Robust Video Compression for Time-Varying Wireless Channels

Shankar L. Regunathan and Kenneth Rose
Dept. of Electrical and Computer Engineering,
University of California, Santa Barbara, CA 93106

ABSTRACT

We investigate joint source-channel coding for transmission of video over time-varying channels. We assume that the channel state is known at the receiver, but only statistical description of the time varying nature of the channel is available at the transmitter. A multimode coder is proposed to efficiently quantize the input video, and generate a quasi fixed-length bit stream of unequal importance. We vary the error protection offered to the individual bits, by matching it to both its importance, and the channel noise statistics. Based on the channel state, the decoder makes the best estimate of the source vector from the received codeword. We present a design algorithm which optimizes the overall rate-distortion performance of the system. Simulation results show that the proposed system outperforms a reference scheme where the multimode (source) codes and the channel codes were designed separately. Further, both the multimode coding schemes provide substantial gains over fixed length JSCC coding.

Keywords: Video Coding, Image Coding, Joint Source-Channel Coding, Wireless , Time-varying Channels

1. INTRODUCTION

Multimedia communications over wireless channels is becoming increasingly important due to emerging applications such as Personal Communication Networks. The fundamental obstacle to the design of such systems is the time-varying characteristics of the channel noise encountered over such links. It has been observed that the received signal to noise ratio may decrease by as much as 20-30 dB from its average value during periods of deep fades.[1] This event though infrequent, must be accounted for in the design as it may lead to drastic degradation in performance.

Conventional approaches to this problem have assumed independent design of the source and channel coders. For example, error correction codes (ECC) can be chosen to make effective channel nearly error free for the worst case channel condition, albeit at the cost of a substantial decrease in the effective rate available to the source coder. Obviously, this approach is too conservative and can significantly undermine the performance during the frequent and long periods of low channel noise. There is an additional drawback to separate source and channel coding. In several practical scenarios, some information about the channel condition is available at the receiver/transmitter. Separate design of source and channel codes does not take advantage of this information.

We consider the important case of a slowly varying channel where the channel error rate can be estimated at the receiver. We propose to achieve superior performance by designing the source and channel coding schemes jointly so as to exploit the channel state information (CSI) available at the receiver. Previous investigations in JSCC image/video coding for time-varying channels have been presented in[234].[5] However, these methods share one common shortcoming in that they restrict the source coder to employ fixed rate quantizers due to their inherent robustness to channel errors. However it is well established that video has highly non-stationary statistics. Fixed length coders are unable adapt to these changing statistics, and hence cannot provide efficient compression. Only a variable length code can offer the desired adaptation. On the other hand, JSCC schemes with variable length codes are difficult to design because of the notorious error propagation problem.

The central line of attack in our work is to use multimode codes[6] in the JSCC schemes to provide efficient compression, without sacrificing error resilience. Further, the source coder generates bits of varying importance, which are unequally protected by appropriate error correcting codes(ECC). The decoder uses the CSI to produce a soft estimate of the source which results in a graceful degradation in performance during periods of high channel noise. The bit allocation and the ECC assignment are adjusted so as to optimize the overall (source+channel) rate-distortion performance.

The paper is organized as follows: The time-varying channel model that we are considering in this work is described in section 2. The multimode source coder and the error protection scheme are described in section 3. We present a design algorithm for the rate-distortion optimization of the coder in section 4. Section 5 presents simulation results that demonstrate the superiority of JSCC designed over separately design source and channel coding systems.

2. CHANNEL MODEL

A sketch of a transmitter/receiver pair communicating over a time-varying channel is given in Figure 1. We consider the "flat" fading channel model which is commonly used to characterize the wireless channel. In flat fading, the spectral characteristics of the received signal are assumed to be fixed, while its strength changes with time. Assuming that the modulation is fixed, we can model the channel as a binary symmetric channel with varying channel bit error rate. Further, we assume that appropriate interleaving is used to render the channel memoryless. The channel bit error rate, ϵ, is assumed to take take one of N values, $\{\epsilon_i | i = 1, 2, ...N\}$ with probability $p(\epsilon_i)$. Note, in particular, that similar models have been used for characterizing GSM channels.

Let the fading be slow enough so that the error rate, ϵ_i, can be estimated accurately at the receiver. This is the channel state information (CSI). The transmitter does not know the exact condition of the channel (no feedback), but has a statistical knowledge of the channel error rate. The objective is to incorporate the CSI in the design of the overall system, so as to optimize its rate-distortion performance.

Note that the above formulation is readily applicable in the broadcast scenario with a single transmitter and multiple receivers. Further, the system can be extended to the case, where the CSI is fed back to the transmitter requiring one JSCC encoder/decoder pair for each transmitter state.

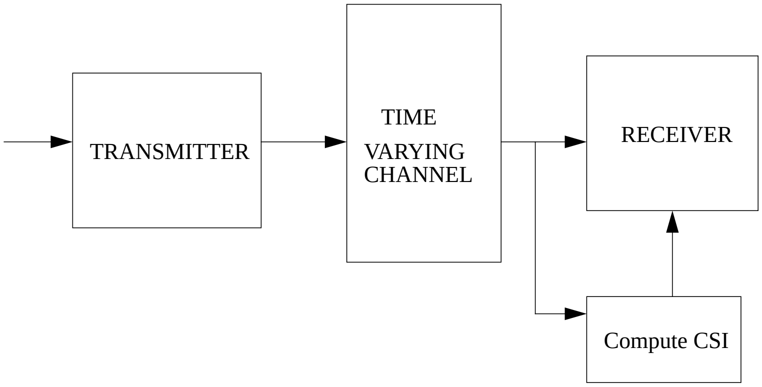

Figure 1. *JSCC scheme communicating over a time varying channel. Channel error rate is estimated at the receiver. Channel state estimation is shown outside the receiver for clarity.*

3. SYSTEM DESCRIPTION

The video signal is quantized by a hierarchical multimode coder (to be described below), whose output is a bit stream of varying importance. Bits of equal importance are grouped together, and an ECC is selected to protect each bit stream, depending on its importance and the channel noise statistics. At the decoder, the CSI is combined with the received codewords to make a soft estimate of the source. Refer to figure 2 for a sketch of the system.

3.1. MULTIMODE SOURCE CODING

Multimode codes were proposed in[6] as an efficient compromise between the compression efficiency of variable length codes and the robustness of fixed length codes. The multimode video coder is constructed from a simple predictive coding scheme. The video sequence is divided into groups of frames (GOF). The first frame in each GOF is intraframe

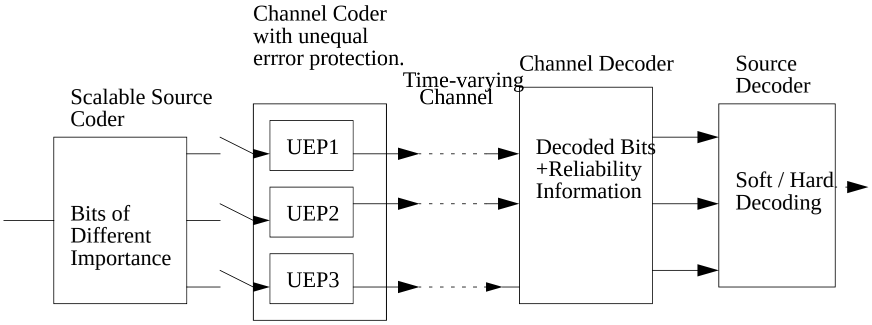

Figure 2. *Structured JSCC: Hierarchical quantization is followed by error protection matched to bit importance. CSI is estimated at the decoder to compute reliability metric, which can be used for soft decoding.*

coded(without recourse to prediction). The other frames in the GOF are predicted from the previous reconstructed frame, and the prediction error is s quantized. Thus, there is no temporal prediction across GOF boundaries. Although we did not use motion compensation in our prediction, it is easy to incorporate motion vectors by sending them as side information.

Each frame is divided into blocks and transformed by DCT. For encoding the DCT blocks, the multimode encoder is allowed to select the "best" bit allocation from a set of possible bit allocations. Each bit allocation, together with its corresponding fixed length quantizers and ECCs, constitutes a single mode of the coder. Each mode is, in fact, a specific fixed length code that is optimized for quantizing a particular subset of source statistics. By allowing the coder to switch modes for encoding each DCT block, we enable adaptation to statistics. The mode index is transmitted to the receiver, as heavily protected side information.

As long as the mode index is received error free, the decoder is perfectly synchronized to the encoder. Thus, any channel error in the fixed length code is confined to a single codeword and there is no error propagation. We use scalar quantizers for encoding the DCT coefficients within each mode. The codewords of the quantizers are indexed by the natural binary code and their bits are naturally of unequal importance. The codeword bits are protected by ECCs matched to their importance , and transmitted. At the decoder, the DCT coefficients are estimated from the received codewords, and inverse DCT is performed to reconstruct the frame.

A few carefully designed modes can provide sufficient flexibility in adaptation to provide efficient compression. At the same time, the quasi fixed-length operation of the coder eliminates error propagation and makes it feasible to directly optimize the overall rate-distortion cost. We confirmed through simulation, that the small loss in performance due to the heavy protection of the mode information is overwhelmingly offset by the gains in compression efficiency achieved through adaptation.

3.2. CHANNEL CODING

The channel encoder selects an error correction code to protect each bit produced by the source codeword. The selection is made from a family of rate compatible punctured convolutional (RCPC) codes, which can provide unequal error protection. Rate compatible puncturing allows the same channel encoder/decoder structure to be used for the entire set of codes, thereby reducing complexity.

Since the optimal assignment of ECCs is determined at the design phase, it is available at the receiver which can decode the ECC to obtain the source bits. Further, a reliability metric is computed for each bit. The reliability metric is defined as the effective error probability of the bit , given the current channel error rate (ϵ), and the RCPC code used for protecting that bit. The reliability metric for the different bits of a quantizer can be used to construct the probability transition matrix between the different codewords. The decoded codeword and the reliability metric are combined to compute the "best" estimate of the source vector.

Consider a DCT coefficient, x, which is quantized by a n-bit scalar quantizer. Let x belong to the encoder partition I, and i be the corresponding index produced by the quantizer. The channel error rate is assumed to be ϵ. Let the received index be j. The probability of transition $p(j|i,\epsilon)$ can be computed from the reliability metrics of the individual bits. Using the received index and transition probabilities, let the source decoder produces soft estimate y_j^ϵ. The total distortion, when the channel is in state ϵ, can be written as

$$D(\epsilon) = E||x - y_j^\epsilon||^2 = \sum_{j=1}^{N}\sum_{i=1}^{N}\int_{x\in I} ||x - y_j^\epsilon||^2 p(j|i,\epsilon)p(x)dx \tag{1}$$

For each received index j and channel error rate ϵ, the source decoder can minimize the distortion by computing the optimal estimate

$$y_j^\epsilon = \frac{\sum_{i=1}^{N} E\{x|x\in I\}P(x\in I)p(j|i,\epsilon)}{\sum_{i=1} P(x\in I)p(j|i,\epsilon)}, \tag{2}$$

and the resulting distortion for state ϵ, is given by

$$D(\epsilon) = \sigma^2 - \sum_{j=1}^{N} |y_j^\epsilon|^2 \sum_{i=1}^{N} P(x\in I)p(j|i,\epsilon). \tag{3}$$

The total distortion D is obtained by averaging over the probabilities of the channel states.

$$D = \sum_{\epsilon} D(\epsilon)p(\epsilon) \tag{4}$$

The design problem consists of selecting the bit allocation and RCPC allocation for all the modes so as to optimize the rate-distortion cost.

4. OPTIMIZATION WITHIN A RATE-DISTORTION FRAMEWORK

The objective is to minimize the total (source+channel) distortion, D, while satisfying the constraint on the total rate, R. Equivalently, the problem is formulated as an unconstrained minimization of the rate-distortion Lagrangian, $L = D + \lambda R$, where λ is the Lagrange parameter. Note that the Lagrangian cost is separable (as there is no error propagation), over the coefficients, and this makes its evaluation simpler.

The design algorithm for minimizing the Lagrangian cost of the overall system is summarized in the following steps:

Step 0. *A set of video frames is divided into GOFs to form the training set. Start with an initial set of mode indices, bit allocations and quantizers/ECCs for the coding modes.*

Step 1. *Partition the training set using the newly designed mode indices and fixed length codes. Form the DCT blocks of each frame. Assign each block to the coding mode which minimizes the Lagrangian cost of encoding its coefficients.*

Step 2. *Given the current partition, determine the optimal bit allocation maps and the corresponding ECC allocation for each mode so as to minimize the Lagrangian contribution of its training subset.. This is done independently for each DCT coefficient position as explained earlier.*

Step 3. *Redesign the mode indices based on the empirical rate of occurrence of each mode. We use Huffman coding followed by a rate 1/3 convolutional code for protecting the mode indices.*

Step 4. *If convergence criteria is met, stop. Otherwise go to step 1.*

In step 2, the optimal bit allocation and the ECC allocation can be determined independently for each coefficient using exhaustive search or greedy descent methods. Steps 1-3 of the algorithm minimize the Lagrangian. Due to the temporal dependencies in encoding the training set, convergence can not be ensured theoretically. However, simulations confirm that convergence is not a practical problem and the algorithm descends in the Lagrangian cost. Note that some of the modes may become unpopulated during the design procedure so that, ultimately, a smaller number of *effective* modes is obtained. Note further that the conventional fixed length code is the special suboptimal case where the coder is constrained to have only a single mode.

5. RESULTS

In this section, we present simulation results to demonstrate the performance of multimode JSCC for video compression over time-varying channels. The multimode source codes and the channel protection were jointly designed to optimize the rate-distortion cost, as described in the previous section. For comparison, we used a multimode video coding system where the source and channel codes were designed independently. As an addition benchmark, we also presents the results for fixed length JSCC of video.

In our simulations, we used a memoryless BSC, whose bit error rate can take the values, $\{\epsilon = 0.0, 10^{-4}, 10^{-3}, 5 * 10^{-3}, 10^{-2}\}$. The time variation in the channel condition is modeled by the probability distribution over the channel states, $p(\epsilon)$, which is summarized in table 1. The receiver knows the exact channel error rate, while the transmitter only has knowledge of the statistics, $p(\epsilon)$.

$\epsilon ->$	0.0	10^{-4}	10^{-3}	$5 * 10^{-3}$	10^{-2}
$p(\epsilon)$	0.3	0.2	0.2	0.15	0.15

Table 1. *Probability of channel error, $p(\epsilon)$, for different channel error rate conditions, ϵ.*

We used the RCPC codes described in[7] for error protection. The mode information was protected by a 1/3 code, which guarantees an error rate of less than 10^{-14}, even for the highest bit error condition of the channel. Therefore, we assumed the mode information to arrive effectively error free at the receiver in our simulations. In practice, some resynchronization mechanism can be used to recover from this extremely unlikely event of corrupted mode information. The results are summarized in table 1.

The sequence *Carphone* was used to design the coder. The performance of the system was tested on the sequence *salesman* and *Grandma* The average PSNR of the reconstructed frames at the decoder for various rates is listed in table 2 and table 3. Note, that the rates are all inclusive and, in particular, include all the bits spent on side information and channel error protection.

Total Rate in kbps	Multimode Joint Source-Channel Coding	Multimode Separate Source-Channel Coding	Fixed Length Source-Channel Coding
64	32.92	32.66	28.50
96	34.49	34.03	28.85
128	35.82	34.65	30.16
160	36.81	35.85	30.71

Table 2. *Performance of multimode JSCC coding on the sequence Grandma. Average PSNR of reconstructed frame at various target rates is shown.. For comparison, performance of multimode coding with separate design of source and channel codes, and JSCC fixed length codes is shown.*

The results show that both the multimode coding schemes substantially outperform the fixed length coding system with gains of up to 5 dB in average PSNR of reconstructed video. Further, joint design of source and channel codes in the multimode video coder can yield an additional 1 dB in PSNR over separate design. There was a corresponding increase in the perceptual quality of the reconstructed frames as seen in figures 3 and 4. Thus JSCC multimode video coding is a very good candidate for video compression over time-varying channels.

6. CONCLUSIONS

We have proposed a multimode JSCC scheme for transmitting video over time-varying wireless channels. The scheme makes use of the channel state information (CSI) available at the decoder to achieve superior performance. Multimode source codes enable efficient trade-off between compression efficiency and robustness. The source quantizer produces bit stream of varying importance, which are unequally protected by RCPC codes. We proposed an algorithm to design the bit allocation and the RCPC allocation so as to optimize the rate-distortion performance. The scheme was shown to outperform separate source and channel coding as well as fixed length JSCC. Future work includes extension to the case where the CSI is fed back to the transmitter.

REFERENCES

1. T.S. Rappaport, "Wireless Communications Principles and Practice," Prentice Hall.
2. I. Kozintsev and k. Ramchandran, "Robust Image Transmission over Energy-Constrained Time-Varying Channels Using Multiresolution Joint Source-Channel Coding," IEEE Transactions on Signal Processing, vol. 46, pp. 1012-26, April 1998.
3. N. Chaddha and S. Diggavi, "A frame-work for joint source-channel coding of images over time-varying wireless channels," *Proc. ICIP' 96*, 1996, vol. 2, pp. 89-92.
4. L. C. Yun and D. G. Messerschmitt, "Digital video in a fading wireless environment," *Proc. ICASSP* , 1996, vol. 2, pp. 1069-72.
5. B. Girod, U. Horn, B. Belzer, "Scalable Video Coding with Multiscale Motion Compensation and Unequal Error Protection," *Proc. Symp. on Multimedia Communications and Video Coding*, New York City, NY, October 11-13, 1995.
6. S.L. Regunathan, K. Rose and S. Gadkari, "Multimode Image Coding for Noisy Channels," *Proc. ICC*, 1997, vol. 3, pp. 1644-8.
7. J. Hagenauer, "Rate-compatible punctured convolutional codes (RCPC codes) and their applications," *IEEE Trans. on Com.*, vol. 36, pp. 389-400, Apr. 1988.

Figure 3. Sample reconstructed frame of “Grandma” for Multimode-JSCC (top), Multimode separate source and channel coding (middle), and Fixed Length-JSCC (bottom) for rate $R = 160$ Kbps

Total Rate in kbps	Multimode Joint Source-Channel Coding	Multimode Separate Source-Channel Coding	Fixed Length Source-Channel Coding
64	29.76	28.79	27.29
96	31.34	30.97	28.24
128	32.75	32.22	29.34
160	33.76	33.06	29.84

Table 3. *Performance of multimode JSCC coding on the sequence Salesman. Average PSNR of reconstructed frame at various target rates is shown.. For comparison, performance of multimode coding with separate design of source and channel codes, and JSCC fixed length codes is shown.*

Figure 4. Sample reconstructed frame of "Salesman" for Multimode-JSCC (top), Multimode separate source and channel coding (middle), and Fixed Length-JSCC (bottom) for rate $R = 160$ Kbps