

# Cooperative Source and Channel Coding for Wireless Multimedia Communications

Hoi Yin Shutoy, Deniz Gündüz, Elza Erkip, Yao Wang

Department of Electrical and Computer Engineering,

Polytechnic University

hoiyin@vision.poly.edu, dgunduz01@utopia.poly.edu, elza@poly.edu, yao@poly.edu

**Abstract**—Past work on cooperative communications has indicated substantial improvements in channel reliability through cooperative transmission strategies. To exploit cooperation benefits for multimedia transmission over slow fading channels, we propose to jointly allocate bits among source coding, channel coding and cooperation to minimize the expected distortion of the reconstructed signal at the receiver. Recognizing that, not all source bits are equally important in terms of the end-to-end distortion, we further propose to protect the more important bits through user cooperation. We compare four modes of transmission that differ in their compression and error protection strategies (single layer or multiple layer source coding with unequal error protection, with vs. without cooperation). Our study includes an i.i.d. Gaussian source as well as a video source employing an H.263+ codec. We present an information theoretic analysis for the Gaussian source to investigate the effects of the modulation scheme, bandwidth ratio (number of channel uses per source sample), and average link signal-to-noise ratios on the end-to-end distortion of the four modes studied. The information theoretic observations are validated using practical channel coding simulations. Our study for video considers error propagation in decoded video due to temporal prediction and jointly optimizes a source coding parameter that controls error propagation, in addition to bits for source coding, channel coding and cooperation. The results show that cooperation can significantly reduce the expected end-to-end distortion for both types of source and that layered cooperation provides further improvements and extends the benefits to a wider range of channel qualities.

**Index Terms**—Joint source channel coding, layered source coding, unequal error protection, user cooperation, wireless video.

## I. INTRODUCTION

THE advance of next generation wireless communication systems are bringing real-time video onto wireless devices. Video applications are bandwidth demanding and error sensitive, while wireless channels are unreliable and limited in bandwidth. Furthermore, the real-time nature of many video applications makes retransmission impossible. The key to improved real-time wireless video transmission therefore lies in increasing channel reliability and error resilience without sacrificing the bandwidth efficiency [1]. The multiple-input, multiple-output (MIMO) technology has brought improvements in robustness to fading by providing diversity in the spatial domain. As an alternative to multiple antenna systems when the devices are limited to a single antenna (because of

either the limitations on the device size, and/or the cost of additional antennas), user cooperation or wireless relaying is now being vigorously researched to provide spatial diversity. In cooperative communication systems, wireless transmission initiated from a source node is overheard by other terminals (called relays). Instead of discarding this overheard signal, the relays process and forward the signal to the intended destination, where different copies of the signals are combined for improved reliability. This spatial diversity scheme provides adaptation to the time-varying channel state by exploiting the network resources instead of limiting itself to the bottlenecks of the point-to-point channel.

Sendonaris et. al [2], [3] introduced the concept of cooperation among wireless terminals for spatial diversity. They showed that user cooperation is able to effectively achieve robustness against fading. Laneman et. al. [4] studied different cooperative protocols to formulate cooperative diversity gains in terms of physical layer outage probability. Hunter and Nosratinia [5], [6] proposed a cooperative scheme using rate compatible punctured convolutional (RCPC) codes and cyclic redundancy check (CRC) for error detection. Stefanov and Erkip [7] presented analytical and simulation results studying the diversity gains of cooperative coding and suggested guidelines on how good cooperative codes can be designed. The fundamental principles of cooperative transmission outlined in [2]-[7] are extended to multiple relays and/or multiple transmitter/receiver pairs in [8]-[10].

Most prior studies focus on the physical layer and use channel outage probability or the residual error probability after channel decoding as a performance measure. For multimedia signals (audio, image and video), a more relevant performance measure is the distortion of the decoded signal compared with the original source. This distortion is contributed from both the source encoder (due mainly to quantization) and residual transmission errors after channel decoding. The latter depends on channel statistics, the channel encoder and decoder, the source coder error resilience features and source decoder error concealment techniques, as well as the diversity method employed. There is a tradeoff between the source coder quality and the channel coder reliability, and the optimal allocation of the resources among these two requires a joint optimization.

In this paper, we optimize bit allocation among source coding (for both compression and error-resilience), channel coding, and cooperation, subject to a total bit rate constraint. In source coding, not all bits are equal in importance since the

rate-distortion function is in general not linear. We therefore propose to protect the more important bits through stronger channel coding and user cooperation. We call this *layered-cooperation*. To show the benefits of cooperation and unequal error protection (UEP), we compare four communication modes: direct transmission (single-layer source and channel coding without cooperation), layered transmission (layered source coding and UEP through channel coding, without cooperation), cooperative transmission (single-layer source and channel coding with cooperation), and layered cooperation (layered source coding with UEP, and cooperation among layers). These strategies provide various levels of adaptivity to the overall network state. We show that optimal allocation of the resources in case of layered cooperation results in the highest performance as multiple layers provide better adaptivity to the unknown channel coefficients at the transmitters, and it can better exploit the network resources through channel cooperation. While our results here focus on a simple network of three nodes, they can easily be extended to larger network scenarios following the results of [8]-[10].

To gain theoretical insights into the performance limits of these four modes of communications, we first examine the achievable minimal distortion for an i.i.d. Gaussian source. For this study, we make use of the well-known rate distortion function and the successive refinability of an i.i.d. Gaussian source [19] to determine the encoding-induced distortion at different source rates with or without layered source coding. In order to model residual transmission error, we perform an information theoretical analysis and use outage probability, defined as the probability of the instantaneous capacity of the fading channel being less than the desired transmission rate [20], which has been shown to be a tight lower bound to the error probability in the limit of infinite length codewords [21]. This provides us an understanding of the theoretical limits achievable by the proposed schemes without assuming any specific channel code. Furthermore, it also enables us to easily observe the effects of the modulation, various link qualities and the bandwidth ratio (number of channel uses per source sample) on the end-to-end performance.

We also study the performance achieved by practical channel codes for transmitting a Gaussian source by performing channel simulations using binary phase shift keying (BPSK) modulation and RCPC channel codes [22]. Noticeably, comparison of these simulation results with the information theoretic bounds for BPSK modulation reveals similar trends in terms of the end-to-end average distortion. In particular, our results, both theoretical and practical, show that in a modulation-constrained environment cooperation can provide significant improvements over no-cooperation when the average channel SNR is in the low to medium range, and that layered cooperation can extend this benefit to a larger range of channel qualities. When the maximum modulation level is not fixed (such as Gaussian codebooks used in the outage approach), layered cooperation always outperforms all other transmission modes, and the gap increases with channel SNR. Furthermore, benefits of cooperative transmission and layered cooperation increase when we have more channel uses per source sample, i.e., higher bandwidth ratio.

Encouraged by our results for the Gaussian source, we extend our study to include real video sequences. We use the H.263+ video codec [23], [24] with SNR scalability option for video coding, and further examine the impact of the encoder intra-block rate, a parameter that controls the encoder error resilience. Overall, we find that benefits obtained from cooperation and layered source coding for video follow similar trends exhibited by the Gaussian source. For the modulation-constrained case, the particular channel SNR range in which cooperation outperforms no-cooperation, and the range in which layered cooperation outperforms single-layer cooperation, however, do differ. These ranges also depend on the motion content of the underlying sequences. Moreover, a video source is more sensitive to the user-to-relay channel quality. A high SNR user-to-relay channel benefits video more than Gaussian source.

The literature on joint source-channel coding for cooperative systems is sparse. In [11] -[16], we introduced the problem of source-channel coding in cooperative relay systems. We considered single and two-layer source coders and through simulation studies, compared achievable minimal distortions by different communication modes, both for the i.i.d. Gaussian source [11], [12], [15] and a real video source [16]. In [13], [14] we studied the high SNR behavior of end-to-end distortion, and provided cooperative source and channel coding strategies that use layered compression with progressive or simultaneous transmission. Results in [13], [14] are significant in the sense that, in the high SNR regime, our cooperative joint source-channel coding schemes provide optimal decay rate of end-to-end distortion (known as distortion exponent [17]) with increasing SNR, for a wide range of bandwidth ratios. The authors in [18] combined multiple description source coding with user cooperation, and studied achievable performances when a Gaussian source is coded into either one or two descriptions, and sent with or without cooperation.

The remaining of the paper is organized as follows: In Section II, we describe the channel model, formulate the four modes of communication and introduce the optimal bit allocation problem. In Section III, we consider transmission of an i.i.d. Gaussian source, and analyze both the information theoretic limits and the practical channel coding performance for various link qualities. In Section IV, we analyze the simulation results for various video sequences using H.263+ codec. Section V concludes this study.

## II. SYSTEM SETUP AND FOUR MODES OF COOPERATION

We consider two terminals,  $T_1$  and  $T_2$ , in the same wireless network. We assume  $T_1$  sends its source,  $S_1$ , to its destination  $D_1$  and  $T_2$  is  $T_1$ 's cooperative relay. As  $T_1$  transmits towards  $D_1$ ,  $T_2$  overhears  $T_1$ 's signal and relays it to  $D_1$ . This setup can be applied to several scenarios, provided that  $T_2$  can overhear  $T_1$ 's transmission; and  $D_1$  can receive and combine the two signals. For example,  $T_1$  could be a 3G mobile phone sending real-time video to its base station,  $D_1$ ; and  $T_2$  is another mobile device in the network. Or,  $D_1$  could be a wireless LAN client receiving streaming video from a server,

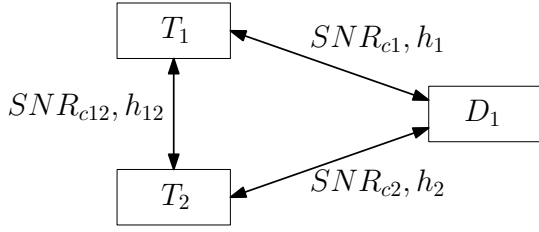


Fig. 1. Cooperative system model. Here  $T_1$  wishes to communicate with  $D_1$  with the help of the relay  $T_2$ . In the figure,  $SNR_{c1}$ ,  $SNR_{c2}$  and  $SNR_{c12}$  denote average received SNRs;  $h_1$ ,  $h_2$  and  $h_{12}$  denote instantaneous fading amplitudes.

then  $T_1$  would be the access point and  $T_2$ , another WLAN client who streams video from  $T_1$  either for his own usage or to help  $D_1$  receive a better quality video. As customary in the literature [4] and to capture the benefits of cooperation, we assume time division multiple access among  $T_1$  and  $T_2$ . While cooperating,  $T_1$  only transmits during a certain part of its time slot (or channel frame); the remaining part of the time slot for  $T_1$  is used by  $T_2$  to send  $T_1$ 's information. Similar to [5], [7], if  $T_2$  cannot decode  $T_1$ 's signal (which we can check using an error detection mechanism such as CRC),  $T_1$  uses the remaining time to continue transmission to  $D_1$ . Note that, this requires a single bit feedback from  $T_2$  to both  $T_1$  and  $D_1$ , which can be supported by a low rate high reliability feedback link. This feedback rate is negligible compared with the data rate used in communication. Furthermore, CRC is already present in communication systems and the overhead due to CRC is small. In general,  $T_2$  may also be sending its own source to either the same or a different destination in the timeslot belonging to  $T_2$ ; and  $T_1$  may or may not help in  $T_2$ 's transmission, but in this paper, we are only interested in investigating the performance from  $T_1$ 's perspective and hence, we concentrate only on the time slot allocated for  $T_1$ 's transmission. Also we do not make use of the correlation between the sources of  $T_1$  and  $T_2$ , even when such correlation exists.

We assume flat, slow Rayleigh fading, with independent fading amplitudes  $h_1$ ,  $h_2$  and  $h_{12}$  for the three wireless links as shown in Figure 1. We define one channel frame as a block of  $N$  channel uses and assume that the channel coherence time (over which the fading amplitude is constant) spans an integer number of channel frames. The fading amplitudes vary independently from one coherence time to the next. The fading amplitude squares,  $|h_1|^2$ ,  $|h_2|^2$  and  $|h_{12}|^2$ , each follow an exponential distribution with unit mean. The additive noise is i.i.d. white Gaussian and independent for each receiver. We use  $SNR_{ci}$ ,  $i=1,2$ , to denote the average received signal to noise ratio at the destination, for channels 1 ( $T_1$  to  $D_1$ ) and 2 ( $T_2$  to  $D_1$ ), respectively. We will call them user channel SNRs. The average received SNR for the channel between  $T_1$  and  $T_2$ , the inter-user channel, is  $SNR_{c12}$ . We assume that average received SNRs, which model the pathloss and/or shadowing in the environment, are known by all terminals, but instantaneous channel realizations can only be measured at the corresponding receivers, which are then used for maximum likelihood channel decoding. In our analysis and simulations,

we consider both an asymmetric scenario, where  $SNR_{c2} = SNR_{c12} = 4SNR_{c1}$ , i.e.,  $T_2$  has better channels to both  $T_1$  and  $D_1$  than the  $T_1 - D_1$  direct link on the average, and a symmetric scenario, where  $SNR_{c1} = SNR_{c2}$ , i.e., both terminals have the same average quality channels to the destination. Asymmetric scenario could happen, for example, if  $T_2$  is closer to  $D_1$  than  $T_1$  or  $T_1 - D_1$  link is severely shadowed; symmetric scenario takes place when  $T_1$  and  $T_2$  are close to each other.

In order to describe the four modes of transmission, we consider a system with a fixed  $M$ -ary modulation with which a total of  $R_t$  bits can be transmitted over the channel in one channel frame, that is,  $R_t = N \log_2 M$ . We assume we have  $K$  source samples to be transmitted in one channel frame, leading to a *bandwidth ratio*,  $b = N/K$ . The  $R_t$  bits in a channel frame are allocated among source coding (number of compressed bits representing source samples), channel coding (channel parity bits sent by the user), and cooperation (channel parity bits sent by the relay). Modes 1 and 2 below are well-known communication schemes neither of which utilizes user cooperation. Mode 2 differs from mode 1 in that it uses the popular UEP method to increase source error resilience. Mode 3 uses cooperative coding [5], [7], while mode 4 further applies cooperation to implement UEP. These communication modes are detailed in Figure 2.

**Mode 1** (direct transmission): Terminal  $T_1$  compresses  $S_1$  at  $R_b$  source bits/frame and applies a rate  $R_b/(R_b + r_b)$  channel code, where  $r_b$  is the number of parity bits/channel frame. It transmits the resulting  $R_b + r_b$  bits directly to destination  $D_1$ . The constraint is  $R_b + r_b \leq R_t$ .

**Mode 2** (layered source coding and transmission): This is the conventional layered source coding with UEP through channel coding.  $S_1$  is compressed into two layers: a base-layer (BL) with  $R_b$  source bits and an enhancement-layer (EL) with  $R_e$  source bits in each channel frame.  $T_1$  applies a channel code of rate  $R_b/(R_b + r_b)$  for BL and one of rate  $R_e/(R_e + r_e)$  for EL. All  $R_b + r_b + R_e + r_e$  bits are sent directly to  $D_1$ . The constraint is  $R_b + r_b + R_e + r_e \leq R_t$ . When  $R_e = 0$  and  $r_e = 0$ , mode 2 operates as mode 1.

**Mode 3** (cooperative channel coding): Terminal  $T_1$  compresses  $S_1$  into a single layer using  $R_b$  bits, and then implements a rate  $R_b/(R_b + r_{b1} + r_{b2})$  channel code that can be punctured to rate  $R_b/(R_b + r_{b1})$ .  $R_b$  source bits and  $r_{b1}$  parity bits are then transmitted directly to  $D_1$ . As  $T_2$  overhears and decodes  $T_1$ 's transmission, it generates and sends the remaining  $r_{b2}$  channel parity bits (the "cooperation bits") for  $T_1$ . If  $T_2$  cannot decode the  $S_1$  source bits, it sends 1-bit "frame error" signal to  $T_1$  and  $D_1$ .  $T_1$  will then continue to send the  $r_{b2}$  parity bits itself. The bit allocation constraint is  $R_b + r_{b1} + r_{b2} \leq R_t$ . Notice that when  $r_{b2} = 0$ , mode 3 operates as mode 1.

**Mode 4** (layered cooperation):  $S_1$  is compressed with  $R_b$  bits for the BL and  $R_e$  bits for the EL. We apply cooperation only for BL. BL is protected by  $r_{b1}$  parity bits sent by  $T_1$  and  $r_{b2}$  cooperative parity bits sent by  $T_2$  if  $T_2$  can decode the base layer after it observes  $R_b + r_{b1}$  channel coded bits. Otherwise, the  $r_{b2}$  parity bits are sent by  $T_1$ . EL is protected by  $r_e$  parity bits sent by  $T_1$ . The bit allocation constraint is

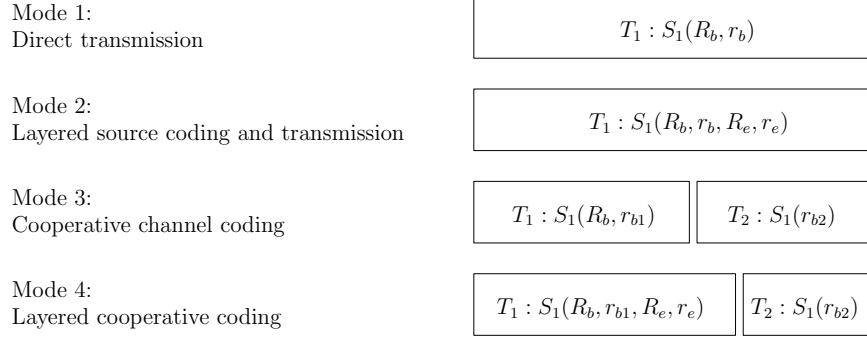


Fig. 2. Four modes of communication described in Section II.

$R_b + r_{b1} + r_{b2} + R_e + r_e \leq R_t$ . Mode 3 is a special case of mode 4 when  $R_e = r_e = 0$ . So, mode 4 includes both mode 3 and mode 1. Mode 2 is also a special case of mode 4 when  $r_{b2} = 0$ . Hence mode 4 is the most general mode under consideration.

For each of these four modes, there is an optimal bit allocation, subject to a constraint on the total number of bits/channel frame  $R_t$ , among source compression (BL and EL), channel coding and cooperation that will lead to the lowest end-to-end source distortion. This optimal allocation depends on the type of the source, the modulation, the frame size, the bandwidth ratio and the source and channel coding strategies as well as the average channel SNRs. In the next few sections, we will study this optimal bit allocation problem and compare the resulting end-to-end distortion for different sources and channel qualities as well as modulation modes and bandwidth ratios.

### III. COOPERATION FOR AN I.I.D. GAUSSIAN SOURCE

To gain insights into the achievable performance of the various communication modes in Sec. II, we first evaluate their performances with an i.i.d. real Gaussian source with unit variance. For a bandwidth ratio of  $b$ , if each source sample is compressed to  $\tilde{R}$  bits,  $\tilde{R}$  is constrained to be  $\tilde{R} \leq R_t/K = b \log_2 M$ . Using the distortion-rate function of a zero mean, unit variance i.i.d. real Gaussian source [25], the distortion per sample in terms of mean squared error (MSE) is  $D(\tilde{R}) = 2^{-2\tilde{R}}$ .

Note that, this distortion-rate function is an upper bound for the distortion achievable for any unit variance memoryless source at the same compression rate and is asymptotically tight as  $\tilde{R}$  goes to infinity. For sources with memory, including video, the precise form of the preceding function is not applicable, but the exponential decay of distortion with rate is generally still true [26], [27].

#### A. Formulation of the Expected Distortion and the Optimal Bit Allocation Problem

For a given communication mode, modulation level, bandwidth ratio and average channel SNRs, the expected distortion (ED) is a function of source rate ( $R_b$  and  $R_e$ ), the amount of channel coding ( $r_{b1}$  and  $r_e$ ) and the level of cooperation ( $r_{b2}$ ).

Below we will discuss how to compute ED for mode 4. The formulation of ED for other modes can be easily derived since they are special cases of mode 4. When the channel decoder cannot correct all the bit errors that occur in a channel frame for either BL or EL, we assume all source bits corresponding to that layer are dropped. For a given channel frame, there are three mutually exclusive events after channel decoding:

- Both BL and EL are correctly decoded.
- Only BL bits are decoded correctly.
- BL is not decodable. Since without the BL, the EL cannot be used, we simply replace all the samples by their mean values.

We denote the average probabilities of cases (a), (b) and (c) (averaged over all fading realizations) as  $P_2$ ,  $P_1$  and  $P_0$  respectively. Due to the successive refinability of the Gaussian source [19], for case (a), the distortion per sample is  $D(\tilde{R}_b + \tilde{R}_e)$ , where  $\tilde{R}_b = R_b/K$  and  $\tilde{R}_e = R_e/K$ . For case (b), it is  $D(\tilde{R}_b)$  and for case (c), it is  $D(0) = 1$ <sup>1</sup> due to unit variance assumption. The expected distortion (ED) for a given channel frame is:

$$ED(R_b, R_e, r_{b1}, r_{b2}, r_e) = P_2 D(\tilde{R}_b + \tilde{R}_e) + P_1 D(\tilde{R}_b) + P_0 D(0). \quad (1)$$

Since the source is memoryless and we assume that source coding and decoding in each channel frame only depends on samples in the current frame, there is no decoding error propagation from one channel frame to the next in the decoded samples. Note that, as discussed in detail in Section IV-A, because of error propagation, (1) does not hold for video sources.

Before we proceed to discuss the formulation of probabilities  $P_2$ ,  $P_1$  and  $P_0$ , we first define the following that describe the residual frame error rates (FER) after channel decoding.

- $P_{in}$  is the average FER of the inter-user channel, averaged over different fading levels  $h_{12}$ . It can be expressed as:

$$P_{in} = E_{h_{12}}[P(R_b, r_{b1}; SNR_{c12}|h_{12})],$$

where  $P(R_b, r_{b1}; SNR_{c12}|h_{12})$  is the probability of incorrectly decoding a channel frame with  $R_b$  source bits and  $r_{b1}$  parity bits when it is sent over a channel with average SNR,  $SNR_{c12}$ , for an instantaneous fading

<sup>1</sup>Although it is possible to perform error concealment of samples in the current frame from reconstructed samples in previous frames for general sources, we do not consider this here since the source is memoryless.

amplitude  $h_{12}$ . Generally,  $P_{in}$  depends on other channel coding parameters such as the channel code used, the modulation level and packet size. Here we assume these parameters are fixed and we do not explicitly show that dependence.

- $P_{b|h_1}$  is the FER of the BL frame at  $D_1$  for a particular fading level  $h_1$  given that the relay  $T_2$  cannot decode the  $R_b$  source bits and thus cooperation does not take place. When BL is transmitted directly by  $T_1$ , the  $R_b$  source bits and  $r_{b1} + r_{b2}$  parity bits in a channel frame are all subject to the fading  $h_1$  and the average channel SNR,  $SNR_{c1}$ .  $P_{b|h_1}$  can be written as:

$$P_{b|h_1} = P(R_b, r_{b1} + r_{b2}; SNR_{c1}|h_1).$$

- $P_{e|h_1}$  is the instantaneous FER of the EL frame at  $D_1$ , which is transmitted entirely by  $T_1$ . The  $R_e$  source bits and  $r_e$  parity bits are sent over a channel with fading amplitude,  $h_1$  and average SNR,  $SNR_{c1}$ :

$$P_{e|h_1} = P(R_e, r_e; SNR_{c1}|h_1).$$

- $P_{b|h_1, h_2}$  is the FER of BL when cooperation takes place. It is the probability of decoding BL incorrectly when  $R_b$  source bits and  $r_{b1}$  parity bits are sent by  $T_1$ ; while  $r_{b2}$  parity bits are sent by  $T_2$  after decoding the  $R_b$  source bits. Thus  $R_b + r_{b1}$  bits are sent under a channel realization of  $h_1$  with average SNR,  $SNR_{c1}$  while  $r_{b2}$  bits are sent under  $h_2$  with average SNR,  $SNR_{c2}$ .

$$P_{b|h_1, h_2} = P(R_b, r_{b1}, r_{b2}; SNR_{c1}, SNR_{c2}|h_1, h_2),$$

For each of the cases (a), (b) and (c), there are two mutually exclusive situations depending on whether  $S_1$  is transmitted with or without cooperation. We define the probability of the six resulting outcomes as follows:

Pr(receive BL and EL, with coop.):

$$P_{2,coop} = (1 - P_{in})E_{h_1, h_2}[(1 - P_{b|h_1, h_2})(1 - P_{e|h_1})], \quad (2)$$

Pr(receive BL and EL, without coop.):

$$P_{2,nocoop} = P_{in}E_{h_1}[(1 - P_{b|h_1})(1 - P_{e|h_1})], \quad (3)$$

Pr(receive BL only, with coop.):

$$P_{1,coop} = (1 - P_{in})E_{h_1, h_2}[(1 - P_{b|h_1, h_2})P_{e|h_1}], \quad (4)$$

Pr(receive BL only, without coop.):

$$P_{1,nocoop} = P_{in}E_{h_1}[(1 - P_{b|h_1})P_{e|h_1}], \quad (5)$$

Pr(do not receive BL, with coop.):

$$P_{0,coop} = (1 - P_{in})E_{h_1, h_2}[P_{b|h_1, h_2}], \quad (6)$$

Pr(do not receive BL, without coop.):

$$P_{0,nocoop} = P_{in}E_{h_1}[P_{b|h_1}], \quad (7)$$

The probabilities  $P_i$ , for  $i = 0, 1, 2$ , in (1) can therefore be expressed in terms of (2)-(7) as:

$$P_i = P_{i,coop} + P_{i,nocoop}. \quad (8)$$

For a given bandwidth ratio, modulation level, channel code and packet size, the goal is to find the optimal bit

allocation  $(R_b, r_{b1}, r_{b2}, R_e, r_e)$  that results in the minimum ED for each  $SNR_{c1}$ ,  $SNR_{c2}$  and  $SNR_{c12}$  under consideration. We also wish to study conditions under which each of the four transmission modes becomes dominant and investigate the choice of system parameters on the end-to-end distortion.

### B. Information Theoretic Formulation of Expected Distortion

In this section we introduce an information theoretic framework to solve the expected distortion optimization problem described above. In order to understand the potentials of the suggested four modes of communication, we will consider *outage probability* which serves as a tight bound for the frame error rate of any practical channel coding scheme in the limit of infinitely long codewords [21]. The outage probability is defined as the probability that the instantaneous channel capacity is lower than the desired transmission rate. Intuitively, the best possible channel code (in Shannon sense) only results in errors when we transmit at a rate higher than the channel capacity, hence errors in a fading environment will only occur if the desired rate is higher than the instantaneous capacity.

For an  $M$ -ary modulation scheme with soft decision decoding, we can model a slow fading channel of the form described in Section II as an  $M$ -input, real-output channel. We denote the instantaneous capacity of this channel at fading level  $h$ , and average signal-to-noise ratio SNR as  $C_M(|h|^2 SNR)$ . When the modulation scheme is not constrained and we are allowed to use real valued inputs (as in Gaussian codebooks), we have  $C_G(|h|^2 SNR) = \log(1 + |h|^2 SNR)$ . Then the outage probability  $P^{out}$  for rate  $R$  becomes  $P^{out} = \Pr\{C_M(|h|^2 SNR) < R\}$ . Note that  $R$  represents the desired number of information bits (or source bits) per channel use that is to be transmitted over the fading channel. For example, for our proposed mode 1, we have

$$R = \frac{R_b}{R_b + r_b} \log_2 M = \frac{R_b}{N}. \quad (9)$$

The corresponding source coding rate in terms of bits per source sample becomes  $\hat{R}_b = R_b/K = bR$ , where  $b$  is the bandwidth ratio.

We now rewrite the various frame error probabilities and the expected end-to-end distortion expression in Section III-A for mode 4, in terms of the outage probabilities. For the base layer, we transmit a total of  $R_b + r_{b1} + r_{b2}$  bits through the channel. For  $M$ -ary modulation, this results in  $(R_b + r_{b1} + r_{b2})/\log_2 M$  channel uses. Therefore, the base layer utilizes  $\alpha$  proportion of the channel uses, where

$$\alpha = \frac{R_b + r_{b1} + r_{b2}}{N \log_2 M} = \frac{R_b + r_{b1} + r_{b2}}{R_t}.$$

Of these  $\alpha N$  channel uses, the original user gets  $\beta$  proportion, where

$$\beta = \frac{R_b + r_{b1}}{R_b + r_{b1} + r_{b2}},$$

and the relay uses the remaining  $(1 - \beta)$  proportion. Similarly, the enhancement layer utilizes  $1 - \alpha = (R_e + r_e)/R_t$  proportion of the channel uses.

The base layer is transmitted at  $R_1 = R_b \log_2 M / (R_b + r_{b1} + r_{b2})$  information bits per channel use and the enhancement layer at  $R_2 = R_e \log_2 M / (R_e + r_e)$  information bits

per channel use towards the destination. The transmission rate of the base layer from  $T_1$  to  $T_2$  becomes  $R_1/\beta = R_b \log_2 M / (R_b + r_{b1})$  information bits per channel use. The corresponding source compression rates for the base and the enhancement layer are  $\tilde{R}_b = b\alpha R_1$  and  $\tilde{R}_e = b(1 - \alpha)R_2$  bits per source sample.

Combining, the outage probabilities corresponding to the FERs in (2)-(7) are:

$$P_{2,coop}^{out} = (1 - P_{in}^{out}) \cdot Pr\{R_1 < \beta C_{M1} + (1 - \beta)C_{M2}, R_2 < C_{M1}\}, \quad (10)$$

$$P_{2,nocoop}^{out} = P_{in}^{out} Pr\{R_1 < C_{M1}, R_2 < C_{M1}\}, \quad (11)$$

$$P_{1,coop}^{out} = (1 - P_{in}^{out}) \cdot Pr\{R_1 < \beta C_{M1} + (1 - \beta)C_{M2}, R_2 \geq C_{M1}\}, \quad (12)$$

$$P_{1,nocoop}^{out} = P_{in}^{out} Pr\{R_1 < C_{M1}, R_2 \geq C_{M1}\}, \quad (13)$$

$$P_{0,coop}^{out} = (1 - P_{in}^{out}) Pr\{R_1 \geq \beta C_{M1} + (1 - \beta)C_{M2}\}, \quad (14)$$

$$P_{0,nocoop}^{out} = P_{in}^{out} Pr\{R_1 \geq C_{M1}\}, \quad (15)$$

with  $P_{in}^{out} = Pr\{\frac{R_1}{\beta} > C_{M12}\}$ . In the above expressions,  $C_{M1}$ ,  $C_{M2}$ , and  $C_{M12}$  correspond to, respectively, the capacities of  $T_1 - D_1$ ,  $T_2 - D_1$ , and  $T_1 - T_2$  links and depend on  $h_i$  and  $SNR_{ci}$  for  $i = 1, 2$  or  $\{12\}$ . Also note that, to model the additional parity bits transmitted by  $T_2$  when cooperation takes place (as in  $P_{2,coop}^{out}$ ), we use independent codebooks at  $T_2$  for retransmission of the decoded information [6]. Hence, the total mutual information is the sum of the mutual information terms from the source and the relay to the destination. This is higher than the rate achievable by repetition based-scheme, which use the same codebook at both the source and the relay ([4]).

The overall outage based end-to-end distortion can be found using the outage probability expressions in (10)-(15) as

$$ED(R_1, R_2, \alpha, \beta) = P_2^{out} D(b\alpha R_1 + b(1 - \alpha)R_2) + P_1^{out} D(b\alpha R_1) + P_0^{out} \quad (16)$$

with  $P_i^{out} = P_{i,coop}^{out} + P_{i,nocoop}^{out}$  for  $i = 0, 1, 2$ .

For fixed  $M, b$ , and link SNRs, the end-to-end distortion can now be minimized over all choices of  $(\alpha, \beta, R_1, R_2)$  with  $0 \leq \alpha \leq 1$ ,  $0 \leq \beta \leq 1$ . Note that  $R_1$  and  $R_2$  are both upper bounded by  $\log_2 M$  and the assumption of large channel frame lengths used in the information theoretic approach enables us to use any real valued  $(\alpha, \beta, R_1, R_2)$ . As special cases of mode 4, if  $\alpha = 1, \beta = 1, R_2 = 0$  we would reduce to mode 1, if  $\beta = 1$  we would reduce to mode 2, and if  $\alpha = 1, \beta = 0, R_2 = 0$  we would reduce to mode 3.

It is possible to use the exact expressions for the outage probabilities and the distortion-rate functions in (16), and obtain an expected distortion expression in terms of the rates  $R_1, R_2$  and the channel allocation variables  $\alpha$  and  $\beta$ . The optimal allocation of these variables is a non-linear optimization problem whose efficient solution is not obvious. Here we will use exhaustive search over a discretized search space to obtain numerical results.

### C. Results for the Information Theoretic Analysis

In this subsection, we present our information theoretical results, which provide an expected distortion lower bound for any practical coding scheme that can be used. For easier comparison with video simulation results presented in Sec. IV, instead of ED we use the decoded signal SNR,  $SNR_g$ , defined below, as our performance measure

$$SNR_g = 10 \log_{10} \frac{1}{ED}.$$

In general, when there are no constraints on the maximum modulation level, the optimal transmission scheme that achieves the Gaussian channel capacity  $C_G(|h|^2 SNR) = \log(1 + |h|^2 SNR)$  requires Gaussian codebooks. We consider two bandwidth ratios  $b = 1$  and  $b = 4$ , and study an asymmetric scenario where  $SNR_{c2} = SNR_{c12} = 4SNR_{c1}$  with Gaussian codebooks. The optimal end-to-end signal SNR ( $SNR_g$ ) as a function of average channel SNR,  $SNR_{c1}$  can be seen in Figure 3 for the four modes of transmission. The four topmost curves in the figure correspond to  $b = 4$  case while the lowest four curves are for  $b = 1$ . We observe a significant improvement in the reconstructed signal SNR, with cooperation (mode 4 and mode 3). While source layering (mode 2) improves the end-to-end distortion compared to direct transmission (mode 1), cooperation results in a larger gain with layered cooperation (mode 4) resulting in the best signal quality. Also note that the difference between the modes increase as channel SNR increases. We also see in Figure 3 that the amount of improvement that modes 2, 3, and 4 bring over mode 1 increases with the bandwidth ratio, which yields more channel uses per sample. In fact, the rate of increase of  $SNR_g$  with channel SNR, gets larger with the bandwidth ratio and mode 4 consistently results in the largest rate of increase. This is consistent with our analysis of the distortion exponent of layered cooperation in [13], [14] where we characterized the rate of decrease of expected distortion with SNR at high channel SNR range. The distortion exponent,  $\Delta$ , is defined as the asymptotic decay rate of the average end-to-end distortion with increasing channel signal-to-noise ratio (SNR), i.e.,  $\Delta \triangleq -\lim_{SNR \rightarrow \infty} \log(\text{Expected Distortion}) / \log SNR$ . In [13], [14] we provide the distortion exponent of a cooperative system for a given number of source coding layers and bandwidth ratio, for various transmission strategies such as progressive, simultaneous and hybrid digital-analog schemes. While [13], [14] only provides distortion exponent results based on an high SNR analysis, our results here confirm their validity for a wide range of channel SNR values.

In order to study the effect of finite order modulation, we also consider BPSK. While it is not possible to obtain an explicit expression for  $C_2(|h|^2 SNR)$ , the capacity of BPSK signaling over AWGN channel, it can be easily computed numerically. Note that, when we constrain the transmission to BPSK, the channel capacity is limited by 1 bits per channel use, no matter how high the channel SNR is. This, in the end, results in a non-zero lower bound for the minimum achievable expected distortion (ED) as opposed to the Gaussian codebook case, where ED approaches to zero (or equivalently  $SNR_g$  to infinity) with increasing channel SNR.

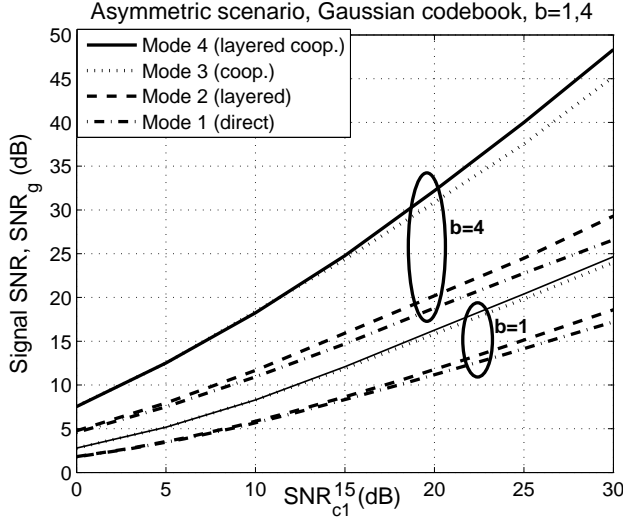


Fig. 3. Signal SNR,  $SNR_g$  (dB) vs. average channel SNR for  $T_1 - D_1$  link,  $SNR_{c1}$  (dB) for different modes of communication with no constraints on modulation. Results are based on information theoretic analysis. The topmost 4 curves are for a bandwidth ratio of  $b = 4$ , and the 4 curves below are for a bandwidth ratio of  $b = 1$ . We assume an asymmetric scenario with  $SNR_{c2} = SNR_{c12} = 4SNR_{c1}$ .

For BPSK modulation we first consider the asymmetric scenario with  $SNR_{c2} = SNR_{c12} = 4SNR_{c1}$  and  $b = 4$ . Figure 4 compares the reconstructed signal quality (in terms of  $SNR_g$ ) of four modes for the asymmetric scenario. We vary  $SNR_{c1}$  (hence  $SNR_{c2}$  and  $SNR_{c12}$ ) to analyze the effect of channel qualities on cooperation. Since the channel capacity is bounded above by 1 even when SNR is high, all four modes converge to a similar  $SNR_g$  for large channel SNRs. However, for smaller SNRs, gains due to cooperation are significant. Cooperative coding (mode 3) improves the end-to-end distortion over modes 1 and 2 up to  $SNR_{c1} = 17$  dB while layered cooperation (mode 4) extends the benefits of cooperative transmission to all SNRs. Comparison with Figure 3 illustrates that fixing the modulation scheme limits the potential improvements one can get with layered cooperation. Therefore, adaptive modulation (based on average channel SNRs) coupled with layered cooperation would be essential for better reconstructed signal quality.

To study the effects of various link qualities, we also consider the symmetric case for which  $SNR_{c1} = SNR_{c2}$  in Figure 5. To understand the effect of inter-user channel between  $T_1$  and  $T_2$ , we consider  $SNR_{c12} = \infty$  in Figure 5(a) (perfect inter-user channel) and  $SNR_{c12} = 14$  dB in Figure 5(b). We chose  $SNR_{c12}$  as an intermediate SNR value. Note that, we examine  $SNR_{c1} = SNR_{c2}$  between 0 – 30 dB. Hence, for  $SNR_{c1} < 14$  dB  $T_1$  and  $T_2$  have worse quality channels to  $D_1$  than the  $T_1 - T_2$  channel, while for  $SNR_{c1} > 14$  dB,  $T_1$  observes a better channel to  $D_1$  than to  $T_2$ . Note that, performances of mode 2 and mode 1 are not affected by the inter-user channel. Comparing the two plots in Figure 5, we observe that both mode 3 and mode 4, perform very close to the perfect inter-user channel case even when the inter-user channel has a fixed average SNR of 14 dB. Note that in Figure 5(b) cooperative coding (mode

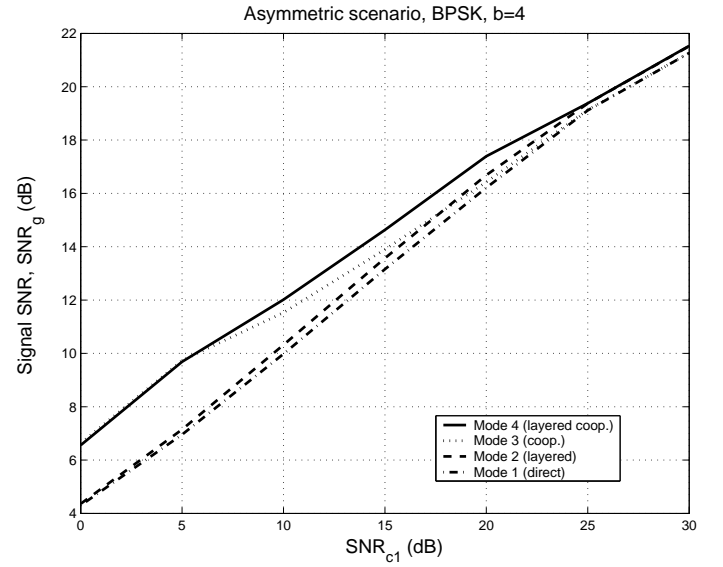


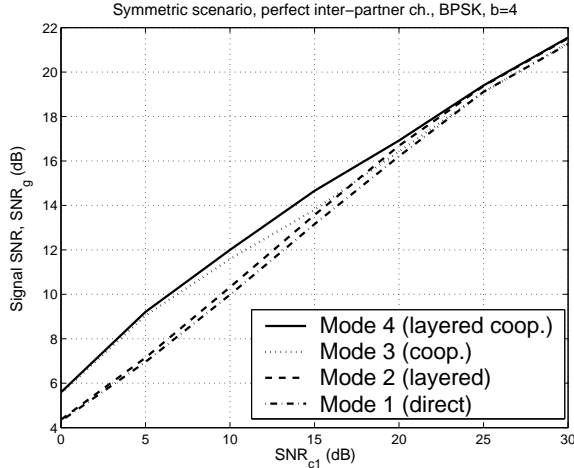
Fig. 4. Signal SNR,  $SNR_g$  (dB) vs. average channel SNR,  $SNR_{c1}$  (dB) for different modes of communication. We assume a bandwidth ratio of  $b = 4$  with BPSK modulation, and an asymmetric scenario with  $SNR_{c2} = SNR_{c12} = 4SNR_{c1}$ . Results are based on information theoretic analysis.

3) performs better than direct transmission (mode 1) beyond  $SNR_{c1} = 14$  dB, that is even when  $T_1 - T_2$  channel has worse quality than  $T_1 - D_1$  channel. This is due to the increased diversity of cooperative transmission [4]. Also, while mode 3 and 4 reconstructed signal qualities are slightly worse than the asymmetric case, we still get significant improvements over no cooperation (modes 1 and 2). Outage or error probability of cooperation is usually smaller for asymmetric scenarios such as the one studied here, mainly due to improvements in pathloss, which is consistent with our distortion analysis.

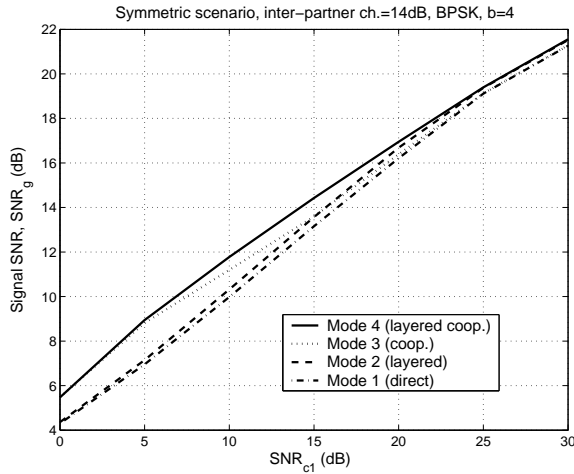
#### D. Simulation Results for Practical Channel Codes

In order to study the performance achievable by practical channel coding and modulation schemes, we use BPSK modulation and RCPC convolutional codes [22]. A rate 1/4 mother code with (17, 13, 13, 15) generator is employed. The code can then be punctured to rates 1, 2/3, 1/2 and 1/3. At the receiver, a Viterbi decoder is used for decoding. We fix our channel frame size,  $R_t$  to be 1728 bits and the number of samples per frame,  $K$  to be 432, resulting in  $\tilde{R}_t = 4$  bits/sample. Since  $M = 2$  this results in a bandwidth ratio of  $b = 4$ . The optimal bit allocation problem explained in Section II is solved through exhaustive search, that is, we obtain the various probabilities in (2)-(7) through channel simulations for every possible bit allocation satisfying the rate constraint and find the optimal bit allocation that minimizes ED.

Similar to the information theoretic analysis, we examine the decoded signal SNR both for an asymmetric and a symmetric cooperation scenario ( $SNR_{c2} = SNR_{c12} = 4SNR_{c1}$  and  $SNR_{c2} = SNR_{c1}$  with perfect inter-user channel, respectively). The results for both scenarios are plotted in Figure 6, where the reconstructed signal quality (in terms of SNRg) obtained with optimal bit allocation using the four communications modes is plotted as a function of increasing



(a) Perfect inter-user channel.



(b) 14 dB inter-user channel.

Fig. 5. Signal SNR,  $SNR_g$  (dB) vs. average channel SNR (dB) for different modes of communication based on information theoretic analysis. We assume a bandwidth ratio of  $b = 4$  with BPSK modulation, and a symmetric scenario with  $SNR_{c2} = SNR_{c1}$ .

$SNR_{c1}$ . The trends are similar to the information theoretic bounds in Figure 4 and 5, and  $SNR_g$  is slightly lower than the bounds as expected. By examining the optimal bit allocation resulting from our study, we find that, at low  $SNR_{c1}$ , the need for more protection through channel coding leads to no bit assignment for EL. So mode 4 operates as mode 3 (and mode 2 as 1), which explains the overlap of mode 4 and mode 3 (mode 2 and 1) at the low  $SNR_{c1}$  region in Figure 4. At high  $SNR_{c1}$ , channel coding bits are no longer necessary, hence the optimal allocation assigns all bits to source compression to reduce compression distortion. Therefore, mode 4 and mode 2 (mode 3 and 1) overlap at high  $SNR_{c1}$  in Figure 4. Layered transmission modes are equivalent, or better than single layer modes (mode 2 is better than 1, and mode 4 is better than 3). We have also carried out simulations for  $SNR_{c12} = 14$  dB,

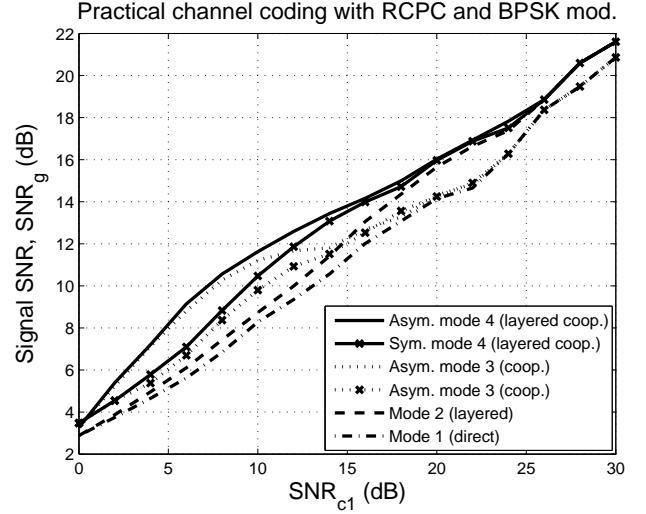


Fig. 6. Signal SNR,  $SNR_g$  (dB) vs. average channel SNR (dB) of practical channel coding scheme (RCPC) for both the asymmetric and the symmetric scenarios, and  $b = 4$ . The symmetric scenario assumes the inter user channel is perfect.

and found out that the optimal bit allocations are identical to perfect inter-user channel case.

Our results for the BPSK modulation scheme (both information theoretic and the practical coding results) show that i) layered compression and transmission (mode 2) brings improvement over direct transmission (mode 1) at high  $SNR_{c1}$  range; ii) Cooperative channel coding (mode 3) improves performance over mode 1 and mode 2 up to medium values of  $SNR_{c1}$ , but mode 3 is not as good as mode 2 at high  $SNR_{c1}$ ; iii) layered cooperation (mode 4), by combining the benefits of mode 2 and mode 3, provides the best performance over all other modes for the entire range of  $SNR_{c1}$  tested; iv) lastly, there is a distinct improvement from mode 4 over all other modes in a mid-range of  $SNR_{c1}$ .

#### IV. COOPERATION FOR VIDEO TRANSMISSION

##### A. The Video Source and Encoder

In the previous section, we showed the fundamental benefits of layered cooperation for i.i.d Gaussian sources, and argued that end-to-end expected distortion obtained with practical codes closely resembles the information theoretic bounds for fixed modulation. Motivated by this, in this section we perform experiments with real video signals and practical channel codes. Note that although the distortion-rate function of a video encoder generally still follows the exponential decay behavior [26], [27], the impact of packet losses on the end-to-end distortion is far more complicated. This is because errors injected at different locations of a video stream contribute differently to the final distortion. Hence the expected distortion expression for Gaussian sources in (1) will not hold for video signals. Although there has been prior research (e.g. [28], [29]) on modeling video distortion from channel packet loss, these models are developed under various assumptions about the video signal characteristics and the channel error statistics. Instead of using analytical models, we choose to code real video sequences and simulate their transmissions.



For video encoding, we use an H.263+ codec [24] with SNR scalability. The H.263+ encoder uses the popular block-based motion compensated prediction (MC). Each macroblock (MB) in the current video frame is expressed in terms of the best matching MB in the past frame through a motion vector (MV), which describes the block translation. In place of the color intensities, only the MVs and the prediction error are transmitted. MC improves coding efficiency by exploiting dependency between adjacent video frames, but it makes the compressed stream very sensitive to channel errors. The use of previously reconstructed video frames at the decoder forms a prediction loop and introduces temporal error propagation. One way to circumvent this problem is to use periodic intra refresh, where an MB is coded periodically without MC using an intra mode. This allows the MB to be decoded independently. Error propagation thus stops at the next received intra MB.

In our experiment, we generate the compressed video bit streams using the H.263+ encoder software [30] and simulate its transmission over wireless medium for all of the four communication modes. We intra-code the first video frame in the test sequence, called I-frame. All the other video frames, called P-frames, are compressed using MC from one previously reconstructed video frame. For each P-frame, we intra-code every  $J$ -th MB, while all other MBs are coded in an inter mode using MC. We define the intra rate as  $\beta = 1/J$ . Note that hereafter, we use just “frame” to refer to a “video frame” while we will use “channel frame” to specifically refer to a channel frame.

Layered or scalable compression is achieved by partitioning video data into different layers. BL contains the essential information, while ELs contain additional information for refinement. Different ways of partitioning the video data constitutes the different scalability modes in H.263+: temporal, SNR and spatial. For the simulations of mode 2 and mode 4, we choose SNR scalability. The BL is a low but acceptable quality representation of video obtained by using a relatively large quantization step size, and the EL is the difference between the original frame and the BL, represented with a smaller quantization step size. In our setup, besides the first I-frame, the BL is dependent on one previous BL frame; while the EL is dependent on the current BL frame as well as one previous EL frame. Thus a loss in the EL bitstream will not affect the reconstruction of the BL frames. Scalability is an error resilience feature since the more essential BL can be prioritized when network or system resources are limited. Other error resilience features [31], [32] offered in H.263+ include the use of resynchronization markers. In our experiment, we group the MBs into group of block (GOB), with each GOB containing one row of MBs. We add a resynchronization marker at the beginning of each GOB. The GOB header is useful when decoding a video frame that spans multiple channel frames. If a channel frame is lost, only the affected GOBs are lost (replaced by “0” bits). The video decoder will search for the next GOB header and resume decoding from there. Hence, instead of losing the entire video frame, some GOBs can still be decoded. For those MBs that cannot be decoded, we use a simple frame copy error

concealment method that copies the corresponding block from the previously reconstructed frame.

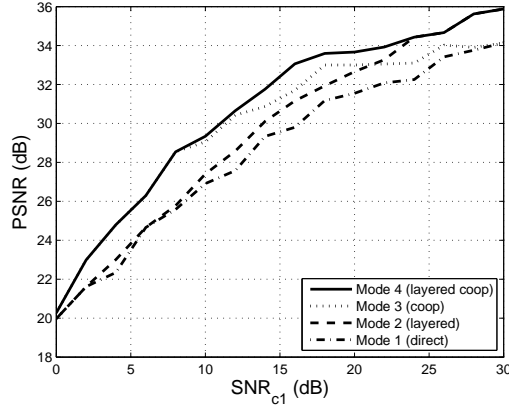
We fill the channel frames with encoded video bits according to the channel frame bit allocation assuming negligible length packet header at the physical layer. We put  $R_b$  successive compressed BL bits and  $R_e$  successive compressed EL bits into their respective layers. Generally there is no alignment between GOBs and channel frames. The channel coder then applies channel coding according to the modes described in Sec. II. Given a fixed channel frame rate  $F$  frames/sec, for a candidate source bit allocation of  $R_b$  and  $R_e$  bits per channel frame, the video bit rates are  $V_b = F \cdot R_b$  bits/sec for BL and  $V_e = F \cdot R_e$  bits/sec for EL. We use the rate control algorithm in [30] to achieve the desired bit rates. However, the resulting bit rate is not exactly the same as the target rate. Although we choose the simulation parameters so that the target number of bits per video frame (including source and channel coding bits) will fit an integer number of channel frames, the number of source bits produced per video frame generally varies. As we packetize the compressed video into channel frames, we start a new video frame in a new channel frame, which implies that the last channel frame in a video frame may not be fully packed. Stuffing the unused bits with “0”s, we actually operate at a bit rate slightly higher than that specified by the bit allocation. Another assumption we make is that the first video frame in the sequence, as well as all video frame headers are uncorrupted during transmission. In practice, this may be achieved by a guaranteed out-of-band channel. Instead of minimizing expected distortion as we did for the Gaussian source, we quantify video quality with the popular objective video quality measure, average peak signal to noise ratio (PSNR) of the decoded video. The framewise PSNR is defined as:

$$PSNR = 10 \log_{10} \frac{255^2}{MSE},$$

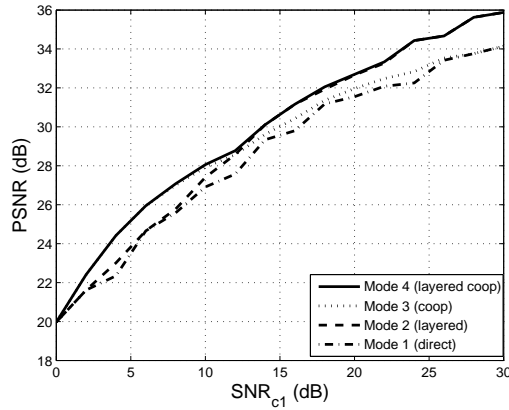
where MSE is the mean squared error between the original and decoded luminance component of a video frame. We average the PSNR over all transmitted and decoded frames. Therefore it also includes the averaging over channel fading. We maximize the decoded video PSNR jointly over source coding ( $R_b$ ,  $R_e$  and  $\beta$ ), channel coding ( $r_{b1}$ ,  $r_e$ ) and cooperation ( $r_{b2}$ ).

## B. Video Simulation and Results

For the FER simulations, we use the same channel code, modulation level, and channel frame length as in Sec. III-D. For the Gaussian source, the coherence time of the channel is not important as long as it is a multiple of channel frames. For video, however, coherence time determines how many video frames observe the same fading level, thus affecting the quality of the reconstructed signal. We assume a coherence time of 0.1 sec. For 3G, 802.11 WLAN and WIMAX systems, each with carrier frequencies in the GHz. range, this would model a pedestrian user. We fix our channel frame rate  $F$  to be 100 frames/sec and our video frame rate to be 10 video frames/sec (fps). Therefore each video frame is sent with approximately 10 channel frames, during which the fading level stays constant. Because each channel frame carries 1728



(a) Perfect inter-user channel.



(b) 14dB inter-user channel.

Fig. 7. Average end to end quality (in terms of PSNR) of “Foreman” video stream vs. average channel SNR,  $SNR_{c1}$  in dB with BPSK modulation and RCPC channel coding for different communication modes. We have symmetric cooperation with  $SNR_{c1} = SNR_{c2}$ .

bits, the target total video bit rate (including source and channel coding) is 172.8 kbits/sec (Kbps). This bit rate is appropriate for low resolution video transmission over wireless channels (where a significant portion of the bit rate may be used for channel coding).

We carry out experiments on three standard test video sequences in QCIF format (176x144 Y pixels per frame). The sequences are: A low-motion sequence “Mother & Daughter” (300 frames); an intermediate motion sequence “Foreman” (300 frames, with large panning at the end of the sequence); and a high-motion sequence “Football” (208 frames). Skipping 2 out of every 3 frames, the original 30 fps test sequences are converted to 10 fps and compressed for transmission. We loop each video sequence as necessary to transmit 10,000 channel frames. Considering total bit rate of 172.8 Kbps, we vary the BL bit rate from 43.2 to 172.8 Kbps, and the EL rate from 0 to 86.4 Kbps.

Our results for “Foreman” using symmetric cooperation

$SNR_{c1}$	$R_b$	$r_{b1}$	$r_{b2}$	$R_e$	$r_e$	$\beta$	PSNR
0	432	432	864	0	0	0.2	20.277
2	432	432	864	0	0	0.2	22.986
4	432	432	864	0	0	0.166	24.800
6	432	432	864	0	0	0.166	26.292
8	432	432	864	0	0	0.04	28.540
10	816	0	816	96	0	0.1	29.345
12	632	0	632	232	232	0.05	30.655
14	608	0	608	256	256	0.05	31.763
16	656	0	656	208	208	0	33.061
18	736	0	736	256	0	0	33.594
20	736	0	736	256	0	0	33.666
22	736	0	736	256	0	0	33.927
24	1520	0	0	208	0	0.1	34.431
26	1536	0	0	192	0	0.083	34.666
28	1576	0	0	152	0	0	35.624
30	1576	0	0	152	0	0	35.878

TABLE I

OPTIMAL BIT ALLOCATION AND  $\beta$  FOR VIDEO “FOREMAN” AT MODE 4, SYMMETRIC COOPERATION, PERFECT INTER-USER CHANNEL.

( $SNR_{c1} = SNR_{c2}$ ) are shown in Figure 7 for perfect inter-user channel and for  $SNR_{c12} = 14$  dB. The trends are similar to the Gaussian source trends. Layered cooperation (mode 4) consistently outperforms direct transmission (mode 1), and improves over other modes at different average  $SNR_{c1}$  ranges. When the inter-user channel is perfect, the improvement of cooperation over direct transmission is more pronounced for video than for Gaussian source. Better channel robustness from cooperation reduces not only distortion in the currently affected video frame, but also that in the future frames which are dependent upon this frame through MC.

Note that noise in the inter-user channel dampens the benefits of cooperation more for video than for i.i.d. Gaussian source where the difference between  $SNR_{c12} = \infty$  and  $SNR_{c12} = 14$  dB was insignificant. The reason behind this still needs to be thoroughly explored. One possible explanation is that the inefficiency of the practical source encoder limits the strength available for error protection in the inter-user channel. Another reason is that any residual channel error lowers the decoded video quality more significantly than for the Gaussian case, because of the error propagation effect.

For video delivery, we are mainly interested in improving the video quality at low to medium  $SNR_{c1}$  range because at higher  $SNR_{c1}$ , the quality is already acceptable without cooperation. As shown in Figure 7, user cooperation (mode 3 and mode 4) brings significant improvement to video quality (in terms of PSNR) at low to medium  $SNR_{c1}$  range. Furthermore, mode 4 (layered cooperation) offers distinct improvement over all other modes in the intermediate  $SNR_{c1}$  range.

We show the corresponding optimal bit allocations in Table I for mode 4 at perfect  $SNR_{c12}$ . As expected, as  $SNR_{c1}$  increases, less error protection is needed and hence generally both channel coding bits (including cooperation bits) and the intra rate decreases. However, when a “sharp transition” occurs in the bit allocation with a significant drop in the channel coding bits (e.g. from  $SNR_{c1} = 8$  to  $SNR_{c1} = 10$  and from  $SNR_{c1} = 22$  to  $SNR_{c1} = 24$  in Table I),  $\beta$  increases slightly

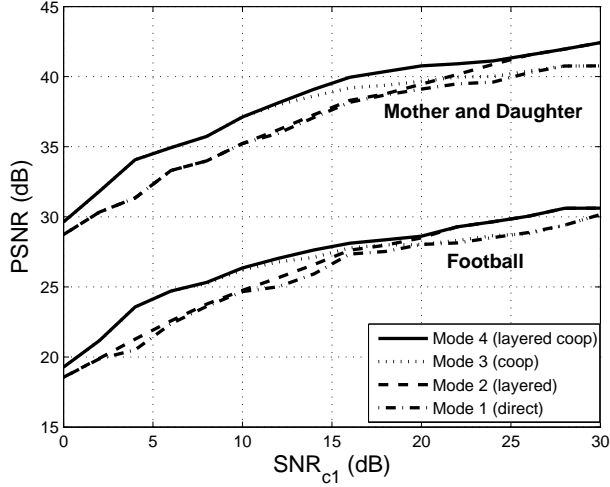


Fig. 8. Average end to end quality (in terms of PSNR) of “Football”, and “Mother & Daughter” video streams as a function of channel SNR,  $SNR_{c1}$  (dB), with BPSK modulation and RCPC channel coding for different communication modes. We have symmetric cooperation ( $SNR_{c1} = SNR_{c2}$ ) with perfect inter-user channel.

to offer more protection against error propagation. We can conclude that the parameter  $\beta$  here acts to “fine tune” the quality. We note that this discontinuity in the optimal values for the coding parameters may be due to the fact that we did not search over a continuous range for each parameter, which is computationally prohibitive.

We carry out the same experiments on two other video sequences: the lower-motion sequence “Mother & Daughter” and the higher-motion sequence “Football”. Figure 8 shows the results for the symmetric scenario with perfect inter-user channel case. Besides the expected higher PSNR for the low-motion “Mother & Daughter” (and vice versa for the faster “Football”), we make two observations on the trend as  $SNR_{c1}$  increases from low to high: First, the layered modes (mode 2 and 4) start to strictly dominate the single-layer modes (mode 1 and 3) at a lower  $SNR_{c1}$  when the video has higher motion. Higher motion video requires more source bits but at the same time, it is also more affected by channel errors because distortion from error-propagation is more severe. Layering therefore offers the best solution by assigning more bits for the video source through EL, while maintaining the channel coding rate to protect the essential BL. We conclude from here the importance of layered compression with UEP for high motion video. Our second observation is that, the cooperation modes (3 and 4) converge to the no-cooperation modes (1 and 2) at a lower  $SNR_{c1}$  for higher motion video. Since the final video distortion is contributed by both channel errors and lossy video compression, as the amount of motion increases, the compression distortion is more severe for the same source coding rate. Compared to low motion video, the channel-error induced distortion thus becomes less dominant for higher motion sequence at a lower channel  $SNR_{c1}$ .

In the case of asymmetric cooperation  $SNR_{c2} = SNR_{c12} = 4SNR_{c1}$  for video, we measure video PSNR for the “Foreman” sequence as  $SNR_{c1}$  varies from 0 to 30 dB.

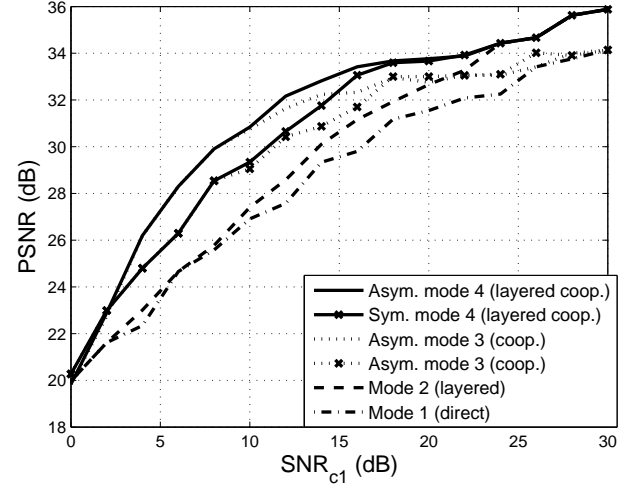


Fig. 9. Video PSNR for Foreman, QCIF, 10fps, vs. channel SNR,  $SNR_{c1}$  (dB) for both asymmetric and symmetric scenarios. The symmetric scenario assumes perfect inter-user channel.

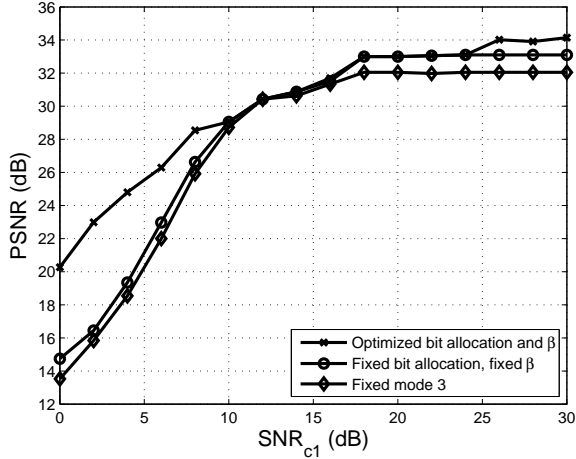
The results are plotted in Figure 9. The trends are similar to the Gaussian source and the end-to-end PSNR is higher than the symmetric case.

It is of interest to find out if the optimal bit allocation provides substantial PSNR gain over one that is not optimized. For both mode 3 and mode 4, we choose the optimal bit allocation and  $\beta$  for “Foreman” for  $SNR_{c1} = 12dB$ , a mid-range  $SNR_{c1}$ . Then, in Figure 10, we plot the PSNR for the entire range of  $SNR_{c1}$  using this parameter set for symmetric cooperation with perfect inter-user channel. Also present in the graph is the PSNR result of optimizing only  $\beta$ , but not the bit allocation. The results show that optimization provides substantial gains (up to 6 dB) in PSNR for both mode 3 and mode 4 at low  $SNR_{c1}$ . When  $SNR_{c1}$  is high, the gain is less, but still significant (up to 3 dB). Therefore adaptation of the source and channel coding and cooperation based on channel conditions can bring significant gains. Furthermore, the system performance is less sensitive to the mismatch between the actual channel SNRs and assumed average channel SNRs in the high SNR range, but is more sensitive to the mismatch in the low SNR range.

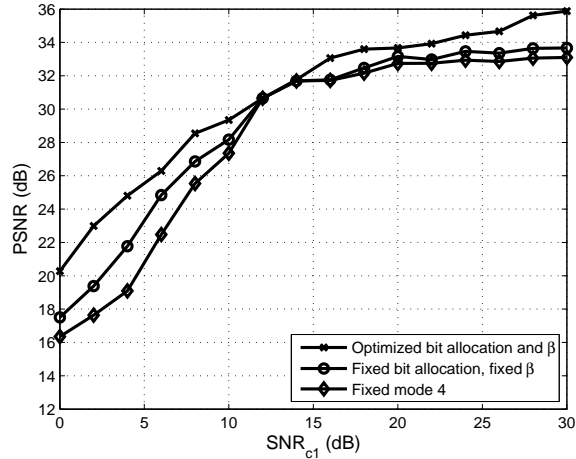
## V. CONCLUSIONS

User cooperation is a popular spatial diversity technique that has been previously explored from a physical layer perspective. In this paper, we explore a cross layer approach and utilize user cooperation to maximize the end-to-end source quality. We optimize the parameters of source coding, channel coding and cooperation jointly to minimize the expected source distortion at the receiver. Making use of the increased channel reliability offered by cooperation, we introduce a layered cooperation scheme that protects the more important base layer through cooperation and stronger channel coding.

Our information theoretic analysis for an i.i.d. Gaussian source and Gaussian channel codebooks (no constraint on maximum modulation level) suggests that cooperative coding and layered cooperation are superior to non-cooperative



(a) Mode 3



(b) Mode 4

Fig. 10. Optimization vs. no Optimization comparison for the quality (in terms of PSNR) of Foreman video sequence. “Fixed” corresponds to the result using parameters optimized for  $SNR_{c1} = 12\text{dB}$  in a symmetric, perfect inter-user channel setup, (a) compares the results for single layer cooperative coding, (b) compares results for layered cooperation.

strategies for a wide range of bandwidth ratios and average link SNRs. In fact layered cooperation outperforms all other modes in terms of end-to-end average distortion and provides a higher rate of decrease in the distortion as a function of channel SNR. The results in this paper are consistent with [13], [14], where we concluded that in communication over fading channels, layered source coding dramatically improves the minimum expected distortion for high SNRs. When the modulation is constrained, we observe that for both i.i.d. Gaussian source and video source, single layer cooperation can significantly improve the source quality for low to medium channel SNR values. At higher channel SNRs, layered compression with UEP through channel coding is necessary for further improvements. By combining user cooperation and layered

compression, we are able to improve the source quality for a larger channel SNR range. The layered cooperation scheme also provides a distinct improvement over either cooperation alone or layering alone for a mid range of channel SNRs. Our results are confirmed by both information theoretic analysis based on outage probability of the transmission and simulation of a practical modulation/channel coding scheme.

The importance of layering is particularly profound for high motion video. Our study illustrates that for symmetric users, when a good quality inter-user cooperative channel is present, cooperation provides a slightly better improvement for a video source when compared with a Gaussian source. However, the presence of temporal error propagation in the video decoder also causes the performance to be more sensitive to the inter-user channel. On the other hand, if the relay terminal can provide a better quality channel to the user’s destination, as in our asymmetric cooperation, video quality can be enhanced greatly by cooperation or layered cooperation even if the inter-user channel is not perfect.

Finally, we would like to comment on how adaptation of bit allocation among source coder, channel coder and cooperation can be accomplished in practical systems. Note that our optimization problem is formulated based on average received channel SNR’s rather than instantaneous fading levels. Therefore, we can precompute the optimal bit allocations for different combinations of  $SNR_{c1}$ ,  $SNR_{c2}$ , and  $SNR_{c12}$ , and store the results in a look up table at the sender. We assume that the values of  $SNR_{c1}$ ,  $SNR_{c2}$ , and  $SNR_{c12}$  can be estimated at the receivers (relay and the destination) and fed back to the sender periodically, and the sender will choose its bit allocation according to the look-up table. We further assume the sender can inform the relay and the destination its bit allocation using some auxiliary data packets. Note that, the transmission of the SNR and bit allocation information requires a very low rate link since average channel SNR’s change at a very low rate. Because the optimal bit allocations do not need to be computed in real time, their computational complexity is not a limiting factor in the system design.

## REFERENCES

- [1] M. Etoh, T. Yoshimura, “Advances in wireless video delivery,” *Proceedings of IEEE*, 93(1): 111-122, Jan. 2005.
- [2] A. Sendonaris, E. Erkip, B. Aazhang, “User cooperation diversity-Part I: System description,” *IEEE Transactions on Communications*, 51(11):1927-1938, Nov. 2003.
- [3] A. Sendonaris, E. Erkip, B. Aazhang “User cooperation diversity-Part II: Implementation aspects and performance analysis,” *IEEE Trans. on Communications*, 51(11):1939-1948, Nov. 2003.
- [4] J. N. Laneman, D.N.C. Tse, G. W. Wornell “Cooperative diversity in wireless networks: Efficient protocols and outage behavior,” *IEEE Trans. on Information Theory*, 50(12): 3062-3080, Dec. 2004.
- [5] T. Hunter, A. Nosratinia, “Diversity through coded cooperation,” *IEEE Trans. on Wireless Communications*, vol. 5, no. 2, pp. 283-289, Feb. 2006.
- [6] T. Hunter, S. Sanayei, and A. Nosratinia, “Outage analysis of coded cooperation,” *IEEE Trans. on Information Theory*, vol. 52, no. 2, pp. 375-391, Feb. 2006.
- [7] A. Stefanov, E. Erkip “Cooperative coding for wireless networks,” *IEEE Trans. on Communications*, 52(9): 1470-1476, September 2004.
- [8] J. N. Laneman and G. W. Wornell, “Distributed Space-Time Coded Protocols for Exploiting Cooperative Diversity in Wireless Networks,” *IEEE Trans. Inform. Theory*, vol. 49, no. 10, pp. 2415-2525, Oct. 2003.
- [9] A. Wittneben, “Coherent Multiuser Relaying with Partial Relay Cooperation,” *IEEE Wireless Communication and Networking Conference*, Las Vegas, NV, USA, Apr. 2006.

- [10] O. Shalvi, "Multiple source cooperation diversity," *IEEE Communications Letters*, vol. 8, no. 12, pp. 712-714, Dec. 2004.
- [11] D. Gündüz, E. Erkip, "Joint source-channel cooperation: Diversity versus spectral efficiency," *Proc. of 2004 ISIT*, Chicago, IL, June 2004.
- [12] X. Xu, D. Gündüz, E. Erkip, Y. Wang, "Layered cooperative source and channel coding," *IEEE International Conference on Communications*, Seoul, Korea, May. 2005.
- [13] D. Gündüz and E. Erkip, "Source and channel coding for cooperative relaying," *In Proceedings of 2005 International Workshop on Signal Processing Advances for Wireless Communications (SPAWC)*, New York, NY, June 2005.
- [14] D. Gündüz, E. Erkip, "Source and channel coding for cooperative relaying," *IEEE Transactions on Information Theory*, submitted August 2006, revised February 2007.
- [15] X. Xu, Y. Wang and E. Erkip, "Layered cooperation for wireless multimedia communications," *In Proceedings of 2004 Picture Coding Symposium*, San Francisco, CA, Dec. 2004.
- [16] H. Shutoy, Y. Wang, E. Erkip, "Cooperative source and channel coding for wireless video transmission," *in Proc. of IEEE International Conference on Image Processing (ICIP)*, Atlanta, Georgia, Oct. 2006.
- [17] J. N. Laneman, E. Martinian, G. W. Wornell, and J. G. Apostolopoulos, "Source-channel diversity for parallel channels," *IEEE Transactions on Information Theory*, vol. 51, no. 10, pp. 3518-3539, Oct. 2005.
- [18] A. Kwasinski, Z. Han, K.J.R. Liu, "Joint source-channel coding for cooperative multimedia communications," *In Proceedings of 2005 International Workshop on Signal Processing Advances for Wireless Communications (SPAWC)*, New York, NY, June 2005.
- [19] W. Equitz, T. Cover, "Successive refinement of information," *IEEE Transactions on Information Theory*, 37(2): 269-275, Mar. 1991.
- [20] L. H. Ozarow, S. Shamai, and A. Wyner, "Information theoretic considerations for cellular mobile radio," *IEEE Trans. Veh. Technol.*, vol. 43, pp. 359-378, May 1994.
- [21] L. Zheng, D. Tse, "Diversity and multiplexing: A fundamental tradeoff in multiple antenna channels," *IEEE Trans. on Information Theory*, vol. 49, pp. 1073-96, May 2003.
- [22] J. Hagenauer, "Rate compatible punctured convolutional codes (RCP codes) and their applications," *IEEE Trans. on Communications*, 36(4), 389-400, April 1998.
- [23] ITU-T. Recommendation H.263, "Video coding for low bit rate communication," 1998.
- [24] G. Cote, B. Erol, M. Gallant, F. Kossentini, "H.263+: Video coding at low bit rates," *IEEE Trans. on Circuits and Systems for Video Technology*, 8(7):849-866, November 1998.
- [25] T.M. Cover and J.A. Thomas, *Elements of Information Theory* John Wiley & Sons, 1991.
- [26] L. Choi, M. T. Ivrla, E. Steinbach, and J. A. Nossek, "Sequence-level models for distortion-rate behavior of compressed video," *Proc. IEEE. Conf. Image Process*, Sept. 2005, vol. II, pp. 486-489.
- [27] Z. He and D. Wu, "Resource Allocation and Performance Analysis of Wireless Video Sensors," *IEEE Trans. Circuits Syst. Video Technol.*, 16(5):590-599, May 2006.
- [28] K. Stuhlmüller, N. Färber, M. Link and B. Girod, "Analysis of video transmission over lossy channels," *IEEE Journal on Selected Areas in Communications*, 18(6): 1012- 1032, June 2000.
- [29] Y. Wang, Z. Wu, J. M. Royce, "Modeling of transmission-loss-induced distortion in decoded video," *IEEE Trans. Circuits and Systems for Video Technology*, vol. 16, no. 6, pp. 716-732, June 2006.
- [30] Signal Processing and Multimedia Lab., Univ. British Columbia, "TMN 8 (H.263+) encoder/decoder, Version 3.1.3," TMN 8 (H.263+) codec, Feb. 1998.
- [31] Y. Wang, S. Wenger, J. Wen, A.K. Katsaggelos, "Error resilient video coding techniques," *IEEE Signal Processing Mag.*, 17(4):61-82, July 2000.
- [32] S. Wenger, G. Knorr, J. Ott, F. Kossentini, "Error resilience support in H.263+," *IEEE Trans. Circuits Sys. Video Tech.*, 8(6): 867-877, Nov. 1998.