002.

[41] C. Kuo, C. Kim, and C-C. J. Kuo, "Robust video transmission over wideband wireless channel using space-time coded OFDM systems," in *IEEE Wireless*

Communications and Networking Conference (WCNC2002), vol. 2, pp. 931-936, March 2002.

[42] D. Dardari, M. G. Martini, M. Milantoni, and M. Chiani, "MPEG-4 video transmission in the 5GHz band through an adaptive OFDM wireless scheme," in *13th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, vol. 4, pp. 1680-1684, 2002.

[43] Q. Zhang, W. Zhu and Y-Q. Zhang, "Network-adaptive scalable video streaming over 3G wireless network," in *IEEE International Conference on Image Processing* , vol. 3, pp. 579 - 582, Oct. 7-10, 2001.

[44] Q. Zhang;, W. Zhu, Z. Ji, and Y-Q. Zhang, "A power-optimized joint source channel coding for scalable video streaming over wireless channel," in *IEEE International Symposium on Circuits and Systems*, vol. 5, pp. 137 - 140, May 6-9, 2001.

[45] P. G. Sherwood, X. Tian, and K. Zeger, "Channel code blocklength and rate optimzation for progressive image transmission," in *Proc. WCNC-1999 IEEE Wireless Communications and Networking Conference,* pp. 978-982, New Orleans, LA, September 1999.

[46] V. Stankovic, R. Hamzaoui, Y. Charfi, and Z. Xiong, "Real-time unequal error protection algorithms for progressive image transmission," *IEEE Journal on Selected Areas in Communications: Special Issue on Recent Advances in Wireless Multimedia,* vol. 21, no. 10, pp. 1526-1535, December 2003.

[47] S. Zhao, Z. Xiong, X. Wang, and J. Hua, "Progressive video delivery over wideband wireless channels using space-time differentially coded OFDM systems," Submitted to *IEEE Trans. on Mobile Computing.*

[48] S. Zhao, Z. Xiong, and X. Wang, "Joint error control and power allocation for video transmission over CDMA networks with multiuser detection," in *IEEE International Conference on Communications (ICC 2002),* vol. 5, pp. 3212-3216, May 2002.

[49] S. Zhao, Z. Xiong, and X. Wang, "Optimal resource allocation for wireless video over CDMA networks," in *Proc. IEEE International Conference on Media and Expo (ICME 2003),* Vol. 1, pp. 277-280, July 2003.

[50] X. Zhou, S. Zhao, and X. Wang, "Adpative and integrated video communications for wireless ATM in WCDMA systems," in *VTC 2001 IEEE Vehicle Technology Conference,* vol. 3, pp. 2147-2151, May 2001.

[51] Y. Sun, S. Zhao, Z. Xiong, and X. Wang, "Error robust video transmission over wireless IP networks with multiuser detection," in *ISCAS 2002 IEEE International Symposium on Circuits and Systems,* vol. 1, pp. 26-29, May 2002.

[52] L. Rizzo, "Effective erasure codes for reliable computer communication protocols," *ACM Computer Communication Review*, vol. 27, no. 2, pp. 24-36, April 1997.

[53] B. Girod and N. Farber, "Feedback-based error control for mobile video transmission," *Proc. of the IEEE*, vol. 87, pp. 1707-1723, October 1999.

[54] J. Hagenauer and T. Stockhammer, "Channel coding and transmission aspects for wireless multimedia," *Proc. of the IEEE*, vol. 87, pp. 1764-1777, October 1999.

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

[56] ITU-T Recommendation H.263, version 2, *Video coding for low bitrate communication,* January 1998.

[57] L. Yang, F. C. M. Martins, and T. R. Gardos, "Improving H.263+ scalability performance for very low bit rate applications," in *Proc. VCIP'99*, San Jose, CA, January 1999.

[58] J. Arnold, M. Frater, and Y. Wang, "Efficient drift-free signal-to-noise ratio scalability," *IEEE Trans. Circuits and Systems for Video Tech.*, vol. 10, pp. 70–82, February 2000.

[59] S. Li, F. Wu and Y.-Q. Zhang, "Study of a new approach to improve FGS video coding efficiency," MPEG-99/m5583, Maui, HI, December 1999.

[60] E. Steinbach, N. Farber, and B. Girod, "Standard compatible extension of H.263 for robust video transmission in mobile environments," *IEEE Trans. Circuits and Systems for Video Tech.*, vol. 7. pp. 872-881, December 1997.

[61] G. Karlsson and M. Vetterli, "Three dimensional subband coding of video," in *Proc. ICASSP'88*, pp. 1100-1103, April 1988.

[62] D. Taubman and A. Zakhor, "Multirate 3-D subband coding of video," *IEEE Trans. Image Processing*, vol. 3, pp. 572-588, September 1994.

[63] J. R. Ohm, "Three-dimensional subband coding with motion compensation," *IEEE Trans. on Image Processing*, vol. 3, pp. 559-571, September 1994.

[64] C. I. Podilchuk, N. S. Jayant, and N. Farvardin, "Three-dimensional subband coding of video," *IEEE Trans. on Image Processing*, vol. 4, pp. 125-139, February 1995.

[65] S. J. Choi and J. W. Woods, "Motion-compensated 3-D subband coding of video," *IEEE Trans. Image Processing*, vol. 8, pp. 155–167, February 1999.

[66] A. Said and W. A. Pearlman, "A new, fast, and efficient image codec based on set partitioning in hierarchical trees," *IEEE Trans. Circuits and Systems for Video Tech.*, vol. 6, pp. 243-250, June 1996.

[67] J. M. Shapiro, "Embedded image coding using zerotrees of wavelet coefficients," *IEEE Trans. on Signal Processing*, vol. 41, no. 12, pp. 3445-3463, December 1993.

[68] University of California at Los Angeles (UCLA) Image Communications Lab, "Wavlet image coding: PSNR results," Web site http://www.icsl.ucla.edu/~ipl/psnr_results.html, 1998.

[69] A. Wang, Z. Xiong, P. A. Chou, and S. Mehrotra, "Three-dimensional wavelet coding of video with global motion compensation," in *Proc. DCC'99*, Snowbird, UT, March 1999.

[70] M. Antonini, M. Barlaud, P. Mathieu, and I. Daubechies, "Image coding using wavelet transform," *IEEE Trans. Image Processing*, vol. 1, pp. 205-221, April 1992.

[71] A. Host-Madsen and X. Wang, "Performance of blind and group-blind multiuser detectors," *IEEE Trans. Information Theory*, vol. 48, no. 7, pp. 1849-1872, July 2002.

[72] A. Host-Madsen and X. Wang, "Performance of subspace-based multiuser detection," in *Proc. ISIT'01*, Washington, DC, June 2001.

[73] A. Host-Madsen and X. Wang, "Performance analysis of subspace-based blind and group-blind multiuser detection," in *Proc. ISITA'00*, Honolulu, HI, November 2000.

[74] A. Host-Madsen and X. Wang, "Performance of blind multiuser detectors," in *Proc. 38th Annual Allerton Conference on Communications, Control and Computing*, Monticello, IL, October 2000.

[75] H. V. Poor and S. Verdu, "Probability of error in MMSE multiuser detection," *IEEE Trans. Inform. Theory*, vol. 43, pp. 858-871, May 1997.

[76] X. Wang and H. Poor, "Blind multiuser detection: A subspace approach," *IEEE Trans. Information Theory*, vol. 44, pp. 677-691, March 1998.

[77] X. Wang, "Blind Multiuser Detection," *Encyclopedia of Telecommunications*, J. G. Proakis, Ed., New York, Wiley & Sons, 2002.

[78] X. Wang and A. Host-Madsen, "Group-blind multiuser detection for uplink CDMA," *IEEE Journal on Selected Areas in Communications*, vol. 17, pp. 1971-1984, November 1999.

[79] X. Wang and H. V. Poor, "Blind equalization and multiuser detection for CDMA communications in dispersive channels," *IEEE Trans. Commun.*, vol. 46, pp. 91-103, January 1998.

[80] M. Ruf and J. Modestino, "Operational rate-distortion performance for joint source and channel coding of images," *IEEE Trans. Image Processing*, vol. 8, pp. 305-320, March 1999.

[81] J. Lu, A. Nosratinia, and B. Aazhang, "Progressive source-channel coding of images over bursty channels," in *Proc. ICIP'98*, Chicago, IL, October 1988.

[82] V. Chande and N. Farvardin, "Progressive transmission of images over memoryless noisy channels", *IEEE Journal on Selected Areas in Communications*, vol. 18, pp. 850-860, June 2000.

[83] R. Fletcher, *Practical Methods of Optimization*, 2nd ed., New York:Wiley, 1987.

[84] Tenkasi V. Ramabadran and Sunil S. Gaitonde, "A Tutorial on CRC computations," *IEEE Micro.*, pp. 62-75, August 1998.

[85] J. Evans and D. Tse, "Large system performance of linear multiuser receivers in multipath fading channels," *IEEE Trans. Inform. Theory*, vol. 46, pp. 2059-2078, September 2000.

[86] D. Taubman, "High performance scalable image compression with EBCOT," *IEEE Trans. Image Processing*, vol. 9, pp. 1158-1170, July 2000.

[87] M. Marcellin and A. Bilgin, "Quantifying the parent-child coding gain in zerotree-based coders," *IEEE Signal Processing Letters*, vol. 8, pp. 67-69, March 2001.

[88] I. Witten, R. Neal and J. Cleary, "Arithmetic coding for data compression," *Communications of the ACM*, vol. 30, pp. 520-540, 1987.

[89] Y. Shoham and A. Gersho, "Efficient bit allocation for an arbitrary set of quantizers," *IEEE Trans. Acoustics, Speech, and Signal Proc.*, vol. 36, pp. 1445-1453, September 1988.

[90] V. Stankovic, R. Hamzaoui, and Z. Xiong, "Packet loss protection of embed ded data with fast local search," in *Proc. ICIP'02*, Rochester, NY, September 2002.

[91] R. Hamzaoui, V. Stankovic, and Z. Xiong, "Joint product code optimization for scalable multimedia transmission over wireless channels," in *Proc. ICME'02*, Lausanne, Switzerland, August 2002.

[92] J. Hua and Z. Xiong, "Optimal rate allocation in progressive joint source-channel coding for image transmission over CDMA networks," in *IEEE International Conference on Multimedia and Expo, 2003*, pp. 661 - 664, vol. 1, July 6-9, 2003.

[93] Z. He, Y. Liang, L. Chen, I. Ahmad, and D. Wu, "Power-rate-distortion analysis for wireless video communication under energy constraint," *IEEE Transactions on Ciruits and Systems for Video Technology*, April, 2004.

[94] D. Chiu and R. Jain, "Analysis of the increase and decrease algorithm for congestion avoidance in computer networks," *Journal of Computer Networks and ISDN*, vol. 17, pp. 1-14, June 1989.

[95] R. Rejaie, M. Handley, and D. Estrin, "Layered quality adaptation for Internet video streaming," *IEEE Journal on Selected Areas in Communications*, vol. 18, pp. 2530-2544, December 2000.

[96] K.M. Mackenthun, "A fast algorithm for multiple-symbol differential detection of MPSK," *IEEE Transactions on Communications*, Vol. 42, No. 2-4, pp. 1471-1474, Feb.-Apr. 1994.

[97] J.G. Proakis, *Digital Communications*, 2nd edition, New York, McGraw-Hill, 1989.

VITA

Shengjie Zhao was born at Fenyang, Shanxi, China. His high school diploma is from Fenyang High School, Shanxi, China. He received his B.S. degree in electrical engineering from the University of Science and Technology of China (USTC), Hefei, China, and M.S. degree in electrical engineering from China Aerospace Institutes, Beijing, China, in 1988 and 1991, respectively. He was a senior research engineer before he entered Texas A&M University to pursue his Ph.D. degree in 1999. His address is Department of Electrical Engineering, #384, Texas A&M University, College Station, TX 77843-3128.

The typist for this dissertation was Shengjie Zhao.