Layered Video Transmission over Multirate DS-CDMA Wireless Systems

Lisimachos P. Kondi, Deepika Srinivasan, Dimitrios A. Pados, Stella N. Batalama

Dept. of Electrical Engineering, SUNY at Buffalo, Buffalo, NY 14260

ABSTRACT

In this paper, we consider the transmission of video over wireless direct-sequence code-division multiple access (DS-CDMA) channels. A layered (scalable) video source codec is used and each layer is transmitted over a different CDMA channel. Spreading codes with different lengths are allowed for each CDMA channel (multirate CDMA). Thus, a different number of chips per bit can be used for the transmission of each scalable layer. For a given fixed energy value per chip and chip rate, the selection of a spreading code length affects the transmitted energy per bit and bit rate for each scalable layer. An MPEG-4 source encoder is used to provide a two-layer SNR scalable bitstream. Each of the two layers is channel-coded using Rate-Compatible Punctured Convolutional (RCPC) codes. Then, the data are interleaved, spread, carrier-modulated and transmitted over the wireless channel. A multipath Rayleigh fading channel is assumed. At the other end, we assume the presence of an antenna array receiver. After carrier demodulation, multiple-access-interference suppressing despreading is performed using space-time auxiliary vector (AV) filtering. The choice of the AV receiver is dictated by realistic channel fading rates that limit the data record available for receiver adaptation and redesign. Indeed, AV filter short-data-record estimators have been shown to exhibit superior bit-error-rate performance in comparison with LMS, RLS, SMI, or "multistage nested Wiener" adaptive filter implementations. Our experimental results demonstrate the effectiveness of multirate DS-CDMA systems for wireless video transmission.

Keywords: Wireless Video Transmission, DS-CDMA, scalable video, multiple-access systems

1. INTRODUCTION

During the past few years there has been an increasing interest in multimedia communications over different types of channels. A significant amount of research has been focused on multimedia transmission over wireless channels. In this paper we consider the problem of operational rate-distortion optimal multirate scalable video transmission over wireless DS-CDMA multipath fading channels.

In a compressed video bit stream the various parts of the bit stream are not equally important to the quality of the decoded video sequence. Thus, instead of protecting them equally, it would be advantageous to protect the most important parts of the bit stream more than the less important parts. This is the idea of data partitioning and unequal error protection (UEP). In this work we apply UEP to the layers of a scalable bit stream. We accomplish UEP through the use of channel codes of variable rates and spreading codes of variable lengths.

The break-up of the bit stream into subsets of varying importance using a scalable codec lends itself naturally to employing an unequal error protection scheme. The base layer is typically better protected than the enhancement layers. This allows for added degrees of freedom in selecting the rates that will minimize the overall distortion. In ref. 1, the benefits of using scalability in an error prone environment are demonstrated by examining all scalability modes supported by MPEG-2 in an ATM network.

In Refs. 2 and 3, we assumed video transmission from one transmitter to one receiver using binary phase-shift keying (BPSK) modulation. The channel model was a non frequency selective Rayleigh fading channel. In Ref. 4, we assumed video transmission over a direct-sequence code-division-multiple-access (DS-CDMA) system. The channel model that we considered was a frequency-selective (multipath) Rayleigh fading channel. At the receiver, we employed an adaptive antenna array auxiliary-vector (AV) linear filter that provides space-time

Correspondence Email: lkondi@eng.buffalo.edu

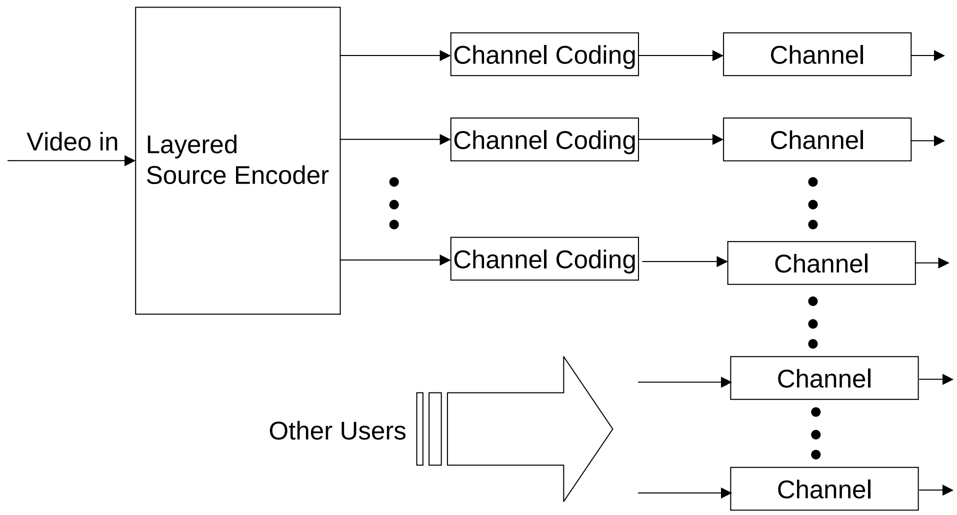

Figure 1: Multilayer wireless video transmission system.

RAKE-type processing (thus taking advantage of the multipath characteristics of the channel) and multiple-access interference suppression. The choice of the AV receiver was dictated by realistic channel fading rates that limit the data record available for receiver adaptation and redesign. Indeed, AV filter short-data-record estimators have been shown to exhibit superior bit-error-rate performance in comparison with LMS, RLS, SMI adaptive or multistage nested Wiener filter implementations.[5–8]

In this paper, we extend our previous results on DS-CDMA video transmission by allowing a single video user to utilize more than one CDMA channels. Thus, each layer of the scalable bitstream is transmitted over a separate CMDA channel. The proposed system is shown in Fig. 1. Each video scalable layer is individually channel-coded and transmitted over a separate CDMA channel. As shown in the figure, other users are simultaneously transmitting date (video, audio or speech). In this paper, we compare scalable video transmission over a single DS-CDMA channel with transmission over two DS-CDMA channels. In both cases, an operational rate-distortion problem is defined and solved in order to optimally select the source coding rates, channel coding rates and spreading code lengths (processing gains) used for the transmission.

The rest of the paper is organized as follows. In Section 2, we describe the elements of the video transmission system, i.e., scalable video coding (Section 2.1), channel encoding (Section 2.2), received signal (Section 2.3), and auxiliary-vector filtering (Section 2.4). In Section 3, the joint source coding optimization algorithm is described. In Section 4, experimental results are presented. Finally, in Section 5, conclusions are drawn.

2. VIDEO TRANSMISSION SYSTEM

2.1. Scalable Video Coding

A scalable video codec produces a bit stream which can be divided into embedded subsets. The subsets can be independently decoded to provide video sequences of increasing quality. Thus, a single compression operation can produce bit streams with different rates and reconstructed quality. A subset of the original bit stream can be initially transmitted to provide a base layer quality with extra layers subsequently transmitted as enhancement layers.

There are three main types of scalability: signal-to-noise ratio (SNR), spatial, and temporal. In SNR scalability, the enhancement in quality translates in an increase in the SNR of the reconstructed video sequence, while in spatial and temporal scalability the spatial and temporal resolution, respectively, is increased. In this paper, we use the MPEG-4[9] video compression standard to obtain SNR scalability.

2.2. Channel Coding

Rate-Compatible Punctured Convolutional (RCPC) codes for channel coding are used in this work. Convolutional coding is accomplished by convolving the source data with a convolutional matrix, G. In essence rather than having a number of channel coded symbols for the corresponding block of source codes as in linear block codes, convolutional coding generates one codeword for the entire source data. Convolution is the process of modulo-2 addition of the source bit with previously delayed source bits where the generator matrix specifies which delayed inputs to add with the current input. This is equivalent to passing the input data through a linear finite-state register where the tap connections are defined by the generator matrix. The rate of a convolutional code is defined as k/n where k is the number of input bits and n is the number of output bits.

Decoding convolutional codes is most commonly done using the Viterbi algorithm,[10] which is a maximum-likelihood sequence estimation algorithm. There are two types of Viterbi decoding, soft and hard decoding. In soft decoding, the output of the square-law detector is the input into the decoder and the distortion metric used is typically the Euclidean distance. In hard decoding a decision as to the received bit is made before the received data are input into the Viterbi decoder. The distortion metric commonly used in this case is the Hamming distance.

If a non fixed rate code is desired, then a higher rate code can be obtained by puncturing the output of the code.[11] Puncturing is the process of removing, or deleting, bits from the output sequence in a predefined manner so that fewer bits are transmitted than in the original coder leading to a higher coding rate. The idea of puncturing was extended to include the concept of rate compatibility.[12] Rate compatibility requires that a higher rate code be a subset of a lower rate code, or that lower protection codes be embedded into higher protection codes. This is accomplished by puncturing a "mother" code of rate $1/n$ to achieve higher rates. Rate compatibility represents a "natural way" to apply unequal error protection to the layers of a scalable bitstream. If a high rate code is not powerful enough to protect from channel errors, lower rates can be used by transmitting only the extra bits.

2.3. Received Signal

The time-varying multipath characteristics of mobile radio communications suggest a channel model that undergoes a process known as fading. The baseband received signal at each antenna element m, $m = 1, \ldots, M$, is viewed as the aggregate of the multipath received SS signal of interest with signature code $\mathbf{s}_0$ of length L (if T is the symbol period and T_c is the chip period then $L = T/T_c$), $K-1$ multipath received DS-SS interferers with unknown signatures $\mathbf{s}_k$, $k = 1, \ldots, K-1$, and white Gaussian noise. For notational simplicity and without loss of generality, we choose a chip-synchronous signal set-up. We assume that the multipath spread is of the order of a few chip intervals, P, and since the signal is bandlimited to $B = 1/2T_c$ the lowpass channel can be represented as a tapped delay line with $P+1$ taps spaced at chip intervals T_c. After conventional chip-matched filtering and sampling at the chip rate over a multipath extended symbol interval of $L+P$ chips, the $L+P$ data samples from the mth antenna element, $m = 1, \ldots, M$, are organized in the form of a vector $\mathbf{r}_m$ given by

$$\mathbf{r}_m = \sum_{k=0}^{K-1} \sum_{p=0}^{P} c_{k,p} \sqrt{E_k} (b_k \mathbf{s}_{k,p} + b_k^- \mathbf{s}_{k,p}^- + b_k^+ \mathbf{s}_{k,p}^+) \mathbf{a}_{k,p}[\mathrm{m}] + \mathbf{n} \ , \ m = 1, \ldots, M, \tag{1}$$

where, with respect to the kth SS signal, E_k is the transmitted energy, b_k, b_k^-, and b_k^+ are the present, the previous, and the following transmitted bit, respectively, and $\{c_{k,p}\}$ are the coefficients of the frequency-selective slowly fading channel modeled as independent zero-mean complex Gaussian random variables that are assumed to remain constant over several symbol intervals. $\mathbf{s}_{k,p}$ represents the 0-padded by P, p-cyclic-shifted version of the signature of the kth SS signal $\mathbf{s}_k$, $\mathbf{s}_{k,p}^-$ is the 0-filled $(L-p)$-left-shifted version of $\mathbf{s}_{k,0}$, and $\mathbf{s}_{k,p}^+$ is the 0-filled $(L-p)$-right-shifted version of $\mathbf{s}_{k,0}$. Finally, $\mathbf{n}$ represents additive complex Gaussian noise and $\mathbf{a}_{k,p}[\mathrm{m}]$ is the mth coordinate of the kth SS signal, pth path, array response vector:

$$\mathbf{a}_{k,p}[\mathrm{m}] = e^{j2\pi(m-1)\frac{\sin\theta_{k,p} d}{\lambda}} \ , \ m = 1, \ldots, M, \tag{2}$$

where $\theta_{k,p}$ identifies the angle of arrival of the pth path of the kth SS signal, λ is the carrier wavelength, and d is the element spacing (usually $d = \lambda/2$).

To avoid in the sequel cumbersome 2-D data notation and filtering operations, we decide at this point to "vectorize" the $(L+P) \times M$ space-time data matrix $[\mathbf{r}_1 \ \mathbf{r}_2 \ \cdots \ \mathbf{r}_M]$ by sequencing all matrix columns in the form of a single $(L+P)M$-long column vector:

$$\mathbf{r}_{(\mathrm{L+P})\mathrm{M}\times 1} = Vec\{[\mathbf{r}_1, \ \mathbf{r}_2, \ \cdots, \ \mathbf{r}_M]_{(\mathrm{L+P})\times\mathrm{M}}\} \ . \tag{3}$$

From now on, $\mathbf{r}$ denotes the joint space-time data in the $\mathcal{C}^{(L+P)M}$ complex vector domain.

For conceptual and notational simplicity we may rewrite the vectorized space-time data equation as follows:

$$\mathbf{r} = \sqrt{E_0} b_0 \mathbf{w}_{\text{R-MF}} + \mathbf{i} + \mathbf{n} \tag{4}$$

where $\mathbf{w}_{\text{R-MF}} = E_{b_0}\{\mathbf{r}\, b_0\} = Vec\{[\ \sum_{p=0}^{P} c_{0,p}\mathbf{s}_{0,p}, \mathbf{a}_{0,p}[1], \ \cdots, \ \cdots, \ \sum_{p=0}^{P} c_{0,p}\mathbf{s}_{0,p}\mathbf{a}_{0,p}[\mathrm{M}]\ ]\}$ is the effective space-time signature of the SS signal of interest (signal *0*) and $\mathbf{i}$ identifies comprehensively both the inter-symbol and the SS interference present in $\mathbf{r}$ ($E_{b_0}\{\cdot\}$ denotes statistical expectation with respect to b_0). We use the subscript R-MF in our effective S-T signature notation to make a direct association with the RAKE matched-filter time-domain receiver that is known to correlate the signature $\mathbf{s_0}$ with $P+1$ size-L shifted windows of the received signal (that correspond to the $P+1$ paths of the channel), appropriately weighted by the conjugated channel coefficients $c_{0,p}, p = 0, \ldots, P$. In our notation, the generalized S-T RAKE operation corresponds to linear filtering of the form $\mathbf{w}_{\text{R-MF}}^{\mathrm{H}}\mathbf{r}$, where H denotes the Hermitian operation.

In this paper, we consider *variable sequence length multirate CDMA*. In variable sequence length multirate CDMA, the spreading sequences of different CDMA channels can have different lengths. Let us assume, for example, two CDMA channels: Channel 1 with a spreading code of length $L = 16$ and channel 2 with a spreading code of length $L = 32$. Assuming equal energy per chip for the two channels, channel 2 transmits twice as many chips per bit (and energy per bit) than channel 1. However, channel 1 transmits with twice the bit rate as channel 2. Thus, channel 2 exhibits a lower bit error rate than channel 1, but channel 1 allows for transmission at a higher data rate.

2.4. Auxiliary-Vector Adaptive Filtering

The AV algorithm generates an infinite sequence of filters $\{\mathbf{w}_n\}_{n=0}^{\infty}$. The sequence is initialized at the S-T RAKE filter

$$\mathbf{w}_0 = \frac{\mathbf{w}_{\text{R-MF}}}{\|\mathbf{w}_{\text{R-MF}}\|^2}, \tag{5}$$

which is here scaled to satisfy $\mathbf{w}_0^H \mathbf{w}_{\text{R-MF}} = 1$. At each step $k+1$ of the algorithm, $k = 0, 1, 2, \ldots$, we incorporate in $\mathbf{w}_k$ an "auxiliary" vector component $\mathbf{g}_{k+1}$ that is orthogonal to $\mathbf{w}_{\text{R-MF}}$ and weighted by a scalar μ_{k+1} and we form the next filter in the sequence

$$\mathbf{w}_{k+1} = \mathbf{w}_k - \mu_{k+1}\mathbf{g}_{k+1}. \tag{6}$$

The auxiliary vector $\mathbf{g}_{k+1}$ is chosen to maximize, under fixed norm, the magnitude of the cross-correlation between its output $\mathbf{g}_{k+1}^H\mathbf{r}$ and the previous filter output $\mathbf{w}_k^H\mathbf{r}$ and is given by

$$\mathbf{g}_{k+1} = \mathbf{R}\mathbf{w}_k - \frac{\mathbf{w}_{\text{R-MF}}^H \mathbf{R}\mathbf{w}_k}{\|\mathbf{w}_{\text{R-MF}}\|^2}\mathbf{w}_{\text{R-MF}} \tag{7}$$

where R is the input autocorrelation matrix, $R = E\{\mathbf{r}\mathbf{r}^H\}$. The scalar μ_{k+1} is selected such that it minimizes the output variance of the filter $\mathbf{w}_{k+1}$ or equivalently minimizes the MS error between $\mathbf{w}_k^H\mathbf{r}$ and $\mu_{k+1}^*\mathbf{g}_{k+1}^H\mathbf{r}$. The MS-optimum μ_{k+1} is

$$\mu_{k+1} = \frac{\mathbf{g}_{k+1}^H \mathbf{R}\mathbf{w}_k}{\mathbf{g}_{k+1}^H \mathbf{R}\mathbf{g}_{k+1}}. \tag{8}$$

The AV filter recursion is completely defined by (5)-(8). Theoretical analysis of the AV algorithm was pursued in Ref. 7. The results are summarized below in the form of a theorem.

THEOREM 1. *Let* $\mathbf{R}$ *be a Hermitian positive definite matrix. Consider the iterative algorithm of eqs. (5)-(8).*
(i) *Successive auxiliary vectors generated through (6)-(8) are orthogonal:* $\mathbf{g}_i^H\mathbf{g}_{i+1}=0$, $i=1,2,3,\ldots$, *(however, in general* $\mathbf{g}_i^H\mathbf{g}_j \neq 0$ *for* $|i-j|\neq 1$*).*
(ii) *The generated sequence of auxiliary-vector weights* $\{\mu_n\}$, $n=1,2,\ldots$, *is real-valued, positive, and bounded:* $0<\frac{1}{\lambda_{\max}}\leq\mu_n\leq\frac{1}{\lambda_{\min}}$, $n=1,2,\ldots$, *where* $\lambda_{\max}$ *and* $\lambda_{\min}$ *are the maximum and minimum, correspondingly, eigenvalues of* $\mathbf{R}$.
(iii) *The sequence of auxiliary vectors* $\{\mathbf{g}_n\}$, $n=1,2,\ldots$, *converges to the* $\mathbf{0}$ *vector:* $\lim_{n\to\infty}\mathbf{g}_n=\mathbf{0}$.
(iv) *The sequence of auxiliary-vector filters* $\{\mathbf{w}_n\}$, $n=1,2,\ldots$, *converges to the MVDR filter:* $\lim_{n\to\infty}\mathbf{w}_n=\frac{\mathbf{R}^{-1}\mathbf{w}_{R\text{-}MF}}{\mathbf{w}_{R\text{-}MF}^H\mathbf{R}^{-1}\mathbf{w}_{R\text{-}MF}}$. □

If $\mathbf{R}$ is unknown and sample-average estimated from a packet data record of D points, then Theorem 1 shows that

$$\hat{\mathbf{w}}_n(D)\underset{n\to\infty}{\longrightarrow}\hat{\mathbf{w}}_\infty(D)=\frac{\left[\hat{\mathbf{R}}(D)\right]^{-1}\mathbf{w}_{\text{R-MF}}}{\mathbf{w}_{\text{R-MF}}^H\left[\hat{\mathbf{R}}(D)\right]^{-1}\mathbf{w}_{\text{R-MF}}} \tag{9}$$

where $\hat{\mathbf{w}}_\infty(D)$ is the widely used MVDR filter estimator known as the sample-matrix-inversion (SMI) filter.[13] The output sequence begins from $\hat{\mathbf{w}}_0(D)=\frac{\mathbf{w}_{\text{R-MF}}}{\|\mathbf{w}_{\text{R-MF}}\|^2}$, which is a *0*-variance, fixed-valued, estimator that may be severely biased ($\hat{\mathbf{w}}_0(D)=\frac{\mathbf{w}_{\text{R-MF}}}{\|\mathbf{w}_{\text{R-MF}}\|^2}\neq\mathbf{w}_{\text{MVDR}}$) unless $\mathbf{R}=\sigma^2\mathbf{I}$ for some $\sigma>0$. In the latter trivial case, $\hat{\mathbf{w}}_0(D)$ is already the perfect MVDR filter. Otherwise, the next filter estimator in the sequence, $\hat{\mathbf{w}}_1(D)$, has a significantly reduced bias due to the optimization procedure employed at the expense of non-zero estimator (co-)variance. As we move up in the sequence of filter estimators $\hat{\mathbf{w}}_n(D)$, $n=0,1,2,\ldots$, the bias decreases rapidly to zero while the variance rises slowly to the SMI ($\hat{\mathbf{w}}_\infty(D)$) levels (cf. (9)).

An adaptive data-dependent procedure for the selection of the most appropriate member of the AV filter estimator sequence $\{\hat{\mathbf{w}}_n(D)\}$ for a given data record of size D is presented in Ref. 14 and Ref. 15. The procedure selects the estimator $\hat{\mathbf{w}}_{n_0}$ from the generated sequence of AV filter estimators that exhibits maximum J-divergence between the filter output conditional distributions given that $+1$ or -1 is transmitted. Under a Gaussian approximation on the conditional filter output distribution, it was shown in Refs. 14, 15 that the J-divergence of the filter estimator with n auxiliary vectors is

$$J(n)=\frac{4\,E^2\left\{b_0\text{Re}\left[\hat{\mathbf{w}}_n^H(D)\mathbf{r}\right]\right\}}{Var\left\{b_0\text{Re}\left[\hat{\mathbf{w}}_n^H(D)\mathbf{r}\right]\right\}}. \tag{10}$$

To estimate the J-divergence $J(n)$ from the data packet of size D, the transmitted information bits b_0 are required to be known. We can obtain a blind *approximate* version of $J(n)$ by substituting the information bit b_0 in (10) by the detected bit $\hat{b}_0=\text{sgn}\left(\text{Re}\left\{\hat{\mathbf{w}}_n^H(D)\mathbf{r}\right\}\right)$ (output of the sign detector that follows the linear filter). In particular, using $\hat{b}_0$ in place of b_0 in (10) we obtain the following J-divergence expression:

$$J_B(n)=\frac{4\,E^2\left\{\hat{b}_0\text{Re}\left[\hat{\mathbf{w}}_n^H(D)\mathbf{r}\right]\right\}}{Var\left\{\hat{b}_0\text{Re}\left[\hat{\mathbf{w}}_n^H(D)\mathbf{r}\right]\right\}}=\frac{4\,E^2\left\{\left|\text{Re}\left[\hat{\mathbf{w}}_n^H(D)\mathbf{r}\right]\right|\right\}}{Var\left\{\left|\text{Re}\left[\hat{\mathbf{w}}_n^H(D)\mathbf{r}\right]\right|\right\}} \tag{11}$$

where the subscript "B" identifies the blind version of the J-divergence function. To estimate $J_B(n)$ from the data packet of size D, we substitute the statistical expectations in (11) by sample averages. The following criterion summarizes the corresponding AV filter estimator selection rule.

CRITERION 1. *For a given data record of size* D, *the unsupervised (blind) J-divergence AV filter estimator* *selection rule chooses the estimator* $\hat{\mathbf{w}}_{n_0}(D)$ *with* n_0 *auxiliary vectors where*

$$n_0=\arg\max_n\left\{\hat{J}_B(n)\right\}=\arg\max_n\left\{\frac{4\left[\frac{1}{D}\sum_{i=0}^{D-1}\left|\mathrm{Re}\left[\hat{\mathbf{w}}_n^H(D)\mathbf{r}_i\right]\right|\right]^2}{\frac{1}{D}\sum_{i=0}^{D-1}\left|\mathrm{Re}\ \left[\hat{\mathbf{w}}_n^H(D)\mathbf{r}_i\right]\right|^2-\left[\frac{1}{D}\sum_{i=0}^{D-1}\left|\mathrm{Re}\left[\hat{\mathbf{w}}_n^H(D)\mathbf{r}_i\right]\right|\right]^2}\right\}. \tag{12}$$

□

Criterion 1 completes the design of the joint S-T auxiliary-vector filter estimator.

3. OPTIMAL RESOURCE ALLOCATION

We next describe the the optimal resource allocation for the cases of video transmission over one or two CDMA channels. The optimization constraint in both cases is the available chip rate R_{budget}^{chip}. A fixed energy per chip is assumed.

3.1. Single CDMA Channel Case

If a single CDMA channel is used for the transmission of all scalable layers of a video user, the layers are time-multiplexed. The available transmission bit rate is

$$R_{budget}=\frac{R_{budget}^{chip}}{L}, \tag{13}$$

where L is the system processing gain (spreading code length). The available bit rate R_{budget} has to be allocated between scalable layers and, within each layer, between source and channel coding.

The formal statement of the problem that we are solving is as follows: Given an overall bit rate R_{budget}, we want to optimally allocate bits between source and channel coding such that the overall distortion D_{s+c} is minimized, that is,

$$\min D_{s+c} \text{ subject to } R_{s+c}\le R_{budget} \tag{14}$$

where R_{s+c} is the total bit rate used for source and channel coding for all layers and D_{s+c} is the resulting *expected* distortion which is due to both source coding errors and channel errors. The distortion that is caused by source coding is due to quantization and is deterministic. However, the distortion due to channel errors is stochastic. Thus, the total distortion is also stochastic and we use its expected value.

For K scalable layers, R_{s+c} is equal to

$$R_{s+c}=\sum_{l=1}^{K}R_{s+c,l} \tag{15}$$

where $R_{s+c,l}$ is the bit rate used for source and channel coding for the scalable layer l. It is equal to

$$R_{s+c,l}=\frac{R_{s,l}}{R_{c,l}}, \tag{16}$$

where $R_{s,l}$ and $R_{c,l}$ are the source and channel rates, respectively, for the scalable layer l. It should be emphasized that $R_{s,l}$ is in bits/s and $R_{c,l}$ is a dimensionless number.

The problem is a discrete optimization problem: $R_{s,l}$ and $R_{c,l}$ can only take values from discrete sets that are defined as part of the problem.

The optimization problem in (14) can be solved by Lagrangian optimization. To reduce the computational complexity of the solution, it is useful to write the overall distortion D_{s+c} as the sum of distortions per scalable layers:

$$D_{s+c}=\sum_{l=1}^{K}D_{s+c,l}. \tag{17}$$

In a subband-based scalable codec, it is straightforward to express the distortion as the sum of distortions per layer since each layer corresponds to different transform coefficients. However, in our scalable codec, we need to redefine distortion per layer as the *differential improvement* of including the layer in the reconstruction. Therefore, in the absence of channel errors, only the distortion for layer 1 (base layer) would be positive and the distortions for all other layers would be negative since inclusion of these layers reduces the mean squared error (MSE).

Another observation that should be made is that the differential improvement in MSE due to a given layer depends on the rates of the previous layers. For example, for a two layer case, an enhancement layer of 28 kbps will cause a different improvement in the MSE depending on the rate used for the base layer. The differential improvement depends on how good the picture quality was before the inclusion of the next scalable layer. Therefore, the distortion per layer is better expressed as

$$D_{s+c} = \sum_{l=1}^{K} D_{s+c,l}(R_{s+c,1}, \ldots, R_{s+c,l}). \tag{18}$$

Utilizing Lagrangian optimization, the constrained problem in (14) is transformed into the unconstrained problem of minimizing

$$J(\lambda) = D_{s+c} + \lambda R_{s+c}. \tag{19}$$

More information on the single CDMA channel video tramsmission optimization can be found in Ref. 2.

Our problem reduces to finding the Operational Rate-Distortion Functions (ORDF)$D_{s+c,l}(\cdot, ..., \cdot)$ for each scalable layer. One way to proceed is to experimentally obtain the expected distortion for each layer for all possible combinations of source and channel rates and all possible channel conditions. However, this may become prohibitively complex for even a small number of admissible source and channel rates and channel conditions. Thus, we choose to utilize the *universal rate-distortion characteristics* (URDC) at the expense of a slight deviation from the optimal performance. These characteristics show the expected distortion per layer as a function of the bit error rate (after channel coding). The use of the URDCs is discussed in Section 3.3.

3.2. Multiple CDMA Channel Case

We next discuss the case where each scalable video layer is transmitted over a separate DS-CDMA channel. In that case, the available bit rate for layer i is

$$R_{s+c,i} = \frac{R_{budget}^{chip}}{L_i}, \tag{20}$$

where L_i is the spreading length for layer i. Thus, if two layers are assumed, the ratio $R_{s+c,1}/R_{s+c,2}$ is fixed and equal to L_2/L_1. This is in contrast with the single CDMA channel case where the allocation of the available bit rate to each scalable layer is part of the optimization.

Our optimization problem is (for the case of K layers):

$$\min D_{s+c} \text{ subject to } L_i R_{s+c,i} \leq R_{budget}^{chip}, \text{ for } i = 1, \ldots, K \tag{21}$$

For each layer, we need to determine the source coding rate $R_{s,i}$, the channel coding rate $R_{c,i}$ and the spreading length L_i

As mentioned previously, the total bit rate allocated to a scalable layer depends only on L_i and not on any decisions made for another layer. Since $D_{s+c} = \sum_{l=1}^{K} D_{s+c,l}(R_{s+c,1}, \ldots, R_{s+c,l})$, our problem can be broken into separate problems for each layer, thus simplifying the optimization when compared to the single CDMA channel case. For the case of two layers, the two problems to be solved are:

$$\{R_{s,1}^*, R_{c,1}^*, L_1^*\} = \arg\min D_{s+c,1}(R_{s+c,1}) \text{ subject to } L_1 R_{s+c,1} \leq R_{budget}^{chip}, \tag{22}$$

and

$$\{R^*_{s,2}, R^*_{c,2}, L^*_2\} = \arg\min D_{s+c,2}(R^*_{s+c,1}, R_{s+c,2}) \text{ subject to } L_2 R_{s+c,2} \leq R^{chip}_{budget}. \tag{23}$$

Thus, the optimization for two-channel video transmission case is simpler than the optimization in the single-channel case. Since there is no dependency between layers, the above two problems can be solved independently using Lagrangian optimization. As in the single-channel case, we first need to obtain the rate-distortion functions $D_{s+c,l}(.), l = 1, 2$. Those can be obtained using Universal Rate-Distortion Characteristics, as explained next.

3.3. Universal Rate-Distortion Characteristics

So far, we have formulated the problem and outlined the solution for the case of dependent layer distortions assuming that the optimal rate-distortion characteristics of the individual layers, $D^*_{s+c,l}(.)$, are given. We now describe the technique used to obtain the $D^*_{s+c,l}(.), l = 1, \ldots, K$.

While it is possible to simulate transmission of the actual source coded data over a channel, gather statistics and develop an optimal strategy based on these results, in practice this leads to extremely high computational complexity and makes the process impractical in many ways. For every bitstream, we will have to simulate transmission of the data over the channel using all combinations of source and channel coding rates per scalable layer and for channel conditions of interest. Clearly, the computational complexity of this approach is prohibitive. To circumvent this problem, *Universal* Rate-Distortion Characteristics of the source coding scheme are utilized.[2,16,17] This approach is described next.

For given channel SNRs, spreading lengths and choice of channel codes, the probability of bit error, P_b, is calculated for the set of channel coding rates of interest. It establishes a reference as to the performance of the channel coding over the particular channel with the given parameters. Furthermore this channel performance analysis needs to be done only once. An example plot showing the performance of the channel coding as a function of a given channel parameter is shown in Fig. 2 for a set of channel coding rates, R_c. We will call these plots *Channel Characteristic Plots.*

Towards calculating the impact of the errors due to both source coding and channel transmission on a set of data, it is realized that for a given set of preceding layer source rates, the distortion for a particular layer, $D_{s+c,l}$, given a particular source coding rate, $R_{s,l}$, is a function of the bit error rate. Thus the rate-distortion function of the layer, (given the preceding layer source rates) for a fixed source rate, $R_{s,l}$, is a function of the bit error rate (after channel decoding), P_b. It is then possible to plot a family of $D_{s+c,l}$ vs. $1/P_b$ curves given a set of source coding rates of interest as shown in Fig. 3. These are defined as the *Universal Rate-Distortion Characteristics* (URDCs) of the source. Due to the use of Variable Length Codes in the video coding standards, it would be extremely difficult if not impossible to analytically obtain the URDCs. Thus, the URDCs are obtained experimentally using simulations. To obtain the URDC for $D_{s+c,l}$, layer l of the bitstream is corrupted with independent errors with bit error rate P_b. Layers $1, \ldots, l-1$ are not corrupted. The bitstream is then decoded and the Mean Squared Error is computed. The experiment is repeated many times (in our case, 30 times). If $l > 1$, i.e., we are calculating the URDC for an enhancement layer, we need to subtract the distortion of the first $l-1$ uncorrupted layers, since $D_{s+c,l}$ in this case is the differential improvement of including layer l, as mentioned earlier.

Using the Channel Characteristic Plots and the Universal Rate-Distortion Characteristics, Operational Rate-Distortion Functions for each scalable layer are constructed as follows. First, for the given channel parameters, we use the Channel Characteristic Plot to determine the resulting bit error rates for each of the available channel coding rates. Then, for each of these probability values, we use the Universal Rate-Distortion Characteristics to obtain the resulting distortion for each available source coding rates. By also obtaining the total rate R_{s+c} for each combination of source and channel codes, we have the Rate-Distortion operating points for the given channel conditions.

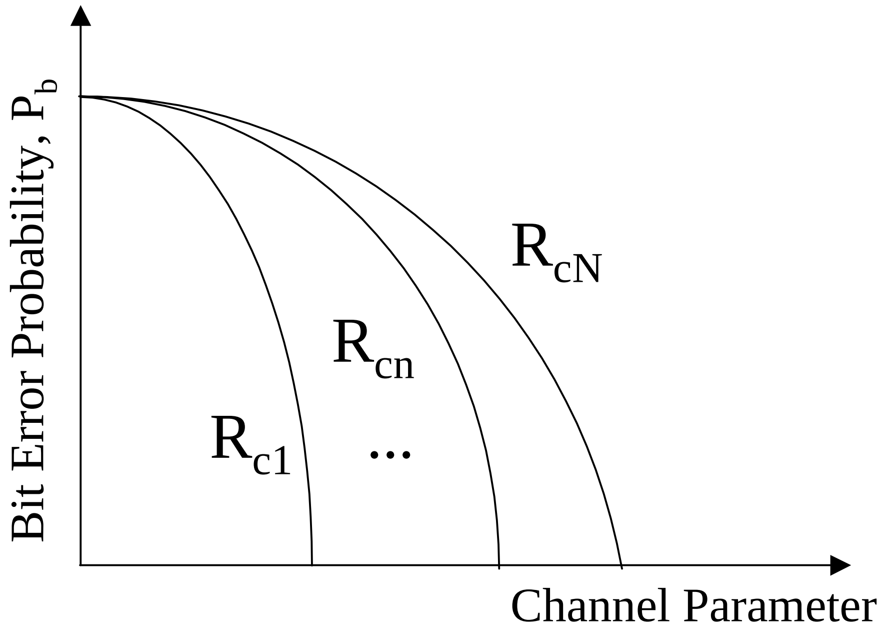

Figure 2. A typical channel characteristic plot, i.e., channel parameters vs. bit error rate for a set of channel coding rates R_{cn} with $n = 1, \ldots, N$.

4. EXPERIMENTAL RESULTS

We next present experimental results that compare scalable video transmission over one and two DS-CDMA channels. An MPEG-4 compatible video source codec was used, along with Rate Compatible Punctured Convolutional (RCPC) channel codes from Ref. 12. Auxiliary-Vector filtering followed by soft decision Viterbi decoding was used at the receiver. For the single CDMA channel case, 10 channels were assumed (the user of interest and nine interferers). All channels had an SNR of 15 dB. For the dual CDMA channel case, 10 channels were also assumed. Two channels were used by the video user of interest and had an SNR of 11.98dB. The eight interferers had an SNR of 15 dB. Thus, the video user was transmitting the same power in the one and two channel case, in order to have a fair comparison. In the two channel case, each video user channel was transmitting at half the power than the single channel.

For both the single and dual CDMA channel cases, the admissible source coding rates were 28000, 42000 and 56000 bits per second for the base layer and 14000, 28000, 42000 and 56000 bits per second for the enhancement layer. The admissible channel coding rates for both layers were 1/2, 2/3 and 4/5 while the admissible spreading code lengths were 16 and 32.

Figure 4 shows a comparison of the performance of scalable video transmission over a DS-CDMA system when the video user of interest occupies one or two CDMA channels. The Mean Squared Error is plotted against the total chip rate. It can be seen that, for the SNRs under consideration, the two-channel case outperforms the one-channel case for all chip rates except 1680 and 2016 kcps. However, as can be seen in the figure, there were no optimal rate-distortion points close to these rate for the two-channel case. If a different selection of admissible source coding rates is used, it is possible that the two-channel case will outperform the one-channel case for all chip rates. However, it is possible that single channel video transmission can perform better under different channel conditions.

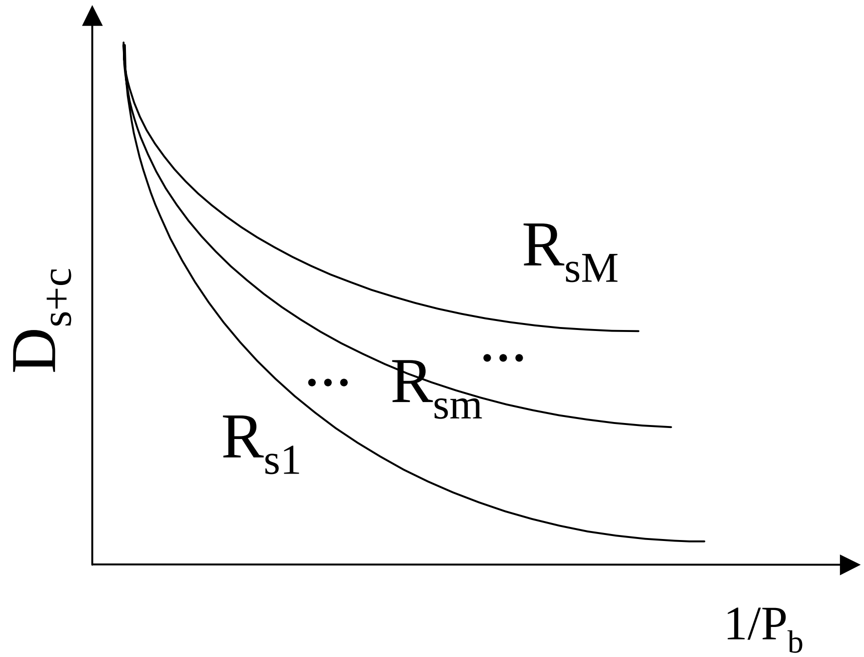

Figure 3. Universal rate-distortion characteristics of the source; typical end-to-end distortion vs. bit error rate response characterizing the source and channel for a set of source coding rates, R_{sm} for $m = 1, \ldots, M$.

5. CONCLUSIONS

We have presented a multiple-access video communications scheme that is based on DS-CDMA. Each video user is allowed to occupy one or two CDMA channels. We have presented resource optimization schemes for both cases. Our experimental results show that, for the particular SNRs chosen, two-channel video transmission outperforms single-channel video transmission for most chip rates.

REFERENCES

1. R. Aravind, M. R. Civanlar, and A. R. Reibman, "Packet loss resilience of MPEG-2 scalable video coding algorithms," *IEEE Trans. on Circuits and Systems for Video Technology* **6**, pp. 426–435, Oct. 1996.
2. L. P. Kondi, F. Ishtiaq, and A. K. Katsaggelos, "Joint source-channel coding for snr scalable video," *IEEE Transactions on Image Processing* **11**, pp. 1043–1054, Sept. 2002.
3. L. P. Kondi, F. Ishtiaq, and A. K. Katsaggelos, "Joint source-channel coding for scalable video," in *Proceedings of the SPIE Conference on Image and Video Communications and Processing*, pp. 324–335, (San Jose, CA), 2000.
4. 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 *Proceedings of the IEEE International Conference on Image Processing*, (Thessaloniki, Greece), 2002.
5. A. Kansal, S. N. Batalama, and D. A. Pados, "Adaptive maximum SINR RAKE filtering for DS-CDMA multipath fading channels," *IEEE J. Select. Areas Commmun.* **16**, pp. 1765–1773, Dec. 1998. Special Issue on Signal Processing for Wireless Communications.
6. D. A. Pados and S. N. Batalama, "Joint space-time auxiliary-vector filtering for DS-CDMA systems with antenna arrays," *IEEE Trans. Commmun.* **47**, pp. 1406–1415, Sept. 1999.

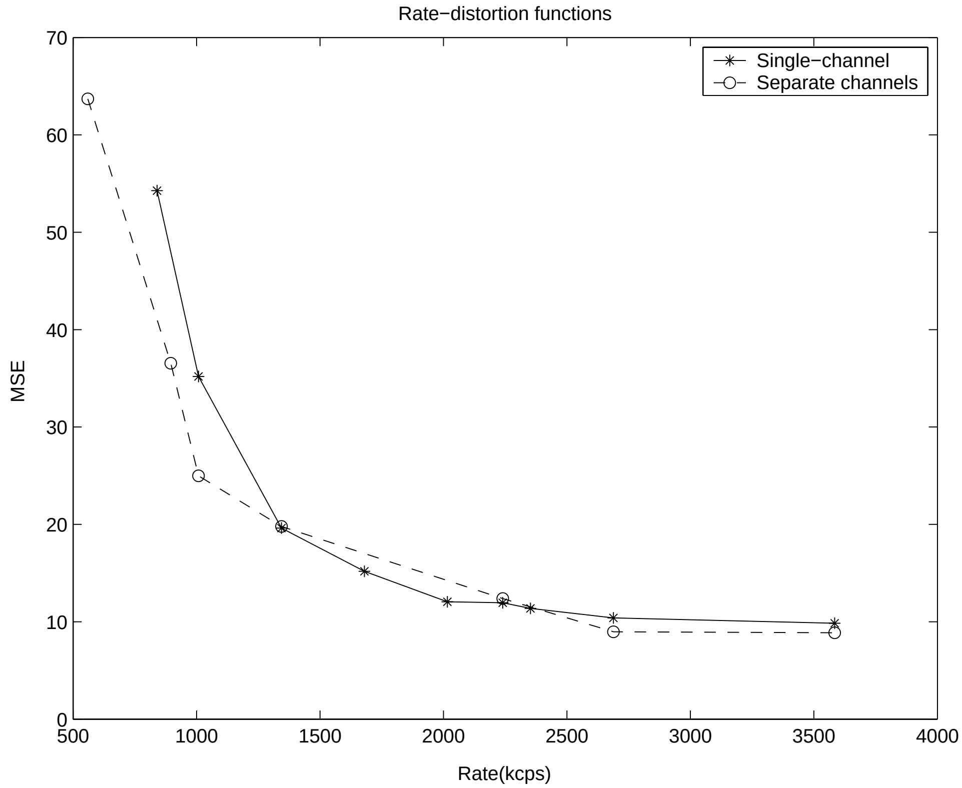

Figure 4: Rate-Distortion performance of scalable video transmission over a DS-CDMA wireless system.

7. D. A. Pados and G. N. Karystinos, "An iterative algorithm for the computation of the MVDR filter," *IEEE Trans. Signal Proc.* **49**, pp. 290–300, Feb. 2001.
8. J. S. Goldstein, I. S. Reed, and L. L. Sharf, "A multistage representation of the Wiener filter based on orthogonal projections," *IEEE Transactions on Information Theory* **44**, pp. 2943–2959, Nov. 1998.
9. ISO/IEC 14496-2 MPEG-4, "Information technology-coding of audio-visual objects: Visual," October. 1997.
10. G. D. Forney, "The Viterbi Algorithm," *Proceedings of the IEEE* **61**, pp. 268–278, March 1973.
11. J. B. Cain, G. C. Clark, and J. M. Geist, "Punctured convolutional codes of rate $(n-1)/n$ and simplified maximum likelihood decoding," *IEEE Transactions on Information Theory* **IT-25**, pp. 97–100, Jan. 1979.
12. J. Hagenauer, "Rate-compatible punctured convolutional codes (RCPC codes) and their applications," *IEEE Transactions on Communications* **36**, pp. 389–400, April 1988.
13. I. S. Reed, J. M. Mallet, and L. E. Brennan, "Rapid convergence rate in adaptive arrays," *IEEE Transactions on Aerospace and Electronic Systems* **10**, pp. 853–863, Nov. 1974.
14. H. Qian and S. N. Batalama, "Data-record-based criteria for the selection of an auxiliary-vector estimator of the MVDR filter," in *Proceedings of the Thirty Fourth Annual Asilomar Conference on Signals, Systems, and Computers*, pp. 802–807, (Pacific Grove, California), 2000.
15. H. Qian and S. N. Batalama, "Data-record-based criteria for the selection of an auxiliary-vector estimator of the MMSE/MVDR filter," *IEEE Transactions on Communications* . Submitted.
16. M. Bystrom and T. Stockhammer, "Modeling of operational distortion-rate characteristics for joint source-channel coding of video," in *Proceedings of the IEEE International Conference on Image Processing*, (Vancouver, Canada), 2000.

17. M. Bystrom and J. W. Modestino, "Combined source-channel coding for transmission of video over a slow-fading rician channel," in *Proceedings of the IEEE International Conference on Image Processing*, pp. 147–151, 1998.