Wyner-Ziv Coding of Multispectral Images for Space and Airborne Platforms

Rane, S.; Wang, Y.; Boufounos, P.; Vetro, A.

Abstract

This paper investigates the application of lossy distributed source coding to high resolution multispectral images. The choice of distributed source coding is motivated by the need for very low encoding complexity on space and airborne platforms. The data consists of red, blue, green and infra-red channels and is compressed in an asymmetric Wyner-Ziv setting. One image channel is compressed using traditional JPEG and transmitted to the ground station where it is available as side information for Wyner-Ziv coding of the other channels. Encoding is accomplished by quantizing the image data, applying a Low-Density Parity Check code to the remaining three image channels, and transmitting the resulting syndromes. At the ground station, the image data is recovered from the syndromes by exploiting the correlation in the frequency spectrum of the band being decoded and the JPEG-decoded side information band. In experiments with real uncompressed images obtained by a satellite, the rate-distortion performance is found to be vastly superior to JPEG compression of individual image channels and rivals that of JPEG2000 at much lower encoding complexity.

Picture Coding Symposium (PCS)

WYNER-ZIV CODING OF MULTISPECTRAL IMAGES
FOR SPACE AND AIRBORNE PLATFORMS

Shantanu Rane, Yige Wang, Petros Boufounos and Anthony Vetro

Mitsubishi Electric Research Laboratories, Cambridge, Massachusetts, U.S.A.

ABSTRACT

This paper investigates the application of lossy distributed source coding to high resolution multispectral images. The choice of distributed source coding is motivated by the need for very low encoding complexity on space and airborne platforms. The data consists of red, blue, green and infra-red channels and is compressed in an asymmetric Wyner-Ziv setting. One image channel is compressed using traditional JPEG and transmitted to the ground station where it is available as side information for Wyner-Ziv coding of the other channels. Encoding is accomplished by quantizing the image data, applying a Low-Density Parity Check code to the remaining three image channels, and transmitting the resulting syndromes. At the ground station, the image data is recovered from the syndromes by exploiting the correlation in the frequency spectrum of the band being decoded and the JPEG-decoded side information band. In experiments with real uncompressed images obtained by a satellite, the rate-distortion performance is found to be vastly superior to JPEG compression of individual image channels and rivals that of JPEG2000 at much lower encoding complexity.

Index Terms— multispectral, LDPC code, Wyner-Ziv coding

1. INTRODUCTION

Multispectral images are important tools for remote sensing of the earth's surface. They provide accurate information about surface features, cloud and weather patterns and topographic changes in natural disasters. The image data consists of multiple bands, each corresponding to a specific frequency range. Typically, the image bands correspond to information in the visible and infra-red region of the spectrum. In a high-resolution multispectral imaging system, a large volume of raw image data is acquired onboard and must be compressed before it is transmitted to the ground station. The onboard compression module must use highly reliable hardware that consumes low power during operation.

To satisfy the low-power requirement, the compression scheme must be implemented with low complexity: It is extremely costly to estimate the dependencies among the multiple image bands and then to apply predictive coding onboard, even though this would improve the compression performance. The state-of-the-art solution is to independently encode each image band using the JPEG2000 compression algorithm [1]. Our aim is to investigate whether compression efficiency can be improved by employing very low complexity encoding and exploiting inter-band dependencies at the decoder.

Distributed source coding schemes are based on two seminal information theoretic results on correlated sources that are encoded independently but decoded jointly: Slepian and Wolf proved that such a scheme achieves the same asymptotic lossless compression performance as joint encoding of the sources [2]. Wyner and Ziv showed that, if the sources are jointly Gaussian, then such a scheme

has the same rate-distortion penalty as joint encoding [3]. Even if the sources are not jointly Gaussian, the rate-distortion penalty with respect to joint encoding is bounded. Thus, the rationale behind distributed coding of multi-band image data is as follows: Rather than employing complex inter-band predictive encoding, it is simpler and almost as efficient to apply a simple encoding scheme to individual bands and exploit inter-band dependencies only at the decoder.

For the closely related problem of compressing hyperspectral imagery, Magli *et al.* have proposed a Slepian-Wolf coding scheme that outperforms independent lossless coding of the image bands [4]. For the lossy case, Cheung *et al.* have demonstrated that Wyner-Ziv coding in the wavelet transform domain gives significant compression gains over independent wavelet-based compression of the image bands [5]. An excellent survey of existing work in distributed hyperspectral image compression appears in [6]. Compared to the 200 or more bands available in hyperspectral imagery, there are typically only four or five image bands covering the same frequency range in multispectral images. Therefore, exploiting inter-band correlations for multispectral data compression is a more daunting task. This is the problem addressed in the present work.

The remainder of this paper is organized as follows. Section 2 describes the overall architecture of the Wyner-Ziv codec applied to multispectral image compression. In Section 3, we present a study of the statistical properties of the multispectral data and the measured correlation among bands. An experimental evaluation of the rate-distortion performance of the codec is presented in Section 4.

2. WYNER-ZIV CODING SCHEME

A block diagram of the Wyner-Ziv codec appears in Fig. 1. This is a transform-domain Wyner-Ziv coding scheme based on a blockwise Discrete Cosine Transform (DCT) and Low Density Parity Check (LDPC) coding. Let the 2D-DCT be applied to $N \times N$ pixel blocks, each pixel having bit-depth B. Let the length of the LDPC codeword be L.

2.1. Side Information Coding

The proposed scheme performs conventional JPEG coding of one image band at the compression module and transmits this bit stream to the ground station. This is similar to the conventional intra coding of key frames in distributed video coding [7]. At the ground station, the JPEG run-length entropy coding is reversed and the quantized DCT coefficients are available as side information for distributed decoding of the remaining image bands. In our studies, it was observed that JPEG coding was most efficient for the blue band, followed by green, red and infra-red. This observation is consistent across all datasets and is most probably a characteristic of the acquisition system. In order to maximize the overall compression ratio for all image bands, it was decided to code the blue band using JPEG coding, and
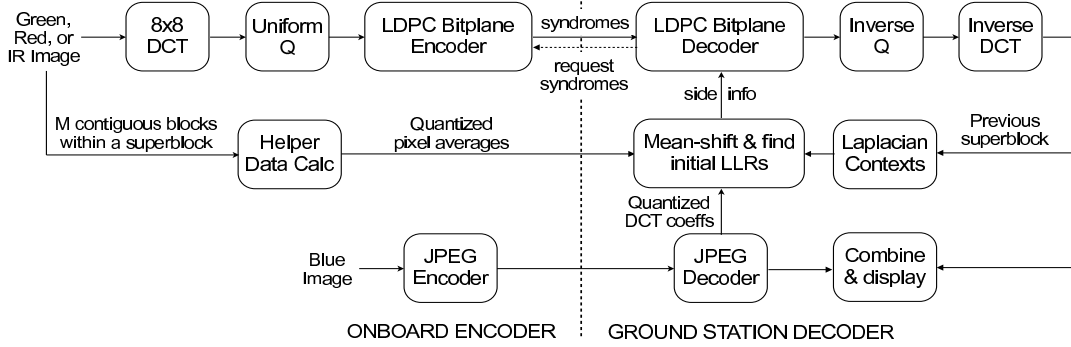
Fig. 1. Low complexity on-board encoder performs JPEG coding of one image band and LDPC syndrome coding of all other bands. At the ground station, the syndromes are decoded using the JPEG-decoded band and locally determined Laplacian context labels as side information. Wyner-Ziv encoding of the different bands can be performed in parallel. A practical final design is expected to use a rate control algorithm to determine the required code rate rather than using a feedback channel to request syndromes.

the remaining bands using Wyner-Ziv coding. We use a uniform quantizer with a constant step size for all DCT coefficients.

2.2. Wyner-Ziv Encoding

First, a blockwise 2D-DCT is applied to the image band. Since the proposed compression module already contains hardware for JPEG compression of the side information band, JPEG's 8×8 DCT is re-used for Wyner-Ziv coding ($N = 8$). The DCT coefficients are quantized by a uniform quantizer with a constant step size[1]. After quantization, the N^2 quantized DCT coefficients are packaged into BN^2 side information bitplanes. The length of each bitplane is L. Then, a channel code with codeword length L is applied separately to each bitplane. This operation involves simple XOR operations and results in BN^2 syndrome vectors, one for each encoded bitplane. The length of the syndrome vector depends on the rate of the channel code. e.g., a high-rate code would be employed for higher significant bitplanes of low-frequency DCT coefficients, resulting in small syndrome vectors. LDPC Accumulate (LDPCA) codes, proposed by Varodayan et al. [8], are used so that new syndromes can be incrementally generated until decoding succeeds at the ground station. In this way, the code rate adapts to the difficulty of decoding a specific bitplane. The request to transmit more syndromes is received via a feedback channel from the ground station. In order to reduce the number of usages of the feedback channel, each feedback message prompts the transmission of a blocks of 40 syndromes until decoding is successful. The syndrome vectors constitute the compressed bit stream transmitted to the ground station.

2.3. Helper Information

The image is divided into several superblocks, each containing L consecutive $N \times N$ DCT transform blocks. In general, there is a mismatch between the values of the DCT coefficients of the side information band and the DCT coefficients of the band to be decoded. Side information decoding aims to correct this mismatch using as few syndromes as possible and as little extra helper information as possible. A compromise must be made between the helper data overhead and accuracy of the side information available for syndrome decoding. The helper data is generated as follows: Each superblock is divided into $M \ll L$ contiguous $N \times N$ pixel blocks, each containing $\frac{LN^2}{M}$ pixels. The average pixel value in each of these M

---

[1]There was no observable improvement in rate distortion performance when JPEG's traditional non-uniform quantization matrix was used.

blocks is transmitted to the ground station as helper information. In our implementation, $N = 8$, $L = 8000$, $M = 125$ and the average pixel values are computed over $\frac{8000}{125} = 64$ contiguous 8×8 blocks, i.e., over 4096 pixels. The average values can be computed very easily from the DC coefficients of the 8×8 DCT. Entropy coding may be applied to reduce the bit rate of this helper information, but this is not done in our implementation. Since $M \ll L$, the extra bandwidth required for transmitting the integer means is already negligible.

2.4. Wyner-Ziv Decoding

Given a syndrome vector, the decoder recovers every bitplane of every DCT coefficient using LDPCA decoding via the belief propagation (BP) algorithm. To initialize BP decoding, the check nodes of the code graph are populated with the received syndrome bits.

Next, the side information is mean-shifted as follows: The average pixel value is computed over $\frac{L}{M}$ contiguous $N \times N$ blocks of the side information and this value is subtracted from the side information. To the result, we add the average pixel values of the corresponding blocks received in the helper data stream. Now, the difference Δ between the DCT coefficients to be decoded and the DCT coefficient of the side information band is modeled as a zero-mean Laplacian random variable. Thus,

$$f(\Delta) = \frac{1}{2\lambda} e^{-\frac{|\Delta|}{\lambda}}, \ \lambda > 0$$

The Laplacian parameter λ is estimated from the previously decoded superblock in the image band. The smaller the value of λ for a DCT coefficient, the easier it is to perform interband prediction for that coefficient, and thus the more reliable the side information. It might initially seem that a given DCT coefficient in position (i, j), $1 \leq i, j \leq N$ should have a constant $\lambda(i, j)$ no matter where in the superblock the transform blocks are located. However, our observation is that, the reliability with which a DCT coefficient at position (i, j) can be predicted does depend on the region being imaged. In other words, $\lambda(i, j)$ varies depending upon the location of the $N \times N$ transform block within the larger superblock. For example, DCT coefficients in regions containing sea or clouds are more reliably predictable across all the bands, while those in city regions are more difficult to predict. To capture this, we classify all the $N \times N$ DCT blocks in a superblock into T groups or contexts according to their reliability. There are related methods in distributed video coding, in which Laplacian parameters are estimated at the block-level, frame-level and video sequence-level [9].

Quantity	Value
Image Size	7100×8000 pixels per band
Bit Depth	8 bits/pixel
Channels	Blue: 0.42-0.50 μm, Green: 0.52-0.60 μm
(Wavelengths)	Red: 0.61-0.69 μm, IR: 0.76-0.89 μm
Spatial Resolution	10m at Nadir
Swath Width	70km at Nadir

Table 1. Specifications of multispectral data used in experiments.

To determine the context of a $N \times N$ DCT transform block, we count the number of insignificant DCT coefficients in that block, i.e., those with absolute value less than a threshold δ. Our observation is that the larger the number of insignificant coefficients, the more reliable the side information. So, context-dependent Laplacian parameters $\lambda(i, j, t), t \in \{1, 2, ..., T\}$ are determined from the previously decoded superblock to be used while decoding the current superblock. These parameters are used to initialize the log-likelihood ratios (LLRs) of each variable node in the LDPC code graph.

Denote the estimate of the bits of a DCT coefficient by $[\hat{b}_1, \hat{b}_2, \cdots, \hat{b}_B]$, where $\hat{b}_B$ is the most significant bit. Contrary to most previous work on distributed video coding, we decode from the least significant bit to the most significant bit, i.e., from $\hat{b}_1$ to $\hat{b}_B$. The rationale is that, when lower significant bitplanes are decoded, the higher bitplanes are easier to decode because candidate values are a greater distance apart — a fact exploited by Ungerboeck for decoding channel codes [10]. While decoding the ith bitplane, the initial LLR of the jth variable node is calculated as follows:

$$R_{ij} = \ln \frac{\Pr[\hat{b}_i = 0 | W_{ij}, \hat{b}_1, \cdots, \hat{b}_{i-1}]}{\Pr[\hat{b}_i = 1 | W_{ij}, \hat{b}_1, \cdots, \hat{b}_{i-1}]}, \ 1 \le j \le L$$

where W_{ij} is the value of the DCT coefficient of the side information that corresponds to the jth variable node. The conditional probability for the ith bit is thus computed from the Laplacian context and from the values of the previously decoded bits.

3. STATISTICS OF MULTISPECTRAL DATA

3.1. Data Source

Although the principles are valid for a variety of data sources, the experiments in this work are carried out on real multipectral data acquired by the Advanced Land Observation Satellite (ALOS) using the Advanced Visible and Near Infra-red Radiometer (AVNIR-2). While compressed data is transmitted during normal operation, a special mode is used to transmit uncompressed raw data which was used in our compression experiments. More details about ALOS-AVNIR-2 and directions to acquire datasets are available online[2]. The specifications of the images used in this work are given in Table 1. Example images of the data appear in Fig. 2. The data was intentionally chosen to contain various topographical features such as vegetation, ocean, coastline and city structures as well as weather conditions such as atmospheric haze and clouds.

3.2. Inter-band Correlation

An initial study of multispectral images demonstrated correlation among bands, which can be exploited in inter-band prediction. Table 2 reports the correlation coefficients among the pixels of the four image bands in data sets 1 and 2. The study demonstrated significant correlation among adjacent spectral bands (e.g. Blue and Green, or Green and Red) and reduced correlation as the bands separate (e.g.

Data 1   Data 2
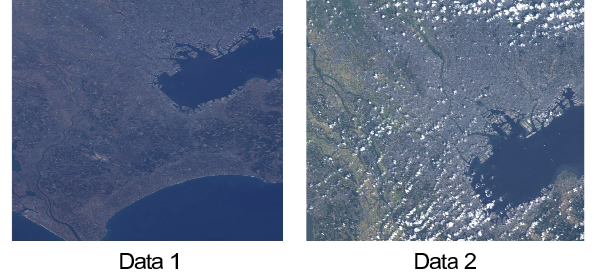
Fig. 2. Examples of ALOS-AVNIR-2 multispectral image data.
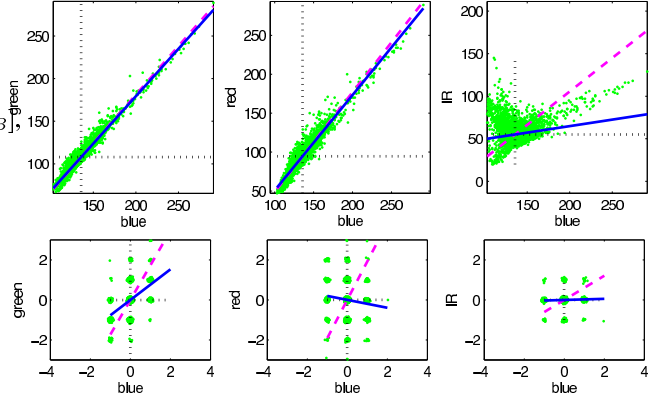
Fig. 3. DCT coefficients prediction using Blue band as reference for DC coefficient (top row) and DCT(8,8) coefficient (bottom row). Raw data shown in green dots. Data averages shown in dotted lines. Ideal LMMSE predictor plotted in solid blue line. Predictor assuming unit correlation coefficient plotted in dashed pink line.

among Blue and IR). The low correlation of all bands with the IR band is partly attributed to the separation of the wavelength of the IR channel from the end of the Red spectrum and the larger bandwidth of the IR receiver hardware. Our evaluation of a variety of other data sets from the same source provided a similar correlation pattern.

A similar correlation pattern exists in the low-frequency DCT coefficients. Specifically, we considered 8×8-DCT blocks and evaluated the correlations among DCT coefficients. As expected, the inter-band correlation in the DC and the low frequency coefficients was similar to that in the raw pixel domain. However, high frequency coefficients did not exhibit significant inter-band correlation.

If the data means and covariances are known, the coefficients of one band can be predicted from those of another band using the classical Linear Minimum Mean Squared Error (LMMSE) predictor:

$$\widehat{y} = \frac{\sigma_{x,y}}{\sigma_x^2}(x - \mu_x) + \mu_y = \rho_{x,y} \frac{\sigma_y}{\sigma_x}(x - \mu_x) + \mu_y \qquad (1)$$

where y is the coefficient to be predicted from coefficient x, $\sigma_{x,y}$ denotes their covariance, σ_y, σ_x denote their standard deviations and $\rho_{x,y}$ is the correlation coefficient. Figure 3 demonstrates the inter-band cross correlation and prediction performance for the DC (top

	G	R	IR			G	R	IR
B	.904	.853	.654		B	.949	.933	.567
G		.932	.595		G		.957	.612
R			.418		R			.464

Table 2. Raw pixel domain inter-band correlation coefficient for Data Set 1 (left) and 2 (right).

row) and the AC(8,8) (bottom row) quantized coefficients using the Blue band as a reference to predict the other three bands. The figure plots the raw data (green dots), the mean of the coefficients (dotted horizontal and vertical lines), the linear predictor using (1)(solid blue line), and the linear predictor assuming $\rho_{x,y} = 1$ and $\sigma_x = \sigma_y$, i.e., that the data are perfectly correlated (dashed pink line). Note that the raw data in the DCT(8,8) quantized coefficient domain have been slightly perturbed to demonstrate the size of each coefficient cluster.

Fig. 3 confirms that at low spatial frequencies, the approximation $\rho_{x,y} \approx 1, \sigma_x \approx \sigma_y$ is good. This approximation incurs no additional helper data overhead, other than the transmission of the quantized means explained in Section 2. For high spatial frequencies, linear prediction provides a small benefit over using the mean of the data as a predictor, i.e., over just using $\rho_{x,y} \approx 0$. However, transmitting the covariances for true linear prediction does incur additional overhead. Our experiments confirmed that using $\rho_{x,y} \approx 1, \sigma_x \approx \sigma_y$ for low spatial frequencies and $\rho_{x,y} \approx 0$ for high spatial frequencies is more efficient in terms of the overall rate-distortion tradeoff.

4. RATE-DISTORTION PERFORMANCE

The rate-distortion performance of the Wyner-Ziv coding scheme was compared against independent JPEG and JPEG2000 coding of the subbands. The PSNR metric is used to measure the quality of the individual subbands. The results are plotted in Fig. 4 for the datasets 1 and 2 pictured in Fig. 2 at various rate-PSNR combinations. We observe that the rate distortion performance is significantly better than traditional JPEG coding and comparable to that of JPEG2000 coding[3]. Further, using multiple context-specific Laplacian predictors improves the image quality at no extra rate cost. For both datasets, the improvement from distributed coding is the least for the IR band. We also experimented with decoding the IR band using the red band as side information, but the rate-distortion improvement is not significant.

5. CONCLUSIONS

We proposed a scheme for Wyner-Ziv compression of multispectral images. A JPEG-coded version of the blue band is used as side information to decode the red, green and infra-red bands. Depending on the amount of spatial detail, 8×8 blocks in an image band are divided into contexts and a context-dependent Laplacian residual model is used to initialize belief propagation decoding of each bitplane. The rate-distortion performance of this scheme is superior to traditional JPEG coding at a comparable encoding complexity. In fact, the proposed scheme achieves the same performance as independent JPEG2000 coding of the image bands without incurring the complexity of multi-pass entropy coding of JPEG2000. The design of a wavelet transform-based distributed compression algorithm for multispectral data is the focus of our current work.

6. REFERENCES

[1] D. Taubman and M. Marcellin, *JPEG2000: Image Compression Fundamentals, Standards and Practice*, Kluwer Academic Publishers, 2002.

[2] D. Slepian and J. K. Wolf, "Noiseless Coding of Correlated Information Sources," *IEEE Trans. Information Theory*, pp. 471–480, July 1973.

[3] A. D. Wyner and J. Ziv, "The rate-distortion function for source coding with side information at the decoder," *IEEE Trans. Information Theory*, vol. 22, pp. 1–10, Jan. 1976.
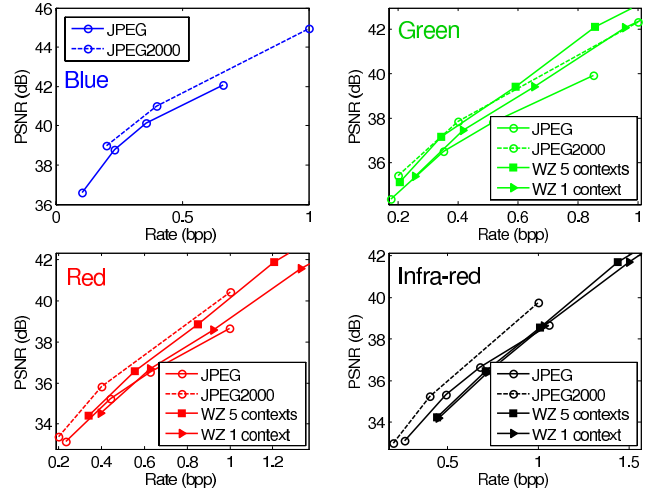
(a) Rate-distortion tradeoffs for Data 1
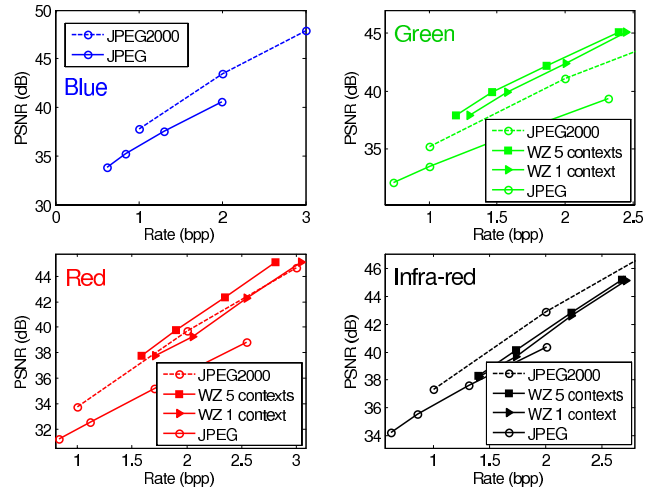
(b) Rate-distortion tradeoffs for Data 2

Fig. 4. Wyner-Ziv scheme compared with JPEG and JPEG2000.

[4] E. Magli, M. Barni, A. Abrardo, and M. Grangetto, "Distributed source coding techniques for lossless compression of hyperspectral images," *EURASIP Journal on Advances in Signal Processing*, vol. 2007, no. 1, pp. 1–13, Jan. 2007.

[5] N-M. Cheung, C. Tang, A. Ortega, and C. Raghavendra, "Efficient wavelet-based predictive Slepian-Wolf coding for hyperspectral imagery," *EURASIP Journal on Signal Processing, Special Issue on Distributed Source Coding*, vol. 86, no. 11, pp. 3180–3195, Nov. 2006.

[6] N-M. Cheung and A. Ortega, "Distributed compression of hyperspectral imagery," *Distributed Source Coding*, pp. 269–292, 2009.

[7] B. Girod, A.Aaron, S. Rane, and D. Rebollo-Monedero, "Distributed video coding," *Proceedings of the IEEE, Special Issue on Advances in Video Coding and Delivery*, vol. 93, no. 1, pp. 71–83, Jan. 2005.

[8] D. Varodayan, A. Aaron, and B. Girod, "Rate-adaptive codes for distributed source coding," *EURASIP Signal Processing Journal*, vol. 86, no. 11, pp. 3123–3130, Nov. 2006.

[9] C. Brites and F. Pereira, "Correlation noise modeling for efficient pixel and transform domain Wyner-Ziv video coding," *IEEE Trans. Circuits and Systems for Video Technology*, vol. 18, no. 9, pp. 1177–1190, Sept. 2008.

[10] G. Ungerboeck, "Channel coding with multilevel/phase signals," *IEEE Trans. Information Theory*, vol. 28, no. 1, pp. 55–67, Jan. 1982.