

Wyner-Ziv Coding of Still Images with Rate Estimation at the Encoder

David P. Varodayan

Department of Electrical Engineering
Stanford University, Stanford, CA 94305
varodayan@stanford.edu

Abstract— We develop a Wyner-Ziv still image codec with decoder feedback, based on a Rate Compatible Punctured Turbo (RCPT) code. Within this framework, two methods of pairing source data with side information are compared and found to be similar in performance. The architecture is modified into a Wyner-Ziv image codec with rate estimation at the encoder. The proposed rate estimator is the mean of the turbo decoder’s correct channel probability estimates of systematic bits. The codec in conjunction with this rate estimator suffers only a rate of loss 0.1 bits per pixel for certain images. However, for images with a prevalence of distinct edges the rate-estimated Wyner-Ziv image codec does not achieve the performance of the Wyner-Ziv image codec with decoder feedback.

I. INTRODUCTION

Source encoders exploit statistical redundancy to compress data to a more concise form. In conventional schemes, encoding is a very computationally expensive task while decoding is relatively simple. This paradigm is well-suited to traditional broadcast and storage scenarios, in which content creation is centralized. However, shifting the computational burden from the encoder to the decoder would greatly benefit many modern distributed applications, such as sensor networks and ‘disposable’ digital cameras.

The theoretical foundations for distributed source coding are the Slepian-Wolf [1] and Wyner-Ziv [2] information theoretic results. Consider a source X correlated with side information Y . Slepian and Wolf proved that X can be losslessly coded at the conditional entropy of X given Y regardless of whether Y is known at both encoder and decoder or at the decoder only. Wyner and Ziv developed the rate-distortion counterpart: with X and Y Gaussian and mean-squared error distortion, the conditional rate-distortion function for X given Y is again the same regardless of whether side information Y is known at both encoder and decoder or at the decoder only. These theorems motivate the design of efficient Wyner-Ziv codecs with simplified encoders that ignore source redundancy coupled with complex decoders that do take it into account.

Recent progress in developing practical Wyner-Ziv coding technology has been in the motion video application domain [3] [4]. These schemes present effective methods for exploiting temporal redundancy in video,

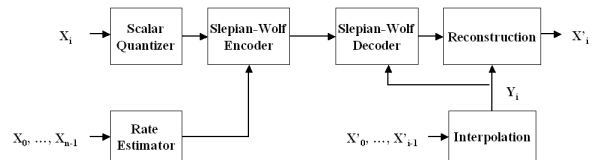


Fig. 1. Wyner-Ziv Image Codec with Decoder Feedback

but neither addresses the spatial Wyner-Ziv coding of still image frames. Another shortcoming is that they do not suggest simplified rate estimation strategies that can be applied at the encoder without feedback from the decoder. In this paper we consider both of these questions.

Section II describes the adaptation of the Wyner-Ziv video codec with decoder feedback of [3] to the case of still images and suggests two different pairings of data with side information. In Section III, the Wyner-Ziv image codec is modified to enable rate estimation and a rate estimator is proposed. Section IV compares the two Wyner-Ziv image codecs with decoder feedback of Section II, and evaluates the effectiveness of the Wyner-Ziv codec with the proposed rate estimation scheme. Finally, Section V concludes with suggestions for future work.

II. WYNER-ZIV STILL IMAGE CODEC

The Wyner-Ziv still image codec with decoder feedback depicted in Fig. 1 follows the architecture of the one in [3] closely. Both make use of an inner turbo-code based Slepian-Wolf codec and an outer quantization-reconstruction pair. The key difference is in the pairing of data with decoder side information. According to the Wyner-Ziv result, it is the correlation between the data and the side information that determines the amount of compression achievable. In the video codec the processing is frame-by-frame, whereas in our image codec it is band-by-band. The consequence is that the video and image codecs exploit temporal and spatial redundancy, respectively, to achieve compression.

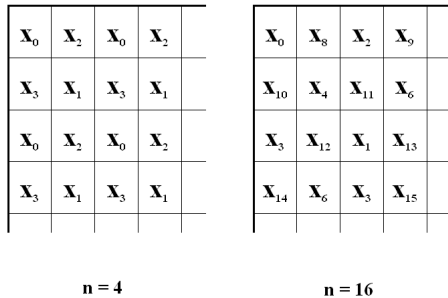


Fig. 2. Subsampling templates for Pairing data and side information using $n = 4$ and $n = 16$ bands

A. Pairing Data with Side Information

While the Wyner-Ziv video codec partitions a raw video sequence temporally, the image codec partitions a raw image spatially into n subsampled bands, X_0, X_1, \dots, X_{n-1} , each of which contains $\frac{1}{n}$ of the image's pixels. In this paper we consider two different pairing strategies, one with $n = 4$ bands and the other with $n = 16$ bands. The subsampling templates for these two schemes are shown in Fig. 2.

In both cases, the first band X_0 is assumed to be entropy coded and sent to the decoder to produce the initial side information. The remaining bands, X_1, \dots, X_{n-1} , are Wyner-Ziv coded one-by-one according to the architecture in Fig. 1. As the pixels in each band are decoded, the reconstructions X'_1, \dots, X'_{n-1} contribute to the side information for subsequent bands.

Thus band X_i is paired with side information which is a function of previously reconstructed bands, $Y_i = f(X'_0, X'_1, \dots, X'_{i-1})$. The side information construction function considered in this paper is bilinear interpolation. The estimate of each pixel of Y_i is calculated from the values of the four nearest pixels of $X'_0 \cup X'_1 \cup \dots \cup X'_{i-1}$. For natural images, we model the resulting interpolation error as a Laplacian distribution.

B. Components in Common with the Video Codec

Both the still image and video codecs with decoder feedback share similar quantization, turbo-code based Slepian-Wolf coding and reconstruction components.

The pixels of an image band X_i are first quantized using a uniform scalar quantizer. The quantization index bits are then grouped by bit plane with the most significant first.

The Slepian-Wolf encoder and decoder are built from a Rate Compatible Punctured Turbo (RCPT) code encoder and decoder [5], respectively. The turbo encoder computes the parity sequence corresponding to each bit

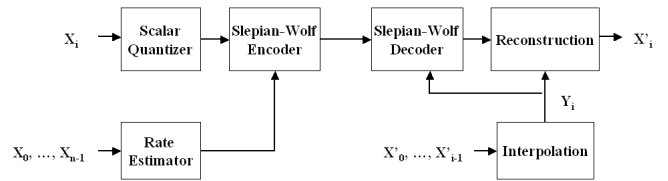


Fig. 3. Wyner-Ziv Image Codec with Rate Estimation at the Encoder

plane of X_i and transmits more and more of this sequence as requested by the decoder via the feedback channel. The turbo decoder initializes probability estimates for the systematic bits based on the interpolation Y_i and the distribution of the interpolation error. Then, using the encoder's parity sequence, the decoder attempts to resolve the probability estimates to their actual quantization index bit plane values. The decoder requests further parity bits until decoding is successful. The decoder also refines its estimates of less significant bit planes with its knowledge of more significant bit planes. Note that in this paper, we assume that the decoder knows the variance of the Laplacian interpolation error. This assumption is reasonable since the turbo decoder can iteratively estimate the variance in a similar way to that of [6] and it is robust to small discrepancies in this value [3] anyway.

Finally, the reconstruction function maps each decoded quantization index to the value in the quantization interval that is nearest the interpolated value. A minor difference between the Wyner-Ziv image codec in this paper and the video codec is that the reconstruction for the image codec does not employ dithering.

III. RATE ESTIMATION AT THE ENCODER

Fig. 3 depicts the Wyner-Ziv image codec modified to enable rate estimation at the encoder. The rate estimator must examine the statistics of the band X_i and estimate what proportion of the parity sequence is required by the decoder to resolve each bit plane of the interpolation Y_i to the correct bit plane of X_i . In this way, the encoder determines the transmission rate for the bit plane in question.

A. Pros and Cons of Rate Estimation at the Encoder

This modification eliminates the need for a feedback channel and enables the Wyner-Ziv image codec to be used in more diverse scenarios. Another important improvement is that transmission rate for a single bit plane of X_i never exceeds 1. If the estimated required rate is greater than 1, the encoder transmits the bits

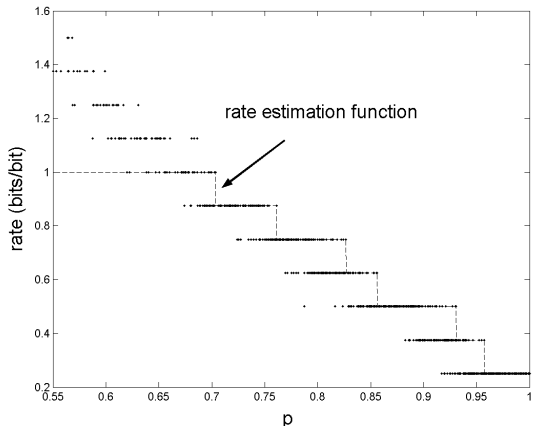


Fig. 4. Required Rate (bits/bit) vs. Parameter p for $n = 4$ bands

uncoded instead. A final benefit of rate estimation is that it reduces the number of computations required at the decoder.

The chief drawback of even simple rate estimation is that it still increases the complexity of the encoder. Furthermore, the rate-distortion performance of the rate-estimated codec is usually inferior to the codec of the previous section unless the estimation is perfect and the minimal rate is always selected.

B. Proposed Rate Estimator

In the Wyner-Ziv codec with decoder feedback, the number of parity bits eventually requested by the decoder to resolve a bit plane of Y_i is directly influenced by the turbo decoder's initial probability estimates for that bit plane. Hence, the proposed rate estimator is the mean of the turbo decoder's correct channel probability estimates of the systematic bits. In other words, it is the expectation of the probability that the bits of the corresponding bit planes of X_i and Y_i match; that is,

$$p = E[Pr\{Q_b(Y_i) = Q_b(X_i)\}], \quad (1)$$

where $Q_b()$ extracts the values of the bits in bit plane b of its argument.

The scatter plot in Fig. 4 shows the decoder's required rates versus p using the Wyner-Ziv codec with decoder feedback for the data and side information pairing with $n = 4$ bands. This figure indicates that the strong correlation between these parameters can be exploited for rate estimation. In this paper, we use the dashed line in Fig. 4 as a rate estimation function defined on p . The rate estimation function is conservative in order to minimize the occurrence of turbo decoder failure.

Notice, however, that this calculation requires the interpolation Y_i . Since $Y_i = f(X'_0, X'_1, \dots, X'_{i-1})$

is unavailable at the encoder, we use $Z_i = f(X_0, X_1, \dots, X_{i-1})$ instead. Moreover, we estimate p using subsampled versions of Z_i to reflect various levels of encoder complexity. Both of these imperfections contribute to errors in the rate estimation process.

Strictly speaking, this rate estimated codec is not a Wyner-Ziv scheme because the encoder makes use of the bands X_0, \dots, X_{i-1} in order to encode X_i . However, this is acceptable since the amount of computation needed for this is insignificant compared to conventional encoding.

IV. SIMULATION RESULTS

The results in this section compare the two data and side information pairing strategies proposed for the Wyner-Ziv codec with decoder feedback in Section II, and evaluate the Wyner-Ziv codec with rate estimation at the encoder suggested in Section III.

A. Comparison of Pairing Strategies

Figs. 5 and 6 show rate-distortion plots for the *bridge* and *harbour* sample images, respectively, using pairing with $n = 4$ bands (solid curves) and $n = 16$ bands (dashed curves) for the Wyner-Ziv image codec with decoder feedback. For the *bridge* image, both strategies perform almost equally well, while the pairing with $n = 16$ bands appears to outperform $n = 4$ bands by 0.4 bits/pixel over a range of PSNR for the *harbour* image. Recall, however, that we assume that the first band X_0 in each of these schemes is entropy coded in these calculations. If X_0 is coded more aggressively, the relative positions of the rate-distortion curves may change. Hence, the chief observation is that for the Wyner-Ziv image codec with decoder feedback, pairing data and side information with $n = 4$ bands permits compression at higher PSNR and correspondingly higher rates, while $n = 16$ bands allows for compression at lower PSNR and lower rates.

B. Evaluation of the Rate Estimation Scheme

The performance of the Wyner-Ziv image codec with rate estimation of different complexities is evaluated against the Wyner-Ziv image codec with decoder feedback. Various levels of encoder complexity are considered by implementing the rate estimation over all, $\frac{1}{16}$, $\frac{1}{64}$ and $\frac{1}{256}$ of the image pixels.

The resulting rate-distortion curves for the *bridge* image with pairing using $n = 4$ bands are depicted in Fig. 7. As expected, none of the Wyner-Ziv codecs with rate estimation (solid lines) perform as well as the Wyner-Ziv codec with feedback (dashed line). This is because the conservative rate estimation function does not always select the minimal allowable rate. Furthermore, observe that performance degrades as the parameter p is estimated at the encoder using fewer and fewer pixels.

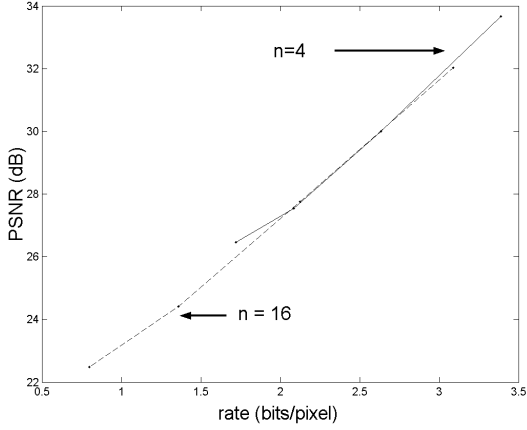


Fig. 5. Rate-Distortion for Wyner-Ziv codec with decoder feedback on *bridge* image

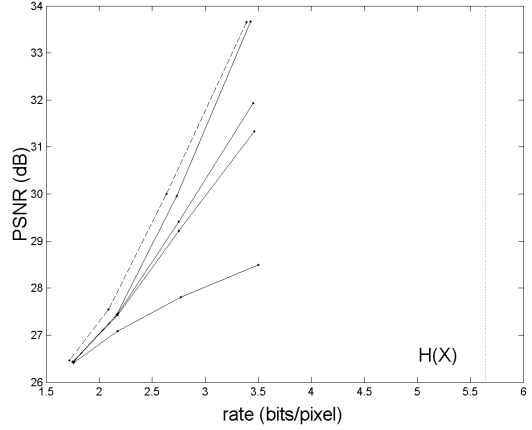


Fig. 7. Rate-Distortion for Wyner-Ziv codec with rate estimation on *bridge* image using $n = 4$ bands

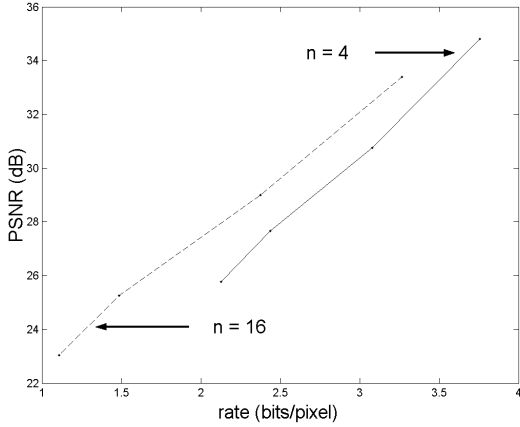


Fig. 6. Rate-Distortion for Wyner-Ziv codec with decoder feedback on *harbour* image

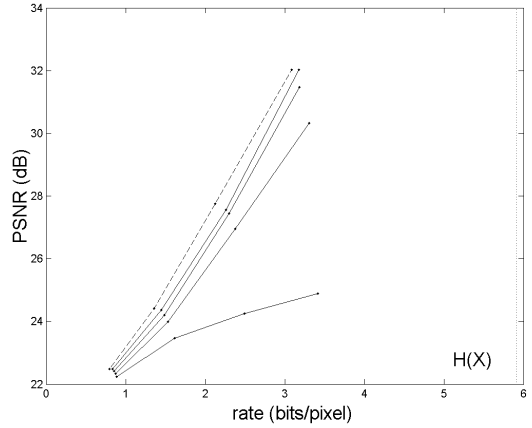


Fig. 8. Rate-Distortion for Wyner-Ziv codec with rate estimation on *bridge* image using $n = 16$ bands

Nevertheless, Wyner-Ziv image coding with rate estimation at the encoder using all the pixels suffers only 0.1 bits/pixel of rate loss compared to the codec with decoder feedback, for all values of PSNR. Interestingly, using either $\frac{1}{16}$ or $\frac{1}{64}$ of the pixels for rate estimation gives very similar results. However, rate estimation using only $\frac{1}{256}$ of the image pixels is very poor. For rate comparison, we have indicated the pixelwise entropy of the image, calculated as

$$H(X) = \frac{1}{n} \sum_{i=0}^{n-1} H(X_i | X_0, \dots, X_{i-1}). \quad (2)$$

For this image with pairing using $n = 4$ bands, all the Wyner-Ziv rates are superior to the pixelwise entropy of 5.6 bits/pixel.

The $n = 16$ band strategy applied to the same *bridge* image yields very similar results, shown in Fig. 8. In

this case, Wyner-Ziv image coding with rate estimation using all the pixels or $\frac{1}{16}$ of them follows the performance of the Wyner-Ziv image codec with decoder feedback closely, with rate losses of only 0.1 and 0.15 bits/pixel, respectively. Once again the performance when $\frac{1}{64}$ of the pixels are used is reasonable, but when $\frac{1}{256}$ of them are used it is very poor. Note also that pixelwise entropy is much larger than any of the rates observed.

The rate-distortion performance of the rate estimation schemes applied to the *harbour* image with pairing using $n = 4$ and $n = 16$ bands is shown in Figs. 9 and 10, respectively. Unfortunately, none of the Wyner-Ziv codecs with rate estimation track the performance of the Wyner-Ziv codec with decoder feedback. All of them incur heavy distortion losses. This poor rate estimation occurs because of the prevalence of distinct

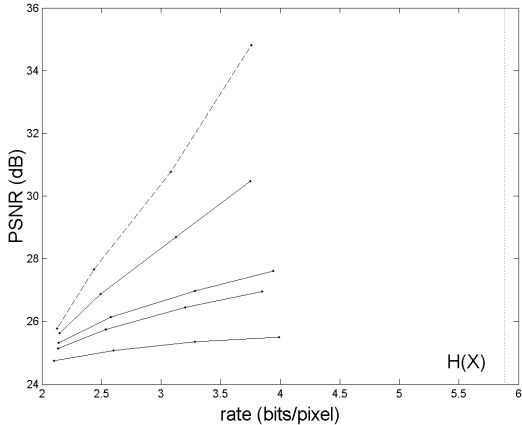


Fig. 9. Rate-Distortion for Wyner-Ziv codec with rate estimation on *harbour* image using $n = 4$ bands

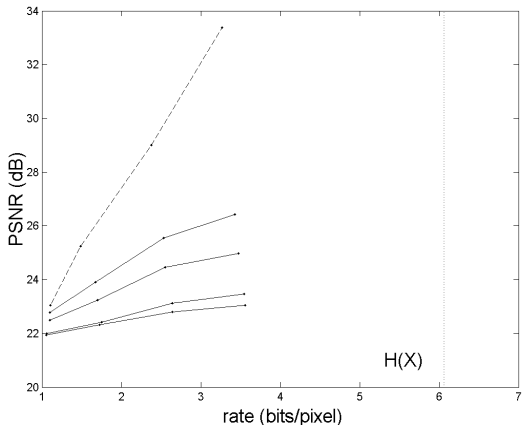


Fig. 10. Rate-Distortion for Wyner-Ziv codec with rate estimation on *harbour* image using $n = 16$ bands

edges in the image. Bilinear interpolation at the decoder fails to reconstruct edges well and consequently the rate estimation at the encoder does not take them into account either.

V. CONCLUSIONS AND FUTURE SUGGESTIONS

In this paper, we have adapted the Wyner-Ziv video codec with decoder feedback from [3] for spatial Wyner-Ziv compression of still images. Two schemes to pair data with side information were compared, one using $n = 4$ bands and the other using $n = 16$ bands. The performance of these pairing strategies was found to be comparable with $n = 4$ bands supporting higher PSNR at higher rates and $n = 16$ bands more suited to lower PSNR at lower rates.

The Wyner-Ziv image codec was then modified to enable rate estimation at the encoder. A parameter p ,

the mean of the turbo decoder's correct channel probability estimates of the systematic bits, was proposed as a rate estimator. The required rate was shown to be highly correlated with p , and Wyner-Ziv image coding with rate estimation was successful for certain images. However, performance was poor when the image contained many distinct edges.

The Wyner-Ziv image codecs proposed in this paper can be improved in several ways. One improvement would be to implement a more sophisticated interpolation method than bilinear interpolation at the decoder. This would improve the compression efficiency of the Wyner-Ziv codec with decoder feedback. If the interpolation includes edge detection, the performance of the Wyner-Ziv codec with rate estimation would likely improve in the case of images with plenty of edges. A better choice of rate estimator is also possible. Investigating the convergence behavior of turbo decoder via Extrinsic Information Transfer (EXIT) charts [7] would probably inspire a parameter better correlated with the required rate.

The insights about spatial Wyner-Ziv coding and rate estimation at the encoder gained from this study can also contribute to the development of a more sophisticated Wyner-Ziv video codec.

ACKNOWLEDGEMENTS

The author would like to thank Anne Aaron for technical assistance during this project as well as Professor Girod and David Rebollo-Monedero for valuable discussions.

REFERENCES

- [1] J. D. Slepian and J. K. Wolf, "Noiseless coding of correlated information sources," *IEEE Trans. on Inform. Theory*, vol. IT-19, pp. 471-480, July 1973.
- [2] A. D. Wyner and J. Ziv, "The rate-distortion function for source coding with side information at the decoder," *IEEE Trans. on Inform. Theory*, vol. IT-22, no. 1, pp. 1-10, Jan. 1976.
- [3] A. Aaron, R. Zhang, and B. Girod, "Wyner-Ziv coding of motion video," in *Proc. Asilomar Conference on Signals and Systems*, Pacific Grove, CA, Nov. 2002.
- [4] R. Puri and K. Ramchandran, "PRISM: A new robust video coding architecture based on distributed compression principles," in *Proc. Allerton Conference on Communication, Control, and Computing*, Allerton, IL, Oct. 2002.
- [5] D. Rowitch and L. Milstein, "On the performance of hybrid FEC/ARQ systems using rate compatible punctured turbo codes," *IEEE Trans. on Commun.*, vol. 48, no. 6, pp. 948-959, June 2000.
- [6] J. Garcia-Frias and Y. Zhao, "Compression of correlated binary sources using turbo codes," *IEEE Commun. Lett.*, vol. 5 no. 10, pp. 417-419, Oct. 2001.
- [7] S. ten Brink, "Convergence behavior of iterative decoded parallel concatenated codes," *IEEE Trans. on Commun.*, vol. 49, no. 10, pp. 1727-1737, Oct 2001.