UC San Diego
UC San Diego Previously Published Works

Title

Error-resilient video communications over CDMA networks with a bandwidth constraint

Permalink

https://escholarship.org/uc/item/86x3d8wx

Journal

IEEE Transactions on Image Processing, 15(11)

ISSN

1057-7149

Authors

Shen, Yushi, Dr.
Cosman, P C, Prof.
Milstein, L B, Prof.

Publication Date

2006-11-01

Peer reviewed

Error-Resilient Video Communications Over CDMA Networks With a Bandwidth Constraint

Yushi Shen, Pamela C. Cosman, *Senior Member, IEEE*, and Laurence B. Milstein, *Fellow, IEEE*

Abstract—We present an adaptive video transmission scheme for use in a code-division multiple-access network, which incorporates efficient bandwidth allocation among source coding, channel coding, and spreading under a fixed total bandwidth constraint. We derive the statistics of the received signal, as well as a theoretical bound on the packet drop rate at the receiver. Based on these results, a bandwidth allocation algorithm is proposed at the packet level, which incorporates the effects of both the changing channel conditions and the dynamics of the source content. Detailed simulations are done to evaluate the performance of the system, and the sensitivity of the system to estimation error is presented.

Index Terms—Bandwidth allocation, channel estimation, code-division multiple-access (CDMA), joint source-channel coding, mode switching, multiuser system, video communications, wireless Internet.

I. INTRODUCTION

WITH the increasing demand for multimedia services on mobile terminals, and the recent advances in mobile computing, video services are expected to be widely deployed. Direct-sequence code-division multiple-access (DS-CDMA) technology is useful because it exhibits robust performance against channel fading and interference, as well as good multiple access capacity [1]. Source coding, channel coding, and spread spectrum are the three main components in most CDMA communications systems. Source coding frees up bandwidth for both channel coding and spreading, while channel coding and spreading protect the transmitted bits from noise, interference and fading [2].

The major resource shared among the three components is the total bandwidth, which is equivalent to the transmission chip rate. We denote by R_s the source bit rate (in bits per second) and by W the chip rate (in chips per second); also, we denote by r_c the channel code rate and by M the processing gain. The variables R_s, r_c, and M are constrained as follows:

$$\frac{M}{r_c} R_s = W. \tag{1}$$

Another constraint in video communications is the total transmission time, which is equivalent to achieving an average target frame rate f_s (frames per second). Our goal is, for a target chip rate W and a target frame rate f_s, to find the optimal bandwidth allocation (R_s^*, r_c^*, M^*) for each packet, such that the expected distortion over all frames between the transmitted video bitstream and the reconstructed video bitstream at the decoder is minimized.

In [3] and [4], bandwidth allocation algorithms were proposed for image transmission, where a simplified source, with either a uniform or a Gaussian distribution, and an optimal scalar quantizer, were assumed. In [5]–[8], bandwidth allocation was discussed for image or video communications over either a single channel or multiple channels, and the allocation of system parameters was chosen for the entire transmission duration. In this paper, we present a robust CDMA scheme with efficient bandwidth allocation for the transmission of packet video over a single wireless channel, which is assumed to be frequency selective with Rayleigh amplitude statistics. The choice of bandwidth allocation is made adaptively at the packet level, and incorporates the effects of both the changing channel characteristics and the current video content. This algorithm can be easily extended to a tandem channel with both wireline and wireless links.

This paper is organized as follows. In Section II, the system model and channel model are briefly introduced. In Section III, we analyze the statistics of the received signal, and present a theoretical bound on the block error rate. Based on the results, Section IV presents strategies for channel coding and/or spreading selection, given a target packet drop rate. In Section V, the algorithm to allocate bandwidth for each packet is presented, which solves the optimization problem by a two step approach based on the results illustrated in Sections II–IV. The statistical characteristics of channel estimates are derived and analyzed in Section VI. System performance and its sensitivity to channel estimation errors are presented numerically in Section VII. Last, Section VIII concludes the paper.

II. SYSTEM MODEL

The system diagram is shown in Fig. 1, and we will discuss the components in detail in this section.

A. Source Coding

In our modified H.263+ video compression, each frame is segmented into macroblocks (MBs) of size 16×16 pixels. The encoding mode and the quantization step are selected for each MB individually. Intercoding and intracoding are the two basic encoding modes: intercoding, which encodes the current block by using the most similar displaced block in the previous frame and sending the difference, compresses efficiently

The authors are with the Department of Electrical and Computer Engineering, University of California, San Diego, La Jolla, CA 92093-0407 USA (e-mail: yushen@code.ucsd.edu; pcosman@code.ucsd.edu; lmilstein@ucsd.edu).
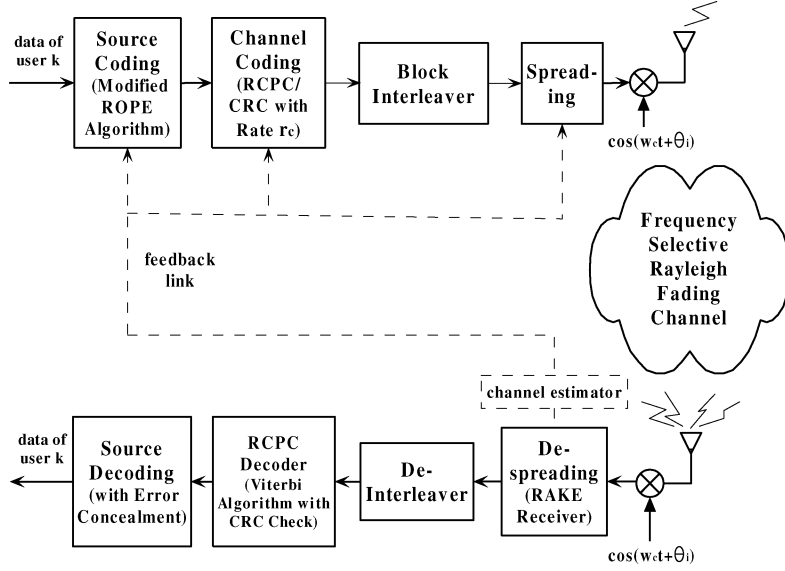
Fig. 1.   System overview.

but may suffer from error propagation; on the other hand, intra-coding, which encodes the current block by itself, can stop error propagation at the expense of inefficient compression. It is desired to switch between intra and inter coding efficiently.

The source encoder in Fig. 1 uses an algorithm with fixed-length packetization called modified-ROPE [9], [10]. Given the packet loss rate at the receiver and a target source bit rate, the modified-ROPE algorithm optimally chooses the inter/intra mode and quantization step for each MB according to the expected distortion at the pixel level.

Over a wireless channel, fading, noise, and interference will cause packets to be dropped, due to uncorrectable bit errors. We denote the packet drop rate by p_p. It is shown in Section III that p_p is a function of r_c and M, given the other system parameters and channel characteristics.

B. Channel Coding

A concatenated code is used for forward error correction (FEC), which consists of a rate-compatible punctured convolutional (RCPC) inner coder and a cyclic redundancy check (CRC) outer coder. The encoder chooses a RCPC code from a family of RCPC code candidates for each packet. Denoting by r_{RP} the rate of a nonpacketized RCPC code, the channel coding rate of the resultant block code is

$$r_c = \frac{y - 16 - Z}{y} r_{\mathrm{RP}} \qquad (2)$$

where y is the fixed packet length at the input of the RCPC encoder, 16 bits are used for the CRC check, and Z is the number of zero tail bits to terminate the packet [10]. In our system, the RCPC code candidates have rate r_{RP} equal to 1/3, 2/3, 8/9, and 1, with memory $Z = 6$ [2], [10]. The serial list-Viterbi algorithm at the channel decoder is used to find the best candidate in the trellis, and the packet is discarded if none of the first 100 paths with the maximal metric satisfies the CRC checksum equations [10].

C. Signal Spreading

Before going through the wireless channel, an interleaver is used to randomize the error bursts in the transmitted data block. The interleaved data stream is then spread using direct sequence with a long spreading code by a factor of M (processing gain), and transmitted using BPSK modulation.

We consider an asynchronous CDMA up-link system with K users [2]. Denote by $a_k(t)$ the signature sequence waveform for the kth user ($0 \leq k \leq K - 1$), and by $a_{k,j}$ the corresponding sequence element, where $a_{k,j} \in (+1, -1)$. Similarly, denote by $b_k(t)$ the data signal waveform for the kth user, and by $b_{k,j}$ the corresponding sequence element, where $b_{k,j} \in (+1, -1)$. Also, denote by T_s and T_c the time duration of each symbol (i.e., channel bit) and each chip, respectively, with $T_s = MT_c$. Then, $a_k(t) = \sum_{j=-\infty}^{\infty} a_{k,j} P_{T_c}(t - jT_c)$ and $b_k(t) = \sum_{j=-\infty}^{\infty} b_{k,j} P_{T_s}(t - jT_s)$, where $P_{T_c}(t)$ is the pulse shape function so that $P_{T_c}(t) = 1$ for $0 \leq t < T_c$ and zero elsewhere, and $P_{T_s}(t)$ is the pulse shape function with $P_{T_s}(t) = 1$ for $0 \leq t < T_s$.

Denote by E_b, E_s, and E_c the energy-per-information bit, per-channel bit and per-chip, respectively. Also, denote by A the amplitude of the transmitted signal. We assume perfect power control, thereby implying that all users have the same received power. Then, we obtain

$$E_c = \frac{E_s}{M} = \frac{r_c}{M} E_b = \frac{1}{2} A^2 T_c. \qquad (3)$$

Because both the total video transmission time and the bandwidth (thus, the chip duration T_c) are fixed, keeping the total transmission energy constant is equivalent to keeping the energy per chip (E_c) constant. Last, the transmitted signal for the kth user is given by

$$s_k(t) = Re[S_k(t)e^{j\omega_c t}], \quad \text{where } S_k(t) = Aa_k(t)b_k(t). \quad (4)$$

D. Wireless Channel Model

A linear tapped delay line filter is used to model the frequency selective Rayleigh fading channel, with a lowpass equivalent impulse response for the kth user $h_k(t) = \sum_{l=1}^{L} c_{k,l}(t) \delta(t - lT_c - \tau_{k,l}(t))$, where $\delta(t)$ is the Dirac delta function, L is the number of resolvable paths (which is assumed constant over time), and $c_{k,l}(t) = \alpha_{k,l}(t)e^{j\theta_{k,l}(t)}$ is the complex gain. The $\alpha_{k,l}$s, $\theta_{k,l}$s and $\tau_{k,l}$s are the random path amplitudes, phases and delays. Assume the $\alpha_{k,l}$s, $\theta_{k,l}$s and $\tau_{k,l}$s are independent for different k (but correlated over time).

We assume slow fading, that is, $T_c/T_{\text{coh}} \ll 1$, where $T_{\text{coh}} = c/(f_c v)$ is the coherence time of the channel, v is the speed of the mobile and c the speed of light. Considering a scenario where f_c is 900 MHz, a typical T_{coh} is 10^{-2} s at highway speeds (75 mph) and 2.5×10^{-2} at local speeds (30 mph). For a packet size of $y = 400$ bits, and a source bit rate R_s greater than 50 kbps, T_{coh} is typically equal to the transmission time of at least 2 packets. It means the fading is slow enough that the bits inside the same packet see the same fading amplitude and phase.

The received signal for the kth user over the kth channel is $r_k(t) = s_k(t) * h_k(t) + n_k(t)$, and the composite signal at the output of the channel is $r(t) = Re[R(t)e^{j\omega_c t}]$, where

$$R(t) = \sum_{k=0}^{K-1} \sum_{l=1}^{L} \alpha_{k,l} S_k(t - lT_c - \tau_{k,l})e^{j\varphi_{k,l}} + N(t). \quad (5)$$

In (5), $\varphi_{k,l} = \theta_{k,l} - \omega_c \tau_{k,l}$, and $N(t)$ is complex Gaussian noise with two-sided power spectral density N_0 [1]. The $\varphi_{k,l}$s are independent identically distributed (i.i.d.) random variables (rvs), uniformly distributed in $[0, 2\pi)$, the $\tau_{k,l}$s are i.i.d. rvs, uniformly distributed in $[0, T_s)$, and the $\alpha_{k,l}$s are independent Rayleigh rvs with density function $p(\alpha_l) = (\alpha_l/\sigma_l^2)\exp(-\alpha_l/2\sigma_l^2)$. Let $\Omega_l = E[(\alpha_l)^2] = 2\sigma_l^2$ and $\Omega_1 = 1$. For an exponential multi-path intensity profile (MIP), i.e., one where the energy of each path is decaying exponentially, we have $\Omega_l = \Omega_1 e^{-\nu(l-1)}$, and, thus, $\sum_{l=1}^{L} \Omega_l = (1 - e^{-\nu L})/(1 - e^{-\nu})$.

III. SYSTEM ANALYSIS

In this section, we derive the statistics of the received signal, and the theoretical bound on the packet drop rate p_p due to wireless fading and interference.

A. Statistics of Received Signal

As shown in Fig. 2, we use an L_R-finger CDMA RAKE receiver [1], [2], where the matched filter is matched to the reference user's signature code and is assumed to have achieved time synchronization. Note that L_R may be less than L due to complexity constraints and the omission of the weakest paths; however, here we choose $L_R = L$ for simplicity. Let the 0th user be the reference user. In the current analysis, we assume the tap weights and phases are perfect estimates, so that $\hat{c}_{k,l}^* = c_{k,l}^* = \alpha_{k,l}e^{-j\theta_{k,l}}$; equivalently, we assume $\theta_{0,l} = \tau_{0,l} = 0$, and, thus, $\varphi_{0,l} = 0$. Later, in Section VI, the effects of estimation errors will be considered.

The lowpass output statistic of the RAKE receiver is given by $g_n(t) = \sum_{m=1}^{L} A c_{0,m}^* q(t - (L - m)T_c)$, where $q(t)$ is the output of the matched filter, given by $q(t) = R(t) * a_0(T_s - nt)$. We can express $g_n(t) = g_S(t) + g_{I_s}(t) + g_{I_m}(t) + g_N(t)$,
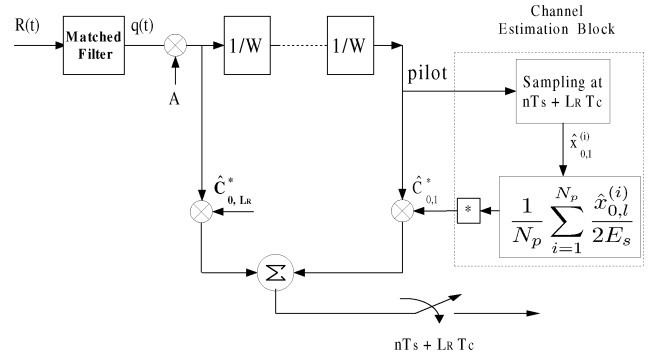

Fig. 2. RAKE receiver, with the channel estimation block.

where it is seen that $g_n(t)$ includes four components: the signal $g_S(t)$, the self-interference $g_{I_s}(t)$, the multiaccess-interference $g_{I_m}(t)$ and the Gaussian noise $g_N(t)$. At the sampling time $t = nT_s + LT_c$, we simplify the notation to $g_n = g_S + g_{I_s} + g_{I_m} + g_N$.

Conditioned on $b_0 = 1$, the signal component g_S is

$$g_S = A^2 M T_c \sum_{m=1}^{L} \alpha_{0,m}^2 = 2E_s \sum_{m=1}^{L} \alpha_{0,m}^2. \quad (6)$$

Conditioned on $\{\alpha_{0,m}\}_{m=1}^{L}$, the noise component g_N is a conditional complex zero-mean Gaussian rv with variance

$$\sigma_{g_N}^2 = E[g_N g_N^* | \{\alpha_{0,m}\}] = 4E_s N_0 \left(\sum_{m=1}^{L} \alpha_{0,m}^2\right). \quad (7)$$

If the number of users K is sufficiently large, g_{I_m} is asymptotically a conditional zero-mean Gaussian rv with variance

$$\sigma_{g_{I_m}}^2 = \frac{8E_s^2}{3M}(K - 1)\left(\sum_{l=1}^{L} \Omega_l\right)\left(\sum_{m=1}^{L} \alpha_{0,m}^2\right) \quad (8)$$

and g_{I_s} becomes negligible compared to g_{I_m} [1], [11].

Thus, for large K, the decision variable, $Re[g_n]$, is a Gaussian rv with conditional mean g_S and conditional variance $\sigma^2/2$, where $\sigma^2 = \sigma_{g_{I_s}}^2 + \sigma_{g_{I_m}}^2 + \sigma_{g_N}^2$. Combining (3)–(8), the signal-to-noise-plus-interference-ratio (SNIR) at the receiver (denoted by γ_r), conditioned on $\{\alpha_{0,m}\}_{m=1}^{L}$, is given by [11]

$$\gamma_r = \frac{g_S^2}{\sigma^2} = \left(\sum_{m=1}^{L} \alpha_{0,m}^2\right)\frac{1}{\frac{2(K-1)}{3M}\frac{1-e^{-\nu L}}{1-e^{-\nu}} + \frac{N_0}{r_c E_b}}$$

$$= \left(\sum_{m=1}^{L} \alpha_{0,m}^2\right)\frac{1}{\frac{2(K-1)}{3M}\frac{1-e^{-\nu L}}{1-e^{-\nu}} + \frac{N_0}{ME_c}}. \quad (9)$$

As a result, given the parameters (K, L, ν and chip-energy-to-noise ratio E_c/N_0) and the current channel conditions $\{\alpha_{0,m}\}_{m=1}^{L}$, γ_r is a function of r_c and M.

B. Packet Error Rate

In [12], a method was presented to find the weight enumerators, $T(x) = \sum_{d=d_{\min}}^{y} A_d x^d$, of any binary linear block codes formed from a family of RCPC codes for a given packet length y, where d is the codeword weight, A_d is the number of codewords with weight d, and $d_{\min}$ is the minimum weight (distance) of the block code. Furthermore, conditioned on the receiver
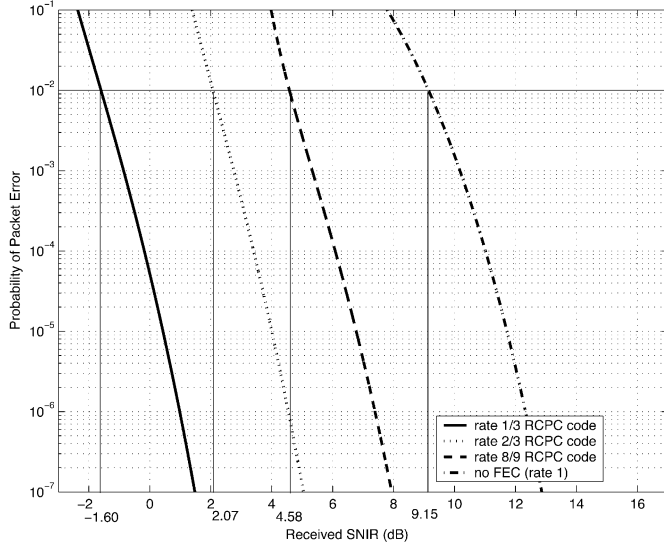
Fig. 3. Probability of packet error versus the received SNIR for the 4 candidate RCPC codes.

SNIR, γ_r, a tight bound on the packet error rate introduced by the wireless link is given by [12]

$$p_p \leq \sum_{d=1}^{y} A_d Q(\sqrt{2d\gamma_r}). \tag{10}$$

For example, for a packet length $y = 400$, the relationship between packet error rate p_p and γ_r for the four RCPC code candidates of our system is shown in Fig. 3.

In short, p_p is determined by the weight enumerator of the truncated RCPC code and the receiver SNIR, where the former is determined by the structure of the RCPC code (characterized by r_c) and packet length y, and the latter is a function of r_c and M, given the system parameters and channel characteristics. We conclude that, given the channel and system parameters, and conditioned on the current channel conditions, p_p is fully determined by r_c and M.

IV. TRADEOFF BETWEEN CHANNEL CODING AND SPREADING UNDER A TARGET PACKET DROP RATE

In this section, the tradeoff between channel coding and spreading is illustrated under a predetermined target packet drop rate (denoted by p_p^*). That is, besides the chip rate constraint (1), we impose another constraint: for each packet, the expected packet drop rate due to the wireless link will not be larger than p_p^*. For example, in [10] and [13], a target $p_p^* = 1\%$ was used. In the next section, we will relax this constraint, and allow the packet drop rate to vary based on video content.

In Section III, it was shown that p_p is determined by r_c and M, given the system parameters and channel characteristics. Also, from the perspective of the source encoder, given the packet drop rate at the receiver, maximizing the source bit rate R_s is an intuitive strategy, although it has not been shown to be optimal. We speculate that, for a video encoder with a convex operational rate-distortion (RD) curve, under a given packet drop rate, distortion would be reduced by transmitting more

source bits. This we confirmed empirically. We conducted 1 000 simulations for the modified ROPE video encoder [10] and we found that, in every case, for the given packet drop rate, the distortion was minimized when the source rate was maximized. Therefore, our baseline strategy for bandwidth allocation, for a given p_p^*, is to choose the $(\hat{r_c}, \hat{M})$ for each packet that maximizes R_s, among all the (r_c, M) combinations that achieve p_p^*. In other words, we will choose the $(\hat{R_s}, \hat{r_c}, \hat{M})$ with maximal R_s among all the combinations that satisfy the p_p^* target.

Assume the number of users K, the packet length y, and the channel parameters L and ν are known in advance. The transmitter precalculates the relationship between p_p and γ_r for each RCPC code before the data transmission and, thus, knows the SNIR threshold, $TH_{\gamma_r}(r_c)$, above which the target p_p^* is achieved. For example, for our system with 4 RCPC codes and $p_p^* = 1\%$, as shown in Fig. 3, the thresholds are as follows:

$$TH_{\gamma_r}(r_c) = \begin{cases} 9.15 \text{ dB}, & r_c = 0.945 \ (r_{RP} = 1) \\ 4.585 \text{ dB}, & r_c = 0.84 \ \left(r_{RP} = \frac{8}{9}\right) \\ 2.075 \text{ dB}, & r_c = 0.63 \ \left(r_{RP} = \frac{2}{3}\right) \\ -1.605 \text{ dB}, & r_c = 0.315 \ \left(r_{RP} = \frac{1}{3}\right). \end{cases} \tag{11}$$

Note that the relationship between r_c and r_{RP} is specified in (2). During the data transmission, the receiver estimates the $\alpha_{0,m}$s, calculates $(\sum_{m=1}^{L} \alpha_{0,m}^2)$ and sends this number back to the transmitter. Depending on whether the processing gain (M) is fixed or variable, the transmitter selects the RCPC code and/or M according to the strategies given in the following subsections, encodes the data, and sends the packet out.

Note that, instead of the receiver sending back the value $(\sum_{m=1}^{L} \alpha_{0,m}^2)$ and the transmitter doing the calculations to select the best RCPC code and/or M, the calculations can be done by the receiver, and, thus, only the RCPC code and/or M must be sent back to the transmitter.

A. Strategy With Fixed Processing Gain

For certain scenarios, the spreading gain M is predetermined and fixed for all packets during the transmission process. Under this situation, given the system parameters and channel characteristics, γ_r is solely determined by r_c, and to maximize R_s we must maximize r_c [see (1)]. Thus, we choose the RCPC code with highest coding rate that achieves $p_p \leq p_p^*$.

Specifically, the encoder will first try to use an uncoded bit stream, calculate $\gamma_r(r_c)$ by (9), and check whether it is larger than the corresponding threshold $TH_{\gamma_r}(r_c)$. If so, this RCPC code is used for the current packet; if not, it will try the RCPC code with second highest r_c, and repeat the procedure until the lowest rate code is considered.

B. Strategy With Variable Processing Gain

If both the channel coding rate r_c and the spreading gain M are variable for each packet, choosing the pair $(\hat{r_c}, \hat{M})$ that maximizes R_s among those that achieve $p_p \leq p_p^*$, is equivalent to choosing $(\hat{r_c}, \hat{M})$ with the highest (r_c/M) ratio among all the (r_c, M) pairs such that $\gamma_r(r_c, M) \geq TH_{\gamma_r}(r_c)$.

Specifically, for each candidate code rate r_c, the encoder finds the $\hat{M}(r_c)$ which is the smallest integer that satisfies the equation $\gamma_r(r_c, M) \geq TH_{\gamma_r}(r_c)$. Then, the pair $(\hat{r_c}, \hat{M})$

is chosen from all the $(r_c, \hat{M}(r_c))$ pairs such that it has the highest $(r_c/\hat{M}(r_c))$ ratio.

This strategy allows the encoder to select both r_c and M for each packet. However, it can be shown that the encoder will actually select the same r_c and only vary M most of the time. To see this, from (9) with constant E_c/N_0, we have $\hat{M}(r_c) = \bar{M}(r_c) + \delta$, where

$$\bar{M}(r_c) = \left(\frac{\frac{2}{3}(K-1)\frac{1-e^{-\nu L}}{1-e^{-\nu}} + \frac{N_0}{E_c}}{\sum\limits_{m=1}^{L} \alpha_{0,m}^2} \right) TH_{\gamma_r}(r_c) \qquad (12)$$

and $\delta \in [0, 1)$ is a fraction that makes $\hat{M}(r_c)$ an integer. Because, typically, $\bar{M}(r_c) \gg \delta$

$$\frac{r_c}{\hat{M}(r_c)} \approx \frac{r_c}{\bar{M}(r_c)} = \left(\frac{\sum\limits_{m=1}^{L} \alpha_{0,m}^2}{\frac{2}{3}(K-1)\frac{1-e^{-\nu L}}{1-e^{-\nu}} + \frac{N_0}{E_c}} \right) \frac{r_c}{TH_{\gamma_r}(r_c)}. \qquad (13)$$

Thus, given K, L, ν and $(\sum_{m=1}^{L} \alpha_{0,m}^2)$, maximizing $(r_c/\hat{M}(r_c))$ is equivalent to using the RCPC code with the highest $(r_c/TH_{\gamma_r}(r_c))$ ratio. For our example of 4 RCPC codes with $p_p^* = 1\%$, the ratio $(r_c/TH_{\gamma_r}(r_c))$ is 0.455, 0.391, 0.292, and 0.115 for the rate-1/3, rate-2/3, rate-8/9, and rate-1 RCPC codes, respectively. So the transmitter will always use the rate-1/3 RCPC code, and only M is adjusted to account for the changing channel conditions. This result illustrates that FEC is more important than spreading in a scenario where the number of users is large enough so that the self-interference due to the RAKE receiver is negligible compared to the multiaccess interference.

V. BANDWIDTH ALLOCATION ALGORITHM

The tradeoff between channel coding and spreading under a predetermined p_p^* was discussed in Section IV. However, packets are not equally important. We want to choose the p_p^* for each packet according to the video content. For example, a packet with static MBs should be relatively less sensitive to errors because of the ease of concealment, and, thus, a higher p_p^* should be acceptable.

In this section, we propose an algorithm for bandwidth allocation adaptively at the packet level, which incorporates the effects of both the changing channel conditions and the current source content. Our target is to find the bandwidth allocation among source coding, channel coding and spreading to optimize the overall performance, under the given bandwidth constraint and a target frame rate. We select a predetermined set of target packet drop rates, denoted $[p_{p1}, p_{p2}, \ldots, p_{pN}]$. During transmission, for each packet, the bandwidth allocation parameters are obtained using the following two steps.

A. Step I

We first trade off r_c and M for a given target packet drop rate due to the wireless link. That is, for each $p_{pi} \in$ $[p_{p1}, p_{p2}, \ldots, p_{pN}]$, we trade off r_c and M using the algorithm of Section IV, so that the target p_{pi} is achieved and R_s is maximized under the chip rate constraint. As a result, we obtain a set of 4-tuples $(p_{pi}, R_{si}, r_{ci}, M_i)$.

It is possible that no (r_c, M) combination can achieve a certain p_{pi} under the current channel conditions. Then, the encoder will discard this p_{pi} as a candidate allocation for the current packet. Also, if no member of $[p_{p1}, p_{p2}, \ldots, p_{pN}]$ can be achieved, the encoder will declare deep fading and temporarily send nothing (except the pilot bits for channel estimation); however, this will not happen often if the range of p_{pi} is large enough.

B. Step II

We now trade off the packet drop rate (p_p) and source coding rate (R_s) to maximize the overall performance under the chip rate constraint and frame rate target.

For each $(p_{pi}, R_{si}, r_{ci}, M_i)$ 4-tuple, the video encoder encodes the current video content (MB by MB) based on the Modified-ROPE algorithm [10] until the encoded bits fill the fixed length packet. That is, for each encoded MB, the encoder exhaustively tries all combinations of inter/intra mode and possible quantization steps, and finds the optimal mode/quantization decision.

Denote the pixel to be encoded by f, and the recovered pixel at the decoder by $\tilde{f}$, where $\tilde{f}$ is treated as a rv at the encoder. The expected distortion for each pixel is given by

$$d = E\{(f - \tilde{f})^2\} = (f)^2 - 2fE\{\tilde{f}\} + E\{(\tilde{f})^2\}. \qquad (14)$$

The formulas to calculate the two moments of $\tilde{f}$ were derived in [10], given the coding mode (inter or intra), the quantization parameter (QP), the packet drop rate, and the concealment method.

For each combination of mode and quantization step, the encoder calculates the distortion of each pixel based on the packet drop rate (p_{pi}) according to (14), and further calculates the total distortion of the current MB (denoted by D_{MB}), the number of source bits used to encode the current MB (denoted by R_{MB}), and the number of chips needed to transmit the current MB (denoted by W_{MB}), where $W_{MB} = (R_{MB}/r_{ci}) M_i$.

The optimal inter/intra mode and QP for the current MB are obtained by minimizing a Lagrange variable that is determined by both expected distortion and rate usage. Unlike [9], [10], our constraint is on chip rate rather than source bit rate. Given the frame rate f_s (frames per second) and the chip constraint W (chips per second), the target average number of chips for each MB over time, denoted W_{Target}, is equal to the constant $W/(N_{MB}f_s)$, where N_{MB} is the number of MBs per frame ($N_{MB} = 99$ for QCIF video). The extended Rate-Distortion (RD) framework is to choose the mode and the QP to minimize

$$\min_{(\text{mode}, QP)} J_{MB} = \min_{(\text{mode}, QP)} (D_{MB} + \lambda W_{MB}). \qquad (15)$$

Assuming the current encoding MB is the nth MB from the beginning, rate control is achieved by updating λ using the following algorithm modified from [9] and [10] to use chips rather than bits:

$$\lambda(n) = \lambda(n-1)[1 + \alpha(W_{used}(n-1) - (n-1)W_{Target})] \qquad (16)$$
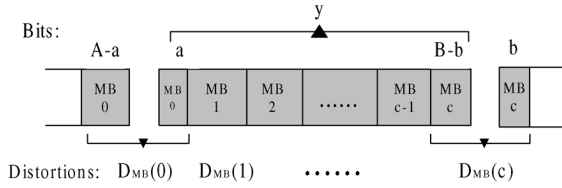
Fig. 4. Illustration of a typical packet after source encoding.

where $\alpha = 1/(5W_{\text{Target}})$, $W_{\text{used}}(n)$ is the total number of chips used up to the nth MB, and $W_{\text{used}}(n) = W_{\text{used}}(n-1) + W_{\text{MB}}(n)$.

As a result, for each 4-tuple $(p_{\text{pi}}, R_{\text{si}}, r_{\text{ci}}, M_i)$, the encoder uses the Modified-ROPE algorithm (with the modified rate control scheme for CDMA) to encode MBs until the encoded bits fill the fixed length packet. Note that this encoding strategy guarantees satisfaction of the frame rate target, under the chip rate constraint, over time.

Because fixed length packetization is employed, the end of the packet usually is not the end of the encoded bits of a MB. A typical diagram of the packetization is shown in Fig. 4. As shown in Fig. 4, y denotes the fixed packet length in bits, A and a denote the number of total bits and tail bits of the last MB of the previous packet, respectively, B and b denote the number of total bits and tail bits of the current MB, respectively, c denotes the number of encoded MBs in the current packet, and $D_{\text{MB}}(j)$ $(j = 0, \ldots, c)$ is the corresponding expected distortion ($D_{\text{MB}}(0)$ is the distortion of the last MB of the previous packet). Note that y is a constant over the entire transmission; A, a and $D_{\text{MB}}(0)$ are the same for all $(p_{\text{pi}}, R_{\text{si}}, r_{\text{ci}}, M_i)$ 4-tuples when encoding the current packet; b and $D_{\text{MB}}(j)$ $(j = 1, \ldots, c)$ are different for each tuple, and c may differ, as well.

The last step is for the encoder to choose among the $(p_{\text{pi}}, R_{\text{si}}, r_{\text{ci}}, M_i)$ 4-tuples so that the performance is optimized. Although the performance one would like to optimize is the PSNR of the entire transmission, this is not feasible both because we do not have access to the entire video and future channel conditions in advance, and because the overall optimization would be computationally intractable in any case. So, instead, our greedy encoding strategy minimizes the distortion-per-time unit. The encoder chooses the 4-tuple that minimizes a parameter defined as the average expected distortion-per-time unit, which is equal to the total expected distortion across the packet divided by the total transmission time of the packet, and is approximately determined by

$$\frac{\text{Packet Distortion}}{\text{Packet Transmission Time}}$$

$$\approx \frac{\left(\frac{a}{A}\right) D_{\text{MB}}(0) + \sum_{j=1}^{c-1} D_{\text{MB}}(j) + \left(\frac{B-b}{B}\right) D_{\text{MB}}(c)}{\frac{y}{R_{\text{si}}}}. \quad (17)$$

After the 4-tuple that minimizes (17) is determined, the corresponding packet will be sent out.

Note that (17) is approximate due to the terms $(a/A)D_{\text{MB}}(0)$ and $((B-b)/B)D_{\text{MB}}(c)$, which approximately represent the corresponding distortion of the partial MBs inside the current packet. Because the tails are much shorter than y, the effect of the imprecision is relatively negligible.

It is worth discussing the computational complexity of the algorithm, as well as the method to choose the set of possible p_p^*s. With the two-step approach, the complexity burden of the first step is negligible compared to both the second step and motion vector (MV) searching. In the second step, for each $(p_{\text{pi}}, R_{\text{si}}, r_{\text{ci}}, M_i)$ 4-tuple, the encoder chooses the optimal encoding mode and QP for each MB, by searching all possible mode/QP combinations (a total of $2 \times 31 = 62$ combinations). This complexity is comparable to that of the original ROPE algorithm and the DCT operations in the H.263+ codec [9], [10]. Assuming the number of possible p_p^*s is N_1, the total computational burden of the second step will be N_1 times higher than that of the algorithm with a single p_p^*. In [10], it was shown that for the same R_s, the performance gain from decreasing p_p diminishes dramatically for $p_p < 1\%$, and becomes negligible for $p_p < 0.1\%$. Also, we have found that if the SNIR falls below a value corresponding to a packet drop rate of roughly 5%, that packet drop rate increases very rapidly as the SNIR decreases further. Thus, we chose the p_p^*s in the range [0.1%, 5%]. In particular, after many simulation runs for which the chip rates varied from 2 to 30 Mcps, and E_c/N_0 varied from -18 to 0 dB, we chose the following set of packet drop rates: [0.2%, 0.6%, 1%, 1.5%, 3%]. Adding additional p_p^*s only yielded trivial performance improvement. As a result, our algorithm uses $N_1 = 5$ and, thus, has about five times higher complexity than that of the original and modified ROPE algorithms described in [9] and [10]. Using a smaller N_1 will decrease the complexity at the expense of performance degradation.

The coherence time of the channel is indicative of the correlation of channel conditions between adjacent packets. In Section II-D, we showed that the coherence time under most situations is at least the transmission time of two packets; thus, we assume a constant channel across one packet. Furthermore, in our system design and analysis, we assume that the estimation feedback is instantaneous and error-free. We also assume instantaneous encoding and zero encoder buffer delay. Under these ideal assumptions, the encoder has the correct channel conditions for each packet instantly, and the encoder does not need knowledge of the coherence time. However, in the absence of these assumptions, knowledge of the coherence time of the channel can be usefully incorporated into the system design for predicting the changing channel conditions [14].

In all our numerical results, we assume that the RAKE receiver employs all resolvable paths of the selective fading channel (i.e., $L_R = L$). Such a receiver is practical in narrowband direct-sequence (DS) systems, where the number of resolvable paths is small. However, in wideband DS systems, and especially in ultrawideband (UWB) systems, the use of less than the full number of resolvable paths (i.e., $L_R < L$), known as generalized selection combining (GSC), is typically employed. For perfect channel state information, the results when $L_R = L$ serve as an upper bound on the performance achievable when $L_R < L$. However, with noisy channel estimates, eliminating the weak paths is usually beneficial, because the larger channel estimation errors experienced by the weak paths often outweigh the enhanced diversity achievable by using them [14], [15].

As to the optimality of the proposed algorithm, *Step II* is optimal in terms of minimizing distortion, since we select the

bandwidth allocation 4-tuple from a set corresponding to the minimum distortion. *Step I*, however, has not been proved mathematically to be optimal. In our simulations, we found that once the packet drop rate was fixed, among those (R_s, r_c, M) triples which achieved that packet drop rate, the one with largest R_s invariably produced the lowest distortion. Note that this is not necessarily an optimal strategy for a variety of reasons: because the video encoder does not strictly have a convex RD curve, since adding one bit does not always decrease the distortion; because time correlation exists between dropped packets; and because the target packet drop rate may not be exactly achieved, since the choice of RCPC codes is finite. So while not provably optimal, it was found empirically that the strategy of *Step I* of choosing the largest R_s among those (R_s, r_c, M) triples which achieve each given packet drop rate was always the best choice.

This algorithm can be also extended to video communications over a tandem channel, which models both packet losses due to network congestion or buffer overflows on the wired link, and burst bit errors due to noise, fading and interference on the wireless component. Denote by p_e the packet erasure rate introduced by the wired link, and assume it is known at the encoder. To use the proposed bandwidth allocation algorithm, instead of using the packet drop rate p_p of the wireless link, we need to use the total packet drop rate due to both wired and wireless links, which is given by

$$P_R = p_e + p_p - p_e \times p_p. \tag{18}$$

VI. EFFECT OF IMPERFECT CHANNEL ESTIMATION

In this section, the bias and variance of the channel estimates are derived, and the system sensitivity to channel estimation errors will be presented numerically in the next section. We assume the channel is characterized by a flat MIP, that is $\nu = 0$, and $\sum_{l=1}^{L} \Omega_l = L$. We also assume the data bits and pilot bits are time multiplexed. Specifically, for each packet, N_p pilot bits are appended to the end of that packet. As indicated in Fig. 2, the sample mean at the output of each tap of the RAKE receiver is used as the channel estimate. That is, if $k = 0$ corresponds to the desired user, then for each resolvable path l ($1 \leq l \leq L$) and each pilot bit i ($1 \leq i \leq N_p$), $\hat{x}_{0,l}^{(i)}$ denotes the output of the sampler, and the estimate of the complex channel gain of the lth resolvable path, $\hat{c}_{0,l}$, is given by

$$\hat{c}_{0,l} = \frac{1}{N_p} \sum_{i=1}^{N_p} \frac{\hat{x}_{0,l}^{(i)}}{2E_s}. \tag{19}$$

Following the analysis of Section III-A, it can be shown that $\hat{x}_{0,l}^{(i)}$ is composed of four components, the signal $\hat{x}_s^{(i)}$, the self-interference $\hat{x}_{I_s}^{(i)}$, the multiaccess-interference $\hat{x}_{I_m}^{(i)}$, and the Gaussian noise $\hat{x}_N^{(i)}$. The signal component $\hat{x}_s^{(i)} = 2(\alpha_{0,l}e^{j\varphi_{0,l}})E_s$. Conditioned on the current channel conditions of the desired user (that is, $c_{0,l} = \alpha_{0,l}e^{j\theta_{0,l}}$, for $1 \leq l \leq L$), $\hat{x}_N^{(i)}$ is a conditional complex zero-mean Gaussian rv with variance $\sigma_N^2 = 4E_s N_0$. In the following analysis, we assume the number of users K is sufficiently large, so that $\hat{x}_{I_m}^{(i)}$ is

asymptotically a conditional zero-mean Gaussian rv with variance $\sigma_M^2 = (8E_s^2/3M)(K-1)L$, and $\hat{x}_{I_s}^{(i)}$ becomes negligible. As a result, for large K, $(\hat{x}_{0,l}^{(i)}/2E_s)$ is asymptotically a conditional complex Gaussian rv with conditional density $\hat{x}_{0,l}^{(i)} \sim N(\alpha_{0,l}e^{j\varphi_{0,l}}, \sigma_x^2)$, where

$$\sigma_x^2 \approx \frac{\sigma_N^2 + \sigma_M^2}{(2E_s)^2} = \frac{2(K-1)L}{3M} + \frac{N_0}{E_s}. \tag{20}$$

Ignoring the self-interference components, the channel estimate is given by

$$\hat{c}_{0,l} \approx \frac{1}{N_p} \sum_{i=1}^{N_p} \frac{1}{2E_s} (\hat{x}_s^{(i)} + \hat{x}_{I_m}^{(i)} + \hat{x}_N^{(i)}) \tag{21}$$

where the $\hat{x}_N^{(i)}$ are independently and identically distributed (i.i.d) conditional zero-mean Gaussian rvs, $\hat{x}_{I_m}^{(i)}$ and $\hat{x}_N^{(i)}$ are independent of each other, and $\hat{x}_{I_m}^{(i)}$ and $\hat{x}_{I_m}^{(j)}$ are independent if $|i - j| \geq 2$, because the two regions of integration will see two independent random data bits. Thus, $\hat{c}_{0,l}$, the average of a set of jointly conditional Gaussian rvs, also is a conditional complex Gaussian rv, with a conditional density given by $\hat{c}_{0,l} \sim N(\alpha_{0,l}e^{j\varphi_{0,l}}, \sigma_c^2)$, where

$$\sigma_c^2 \approx \frac{1}{4N_p^2 E_s^2} \left(\sum_{i=1}^{N_p} (\sigma_N^2 + \sigma_M^2) + \sum_{i=1}^{N_p-1} E(\hat{x}_{I_m}^{(i)} \hat{x}_{I_m}^{(i+1)}) \right). \tag{22}$$

Note that the correlation between $\hat{x}_{I_m}^{(i)}$ and $\hat{x}_{I_m}^{(i+1)}$ (i.e., $E(\hat{x}_{I_m}^{(i)} \hat{x}_{I_m}^{(i+1)})$) is not zero, but is negligible compared to σ_M^2 because of the relatively low values of the partial autocorrelations of the spreading sequences [1], so we approximate σ_c^2 as

$$\sigma_c^2 \approx \frac{1}{N_p^2} \sum_{i=1}^{N_p} \sigma_x^2 \approx \frac{1}{N_p} \left(\frac{2(K-1)L}{3M} + \frac{N_0}{E_s} \right). \tag{23}$$

In Section IV, $(\sum_{l=1}^{L} \alpha_{0,l}^2)$ was the parameter that the decoder feeds back to the encoder. Now, the decoder feeds back $F = \sum_{l=1}^{L} |\hat{c}_{0,l}|^2$. If we assume the resolvable paths are independent, the $\hat{c}_{0,l}$s are i.i.d conditional complex Gaussian rvs, and F follows a noncentral chi-square distribution with $2L$ degrees of freedom. It is shown in the Appendix that

$$E\{F\} = \left(\sum_{l=1}^{L} \alpha_{0,l}^2 \right) + L\sigma_c^2 \tag{24}$$

$$Var\{F\} = 2\sigma_c^2 \left(\sum_{l=1}^{L} \alpha_{0,l}^2 \right) + L\sigma_c^4. \tag{25}$$

As a result, F is a biased estimator of $(\sum_{l=1}^{L} \alpha_{0,l}^2)$ with a deviation

$$\Delta = F - \left(\sum_{l=1}^{L} \alpha_{0,l}^2 \right) = L\sigma_c^2 \approx \frac{L}{N_p} \left(\frac{2(K-1)L}{3M} + \frac{N_0}{E_s} \right). \tag{26}$$
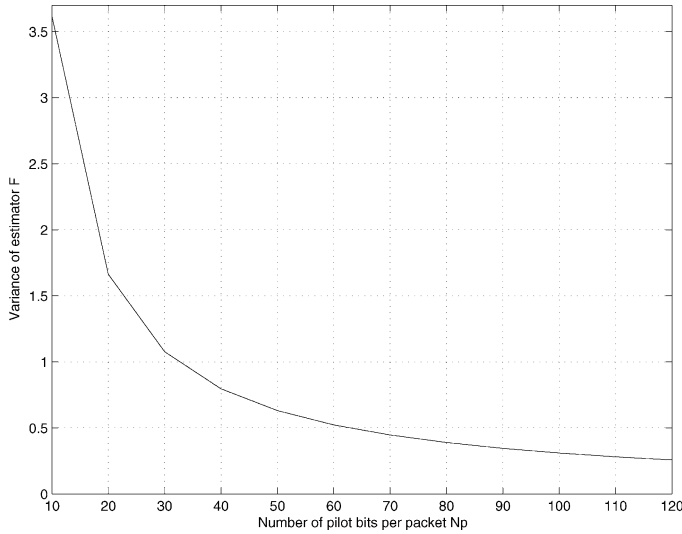
Fig. 5. Typical illustration of variance of the estimator F versus the number of pilot bits per packet N_p.

As the number of pilot bits per packet (denoted by N_p) increases, both the deviation Δ and the variance $Var\{F\}$ decrease and eventually go to zero. A typical illustration of the relationship between $Var\{F\}$ and N_p is shown in Fig. 5, under the scenario $y = 400$, $K = 20$, $L = 4$, $E_c/N_0 = -8$ dB, $r_{\mathrm{RP}} = 1/3$, and $M = 15$. If we use the estimator $\hat{F} = F - L\sigma_c^2$, it will be unbiased, but a given estimate could be negative. Since the term being estimated is nonnegative, a better estimate is $F^* = \max(\hat{F}, 0)$.

The detrimental effect of using a larger N_p is that pilot bits will occupy more bandwidth, reducing that available to data transmission and, thus, yielding a lower source rate. Comparing to (2), the relationship between r_c and r_{RP} with N_p pilot bits per packet is determined by

$$r_c = \frac{y - 16 - Z}{y + N_p} r_{\mathrm{RP}}. \tag{27}$$

The best choice of N_p is determined by a tradeoff between the estimation accuracy and the rate efficiency.

VII. SIMULATION RESULTS

The proposed system was evaluated using a modified H.263+ codec. In our simulation, we set the packet length $y = 400$ bits, the number of users $K = 20$, the number of resolvable paths $L = 4$, and the normalized Doppler spread $f_D T_c = 3 \times 10^{-6}$. We also assumed the channel was characterized by a flat MIP. The Jakes model [16] was used to generate time-correlated Rayleigh fading parameters for each independent path of each user. Standard QCIF (176×144) video sequences were used at frame rate 15 fps. The end-to-end distortion was measured by the peak signal-to-noise ratio (PSNR), which is defined as $10 \log_{10}(\mathrm{Peak}^2/\mathrm{MSE})$, where Peak is the peak value of the source and MSE is the mean-square error over all frames.

Figs. 6–8 illustrate the PSNR performance of the system versus the transmission chip rate W, which varies from 2 Mcps to 30 Mcps. In all the figures, $E_c/N_0 = -8$ dB. The figures are for the relatively high motion sequences "Carphone" and
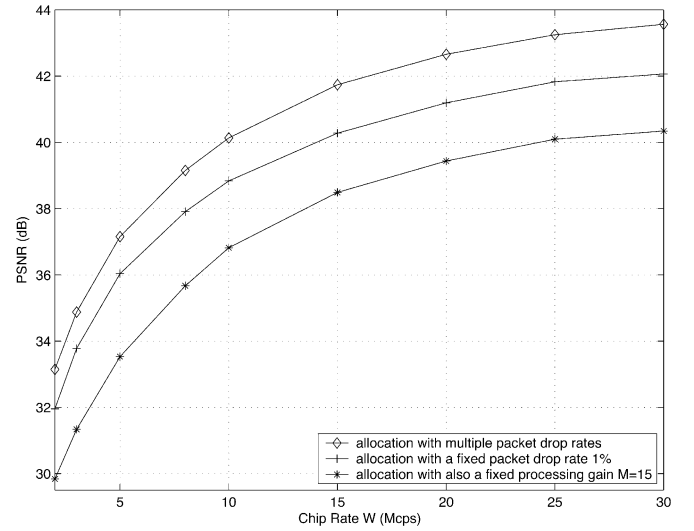

Fig. 6. PSNR performance versus target transmission chip rate, Carphone QCIF, and $E_c/N_0 = -8$ dB.
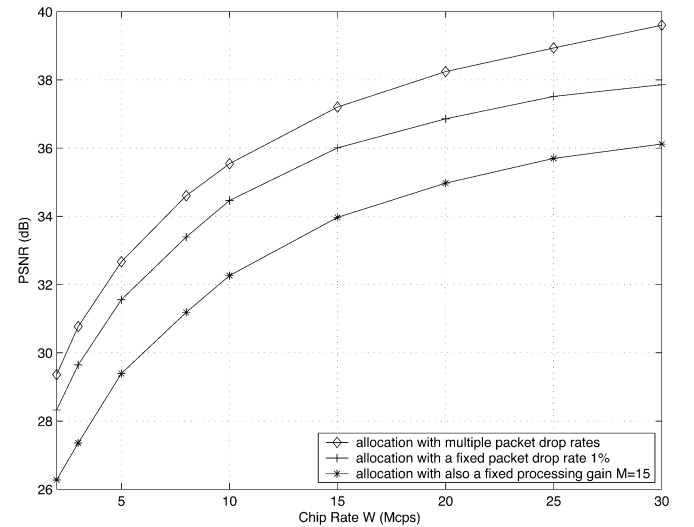

Fig. 7. PSNR performance versus target transmission chip rate, Coastguard QCIF, and $E_c/N_0 = -8$ dB.

"Coastguard," and for a very low motion sequence "Akiyo." In each figure, the performance is shown for a system using the proposed bandwidth allocation with multiple target packet error rates, a system having a bandwidth allocation with a fixed target packet error rate of 1%, and a system having a bandwidth allocation with both a fixed target packet error rate of 1% and a fixed spreading gain $M = 15$. The system with multiple packet drop rates outperforms the one with a fixed target packet error rate by about 1.0 to 1.7 dB, and outperforms the one with both fixed target packet error rate and fixed spreading gain by about 2.1 to 3.6 dB.

Fig. 9 illustrates the performance versus E_c/N_0, as E_c/N_0 varies from -18 dB to 0 dB, for "Carphone" QCIF at the transmission chip rate $W = 15$ Mcps. The system with multiple packet drop rates outperforms the one with a fixed target packet rate by about 1.4 to 1.7 dB, and outperforms the one with both fixed target packet error rate and fixed spreading gain by about
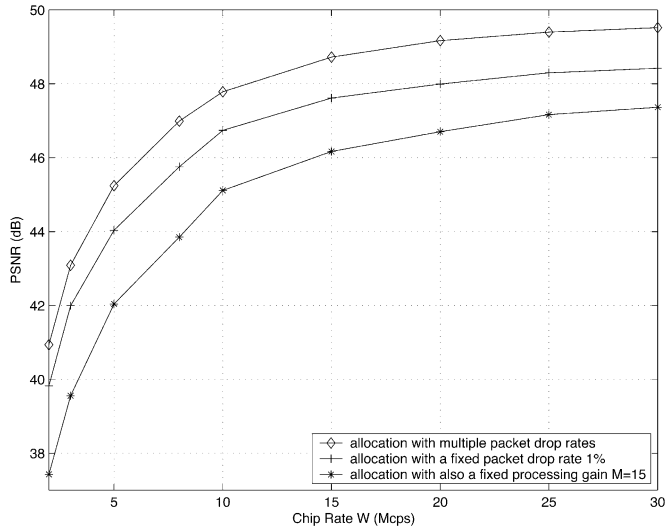
Fig. 8. PSNR performance versus target transmission chip rate, Akiyo QCIF, and $E_c/N_0 = -8$ dB.
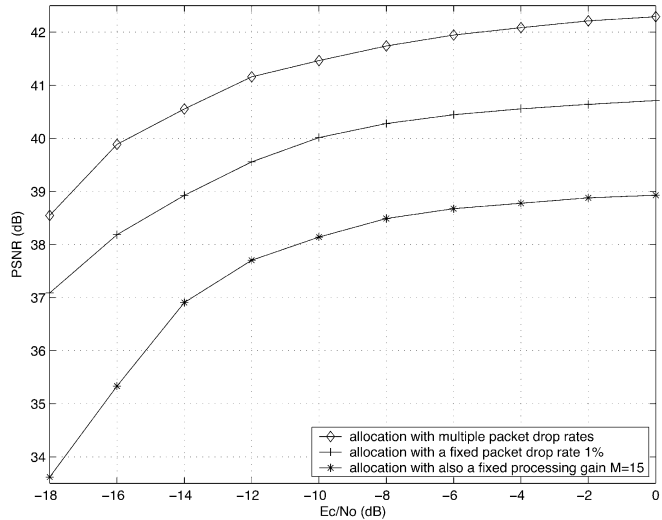

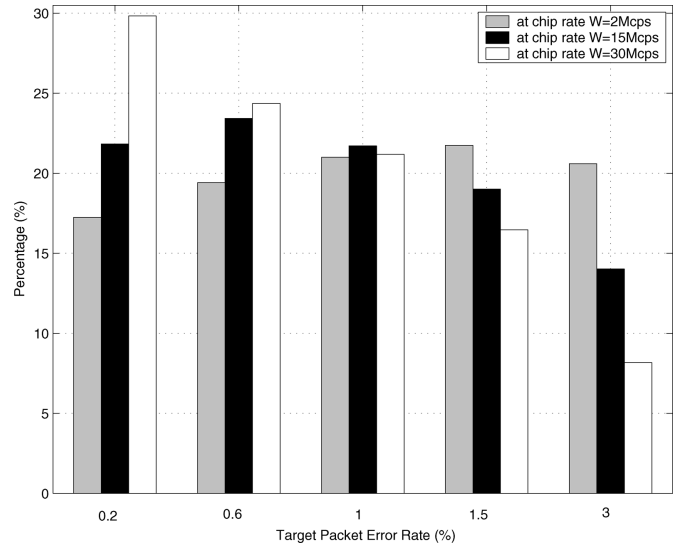Fig. 9. PSNR performance versus E_c/N_0, Carphone QCIF, and chip rate 15 Mcps.

Fig. 10. Percentage of the total packets that employ the corresponding target drop rate for Carphone QCIF at $E_c/N_0 = -8$ dB, and chip rate 2, 15, and 30 Mcps.
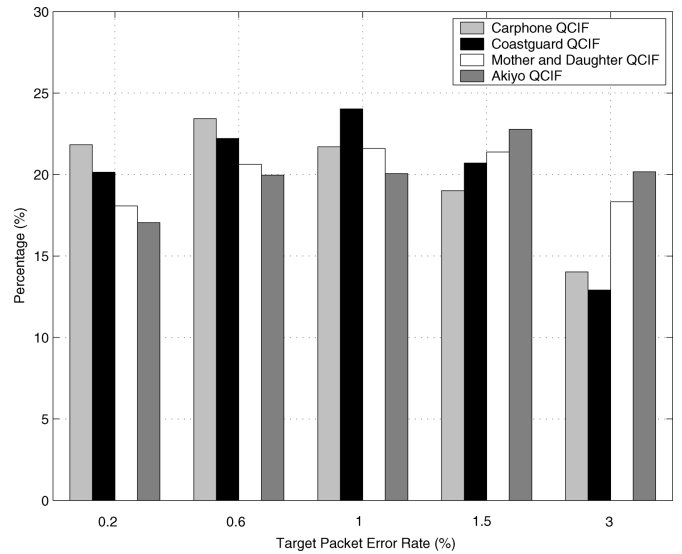

Fig. 11. Percentage of the total packets that employ the corresponding target drop rate at chip rate 15 Mcps and $E_c/N_0 = -8$ dB, for Carphone QCIF, Coastguard QCIF, and Mother and Daughter QCIF, and Akiyo QCIF.

3.2 to 4.8 dB. Also, the benefit of a higher E_c/N_0 diminishes when $E_c/N_0 \geq -8$ dB.

During the simulations, it was observed that for the same test sequence, and with other conditions the same, operating at a higher chip rate usually leads to a greater use of low target packet error rates. This is because the larger available bandwidth improves the performance by increasing both the source encoding accuracy and the error correction capability. Fig. 10 illustrates this trend, by showing the percentage of total packets that end up employing the corresponding target drop rate in the bandwidth allocations for "Carphone" QCIF with $E_c/N_0 = -8$ dB, operating at chip rate equal to 2, 15, and 30 Mcps. Similarly, under the same scenario, increasing E_c/N_0 also leads to a greater use of low target packet error rates.

In Fig. 11, histograms for the target drop rate are shown for two high motion sequences "Carphone" and "Coastguard," and two low motion sequences "Mother and Daughter" and "Akiyo,"

all at chip rate 15 Mcps and $E_c/N_0 = -8$ dB. It is seen that, for this particular scenario, the low motion sequences use relatively more high packet error rates, and the high motion sequences use relatively more low packet error rates. This may be because low motion sequences are more tolerant of packet errors due to the ease of concealment.

We also compare our system with a recent system [17] which did the bandwidth allocation for scalable video transmission over single or dual DS-CDMA channels with universal rate-distortion characteristics (URDC) [6], [17]. In [17], bandwidth allocations were chosen from a finite set of possible source bit rates, a finite set of possible coding rates, and a finite set of possible spreading lengths. This was done once for the whole transmission period. We compare the performance of our system with the results given in [17, Fig. 1], where the comparison system is
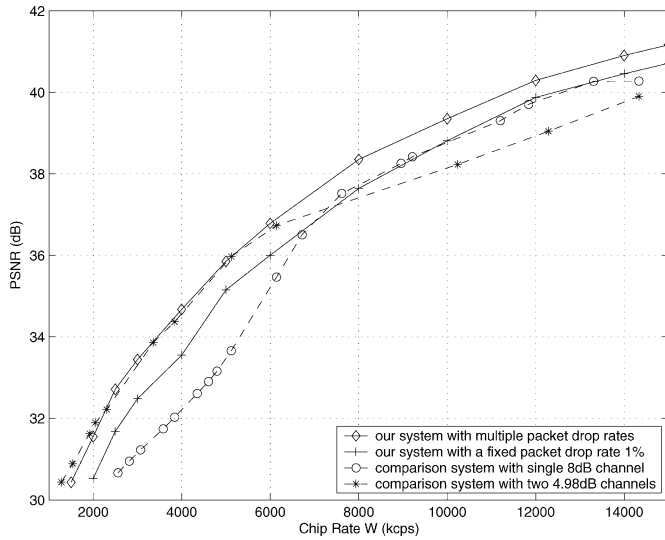
Fig. 12. PSNR performance versus target transmission chip rate for our system and comparison system, Foreman QCIF, frame rate 10 fps.
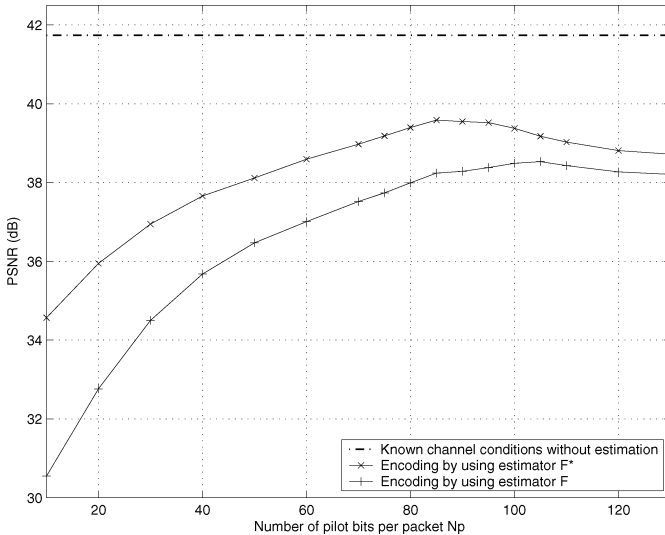

Fig. 13. PSNR performance versus the number of pilot bits per packet N_p, for "Carphone" QCIF at chip rate 15 Mcps and $E_c/N_0 = -8$ dB, with the proposed bandwidth allocation algorithm using the channel estimators F and F^*. The reference line is the system performance with known channel conditions.

operated over block fading channels with the number of resolvable paths $L = 3$ and 8 interferers. To keep the same total transmission power, a single 8-dB channel and two 4.98-dB channels were compared. The simulation results were for "Foreman" at 10 fps, and the performance was measured by MSE (which we converted to PSNR). As shown in Fig. 12, our system with multiple packet drop rates outperforms the comparison systems of both single and dual channels over most chip rates.

Finally, the system performance with imperfect channel estimation is presented. In Fig. 13, PSNR versus the number of pilot bits per packet, N_p, is shown, for "Carphone" QCIF at chip rate 15 Mcps and $E_c/N_0 = -8$ dB, with the proposed algorithm using the channel estimators F and F^*. The top dashed line is an upper bound on performance, under the assumption that a genie informs the transmitter of the actual channel conditions

($N_p = 0$). It is seen that the estimator F^* outperforms the biased estimator F, especially when N_p is small. The optimal choices of N_p under this scenario are 85 and 105 for the algorithm using F^* and F, respectively, which accounts for about 17.5% and 20.8% of the total packet length, respectively (the number of information bits is $y = 400$ per packet). It is also seen that the performance degradation at the optimal N_p is about 2.15 dB for the system using F^*, and about 3.2 dB for the system using F. This suggests that the simple pilot averaging technique for channel estimation is not very efficient, and more elaborate techniques (e.g., Wiener filtering) may be needed.

VIII. CONCLUSION

In this paper, we proposed a robust video transmission scheme with efficient bandwidth allocation among source coding, channel coding, and spreading, that operates at the packet level. The algorithm is proposed for use in a CDMA network, and can be extended to operate over a tandem channel with both packet erasures and burst bit errors. The optimality problem among three parameters is solved using a two-step tradeoff strategy, by introducing the target packet drop rate at the decoder.

Based on the statistics of the received signal, and a theoretical bound on the packet error rate due to the wireless link, p_p, we found that, conditioned on appropriate system parameters and channel conditions, p_p is fully determined by the channel coding rate, r_c, and spreading gain, M. Furthermore, the algorithms to trade off channel coding and spreading for a given packet drop rate p_p^* were derived for systems both with fixed and with variable M. It was also shown that for a system with variable r_c and M, in a scenario where the number of users is large so that the self-interference is negligible, the transmitter should use the lowest rate channel code and only vary M to adapt to different channel conditions.

As to the video source encoding, we used the modified-ROPE algorithm [9], [10], and extended the rate control scheme to operate on chips, rather than bits, so that the effect of variable spreading gain could be accounted for. Likewise, we extended the rate-distortion framework for optimal mode selection (and QP selection) to operate at the chip rate, rather than the bit rate.

We then developed an algorithm with a two-step tradeoff for choosing the bandwidth allocation parameters. The first step was to trade off between channel coding and spreading, given a target packet drop rate, such that we maximized the source bit rate. The second step was to trade off between the packet drop rate and the source bit rate, based on the channel coding and spreading determined by the first step, such that the overall expected distortion per time unit was minimized. This optimization process was done for each packet, and incorporated the effects of both the changing channel characteristics and the current video content. The effect of imperfect channel estimation on system performance was also studied.

Results were simulated for a variety of sequences. It was shown that the proposed system which allows all components to vary offered about a 1.4-dB gain over a scheme using a fixed packet drop rate, and up to 4-dB gain over a scheme using a fixed spreading gain. The scheme also outperformed a comparable system in the literature [17] that adapted source coding

rate, channel coding rate, and spreading gain, but which did not operate on a packet basis.

APPENDIX

In this Appendix, we derive the conditional mean and variance of the term $F = \sum_{l=1}^{L} |\hat{c}_{0,l}|^2$, where the $\hat{c}_{0,l}$s are i.i.d conditional complex Gaussian rvs, with conditional mean $E\{\hat{c}_{0,l}\} = c_{0,l} = \alpha_{0,l} e^{j\varphi_{0,l}}$ and variance σ_c^2, as given by (23). Thus, F has a conditional noncentral chi-square distribution with $2L$ degrees of freedom. The conditional mean of F is given by

$$
\begin{aligned}
E\{F\} &= \sum_{l=1}^{L} E\{|\hat{c}_{0,l}|^2\} = \sum_{l=1}^{L} \left((E\{|\hat{c}_{0,l}|\})^2 + \sigma_c^2 \right) \\
&= \left(\sum_{l=1}^{L} \alpha_{0,l}^2 \right) + L\sigma_c^2
\end{aligned} \tag{28}
$$

and the conditional variance is

$$
\begin{aligned}
Var\{F\} &= E\{F^2\} - (E\{F\})^2 \\
&= \sum_{l=1}^{L} E\{|\hat{c}_{0,l}|^4\} + \sum_{l_1=1}^{L} \sum_{l_2=1, l_2 \neq l_1}^{L} \\
&\quad \left\{ E\{|\hat{c}_{0,l_1}|^2 |\hat{c}_{0,l_2}|^2\} - \left(\sum_{l=1}^{L} E\{|\hat{c}_{0,l}|^2\} \right)^2 \right\}.
\end{aligned} \tag{29}
$$

Each $\hat{c}_{0,l}$ can be expressed as $\hat{c}_{0,l} = \hat{c}_{0,l}^Q + j\hat{c}_{0,l}^I$, where $\hat{c}_{0,l}^Q \sim N(\alpha_{0,l}\cos(\varphi_{0,l}), \sigma_c^2/2)$, and $\hat{c}_{0,l}^I \sim N(\alpha_{0,l}\sin(\varphi_{0,l}), \sigma_c^2/2)$. Also, $\hat{c}_{0,l}^Q$ and $\hat{c}_{0,l}^I$ are conditionally independent, because, from (21), $\hat{c}_{0,l}$ is the average of three components, among which $\hat{x}_s$ is a real number, $\hat{x}_N$ is conditionally a symmetric complex Gaussian rv, and $\hat{x}_{I_m}$ is an asymptotically, conditionally symmetric, complex Gaussian rv.

It is straightforward to show that for any Gaussian rv $x \sim (\mu, \sigma^2)$, $E\{x^3\} = \mu^3 + 3\mu\sigma^2$, and $E\{x^4\} = \mu^4 + 6\mu^2\sigma^2 + 3\sigma^4 = (E\{x^2\})^2 + 4\mu^2\sigma^2 + 2\sigma^4$. Also, $E\{(\hat{c}_{0,l}^Q)^2\} + E\{(\hat{c}_{0,l}^I)^2\} = \alpha_{0,l}^2 + \sigma_c^2 = E\{|\hat{c}_{0,l_1}|^2\}$. Thus, $E\{|\hat{c}_{0,l}|^4\}$ is given by

$$
\begin{aligned}
E\{|\hat{c}_{0,l}|^4\} &= E\left\{ \left((\hat{c}_{0,l}^Q)^2 + (\hat{c}_{0,l}^I)^2 \right)^2 \right\} \\
&= E\{(\hat{c}_{0,l}^Q)^4\} + E\{(\hat{c}_{0,l}^I)^4\} + 2E\{(\hat{c}_{0,l}^Q)^2 (\hat{c}_{0,l}^I)^2\} \\
&= (E\{(\hat{c}_{0,l}^Q)^2\})^2 + 2\sigma_c^2\alpha_{0,l}^2 \cos^2(\varphi_{0,l}) + \frac{\sigma_c^4}{2} \\
&\quad + (E\{(\hat{c}_{0,l}^I)^2\})^2 + 2\sigma_c^2\alpha_{0,l}^2 \sin^2(\varphi_{0,l}) \\
&\quad + \frac{\sigma_c^4}{2} + 2E\{(\hat{c}_{0,l}^Q)^2\}E\{(\hat{c}_{0,l}^I)^2\} \\
&= 2\sigma_c^2\alpha_{0,l}^2 + \sigma_c^4 + \left(E\{(\hat{c}_{0,l}^Q)^2\} + E\{(\hat{c}_{0,l}^I)^2\} \right)^2 \\
&= 2\sigma_c^2\alpha_{0,l}^2 + \sigma_c^4 + (E\{|\hat{c}_{0,l_1}|^2\})^2.
\end{aligned} \tag{30}
$$

Plugging (30) into (29), we get

$$
\begin{aligned}
Var\{F\} &= \sum_{l=1}^{L} \left(2\sigma_c^2\alpha_{0,l}^2 + \sigma_c^4 \right) + \sum_{l=1}^{L} (E\{|\hat{c}_{0,l_1}|^2\})^2 \\
&\quad + \sum_{l_1=1}^{L} \sum_{l_2=1, l_2 \neq l_1}^{L} E\{|\hat{c}_{0,l_1}|^2 |\hat{c}_{0,l_2}|^2\} \\
&\quad - \left(\sum_{l=1}^{L} E\{|\hat{c}_{0,l}|^2\} \right)^2 \\
&= 2\sigma_c^2 \left(\sum_{l=1}^{L} \alpha_{0,l}^2 \right) + L\sigma_c^4.
\end{aligned} \tag{31}
$$

REFERENCES

[1] T. Eng and L. B. Milstein, "Coherent DS-CDMA performance in Nakagami multipath fading," *IEEE Trans. Commun.*, vol. 43, no. 1, pp. 1134–1143, Feb./Mar./Apr. 1995.

[2] Q. Zhao, P. C. Cosman, and L. B. Milstein, "Tradeoffs of source coding, channel coding and spreading in frequency selective Rayleigh fading channels," *J. VLSI Signal Process.*, vol. 30, pp. 7–20, 2002.

[3] ——, "Optimal allocation of bandwidth for source coding, channel coding and spreading in CDMA systems," *IEEE Trans. Commun.*, vol. 52, no. 10, pp. 1797–1808, Oct. 2004.

[4] R. Annavajjala, P. C. Cosman, and L. B. Milstein, "On source coding, channel coding, and spreading tradeoffs in a DS-CDMA system operating over frequency selective fading channels with narrow-band interference," *IEEE J. Sel. Areas Commun.*, vol. 23, no. 5, pp. 1034–1044, May 2005.

[5] D. G. Sachs, R. Anand, and K. Ramchandran, "Wireless image transmission using multiple-description based concatenated codes," *Proc. SPIE*, Jan. 2000.

[6] L. P. Kondi, S. N. Batalama, D. A. Pados, and A. K. Katsaggelos, "Joint source/channel coding for scalable video over DS-CDMA multipath fading channels," in *IEEE Int. Conf. Image Processing*, Oct. 2001, pp. 994–997.

[7] V. Stankovic, R. Hamzaoui, and Z. Xiong, "Real-time error protection of embedded codes for packet erasure and fading channels," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 14, no. 8, pp. 1064–1072, Aug. 2004.

[8] S. Zhao, Z. Xiong, and X. Wang, "Optimal resource allocation for wireless video over CDMA networks," *IEEE Trans. Mobile Comput.*, vol. 4, no. 1, pp. 56–67, Jan./Feb. 2005.

[9] R. Zhang, S. L. Regunathan, and K. Rose, "Video coding with optimal inter/intra-mode switching for packet loss resilience," *IEEE J. Sel. Areas Commun.*, vol. 18, no. 6, pp. 966–976, Jun. 2000.

[10] Y. Shen, P. C. Cosman, and L. B. Milstein, "Video coding with fixed length packetization for a tandem channel," *IEEE Trans. Image Process.*, vol. 15, no. 2, pp. 273–288, Feb. 2006.

[11] M. K. Simon and M-S Alouini, *Digital Communications Over Fading Channels: A Unified Approach to Performance Analysis*. New York: Wiley, 2004.

[12] Y. Shen, P. C. Cosman, and L. B. Milstein, "Weight distribution of a class of binary linear block codes formed from RCPC codes," *IEEE Commun. Lett.*, vol. 9, no. 5, pp. 811–813, Sep. 2005.

[13] P. G. Sherwood and K. Zeger, "Progressive image coding on noisy channels," *IEEE Signal Process. Lett.*, vol. 4, no. 7, pp. 189–191, Jul. 1997.

[14] L. L. Chong and L. B. Milstein, "The effects of channel-estimation errors on a space-time spreading CDMA system with dual transmit and dual receive diversity," *IEEE Trans. Commun.*, vol. 52, no. 7, pp. 1145–1151, Jul. 2004.

[15] T. Eng, N. Kong, and L. B. Milstein, "Comparison of diversity combining techniques for Rayleigh fading channels," *IEEE Trans. Commun.*, vol. 44, no. 9, pp. 1117–1129, Sep. 1996.

[16] P. Dent, G. E. Bottomley, and T. Croft, "Jakes' model revisited," *Electron. Lett.*, vol. 29, no. 13, pp. 1162–1163, 1993.

[17] D. Srinivasan, L. P. Kondi, and D. A. Pados, "Scalable video transmission over wireless DS-CDMA channels using minimum TSC spreading codes," *IEEE Signal Process. Lett.*, vol. 11, no. 10, pp. 836–840, Oct. 2004.

Yushi Shen received the B.S. degree in electrical engineering from Tsinghua University, Beijing, China, in 2001, and the M.S. and Ph.D. degrees in electrical and computer engineering from the University of California at San Diego (UCSD), La Jolla, in 2003 and 2006, respectively.

He is currently with Microsoft, Inc., Seattle, WA. His research interests are in the area of wireless video communications, joint source/channel coding, communication and information theory, and video/multimedia processing.

Mr. Shen is the recipient of a UCSD Electrical and Computer Engineering Departmental Fellowship (2001 to 2002), a California Institute for Telecommunications and Information Technology (Cal-IT2) Fellowship (2003 to 2005), and the Chinese Government Award for Outstanding Students Abroad (2006).

Pamela C. Cosman (S'88–M'93–SM'00) received the B.S. degree (with honors) in electrical engineering from the California Institute of Technology, Pasadena, in 1987, and the M.S. and Ph.D. degrees in electrical engineering from Stanford University, Stanford, CA, in 1989 and 1993, respectively.

She was a National Science Foundation (NSF) Postdoctoral Fellow at Stanford University and a Visiting Professor at the University of Minnesota, Minneapolis, from 1993 to 1995. Since July 1995, she has been with the faculty of the Department of Electrical and Computer Engineering, University of California at San Diego, La Jolla, where she is currently a Professor and Director of the Center for Wireless Communications. Her research interests are in the areas of image and video compression and processing.

Dr. Cosman is the recipient of the ECE Departmental Graduate Teaching Award (1996), a Career Award from the NSF (1996 to 1999), and a Powell Faculty Fellowship (1997 to 1998). She was a Guest Editor of the June 2000 special issue of the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS on "Error-resilient image and video coding," and was the Technical Program Chair of the 1998 Information Theory Workshop in San Diego. She was an Associate Editor of the IEEE COMMUNICATIONS LETTERS (1998 to 2001), and an Associate Editor of the IEEE SIGNAL PROCESSING LETTERS (2001–2005). She was a Senior Editor (2003–2005) and is now the Editor-in-Chief of the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS. She is a member of Tau Beta Pi and Sigma Xi. Her web page address is http://www.code.ucsd.edu/cosman/.

Laurence B. Milstein (S'66–M'68–SM'75–F'85) received the B.E.E. degree from the City College of New York, New York, in 1964, and the M.S. and Ph.D. degrees in electrical engineering from the Polytechnic Institute of Brooklyn, Brooklyn, NY, in 1966 and 1968, respectively.

From 1968 to 1974, he was with the Space and Communication Group of Hughes Aircraft Company, and from 1974 to 1976, he was a Member of the Department of Electrical and Systems Engineering, Rensselaer Polytechnic Institute, Troy, NY. Since 1976, he has been with the Department of Electrical and Computer Engineering, University of California at San Diego (UCSD), La Jolla, where he is a Professor and former Department Chairman, working in the area of digital communication theory, with special emphasis on spread-spectrum communication systems. He has also been a Consultant to both government and industry in the areas of radar and communications.

Dr. Milstein was an Associate Editor for communications Theory for the IEEE TRANSACTIONS ON COMMUNICATIONS, an Associate Editor for Book Reviews for the IEEE TRANSACTIONS ON INFORMATION THEORY, an Associate Technical Editor for the IEEE Communications Magazine, and the Editor-in-Chief of the IEEE JOURNAL ON SELECTED AREA IN COMMUNICATIONS. He was the Vice President for Technical Affairs in 1990 and 1991 of the IEEE Communications Society, has been a member of the Board of Governors of both the IEEE Communications Society and the IEEE Information Theory Society, and is a former Chair of the IEEE Fellows Selection Committee. He is a recipient of the 1998 Military Communications Conference Long-Term Technical Achievement Award, an Academic Senate 1999 UCSD Distinguished Teaching Award, an IEEE Third Millenium Medal in 2000, the 2000 IEEE Communication Society Armstrong Technical Achievement Award, the 2002 MILCOM Fred Ellersick Award, and the Ericsson Endowed Chair in Wireless Communication Access Techniques in 2005.