Efficient Inter-band Prediction and Wavelet Based Compression for Hyperspectral Imagery: A Distributed Source Coding Approach

Caimu Tang[‡], Ngai-Man Cheung[†], Antonio Ortega[†], Cauligi S. Raghavendra[§]
[† §] Dept. of Electrical Engineering, Univ. of Southern California, Los Angeles, CA
[‡ §] Dept. of Computer Science, Univ. of Southern California, Los Angeles, CA
{caimut@usc.edu, ncheung@usc.edu, ortega@sipi.usc.edu, raghu@usc.edu}

Abstract

Hyperspectral images have correlation at the level of pixels; moreover, images from neighboring frequency bands are also closely correlated. In this paper, we propose to use distributed source coding to exploit this correlation with an eye to a more efficient hardware implementation. Slepian-Wolf and Wyner-Ziv based correlated coding theorems have quantified how much additional rate reduction can be obtained. In order to better exploit these correlations, we first propose a prediction model to align images. This model is based on linear prediction techniques and it is simple and shown to be effective for hyperspectral images. We then propose a coding scheme to exploit these correlations. A set-partitioning approach is used on wavelet transformed data to extract bitplanes. Under our correlation model, bitplanes from neighboring bands are correlated and we then use a low-density parity-check based Slepian-Wolf code to exploit this bitplane level correlation. This scheme is appealing for hardware implementation as it is easy to parallelize and it has modest memory requirements. As for coding performance, our preliminary results for high correlation spectral bands from the NASA AVIRIS dataset show, at medium to high reconstructed qualities, gains of about a factor of 3 in compression efficiency as compared to encoding the spectral bands independently using SPIHT.

1. Introduction

Hyperspectral images are image data consisting of hundreds of spectral bands, leading to very large raw data size. For example, the images captured by AVIRIS (Airborne Visible/Infrared Imaging Spectrometer, operated by NASA) include 224 bands, and each single hyperspectral image contains up to 140 Mbytes of raw data; therefore, a high performance compression codec is necessary for hyperspectral imagery. In addition, hyperspectral images are usually captured by satellites that use embedded processors with limited resources, so encoding complexity is critical. In this paper, we propose a novel low complexity lossy and lossless hyperspectral image compression scheme based on distributed source coding techniques [10]. This scheme also contains an efficient inter-band prediction model. Our proposed approach achieves excellent compression performance when neighboring bands are used as side-information during decoding.

Proceedings of the 2005 Data Compression Conference (DCC'05)
1068-0314/05 $ 20.00 **IEEE**

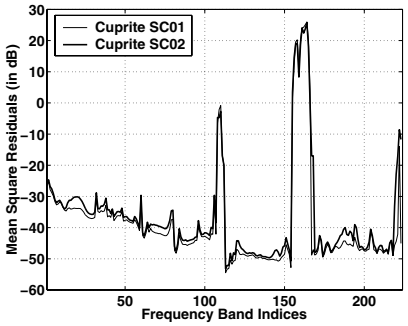


Figure 1: Mean square residuals after simple image alignment and subtraction

In hyperspectral imagery, many spectral bands are highly correlated as shown in Figure 1, where image mean-square residuals after simple alignment are shown for two different views on a site. Images from neighboring bands are correlated and the variation of correlation in many spectral regions is small. This motivates exploiting inter-band correlation in order to achieve good compression performance. Previous work on compression of hyperspectral imagery includes inter-band predictive approaches [4] and 3D wavelet approaches [19]. In inter-band prediction approaches, each band is predicted by another band, and then the residuals of prediction are compressed by intra-picture compression techniques. Compared to inter-band prediction approaches, our proposed distributed source coding approach has the following advantages. First, inter-band prediction methods need to generate exact copies of the decoded bands at the encoder, so encoders need to perform decoding as well, and decoding complexity could be as high as that of encoding. In contrast, distributed source coding needs only correlation statistics to perform encoding, and these statistics can be reliably estimated with low complexity, as we will show. Second, inter-band predictive methods are inherently serial, since encoding of each band must wait for the predictors obtained from processing previous bands. A distributed source coding approach would enable all the correlated bands to be encoded in parallel after inter-band correlations have been estimated, which requires only limited data exchange across bands. This inherent parallelism can facilitate hardware implementations and greatly increase the onboard encoding speed. Third, the proposed approach applies distributed source coding to encode wavelet coefficient bitplane data. A given bitplane in a given subband depends only on the same bitplane in a previously encoded image. Once the data has been encoded efficient rate scalability can be achieved by decoding all images up to the same bitplane resolution level. In contrast, closed loop inter-band prediction makes it difficult to achieve efficient rate scalability.

3D wavelet methods including 3D-SPECK and 3D-SPIHT [19] need to operate on several spectral bands in memory at the same time. In contrast, a distributed source coding approach would need to store only a single spectral band in memory at a time once correlation statistics are estimated. Reduction in memory requirements could potentially lead to low encoding power consumption, since off-chip memory access would be avoided. As a matter of fact, off-chip memory accesses often consumes up to one order of magnitude higher power than that for accessing on-chip data [18].

Our proposed approach for hyperspectral imagery is based on fundamental information-theoretic results from the 1970s. Slepian and Wolf [10] proved that two correlated sources

can be optimally encoded even if the encoder only has access to the two sources separately, as long as both encoded streams are available at the decoder. This counterintuitive result permits (in principle) significant complexity reductions at the encoder, where low complexity is most needed for hyperspectral imagery, while preserving the encoder's ability to optimally compress the data by exploiting the redundancy in the correlated sources. Wyner [11] suggested a practical scheme using syndrome binning. Slepian-Wolf coding was a dormant niche of information theory for nearly three decades, until the recent development of low-complexity, capacity-approaching (turbo or LDPC) channel codes. Now practical applications seem feasible, codes designed for Slepian-Wolf coding problem have been reported in a number of articles [2, 13, 14, 8]. Applications of Slepian-Wolf coding include data aggregation in sensor networks [12, 15] and video coding, e.g., [1, 3]. In the video coding application, the correlated sources are successive video frames. In this paper, correlated sources will be successive bands of hyperspectral imagery.

We propose a scheme called set-partitioning in hierarchical tree with Slepian-Wolf coding (SW-SPIHT) for hyperspectral imagery, as an extension of the well-known SPIHT algorithm [16]. SW-SPIHT first uses an iterative set-partitioning algorithm to extract bitplanes. Bitplanes at the same bit position in neighboring bands are correlated under our prediction model. Once the bitplanes from the first band, which is intra-encoded and intra-decoded, are available to the joint decoder, successive bitplanes at corresponding bit positions from other bands can be decoded. All bitplanes other than those from the first band are intra-encoded by an LDPC based Slepian-Wolf code [8, 9] and jointly decoded by a sum-product decoding algorithm. As an example of coding performance, for the NASA AVIRIS hyperspectral images data set, at medium to high quality SW-SPIHT can achieve gains of up to a factor of three (compared to SPIHT) for highly correlated spectral bands. These gains can be higher at lower reconstruction qualities. Note that when all bitplanes are encoded SW-SPIHT can also provide lossless compression. To the best of our knowledge, distributed source coding techniques have not been investigated for compression of hyperspectral imagery; This paper demonstrates that this approach is promising. Preserving the spectral signature is important in some applications (e.g., in applications where hyperspectral images are classified and the percentage of correct classification is important [19]). SW-SPIHT is flexible on selection of rate and distortion and it supports progressive transmission, so that the rate can be selected so that the spectral signature can be preserved. Detailed analysis on how and to what level SW-SPIHT preserves these application specific spectral signatures will be part of our future work.

This paper is organized as follows: we present the proposed codec and our prediction and estimation model in Section 2 and Section 3, respectively. Implementation and experimental results are shown in Section 4, and Section 5 concludes this paper.

2. Codec Design

In this section, we will present our codec for compression of two hyperspectral bands X and Y. Assume that reconstruction $\hat{X}$ of X will be used to form side-information for decoding of Y. We use a linear predictor to generate side-information from $\hat{X}$. The filtered version $\hat{X}'$ of $\hat{X}$ is as follows: $\hat{X}' = \alpha\hat{X} + \beta$, where α and β are known. Furthermore, in this section, we assume that correlation statistics in terms of crossover probabilities are also

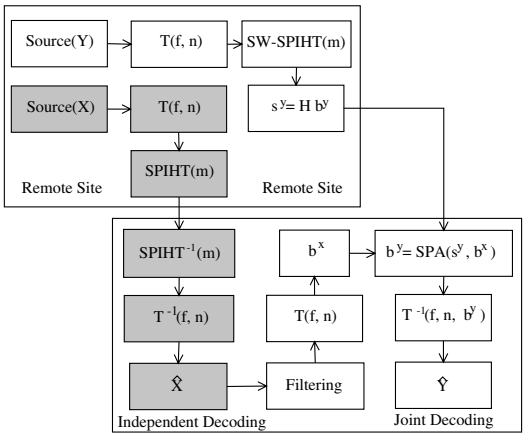


Figure 2: Block Diagram of SW-SPIHT

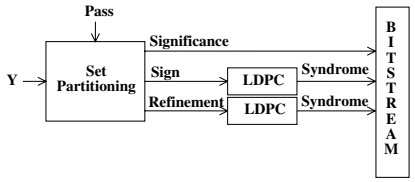


Figure 3: Process of SW-SPIHT

known to both the encoder and the decoder. Section 3 will present techniques to efficiently estimate these parameters from input data.

We first briefly describe SPIHT [16]. SPIHT uses a significance test on wavelet coefficients to partition them into two sets at each iteration: significant set and insignificant set. Bitplanes formed from the significant coefficients are directly outputted to a bitstream. The significance/structure bits convey the significance tree information to the decoder and this is used by the decoder to identify these significant coefficients. Bits representing this significance tree are entropy-coded.

Referring to Figure 2, source band X is intra-encoded by the encoder and independently decoded by the decoder as shown by the gray boxes. The reconstructed band $\hat{X}$ will be used as side-information to decode Y. The other branch consisting of white boxes in Figure 2 shows coding of Y. Image Y is first transformed by a wavelet transform $T(f, n)$ where f is the filter used in the transform and n is the number of transformation levels. Then SW-SPIHT iteratively identifies significant sets of wavelet coefficients of Y. As shown in Figure 3, at the end of each iteration, a sign bitplane, a refinement bitplane and corresponding significance bits are generated. Sign bits and refinement bits are further encoded by an LDPC based Slepian-Wolf code and corresponding syndrome bits are output to the bitstream, and significance bits are intra-coded.

At decoder, $\hat{X}'$ is transformed by $T(f, n)$ which is the same transformation used on Y at the encoder. Then the significance tree of Y (*not X*) is used to parse the wavelet coefficients and extract the side-information bitplane. Note that this significance tree is transmitted in intra-mode to the decoder. We experimented with longer predictor for better side information for decoding, but settled on a first-order predictor as it provided good enough results. Since SW-SPIHT can use a small fraction of pixels in band X and Y to estimate prediction

coefficients α and β, the computation overhead is not really an issue and details on this are shown in Section 3. LDPC sum-product algorithm (SPA) is used to decode the bitplanes of Y given syndrome bits and side-information bitplanes from $\hat{X}$. When these bitplanes are decoded, refinements to the corresponding coefficients from which these bitplanes are extracted are added. This can be done because the significance tree of Y is sent to the decoder in intra-mode.

When all bitplanes are decoded and coefficient refinements are performed, the decoder applies the inverse wavelet transform $T^{-1}(f, n)$ to reconstruct $\hat{Y}$, an estimate of Y. Since Slepian-Wolf coding is used to code these bitplanes, no or negligible information loss is introduced on these bitplanes. As long as the correlation model is correct, all bitplanes can be decoded correctly and information loss is only introduced by not transmitting some of the least significant bitplanes. This means that SW-SPIHT can also provide lossless compression for hyperspectral imagery when all bitplanes are coded, provided that an integer-to-integer wavelet transform [6] is used. Note that the least significant bitplanes tend to be uncorrelated from image to image and have near maximum entropy, thus, in lossless applications, these bitplanes can be sent uncoded.

Crossover probabilities are used by the encoder to determine the compression rate. This rate determines which parity-check matrix should be used for a bitplane. In SW-SPIHT, irregular Gallager codes are used. A table is built offline and it associates different crossover probabilities with random seeds for proper parity-check matrices. Once the crossover probability between a bitplane and its corresponding side-information bitplane is obtained, a proper parity-check matrix can be selected at run-time.

In the coding process, the wavelet decomposition scheme (e.g. dyadic, pyramidal or other wavelet packet bases [7]), the filter function f, and the number of transformation levels n should be kept the same for each band during encoding and decoding. To make sure the same parity-check matrix is used at the decoder, the random seed used by the encoder to generate the parity-check matrix is sent to the decoder. To match the exact bitplane width, column puncturing and splitting is used on the parity-check matrix.

When multiple bands are compressed together, since

$$H(b_i^z|b_i^y, b_i^x) \leq \min\{H(b_i^z|b_i^x), H(b_i^z|b_i^y)\}$$

(i.e., conditioning does not increase entropy), where b_i^z is a bitplane from a third source Z to be compressed, it could possibly decrease the compression rate to use both side information from X and Y. It is quite straightforward to extend SW-SPIHT for compression of any number of correlated bands. We have also implemented this extension, based on our test on three bands, the additional rate reduction on the third band is quite substantial. Due to the lack of space, we omitted here the description of this multi-band extension of SW-SPIHT.

3. Model of Correlation Estimation and Inter-band Prediction

Performance of distributed source coding depends strongly on the estimation of correlation and prediction parameters. In our system, we need to estimate two sets of parameters: 1) linear prediction coefficients, since our proposed system obtains side-information by linear prediction, 2) the crossover probabilities of bitplanes (*a priori* probabilities), which are used to determine the rates of the Slepian-Wolf codes. In this section, we will first outline

the system model on which the proposed codec is based. Then we will present techniques to estimate these parameters with limited data exchange and simple computations. Preliminary results show that these low-complexity techniques are effective.

Denote bitplanes from source W by b^w and a particular bitplane at bit position i of source W by b_i^w, and denote the l-th bit of bitplane b_i^w by $b_i^w(l)$. Estimation of prediction coefficients is performed at the encoder for the purpose of determining of the level of correlation. We list the steps for a two-source case:

1. **Encoder**: 1-1. Estimation of predictor coefficients α and β using pixels in X and Y; 1-2. Application of the prediction coefficients to obtain $X' = \alpha X + \beta$; 1-3. Transformation of X' using the same wavelet transform used for Y; 1-4. Application of the significance tree of Y to the wavelet coefficients tree of X' to extract m bitplanes b_i^x $(1 \leq i \leq m)$; 1-5. Estimation of crossover probability p_i of bitplane pair (b_i^x, b_i^y) $(1 \leq i \leq m)$; 1-6. Determination of the Slepain-Wolf coding rate; 1-7. Generation of parity-check matrix for b_i^y $(1 \leq i \leq m)$.

2. **Decoder**: 2-1. Application of prediction coefficients to obtain $\hat{X}' = \alpha \hat{X} + \beta$; 2-2. Transformation of $\hat{X}'$ using the same wavelets used for Y at the encoder; 2-3. Application of the significance tree of Y to the wavelet coefficients of $\hat{X}'$ to extract m bitplanes b_i^x $(1 \leq i \leq m)$; 2-4. Computation of *a priori* probability $\Pr(b_i^y(j) = 0 | b_i^x(j))$ for $b_i^x(j) = 0$ or 1.

The encoder can determine a rough level of correlation after it estimates the prediction coefficients (i.e., by measuring the mean-square residual). When the mean-square residual is above a threshold, the model can simply flag the codec to code this bitplane in intra-mode and flag it back to inter-coding mode when the mean-square residual is below a threshold. For example, Band 162 in Figure 1 should be coded in intra-mode. From real data sets, a majority of bands can be inter-coded, e.g., more than 95% bands in the Cuprite data set which is one of data sets we used in our experiment. The predictor coefficients used in (2-1.) are transmitted from the encoder. In Step (2-4.), the *a priori* probability is p_i if $b_i^x(j) = 1$, and $1 - p_i$ otherwise.

Each pixel in band Y is filtered by a linear predictor formed by pixels in X. We use the first order linear predictor $X' = \alpha X + \beta$. Least-square technique can be used to find the coefficients α and β. The linear predictor will be used to determine correlation statistics, and this filtering is also performed by the decoder to form side-information.

In addition, we also need the estimates of the crossover probabilities at the encoder for the selection of an appropriate Slepian-Wolf coding rate. The estimated probabilities are also used by the decoder for the initialization of SPA. Next we will outline the techniques to perform this estimation.

To estimate linear predictor coefficients α and β, we down-sample the image bands and use only the pixels in the down-sampled bands for estimation. As shown in Figure 4, with 0.32% of pixels, the resultant predictor could achieve a prediction mean square error (MSE) within 0.05 to that of optimal predictor, which is formed by using all pixels in X and Y.

To estimate the crossover probability we exchange a small portion of bitplane bits formed by set-partitioning. This is viable since set-partitioning can be considered as a scrambling

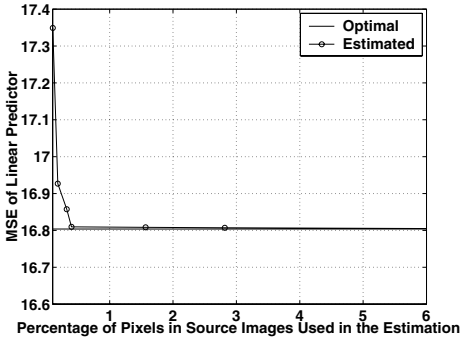


Figure 4: MSE under First-order Linear Predictor

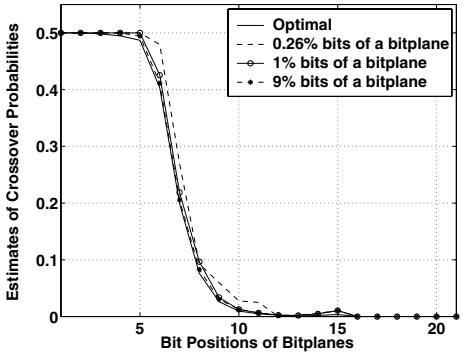


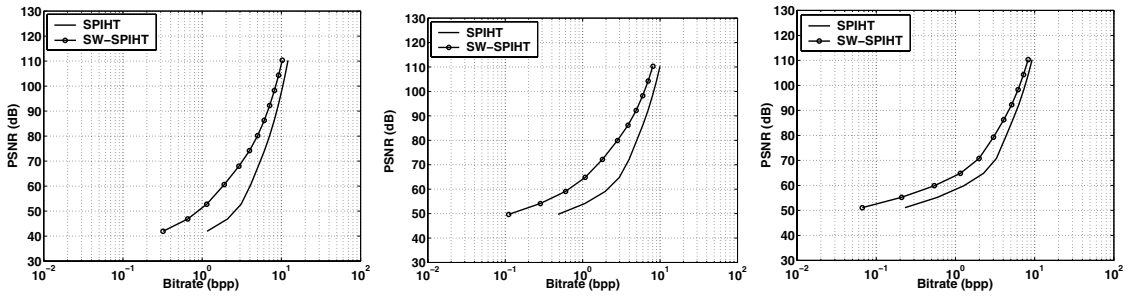
Figure 5: Estimation of Crossover Statistics

process on the ordering of coefficients and the probability mass of bits in a bitplane extracted from significant coefficients tends to be even in different segments within the bitplane. We use the upper bound of the 95 percent statistical confidence interval as the estimate. Since the estimation is simply counting of pairs of crossover in small portions of two bitplanes, thus, the overall estimation overhead is small. Figure 5 shows the estimated *a priori* probabilities by different percentages of bits of bitplanes. As an example, with 9077 bits out of 181554 bits of two bitplanes, the crossover probability estimate is within 0.003 to the actual crossover probability. This demonstrates that accurate estimates of crossover probabilities are possible with little computation and data access across bands. In addition, since the compression rate is set to have some margin of about 0.05 bits from the Slepian-Wolf limit (i.e., conditional entropy), so this estimation accuracy is sufficient.

4. Implementation and Experimental Results

We have implemented SW-SPIHT with our SPA, and hyperspectral images used in our test are 16-bit images. The SPA we implemented for SW-SPIHT is based on the algorithm in Section (III-A) of [17] and similar notations are also used here, and we made two changes to it in order for it to decode Slepian-Wolf based LDPC codes. The first change is on the initialization step to make it support for multiple side-information bitplanes. The initial posteriori probabilities of variable nodes are given as follows:

$$(q_{ml}^0, q_{ml}^1) = (p^0, 1 - p^0), 1 \leq l \leq n$$



(a) Band 40 (b) Band 133 (c) Band 190

Figure 6: Rate-Distortion Curves (Site: Cuprite, View: SC01)

where, n is the total number of check nodes, $p^0 = \Pr(b_i^z(l) = 0|\mathbf{x}_l)$, for these check nodes (w_m) which have an edge in the bipartite graph to variable node v_l. $\mathbf{x}_l$ is a vector consisting of all j-th bits from these side-information bitplanes (it is a scalar 1 or 0 in the 2-band case). Note that for the two-band case, $\Pr(b_i^y(l) = 0|b_i^x(l))$ equals to the crossover probability p_i if $b_i^x(l) \neq 0$. The second change is on the check-node update step in which we introduce a new local kernel function to force the search of the most probable codeword in a designated bin specified by the syndrome while the standard LDPC SPA searches the most probable codeword in the bin corresponding to the syndrome with all zero bits.

Experimental results use SNR and PSNR which are defined as follows:

$$\text{MSE} = E[(x - \hat{x})^2]$$
$$\text{SNR} = 10 \log_{10} \left(\frac{E[x^2]}{\text{MSE}} \right)$$
$$\text{PSNR} = 10 \log_{10} \left(\frac{(65536)^2}{\text{MSE}} \right)$$

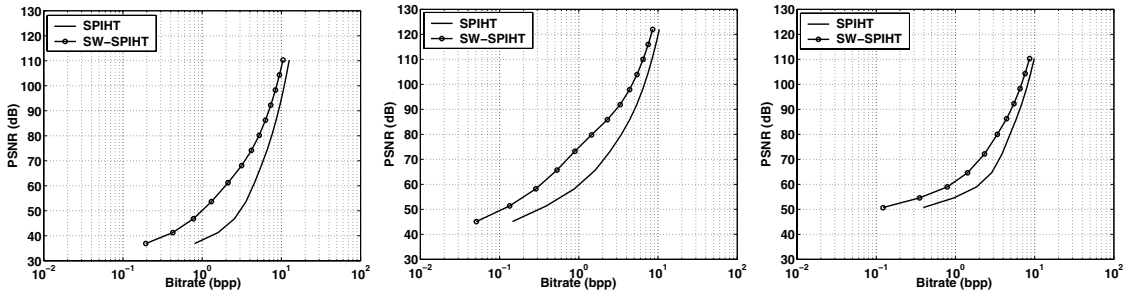
where, $E(.)$ is the expectation operator, x is the 16-bit value of a source pixel and $\hat{x}$ is the 16-bit value of reconstructed pixel of x. From our experimental data, the performance in terms of bit per pixel per band (bpppb) and multi-band SNR (MSNR) can be derived.

Figure 6 compares the performance of SPIHT and SW-SPIHT, where the X-axis in these plots is in logarithmic scales for an easy comparison at low bitrates. We selected three pairs of bands from different spectral regions where the levels of correlations are different as also shown in Figure 1. We did not select bands in spectral regions where the predictor sees large surges on mean square residuals and these bands have low correlation and intra-coding on these bands is used instead. Note that application specific metric (e.g. percent correctly identified for detection) could be used for the selection of different coding modes. In Figure 6, SW-SPIHT outperforms SPIHT significantly especially in the low bitrate regime. In the high bitrate regime, the performance of SW-SPIHT degrades somewhat, and this is due to the fact that these bits from least significant bitplanes are random and uncorrelated, and they are sent in raw mode. There are some variations on the PSNR gain due to variations of the energy among these images and correlations between images in these pairs. Table 1 shows rates obtained with different numbers of bitplanes encoded.

Figure 7 shows a set of rate-distortion curves from another view. Compared to these corresponding plots in Figure 6, the performance gains of SW-SPIHT are slightly degraded

Table 1: Rate Comparison of SW-SPIHT and SPIHT

SNR (dB)	PSNR (dB)	No. of Bitplanes	SW-SPIHT Bitrate	SPIHT Bitrate
61.96	97.93	17	4.377 bpp	6.39 bpp
37.27	73.21	13	0.89 bpp	2.48 bpp
22.244	58.19	11	0.287 bpp	0.889 bpp



(a) Band 40 (b) Band 133 (c) Band 190

Figure 7: Rate-Distortion Curves (Site: Cuprite, View: SC02)

comparing to what have been obtained for view SC01. This actually conforms to the correlation statistics of these two sets of bands as shown in Figure 1 where the residuals from view SC02 are slightly bigger than those from view SC01.

5. Conclusions and Future Work

In this paper, we have demonstrated a viable approach for compression of hyperspectral imagery. A novel scheme called SW-SPIHT on correlating and compressing hyperspectral imagery is proposed. Our scheme obtains significantly lower bitrates compared to existing techniques. Encoding of bands under SW-SPIHT can proceed in parallel once the correlation statistics are estimated, which enables a possible efficient parallel hardware implementation. Estimation of correlation statistics is simple and requires limited data exchange across bands. Encoding of bitplanes takes linear time on bitplane of width.

Currently, our prediction model is based on two images, and our future work will build a prediction model for multiple images at low cost. For the prediction part, further work also includes efficient estimation methods on prior probabilities with minimum bit exchange between bitplanes. Different applications in hyperspectral imagery (e.g. classification) have different interpretations on distortion (e.g. percentage of correctly classified). Another research direction is on analysis and experimental study of the effect of SW-SPIHT on application performance when the codec is configured to be lossy and the predictor optimization is based on application specific requirement.

References

[1] B. Girod, A. Aaron, S. Rane, and D. Rebollo-Monedero, "Distributed video coding," *Proc. IEEE, Special Issue on Advances in Video Coding and Delivery*, pp. 71–83, Jan

2005.

[2] S.S. Pradhan, and K. Ramchandran, "Distributed source coding using syndromes (DISCUS)". *IEEE Trans. on Inform. Theory*, 49(3): 626–643, March 2003.

[3] R. Puri and K. Ramchandran , "PRISM: A 'Reversed' Multimedia Coding Paradigm," *Proc. of IEEE ICIP*, Sept. 2003.

[4] A. C. Miguel, A. R. Askew, A. Chang, S. Hauck, and R. E. Ladner, "Reduced complexity wavelet-based predictive coding of hyperspectral images for FPGA implementation," *Proc. IEEE DCC*, 2004.

[5] S. R. Tate, "Band ordering in lossless compression of multispectral images," *IEEE Transactions on Computers*, vol. 46, pp. 477–483, Apr. 1997.

[6] A. R. Calderbank, I. Daubechies, W. Sweldens, and D. Yeo, "Wavelet transforms that map integers to integers," *Applied Comp. Harmonic Analy.*, 5: 332–369, 1998.

[7] M. V. Wicherhauser, "INRIA lectures on wavelet packet algorithms," http://www.math.wustl.edu/∼ victor/papers/lwpa.pdf, Nov. 1991.

[8] A. D. Liveris, Z. Xiong, and C. N. Georghiades, "Distributed compression of binary sources with side information at the decoder using LDPC codes," *IEEE Communication Letters,* 6:440–442, Oct. 2002.

[9] A. D. Liveris, C-F. Lan, K.R. Narayanan, Z. Xiong, and C. N. Georghiades, "Slepian-Wolf coding of three binary sources using LDPC codes," *Proc. Intl. Symp. Turbo Codes and Related Topics*, Brest, France, September 2003.

[10] D. Slepian, and J. K. Wolf. "Noiseless coding of correlated information sources", *IEEE Trans. on Inform. Theory*, IT-19(4): 471–480, July 1973.

[11] A. Wyner, "Recent results in the Shannon theory", *IEEE Trans. on Inform. Theory*, IT-20(1): 2–10, 1974.

[12] Z. Xiong, A. Liveris, and S. Cheng, "Distributed source coding for sensor networks," *IEEE Signal Processing Magazine*, 21: 80-94, September 2004.

[13] J. Garcia-Frias and Y. Zhao, "Compression of correlated binary sources using Turbo codes," *IEEE Comm. Letters*, 5:417–419, Oct. 2001.

[14] A. Aaron and B. Girod, "Compression with side information using Turbo codes, " *Proc. IEEE DCC*, pp. 252–261, Apr. 2002.

[15] S.S. Pradhan, J. Kusuma, and K. Ramchandran, "Distributed compression in a dense microsensor network," *IEEE Signal Processing Magazine*, 19:51–60, Mar. 2002.

[16] A. Said and W. A. Pearlman, "A new, fast, and efficient image codec using set partitioning in hie rarchical trees," *IEEE Trans. on Cir. and Sys. for Video Tech.*, 6: 243–250, June 1996.

[17] D. J. C. MacKay, "Good error-correcting codes based on very sparse matrices", *IEEE Trans. of Inform. Theory*, 45:399–431, Mar. 1999.

[18] M. Pedram, and J. M. Rabaey, Eds, "Power Aware Design Methodologies", Kluwer Academic Publishers, 2002.

[19] X. Tang, W. A. Pearlman, and J. W. Modestino, "Hyperspectral Image Compression Using Three-Dimensional Wavelet Coding: A Lossy-to-Lossless Solution," submitted to IEEE Transactions on Geoscience and Remote Sensing.