Robust Image and Video Transmission Over Spectrally Shaped Channels Using Multicarrier Modulation

Haitao Zheng and K. J. Ray Liu, *Senior Member, IEEE*

***Abstract*— This paper presents a new parallel transmission framework for reliable multimedia data transmission over spectrally shaped channels using multicarrier modulation. We propose to transmit source data layers of different perceptual importance in parallel, each occupying a number of subchannels. New loading algorithms are developed to efficiently allocate the available resources, e.g., transmitted power and bit rate, to the subchannels according to the source layers they transmit. Instead of making the bit error rate of all the subchannels equal as in most existing loading algorithms, the proposed algorithm assigns different error performance to the subchannels to achieve unequal error protection for different layers. The channel induced distortion in mean-square sense is minimized. We show that the proposed system can be applied nicely to both fixed length coding and variable-length coding. Asymptotic gains with respect to channel distortion are also derived. Numerical examples show that the proposed algorithm achieves significant performance improvement compared to the existing work, especially for spectrally shaped channels commonly used in in ADSL systems.**

***Index Terms*—Combined source coding and channel modulation, entropy coding, layered coding, multicarrier modulation, multimedia transmission, spectrally shaped channel, subband coding.**

I. Introduction

RECENT developments in multimedia data coding and compression techniques permit real-time multimedia communications, such as video telephony and conferencing. Asymmetric digital subscriber lines (ADSL's) [1], a transmission system capable of realizing very high bit-rate services over existing telephone lines, are well suited for these types of applications. Typical channels in ADSL are spectrally shaped channels with characteristics shown in Fig. 1. Multicarrier modulation (MCM) [2], also referred to as orthogonal frequency division multiplexing (OFDM) or discrete multiton (DMT), is currently considered as a standard channel-coding scheme for ADSL. MCM has elegant waveform properties that make it useful for a wide variety of applications. In addition to ADSL, it is widely used in digital audio broadcasting [3], [4], high-rate digital high-definition television (HDTV), and wireless personal communication systems [1], [5]. By applying the discrete Fourier transform (DFT) or fast Fourier transform (FFT) and their inverse, the available channel bandwidth is subdivided into a number of subchannels that achieves bandwidth efficiency.

A crucial aspect in the design of MCM system is the need to optimize the system transmission bandwidth and power through an optimal loading algorithm. Each subchannel in MCM system has two variables: bit rate or modulation rate (number of bits per transmission, e.g., 4 for QAM16 and 6 for QAM64), and transmitted power. Since channel transfer function and noise are nonflat for spectrally shaped channels, assigning equal bit rate and power to all the subchannels is not optimal. The loading algorithms in literature can be divided into two categories. Category one computes bit rate and power distribution for given bit error rate (BER). "Water-filling" has been well known as a capacity achieving energy distribution, but it assumes infinite granularity and is often difficult to compute. Hughes-Hartog suggested a greedy algorithm which assigns the bits successively to the subchannels until the data throughput is reached. For ADSL applications this requires extensive computation. An improved version now known as Campello Algorithms has recently been developed in [6]. In [8], the bit distribution is computed by rounding of approximate water-fill results. Using efficient lookup table searches and a Lagrange multiplier bisection search, an optimal power allocation algorithm for given BER is proposed in [9]. Category two is based on minimizing the overall BER while reaching the data throughput under a power constraint. The overall BER is unknown until the final stage of optimization. In [7], considering the fact that transmitted power and bit rate are directly related, a loading algorithm assigns high bit rate to subchannels with high channel gain-to-noise ratio (CGNR) and low bit rate to subchannels with low CGNR. The subchannels with very low CGNR are not even used. These algorithms are all constrained by the assumption that BER's are equal across the usable subchannels. In [10], the assumption of uniform performance is eliminated. It concludes that the uniform BER assumption is, at best, approximate, which is fair only at high SNR and poor at low and intermediate SNR.

To transmit multimedia data, such as image, video, and audio through noisy channels, BER need not be very low to achieve adequate quality. Usually, the BER may range from 10^{-5} to 10^{-1}, and the channel-induced distortion cannot be ignored. Combined source/channel coding approaches in

Manuscript received August 26, 1998; revised November 18, 1998. This work was supported in part by NSF NYI Award MIP9457397. The associate editor coordinating the review of this paper and approving it for publication was Dr. Raj. K. Talluri.

The authors are with the Institute for Systems Research and Electrical Engineering Department, University of Maryland, College Park, MD 20742 USA (e-mail: haitaoz@eng.umd.edu; kjrliu@eng.umd.edu).

Publisher Item Identifier S 1520-9210(99)01584-9.

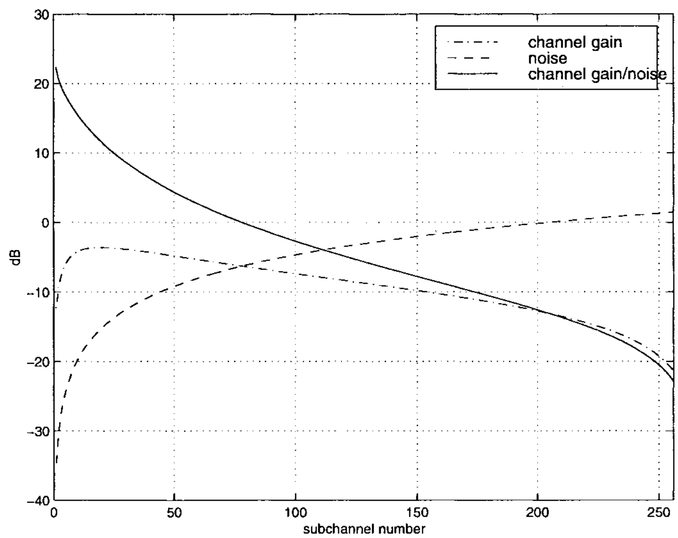


Fig. 1. Typical spectrally shaped channel.

conjunction with scalable or layered source coding schemes can be used to minimize the overall distortion [13]–[15]. More specifically, source data can be decomposed into hierarchical perceptually relevant layers, each of which has different perceptual importance. Layers containing high spatial frequency components are generally more tolerant to channel error effects than layers containing low spatial frequency components and, thus, may receive less channel protection. Unequal error protection (UEP) for different layers achieves more robustness compared to equal error protection.

Recently, MCM has become popular in image and video transmission. A combined source-channel coding scheme using MCM to provide unequal error protection for an additive white Gaussian noise channel (AWGN) is developed in [11]. Typical quantization will make some bits in the quantized signal (e.g., the most significant bit of a scalar quantized codeword) far more important than others. MCM allows different bits to be transmitted at different subchannels while proper power allocation provides different error performance. Predefined bit rate distribution is used. The above idea was later extended to image transmission over AWGN [12]. In [16], multiresolution modulation (MR) is used instead of QAM modulation to transmit "coarse" and "fine" descriptions of the source at predefined BER's. The proposed embedded multicarrier modulation (EMCM) introduces a power and bit rate allocation scheme to maximize the data throughput for given power constraint and target BER's.

It is well known that loading algorithms [2], [6]–[9] are developed for data transmission. If we want to deliver layered coded multimedia data through spectrally shaped channels using these loading algorithms, the layers are transmitted consecutively with the same BER as shown in Fig. 3, Type I, resulting in equal error protection. The algorithm in [10] only optimizes the averaged BER rather than the BER at a particular subchannel. Therefore, the layers are still transmitted at the same averaged BER and receive equal error protection. On the other hand, the algorithm in [11] and [12] assigns the same amount of power to the layers, although the bits within each codeword receive unequal error protection. The number of subchannels used is decided by layer's codeword length. As such, the transmitter/receiver has to frequently update the transmission parameters including the number of subchannels, subchannel power, and bit rate, which is not realizable. The algorithms developed for AWGN channels are also not applicable for spectrally shaped channel. For [16], unequal error protection is achieved essentially by MR, and the designed loading algorithm still assigns the same error performance across the usable subchannels.

This paper aims to develop an efficient, powerful, yet simple scheme to transmit multimedia data through typical spectrally shaped ADSL channels. We consider a fixed number of subchannels and data throughpu independent of source input data. Such an assumption is different from the existing work [11], [12] in which the above parameters are source dependent. The essence of the proposed approach is to allow the source layers of different perceptual importance to be transmitted in parallel, each occupying a set of subchannels. It is a joint source/channel design that allows the coding and modulation at each subchannel to be tailored to the importance of the layer it transmits while attempting to effectively allocate the communication resources among the subchannels. This results in different BER at different subchannels. The proposed loading algorithm aims to optimally assign BER to the layers.

The remainder of the paper is organized as follows. In Section II, we present the advantages of layered coding and MCM and give an idea of how to combine these advantages together to construct a robust, parallel multimedia data transmission system. After describing the channel distortion analysis and the assignment of subchannels to the layers, we derive a mathematical model for the optimization in Section III. In Section IV, a loading algorithm and the simplified version are derived as well as the asymptotic gains in terms of channel distortion. Section V applies the proposed loading algorithm to subband coded image and H.263 video transmissions, which correspond to fixed-length coding and variable-length coding, respectively. In Section VI, we demonstrate how to apply coded modulation to further improve the performance. Finally, we reach our conclusions in Section VII.

II. Combined Layered Source Coding and Multicarrier Modulation

While much recent work on both layered source coding and MCM has demonstrated considerable success, considerable work remains in developing effective and robust schemes which can be useful in the transmission of multimedia data over telephone lines or spectrally shaped channels. In this section, we first introduce the layered coding and MCM as well as their advantages and then discuss how to combine these advantages together to construct a robust system.

A. Layered Coding

The layered or scalable source coding schem, in conjunction with the joint source/channel coding approach, offers the potential for robust performance in the presence of channel error effects. The encoded data is represented in terms of a

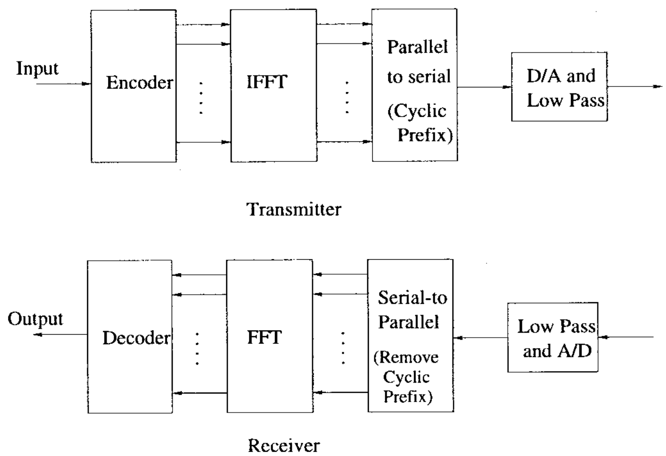


Fig. 2. Multicarrier basic model.

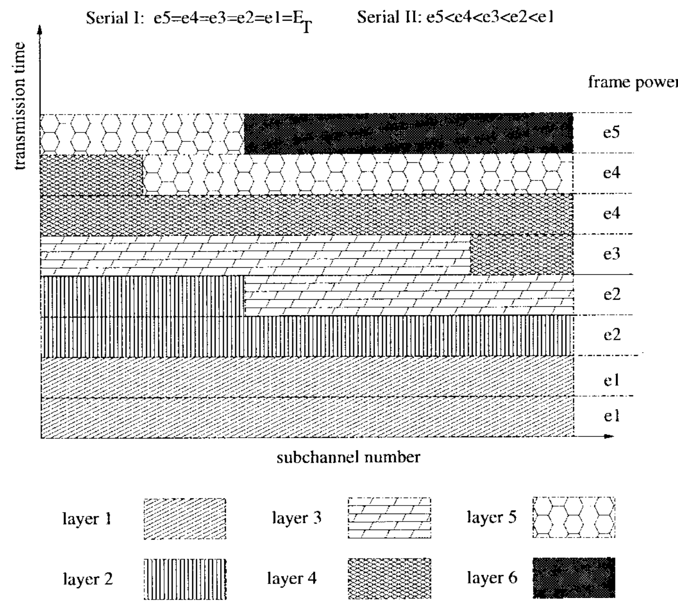


Fig. 3. Serial transmission.

number of layers, each resulting in a distinct data stream, which represent different perceptually relevant components of source material. The different layers may have distinctly different tolerances to channel errors, and the corresponding data streams can then be handled differently by the transmitter or network. Typical layered coding includes subband coding [17] which is initially developed for image compression and later extended to video compression [18]. The image and video is decomposed into a set of subbands. The low spatial and temporal subbands carried more information and, thus, are more important compared to the high spatial or temporal subbands. The bit error happening to the subbands with higher importance will cause larger amplitude change and, thus, more distortion in a mean-square sense. Unequal error protection was adopted by many schemes [13], [15] and yielded unquestionable performance improvement over equal error protection.

B. Multicarrier Modulation

MCM is a form of FDM. The basic MCM system is illustrated in Fig. 2. At the transmitter side, a block of bits (equal to the number of bits each subchannel can support) is then encoded as a set of quadrature amplitude modulated (QAM) symbols and sent to an inverse Fourier transform which combines the complex subsymbols into a set of real-valued time domain samples. In practice, the subchannels are not completely independent, and a so-called cyclic prefix is added to remove intersymbol interference (ISI) among the subchannels [20]. At the receiver side, the reverse process is performed. The block diagram for the receiver is shown in Fig. 2. By carefully dividing the channel into subchannels and allocating the maximum allowable number of bits to each subchannel, MCM or DMT can be shown to be an optimal code [21]. This means that such a system can perform at the theoretical limit and that no other system can exceed its performance.

MCM has the advantage of allowing the transmitted power, bit rate and even the channel encoder of each subchannel to be changed flexibly without affecting other subchannels. Optimum use of the channel can be obtained by making optimum use of each subchannel. As we will show, to transmit layered coded image and video, this property shows more promising advantages such as providing unequal error protection naturally without any increasing complexity at both the transmitter and receiver.

C. Unequal Error Protection versus Equal Error Protection

Typically, the multimedia data layers are transmitted in consecutive order as in data transmission through spectrally shaped channels, namely, serial transmission Type I as shown in Fig. 3. The loading algorithm assigns the same BER to all subchannels by adapting the power and bit rate, similar to that of [2], [7], and [8], or the same averaged BER to the layers if use the algorithm in [10] is used. Due to the fact that unequal importance exists among the layers, this type of transmission which achieves equal error protection is not optimal.

We define the transmitted power sum over all the usable subchannels during a single transmission as frame power. Usually, larger frame power results in lower BER. Therefore, unequal error protection can be achieved in serial transmission by varying the frame power during each layer's transmission, shown in Fig. 3 as serial Type II. Such approach falls into a class of multiple-level optimization which requires extensive computations. For a given frame power distribution $\{e_m\}_{m=1}^N$, the first level optimizes the transmitted power and bit rate assignment for given frame power, which is the same as the loading algorithms in [7] and [10]. The second level assigns frame power $\{e_m\}_{m=1}^N$ to N layers. It is obvious that $e_i > e_j$ if layer i is more important than j. We describe this loading algorithm in detail later in Section IV-C. Since generally the layers differ greatly in bit data size and importance, the subchannel power and bit rate distributions during each layer's transmission period differ greatly. As such, frequent change of channel parameters during the transmission becomes another disadvantage of this approach. For some ADSL channels where the dominant line impairment

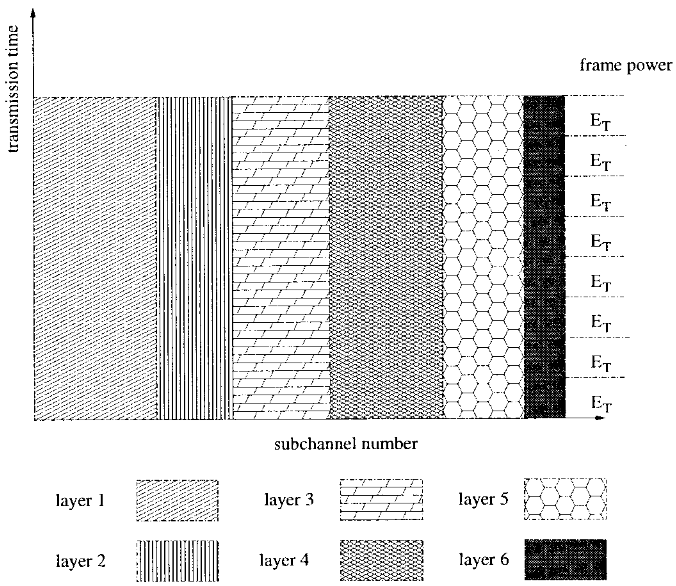


Fig. 4. Parallel transmission.

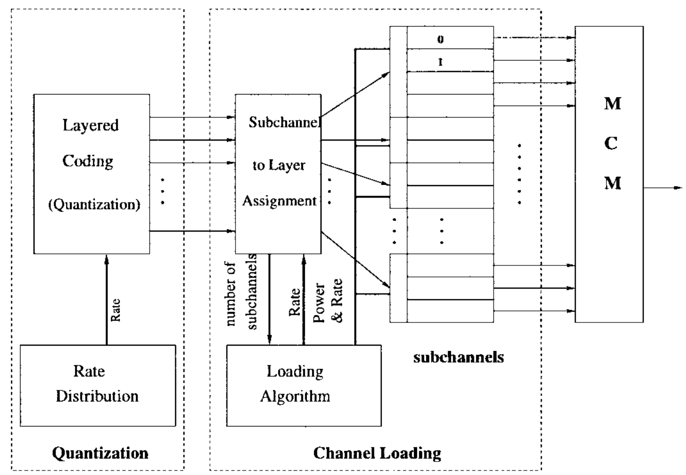


Fig. 5. Combined layered source codingand MCM.

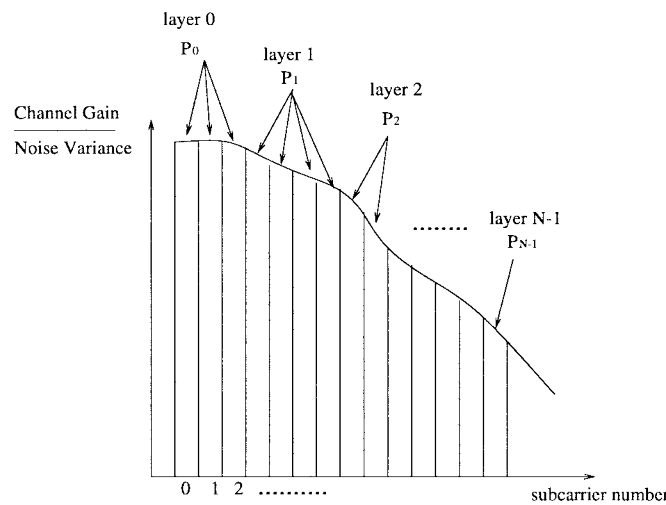


Fig. 6. Subchannel to layer assignment.

is near-end crosstalk (NEXT), such an approach fails since increasing frame power also increases the interfering power of the NEXT.

The serial transmissions share the same criterion, namely during each single transmission, removing the difference in channel gain and noise variance, as well as assigning the same BER across the subchannels. This is still true for serial type II, although the BER's at a particular subchannel during two different transmissions may be different. In contrast, we believe that the existence of different channel gain and noise effect at different subchannels offers the potential for robust transmission, by considering the possibility of providing unequal error protection through optimally allocating the available communication resources. A combined layered coding and MCM would serve as a useful approach. Toward this, we propose to transmit the layers simultaneously, each occupying a number of subchannels, as illustrated in Fig. 4. The transmission time of all the layers is forced to be the same so that the number of subchannels occupying a layer can be decided by the number of data bits of this particular layer as well as the data throughput required per transmission.

The proposed combined MCM and layered coding scheme is shown in Fig. 5. The subchannel-to-layer assignment first computes the number of subchannels that each layer requires given the bit rate distribution. Since the BER varies from subchannel to subchannel, a new loading algorithm is required to achieve such a goal. To a particular subchannel, the error performance may not be the best, but the overall distortion is minimized under the power and bit rate constraints. The power and bit rate distribution is decided by the current source input (image, video frames). For channels exhibiting invariable characteristics, the distribution stays unchanged during the transmission of this particular input. We have already mentioned that the number of subchannels allowed and frame power are decided by the transmitter/receiver rather than source data. Although the receiver must wait until the end of the transmission to get all the layers, the proposed algorithm tries to allocate adequate communication resources to the most important layers to insure a base-level quality.

III. The Optimization Problem

This section provides a mathematical model for the above optimization. We first discuss the analysis of the channel distortion using the weighting factor of the layers and propose a power-efficient scheme to assign the subchannels to the layers. Finally, an optimization function in terms of minimizing the channel distortion in the mean-square sense by finding the optimal power and bit rate distribution across the subchannels is derived.

A. Channel Distortion

Assuming the orthogonality of source and channel distortion, the overall distortion can be written as the sum of source distortion and channel distortion. The bit allocation algorithm [19] can be used in layered coding to optimally distribute the source target data rate among the layers. It is noted that source rate distribution and channel modulation can be jointly optimized to reduce the overall distortion [15]. For simplicity, we choose to allocate the source rates separately from the channel design. Therefore, the objective becomes minimization of the channel distortion by finding the overall optimal error performance for all the layers under power and bit rate constraints.

For layered image and video transmission, if applying vector quantization (VQ), the channel distortion is defined as [12]

$$D_c = \sum_{m=1}^{N} \sum_{i=0}^{N_m-1} \sum_{j=0}^{N_m-1} P(i)P(j \mid i)(y_i - y_j)^2 \qquad (1)$$

where N is the total number of layers (not including the layers thrown out by the source bit rate distribution algorithm), N_m is the cardinality of the layer m, and y_i and i are the codewords for VQ and binary codeword for transmission. We assume that only single bit error within one binary codeword with probability P_m, i.e.,

$$P(j \mid i) \approx \begin{cases} P_m, & \text{if } i, j \text{ differ in 1 bit} \\ 1-(R-1)P_m, & \text{if } i = j \\ 0, & \text{otherwise} \end{cases} \qquad (2)$$

where R is the length of codeword i. Thus, (1) becomes

$$D_c = \sum_{m=1}^{N} P_m \sum_{i=0}^{N_m-1} D_{i,m} = \sum_{m=1}^{N} P_m W_m \qquad (3)$$

where $D_{i,m}$ represents the sum of the distortion between codeword i and the other codeword, which differs 1 bit from i, and $W_m = \sum_{i=0}^{N_m-1} D_{i,m}$ represents the average distortion caused by a single bit error at layer m, deemed the weighting factor of layer m. Usually, layers with high energy have larger weighting factor and, therefore, higher importance. P_m is the BER of layer m, which is a function of power, bit rate, and CGNR of the subchannels assigned to layer m.

B. Assignment of Subchannels to Layers

Since CGNR differs quite a lot at each subchannel, the subchannels-to-layers assignment has to be carefully designed to achieve maximum power efficiency. We want to assign the subchannels with higher CGNR to the layers of higher importance, as shown in Fig. 6. Such assignment will ensure the most important layers are transmitted over reliable channels without large power usage. It becomes more advantageous under low power constraint. The following subchannel-to-layer assignment algorithm is used in conjunction with the loading algorithm during the optimization.

- Step 1: Sort the subchannel in increasing CGNR order, and sort the layers in increasing importance (weighting factor) order. Define $\{g_i\}_{i=1}^{C_T}$ as the CGNR of the subchannels. Now we get $g_1 \geq g_2 \geq g_3 \cdots \geq g_{C_T}$ and $W_1 \geq W_2 \geq W_3 \cdots \geq W_N$.
- Step 2: Given the bit rate $\{R_m\}_{m=1}^{N}$, compute the number of subchannels that each layer needs

$$c_m = \left\lfloor \frac{\text{Bit}_m B_T}{R_m \,\text{Bit}_{\text{total}}} \right\rfloor, \quad m = 1 \ldots M \qquad (4)$$

where Bit_m is the total number of bits of layer m, $\text{Bit}_{\text{total}}$ is that of the whole image, and B_T is the number of bits that must be transmitted in every transmission. For given $\{R_m\}_{m=1}^{N}$, $\{c_m\}_{m=1}^{N}$ are adjusted until

$$\sum_{m=1}^{N} c_m R_m = B_T \qquad (5)$$

if $C = \sum_{m=1}^{N} c_m > C_T$, then this $\{R_m\}_{m=1}^{N}$ is not applicable. Otherwise, assign subchannel 1 through c_1 to layer 1 and subchannel $c_1 + 1$ through $c_1 + c_2$ to layer 2, etc. For $m = 1$ to N, compute the CGNR of the kth subchannel transmitting layer m as

$$G_{k,m} = g_{k+\sum_{n=1}^{m-1} c_n}, \quad k = 1 \ldots c_m. \qquad (6)$$

The optimal design requires all the subchannels belong to the same layer have the same error performance. This can be accomplished in two ways.

Approach I:
Assign the same bit rate to the subchannels belonging to the same layer, and allocate transmitted power to these subchannels to achieve the same SNR. The c_m subchannels corresponding to layer m have the same channel $\text{SNR} = \frac{E_m G_m}{c_m}$ and bit rate R_m. Here, we define E_m as the power sum of the subchannels transmitting layer m, G_m as the CGNR averaged over c_m subchannels

$$E_m = \sum_{k=1}^{c_m} E_{k,m}, \quad G_m = \frac{c_m}{\sum_{k=1}^{c_m} \frac{1}{G_{k,m}}}. \qquad (7)$$

The loading algorithm needs to find the optimal $\{E_m, R_m, c_m(R_m), G_m(R_m)\}_{m=1}^{N}$ to minimize the channel noise effect. After deriving E_m, the power allocated to a particular subchannel is computed as

$$E_{k,m} = \frac{E_m}{c_m} \frac{G_m}{G_{k,m}}, \quad k = 1 \ldots c_m. \qquad (8)$$

Approach II:
Use the algorithm in [7] to assign bit rate and power among the subchannels transmitting the same layer so that they perform at the same BER. We define $R_{m,T}$ as the number of bits of layer m per transmission, which can be decided by Bit_m and B_T. The bit rate and power of a particular subchannel are computed as [7]

$$\begin{aligned} R_{k,m} &= \text{QUANT}\left\{ \frac{R_{m,T}}{c_m} + \frac{1}{c_m} \log_2 \left(\frac{G_{k,m}^{c_m}}{\prod_{i=1}^{c_m} G_{i,m}} \right) \right\} \\ E_{k,m} &= \frac{E_m 2^{R_{k,m}} / G_{k,m}}{\sum_{i=1}^{c_m} 2^{R_{i,m}} / G_{i,m}}, \quad k = 1 \ldots c_m. \end{aligned} \qquad (9)$$

QUANT is defined in [7]. The error probability function of the subchannels transmitting layer m, using QAM modulation, is then

$$\begin{aligned} Pe &\approx 4Q\left(\sqrt{ \frac{3E_m / c_m}{\left(2^{\frac{R_{m,T}}{c_m}} - 1 \right) \frac{1}{c_m} \sum_{i=1}^{c_m} \frac{2^{\left(R_{i,m} - \frac{R_{m,T}}{c_m} \right)}}{G_{i,m}}} } \right) \\ &\approx 4Q\left(\sqrt{ \frac{3E_m G_m / c_m}{(2^{R_m} - 1)} } \right) \end{aligned} \qquad (10)$$

where $R_m = R_{m,T}/c_m$ is the averaged bit rate, $G_m = \frac{c_m}{\sum_{i=1}^{c_m} 2^{(R_{i,m}-R_m)}/G_{i,m}}$ is the averaged CGNR. Therefore, the optimization is to find the optimal $\{E_m, R_m, c_m, G_m\}_{m=1}^N$. We limit R_m to be integer value so that the optimization procedure is as same as approach A. The difference between these two approaches is G_m.

C. Mathematical Problem

Under the above assumption, the optimization function for the loading algorithm can be formulated as

$$\begin{aligned}
&\text{Given} \quad \{\text{Bit}_m\}_{m=1}^N, B_T,\\
&\text{Find} \quad \{E_m, R_m, c_m(R_m), G_m(R_m)\}_{m=1}^N, \text{ by}\\
&\text{Min} \quad D_c = \sum_{m=1}^N Pe\left(R_m, \frac{E_m G_m}{c_m}\right) W_m\\
&\text{subject to} \quad \sum_{m=1}^N E_m \le E_T, \quad \sum_{m=1}^N c_m \le C_T
\end{aligned} \tag{11}$$

where E_T is the frame power constraint, C_T is the maximum number of subchannels allowed to use, and Pe is the BER function. For given $\{R_m\}_{m=1}^N$, $\{c_m\}_{m=1}^N$ are selected until the bit data rate constraint in (5) are satisfied. Therefore, the data rate constraint is turned into the number of subchannel constraint $\sum_{m=1}^N c_m \le C_T$. Under low power constraint, we may consider transmitting fewer numbers of layers and allocating the saved power to the layers left. This would decrease channel distortion but increase source distortion. Therefore, joint design of source rate distribution and channel loading algorithm can further improve the performance. Instead of fixing data throughput, we choose a fixed transmission time and allow the system to choose data throughput, the number of layers to transmit, and the layer's source bit rate optimally to minimize the overall distortion. Given transmission time limit T, the joint optimization problem is then

$$\begin{aligned}
&\text{Find} \quad \{\text{Bit}_m, B_T, E_m, R_m, c_m(R_m), G_m(R_m)\}_{m=1}^N\\
&\text{Min} \quad D = \sum_{m=1}^N D_s(\text{Bit}_m) + Pe\left(\text{Bit}_m, R_m, \frac{E_m G_m}{c_m}\right) W_m\\
&\quad \text{subject to} \quad \sum_{m=1}^N E_m \le E_T, \quad \sum_{m=1}^N c_m \le C_T\\
&\quad t_m = \frac{\text{Bit}_m}{R_m c_m} < T, \quad m = 1 \dots N.
\end{aligned} \tag{12}$$

IV. New Loading Algorithms

Our goal in this section is to develop a computationally efficient solution to the optimization problem (11). Equation (12) can be solved similarly. We start from power allocation for given bit rate distribution, then extend it to a complete loading algorithm for spectrally shaped channels. Asymptotic gains in terms of channel distortion are also derived.

A. Power Allocation

Assuming that all the subchannels use QAM modulations and the $\{R_m\}_{m=1}^N$ are preselected, the probability of error function is given as [23]

$$Pe\left(R_m, \frac{E_m G_m}{c_m}\right) \approx 4Q\left(\sqrt{\frac{3E_m G_m}{c_m(2^{R_m}-1)}}\right). \tag{13}$$

Then, (11) is simplified into

$$\begin{aligned}
&\text{Min} \quad \sum_{m=1}^N 4Q\left(\sqrt{\frac{3E_m G_m}{c_m(2^{R_m}-1)}}\right) W_m\\
&\text{subject to} \quad \sum_{m=1}^N E_m \le E_T.
\end{aligned} \tag{14}$$

By applying the Lagrange multiplier, the optimal solution satisfies

$$W_m \sqrt{\frac{3G_m}{2E_m c_m(2^{R_m}-1)}} \exp \frac{-3E_m G_m}{2c_m(2^{R_m}-1)} = \lambda. \tag{15}$$

If we define

$$\Phi_\alpha(x) = \sqrt{\frac{\alpha}{x}} \exp(-\alpha x) \tag{16}$$

the optimal λ can be found by solving

$$\sum_{m=1}^N \Phi_{\alpha_m}^{-1}(\lambda/W_m) = E_T, \quad \text{with} \quad \alpha_m = \frac{3G_m}{2c_m(2^{R_m}-1)}. \tag{17}$$

Since $\Phi_\alpha(x)$ is a monotonic function of x for $\alpha > 0$, Φ^{-1} can be solved simply by bisection method. Furthermore, Φ_α^{-1} is monotonic as well and the optimal power distribution can be computed as

$$E_m(\lambda_{\text{opt}}) = \Phi_{\alpha_m}^{-1}(\lambda_{\text{opt}}/W_m). \tag{18}$$

For the AWGN channel, assumingthe same CGNR across the subchannels and BPSK modulation, the optimization is simply to find λ satisfying

$$\sum_{m=1}^N \Phi_{\alpha_m}^{-1}(\lambda/W_m)|_{\alpha_m = \frac{1}{c_m}} = E_T. \tag{19}$$

B. Bit Rate and Power Allocation

To derive the optimal bit rate distribution, the Lagrange method cannot be applied due to the integer number requirement of bit rate. The optimization is then an integer programming problem. Usually, in practical applications, the allowed bit rates are limited to the range $R_{\min} \le R_m \le R_{\max}$, where $R_{\min}, R_{\max}$ are the allowed upper bound and lower bound, respectively. $R_{\max}$ must satisfy that $R_{\max} C_T \le B_T$ to achieve data throughput B_T. Greedy search will go through all the $N^{R_{\max}-R_{\min}+1}$ $\{E_m, c_m\}_{m=1}^N$ computations. We name it **Loading Algorithm C**. Observed from (13), BER performance is decided by $\frac{E_m G_m}{c_m(2^{R_m}-1)}$. Decreasing R_m will increase c_m, therefore decreasing $\{G_k\}_{k=m}^N$ while increasing R_m will decrease c_m and increase $\{G_k\}_{k=m}^N$. Therefore, the optimal rate distribution may not be $\{R_m = R_{\max}, m = 1 \dots N\}$ and $\{R_m = R_{\min}, m = 1 \dots N\}$ but something in between. We propose to start from $\{R_m = R_{\max},$

$m = 1 \ldots N\}$ and approach the solution step by step, as described in **Loading Algorithm A**. Each step we select a layer and decrease the bit rate by one. This requires at most $N^2 \times (R_{\max} - R_{\min} + 1)$ $\{E_m, c_m\}_{m=1}^N$ computations [31].

Loading Algorithm A

1) **Initialization:**
   a) Sort CGNR and the layers as in Section III-B.
   b) $R_m = R_{\max}$, $m = 1 \ldots N$.
   c) Compute $\{E_m, c_m\}_{m=1}^N$ as in Section IV-A and $D_c = \sum_{m=1}^N Pe(R_m, \frac{E_m G_m}{c_m}) W_m$.
   d) $D_c^{\min} = D_c$, $n = 0$.
2) **Approach the solution:** Pick one layer, which yields minimum D_c, by reducing the bit rate by one.
   a) $D_c(0) = D_c^{\min}$, $n = n + 1$.
   b) For $l = 1$ to N
      i) $\{R_m^n\}_{m=1}^N = \{R_m^{n-1}\}_{m=1}^N$
      $R_l^n = R_l^n - 1$.
      ii) Compute $\{c_m, E_m\}_{m=1}^N$ and $C = \sum_{m=1}^N c_m$

$$D_c(l) = \begin{cases} \sum_{m=1}^N W_m Pe(R_m, E_m), & C \le C_T \\ +\infty, & C > C_T \end{cases}$$

      loop end.
   c) find $k = \arg\min_{l=0 \ldots N} D_c(l)$.
   d) if $k > 0$ $R_k^n = R_k^n - 1$, $D_c^{\min} = D_c(k)$.
      else $k = 0$, $\rightarrow$ cannot reduce D_c anymore. *Stop!*
3) Continue 2 until $\{R_m\}_{m=1}^N = \{R_{\min}, \ldots, R_{\min}\}$ or reach *Stop!*.

After changing any R_l, the whole $c_m, m = 1 \ldots N$ instead of c_m itself has to be rearranged to satisfy $\sum_{m=1}^N c_m R_m = B_T$. Thus, $C = \sum_{m=1}^N c_m \le C_T$ has to be checked every time. In **A**, at every stage 2, $\{E_m\}_{m=1}^N$ has to be computed for N times to decide which layer's bit rate to decrease. The following observations allow us to use a simple comparison scheme to pick the right layer, with only one $\{E_m\}_{m=1}^N$ computation.

Observation 1:

Assume layers $1 \ldots N$ are transmitted with power and rate distribution $D = \{E_m, R_m\}_{m=1}^N$. Let frame power $\sum_{m=1}^N E_m = E_T$, $\sum_{m=1}^N R_m = R_T$. Assume $\{E_m\}_{m=1}^N$ is optimized for $\{R_m\}_{m=1}^N$ to produce BER distribution $P = P_1, \ldots, P_N$. If another power and rate distribution $D' = \{E'_m, R'_m\}_{m=1}^N$ can achieve the same BER distribution P and $R'_T = R_T$, but the sum of power $E'_T = \sum_{m=1}^N E'_m < E_T$, then D' outperforms D based on P.

Observation 2:

Under the same circumstance in Observation 1, if power and rate distribution $D'' = \{E''_m, R''_m\}_{m=1}^N$ achieve the same BER distribution P, and the sum of power $E''_T = \sum_{m=1}^N E''_m < E'_T$, then D'' outperforms D' based on P.

From the above observations, at iteration n, instead of computing power distribution and distortion, the layer to decrease the bit rate is the one which yields the largest overall power reduction by decreasing the bit rate, while maintaining the same BER. We use the following property to approximate the new power distribution $\{\hat{E_m^n}\}_{m=1}^N$ for given $\{R_m^n\}_{m=1}^N$:

$$4Q\left(\sqrt{\frac{3\hat{E_m^n} G_m^n}{c_m^n (2^{R_m^n} - 1)}}\right) = 4Q\left(\sqrt{\frac{3E_m^{n-1} G_m^{n-1}}{c_m^{n-1} (2^{R_m^{n-1}} - 1)}}\right)$$
$$\longleftrightarrow \frac{\hat{E_m^n} G_m^n}{c_m^n (2^{R_m^n} - 1)} = \frac{E_m^{n-1} G_m^{n-1}}{c_m^{n-1} (2^{R_m^{n-1}} - 1)}$$
$$\longleftrightarrow \hat{E_m^n} = E_m^{n-1} \frac{G_m^{n-1}}{G_m^n} \frac{c_m^n (2^{R_m^n} - 1)}{c_m^{n-1} (2^{R_m^{n-1}} - 1)}. \tag{20}$$

The following is the revised loading algorithm which will only compute optimal power distribution for at most $N \times (R_{\max} - R_{\min} + 1)$ times.

Loading Algorithm B:

1) **Initialization**: same as A.
2) **Approach the solution**: at iteration n:
   a) For $l = 1$ to N, $\{R_m^n\}_{m=1}^N = \{R_m^{n-1}\}_{m=1}^N$, $R_l^n = R_l^{n-1} - 1$.
      Compute $\{G_m^{n+1}, c_m^{n+1}\}_{m=1}^N$
      if $(\sum_{m=1}^N c_m < C_T)$

$$E_s(l) = \sum_{m=1}^N \left(E_m^{n-1} - \hat{E_m^n}\right)$$
$$= \sum_{m=1}^N E_m^{n-1} \left(1 - \frac{G_m^{n-1}}{G_m^n} \frac{c_m^n (2^{R_k^n} - 1)}{c_m^{n-1} (2^{R_k^{n-1}} - 1)}\right). \tag{21}$$

      loop end;
   b) Find $k = \arg\max_{l=1 \ldots N} E_s(l)$.
   c) If $E_s(k) > 0$, set $R_k^n = R_k^n - 1$. Compute the new $\{E_m^n\}_{m=1}^N$ optimized for $\{R_m^n\}_{m=1}^N$. Else $\rightarrow$ cannot reduce power anymore. *Stop!*
3) Continue 2 until $\{R_m\}_{m=1}^N = \{R_{\min}, \ldots, R_{\min}\}$ or reach *Stop!*

C. Serial Loading Algorithm

We develop a loading algorithm for serial transmission Type II as shown in Fig. 3. We use the loading algorithm in [7] to achieve consistent BER performance among the subchannels. The frame powers during each layer transmission are allocated to provide different BER performance for different layers. Let the frame power during layer m transmission be e_m and the number of transmission layer m need to be t_m. Assume C subchannels are used. It is possible that data from two or more layers are transmitted together. Instead of assigning different BER's to different parts, it is reasonable to transmit them at the same BER of the data belonging to the more important layer. From the above assumption, the optimization function is as follows:

$$\text{Min} \sum_{m=1}^N 4Q\left(\sqrt{\frac{3e_m/C}{(2^{\frac{B_T}{C}} - 1)\frac{1}{C}\sum_{i=1}^C \frac{1}{g_i} 2^{(R_i - B_T/C)}}}\right)$$
$$(W_m - \rho_m W_m + \rho_{m+1} W_{m+1})$$
$$\text{subject to} \sum_{m=1}^N e_m t_m \le E_T \sum_{m=1}^N t_m \tag{22}$$

where ρ_m is the portion of the bits which are transmitted with layer $m-1$, g_i is the CGNR at subchannel i, and R_i is bit rate at subchannel i.

Loading algorithm D:

1) Derive the bit rate distribution:

$$R_i = \mathrm{QUANT}\left\{\frac{B_T}{C} + \frac{1}{c}\log_2\left(\frac{g_i^c}{\prod_{i=1}^{C} g_i}\right)\right\},$$

where QUANT is defined in [7].

2) Allocate the power: Similarly to power allocation in IV-A, the optimal $\{e_m\}_{m=1}^N$ can be resolved by applying Lagrange multiplier and solving

$$E_T\sum_{m=1}^{N} t_m = \sum_{m=1}^{N}\Phi_{\beta_m}^{-1}\left(\frac{\lambda}{(1-\rho_m)W_m + \rho_{m+1}W_{m+1}}\right) \tag{23}$$

where

$$\beta_m = \frac{3}{C\left(2^{\frac{B_T}{C}}-1\right)\frac{1}{C}\sum_{i=1}^{C}\frac{1}{g_i}2^{(R_i - B_T/C)}}.$$

The power of Layer m is then

$$e_m = \Phi_{\beta_m}^{-1}\left(\frac{\lambda_{\mathrm{opt}}}{(1-\rho_m)W_m + \rho_{m+1}W_{m+1}}\right). \tag{24}$$

D. Asymptotic Gains for AWGN Channel

Due to the nature of the above optimization problem, the asymptotic gain in channel distortion for the spectrally shaped channel cannot be derived as a close function. Without loss of generality, we derive the asymptotic gain in channel distortion for a Gaussian channel. Assume fixed BPSK modulation and high SNR value. Similar to that of [11], for high SNR and BPSK, the probability of error function can be estimated as

$$Pe_{\mathrm{BPSK}}(E,G) = \frac{e^{-EG}}{\sqrt{EG}} \tag{25}$$

where E and G represent the transmitted power and CGNR, respectively. Assuming G is normalized to 1, (15) is equivalent to

$$\frac{W_m}{c_m}\frac{e^{-\frac{E_m}{c_m}}}{\sqrt{\frac{E_m}{c_m}}} = \text{const.} \Rightarrow \frac{E_m}{c_m} - \frac{E_n}{c_n} \propto \log\frac{W_m/c_m}{W_n/c_n}. \tag{26}$$

Based on this, we compute the asymptotic channel induced distortion of different systems based on

$$D_c = \sum_{m=1}^{N} Pe_{\mathrm{BPSK}}(E_m, G_m)W_m.$$

1) The proposed combined system: from the power constraint and the above estimation, we get

$$E_T = \sum_{m=1}^{N} E_m = E_1\frac{C_T}{c_1} + \log\frac{\prod_{m=1}^{N}\left(\frac{W_m}{c_m}\right)^{c_m}}{\left(\frac{W_1}{c_1}\right)^{C_T}}. \tag{27}$$

After some manipulations, the channel distortion can be derived as

$$D_1 = C_T\frac{e^{-\frac{E_T}{C_T}}}{\sqrt{\frac{E_T}{C_T}}}\prod_{m=1}^{N}\left(\frac{W_m}{c_m}\right)^{\frac{c_m}{C_T}}. \tag{28}$$

2) Single carrier system: the power and modulation are the same for all layers. The channel distortion can be easily derived as

$$D_2 = \sum_{m=1}^{N} W_m\frac{e^{-\frac{E_T}{C_T}}}{\sqrt{\frac{E_T}{C_T}}}. \tag{29}$$

3) System of [11]: the power distribution within a codeword of a layer can be computed the same as the proposed system. Assume that the weighting factors of layer m are $\{W_{i,m}\}_{i=1}^{r_m}$ and $\sum_{i=1}^{r_m} W_{i,m} = W_m$, where r_m is the length of the binary codeword for layer m. The channel distortion can be written as

$$D_3 = \frac{\sum_{m=1}^{N} r_m\left\{\prod_{i=1}^{r_m} W_{i,m}\right\}^{\frac{1}{r_m}} e^{-\frac{E_T}{C_T}}}{\sqrt{\frac{E_T}{C_T}}}. \tag{30}$$

TABLE I
ASYMPTOTIC GAINS

Bit Rate	$C_T = 256$			$C_T = 128$		
(bpp)	$G_{1,2}$	$G_{1,3}$	$G_{3,2}$	$G_{1,2}$	$G_{1,3}$	$G_{3,2}$
1.0	7.622	3.587	2.16	7.633	3.592	2.130
0.4	4.92	2.291	2.148	5.067	2.359	2.148
0.25	3.220	1.683	2.10	3.24	1.60	2.023
0.1	2.39	1.58	1.50	2.41	1.60	1.506

Therefore, we can derive the asymptotic gains $G_{i,j}$ defined as gain of system i over system j, e.g., $G_{i,j} = \frac{D_j}{D_i}$.

1) The proposed system over the single carrier system:

$$G_{1,2} = \frac{\sum_{m=1}^{N} W_m}{\prod_{m=1}^{N}\left(\frac{W_m}{\frac{c_m}{C_T}}\right)^{\frac{c_m}{C_T}}}. \tag{31}$$

Usually, changing C_T will not change $G_{1,2}$ since $\frac{c_m}{C_T}$ will not change. However, since c_m are computed to satisfy $\sum_{m=1}^{N} c_m R_m = B_T$, changing C_T may cause different $\frac{c_m}{C_T}$; therefore, it has a tiny impact on $G_{1,2}$.

2) The proposed system over system of [11]

$$G_{1,3} = \frac{\sum_{m=1}^{N} r_m\left(\prod_{i=1}^{r_m} W_{i,m}\right)^{\frac{1}{r_m}}}{\prod_{m=1}^{N}\left(\frac{W_m}{\frac{c_m}{C_T}}\right)^{\frac{c_m}{C_T}}}. \tag{32}$$

3) The system of [11] over the single carrier system is

$$G_{3,2} = \frac{\sum_{m=1}^{N} W_m}{\sum_{m=1}^{N} r_m \prod_{i=1}^{r_m} W_{i,m}^{\frac{1}{r_m}}}. \tag{33}$$

Table I shows the asymptotic gains at different source coding rates for "Lena" image with $C_T = 256$ and 128. As can be seen, C_T will have a tiny impact on the gains. Also, all the gains decrease as source coding rate decreases. This can be explained as follows. Decreasing source rate results in fewer layers to be transmitted; therefore, less importance difference exists among the layers. Assigning different Pe to different layers will not achieve the same performance improvement over equal error protection. It is expected that when the source rate decreases to produce only one layer, all the gains converge to 1. Since overall distortion can be approximated as the sum of source distortion and channel distortion, the overall distortion gain of the above systems can be easily obtained.

V. Application to Image and Video Transmission

The proposed loading algorithm is designed for any layered source coding applications that produce layers with different importance. In this section, we try to apply the proposed loading algorithm to transmitthe subband coded image and H.263 entropy coded video over spectrally shaped channels. These two applications correspond to fixed-length source coding and variable-length source coding, respectively. We show that MCM combines nicely with different types of layered source coding.

A. Application to Fixed-Length Coded Image

Subband coding combined with VQ has been a well-known scheme for image and video compression [12], [15], [17]. For an octave-band decomposition, the lowest subband contains most of the information and, therefore, is the most important. In our design, the images are first four-level subband decomposed using the Daubechies 16 wavelet filter and then vector quantized using full-search LBG algorithm [22]. The quantized results are fixed-length coded in order to achieve more robustness against error. The weighting factor of the subbands are computed as the averaged distortion in terms of the mean-square sense caused by a single bit error. We assume a fixed source quantizer. Error control codes can be applied to subchannels with low SNR. However, adding extra bits also increases the bit rate or the number of bits that the particular subchannels should carry at each transmission. The impact of the error correction codes are discussed in next section.

To demonstrate the advantage of the proposed loading algorithm, we compare the result to that of [11] and single carrier modulation on the AWGN channel. We use a total of 512 subchannels, employing BPSK modulation. Fig. 7 plots the PSNR of reconstructed "Lena" and "Pepper" grayscale images as a function of averaged channel SNR. For channel SNR below 8 dB, channel noise in BPSK modulation becomes observable and causes large degradation on the received image. For "Lena", at 4 dB channel SNR, and 0.1 and 1.0 b/pixel source rate, our system performs 1 and 4.01 dB better than that of [12], 1.96 and 6.98 dB better than single carrier modulation. As can be seen, the proposed system yields more performance improvement as the source rate increases and the channel SNR decreases. Larger source rate allows us to transmit more subbands so that more tradeoff in terms of BER can be achieved among the subbands. However, this is only true for high SNR's. If channel noise effect is large, transmitting more information also introduces larger channel distortion. If the increase in channel distortion overcomes the decrease in source distortion, increasing the source rate will degrade the performance. In this case, joint source and

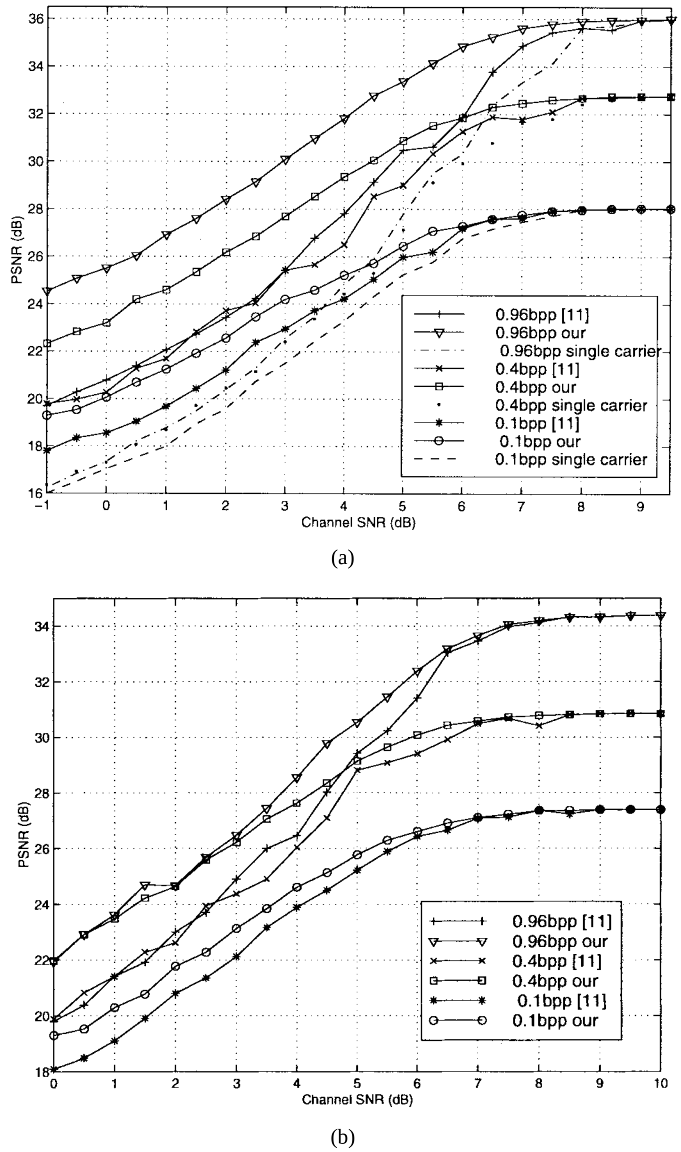
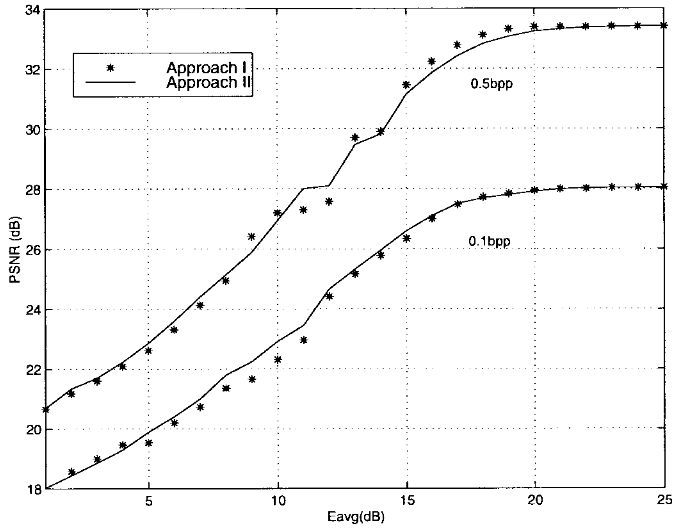


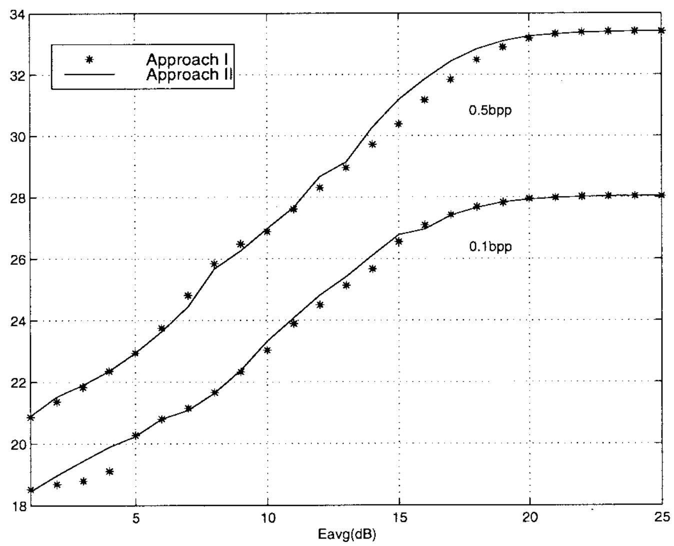
Fig. 7. Comparison of proposed algorithm to [11] and single carrier modulation using (a) "Lena" and (b) "Pepper."

channel rate allocation will compute the optimal source rates and achieve better performance. On the other hand, decreasing channel SNR results in increasing noise effect. Assigning most of the power to the lowest subband will ensure the lowest subband and receive adequate protection, and the performance degrades gracefully as the averaged SNR decreases.

For spectrally shaped channels, we assume a total of 256 subchannels, and each MCM symbol carries 512 bits, e.g., $C_T = 256$, $B_T = 512$. The bit rate is bounded by $R_{\max} = 6$ and $R_{\min} = 2$. We compare the loading algorithm $\mathbf{A}, \mathbf{B}, \mathbf{C}$ **and** $\mathbf{D}$ in Section IV by transmitting the "Lena" image. The results are measured as received image PSNR versus $E_{\text{avg}} = \frac{E_T}{C_T}$ (power per subchannel). Fig. 8 compares the approach I and II for loading algorithm $\mathbf{A}$ and $\mathbf{B}$ where approach II performs 0.2–1.0 dB better for both algorithms. We also compare loading algorithm $\mathbf{A}, \mathbf{B}$ using approach II to $\mathbf{D}$, as shown in Fig. 9(a). Parallel transmission system optimized using $\mathbf{A}, \mathbf{B}$ achieves 2–5 dB gain over the optimized serial

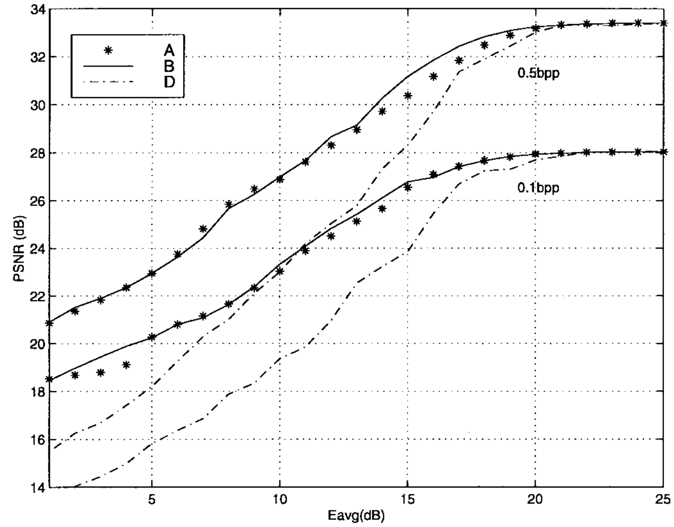


(a)

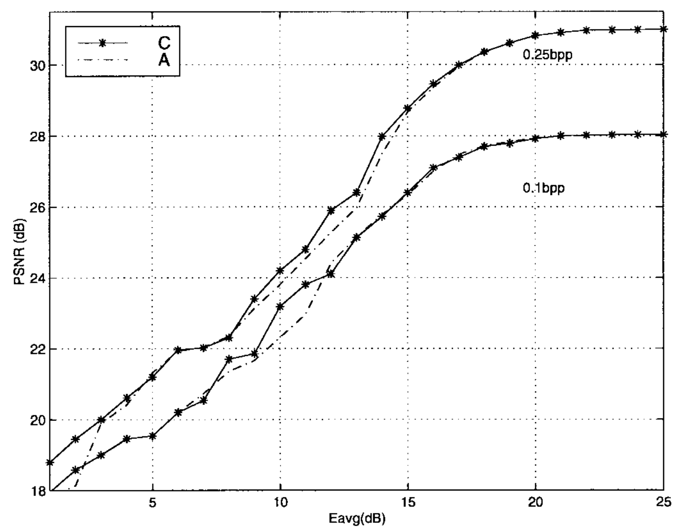


(b)

Fig. 8. Comparison of approach I and II using "Lena." (a) Algorithm A. (b) Algorithm B.



(a)



(b)

Fig. 9. Comparison of proposed loading algorithm (a) A, B versus D and (b) A versus C using "Lena."

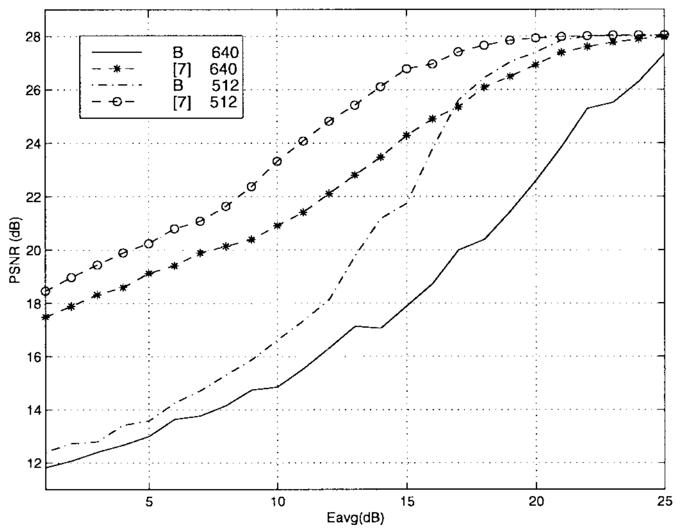
transmission system for $E_{\rm avg}$ ranging from 1 to 20 dB. In the simulation, we found $\mathbf{B}$ sometimes performs even better than $\mathbf{A}$, but not always. This may be explained as BER estimation in low SNR is not as correct as in high SNR, and algorithm $\mathbf{B}$ avoids $\frac{N^2-N}{N^2}$% BER computations compared to $\mathbf{A}$. From Fig. 9(b), we conclude that $\mathbf{A}$ performs nearly as good as $\mathbf{C}$ with much lower complexity.

Meanwhile, Fig. 10 illustrates the received PSNR versus $E_{\rm avg}$ curves of proposed loading algorithm $\mathbf{A}$ and that of [7]. We use the same system configuration as the above. For $B_T = 512$, loading algorithm $\mathbf{A}$ achieves 8–10 dB PSNR improvement at 0.5 b/pixel source rate and 4–6 dB at 0.1 b/pixel source rate. Increasing B_T also increase the subchannel bit rates. With the same frame power, this would degrade overall performance. At the same time, the PSNR improvement over that of [7] increases, especially for $E_{\rm avg}$ above 10 dB. The image results for $E_{\rm avg} = 14$ dB, $B_T = 512$ are shown in Fig. 11.

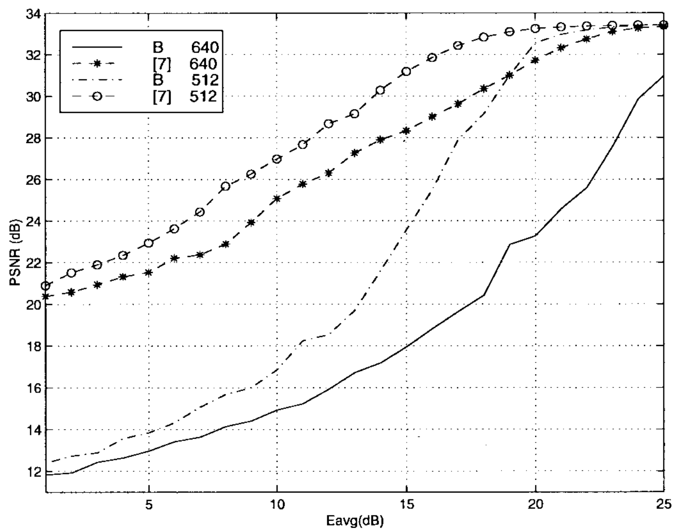
Fig. 12 plots the asymptotic AWGN channel distortion of different systems using the "Lena" images. Asymptotic gains from Table I and simulation results are both sketched for comparison. As can be seen, the asymptotic gain derived for high SNR value serve as a reasonable estimate for that of others. In Fig. 13, we plot the estimated distortion gains as the ratio of distortion computed as $D_c = \sum_{m=1}^{N} Pe(m)W_m$ compared to the simulation results using "Pepper" image. This further proves the correctness of the channel distortion estimation as in (3). We plot the estimated channel distortion gains for spectrally shaped channel applications in Fig. 14 at different target data bit rate B_T. It further proves that our system outperforms the existing systems, especially for larger data throughput.

B. Application to Variable-Length Coded Video

From the above, the proposed algorithm achieves significant performance improvement compared to the existing loading



(a)



(b)

Fig. 10. Comparison of proposed loading algorithm A to [7] using "Lena" at (a) 0.1 b/pixel and (b) 0.5 b/pixel.

algorithms. However, all these schemes are developed for fixed-length coding. Many compression standards, such as JPEG, MPEG, H.261, and H.263 [24] involve entropy coding: a type of variable length coding. The compressed data is highly sensitive to the propagation of channel errors, especially when the errors cause the decoder to lose synchronization and produce dramatic performance degradation. Possible solutions such as inserting extra synchronization codewords and error correction codes [25], [26] introduce more redundancy and, thus, less compression efficiency. UEP can be achieved by assigning different error correction codes to different video layers [27]. But due to the variable length property, the locations of the layers are unknown. Not only would the extra bits carrying the layer location information for every frame decrease the compression efficiency, they would also need to be highly protected, as this is very important information.

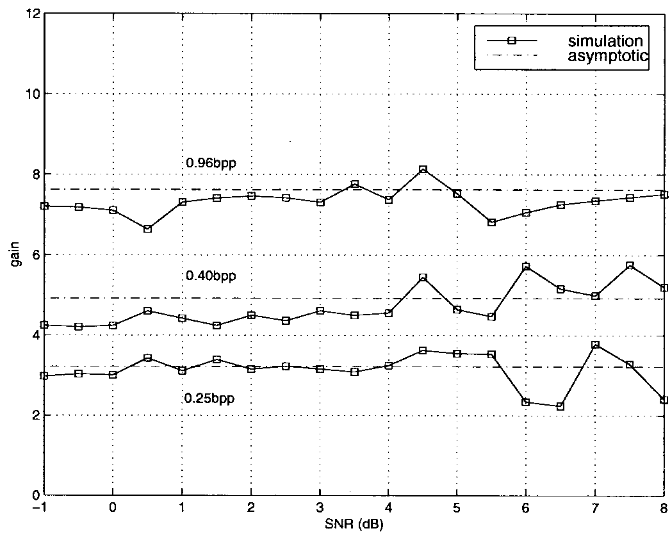
The error-resilient entropy code (EREC) was applied to image and video coding [28], [29] to give increased resilience to random and burst errors while maintaining high compression.

(a)

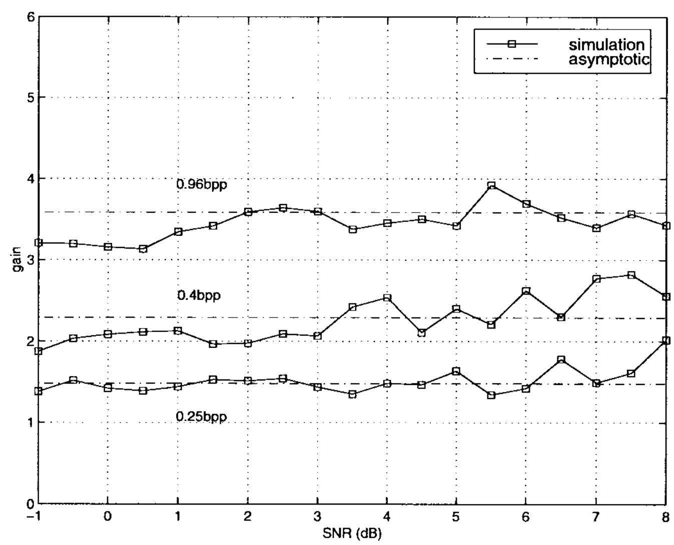
(b)

Fig. 11. (a) Received "Lena" image comparison and (b) proposed system [7].

The basic idea is to reorganize variable-length blocks to fixed length slots such that each block starts at a known position. This ensures that the beginning of each block is more immune to error propagation than those at the end. The redundancy introduced by using EREC is negligible. In [29], it showed that such a scheme can prevent the synchronization loss and the performance degrades gracefully as BER increases. EREC scheme has been applied to H.263 transmission over wireless networks [30]. We observe that although the EREC algorithm can remove the synchronization loss, it cannot prevent or reduce channel distortion in the mean-square sense caused

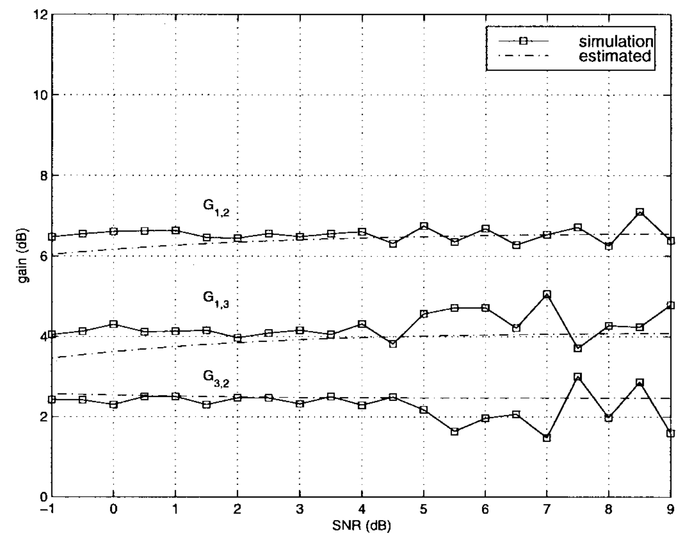


(a)

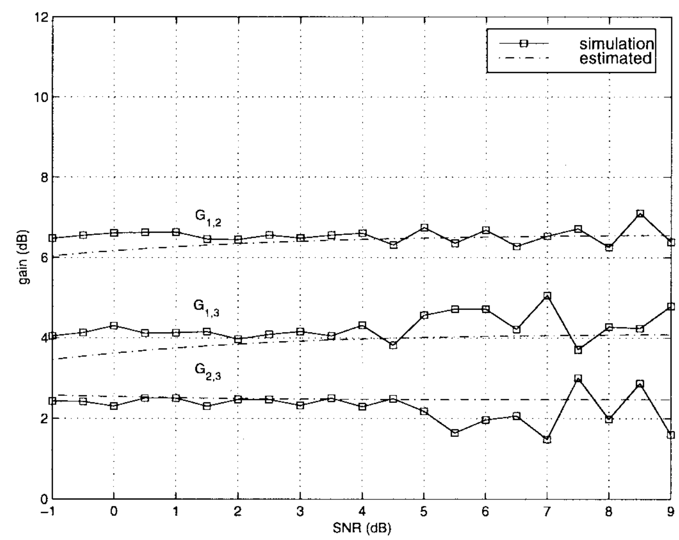


(b)

Fig. 12. Channel distortion gains for (a) $G_{1,2}$ and (b) $G_{1,3}$ using "Lena."



(a)



(b)

Fig. 13. Estimated channel distortion gains for (a) 0.1 b/pixel and (b) 0.4 b/pixel using "Pepper."

by errors happening to the data inside each slot. The current EREC scheme is designed to transmit the bits of each slot with the same error designate performance. However, errors that happen to the data bits of higher importance will lead to more channel distortion and longer error propagation to other blocks if EREC is employed. For example, coded macroblock indication (COD) of each intercoded block indicates whether or not the block is coded and is thus much more important than DCT coefficients. Typically in EREC, each slot may contain information from multiple blocks. COD error in one block would make the decoder use the data belonging to this particular block to reconstruct the next block or use data belonging to the next block to reconstruct the current block. In order to remove such shortcomings, we propose to combine EREC scheme with the MCM system into an efficient, robust transmission system [32].

As shown in Fig. 16, the essence of the proposed approach is first block coding the video frames using the same algorithms as in H.263. Each macroblock layer is deemed a data block. The data blocks are further reorganized to put into fixed length slots. The bits belonging to each slot are classified into several layered components of different importance. For example, COD bit and intra-dc's are of highest importance, while transform coefficients (TCOEF) are of lowest importance. The slot size and the frame type should be considered for classification. The receiver performs an inverse EREC to reorganize the slots to produce H.263 data blocks so that they can be decoded using the original decoder. For robust H.263 [24] video transmission, we suggest the following changes.

1) **Intraframes**: H.263 transmits six blocks corresponding to the same macroblock consecutively, resulting in unknown location of intra-dc's. However, intra-dc's are the most important information of the intraframe and should receive the highest protection. Therefore, we propose to code these six intra-dc's each as an 8-bit codeword and place them inthe beginning of each macroblock.

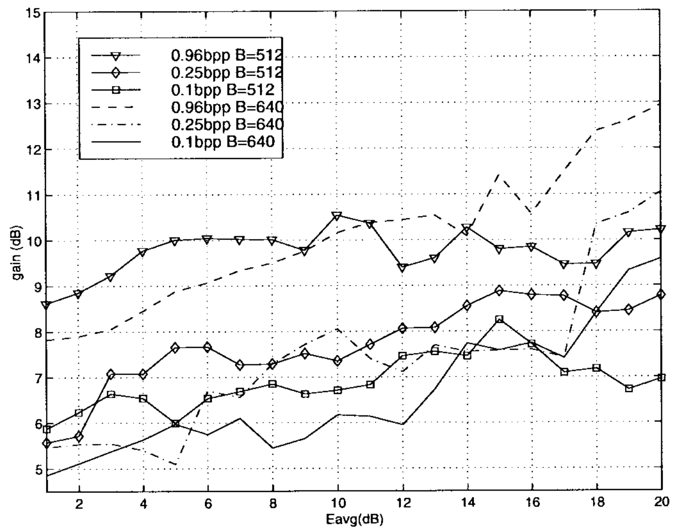


Fig. 14. Channel distortion gain of proposed algorithm over [7] using "Lena."

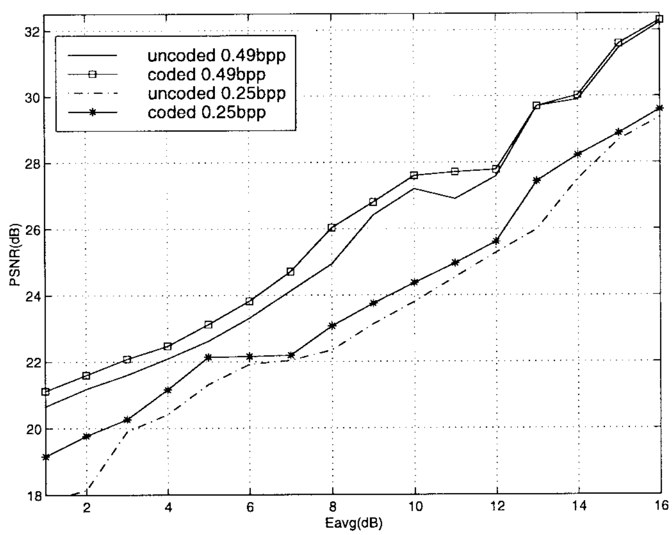


Fig. 15. Performance comparison of coded and uncoded system using "Lena."

This aims to obtain more precise layer classification and weighting factor computation to ensure the correctness and efficiency of the protection.

2) **Interframes**: Each macroblock begins with COD bit which is assigned with the highest importance. Macroblock type and block pattern (MCBPC) and motion vector data (MVD) are classified into the middle importance class. Inside each block layer, the inter-dc's and inter-ac's are entropy coded together to further reduce the bit rate and both are classified into the lowest importance class. Due to unpredictable length property of entropy coding, the importance weighting factor associate with the layers require intensive computation which is not quite practical. Instead, approximations are used. All the schemes associated with UEP face the same challenge, but the proposed scheme has the advantage of fixed-layer locations.

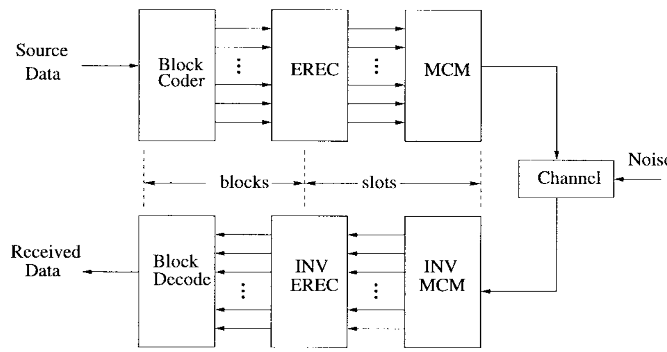
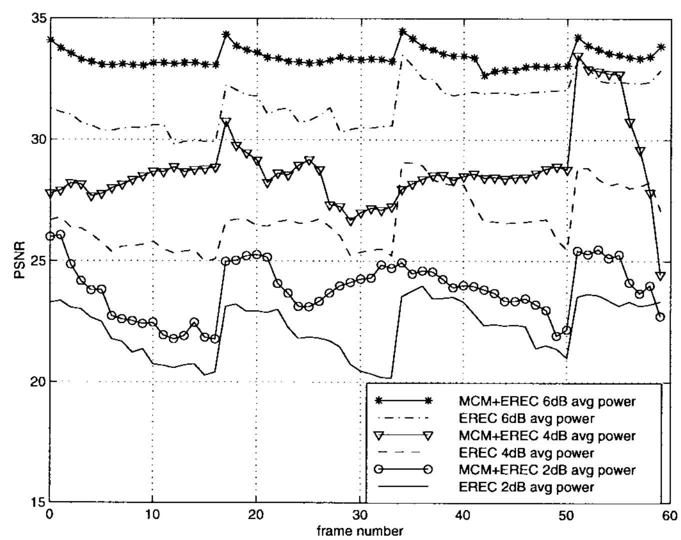
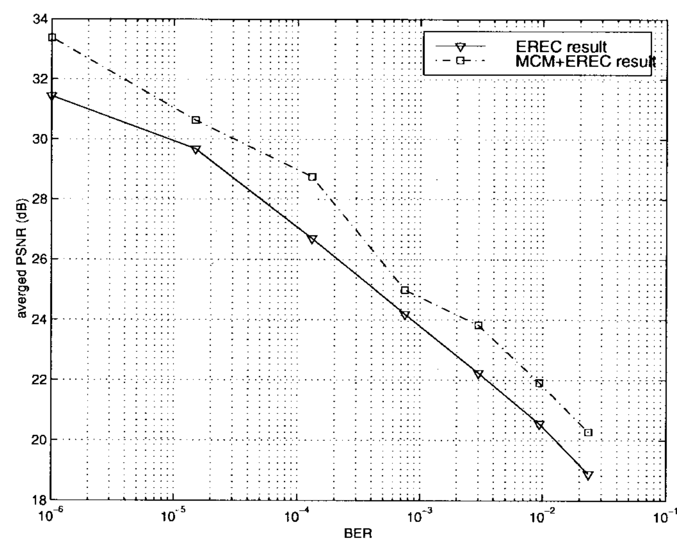


Fig. 16. Block diagram of combined MCM and EREC system.



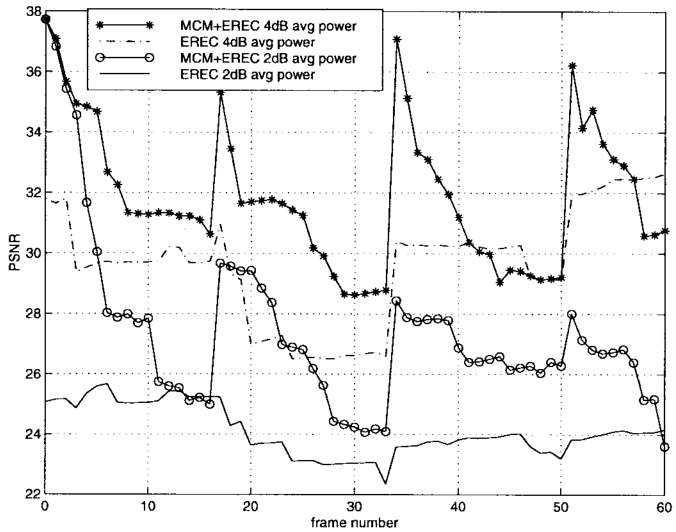
(a)



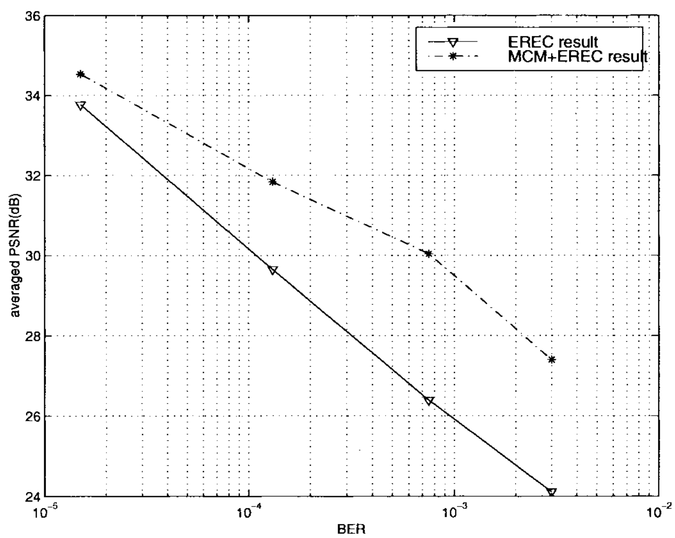
(b)

Fig. 17. Performance comparison of the proposed system and [29] using "Trevor" sequence: (a) PSNR performance versus frame number and (b) averaged PSNR versus BER.

The simulations are carried out on the combined system and the EREC system based on 60 frames of the standard QCIF (176 $\times$ 144 pels) color sequences "Trevor" and "Miss America." The coding, motion estimation, and prediction



(a)



(b)

Fig. 18. Performance comparison of the proposed system and [29] using "Miss America" sequence: (a) PSNR performance versus frame number and (b) averaged PSNR versus BER.

procedures are as the same as the H.263 standard unless described above. We use the following parameters: 512 real FFT, QAM modulation, 256 subchannels each occupying 4.3125 kHz required throughput of 512 b/FFT symbol. The PSNR is an average of the Y, C_b, and C_r component, defined as

$$\mathrm{PSNR} = \frac{4\mathrm{PSNR}_Y + \mathrm{PSNR}_{C_b} + \mathrm{PSNR}_{C_r}}{6}. \qquad (34)$$

An intraframe is inserted every 16 frames at 30 frames/s. Figs. 17(a) and 18(a) show the performance comparison in terms of the PSNR versus frame number. As can be seen, the combined system achieves 2–7 dB gain over the EREC system. Figs. 17(b) and 18(b) sketche the system performances as averaged PSNR over 60 frames versus BER, where approximately 2 dB gain for "Trevor" and 4 dB gain for "Miss America" are achieved by the combined system. Subjective results for frame 10 of the "Trevor" sequence under different average subchannel power are shown in Fig. 19 compared to the original

(a)

(b)

(c)

(d)

(e)

Fig. 19. Received Frame 10 of "Trevor." (a) H.263 no error. (b) EREC $E_{\mathrm{avg}} = 4$ dB. (c) EREC + MCM $E_{\mathrm{avg}} = 4$ dB. (d) EREC $E_{\mathrm{avg}} = 6$ dB. (e) EREC + MCM $E_{\mathrm{avg}} = 6$ dB.

frame. As we can see, for 6 dB average subchannel power, corresponding to 10^{-6} BER, the proposed combined MCM and EREC system performs nearly as good as error free, and for 4 dB average subchannel power, corresponding to 10^{-4} BER, the proposed system maintains preferable performance compared to the huge block effect and color distortion of the EREC system.

VI. Coded Modulation

Powerful coding techniques, such as trellis codes, are feasible means of further improving an MCM system performance [33]. There can be many ways to concatenate trellis codes on an MCM system. Most of the existing approaches use a single serial encoder/decoder which codes across the subchannels [34]. It has the advantage of low decoding latency and complexity. However, it also causes the error happening to a subchannel to propagate to the other subchannels. As a result, subchannel performances depend upon each other, which our system tries to avoid.

We propose to use separate encoder/decoders for the subchannels transmitting different source layers. Since the number of source layers are usually small, this will not cause great increase in complexity and latency. We use Wei's 4-D 16 state codes combined with trellis shaping. The error probability bound can be written as

$$P_{4D} \leq Q(\sqrt{\gamma_f \omega}) e^{0.5(\gamma_f \omega)} (24W^4 + 1536W^5 + 21\,504W^6 + 247\,808W^7 + 2\,824\,200W^8 + \cdots)\big|_{W=e^{-\frac{\omega}{2^{\hat{r}+1}}}} \tag{35}$$

where $\omega = 3\frac{EG}{2^R-1}$, $\gamma_{f,\mathrm{dB}} = 4.5$ dB, and $\hat{r} = 0.5$. This trellis code has a performance margin of approximately 7.5 dB, which means it outperforms the uncoded system only when ω are above this value. Therefore, in order to decide whether or not to apply trellis-coded modulation to the subchannels transmitting a particular layer, the proposed loading algorithm will first compute the power and bit rate while applying trellis coded modulation to all the layers. After comparing the derived ω to the performance margin, the loading algorithm assigns the uncoded modulation to be used on the layers whose ω is below the margin. The power and modulation allocation is reoptimized until the optimal solution is reached.

We want to point out that the sensitivity of P_{4D} in (35) to ω leads to the difficulty of deriving the optimal power distribution. A feasible solution is to first compute optimal power and bit rate distribution using uncoded modulation, and then apply coded modulation to the most important layers to reduce the power consumption of these layers. The BER's of those layers are small; thus, coding gain achieved by the trellis codes can be applied to reduce BER of other layers. From Fig. 15, we can see the improvement of the coded system over the uncoded system.

VII. Conclusion

We have proposed a robust multimedia data transmission framework by combining layered source coding and MCM. Unlike the existing schemes, we propose to transmit all the layers in parallel through different subchannels. Several simple yet powerful loading algorithms are presented which efficiently allocate transmitted power and bit rate to the subchannels according to the importance of the information they transmit. It achieves significant PSNR improvement for spectrally shaped channels based on the idea that "good" channels transmit more important information and "bad" channels transmit less important information.

The proposed algorithm can be used in both fixed-length coding and variable-length coding applications. For typical fixed-length coding applications, such as subband coding with VQ, numerical examples show that on a very noisy AWGN channel (channel SNR ranging from 1 to 6 dB), our result achieves 0.5–4 dB and 1–8 dB PSNR improvement over that of [11] and single carrier modulation, respectively. For the spectrally shaped channel, our scheme yields about 8–10 dB PSNR improvement over that of [7]. We also discover that the performance improvement increases as data throughput and source bit rate increase. The performance comparison under different parameters is also presented. Although most existing schemes are developed for fixed length coding applications, in this paper by combining MCM and EREC, we have proposed a robust and efficient system for entropy coded H.263 video transmission. Under noisy channels, not only synchronization-loss free, the proposed system achieves significant performance improvement.

Most of the work to date in this area has been very specific to the choice of source encoder, channel modulation/coding, channel conditions, transmission parameters, and channel models. The proposed scheme aims to remove this limit by choosing the channel parameters adaptively according to channel or network requirements. The performance comparison of the uncoded system and the system with coded modulation are also presented. We believe that source coding and channel modulation can be considered together as described in (12) to further improve the performance.

The implementation of the proposed algorithm requires only a small amount of computation. For a given input, the power and bit rate distribution remains the same for all subchannels. Another important improvement over [11] and [12] is that the number of subchannel allowed and data throughput can be selected flexibly by the system, independent of source input. The complexity of the combined MCM and EREC system mainly depends on the implementation of the EREC scheme, while the complexity of the loading algorithm can be ignored. It should be pointed out that the proposed combined system can be used for all types of data which can be decomposed to produce layers of different importance.

Acknowledgment

The authors would like to express their sincere appreciation to the reviewer for the comments and criticisms that have significantly enhanced the quality of this paper. The would like to thank X. Wang for providing the trellis code for simulation.

REFERENCES

[1] K. Maxwell, "Asymmetric digital subscriber line: Interim technology for the next forty years," *IEEE Commun. Mag.*, pp. 100–106, Oct. 1996.

[2] J. A. C. Bingham, "Multicarrier modulation for data transmission: An idea whose time has come," *IEEE Commun. Mag.*, May 1990.

[3] M. Alard and R. Lassalle, "Principles of modulation and channel coding for digital broadcasting for mobile receivers," *EBU Tech. Rev.*, no. 224, pp. 168–190, Aug. 1987.

[4] B. L. Floch, R. Halbert-Lassalle, and D. Castelain, "Digital sound broadcasting to mobile receivers," *IEEE Trans. Consumer Electron.*, vol. 73, pp. 30–34, Aug. 1989.

[5] C. W. Baum and K. F. Conner, "A multicarrier transmission scheme for wireless local communications," *IEEE J. Select. Areas Commun.*, pp. 521–529, Apr. 1996.

[6] J. Campello de Souza, "Optimal discrete bit loading for multicarrier modulation systems," in *Proc. IEEE Symp. Information Theory*, Boston, MA, 1998.

[7] R. F. H. Fisher and J. B. Huber, "A new loading algorithm for discrete multiton transmission," in *Proc. GlobalCOM'96*, pp. 724–728.

[8] P. S. Chow, J. M. Cioffi, and J. A. C. Bingham, "A practical discrete multiton transceiver loading algorithm for data transmission over spectrally shaped channels," *IEEE Trans. Commun.*, vol. 43, pp. 773–775, Feb.–Apr. 1995.

[9] B. S. Krongold, K. Ramchandran, and D. L. Jones, "Computationally efficient optimal power allocation algorithm for multicarrier communication systems," in *Proc. ICC'98*, Atlanta, GA, June 1998.

[10] T. J. Willink and P. H. Wittke, "Optimization and performance evaluation of multicarrier transmission," *IEEE Trans. Inform. Theory*, vol. 43, pp. 426–440, Mar. 1997.

[11] K. P. Ho and J. M. Kahn, "Transmission of analog signals using multicarrier modulation: A combined source-channel coding approach," vol. 44, *IEEE Trans. Commun.*, pp. 1432–1443, Nov. 1996.

[12] K. P. Ho and J. M. Kahn, "Image transmission over noisy channels using multicarrier modulation," *Signal Process.: Image Commun.*, July 1997.

[13] G. Cheung and A. Zakhor, "Joint source/channel coding of scalable video over noisy channel," in *Proc. IEEE Int. Conf. Image Processing (ICIP)*, 1996, pp. 767–770.

[14] I. Kozintsev and K. Ramchandran, "Multiresolution joint source-channel coding using embedded constellations for power-constrained time-varying channels," in *IEEE Proc. ICIP,*, 1996, pp. 2343–2346.

[15] H. Zheng and K. J. R. Liu, "Image and video transmission over wireless channel: A subband modulation approach," in *Proc. IEEE ICIP,* Chicago, IL, Oct. 1998.

[16] S. Pradhan and K. Ramchandran, "Efficient layered video delivery over multicarrier systems using optimized embedded modulation," in *Proc. IEEE ICIP*, Santa Barbara, CA, Oct. 1997.

[17] J. W. Woods and S. D. O'Neil, "Subband coding of images," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-34, pp. 1278–1288, Oct. 1986.

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

[19] Y. Shoham and A. Gersho, "Efficient bit allocation for an arbitrary set of quantizers," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 36, pp. 1445–1453, Sept. 1988.

[20] M. Barton and M. L. Honig, "Optimization of discrete multitone to maintain spectrum compatibility with other transmission systems on twisted copper pairs," *IEEE J. Select. Areas Commun.*, vol. 13, Dec. 1995.

[21] J. S. Chow, J. C. Tu, and J. M. Cioffi, "A discrete multitone transceiver system for HDSL applications," *IEEE J. Select. Areas Commun.*, vol. 9, Aug. 1991.

[22] A. Gersho and R. M. Gray, *Vector Quantization and Signal Compression*. Norwell, MA: Kluwer, 1992.

[23] J. G. Proakis, *Digital Communications*, 3rd ed. New York: McGraw-Hill.

[24] ITU-T DRAFT H.263, Video Coding for Low Bitrate Communication, May 1996.

[25] E. Steinbach, N. Farber, and B. Girod, "Standard compatible extension of H.263 for robust video transmission in mobile environments," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 7, Dec. 1997.

[26] A. Andreadis, G. Benelli, A. Garzelli, and S. Susini, "FEC coding for H.263 compatible video transmission," in *Proc. Int. Conf. Image Processing*, 1997.

[27] C. W. Yap *et al.*, "Error protection scheme for the transmission of H.263 coded video over mobile radio channels," *Vis. Commun. Image Process.*, 1997.

[28] N. T. Cheng and N. G. Kingsbury, "The ERPC: An efficient error resilient technique for encoding positional information of sparse data," *IEEE Trans. Commun.*, vol. 40, pp. 140–148, 1992.

[29] D. W. Redmill and N. G. Kingsbury, "The EREC: An error-resilient technique for coding variable-length blocks of data," *IEEE Trans. Image Processing*, vol. 5, pp. 565–574, Apr. 1996.

[30] T. Kawahara and S. Adachi, "Video transmission technology with effective error protection and tough synchronization for wireless channels," in *Proc. ICIP*, 1996, pp. 101–104.

[31] H. Zheng and K. Liu, "A new loading algorithm for image transmission over noisy channel," presented at 32nd Asilomar Conf. Signal, Systems, Computers, Pacific Grove, CA, Nov. 1998.

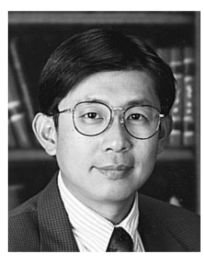
[32] H. Zheng and K. J. R. Liu, "Robust H.263 video transmission over spectrally shaped channels using multicarrier modulation," in *IEEE Proc. Workshop on Multimedia Signal Processing*, Los Angeles, CA, Dec. 1998.

[33] E. Biglieri, D. Divsalar, P. J. McLane, and M. K. Simon, *Introduction to Trellis-Coded Modulation with Applications*. New York: Macmillan, 1991.

[34] J. C. Tu and J. M. Cioffi, "A loading algorithm for the concatenation of coset codes with multichannel modulation methods," in *Proc. Globecom'90*, San Diego, CA, pp. 1664–1668, Nov. 1990.

Haitao Zheng received the B.S. degree in electrical engineering from Xian Jiaotong University, Xian, China, in 1995, and the M.S. degree in electrical engineering from the University of Maryland, College Park, in 1998, where she is currently pursuing the Ph.D. degree.

From 1994 to 1995, she was a Member of the Video Technology Institute, Xian Jiaotong University. From 1995 to 1998, she was an Institute for Systems Research Fellow and is now a Research Assistant with the Digital Signal Processing Labortory, Institute for Systems Research and Electrical Engineering Department, University of Maryland. Her research interests include joint source channel coding, robust image and video transmission, multicarrier modulation, and wireless communications.



K. J. Ray Liu (S'86–M'90–SM'93) received the B.S. degree from the National Taiwan University, Taipei, Taiwan, R.O.C., in 1983 and the Ph.D. degree from the University of California, Los Angeles, in 1990, both in electrical engineering.

Since 1990, he has been with Electrical Engineering Department and Institute for Systems Research, University of Maryland, College Park, where he is an Associate Professor. During his sabbatical leave in 1996–1997, he was Visiting Associate Professor at Stanford University, Stanford, CA. His research interests span various aspects of signal/image processing and communications. He has published over 160 papers, of which over 60 are in archival journals and book chapters. His research web page is at http:/dspserv.eng.umd.edu.

Dr. Liu has received numerous awards, including the 1994 National Science Foundation Young Investigator Award, the IEEE Signal Processing Society's 1993 Senior Award (Best Paper Award), the George Corcoran Award in 1994 for outstanding contributions to electrical engineering education, and the 1995–1996 Outstanding Systems Engineering Faculty Award in recognition of outstanding contributions in interdisciplinary research, both from the University of Maryland, and many others. He has been an Associate Editor of IEEE TRANSACTIONS ON SIGNAL PROCESSING. He was a Guest Editor of special issues on Multimedia Signal Processing of the PROCEEDINGS OF THE IEEE, a Guest Editor of special issue on Signal Processing for Wireless Communications of the IEEE JOURNAL OF SELECTED AREAS IN COMMUNICATIONS, and a founding member of Multimedia Signal Processing Technical Committee of the IEEE Signal Processing Society. He is an editor of the *Journal of VLSI Signal Processing Systems*, and the book series editor of the Marcel Dekker series on signal processing. He is the co-editor of the books *High Performance VLSI Signal Processing: Volume I: System Design and Methodology; Vol II: Algorithms, Architectures, and Applications* (Piscataway, NJ: IEEE Press).