# Scalable Video Transmission Over Wireless Channels Using 3-D SPIHT and LDPC Codes

Manu Bansal, Lisimachos Kondi,
Dept. of Electrical Engineering, SUNY at Buffalo, Buffalo, NY 14260
{mbansal, lkondi}@eng.buffalo.edu

*Abstract*— **In this paper, an important problem of efficient utilization of the available resources for video transmission over wireless channels is investigated. The popular 3-D Set Partitioning In Hierarchical Trees (3-D SPIHT) is used to produce a completely embedded video bit stream. A new channel coding technique employing low-density parity check (LDPC) codes and rate-compatible punctured convolutional (RCPC) codes/ cyclic redundancy check (CRC) is proposed for error-detection and correction. A flat-fading Rayleigh channel with additive white gaussian noise (AWGN) is modeled for transmission.**

**The embedded bit stream is given unequal error protection (UEP) and the rate-distortion optimization algorithm is developed and carried out for the selection of source coding rate and the channel coding rate. The experimental results demonstrate the effectiveness of this system under different wireless channel conditions for video transmission and are seen to be better than the classical technique of using RCPC/CRC.**

## I. INTRODUCTION

In the recent past, there has been a tremendous increase in the capabilities of wireless multimedia devices and services. The demand to improve quality of such systems within the limited bandwidth resources motivates the interest in error-resilient multimedia coding methods. This research problem is best dealt by choosing, 1) a scalable source codec with good compression efficiency. A scalable bit stream can be produced by encoding the source only once for different quality requirements and can be decoded at different rates with progressive reconstruction quality, and, 2) an appropriate set of channel codes for efficient forward error correction (FEC).

Various methods for wavelet-based image and video transmission over wireless channels have been discussed in [1], [2], [3], [4]. Mostly, these are based on 2-D/3-D set partitioning in hierarchical trees (SPIHT) source codec [5], [6] and channel coding employing rate compatible punctured convolutional (RCPC) codes [7].

In this paper, we propose a wireless transmission system for scalable video over flat-fading Rayleigh channel. The embedded 3-D SPIHT codec [6] is used for video coding and a new scheme for channel coding using the serial concatenation of low-density parity check (LDPC) codes [8], [9] and RCPC/ cyclic redundancy check (CRC) [10] codes is proposed. This channel coding method provides unequal error protection (UEP) to the bit stream divided into packets of constant size. The problem of optimal resource allocation is then formulated and solved using the Lagrangian multiplier method [11], [12], [13] for the choice for different transmission rates. Performance of this system is evaluated for different channel conditions and simulation results outperform those using the classical technique of RCPC/CRC.

The rest of the paper is organized as follows. Section II and III gives an overview of the source and channel codes, respectively, used in this work. Section IV discusses the optimization technique for efficient bit allocation and Section V, presents experimental results. Section VI concludes the paper.

## II. EMBEDDED VIDEO CODING USING 3-D SPIHT

3-D SPIHT [6] is a wavelet-based embedded video coder employing subband coding (SBC) technique. Basically, it is a 3-D extension of the highly successful 2-D SPIHT [5] image codec. The underlying SPIHT algorithm is based on the concepts of first forming the *spatial orientation trees*, which are groups of wavelet transform coefficients organized into trees rooted in the lowest frequency or coarsest scale subband with offspring in several generations along the same spatial orientation in the higher frequency (resolution) subbands. This structure of a tree exploits the self-similarity and magnitude localization properties of the wavelet transformed image. It is assumed that if a coefficient magnitude in a certain node of a spatial orientation tree does not exceed a given threshold, it is very likely that none of its descendants will exceed that threshold. Further, the wavelet transform coefficients in these trees are tested against a magnitude threshold $2^n$, where $n$ is called the level of significance (i.e., bit-plane coding, starting with the most-significant bit) and partitioned into different sets according to their significance. Finally, the sign bits and refinement bits (for the coefficients that are tested to be significant earlier) are coded and transmitted for that value of $n$. The value of threshold is successively lowered by power of 2 and the process terminates when the desired rate or quality level is reached.

The basic algorithm for 3-D SPIHT is same as 2-D SPIHT, except that a 3-D spatial-temporal tree structure is formed here using the wavelet transform coefficients for a group of frames (GOF). Under certain conditions, 3-D SPIHT codec (even without motion-compensation) performs comparable to H.263 and outperforms MPEG-2. As mentioned earlier, 3-D SPIHT codec produces an embedded bit stream which means that every lower rate bit stream is a prefix of higher rate bit streams and hence, is totally rate scalable. This helps in progressive transmission of the bit stream. It also provides multiresolution scalability, precise rate control and low complexity that makes it an attractive codec for wireless transmissions.

## III. CHANNEL MODEL AND CODING

A flat-fading Rayleigh channel with additive white gaussian noise (AWGN) and perfect interleaving (i.e., we assume that the samples of the Rayleigh random process $\alpha$ are independent identically distributed (i.i.d.)) is modeled for the transmission.

$$r(t) = \alpha(t)s(t) + n(t) \tag{1}$$

where $r(t)$ is the received signal, $s(t)$ is the transmitted signal, $\alpha$ is a Rayleigh distributed random process and $n(t)$ is the AWGN.

### *A. Forward Error Correction*

While 3-D SPIHT codec gives a very good compression efficiency, due to it's embedded nature, even a single bit error can lead to the unrecoverable video at the decoder. When transmitting video over an unreliable channel, channel coding in the form of FEC is needed for error detection and correction. As mentioned earlier, we have used a new channel coding technique employing the *serial concatenation* of LDPC codes [8], [9] and RCPC codes [7]/ CRC code [10]. This coding scheme is used to provide UEP to the bit stream divided into packets of constant size.

*1) LDPC Codes:* LDPC codes were first proposed by Gallager in his 1960 PhD. dissertation at MIT [8] and was scarcely considered in many years that followed due to heavy hardware requirements. The study of LDPC codes was resurrected in the mid-1990's with the work of Mackay [9]. Basically, LDPC codes are a class of block codes which provide near-capacity performance on a large collection of data transmissions while simultaneously admitting an implementable decoder. LDPC codes have a very sparse parity check matrix $H$ (i.e., very few 1's and mostly 0's) and it is generated by applying random perturbations to the zero matrix until a specified number of ones appear in each column and roughly fixed equal number of ones appear in each row. Such LDPC codes are called *regular* LDPC codes. The associated generator matrix $G$ is obtained by Gaussian elimination of $H$. The sparseness of $H$ eliminates the need for interleavers at both encoder and decoder and also, facilitates the faster decoding of such codes for even large block lengths.

The decoding of LDPC codes is done by an iterative probabilistic algorithm known as the *belief-propagation* or *sum-product* algorithm. This algorithm is better understood with the help of Tanner graphs [14]. It starts with some initial probabilities of code bits and iteratively updates these probabilities based on message-passing and performes parity checks until all the parity checks are satisfied or a predetermined maximum number of iterations are done. The decision is then taken on all the received bits based on the final probabilities values.

*2) RCPC Codes:* In this coding scheme, RCPC codes [7] are used to provide the UEP to the bit stream. RCPC codes form a class of convolutional codes that are obtained by puncturing the output of a "mother" convolutional code. More details on RCPC codes can be obtained from [7]. RCPC codes are decoded using the Viterbi algorithm which is a maximum-likelihood sequence estimation procedure.

*3) CRC Codes:* CRC codes are used in this work for error detection. Details on CRC codes can be obtained from [10].

## IV. OPTIMAL RESOURCE ALLOCATION

*Objective*: Given an overall bit rate $R_{budget}$, the aim is to distribute the bits between source and channel coding such that the overall mean square distortion $D_{s+c}$ (mean square error (MSE)) is minimized or on the other hand, the overall peak signal-to-noise ratio (PSNR) is maximized $(PSNR = 10log_{10}(255^2/MSE))$. This is denoted by the following equations:

$$\min D_{s+c} \text{ subject to } R_{s+c} \leq R_{budget} \tag{2}$$

and

$$\max P_{s+c} \text{ subject to } R_{s+c} \leq R_{budget} \tag{3}$$

The $R_{s+c}$ term symbolizes the total bit rate used by source and channel coding for all (constant size) packets and the $D_{s+c}$ is the resulting distortion for both channel and source coding and it depends on both the channel and the encoder - decoder configuration. On the other hand, $P_{s+c}$ is the resulting video quality and is inversely proportional to the distortion. To be more specific, the distortion caused by the source coding is due to quantization and is deterministic. The distortion due to channel errors is stochastic. Therefore, the total distortion is also stochastic and the expected value of it is used, which is given as:

$$E_N(D_{s+c}) = \sum_{i=0}^{N} P_i(R)D_i(R), \tag{4}$$

where $N$ is the total number of packets,
$R$ is the rate allocation vector, $(rv_1, rv_2, \ldots, rv_N)$, which assigns to each packet $i$, $i = 1, 2, \ldots, N$, a channel code rate $rv_i$. Each of these channel code rates are chosen from a set $RV$ of $m$ channel code rates $(r_1 < r_2 < \ldots < r_m)$ used by the combination of LDPC and RCPC,
$D_0(R)$ is the distortion when none of the packets are received error free and is equal to the source variance, and for $i \geq 1$, $D_i(R)$ is the reconstruction distortion using the first $i$ packets,
For $i = 1, \ldots, N-1$, $P_i(R) = (\prod_{j=1}^{i}(1 - p(rv_j)))p(rv_{i+1})$, is the probability of no errors in the first $i$ packets but with an error in the next one, $P_0(R) = p(rv_1)$ is the probability of an error in the first packet, and $P_N(R) = \prod_{j=1}^{N}(1 - p(rv_j))$ is the probability that all the $N$ packets are correctly received.
For $i = 1, 2, \ldots, m$, $p(r_i)$ is the probability of a decoding error in a packet protected by the channel code rate $r_i$ and $p(r_i) = 1 - (1 - BER)^l$, where $l$ is the packet size and $BER$ is the bit error rate after channel decoding.

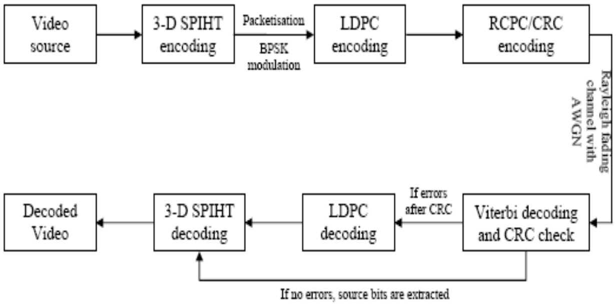

Fig. 1. Block diagram of system employing concatenated LDPC+RCPC/CRC coding scheme.

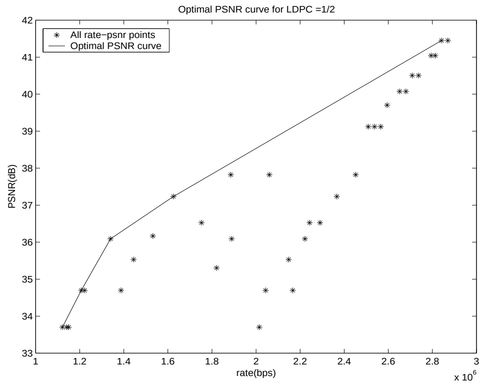

Fig. 2. R-D operating points and the optimal curve for PSNR at LDPC =1/2 rate and channel SNR=5 dB.

As discussed above, the problem of efficient UEP and hence the optimal bit allocation for this scheme is formulated as a constrained optimization problem and is then converted into an unconstrained one by using the Lagrangian optimization [11], [12], [13], i.e.,

$$\min J(\lambda) = D_{s+c} + \lambda R_{s+c} \tag{5}$$

where $\lambda$ is the Lagrange multiplier.

The optimization works on minimizing $J$ and hence the overall expected distortion (MSE) for a given channel condition by efficiently allocating the channel coding rates (from the set $RV$ of channel code rates) to the packets to meet a target transmission rate, $R_{budget}$. The rate-distortion (R-D) operating points are found by allocating combinations of channel codes across the source packets (protection strictly descending in nature) and using Eq. (4). Then the correct Lagrangian multiplier, $\lambda$ is found (using the bisection algorithm) to get the optimal operating R-D point lying on the convex hull of the overall R-D characteristic plot.

## V. EXPERIMENTAL RESULTS

The 3-D SPIHT bitstream for each group of frames (GOF) is partitioned into “fixed-length” packets . At the receiver, if after Viterbi and CRC decoding, the packet is found to be corrupted, LDPC decoding is done to recover the data, else, due to the systematic form of the LDPC codes, the source bits are immediately extracted and source decoded. This reduces the overall complexity of the system. Fig. 1 shows the block diagram of the whole system.

Following the block diagram, a color “Foreman” sequence (QCIF format) is encoded using 3-D SPIHT codec at the source coding rate of 760 kbps. The sequence consists of 300 frames, each of size 176 $\times$ 144 and is encoded on GOF-by-GOF basis at 30 frames/sec, i.e., each group of 16 frames is encoded and transmitted independently. The bit stream corresponding to each GOF is packetised into packets of equal length of 4110 bits. The operational R-D data for the source are found by decoding the same bit stream packet-by-packet. This data is later used to find the expected distortion at receiver using Eq. (4). All the packets are then channel coded, first by LDPC code of constant rate = 1/2 or 2/3, and then followed by RCPC code of rates (8/9, 8/10, 8/16 and 8/24) for UEP. The puncturing tables and other configurations for these RCPC code rates are taken from [7]. 16-bit CRC is then generated and appended to each packet. Binary phase shift keying (BPSK) is used for modulating each packet and transmission through a Rayleigh fading channel with AWGN is then simulated [15]. The performance of the system is evaluated at three different channel SNR values, i.e., 5, 10 and 15 dB. At the receiver, if needed, LDPC decoding is performed with maximum of 500 iterations. All the distortion and PSNR results obtained here are computed on GOF-by-GOF basis and averaged over the GOFs.

Fig. 2 shows all the operating R-D points and the curve joining the optimal PSNR points for wireless video transmission with channel SNR = 5 dB. We can see that the optimal points lie on the convex hull of the R-D curve. Also, some of the optimal points are missed in the output of the optimization as the Lagrangian method always chooses the points lying on the convex hull. Fig. 3 displays the comparison between the PSNR optimization curves obtained at channel SNR = 5, 10 and 15 dB and constant rate LDPC = 1/2. The comparison clearly shows the improvement in the performance of the system as the channel improves, i.e., as the channel improves, higher PSNR is obtained for the same transmission rate. A similar comparison is shown in Fig. 4 for constant rate LDPC = 2/3 rate. Figs. 5 and 6 compare the performances and optimization curves obtained from this scheme of concatenated LDPC+RCPC/CRC code and the classical scheme of RCPC/CRC code at SNR = 15 dB and LDPC = 1/2. It is clearly evident that our scheme significantly outperforms the classical scheme for the same overall channel protection and at a given transmission rate.

## VI. CONCLUSIONS

The performance of the fully scalable video coding (using 3-D SPIHT algorithm) and transmission over flat-fading Rayleigh wireless channel was developed and studied. A new channel coding scheme, i.e., serially concatenated LDPC and RCPC/CRC code was proposed and its performance was compared against classical scheme employing RCPC/CRC. A rate-

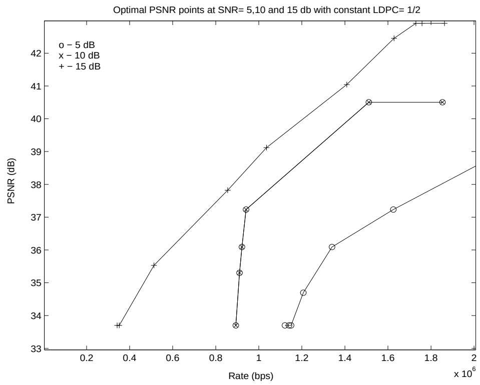

Fig. 3. Comparison of R-D optimal curves for PSNR at LDPC =1/2 rate at channel SNR=5 ,10 and 15 dB.

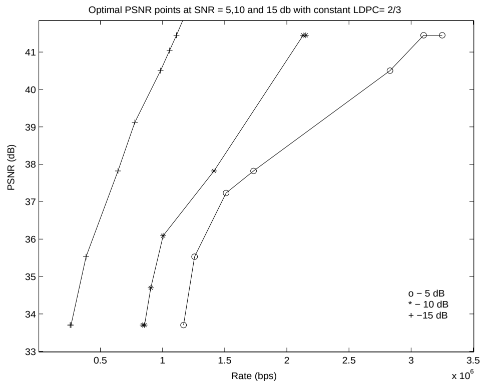

Fig. 4. Comparison of R-D optimal curves for PSNR at LDPC =2/3 rate at channel SNR=5 ,10 and 15 dB.

distortion (R-D) optimization method using the Lagrangian multiplier was developed and carried out for the choice of number of packets (i.e., source rate) to be transmitted and the channel coding rates. The R-D optimization was performed for different channel conditions and the results showed a significant improvement in the performance as the channel improves. It was also shown that the proposed scheme outperforms the RCPC/CRC scheme for the same overall channel protection and transmission rate.

## REFERENCES

[1] Z. Xiong, B. Kim and W.A. Pearlman, “Progressive video coding for noisy channels,” *Proceedings of International Conference on Image Processing*, Volume:1, pp. 334-337, October 1998.

[2] P.G. Sherwood and K. Zeger, “Error protection for progressive image transmission over memoryless and fading channels,” *IEEE Transactions on Communications*, Volume:46, Issue:12, pp. 1555-1559, December 1998.

[3] D.G. Sachs, R. Anand and K. Ramchandran, “Wireless image transmission using multiple-description based concatenated codes,” *Data Compression Conference Proceedings*, March 2000.

[4] V. Chande and N. Farvardin, “Progressive transmission of images over memoryless noisy channels,” *IEEE Journal on Selected Areas in Communications*, Volume:18, Issue:6, June 2000.

[5] A. Said and W.A. Pearlman, “A new fast, and efficient image codec based on set partitioning in hierarchical trees,” *IEEE Transactions on Circuits and Systems for Video Technology*, Volume:6, Issue:3, pp. 243-250, June 1996.

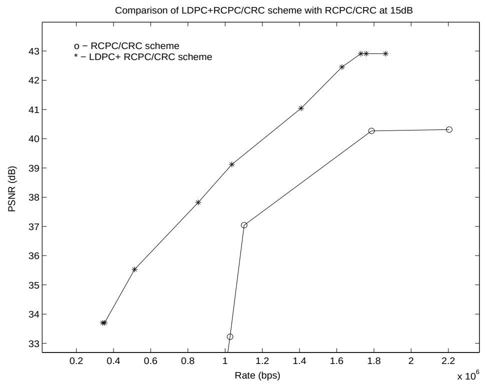

Fig. 5. Comparison of R-D optimal curves for PSNR between LDPC+RCPC/CRC scheme and RCPC/CRC scheme.

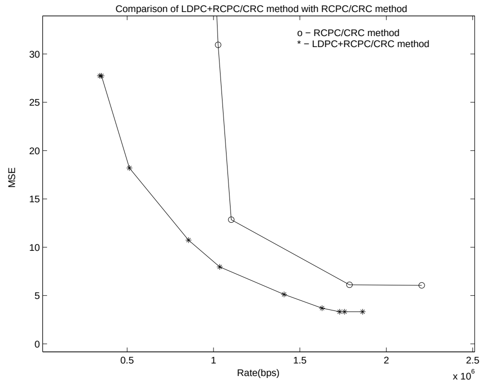

Fig. 6. Comparison of R-D optimal curves for MSE between LDPC+RCPC/CRC scheme and RCPC/CRC scheme.

[6] B. Kim, Z. Xiong and W.A. Pearlman, “Very low bit-rate scalable video coding with 3-D set partitioning in hierarchical trees (3-D SPIHT),” *IEEE Transactions on Circuits and Systems for Video Technology*, Volume:10, Issue:8, December 2000.

[7] J. Hagenauer, “Rate compatible punctured convolutional codes (RCPC codes) and their applications,” *IEEE Transactions on Communications*, Volume 36, pp. 389-399, April 1988.

[8] R.G. Gallager, “Low density parity check codes,” *IEEE Transactions on Information Theory*, Volume IT-8, pp. 2128, January 1962.

[9] D.J.C. Mackay, “Good error-correcting codes based on very sparse matrices,” *IEEE Transactions on Information Theory*, pp. 543-547, March 1999.

[10] W.H. Press, B.P. Flannery, S.A. Teukolsky and W.T. Vetterling, “Cyclic redundancy and other checksums,” *Ch. 20.3 in Numerical Recipes in FORTRAN: The Art of Scientific Computing*, 2nd ed. Cambridge, England: Cambridge University Press, pp. 888-895, 1992.

[11] A. Ortega and K. Ramchandran, “Rate distortion methods for image and video compression,” *IEEE Signal Processing Magazine*, Volume:15, Issue:15 , pp. 23-50, November 1998.

[12] L.P. Kondi, F. Ishtiaq, A.K. Katsaggelos, “Joint source-channel coding for motion-compensated DCT-based SNR scalable video,” *IEEE Transactions on Image Processing*, Volume:11, pp. 1043-1052., September 2002.

[13] L.P. Kondi, D. Srinivasan, D.A. Pados, S.N. Batalama, “Layered video transmission over multirate DS-CDMA wireless systems,” *IEEE Transactions on Circuits and Systems for Video Technology*, under review.

[14] R.M. Tanner, “A recursive approach to low complexity codes,” *IEEE Transactions on Information Theory*, pp. 399-431, September 1981.

[15] G. Partasides, L.P. Kondi, “Scalable video transmission over orthogonal frequency division multiplexing wireless channels,” *Proceedings IEEE Canadian Conference on Electrical and Computer Engineering*, Niagara Falls, Canada, May 2004.