Joint source-channel coding for scalable video

Lisimachos P. Kondi[a], Faisal Ishtiaq[b], Aggelos K. Katsaggelos[a]

[a]Northwestern University
Dept. of Electrical and Computer Engineering
Evanston, IL 60208 USA

[b]Motorola, Inc.
Schaumburg, IL 60196 USA

ABSTRACT

An approach towards joint source-channel coding for wireless video transmission at low bit rates is proposed. An SNR scalable video coder is used and a different amount of error protection is allowed for each scalable layer. Our problem is to allocate the available bit rate across scalable layers and, within each layer, between source and channel coding, while minimizing the end-to-end distortion of the received video sequence. The distortion is due to both source coding (quantization) errors and channel noise errors, and is measured in terms of the Mean Squared Error (MSE). The optimization algorithm utilizes rate-distortion characteristic plots. These plots show the contribution of each layer to the total distortion as a function of the source rate of the layer and the residual bit error rate (the error rate that remains after the use of channel coding). The plots are obtained experimentally using representative video sequences and show the sensitivity of the source encoder and decoder to channel errors. Our algorithm is operationally optimal given the rate-distortion characteristic plots. These plots are used in conjunction with plots that show the bit error rate achieved by the allowable channel coding schemes for given channel conditions in order to obtain the operational rate-distortion curve of each layer. Then, dependent Lagrangian optimization is used to determine the overall bit allocation across all layers.

Keywords: Joint source-channel coding, wireless video, scalable video

1. INTRODUCTION

During the past few years there has been an increasing interest in multimedia communications over different types of channels. In recent days a significant amount of research has been focused on multimedia transmission over wireless channels. This is a complex and challenging problem due to the multipath fading characteristics of the channel. It is made even more complex due to the transmission of compressed video data over these channels where distortions at the receiver are a function of both lossy source coding and the errors introduced by the channel. Hence it is not clear of how to best allocate a given bit budget between source and channel coding.

Source coding is concerned with the efficient representation of a signal. This is accomplished by reducing the redundancy of the signal as much as possible. While bit errors in the uncompressed signal can cause minimal distortions, in its compressed format a single bit error can lead to significantly larger errors. Although source coding can be very effective in reducing the rate of the original sequence, it renders the compressed signal very sensitive to the impact of errors. Hence for transmission over an error prone channel, it is imperative that channel coding be employed to make the data more resilient to channel errors by increasing the redundancy.

Traditionally, source and channel coding have been considered independently. The reason behind this is Shannon's important information-theoretic result establishing the *principle of separability*. The separability principle states that source and channel coding can be separated without any loss in optimality as long as the source coding produces a bit rate that can be carried by the channel and channel coding. While being an important theoretical derivation, this principle relies on the crucial assumption that the source and channel codes can be of arbitrary long lengths. In practical situations, due to limitations on the computational power and processing delays this assumption does not hold. It is then of benefit to consider the problem of source and channel coding jointly.

Joint video source-channel coding has been an active area of research. The focus of the techniques reported in the literature has been to minimize a given criterion (distortion, power consumption, delay, etc.) based on a given constraint. Among the many research directions, studies have focused on subband coding, scalable 3D subband

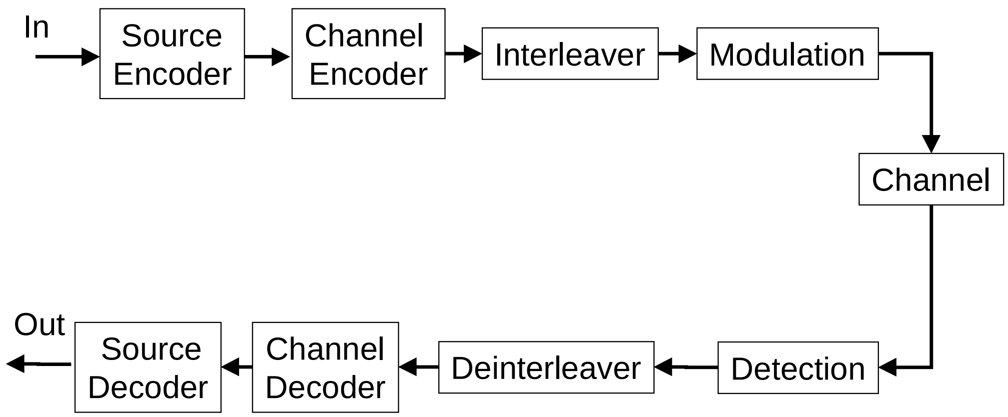

Figure 1. Block diagram of a typical video transmission system.

coding, and cosine transformed video.[1–6] Joint source-channel coding for High Definition Television (HDTV) has been reported in.[7] While most techniques deal with the "digital" (source and channel coding in terms of bits) aspects of source-channel coding, "analog" aspect have also been investigated where modulation techniques are developed in conjunction with the source code and the channel in mind. Studies on embedded modulation techniques and embedded constellations are discussed in Ref. 8 and Ref. 9 among others.

In a compressed video bit stream the various parts of the bit stream are not equally important. A natural way of protecting such a bit stream is by protecting more the bits that will impact the quality the most, such as frame headers, while data, such as the DCT texture data, can be protected less. This concept of varying the protection according to importance is called Unequal Error Protection (UEP). In this work we apply UEP to the layers of a scalable bit stream.

While scalability adds many functionalities to a bit stream, especially for heterogeneous networks, it is also well suited for enhanced error resilience. The break-up of a frame into subsets of varying quality lends itself naturally to employing an unequal error protection scheme. The base layer is typically better protected than the enhancement layers. In most cases it can be beneficial to distribute the protection bits unequally over the layers than to simply provide one protection level to the entire bit stream. This allows for added degrees of freedom in selecting the rates that will minimize the overall distortion. In Ref. 10, the benefits of using scalability in an error prone environment are shown by examining all the scalability modes supported by MPEG-2 in an ATM network.

Joint source-channel coding of video involves many facets of communications, information theory, and signal processing. The basic block structure of the entire system with the individual elements involved is shown in Fig. 1. As we progress through the paper in the development of an optimal rate allocation algorithm for scalable video, we will define each in more detail.

The paper is organized as follows. In section 2 we discuss scalable video coding in general and present our SNR scalable video codec that is used in this paper. In section 3 we discuss the main characteristics of wireless channels while in section 4 we discuss Rate-Compatible Punctured Convolutional Codes which are used in the paper. In section 5 we present the formulation of our problem as well as the algorithm for its solution. In section 6 experimental results are presented, while in section 7 conclusions are drawn.

2. SCALABLE VIDEO CODING

A scalable video codec produces a bit stream which can be divided into embedded subsets, which 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 small 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 types of scalability supported in the H.263 video compression standard: SNR, spatial and temporal. In SNR scalability, the enhancement in quality translates in an increase in the SNR of the reconstructed video

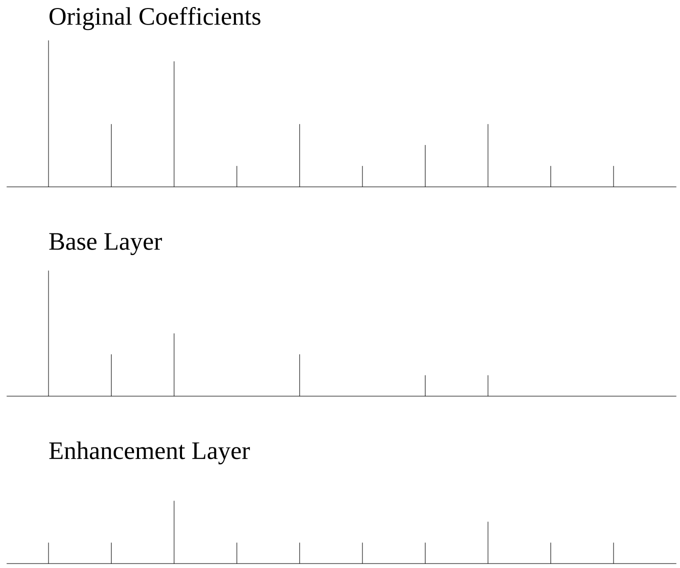

Figure 2. Proposed partitioning of DCT coefficients for SNR scalability.

sequence, while in spatial and temporal scalability the spatial and temporal resolution, respectively, is increased. We describe next a method for SNR scalability which requires only a single DCT and quantization step.[11,12]

The SNR scalability method of H.263 produces satisfactory results but requires an extra forward and inverse Discrete Cosine Transform (DCT) for each enhancement layer. Furthermore, since the residual error is transmitted like a regular frame, it also carries a significant overhead. Another method we have proposed earlier[13,14] involves only a single DCT and quantization operation. Then, the coefficients are partitioned in a number of sets which form the scalable layers. If three layers are involved, the base layer involves the transmission of coefficients (actually their quantization levels) $0-X$ in a zig-zag scan, without their A least significant bits. The first enhancement layer consists of coefficients $X+1$ to 63 without their B least significant bits. All remaining bits of all coefficients are transmitted with the second enhancement layer. The parameters X, A and B are adjusted by a rate control algorithm. This SNR scalability algorithm has a much lower computational complexity and overhead than the method supported by the H.263 standard.

A generalization of the method in[13,14] has been proposed in.[11,12] Clearly, setting the least significant bits of a coefficient to zero is equivalent to subtracting a certain value from it. The Variable Length Code (VLC) tables used in the standards use fewer bits for smaller coefficient magnitudes. Thus, subtracting a value from a coefficient reduces the number of bits required for its transmission but clearly increases the distortion. The decoder reconstructs the quantized DCT coefficients by adding the subtracted values (if available) to the values it received with the base layer. In the generalized approach therefore the base layer is constructed by subtracting a value from each DCT coefficient. The subtracted values then form the enhancement layer (see Fig. 2). If more than two scalable layers are required, the values subtracted for the creation of the base layer are further broken into other values. For example, if we want to transmit a coefficient with magnitude of level 9 using three layers, we can transmit level 5 as base layer, level 2 as first enhancement layer and level 2 as second enhancement layer. We have developed an optimal algorithm for determining the partitioning of the DCT coefficients which is based on Lagrangian optimization and Dynamic Programming.[11,12] The base layer contains the base layer DCT coefficients as well as all other information that is necessary for video decoding such as motion vectors, quantization parameters, coding modes, etc. The enhancement layers contain only the enhancement DCT coefficients.

3. WIRELESS CHANNELS

In this section we present essential elements on wireless channels for the further development and implementation of the scheme in Fig. 1. Wireless, or mobile, channels differ from the traditional Additive White Gaussian Noise (AWGN) and wired computer networks in the types of errors they introduce, as well as, in the severity of these errors. A characteristic feature of wireless channels is *multipath fading.* It is the resulting degradation when multiple versions of a signal are received from different directions at different times. Multipath fading occurs due to a number of factors of which some important ones are the presence and motion of objects reflecting the transmitted signal, the speed and motion of the receiver through this medium, and the bandwidth of the channel. Statistically, a wireless channel in which direct line-of-sight is available is referred to as a Rician fading channel where the distribution of the received signal follows a Rician probability density function. For the case in which no direct line-of-sight is available as in most urban areas we refer to this channel as a Rayleigh fading channel in which the distribution of the received signal follows a Rayleigh distribution. The Rayleigh probability density function (pdf) is given by

$$p(r) = \begin{cases} \frac{r}{\sigma^2} \exp(-\frac{r^2}{2\sigma^2}) & (0 \leq r \leq \infty) \\ 0 & (r \leq 0), \end{cases} \tag{1}$$

where σ^2 is the time averaged power of the signal before detection. In the work presented in this paper a Rayleigh fading channel is assumed.

Due to multiple reflections of the transmitted signal and the delay incurred with each reflected signal, the received signal is attenuated and delayed. Thus given a transmitted signal $u(t)$ over a slowly fading AWGN channel using Binary Phase Shift Keying (BPSK) modulation, the received signal $r(t)$ over a signaling period can be represented as

$$r(t) = \alpha(t) \exp^{-j\phi(t)} u(t) + z(t), \quad 0 \leq t \leq T, \tag{2}$$

where $z(t)$ is white gaussian noise, $\alpha(t)$ is the attenuation factor due to fading over the signaling period T and $\phi(t)$ is the phase shift of the received signal. For this signal the attenuation factor $\alpha(t)$ is a Rayleigh random process with the phase shift $\phi(t)$ being uniformly distributed over the interval $(-\pi, \pi)$.

In this paper, we assume that $\alpha(t)$ and $\phi(t)$ are constant over a signaling period. In the case of BPSK modulation over a fading channel, if the received signal's phase can be estimated from the signal for coherent demodulation, the received signal can be recovered with the use of a matched filter.[15]

4. 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,[16] 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 to the decoder and the distortion metric used is typically the Euclidean distance. In hard decoding, a decision for each received bit is made before Viterbi decoding. Thus, values of "0" and "1" are input to the 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.[17] 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.[18] 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 bit stream. 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.

5. OPTIMAL BIT ALLOCATION BETWEEN SOURCE AND CHANNEL CODING

It has now become clear that channel coding is necessary in order to provide reliable visual communications over a wireless channel. Thus, the available bit budget should be shared between source and channel coding. However, it is not obvious how the bit allocation should be performed. In this section we propose a way of optimally allocating the available bits between source and channel coding.[19]

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

$$\min D_{s+c} \text{ subject to } R_{s+c} \leq R_{budget}, \tag{3}$$

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 noise 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 L scalable layers, R_{s+c} is equal to

$$R_{s+c} = \sum_{l=1}^{L} R_{s+c,l}, \tag{4}$$

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

$$R_{s+c,l} = \frac{R_{s,l}}{R_{c,l}}, \tag{5}$$

where $R_{s,l}$ and $R_{c,l}$ are the source and channel rates, respectively, for 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, that is, $R_{s,l}$ and $R_{c,l}$ can only take values from discrete sets $\mathbf{R}_s^l$ and $\mathbf{R}_c^l$, respectively, i.e.,

$$\begin{aligned} R_{s,l} &\in \mathbf{R}_s^l \\ R_{c,l} &\in \mathbf{R}_c^l. \end{aligned} \tag{6}$$

We will now utilize Lagrangian optimization to solve the problem of Eq. (3). In order 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:

$$D_{s+c} = \sum_{l=1}^{L} D_{s+c,l}. \tag{7}$$

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 would be positive and the distortions for all other layers would be negative since inclusion of these layers reduces the Mean Squared Error (MSE). Of course, in the presence of channel errors, it is possible for the distortion of any layer to be positive since inclusion of a badly damaged enhancement layer can increase the MSE.

Another observation that should be made is that the differential improvement in the MSE that the inclusion of a scalable layer causes 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 to start with before the inclusion of the next scalable layer. Therefore, Eq. (7) can be written as

$$D_{s+c} = \sum_{l=1}^{L} D_{s+c,l}(R_{s+c,1}, \ldots, R_{s+c,l}). \tag{8}$$

We are now ready to utilize Lagrangian optimization in order to solve the optimization problem. The constrained problem of Eq. (3) is transformed into the unconstrained problem of minimizing

$$J(\lambda) = D_{s+c} + \lambda R_{s+c}. \tag{9}$$

Eq. (9) can be written as

$$\begin{aligned} J(\lambda) &= D_{s+c,1}(R_{s+c,1}) + D_{s+c,2}(R_{s+c,1}, R_{s+c,2}) + \dots \\ &+ D_{s+c,l}(R_{s+c,1}, R_{s+c,2}, \dots, R_{s+c,l}) + \lambda(R_{s+c,1} + R_{s+c,2} + \dots R_{s+c,l}) \qquad (10) \\ &= J_1(\lambda) + J_2(\lambda) + \dots J_l(\lambda), \qquad (11) \end{aligned}$$

where

$$\begin{aligned} J_1(\lambda) &= D_{s+c,1}(R_{s+c,1}) + \lambda R_{s+c,1} \\ J_2(\lambda) &= D_{s+c,2}(R_{s+c,1}, R_{s+c,2}) + \lambda R_{s+c,2} \\ &\vdots \\ J_l(\lambda) &= D_{s+c,l}(R_{s+c,1}, R_{s+c,2}, \dots, R_{s+c,l}) + \lambda R_{s+c,l}. \qquad (12) \end{aligned}$$

Assuming $J(\lambda)$ can be minimized, the value of λ for which $R_{s+c} = R_{budget}$ needs to be found. This is done iteratively using, for example, the bisection method, or more sophisticated algorithms, such as, the fitting of a Bezier curve.[20]

In order to reduce the computational complexity of the solution without losing optimality, we can exclude from consideration suboptimal Rate-Distortion pairs. That is, in minimizing $J(\lambda)$, we only use Rate-Distortion pairs that belong to the Operational Rate-Distortion Function (ORDF) of the corresponding layer. Thus, before minimization, we construct the ORDFs for each scalable layer, that is,

$$D^*_{s+c,1}(R_{s+c,1}), D^*_{s+c,2}(R_{s+c,1}, R_{s+c,2}), \dots, \text{ and } D^*_{s+c,l}(R_{s+c,1}, R_{s+c,2}, \dots, R_{s+c,l}).$$

This is done by obtaining all operating points and pruning out the suboptimal ones. Then, in the minimization, we only use operating points that belong to the ORDFs.

Our problem now reduces to finding the ORDFs for each scalable layer. Thus, we should experimentally obtain the expected distortion for each layer for all possible combinations of source and channel rates and all possible channel conditions. This would become prohibitively complex for even a small number of admissible source and channel rates and channel conditions. Thus, we have chosen to relax the optimality of the algorithm and utilize *Universal Rate-Distortion Characteristics* (URDC). These characteristic functions show the sensitivity of the scalable layers to channel errors. More specifically, they show the expected distortion per layer as a function of the bit error rate (after channel coding). These characteristics are obtained experimentally using representative video sequences. In order to obtain the characteristics, we introduce random errors on the bit stream using a Binary Symmetric Channel with a given bit error probability. As we can see in Fig. 3, the URDC is a set of curves, each one of them corresponding to a particular source coding rate for the particular scalable layer. The curves show the expected distortion plotted against $1/P_b$, the inverse of the bit error rate. The decision to use $1/P_b$ was made simply to make the plots convex, reminding of Rate-Distortion plots. It should be pointed out that since the ORDFs for all enhancement scalable layers also depend on the coding rates of the previous layers, the corresponding URDCs should also be conditioned on the source rates of the previous layers. As explained in the previous paragraph, the optimal solution of the problem as originally stated is prohibitively complex to solve. Thus, with the use of the Universal Rate-Distortion Characteristics we relax the optimality of the solution and search for the optimal solution given the Universal Rate-Distortion Characteristics.

We also need to obtain plots that show the channel bit error probability (bit error rate after channel coding) as a function of source coding rates and channel parameters. These should only depend on the channel model used and the channel codes. For the case of a Rayleigh fading channel assuming perfect interleaving, the channel parameter is the SNR per bit $\frac{E_b}{N_0}E[\alpha^2]$. We will call these plots *Channel Error Plots* (CEP). Fig. 4 shows such a plot. Each of the curves corresponds to a certain source coding rate. These plots can be obtained either experimentally or theoretically.

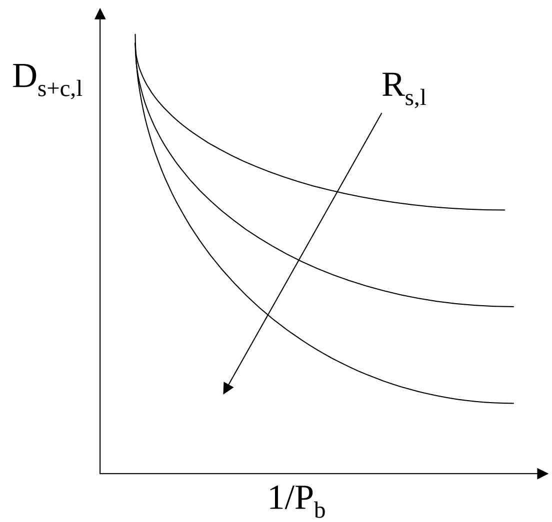

Figure 3. An example of Universal Rate-Distortion Characteristics.

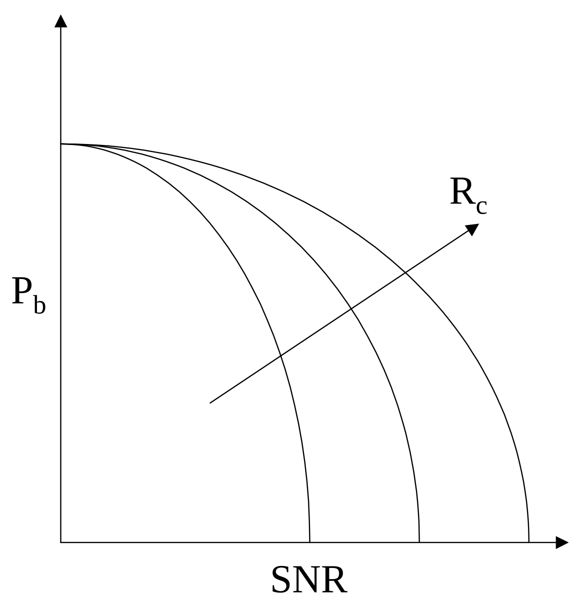

Figure 4. An example of a channel error plot.

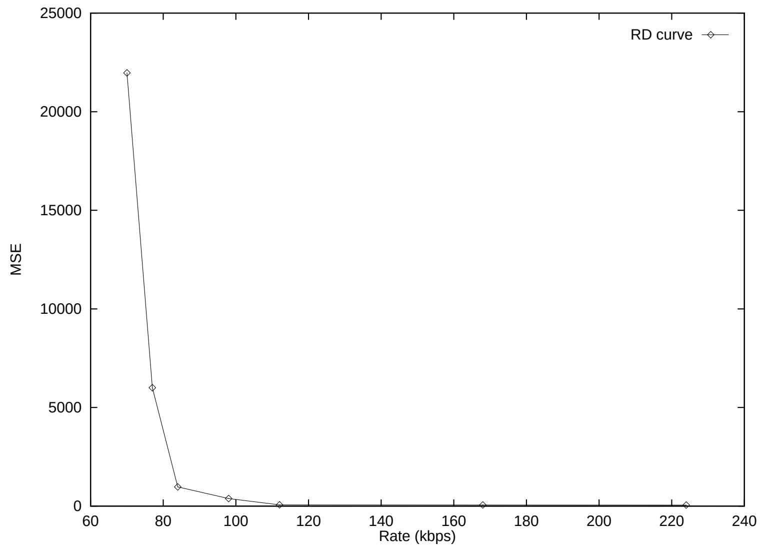

Figure 5. Overall Rate-Distortion plot for a 10 dB Rayleigh fading channel.

We now proceed to explain how to use the Universal Rate-Distortion Characteristics and the Channel Error Plots to construct the Operational Rate-Distortion Functions for each scalable layer. First, for the given channel parameters, we use the Channel Error 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. The Operational Rate-Distortion Function is found by determining the optimal Rate-Distortion operating points. It should be emphasized that the optimality is conditional on the Universal Rate-Distortion Characteristics.

6. EXPERIMENTAL RESULTS

We now present experimental results using the above methodology. We assume a Rayleigh fading channel with additive Gaussian noise and perfect interleaving (i.e., we assume that the samples of the Rayleigh random process α_k are independent identically distributed (i.i.d.)). The modulation scheme is Binary Phase Shift Keying (BPSK) with coherent detection. Rate-Compatible Punctured Convolutional codes (RCPC codes) are used for channel coding. We next present the results for the two-layer scalable video codec of section 2.

We chose the sets of admissible source rates for both layers to be

$$\mathbf{R}_s^1 = \mathbf{R}_s^2 = \{28, 42, 56\} \text{ kbps}. \tag{13}$$

The set of admissible channel rates was chosen to be

$$\mathbf{R}_c^1 = \mathbf{R}_c^2 = \{1/2, 2/3, 4/5\}. \tag{14}$$

We used RCPC codes with mother code rate 1/2 from Ref. 21.

The URDC plots we used were obtained by running 20 simulations for each bit error rate using a Binary Symmetric Channel. Fig. 5 shows the overall Rate-Distortion plot using the URDCs and the proposed algorithm for the 10 dB Rayleigh fading channel, while, Table 1 shows the selection of source and channel codes for the optimal operating points. It can be seen that most of the usable rate distortion points correspond to 1/2 channel code rate. The reason for this is that higher channel rates result in channel errors that cause much higher distortion than the quantization errors that are introduced by the lossy compression. Thus, using the 1/2 rate channel codes at the expense of lower source rates gives the optimal results.

Fig. 6 shows the overall Rate-Distortion plot for the 10 dB Rayleigh fading channel, while, Table 2 shows the selection of source and channel codes for the optimal operating points.

R^*_{s+c}	D^*_{s+c}	$R_{s,1}$	$R_{c,1}$	$R_{s,2}$	$R_{c,2}$
70	21968.16	28	4/5	28	4/5
77	6005.09	28	2/3	28	4/5
84	974.40	28	2/3	28	2/3
98	385.11	28	1/2	28	2/3
112	62.79	28	1/2	28	1/2
168	58.24	56	1/2	28	1/2
224	54.94	56	1/2	56	1/2

Table 1. Optimal rate allocation for two layer single-pass SNR scalable video over a 10 dB Rayleigh fading channel.

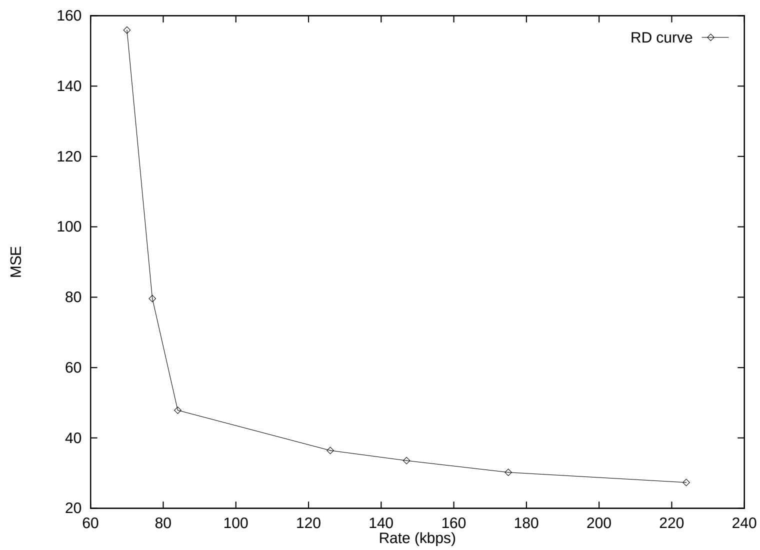

Figure 6. Overall Rate-Distortion plot for a 20 dB Rayleigh fading channel.

R^*_{s+c}	D^*_{s+c}	$R_{s,1}$	$R_{c,1}$	$R_{s,2}$	$R_{c,2}$
70	155.91	28	4/5	28	4/5
77	79.6	28	2/3	28	4/5
84	47.25	28	2/3	28	2/3
126	36.43	56	2/3	28	2/3
147	33.56	56	2/3	42	2/3
175	30.23	56	1/2	42	2/3
224	27.34	56	1/2	56	1/2

Table 2. Optimal rate allocation for two layer single-pass SNR scalable video over a 20 dB Rayleigh fading channel.

Figure 7. Frame 45 of the "Foreman" sequence from an experiment using the 10 dB Rayleigh channel.

Figure 8. Frame 175 of the "Foreman" sequence from an experiment using the 10 dB Rayleigh channel.

We now report some experimental results that were obtained using the Rayleigh channel model with perfect interleaving and the optimal bit allocation determined by our algorithm. Thus, the URDCs were not used for these results. For a 10 dB Rayleigh fading channel, we ran 20 experiments for the 112 kbps rate for which our algorithm suggests that we use 28 kbps source rate and 1/2 channel rate for both the base and the enhancement layer. The resulting MSE was 37.14. We repeated the same procedure for the 20 dB Rayleigh channel and the 126 kbps case. The algorithm suggests that we use 56 kbps source rate for the base layer, 28 kbps source rate for the enhancement layer and 2/3 channel rate for both layers. The resulting MSE was 29.9. It can be seen that these results produce a smaller MSE than the one predicted by the algorithm, therefore the algorithm overestimates the MSE by a small margin.

Fig. 7 shows frame 45 of the "Foreman" sequence from an experiment using the 10 dB Rayleigh channel. In this particular frame, all distortion is caused by source coding. Fig. 8 shows frame 175 of the same sequence. In this frame, distortion due to channel errors is clearly visible.

Figs. 9 and 10 show frames 45 and 175, respectively from an experiment using the 20 dB Rayleigh channel. We can see that there are no visible channel coding errors and the MSE is dominated by the lossy compression (quantization errors).

7. CONCLUSIONS

In this paper we considered video transmission over wireless channels and the problem of bit allocation between source and channel coding. This problem is a very important one. If, for given channel conditions, we allocate too few bits to channel coding, the quality of the received sequence will be very poor due to the channel errors. On the other hand, too many bits allocated to channel coding will reduce the available bits for source coding and the MSE will be dominated by quantization errors. Clearly, a bad allocation between source and channel coding can have severe effects on the quality of the received video sequence.

Figure 9. Frame 45 of the "Foreman" sequence from an experiment using the 20 dB Rayleigh channel.

Figure 10. Frame 175 of the "Foreman" sequence from an experiment using the 20 dB Rayleigh channel.

We introduced an algorithm for the bit allocation between source and channel coding and between scalable layers. This algorithm is optimal given the Universal Rate-Distortion Characteristics used. It also assumes that the distortion, as defined in section 5, is additive, i.e. the sum of the distortions per layer give us the total distortion. Also, for the construction of the URDCs, we assumed that the residual bit errors after channel coding can be simulated by a Binary Symmetric Channel with the same overall probability of error. Without these assumptions, the solution of the problem has a very high computational complexity.

Our results show that channel errors have a much larger effect on the MSE than quantization errors. Thus, the algorithm tends to recommend channel coding rates that keep bit errors to a minimum and reduce the artifacts that significantly increase the MSE.

REFERENCES

1. M. Bystrom and J. W. Modestino, "Combined source-channel coding schemes for video transmission over an additive white gaussian noise channel," Submitted to *IEEE Transactions on Image Processing*, 1997.
2. M. Bystrom and J. W. Modestino, "Combined source-channel coding for transmission of video over a slow-fading rician channel," in *Proceedings of the International Conference on Image Processing*, pp. 147–151, 1998.
3. G. Cheung and A. Zakhor, "Joint source/channel coding of scalable video over noisy channels," in *Proceedings of the International Conference on Image Processing*, vol. 3, pp. 767–770, IEEE, 1996.
4. J. Streit and L. Hanzo, "An adaptive discrete cosine transformed videophone communicator for mobile applications," in *Proceedings of the International Conference on Acoustics, Speech and Signal Processing*, pp. 2735–2738, IEEE, 1995.
5. L. Hanzo and J. Streit, "Adaptive low-rate wireless videophone schemes," *IEEE Transactions on Circuits and Systems for Video Technology* **5**, pp. 305–318, Aug. 1995.

6. M. Khansari, A. Jalali, E. Dubois, and P. Mermelstien, "Low bit-rate video transmission over fading channels for wireless microcellular systems," *IEEE Transactions on Circuits and Systems for Video Technology* **6**, pp. 1–11, Feb. 1995.
7. K. Ramchandran, A. Ortega, K. M. Uz, and M. Vetterli, "Multiresolution broadcast for digital HDTV using joint source-channel coding," *IEEE Journal on Selected Areas of Communications* **11**, pp. 6–23, Jan. 1993.
8. S. S. Prahdan and K. Ramchandran, "Efficient layered video delivery over multicarrier systems using optimized embedded modulations," in *Proceedings of the International Conference on Image Processing*, pp. 452–455, IEEE, 1997.
9. I. Kozintsev and K. Ramchandran, "Multiresolution joint source-channel coding using embedded constellations for power-constrained time-varying channels," in *Proceedings of the International Conference on Acoustics, Speech and Signal Processing*, pp. 2345–2348, IEEE, 1996.
10. 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.
11. L. P. Kondi, *Low Bit Rate SNR Scalable Video Coding and Transmission.* PhD thesis, Northwestern University, December 1999.
12. L. P. Kondi and A. K. Katsaggelos, "An optimal single pass SNR scalable video coder," in *Proceedings of the International Conference on Image Processing*, 1999.
13. M. A. Robers, L. P. Kondi, and A. K. Katsaggelos, "SNR scalable video coder using progressive transmission of DCT coefficients," in *Proc. SPIE Conference on Visual Communications and Image Processing*, pp. 201–212, 1998.
14. L. P. Kondi, F. Ishtiaq, and A. Katsaggelos, "A comparision of two SNR scalable video codecs," in *Proceedings of the International Conference on Image Processing*, pp. 934–938, IEEE, 1998.
15. T. S. Rappaport, *Wireless Communications*, Prentice-Hall, 1996.
16. G. D. Forney, "The Viterbi Algorithm," *Proceedings of the IEEE* **61**, pp. 268–278, March 1973.
17. 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.
18. J. Hagenauer, "Rate-compatible punctured convolutional codes (RCPC codes) and their applications," *IEEE Transactions on Communications* **36**, pp. 389–400, April 1988.
19. F. Ishtiaq, L. P. Kondi, and A. K. Katsaggelos, "Joint source-channel coding for scalable video." To be submitted to *IEEE Transactions on Image Processing*, 1999.
20. G. M. Schuster and A. K. Katsaggelos, *Rate-Distortion Based Video Compression, Optimal Video Frame Compression, and Object Boundary Encoding*, Kluwer Academic Publisher, 1997.
21. J. Hagenauer, "Rate-compatible punctured convolutional codes (RCPC codes) and their applications," *IEEE Transactions on Communications* **36**, Apr. 1988.