SCALABLE VIDEO TRANSMISSION OVER ORTHOGONAL FREQUENCY DIVISION MULTIPLEXING WIRELESS CHANNELS

George Partasides, Lisimachos P. Kondi

Dept. of Electrical Engineering
University at Buffalo
The State University of New York
Buffalo, NY 14260
e-mail: {gap4,lkondi}@eng.buffalo.edu

Abstract

In this paper, we optimize a wireless system for video transmission over Orthogonal Frequency Division Multiplexing (OFDM) systems. A frequency-selective Rayleigh fading channel is assumed. Block transmissions using a cyclic prefix are used to eliminate Inter Block interference (IBI). Quadrature Amplitude Modulation (QAM) is used as modulation scheme. The MPEG-4 video coder is used to provide a two-layer SNR scalable video bit stream. Rate Compatible Punctured Convolutional codes (RCPC codes) are used for error protection.

In this work we apply Unequal Error Protection (UEP) to the layers of the scalable bitstream. We solve the problem of optimal bit allocation between source and channel coding and across scalable layers such that the overall distortion is minimized. The algorithm utilizes universal rate distortion characteristics which are obtained experimentally and show the sensitivity of the source encoder and decoder to channel errors.

Keywords: *Wireless Video Transmission, OFDM, scalable video.*

1. INTRODUCTION

Recently there was a rapid change in technology in the field of computation speed as well as an increasing interest in multimedia communications over wireless channels. These two factors give rise to new wireless techniques such as orthogonal frequency division multiplexing (OFDM). Nowadays OFDM is used in Digital Video and Audio Broadcasting (DVAB) and high speed digital subscriber line (DSL) modems over twisted pair lines. OFDM has also been used in wireless networks such as the IEEE802.11 a/b/g and IEEE802.16.

In this paper MPEG-4 scalable video transmission is considered over a wireless OFDM channel. Similar work was done before for single user Rayleigh fading channels [2] and direct-sequence code division multiple access DS-CDMA wireless channels [3], where an approach toward joint source-channel coding for motion-compensated DCT-based scalable video coding and transmission was developed. In this work we use Unequal Error Protection (UEP) in our video sequence stream because the various parts of the compressed video bit stream are not equally important to the quality of the reconstructed video sequence.

In addition to this, in our wireless OFDM channel we use block transmissions and cyclic prefix to prevent our symbols from the inter block interference (IBI) caused by multipath. Furthermore, we use rate-compatible punctured convolutional (RCPC) codes for channel coding.

For the receiver we use a slicer detector since we utilize a QAM4 modulation scheme, and the Viterbi algorithm for channel decoding.

2. VIDEO TRANSMISSION SYSTEM

2.1 Scalable Video Coding

A scalable video encoder produces a bit stream that consists of embedded subsets. Each embedded subset can be decoded and produce a video sequence of a certain quality. If the video stream is scalable a user with high bandwidth connection can download the entire bit stream and have a high quality video,

CCECE 2004- CCGEI 2004, Niagara Falls, May/mai 2004
0-7803-8253-6/04/$17.00

while a user with a lower bandwidth can download a part of the video bit stream and have video with a lower quality. There are three types of scalability. (i) signal to noise ratio (SNR) scalability. (ii) spatial scalability and (iii) temporal scalability. *SNR scalability or quality scalability* is the representation of video with varying resolutions. This is done by quantizing the pixel values with increasingly finer quantization step sizes leading to different peak SNRs between the original and quantized video. Decoding only the base layer we get a low quality version of the video sequence. By also decoding the remaining layers (Enhancement layers) we increase the video quality. *Spatial Scalability* represents the same video sequence in varying spatial resolutions. By decoding the first layer we can display a version of the decoded video frame at low resolution. By keeping decoding more layers we get a higher resolution video sequence. *Temporal scalability* is the procedure where the same video sequence is represented in a varying resolutions or frame rates. Thus, with temporal scalability we can have different frame rates for different layers of the contents.

An important application of scalability is in error resilient video transmission where stronger error protection is applied to the base layer than to the enhancement layers. This way we can decode the base layer with higher probability.

2.2 Channel Coding

Rate-compatible punctured convolutional (RCPC) codes are used in this work for channel coding. RCPC can provide more redundancy in some groups of information than other groups.

Convolution is the process of module-2 addition of the current source bit with the previously delayed source bits. Convolutional coding generates one codeword for the entire source data. The rate of a Convolutional code is defined as $k/n < 1$ where k is the number of input bits and n is the number of output bits.

In the decoder we used the Viterbi algorith which is a maximum-likehood sequence estimation procedure.

3. OFDM CHANNEL

The channel assumed is a frequency selective Raleigh fading channel. This causes inter-symbol interference (ISI) and fading effects that arise due to mobility and carrier offsets. This can be adjusted by using block transmissions and cyclic prefix (CP).

3.1 Block Transmissions and Equalization

To tone down the time-domain effects that give rise to frequency selectivity is better to transmit the bitstreams in blocks [1]. To be specific, the block size (P) should be much greater than the maximum delay spread (L). The goal is to obtain blocks that are IBI free. By achieving that we can process each block independently as an additive white Gaussian noise channel (AWGN). Here we introduce the “guard chips”, which are used through cyclic prefix in order to provide IBI free blocks. Thus, now we use N symbols to transmit $P - L$ symbols. The system model (see Fig. 1) is:

$$\hat{s}(i) = s(i) F^{H} H F D_{H}^{-1} + n(i) F D_{H}^{-1} \qquad (1)$$

Where F^H corresponds to IFFT matrix while H is the $N \times N$ circulant matrix that depends on the fading coefficients and D_H is a diagonal matrix that is invertible if and only if H is invertible as was proved in [1]. The AWGN is denoted by $n(i)$ and the FFT operation is denoted by F. The block length is symbolized by i, and $s(i)$ is the transmitted symbols in blocks while $\hat{s}(i)$ is the received blocks.

The assumption made here is that the channel transfer function has no zero on the FFT grid, i.e. $H(e^{j2pk/N})$ ¹ 0, " k Î $[0,N-1]$.

It is clear from the model above, that the receiver is very simple and easy to implement since D_H^{-1} is a diagonal matrix.

4. OPTIMAL BIT ALLOCATION

Channel coding is necessary in order to provide reliable visual communication over a wireless channel. Thus, the available bit budget or symbol budget should be shared between source and channel coding. As in [2-4] the aim is to optimally minimize the overall mean square distortion D_{s+c}.

$$\min D_{s+c} \text{ subject to } R_{s+c} = R_{budget} \qquad (2)$$

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 square error distortion which is due to both source coding errors and channel errors. The source coding distortion is due to quantization and it is deterministic while the distortion due to channel errors is stochastic. Thus the overall distortion is stochastic too.

The solution of this problem can be found using Lagrangian optimization and a derivation of this method is in [2-4].

In order to solve the above problem it is necessary to determine the operational rate distortion functions (ORDF) for each scalable layer. It was shown in [3] that the distortion per layer can be expressed as

$$D_{s+c} = \sum_{y=1}^{T} D_{s+c,y}(R_{s+c,1}\ldots, R_{s+c,y}) \qquad (3)$$

where T is the number of scalable layers, $R_{s+c,y}$ is the bit rate used for source and channel coding for scalable layer y and $D_{s+c,y}$ is the distortion of scalable layer y . The ORDF can be found by experimentally obtaining the expected distortion for each layer for all possible combinations of source and channel coding rates and all possible channel conditions. It is obvious though that this becomes very complex. Thus, we choose instead to utilize the universal rate distortion characteristics (URDC) at the expense of a small performance penalty. URDC characteristics show the expected distortion per layer as a function of the bit error rate. Their use is discussed next.

4.1 Universal Rate Distortion Characteristics (URDC) and Channel Characteristics Plots (CCP)

In this section we describe the technique used to obtain the rate distortion characteristics of the individual layers $D_{s+c,y(.)}$, $y = 1,\ldots,T$.

For a given channel SNR, the probability of bit error P_b, is calculated for the set of channel coding rates of interest. P_b establishes a reference as to the performance of channel coding over the particular channel with the given parameters. These plots are called *Channel Characteristic Plots* (CCP) (Fig. 2).

Towards calculating the impact of the errors due to both lousy source coding and channel disturbance on a set of data, it is realized that for a given set of preceding layer source rates the distortion for a particular layer, $D_{s+c,y}$ given a particular source coding rate, $R_{s,y}$, is a function of the bit error rate (BER). Thus, the rate-distortion function of the layer for a fixed source rate, $R_{s,y}$ (given the preceding layer source rates), is a function of the bit error rate (after channel decoding), P_b. It is then possible to plot a family of $D_{s+c,y}$ versus 1/ P_b curves. These plots are called URDC (Fig. 3) and do not depend on channel.

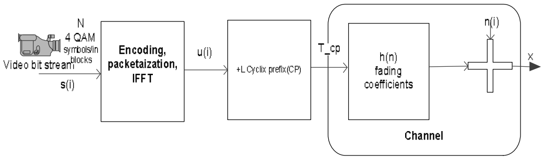
Using the CCP and the URDC operational rate distortion functions for each scalable layer are constructed as follows: (1) For the given channel parameters, the CCP is used to determine the resulting BER for each available channel coding rate. (2) For each of these probability values (P_b) (to be accurate $1/P_b$), the URDC is used to obtain the resulting distortion for each available source coding rate. (3) By obtaining the total rate R_{s+c} for each combination of source and channel codes, the rate-distortion operating points are generated for a given channel condition.

5. EXPERIMENTAL RESULTS

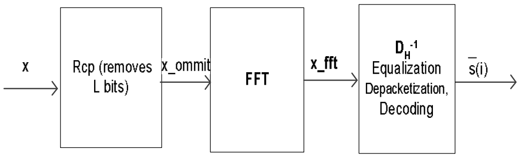
The signal model (Fig. 1) was implemented (interleaver / deinterleaver not shown). "The Foreman" video sequence is encoded using scalable MPEG-4. The bitstream it then QAM-4 modulated, then, inverse fast Fourier transform (IFFT) is performed and the resulting symbols are split into blocks. Cyclic prefix is added to each block and then the bistream is convolved with the channel fading coefficients. AWGN is also added.

At the receiving end we omit the CP symbols that were added from cyclic prefix and after that the Fast Fourier Transform (FFT) operation and equalization are performed. Packets are converted back into bit stream and they are passed to the deinterleaver and decoder.

We used source coding rates of *64kbps*, *96kbps* and *128kbps* and channel coding rates of *1/2* , *2/3* and *4/5* for all combinations for base and enhancement layers. The channel SNR was set to *23dB*. The optimal rate-distortion, convex hull curve, is shown in (Fig. 4). Table I represents the corresponding values of the optimal operational points.



(a) OFDM transmitter using block coding and cyclic prefix over a Raleigh fading Channel.



(b) OFDM receiver.

Fig. 1: OFDM system Model used.

Table I

Optimal rate breakdown of base and enhancement layer using MPEG-4 scalable video over 23-dB Raleigh fading channel.

Source Rate Base Layer In kbps	Channel Rate of the Base Layer	Source Rate Enhan-cement Layer In kbps	Channel Rate of the Enhance-ment Layer	Total Rate kbps	Mean Square Error (MSE)
64	4/5	64	4/5	160	15315
64	2/3	64	4/5	176	779.61
128	2/3	64	4/5	272	285.94
128	1/2	64	2/3	288	283.97
128	2/3	64	1/2	320	49.815

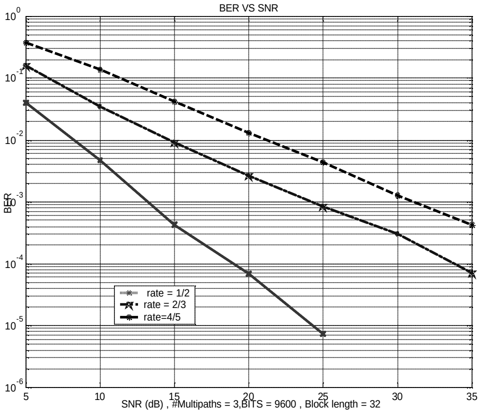


Fig. 2: Channel Characteristics Plots BER versus SNR with block length set to 32.

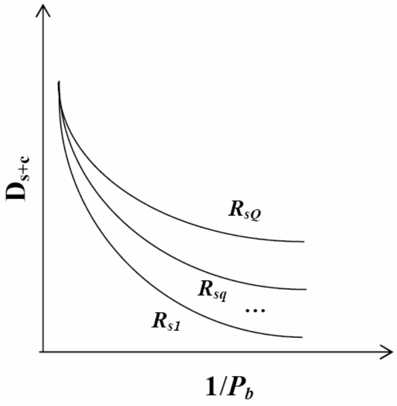


Fig. 3: Universal rate distortion characteristics plots.

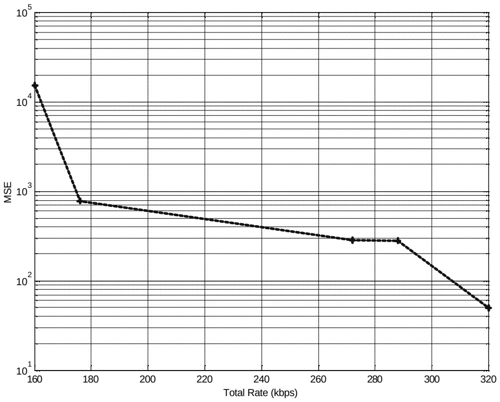


Fig. 4: Optimal rate - distortion performance of scalable MPEG-4 "foreman" video sequence transmission over a single user OFDM wireless channel at 23dB.

6. CONCLUSIONS

The performance of scalable video transmission over an OFDM channel was examined and tested in this work.

An end-to-end scalable video transmission system over a multipath Raleigh fading wireless OFDM single user channel was considered and rate-distortion optimized allocation was performed. The select of source coding rates and channel coding rates for each of the scalable layer was performed. The proposed optimization algorithm guarantees the best possible received video quality for the given channel conditions.

References

[1] Z. Wang and G.B. Giannakis, "Wireless multicarier communications, where Fourier meets Shannon", *IEEE Signal Processing magazine,* vol 17, pp.29-48, May 2000

[2] L.P. Kondi, F. Ishtiaq and A.K. Katsaggelos, "Joint Source-Channel Coding for Motion-Compensated DCT-Based SNR Scalable Video", *IEEE Transactions on Image Processing, Vol. 11, No 9, September 2002.*

[3] L. P. Kondi, D. Srinivasan, D. A. Pados, and S.N. Batalama, "Layered video transmission over multi-rate DS-CDMA wireless systems," in *Proceedings of the SPIE Conference on Image and Video Communications and Processing,* San Jose, CA, 2003,pp. 289–300.

[4] L.P.Kondi, F. Ishtiaq, and A.K. Katsaggelos, "Joint source-channel coding for scalable video," *in Proceedings of the SPIE Conference on Image and Video Communications and Processing,* San Jose, CA, 2003, pp. 324–335.