Performance of practical Wyner-Ziv video codec under Flat-fading Rayleigh Channel

Geno James, Lisimachos P. Kondi, *Member, IEEE*
Department of Electrical Engineering
University at Buffalo
332, Bonner Hall, Buffalo, NY 14260
ggjames,lkondi@eng.buffalo.edu

***Abstract*— In this paper; we analyze the performance of a practical Wyner-Ziv video codec over a Flat-fading Rayleigh Channel. The performance of the same codec is also analyzed for different source coding rates.The performance Wyner-Ziv video coding is compared with pure H.263 Intraframe coding which has similar complexity at the encoder and the superiority of the Wyner-Ziv codec over pure H.263 intra coding is shown. Current video compression standards perform interframe predictive coding to exploit the similarities among successive frames. Since predictive coding makes use of motion estimation, the video encoder is typically five to ten times more complex than the decoder. This asymmetry in complexity is desirable for broadcasting or for streaming video-on-demand systems where video is compressed once and decoded many times. However, some future systems may require the dual scenario. Wyner-Ziv video coding is a step towards achieving this. We present our simulation results that show that Wyner-Ziv video coding performs much better than pure H.263 intracoding under noisy conditions. The dependency of the decoder on the quantization of key (H.263 Intra coded) frames is analyzed. The error resilience of the decoder to missing blocks in the key frames is also shown.**

I. INTRODUCTION

Today's digital video coding architectures have been driven primarily by the *downlink* television broadcast model of a heavy encoder and a multitude of light decoders. However, with the expected proliferation of multiple video sources ranging from digital cameras to low-power video sensor networks to multimedia equipped cellular phones to Webcams, the days of typecasting digital video transmission as a predominantly downlink experience are over. In a typical application scenario, we expect future multimedia systems to use multiple video input and output streams to enhance user experience. These streams need to be captured using a network of distributed devices and transmitted over a bandwidth-constrained, noisy wireless transmission medium, to a central location for processing. For this type of system what we desire is a low-complexity encoder, possibly at the expense of a high-complexity decoder, that nevertheless compresses efficiently.

One such popular system is Wyner-Ziv video codec proposed in [1]. This is an asymmetric video compression scheme where individual frames are encoded independently *(intraframe encoding)* but decoded conditionally *(interframe decoding)*. Two results from information theory suggest that an intraframe encoder - interframe decoder system can come close to the efficiency of an interframe encoder-decoder system. Consider two statistically dependent discrete signals, X and Y, which are compressed using two independent encoders but are decoded by a joint decoder. The Slepian-Wolf Theorem on distributed source coding states that even if the encoders are independent, the achievable rate region for probability of decoding error to approach zero is $R_x \geq H(X|Y), R_y \geq H(Y|X)$ and $R_x + R_y \geq H(X,Y)$ [2]. The counterpart of this theorem for lossy source coding is Wyner and Ziv's work on source coding with side information [3]. Let X and Y be statistically dependent Gaussian random processes, and let Y be known as side information for encoding X. Wyner and Ziv showed that the conditional Rate-Mean Squared Error Distortion function for X is the same whether the side information Y is available only at the decoder, or both at the encoder and the decoder.

In [4] the Wyner-Ziv video coding was applied to the pixel values of a video signal. In [5] this was extended to transform domain. A subset of frames from the video sequence are designated as *key frames* which are compressed using a conventional intraframe codec. The remaining frames, the *Wyner-Ziv frames*, are intraframe encoded using a Wyner-Ziv encoder. To decode a Wyner-Ziv frame, previously decoded frames (both key frames and Wyner-Ziv frames) are used to generate side information. Interframe decoding of the Wyner-Ziv frames is performed by exploiting the inherent similarities between the Wyner-Ziv frame and the side information. A similar video compression system, using distributed source coding principles, was proposed independently by Puri and Ramchandran [6] [7].

This new technology being in its very early stages, needs lot more evaluation and performance testing before it can be practically implemented. This work is one such step towards achieving this goal. Here we have considered one of the widely known codecs, viz., Wyner-Ziv video codec [1] and tested and analysed its performance under a flat fading Rayleigh channel.

In section II, we explain the implemented system and the working of each block. The simulation results and the conclusion are presented in section III.

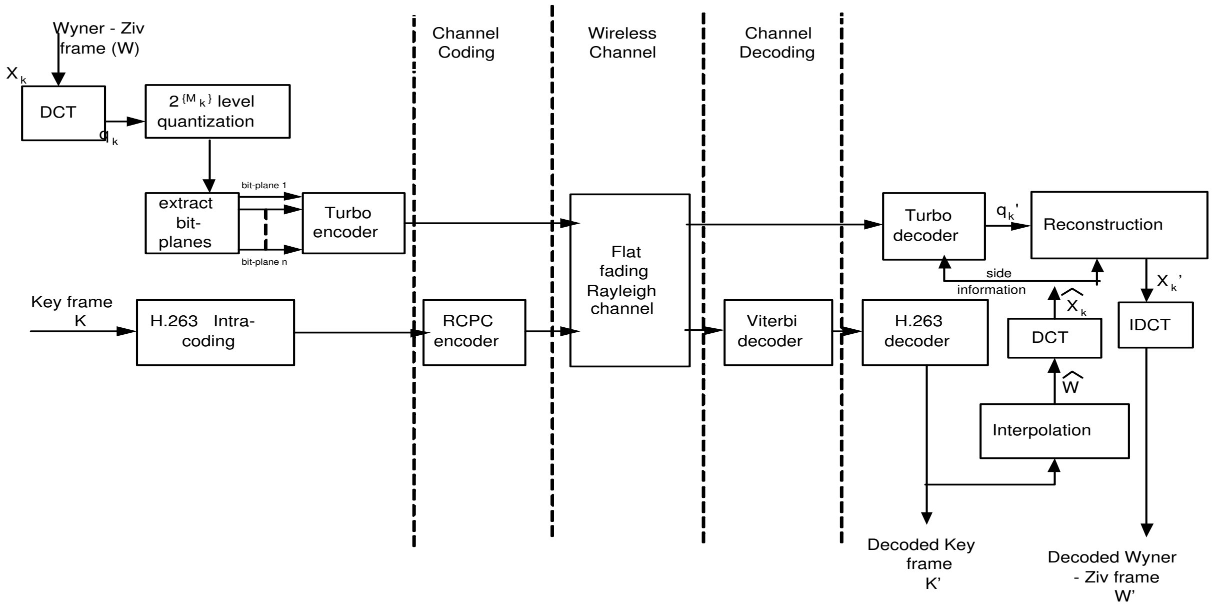


Fig. 1. Block diagram of Wyner-Ziv codec.

II. WYNER-ZIV CODEC

The proposed intraframe encoder and interframe decoder system for video compression is as shown in Fig. 1. A subset of frames from the sequence are designated as key frames and the rest of the frames to be encoded as Wyner-Ziv frames. Here, every even frame is considered a key frame and every odd frame a Wyner-Ziv frame. The key frames, K, are encoded and decoded using a conventional (H.263) [8] intraframe codec. In between the key frames are Wyner-Ziv frames, W, which are intraframe encoded but interframe decoded.

A. Encoding

The encoding is done in two different ways. The Key frames are encoded using a standard H.263 encoder. Since there is no interframe coding involved at the encoder only the intracoding feature of the H.263 encoder is utilized. The bitstream thus generated is protected using RCPC codes. RCPC code with a mother rate of 1/4 was used and they were further punctured to get rates 1/3 and 1/2.

Even though the encoding of Wyner-Ziv frames is called Intraframe coding, it is done is done in a completely different way. As shown in Fig.1, a blockwise(4×4) DCT is applied to the Wyner-Ziv frame W to generate X. The transform coeffcients are grouped together to form coeffcient bands X_k, where k denotes the coeffcient number. Each transform coeffcient band is then quantized and encoded independently. For each band X_k, the coeffcients are quantized using a uniform scalar quantizer with 2^{M_k} levels where $2^{M_k} \in \{0, 2, 4, 8, 16, 32, 64\}$. $2^{M_k} = 0$ means that no bits are sent for coefficient band k and the side information $\hat{X}_k$ is used as the reconstruction $\grave{X}_k$.

The quantized symbols, q_k, are converted to fixed-length binary codewords, and corresponding bit-planes are blocked together forming M_k bit-plane vectors. Each bit-plane vector is then sent to the Slepian-Wolf encoder. The Slepian-Wolf coder is implemented using a rate-compatible punctured turbo code (RCPT). The RCPT, combined with different quantization matrices, provide rate flexibility. After coding, only the parity bits are transmtted and the rest of the bits are discarded.

The turbo encoder is composed of two identical constituent convolutional encoders of rate 1/2 (so, rate of the turbo encoder is 1/3) and generator matrix $\left(1, \frac{1+D+D^3+D^4}{1+D^3+D^4}\right)$ For our simulations we punctured the turbo code to get a final rate of 1/2. Once the bit planes are turbo encoded and the codes were punctred, the information bits are discarded and only the parity bits are transmitted. The Wyner-Ziv coding works as combined source-channel coding. The error protection is built into the source coding itself. So there is no further requirement of protecting this bitstream using any other error correcting techniques. This is where we get a clear advantage over the convolutional Intra frame coding. The H.263 Intra coded frames were protected using RCPC codes of different rates but the Wyner-Ziv coded frames do not need any protection.

B. 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 the channel is referred to as a Rayleigh fading channel in which the distribution of the received signal follows a Rayleigh distribution. 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 channel with additive white Gaussian noise using binary phase shift keying (BPSK) modulation, the received signal $r(t)$ over a signaling period can be represented as

$$r(t) = \alpha(t)\exp^{-j\phi(t)} u(t) + z(t) \qquad 0 \leq t \leq T \tag{1}$$

where is $z(t)$ the complex valued white Gaussian noise, $\alpha(t)$ is the attenuation factor due to fading over the signaling period, and $\phi(t)$ is the phase shift of the received signal. For this signal the attenuation factor $\alpha(t)$ is a Rayleigh distributed random process with the phase shift $\phi(t)$ being uniformly distributed over the interval $(-\pi, \pi)$. For a slow fading Rayleigh channel, we can assume that $\alpha(t)$ and $\phi(t)$ are constant over a signaling interval. In the case of binary phase shift keying modulation (BPSK) over a fading channel, if the received signals phase can be estimated from the signal for coherent demodulation, the received signal can be recovered with the use of a matched filter

C. *Decoding*

At the decoder end, both interframe and intraframe decoding is performed. All the key frames are decoded using the standard h.263 decoder and the Wyner-Ziv frames are interframe decoded. For each W, the decoder takes previously reconstructed adjacent key frames to form the side information, $\hat{W}$, which is an estimate of W. The way in which side information is generated using two adjacent key frames is explained later in the chapter. The decoder applies a blockwise DCT on $\hat{W}$ to generate $\hat{X}$. The transform coefficients from $\hat{X}$ are grouped together to form coefficient bands $\hat{X}_k$, the side information corresponding to X_k. This $\hat{X}_k$ is used along with the parity bits recieved to get $\grave{q}_k$. To be able to use $\hat{X}_k$ reconstruction block, the decoder assumes a Laplacian residual distribution between X_k and $\hat{X}_k$. Let d be the difference between corresponding elements in $\hat{X}$ and $\hat{X}_k$. It is observed that the distribution of d can be approximated as

$$f(d) = \frac{\alpha}{2} e^{-\alpha|d|} \tag{2}$$

The parameter α can be estimated at the decoder using the residual between the key frame K and $\hat{W}$. The reconstructed coefficient band $\hat{X}_k$ is calculated as $E(X_k|\grave{q}, \hat{X}_k)$ Taking 4×4 IDCT of the output of the reconstruction block gives us the decoded Wyner-Ziv frame.

D. *Side Information*

The side information for an even frame at time index t is generated by performing motion-compensated interpolation using the decoded key frames at time $t-1$ and $t+1$. This interpolation technique involves symmetrical bidirectional block matching and motion estimation. The motion vectors thus got are used to get two separate motion compensated estimates of the side informaion. The average of the these two frames serves as the final side information $\hat{W}$ used in the simulation. Since the next key frame is needed for interpolation, the frames have to be decoded out-of-order, similar to the decoding of B frames in predictive video coding.

III. SIMULATION RESULTS

Our work concentrates on the transmission of Wyner-Ziv coded video over wireless cahnnel. Toward this end, we simulate the transmission of Wyner-Ziv coded bitstream over a Rayleigh fading channel with additive white Gaussian noise. The key frames were encoded using standard H.263 (intra) codec using different quantization parameters and their influence over the $psnr$ of reconstructed Wyner-ziv frames were observed and analysed. The performance of the codec was studied under different values of snr for the flat fading Rayleigh channel. For different values of snr, varying amount of protection was given using RCPC codes. Fig. 2 shows the performance improvement of Wyner-Ziv coding acheived over the H.263 (intra) coding and also the $psnr$ gain after decoding the Wyner-Ziv frames using the side information. As explained in section II, the Intra coded frames were used to generate the side information. We see a clear $2dB$ $psnr$ gain from the side information to the decoded frames. 30 frames of Foreman sequence were used for this simulation. In Fig. 3 we see the impact of varying channel snr and Quantization parameter (QP) on the average $psnr$ and the bitrate. It is evident from the plot that at low channel snr Wyner-Ziv coding performes better by reducing the bitrate and increasing the average $psnr$.

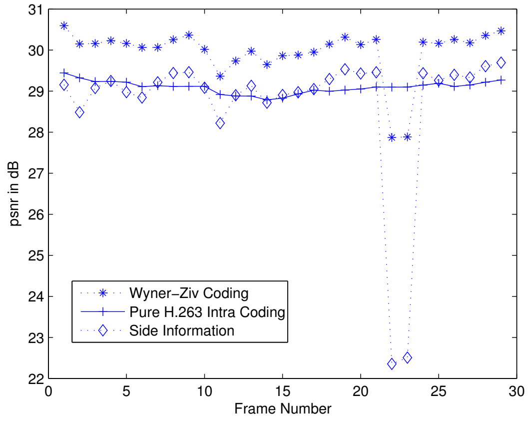


Fig. 2. Plot showing the $psnr$ of the Wyner-Ziv decoded frame, side information and the Intra coded frames.

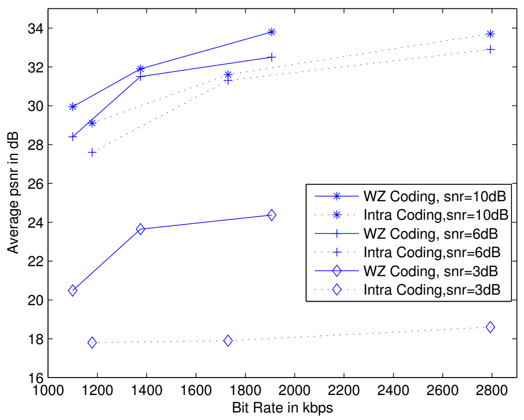
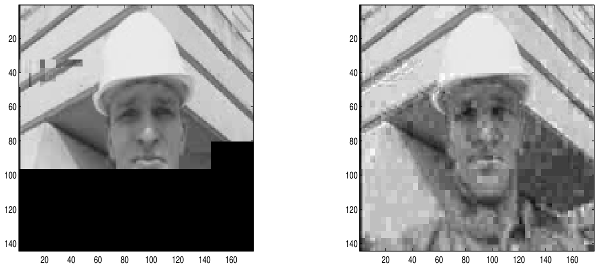


Fig. 3. Rate-distortion performance of the two codecs for different channel snr and fixed channel coding rate of 1/3. The rate is varied by varying the quantization parameter.

At higher channel snr, even though there is not a considerable gain in average $psnr$ the bitrate savings remain the same.



(a) **(b)**

Fig. 4. Predicted frame - Side information (a) and the reconstructed frame (b)

Since the *key frames* used to generate the side information itself have some missing blocks, the predicted frame (side information) Fig. 4(a) also has some missing blocks. In Fig. 4(b), we can see that the reconstructed Wyner-Ziv frame has no missing blocks. This shows the superior performance of Wyner-Ziv coding.

IV. CONCLUSIONS

A Wyner-Ziv video codec was implemented and its performance under flat fading Rayleigh channel was analysed and studied. It was observed that the Wyner-Ziv codec outperforms pure H.263 Intra coding. Even though the complexity of implementing both the codecs are same the rate savings achieved is very considerable. When the QP is high there is a considerable amount of bitrate savings but, the gain in $psnr$ is not very noticeable. For lower QPs the gain in $psnr$ is high but the bitrate savings is not much. We have shown that the Wyner-Ziv Codec performes much better than H.263 Intra coding under a flat fading Rayleigh channel. It was observed that there was gain of upto 2 dB gain in $psnr$ and rate savings of 200 - 1200 kbps.

REFERENCES

[1] B. Girod, A. Aaron, S. Rane, and D. Rebollo-Monedero, “Distributed Video Coding,” *IEEE*, vol. 93, pp. 71–83, January 2005.

[2] D. Slepian and J. K. Wolf, “Noiseless coding of correlated information sources,” *IEEE Trans. Inform. Theory*, vol. IT-19, pp. 471–480, 1973.

[3] A. Wyner and J. Ziv, “The rate-distortion function for source coding with side information at the decoder.,” *IEEE Trans. Inform. Theory*, vol. 22, pp. 1–10, 1976.

[4] A. Aaron, R. Zhang, and B. Girod, “Wyner-Ziv coding of motion video,” in *Asilomer Conference on Signals and Systems*, Nov 2002.

[5] A. Aaron, S.Rane, E. Setton, and B. Girod, “Transform-Domain Wyner-Ziv codec for video,” in *SPIE Visual Communications and Image Processing*, Jan 2004.

[6] R. Puri and K. Ramachandran, “PRISM: A new robust video coding architecture based on distributed compression principles,” in *Allerton Conference on Commnications*, Oct 2002.

[7] R. Puri and K. Ramachandran, “PRISM: An uplink friendly mltimedia coding paradigm,” in *IEEE International Conference on Acoustics, Speech and Signal Processing*, Apr 2003.

[8] G. Cote, B. Erol, M. Gallant, and F. Kossentini, “H.263+: Video coding at low bit rates,” *IEEE Trans on Circuits and Systems on Video Technology*, vol. 8, pp. 849–866, Nov 1998.