Title
Motion-Compensated Scalable Video Transmission Over MIMO Wireless Channels

Journal
IEEE Transactions on Circuits and Systems for Video Technology, 23(1)

ISSN
1051-8215 1558-2205

Authors
Kim, Hobin
Cosman, Pamela C
Milstein, Laurence B

Publication Date
2013

DOI
10.1109/TCSVT.2012.2203202

Peer reviewed

Motion-Compensated Scalable Video Transmission Over MIMO Wireless Channels

Hobin Kim, Pamela C. Cosman, *Fellow, IEEE,* and Laurence B. Milstein, *Fellow, IEEE*

***Abstract*—We study motion compensated fine granular scalable (MC-FGS) video transmission over multiinput multioutput (MIMO) wireless channels applicable to video streaming, where leaky and partial prediction schemes are applied in the enhancement layer of MC-FGS to exploit the tradeoff between error propagation and coding efficiency. For reliable transmission, we propose unequal error protection (UEP) by considering a tradeoff between reliability and data rates, which are controlled by forward error correction and MIMO mode selection to minimize the average distortion. In a high Doppler environment, where it is hard to get an accurate channel estimate, we investigate the performance of the proposed MC-FGS video transmission scheme with joint control of both the leaky and partial prediction parameters, and the UEP. In a slow fading channel, where the channel throughput can be estimated at the transmitter, adaptive control of prediction parameters is considered.**

***Index Terms*—Error resilience, motion compensation, scalable video transmission, source-channel adaptation.**

I. Introduction

RECENTLY, the transmission of multimedia over wireless channels has been in high demand. However, due to the high error probability and fluctuating channel bandwidth in a high Doppler environment, it becomes challenging to maintain the quality of service when a multimedia stream is transmitted over a wireless channel. Fine granularity scalable (FGS) video coding is suitable for mobile users with variable channel bandwidth, since it makes decoding possible even in the case of partial loss of the bitstream, where the FGS bitstream is encoded in a progressive manner. For example, in MPEG-4 FGS video coding, the base layer (BL) contains the most important information, such as coding modes and motion vectors. A scalable enhancement layer (EL) is generated using bit-plane coding on the DCT coefficients. When the BL is transmitted reliably, the scalable EL bitstream can be decoded, even though it is truncated at any point.

In motion compensated prediction (MCP) of conventional MPEG-4 FGS coding, the current BL and EL are only predicted from the BL of the previous frame. By excluding the EL from the MCP loop, MPEG-4 FGS coding can avoid error propagation which can be caused by the corruption of the EL. However, this can decrease the coding efficiency due to the use of a low quality reference frame. To enhance the compression efficiency, motion compensated fine granularity scalable (MC-FGS) coding was proposed in [1]. In this video coding scheme, a high quality reference is generated using the EL as well as the BL, which allows the system to achieve a high coding efficiency. However, the loss of the EL can result in severe error propagation, since there can be a mismatch (drift) between the reconstructed references at the encoder and the decoder.

In [2], progressive FGS (PFGS) was introduced to improve the coding efficiency and alleviate error propagation simultaneously. For higher coding efficiency, PFGS uses a separate prediction loop that contains a high quality reference frame in the encoding of the EL video. In order to address the drift problem, PFGS keeps a prediction path from the BL to the least significant bitplanes of the EL across several frames to ensure graceful recovery from errors over a few frames. Robust FGS (RFGS) [3] uses a different approach to control the tradeoff between the coding efficiency and the error propagation, where the two distinct parameters of leaky and partial prediction are used jointly. UEP can also be used to enhance the error resilience of MC-FGS by reducing the loss probability of the EL to be used for the reconstruction of the reference. As in [4], different numbers of parity bits are allocated to the packets of the EL according to their impact on average distortion.

To enable reliable transmission of video over wireless channel, MIMO has been widely considered. In [5], the authors studied joint source-channel coding for video transmission over a MIMO channel in the context of a diversity-multiplexing tradeoff for the MIMO and compression-error resilience tradeoff for the video coding. Another approach for video transmission over MIMO with loss-aware rate distortion algorithm has introduced in [6]. In [7], Holliday proposed a framework for optimizing the tradeoff between diversity, multiplexing and delay in the transmission of video over a MIMO channel.

In this paper, we study the transmission of an MC-FGS bitstream over a MIMO channel with joint control of the UEP and the prediction parameters. Specifically, we propose a

Manuscript received May 25, 2011; revised October 12, 2011, December 31, 2011; accepted March 27, 2012, April 19, 2012. Date of publication June 6, 2012; date of current version January 9, 2013. This work was supported in part by the Center for Wireless Communications of UCSD, the National Science Foundation under Grant CCF-0915727, LG Electronics, Seoul, South Korea, and the UC Discovery Grant Program. This paper was partially presented at the IEEE Global Communications Conference (GLOBECOM'09). This paper was recommended by Associate Editor D. O. Wu.

H. Kim is with Qualcomm Inc., San Diego, CA 92121 USA (e-mail: mahabin@gmail.com).

P. C. Cosman and L. B. Milstein are with the Department of Electrical and Computer Engineering, University of California, San Diego, La Jolla, CA 92093 USA (e-mail: pcosman@ece.ucsd.edu; milstein@ece.ucsd.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TCSVT.2012.2203202

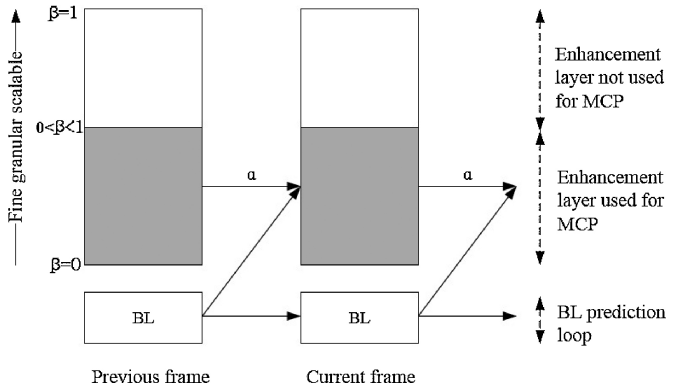


Fig. 1. Motion compensated FGS coder with leaky and partial prediction [4].

UEP policy consisting of FEC and MIMO mode selection to exploit the fundamental tradeoff between spatial diversity and multiplexing. In the MIMO mode selection, we consider spatial multiplexing to provide high data rates, spatial diversity which guarantees high reliability, and a hybrid mode based on double space time transmit diversity [8]. Note that finer control of the multiplexing-diversity tradeoff, such as diversity embedding space-time coding [9], can be also applied. Originally, the idea of combining cooperative diversity gain with UEP for a progressive image bitstream was proposed by Kwasinski in [10], where additional diversity was applied to high priority packets. We extend this to a MIMO system which can provide higher diversity orders. That is, for each packet, we choose an appropriate MIMO mode and FEC code rate to minimize average distortion.

The rest of this paper is organized as follows. In Section II, the source and channel models are described. The UEP scheme, based on a tradeoff between reliability and data rate, is proposed in Section III. In Section IV, we provide simulation results and a discussion regarding the selection of the prediction parameters based on the instantaneous channel information which can be fed back to the transmitter. Finally, in Section V, conclusions are presented.

II. System Model

A. Source Model

We consider an MC-FGS video coder employing leaky and partial prediction, as introduced in [3] and [4]. In conventional MC-FGS video coding, both the BL and EL are used to reconstruct a high-quality reference. However, this can result in error propagation. To compensate for the error propagation, RFGS introduced leaky prediction as presented in Fig. 1, where the EL is scaled by the leaky prediction parameter, $\alpha \in [0, 1]$, before it is incorporated into the MCP loop. That is, at the encoder, the reference for the prediction of the current EL at time n, $\hat{F}_n^{\mathrm{EL,enc}}$, is a weighted sum of the previous BL, F_{n-1}^{BL} and EL, F_{n-1}^{EL}, i.e., [4]

$$\hat{F}_n^{\mathrm{EL,enc}} = (1-\alpha)F_{n-1}^{\mathrm{BL}} + \alpha F_{n-1}^{\mathrm{EL}}. \tag{1}$$

Therefore, if the leaky prediction parameter is set to 0, the scheme becomes the MPEG-4 FGS video coding, where the EL is entirely excluded from the MCP loop. In contrast, if the leaky prediction parameter is fixed at unity, then it works as MC-FGS video coding. However, by choosing the parameter less than unity, the effect of error propagation can be reduced at the price of coding efficiency.

In partial prediction, the encoder designates the amount of the EL to be used for the reconstruction of the reference frame in the MCP loop. For example, the encoder could designate the number of bitplanes to be used. By including more bitplanes of the EL into the MCP loop, better coding efficiency can be achieved. However, if the instantaneous channel bandwidth cannot support the number of bitplanes used in the MCP loop, then it can result in error propagation. Therefore, the partial prediction parameter needs to be chosen based on knowledge of the channel bandwidth. In this paper, we allow an arbitrary number of symbols in the EL bitstream to be used in the MCP loop. Because the EL of MC-FGS coding can be truncated at any point, we do not need to operate with the coarse granularity of whole bitplanes, but can operate with the fine granularity of individual symbols. We define the partial prediction parameter, β, to be the ratio of the number of EL symbols used for MCP and the maximum number of EL symbols for that frame. Then, (1) can be represented as

$$\hat{F}_n^{\mathrm{EL,enc},\beta} = (1-\alpha)F_{n-1}^{\mathrm{BL}} + \alpha F_{n-1}^{\mathrm{EL},\beta} \tag{2}$$

where $\hat{F}_n^{\mathrm{EL,enc},\beta}$ is the reference for the prediction of the current EL with partial prediction, and $F_{n-1}^{\mathrm{EL},\beta}$ is the partially reconstructed EL of the previous frame at the encoder.

B. Channel Model

In this paper, we assume an $M_r \times M_t$ wireless MIMO channel, where M_r and M_t represent the number of receive and transmit antennas, respectively. Assuming perfect matched filter reception, we can use the following discrete-time channel model:

$$r[l] = H[l]s[l] + n[l] \tag{3}$$

where $r[l]$ represents $M_r \times 1$ received signal vector, $s[l]$ is $M_t \times 1$ transmitted signal vector, and $n[l]$ is $M_r \times 1$ independent and identically distributed (i.i.d.) additive Gaussian noise vector for the lth symbol. $H[l]$ is the $M_r \times M_t$ MIMO channel coefficients matrix, whose elements are i.i.d. zero-mean complex Gaussian random variables with variance σ_h^2. We consider a high Doppler environment, which results in rapidly time-varying channels. Given the Doppler spread, to model the channel estimation error, we consider the following system:

1) pilot symbol assisted modulation (PSAM); [11]
2) orthogonal pilot symbols for each transmit antenna;
3) channel estimation using the K nearest pilot samples in conjunction with a Wiener filter.

If we denote the estimation error of the channel from the jth transmit antenna to the ith receive antenna by $\varepsilon_{ij}[l]$, then it can be modeled as a complex Gaussian random variable, and its variance, $\sigma_\varepsilon^2[l]$, can be expressed as [11]

$$\sigma_\varepsilon^2[l] = 1 - w^\dagger[l]R^{-1}w[l] \tag{4}$$

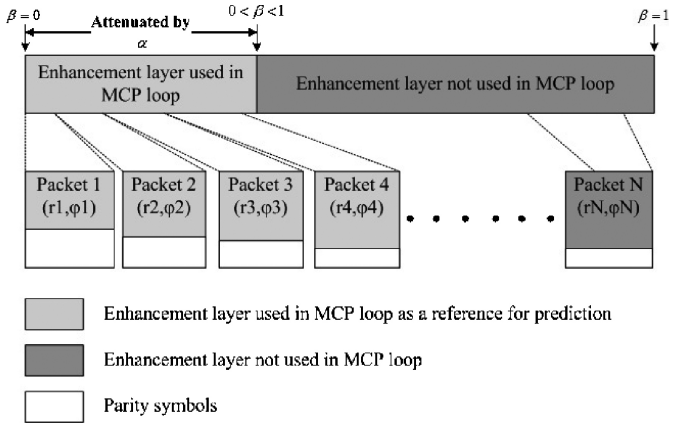


Fig. 2. Packetization of the MC-FGS enhancement layer.

where R represents the $K \times K$ autocorrelation matrix of the nearest K received pilot samples, $w[l]$ is the $K \times 1$ cross-covariance vector between the received pilot samples and $h_{ij}[l]$, and we assume $\sigma_h^2 = 1$. Note that $w[l]$ and R are dependent on the pilot signal-to-noise ratio (SNR), Doppler spread, and pilot spacing. In this paper, we assume that the pilot SNR and its spacing are selected to be equal to the data symbol SNR and the channel coherence time, respectively.

Following the assumption that all MIMO channels are independent, channel estimation errors in the MIMO system can be modeled as a matrix consisting of i.i.d. complex Gaussian random variables with the variance of (4). To maximize the signal-to-interference-and-noise ratio (SINR) under the condition of imperfect channel estimation, we consider the modified minimum-mean-squared-error (MMSE) detection scheme proposed in [12]. We assume BPSK modulation, and that N packets, whose size is fixed at m symbols, are allocated to a frame of the FGS enhancement layer bitstream, as presented in Fig. 2. Note that each packet is protected by FEC. For the transmission of the packet, we choose either spatial diversity, spatial multiplexing or a hybrid of the two. We have a total transmit power constraint to make the system consume a constant power for any selection of MIMO modes. The specific MIMO mode (multiplexing, diversity, hybrid) is chosen at the transmitter on a packet by packet basis in order to minimize distortion. The selection of UEP and MIMO mode per packet is performed at the transmitter because the source statistics and encoder characteristics are known and are based on the average channel gain. The following three MIMO modes described below are considered in this paper.

1) *Spatial Diversity:* Spatial diversity is achieved by transmitting and receiving symbol streams with the same information content through multiple transmit and receive antennas. There are various ways to implement spatial diversity, but quasiorthogonal space time block coding (QOSTBC) [13] is considered in this paper.

2) *Spatial Multiplexing:* Spatial multiplexing makes it possible to increase the transmission rate proportional to $min(M_t, M_r)$ without allocating additional bandwidth or transmit power [14]. Spatial multiplexing can be achieved by transmitting different data streams through multiple transmit antennas over independently fading channels. However, under the constraint of the fixed total transmit power, spatial multiplexing assigns less energy per bit compared to spatial diversity.

3) *Hybrid Mode:* We consider the hybrid mode to have twice the data rate and half the diversity order, as compared to the spatial diversity mode. For the purpose of simulation, we implement Double Space Time Transmit Diversity (D-STTD), as proposed in [8].

III. Unequal Error Protection with FEC and MIMO Mode Selection

In this section, we investigate how to find a UEP policy for minimizing the average distortion by choosing the FEC code rate and the appropriate MIMO mode in a fast varying channel. We assume that N fixed size packets are available for the transmission of the FGS enhancement layer in a frame, where the packet length is equal to the duration of m symbols. For the ith packet, the channel code rate and MIMO mode are denoted r_i and ϕ_i, respectively, where r_i is chosen from the possible channel code rate set and ϕ_i represents one of the MIMO modes. If we denote the packet error rate of the ith packet by $p_i(r_i, \phi_i)$, then the probability that the first k packets are received successfully and the first packet error happens at packet $(k+1)$ is

$$\pi_k \triangleq \Pr(\text{first packet error occurs at packet } k{+}1) = \begin{cases} p_1(r_1, \phi_1), & k{=}0 \\ \prod_{j=1}^{k}(1 - p_j(r_j, \phi_j))p_{k+1}(r_{k+1}, \phi_{k+1}), & 0 < k < N \\ \prod_{j=1}^{N}(1 - p_j(r_j, \phi_j)), & k{=}N. \end{cases} \quad (5)$$

When the first k packets are successfully received, the number of information bits available at the receiver is $\sum_{j=1}^{k} r_j M_j$, where M_j represents the total number of symbols in the jth packet, which is determined by the MIMO mode. Assume that the rate-distortion function of the source input is available at the transmitter. Then, the average distortion, $E[D(\alpha, \beta)]$, can be computed as

$$E[D(\alpha, \beta)] = D(0, \alpha, \beta)\pi_0 + \sum_{k=1}^{N} D(\sum_{j=1}^{k} r_j M_j, \alpha, \beta)\pi_k, \quad (6)$$

where α and β are the leaky and partial prediction parameters, respectively, and $D(R, \alpha, \beta)$ is the distortion of the decoded MC-FGS video when the entire base layer and R bits of enhancement layer are received successfully.

At this point, we focus on the choice of $\underline{r} = [r_1, r_2, \ldots, r_N]$ and $\underline{\phi} = [\phi_1, \phi_2, \ldots, \phi_N]$ for minimizing $E[D(\alpha, \beta)]$, given α and β. In [15], the authors proposed a local search algorithm to get a suboptimal distortion-minimizing UEP policy when the length of the channel codeword is fixed. By using this algorithm, we can get a suboptimal rate allocation and MIMO mode selection to transmit the FGS enhancement layer bitstream. To do this, we define the effective code rate, s_i, as the ratio of the number of information symbols to the

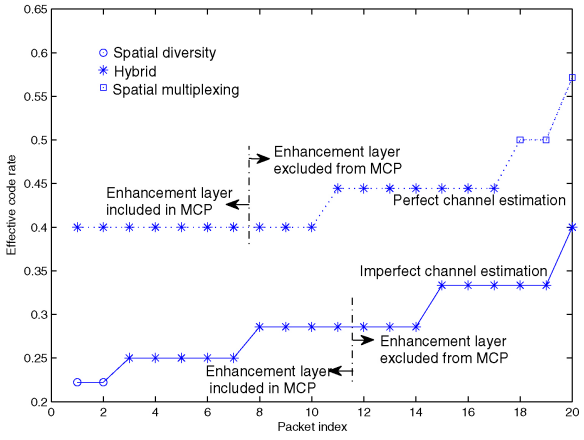


Fig. 3. UEP policy and MIMO mode selection for the *Foreman* sequence where $\alpha = 1.0$ and $\beta = 0.15$. Total transmit signal to noise power ratio is fixed at 8 dB.

maximum number of transmitted symbols in a packet, where the maximum number of transmitted symbols is defined as that in a packet where spatial multiplexing is used. For example, in a 4×4 MIMO system, $4 \times m$ symbols can be transmitted with spatial multiplexing. Then, the effective code rate of the ith packet, s_i, is given by

$$s_i = \begin{cases} r_i/4, & \text{if spatial diversity is used} \\ r_i/2, & \text{if hybrid mode is used} \\ r_i, & \text{if spatial multiplexing is used.} \end{cases}$$

The computational complexity of finding $\underline{s} = [s_1, s_2, s_3, \ldots, s_N]$ using an existing local search algorithm from [15] is lower than the complexity of finding $\underline{r}$ and $\underline{\phi}$ separately. The performance is the same either way, and is suboptimal.

In Fig. 3, the UEP policy and MIMO mode selection versus packet index are shown for the *Foreman* sequence, with $\alpha = 1$ and $\beta = 0.15$. For the FEC coding, rate-compatible punctured convolutional (RCPC) coding is used. Average channel SNR, P_T/σ_n^2, is fixed at 8 dB, where P_T is the total transmit power. Then, for perfect channel estimation, the first eight packets transmit the enhancement layer used for MCP. Since this part of the enhancement layer can result in error propagation if lost, additional protection is required. When there is imperfect channel estimation, more diversity gain and a lower channel code rate are required to protect the enhancement layer used for MCP, where the first 12 packets are involved. In particular, to reduce the loss probability for the beginning of the bitstream, spatial diversity is applied for the first two packets.

IV. Selection of Leaky and Partial Prediction Parameters

In this section, we discuss the selection of leaky and partial prediction parameters based on the available channel information at the transmitter. In [3] and [16], an analysis of error propagation with leaky prediction was presented. According to these analyses, at the encoder, the EL bitstream is generated by subtracting the BL residue, $(F_n^{\mathrm{BL}} - F_{n-1}^{\mathrm{BL}})$, and the high quality

TABLE I
Symbol Definitions

	Encoder	Decoder
Original/reconstructed frame	F_n	$\tilde{F}_n$
BL reconstructed frame	F_n^{BL}	F_n^{BL}
Original EL residue	ϵ_n^{EL}	–
Reconstructed EL residue	$\hat{\epsilon}_n^{\mathrm{EL}}$	$\tilde{\epsilon}_n^{\mathrm{EL}}$
Partially reconstructed EL residue	$\epsilon_n^{\mathrm{EL},\beta}$	$\tilde{\epsilon}_n^{\mathrm{EL},\beta}$
High quality reference frame	$\hat{F}_n^{\mathrm{EL,enc},\beta}$	$\tilde{F}_n^{\mathrm{EL,dec},\beta}$
Partially reconstructed frame	$F_n^{\mathrm{EL},\beta}$	$\tilde{F}_n^{\mathrm{EL},\beta}$

reference frame, $\hat{F}_n^{\mathrm{EL,enc},\beta}$, from the original frame, F_n. That is, the following residue is progressively encoded as an EL of the current frame

$$\epsilon_n^{\mathrm{EL}} = F_n - (F_n^{\mathrm{BL}} - F_{n-1}^{\mathrm{BL}}) - \hat{F}_n^{\mathrm{EL,enc},\beta}.$$

Note that the progressively encoded EL can be truncated anywhere. For example, in MPEG-4 FGS, the EL is encoded using bit-plane coding, and less significant bits can be discarded, at the cost of lower quality. When transmitted, the EL bitstream can be truncated due to a limitation of channel bandwidth or transmit power. We denote the reconstructed EL residue at the encoder using the transmitted EL bitstream as $\hat{\epsilon}_n^{\mathrm{EL}}$. At the receiver, additional truncation can occur in the EL bitstream caused by channel error. Then, $\hat{\epsilon}_n^{\mathrm{EL}}$ and $\tilde{\epsilon}_n^{\mathrm{EL}}$, which represent the reconstructed EL residue at the decoder, are the coarsely quantized versions of ϵ_n^{EL} due to the truncation of the original EL bitstream at the encoder and decoder. We assume no mismatch for the BL at the encoder and decoder. The reconstructed frame at the decoder, $\tilde{F}_n$, is given by

$$\tilde{F}_n = \tilde{\epsilon}_n^{\mathrm{EL}} + (F_n^{\mathrm{BL}} - F_{n-1}^{\mathrm{BL}}) + \tilde{F}_n^{\mathrm{EL,dec},\beta},$$

where $\tilde{F}_n^{\mathrm{EL,dec},\beta}$ is the high quality EL reference for frame n at the decoder. That is

$$\tilde{F}_n^{\mathrm{EL,dec},\beta} = (1-\alpha) F_{n-1}^{\mathrm{BL}} + \alpha \tilde{F}_{n-1}^{\mathrm{EL},\beta}. \tag{7}$$

Hence

$$\tilde{F}_n = \tilde{\epsilon}_n^{\mathrm{EL}} + \alpha(\tilde{F}_{n-1}^{\mathrm{EL},\beta} - F_{n-1}^{\mathrm{BL}}) + F_n^{\mathrm{BL}} \tag{8}$$

where $\tilde{F}_{n-1}^{\mathrm{EL},\beta}$ is the reconstructed $(n-1)$-st frame using the BL and a partial EL specified by β, which is stored in the buffer at the decoder. Therefore

$$\tilde{F}_{n-1}^{\mathrm{EL},\beta} = \tilde{\epsilon}_{n-1}^{\mathrm{EL},\beta} + (F_{n-1}^{\mathrm{BL}} - F_{n-2}^{\mathrm{BL}}) + \tilde{F}_{n-1}^{\mathrm{EL,dec},\beta} \tag{9}$$

where $\tilde{\epsilon}_{n-1}^{\mathrm{EL},\beta}$ represents the reconstructed partial EL residue at the $(n-1)$-st frame at the decoder. This may not be the same as the reconstructed partial EL residue at the encoder, $\epsilon_{n-1}^{\mathrm{EL},\beta}$, if the successfully received number of EL bits is less than the amount of partial EL specified by β. Also, the partially reconstructed $(n-1)$-st frame at the encoder, $F_{n-1}^{\mathrm{EL},\beta}$, is

$$F_{n-1}^{\mathrm{EL},\beta} = \epsilon_{n-1}^{\mathrm{EL},\beta} + (F_{n-1}^{\mathrm{BL}} - F_{n-2}^{\mathrm{BL}}) + \hat{F}_{n-1}^{\mathrm{EL,enc},\beta}. \tag{10}$$

The error between the original frame at the encoder, F_n, and the reconstructed frame at the decoder, $\tilde{F}_n$, is

$$\begin{aligned} F_n - \tilde{F}_n &= \epsilon_n^{\mathrm{EL}} - \tilde{\epsilon}_n^{\mathrm{EL}} + \hat{F}_n^{\mathrm{EL,enc},\beta} - \tilde{F}_n^{\mathrm{EL,dec},\beta} \\ &= \epsilon_n^{\mathrm{EL}} - \tilde{\epsilon}_n^{\mathrm{EL}} + \alpha(F_{n-1}^{\mathrm{EL},\beta} - \tilde{F}_{n-1}^{\mathrm{EL},\beta}). \end{aligned} \tag{11}$$

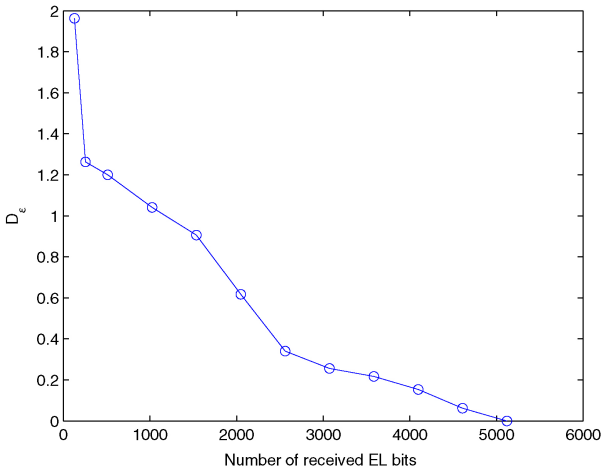


Fig. 4. Distortion corresponding to the reference mismatch, D_ϵ, where $\alpha = 1.0$ and the amount of EL-MCP is 5120 bits.

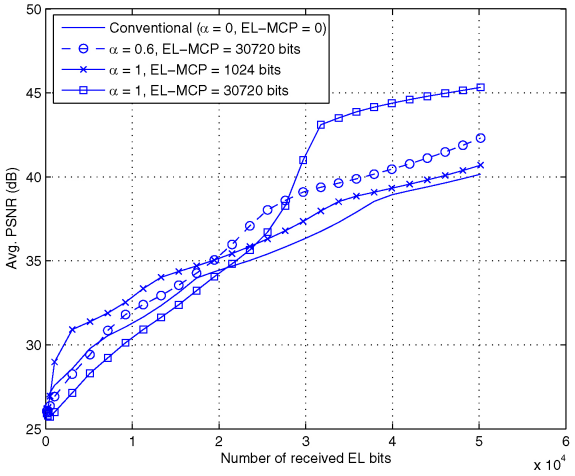


Fig. 5. Rate-PSNR characteristics for MC-FGS video codec and for conventional FGS video codec, generated from the first 150 frames of *Foreman*.

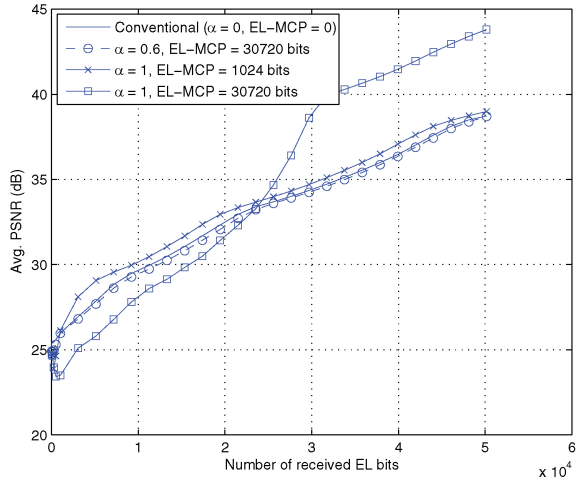


Fig. 6. Rate-PSNR characteristics for MC-FGS video codec and for conventional FGS video codec, generated from the first 150 frames of *Coast guard*.

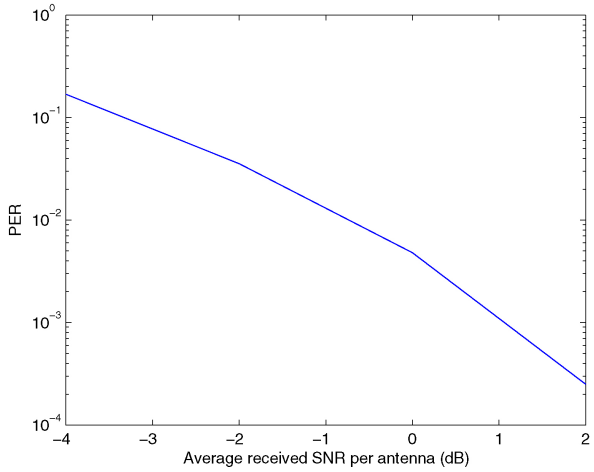


Fig. 7. Packet error rate (PER) for the BL with the lowest code rate and spatial diversity, where $f_{\text{nd}} = 10^{-2}$.

Note that the difference between $F_{n-1}^{\text{EL},\beta}$ and $\tilde{F}_{n-1}^{\text{EL},\beta}$ results in reference mismatch while reconstructing the current frame. From (9) and (10), $(F_{n-1}^{\text{EL},\beta} - \tilde{F}_{n-1}^{\text{EL},\beta})$ can be represented as

$$F_{n-1}^{\text{EL},\beta} - \tilde{F}_{n-1}^{\text{EL},\beta} = \epsilon_{n-1}^{\text{EL},\beta} - \tilde{\epsilon}_{n-1}^{\text{EL},\beta} + \alpha(F_{n-2}^{\text{EL},\beta} - \tilde{F}_{n-2}^{\text{EL},\beta}). \quad (12)$$

Then, the distortion of the nth frame, $D_n(\underline{R}, \alpha, \beta)$, is

$$\begin{aligned} D_n(\underline{R}, \alpha, \beta) &= E[(F_n - \tilde{F}_n)^2] \\ &= E[(\epsilon_n^{\text{EL}} - \tilde{\epsilon}_n^{\text{EL}} + \alpha(F_{n-1}^{\text{EL},\beta} - \tilde{F}_{n-1}^{\text{EL},\beta}))^2] \end{aligned}$$

where $\underline{R}$ is an $(n \times 1)$ vector of the throughputs from the first to the nth frames. If we assume $(\epsilon_n^{\text{EL}} - \tilde{\epsilon}_n^{\text{EL}})$, which depends on the channels corresponding to the nth frame, is independent of $(F_{n-1}^{\text{EL},\beta} - \tilde{F}_{n-1}^{\text{EL},\beta})$ for rapidly varying channels, the above expression becomes

$$\begin{aligned} D_n(\underline{R}, \alpha, \beta) &= E[(\epsilon_n^{\text{EL}} - \tilde{\epsilon}_n^{\text{EL}})^2] \\ &\quad + 2\alpha E[\epsilon_n^{\text{EL}} - \tilde{\epsilon}_n^{\text{EL}}]E[F_{n-1}^{\text{EL},\beta} - \tilde{F}_{n-1}^{\text{EL},\beta}] \\ &\quad + \alpha^2 E[(F_{n-1}^{\text{EL},\beta} - \tilde{F}_{n-1}^{\text{EL},\beta})^2]. \end{aligned} \quad (13)$$

In [17], the residue between the original frame and the reconstructed frame in the BL was well modelled as a zero-mean generalized Gaussian random variable. And, in MC-FGS, to eliminate the temporal redundancy, the previous frame's residue is subtracted from that of the current frame, either partially or entirely. Since both residues are considered as zero-mean generalized Gaussian random variables, the difference between them, ϵ_n^{EL}, can also be considered a zero-mean generalized Gaussian random variable. Note that $\tilde{\epsilon}_n^{\text{EL}}$ is the quantized version of ϵ_n^{EL}, and the distributions of both variables are symmetric around zero. Therefore, the mean of ϵ_n^{EL} and $\tilde{\epsilon}_n^{\text{EL}}$ will be zero and (13) can be represented as

$$\begin{aligned} & D_n(\underline{R}, \alpha, \beta) \\ &= E[(\epsilon_n^{\text{EL}} - \tilde{\epsilon}_n^{\text{EL}})^2] + \alpha^2 E[(F_{n-1}^{\text{EL},\beta} - \tilde{F}_{n-1}^{\text{EL},\beta})^2] \\ &= E[(\epsilon_n^{\text{EL}} - \tilde{\epsilon}_n^{\text{EL}})^2] \\ &\quad + \alpha^2 E[(\alpha(F_{n-2}^{\text{EL},\beta} - \tilde{F}_{n-2}^{\text{EL},\beta}) + \epsilon_{n-1}^{\text{EL},\beta} - \tilde{\epsilon}_{n-1}^{\text{EL},\beta})^2]. \end{aligned} \quad (14)$$

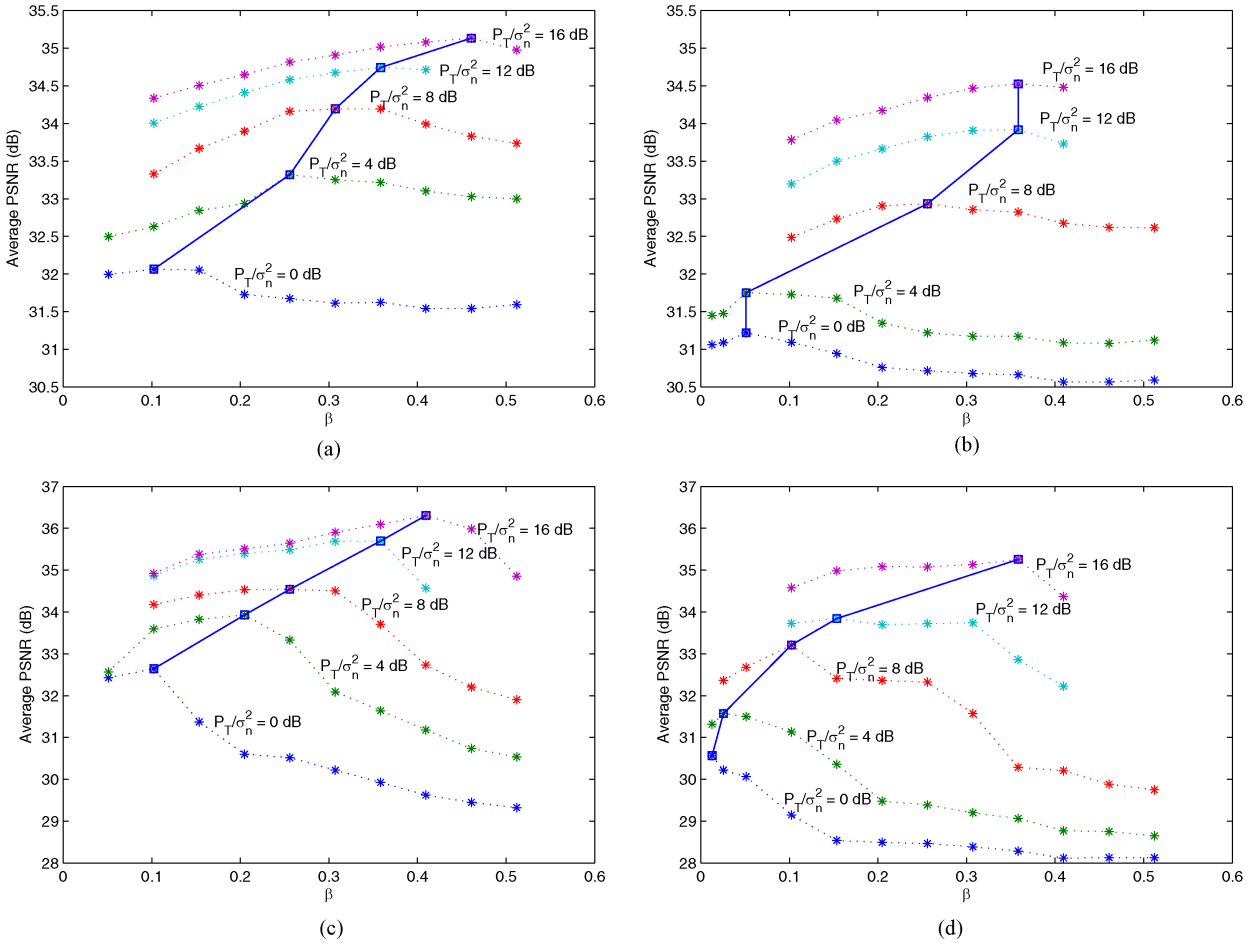


Fig. 8. Average PSNR performance versus β for the *Foreman* sequence. The leaky prediction parameter is fixed at 0.7 and 1.0, and both perfect and imperfect channel estimation are considered. (a) $\alpha = 0.7$, perfect channel estimation. (b) $\alpha = 0.7$, imperfect channel estimation. (c) $\alpha = 1.0$, perfect channel estimation. (d) $\alpha = 1.0$, imperfect channel estimation.

We define $D_\epsilon(R_{n-1}, \beta)$ as

$$D_\epsilon(R_{n-1}, \beta) = E[(\epsilon_{n-1}^{\mathrm{EL},\beta} - \tilde{\epsilon}_{n-1}^{\mathrm{EL},\beta})^2], \tag{15}$$

where R_{n-1} is the number of successfully received bits for the EL at the $(n-1)$st frame and $D_\epsilon(R_{n-1}, \beta)$ becomes zero if R_{n-1} is larger than the amount of EL specified by β. With (12) and (15)

$$\begin{aligned} D_n(\underline{R}, \alpha, \beta) \;=\; & E[(\epsilon_n^{\mathrm{EL}} - \tilde{\epsilon}_n^{\mathrm{EL}})^2] \\ & + \alpha^{2(n-1)} E[(F_1^{\mathrm{EL},\beta} - \tilde{F}_1^{\mathrm{EL},\beta})^2] \\ & + \sum_{i=1}^{n-2} \alpha^{2i} D_\epsilon(R_{n-i}, \beta) \end{aligned} \tag{16}$$

where the second term results from the partial EL reference mismatch at the first frame, and the last term represents the accumulated distortion caused by the reconstructed EL reference mismatch from the subsequent frames. It follows from (2) that large α and β provide better coding efficiency. However, in (16), when α is chosen to be large, the error resilience becomes worse, since the accumulated distortion, $E[(F_1^{\mathrm{EL},\beta} - \tilde{F}_1^{\mathrm{EL},\beta})^2]$ and $\sum_{i=1}^{n-2} D_\epsilon(R_{n-i}, \beta)$ will not decay rapidly. By choosing a large β, the probability that the second and third terms are greater than zero will be increased. In Fig. 4, we present the actual value of D_ϵ for the 10th frame in the *Foreman* video sequence when 5120 bits in the EL bitstream are used for MCP. As seen, if the number of successfully received bits is more than that, D_ϵ becomes zero. However, as the amount of the received EL reduces, D_ϵ increases proportional to the amount of reference mismatch. Therefore, to select the prediction parameters, α and β, to appropriately trade off coding efficiency and error resilience capability, the statistics of R_{n-i} need to be considered.

In Figs. 5 and 6, we present the rate-PSNR characteristics for the MC-FGS video codec with various leaky and partial prediction parameters, and for the conventional FGS video codec. As shown, as leaky prediction parameter, α, increases, the quality of reconstructed video is significantly improved if the video does not suffer from any error propagation. However, if the number of received EL bits is less than the amount of EL-MCP, then the use of a larger leaky parameter results in poor error resilience. Therefore, it can be worse than either conventional FGS video codec or MC-FGS video codec with a small leaky prediction parameter.

In the following, we have two different strategies to select α and β based on the available channel information. If accurate instantaneous channel information is available at the transmitter to choose β based on the expected number of successfully received EL bits to avoid reference mismatch, then α can be

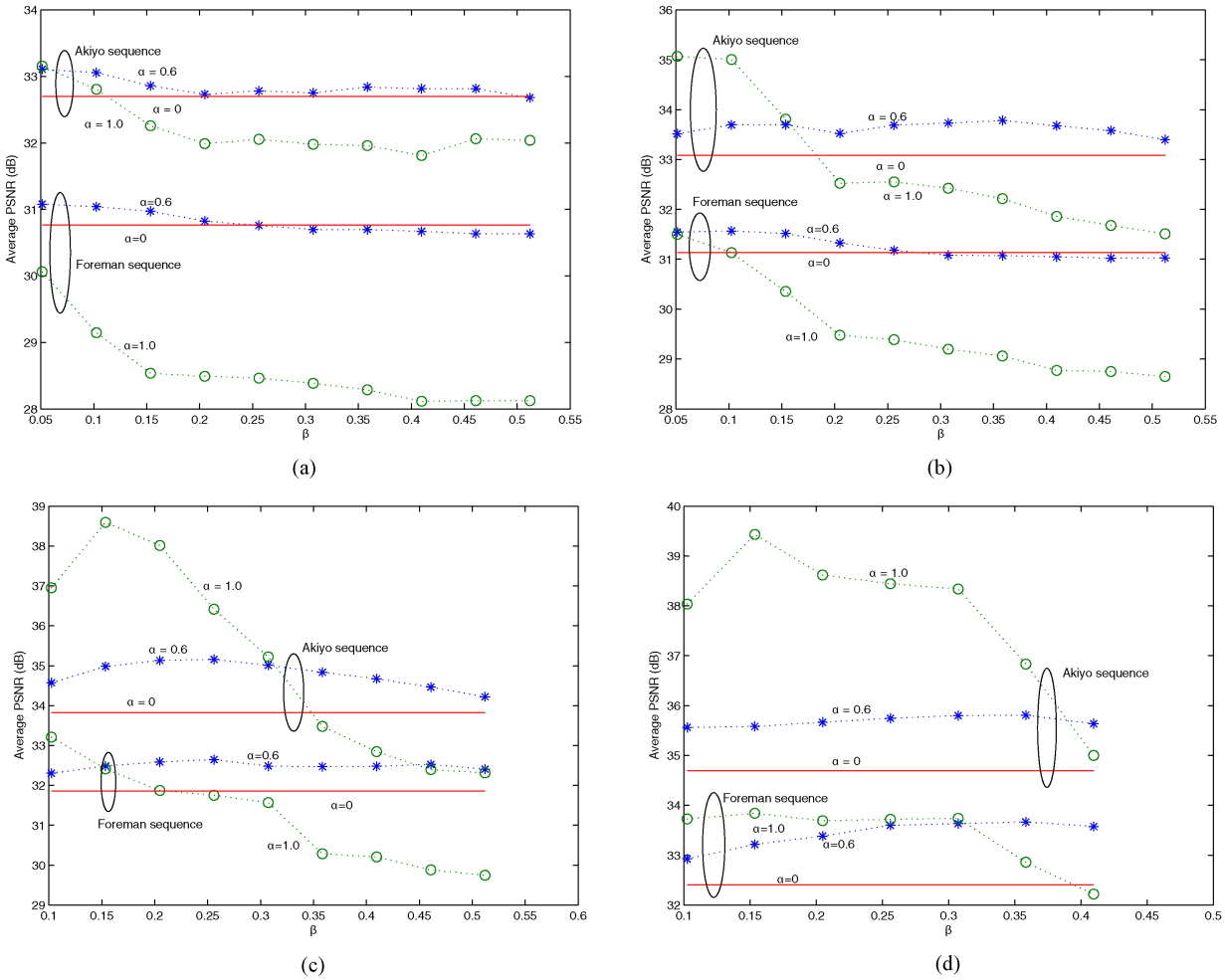


Fig. 9. Average PSNR performance versus β for the *Foreman* and *Akiyo* sequences where P_T/σ_n^2 represents the ratio of the total transmit power and noise variance. The solid line represents the average PSNR without using MC-FGS. Imperfect channel estimation is considered. (a) $P_T/\sigma_n^2 = 0$ dB. (b) $P_T/\sigma_n^2 = 4$ dB. (c) $P_T/\sigma_n^2 = 8$ dB. (d) $P_T/\sigma_n^2 = 12$ dB.

chosen to be unity, since the probability of error propagation will be reduced. However, in a high Doppler environment, where instantaneous channel information may not be available, the average channel information will be used to decide the parameters, where these parameters will be maintained for multiple frames.

A. Selection of Fixed α and β—Fast Fading Channel

In a high Doppler environment, it is hard to get instantaneous channel information at the transmitter. Instead, we use the average channel SNR to select the prediction parameters, as introduced previously to decide UEP. Due to the lack of an analytical approach for the leaky and partial prediction schemes, we find the prediction parameters based on simulation. For the simulation, we use MC-FGS video, which is implemented with an H.264 TML-9 codec for the base layer and an MPEG-4 FGS codec for the EL. Both partial and leaky prediction schemes are incorporated in the enhancement layer MCP loop. The first 150 frames of the video sequence, consisting of a single intra-frame (I-frame) and 149 predicted frames (P-frames), are encoded. Before transmitting the bitstream through the MIMO wireless channel, the bitstream is packetized and channel-encoded by rate compatible punctured convolutional (RCPC) codes. In this paper, we assume the lowest RCPC code rate and spatial diversity are used for the delivery of the BL to achieve the highest reliability. In the Fig. 7, the PER for the BL is plotted with a normalized Doppler frequency of 10^{-2}. For the EL bitstream, its UEP policy and leaky/partial prediction parameters are jointly selected. We fix the size of a packet at 256 BPSK symbols, and each frame is encoded using 20 packets. A 4×4 MIMO antenna configuration is considered, and a minimum mean square error (MMSE) detection scheme is used for all the MIMO modes. We consider Jakes' model with a normalized Doppler frequency of 10^{-2}, which is smaller than the transmission time for a frame for the Rayleigh fading channel.

In Fig. 8, the average PSNR versus various partial prediction parameters is shown when the leaky prediction parameter is fixed at 0.7 and 1.0, under both perfect and imperfect channel estimation. Given the average channel SNR, the optimal selection of the partial prediction parameter, β, is highlighted by the square (□). For all scenarios, it is seen that a larger β can be chosen as the average received SNR increases. This is because a higher channel SNR can enhance the number of successfully received bits, R_j, for the frame and allow the use of larger

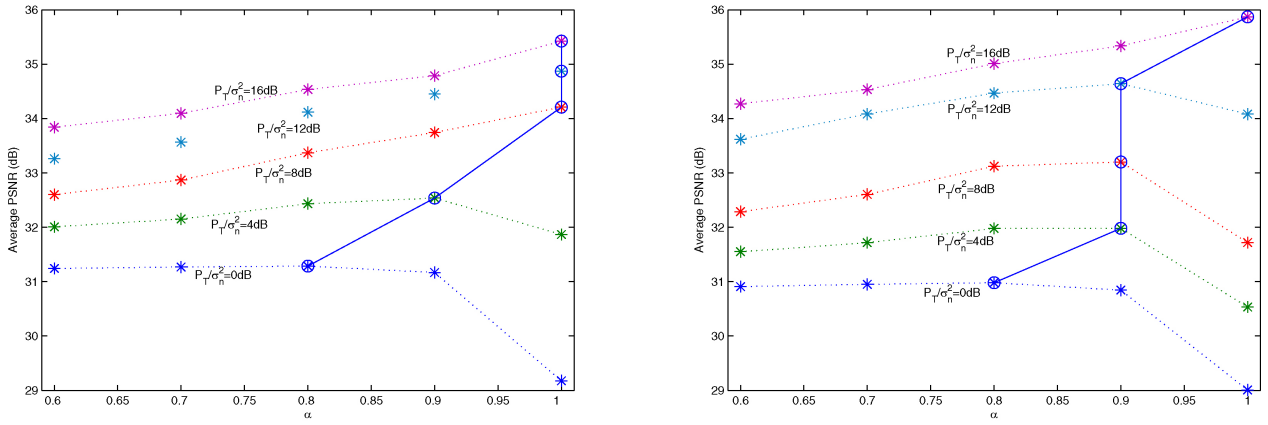


Fig. 10. Average PSNR performance versus α for *Foreman*. The partial prediction parameter is fixed at 0.15 and 0.35, and imperfect channel estimation is considered. (a) $\beta = 0.15$. (b) $\beta = 0.35$.

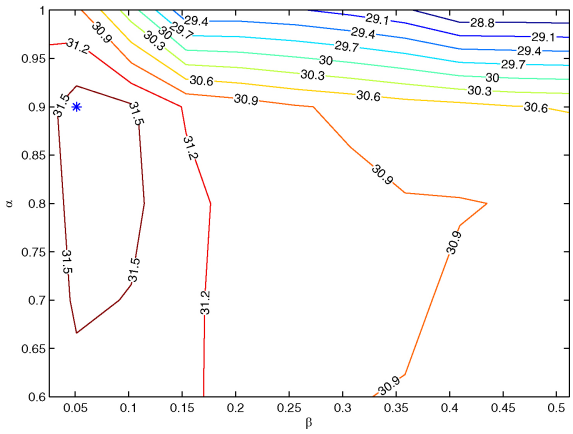


Fig. 11. Contour plot of the PSNR performance for different (β, α) with the average channel SNR of 0 dB, where the normalized Doppler frequency $f_{nd} = 10^{-2}$.

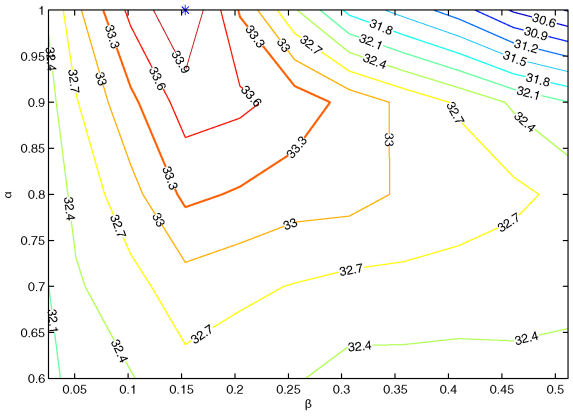


Fig. 12. Contour plot of the PSNR performance for different (β, α) with the average channel SNR of 8 dB, where the normalized Doppler frequency $f_{nd} = 10^{-2}$.

β without increasing $D_\epsilon(R_j, \beta)$. If we compare Fig. 8(a) with Fig. 8(b), it can be observed that imperfect channel estimation results in decreasing partial prediction parameters. However, the choice of β is not sensitive, since the propagated error can be forced to decay by choosing $\alpha < 1$. In contrast, in Fig. 8(c) and (d) with α set to 1, the performance is more sensitive to the choice of β. Because the propagated error is not being

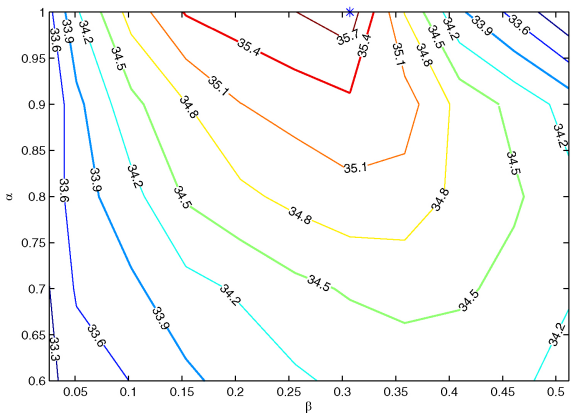


Fig. 13. Contour plot of the PSNR performance for different (β, α) with the average channel SNR of 16 dB, where the normalized Doppler frequency $f_{nd} = 10^{-2}$.

forced to decay, a small value of β is preferred, compared to the case of $\alpha = 0.7$. If we choose a larger β than the optimal one, the corresponding performance is significantly degraded due to the lack of error resilience capability.

In Fig. 9, the average PSNR performance of the MC-FGS video with partial and leaky prediction is compared to that of a conventional FGS video under imperfect channel estimation for *Foreman* and *Akiyo* video sequences. As a point of comparison, we also implement conventional FGS video coding without partial and leaky prediction by setting α to 0. At low SNR, the MC-FGS video coding does not provide significant gain over conventional FGS video coding, since it suffers from error propagation even though a small value of β is chosen. However, by choosing a small α, performance can be enhanced by reducing the propagated error. As the channel SNR increases, MC-FGS outperforms conventional FGS, even if a large value of β is used.

In Fig. 10, the average PSNR versus various leaky prediction parameters is shown when the partial prediction parameter is fixed at a specific value. As can be seen, as a larger value of β is chosen, a relatively smaller value of α needs to be selected. For example, at 8 dB of average channel SNR, if β is 0.15, then the best performance is achieved when α is chosen to be unity. In contrast, for the same conditions, if β is 0.35, then α

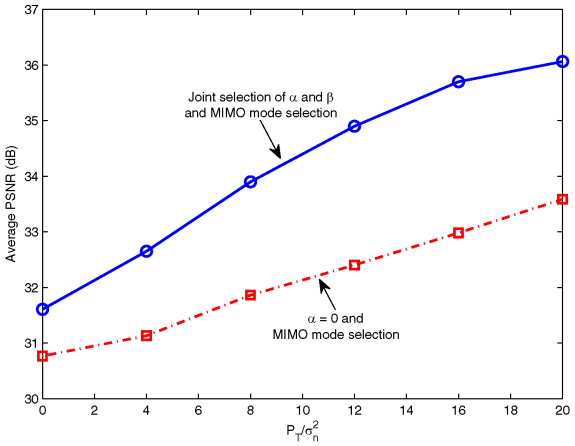


Fig. 14. Comparison of average PSNR for the conventional FGS and MC-FGS video coding with fixed α and β under imperfect channel estimation, where the normalized Doppler frequency $f_{\text{nd}} = 10^{-2}$.

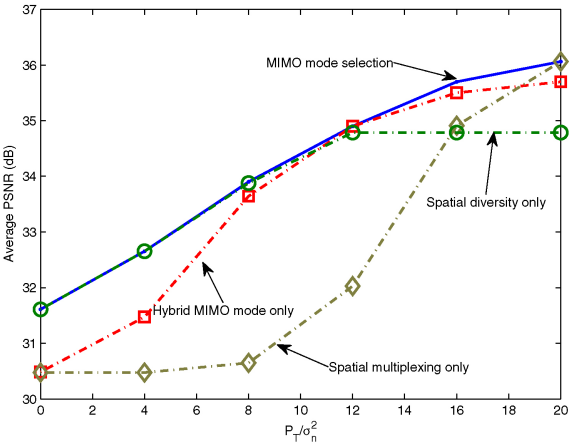


Fig. 15. Comparison of average PSNR for the MIMO mode selection and single MIMO mode with fixed α and β under imperfect channel estimation, where the normalized Doppler frequency $f_{\text{nd}} = 10^{-2}$.

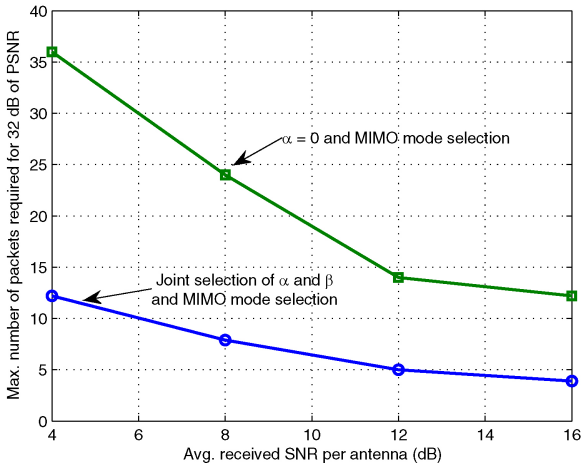


Fig. 16. Maximum number of packets required for 32 dB of PSNR when *Foreman* is transmitted over the channel whose normalized Doppler frequency $f_{\text{nd}} = 10^{-2}$.

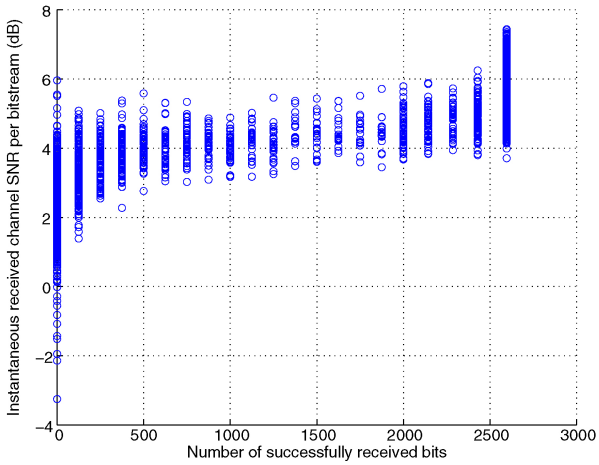


Fig. 17. Distribution of the throughput and corresponding instantaneous received channel SNR per bitstream from the investigation of transmission of 2500 frames in 4×4 MIMO system. $P_T/\sigma_n^2 = 0$ dB and ZF detection is used and $f_{\text{nd}} = 10^{-5}$.

is required to be less than unity because 8 dB of channel SNR is not enough to reliably deliver the amount of EL designated by β of 0.35. In addition, higher channel SNR allows the use of a higher α, and achieves a better compression efficiency, since the EL can be delivered more reliably.

In Fig. 11, we show a contour plot of constant PSNR performance when choosing α and β jointly when the average channel SNR is fixed at 0 dB. Each contour represents a constant PSNR achieved by different combinations of α and β. Note that we indicated the optimal pair of parameters with the symbol (*). In Figs. 12 and 13, the same contour plots with 8 and 16 dB of average channel SNR are presented. Compared to the previous figure, the optimal values of both α and β are increased. In particular, the figures show that β is closely related to the average channel SNR.

In Fig. 14, the PSNR performance of conventional FGS video coding is compared with that of MC-FGS video coding using the leaky and partial predictions. For both systems, UEP and MIMO mode selection are applied. In the conventional FGS video coding, the FGS EL is excluded in MCP, so error propagation due to the corruption of the EL is entirely prohibited. Instead, it can suffer from poor compression efficiency. Our simulation results show that the leaky and partial predictions can result in a significant gain over a conventional FGS over a wide range of channel SNR in fast fading channels by controlling the tradeoff between compression efficiency and error resilience. In Fig. 15, the PSNR performances for MC-FGS video coding with and without MIMO mode selection are presented. Note that, at low SNR, the selected MIMO mode for all packets is spatial diversity. Therefore, the performance of MIMO mode selection is equivalent to that of spatial diversity only. For the same reason, at high SNR, the performance of MIMO mode selection is the same as that of spatial multiplexing. However, in the intermediate region, different MIMO modes can be assigned based on the packet's priority, as seen in Fig. 3, and MIMO mode selection provides better performance than any single choice of MIMO mode.

Fig. 16 shows the bandwidth reduction achieved by using the joint selection of leaky/partial prediction parameters as well as FEC coding rate/MIMO mode selection, compared to the only the selection of the channel parameter with conventional video codec. In this plot, we assume 256 BPSK symbols per packet and 4×4 MIMO channels. As shown, to achieve 32 dB PSNR for the *Foreman* sequence, for the entire

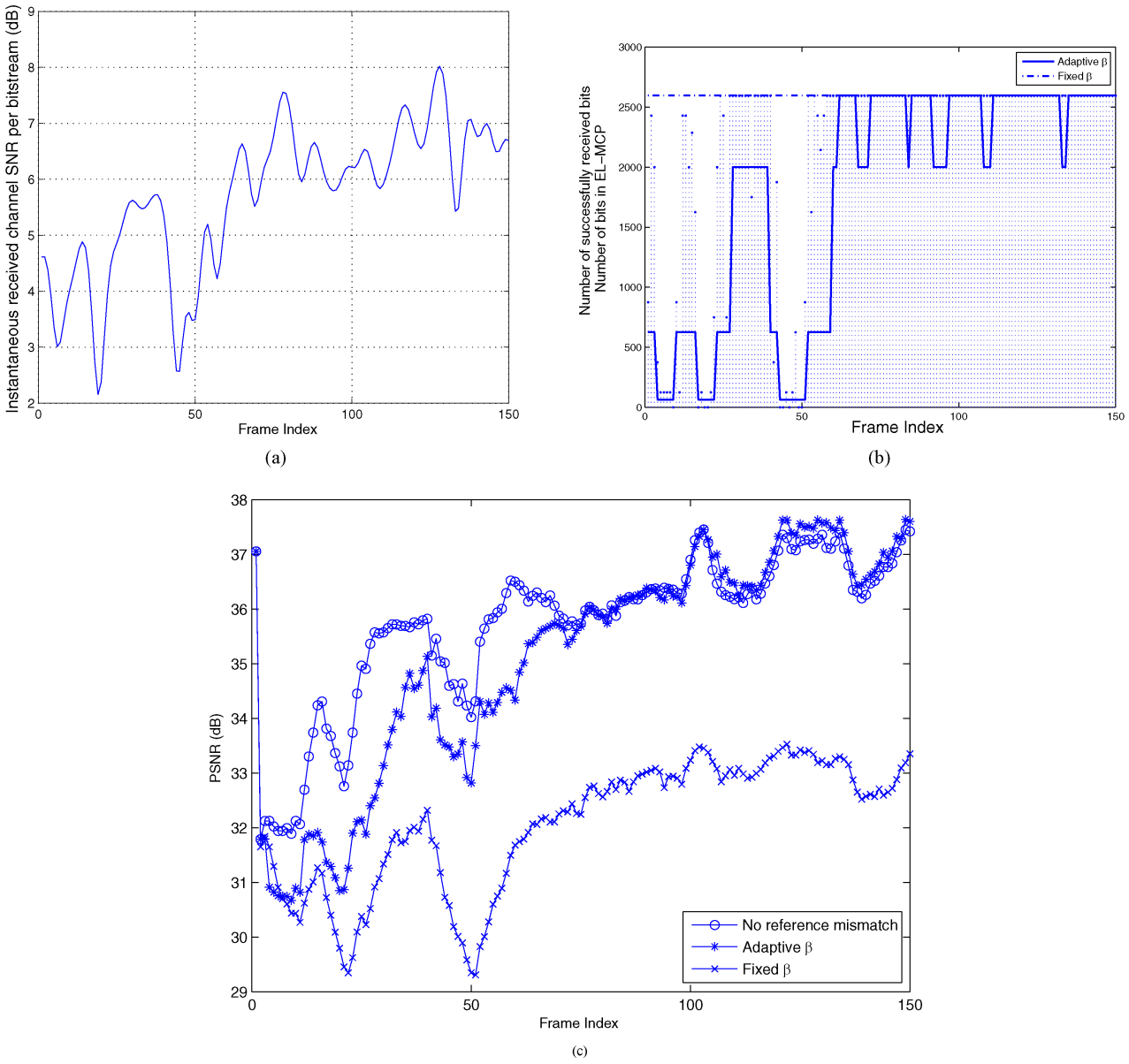


Fig. 18. $P_T/\sigma_n^2 = 0$ dB. Video simulation using the *Akiyo* sequence. (a) Instantaneous channel received SNR. (b) Actual throughput and selected β. (c) PSNR profiles.

range of SNR, joint selection of source and channel parameters requires at most half as many packets compared to the case of optimizing only over the channel parameters.

B. Adaptive Selection of β—Slow Fading Channel

Previously, we saw that β is closely related to the average channel SNR or, equivalently, the number of successfully received bits. That is, if the number of successfully received EL bits is less than the amount of the EL specified by β, then reference mismatch between encoder and decoder can result. Therefore, if the transmitter can estimate the instantaneous throughput for the frame which is encoded currently, then it can be used to choose a value of β to avoid reference mismatch. In this section, by utilizing channel feedback from the decoder, we propose adaptive selection of β. To do that, we assume a slowly varying channel, where the channel is highly correlated over multiple frames.

Before the encoding of the nth frame, it is assumed that the instantaneous channel information corresponding to the transmission of frame $(n-1)$ is available at the encoder. Then, based on this instantaneous channel information, we decide the appropriate β. At this point, we are not assuming adaptive UEP based on the instantaneous channel information in order to avoid high computational complexity to find the UEP policy frame-by-frame. That is, the UEP is only dependent on the average channel SNR and is kept fixed for some predetermined number of frames (in our simulation, 150 frames). Then, the transmitter finds the distribution of the throughput, T, when the instantaneous channel information, especially the instantaneous received SNR per bitstream, is given, where the instantaneous received SNR per bitstream is defined as the SNR at the receiver after signal combining, and is presented in [14] and [18]. This will be possible by assuming that the transmitter knows the throughput profiles

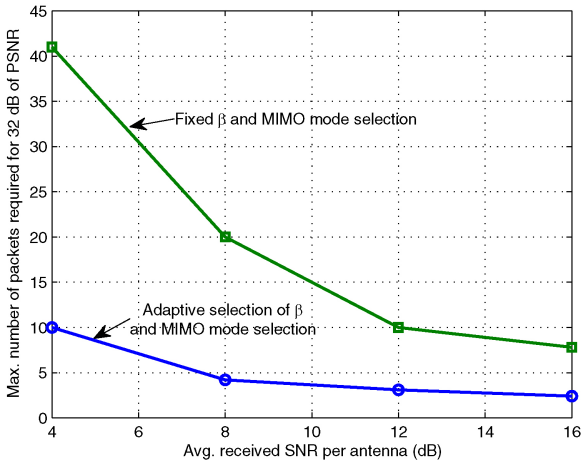


Fig. 19. Maximum number of packets required for 32 dB of PSNR when *Foreman* is transmitted over the channel whose normalized Doppler frequency $f_{nd} = 10^{-5}$.

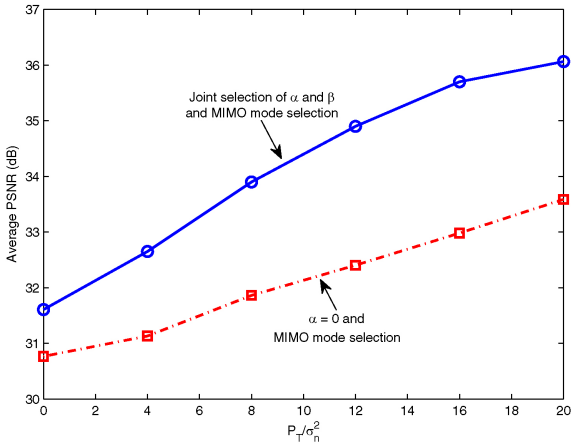


Fig. 20. Performance comparison of different strategies to select β.

over the previous frames and corresponding instantaneous channel SNRs as shown in Fig. 17.

To compute the instantaneous received SNR per bitstream, we follow the expressions in [14] and [18], where the computation is different for each of the different MIMO modes. Due to the important role of the first packet in the progressive encoded EL, we compute the instantaneous received channel SNR per bitstream for the MIMO mode of the first packet.

We denote the probability mass function of the throughput by $P_T(T)$, where $0 \le T \le \sum_{j=1}^{N} r_j M_j$. Then we consider how to choose β from this information. Denote the average PSNR from frame n to $(N+n-1)$ by $PSNR_n(\underline{\beta}, \underline{T})$, where the vectors $\underline{\beta} = [\beta_n, \beta_{n+1}, \cdots, \beta_{N+n-1}]^T$ and $\underline{T} = [T_n, T_{n+1}, \cdots, T_{N+n-1}]^T$ represent the partial prediction parameter β and the throughput for N frames. We assume $PSNR_n(\underline{\beta}, \underline{T})$ is available at the encoder. We choose β to maximize the expected $PSNR_n$ when the instantaneous received channel SNR per bitstream is given. If β_n can be chosen from the set of $\{\beta^1, \beta^2, \cdots, \beta^K\}$, then for each possible value of β_n, the expected $PSNR_n$ is computed as follows using the probability mass function of the throughput

$$\begin{aligned} &E[PSNR_n([\beta_n = \beta^i, \beta_{n+1}, \cdots, \beta_{N+n-1}], \underline{T})] \\ &= \sum_{t=0}^{r_j M_j} P_T(T_n = t) \\ &\cdot PSNR_n([\beta_n = \beta^i, \beta_{n+1}, \cdots, \beta_{N+n-1}], \\ &\cdot [T_n = t, T_{n+1}, \cdots, T_{N+n-1}]) \end{aligned} \tag{17}$$

where all elements in the vectors $\underline{\beta}$ and $\underline{T}$ except β_n and T_n are fixed at the specific value not to result in reference mismatch in the subsequent $N-1$ frames. That is

$$\begin{aligned} &E[PSNR_n([\beta_n, \beta_{n+1}, \cdots, \beta_{N+n-1}], \underline{T})] \\ &= E[PSNR_n(\beta_n, T_n)]. \end{aligned}$$

Then, β_n is chosen to satisfy the following:

$$\max_{\beta_n \in \{\beta^1, \beta^2, \cdots, \beta^K\}} E[PSNR_n(\beta_n, T_n)]. \tag{18}$$

An example of adaptive choice of β is presented in Fig. 18(b) when the instantaneous channel is given as shown in Fig. 18(a) with $f_{nd} = 10^{-5}$, which makes the channel correlated over a time period of three frames. As can be seen, when β is selected adaptively, the possibility of reference mismatch is reduced. Note that α is fixed at unity for the case of adaptive selection of β due to the lack of reference mismatch, and less than unity to compensate for the propagated error when β is fixed. Their performances are compared in Fig. 18(c), where PSNR profiles for all strategies are presented.

Fig. 19 shows the reduction in the number of packets required to achieve 32 dB PSNR by using the adaptive selection of the partial prediction parameter per frame compared to using a fixed partial prediction parameter for the 150 frames. As shown here, over the entire range of SNR, adaptive selection of the partial prediction parameter yields a significant reduction in the number of packets required to achieve a specific video quality.

In Fig. 20, a performance comparison of different selection strategies for β is presented for *Akiyo*. To choose β for the current frame, we assume that the instantaneous channel information when the previous frame was transmitted is known at the transmitter before the encoding of the current frame. In the figure, the case of no reference mismatch represents the ideal case where the encoder knows the instantaneous throughput for every frame. By choosing β equal to the instantaneous throughput, reference mismatch never happens. In contrast, fixing β represents the case where both α and β are jointly selected for the whole video using average instead of instantaneous channel information. As can be seen, when β is fixed, the PSNR loss is more than 3 dB compared to the no reference mismatch case, especially in the low SNR region. However, based on the simulation, the adaptive approach to select β frame-by-frame showed less than 0.5 dB PSNR reduction compared to the no reference mismatch case. Note that we fixed $\alpha = 1$ when β is selected adaptively, since the probability of reference mismatch is reduced significantly.

V. Conclusion

In this paper, we studied the transmission of an MC-FGS video bitstream over a MIMO wireless channel under the condition of imperfect channel estimation. We proposed a UEP policy consisting of FEC and MIMO mode selection per packet for the enhancement layer of MC-FGS by exploiting

the fundamental tradeoff between multiplexing and diversity. To compensate for reference mismatch and the resulting error propagation, leaky and partial prediction schemes are applied in the MCP loop. Then we investigated the average PSNR performance with various choices of the leaky and partial prediction parameters for rapidly varying channels. Simulation results show that the joint control of prediction parameters and UEP can enhance the system performance significantly. In addition, we investigated the adaptive selection of the partial prediction parameter in a slow fading channel. If instantaneous channel information is available at the transmitter, the partial prediction parameter can be chosen based on the throughput which is estimated using the instantaneous channel information. Simulation results show that the adaptive selection of the partial prediction parameter is useful if the channel is varying slowly to allow the transmitter to obtain instantaneous channel information. Therefore, for the selection of prediction parameters, channel information, such as the channel SNR and the coherence time, needs to be considered.

REFERENCES

[1] M. van der Schaar and H. Radha, "Adaptive motion-compensation fine granular-scalability (AMC-FGS) for wireless video," *IEEE Trans. CSVT.,* vol. 12, no. 6, pp. 360–371, Jun. 2002.

[2] F. Wu, S. Li, and Y-Q. Zhang, "A framework for efficient progressive fine granularity scalable video coding," *IEEE Trans. CSVT.,* vol. 11, no. 3, pp. 332–344, Mar. 2001.

[3] H.-C. Huang, C.-N. Wang, and T. Chiang, "A robust fine granularity scalability using trellis-based predictive leak," *IEEE Trans. CSVT.,* vol. 12, no. 6, pp. 372–385, Jun. 2002.

[4] Y. S. Chan, P. C. Cosman, and L. B. Milstein, "A multiple description coding and delivery scheme for motion-compensated fine granularity scalable video," *IEEE Trans. Image Process.,* vol. 17, no. 8, pp. 1353–1366, Aug. 2008.

[5] S. Lin, A. Stefanov, and Y. Wang, "On the performane of space-time block-coded MIMO video communications," *IEEE Trans. Veh. Tech.,* vol. 56, no. 3, pp. 1223–1229, May 2007.

[6] R. Hormis, E. Linzer, and X. Wang "Joint diversity- and rate-control for video transmission on multi-antenna channels," in *Proc. Globecom,* pp. 2064–2069, Nov. 2007.

[7] T. Holliday, A. J. Goldsmith, and H. V. Poor, "Joint source and channel coding for MIMO systems: Is it better to be robust or quick?" *IEEE Trans. IT.,* vol. 54, no. 4, pp. 1393–1405, Apr. 2008.

[8] 3GPP, "Technical Specification Group Radio Access Network : Multiple Input Multiple Output in UTRA," 3GPP, Tech. Rep., 25.876 V1.7.0, Aug. 2004.

[9] S. N. Diggavi, A. R. Calderbank, S. Dusad, and N. Al-Dhahir, "Diversity embedded space-time codes," *IEEE Trans. IT.,* vol. 54, no. 1, pp. 33–50, Jan. 2008.

[10] A. Kwasinski and K. J. R. Liu, "Optimal unequal error protection with user cooperation for transmission of embedded source-coded images," in *Proc. IEEE Int. Conf. Image Process.,* Oct. 2006, pp. 717–720.

[11] J. K. Cavers, "An analysis of pilot symbol assisted modulation for rayleigh fading channels," *IEEE Trans. Veh. Tech.,* vol. 40, no. 4, pp. 686–693, Nov. 1991.

[12] K. Lee and J. Chun, "Symbol detection in V-BLAST architectures under channel estimation errors," *IEEE Trans. Wireless Comm.,* vol. 6, no. 2, pp. 593–597, Feb. 2007.

[13] H. Jafarkhani, "A quasi-orthogonal space-time block code," *IEEE Commun. Letters,* vol. 49, no. 1, pp. 1–4, Jan. 2001.

[14] A. Paulraj, R. Nabar, and D. Gore, *Introduction to Space-Time Wireless Communications,* U.K. : Cambridge Univ. Press, 2003

[15] V. Stanković, R. Hamzaoui, and D. Saupe, "Fast algorithm for rate-based optimal error protection of embedded codes," *IEEE Trans. Commun.,* vol. 51, no. 11, pp. 1788–1795, Nov. 2003.

[16] J. Wu and J. Cai, "Rate-distortion analysis of leaky prediction based FGS video for constant quality constrained rate adaptation," *J. Vis. Comm. Image Representation,* vol. 18/1, pp. 45–58, Feb. 2007.

[17] J. Sun, W. Gao, D. Zhao, and Q. Huang, "Statistical model, analysis and approximation of rate-distortion function in MPEG-4 FGS videos," *IEEE Trans. CSVT.,* vol. 16, no. 4, pp. 535–539, Apr. 2006.

[18] H. Lee, R. W. Heath, and E. J. Powers, "Information outage probability and diversity order of alamouti transmit diversity in time-selective fading channels," *IEEE Trans. VT.,* vol. 57, no. 6, pp. 3890–3895, Nov. 2008.

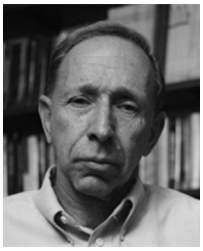
Hobin Kim received the B.S. degree in electrical engineering from Hanyang University, Seoul, Korea, in 2002, the M.S. degree in electrical engineering from the University of Michigan, Ann Arbor, in 2004, and the Ph.D. degree in electrical engineering from the University of California, San Diego, in 2011.

From January 2005 to June 2006, he was with LG Electronics Mobile Research, San Diego, CA, where he was focused on 1x EV-DO standardization. Since June 2011, he has been with Qualcomm, Inc., San Diego, as a System Engineer. His current research interests include channel estimation, multiuser MIMO, beamforming and joint source-channel coding.

Pamela C. Cosman (S'88–M'93–SM'00–F'08) received the B.S. degree (Hons.) 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 an NSF Post-Doctoral Fellow with Stanford University and a Visiting Professor with the University of Minnesota, Minneapolis, from 1993 to 1995. In 1995, she joined the Faculty of the Department of Electrical and Computer Engineering, University of California, San Diego, where she is currently a Professor and the Vice Chair. She was the Director of the Center for Wireless Communications, University of California at San Diego, La Jolla, from 2006 to 2008. Her current research interests include image and video compression and processing, and wireless communications.

Dr. Cosman was a recipient of the ECE Departmental Graduate Teaching Award, a Career Award from the National Science Foundation, a Powell Faculty Fellowship, and a Globecom 2008 Best Paper Award. 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 has been an Associate Editor of the IEEE COMMUNICATIONS LETTERS from 1998 to 2001, and an Associate Editor of the IEEE SIGNAL PROCESSING LETTERS from 2001 to 2005. She was the Editor-in-Chief from 2006 to 2009 as well as a Senior Editor from 2003 to 2005, and 2010-present of the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS. She is a member of Tau Beta Pi and Sigma Xi.



Laurence B. Milstein (S'66–M'68–SM'77-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 Communications Group, Hughes Aircraft Company, Culver, CA, 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, San Diego, La Jolla, where he is the Ericsson Professor of Wireless Communications Access Techniques and the 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 with both the government and the industry in the areas of radar and communications.

Dr. Milstein was an Associate Editor of communication theory for the IEEE TRANSACTIONS ON COMMUNICATIONS, an Associate Editor of 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 AREAS IN COMMUNICATIONS. He was the Vice President for Technical Affairs in 1990 and 1991 of the IEEE Communications Society, and is a Former Chair of the IEEE Fellows Selection Committee. He was a recipient of the 1998 Military Communications Conference Long Term Technical Achievement Award, the Academic Senate 1999 UCSD Distinguished Teaching Award, the IEEE Third Millennium Medal in 2000, the 2000 IEEE Communication Society Armstrong Technical Achievement Award, and various Prize Paper Awards, including the 2002 MILCOM Fred Ellersick Award.