DISTRIBUTED CODING OF MULTISPECTRAL IMAGES: A SET THEORETIC APPROACH

Xin Li

Lane Department of Comp. Sci. and Elec. Engr.
West Virginia University, Morgantown, WV 26506-6109

ABSTRACT

Distributed coding problem poses the challenge of how to shift the exploitation of the correlation structure of source from encoder to decoder with minimal degradation on coding efficiency. In this paper, we propose a novel convex-set theoretic framework for distributed coding of multispectral images. Alternating projection based decoding algorithms are developed to exploit the correlation among different spectral channels at the centralized decoder. Both asymmetric and symmetric protocols are studied and compared. Experiment results have shown that the proposed symmetric distributed coder only falls behind standard wavelet coders (e.g., JPEG2000) by less than 2dB at the bit rate of 1-2bpp.

1. INTRODUCTION

Due to rapid development of sensor network technologies, distributed source coding (DSC) has attracted much attention recently. Practical limitations in sensor networks (e.g., power and bandwidth) determine that each distributed sensor has to employ a low-complexity encoding algorithm but relies on a centralized node (base station) to collect and jointly process the information at the decoder. The grand challenge with DSC is how to intelligently shift the exploitation of the correlation structure of the source from encoder to decoder with minimal coding efficiency loss. Theoretical framework of DSC was first established by Slepian and Wolf [1] in 1973 and then extended into lossy coding case by Wyner and Ziv [2] in 1976. Only recently, several practical DSC algorithms for *binary* source whose performance is close to the fundamental limit were proposed [10, 12, 13].

However, distributed coding problem for *image* source is still largely open. Due to the curse of dimensionality, the simplified assumption with correlated binary source does not hold in the practice of image sensing systems. Existing approaches to distributed image coding (e.g., [9, 11]) indicate significant coding efficiency loss when compared to the traditional image/video coding algorithms. It is unknown whether theoretically appealing results of Slepian-Wolf Theorem (or its counterpart in the lossy scenario: Wyner-Ziv Theorem [2]) can be achieved for image source by any practical distributed coding algorithms.

In this paper, we address the importance of modeling the correlation structure of image source for distributed coding. There are primarily two types of correlation structure with images acquired by a sensor network: *photometric* (sensors are spatially aligned and capture the same scene at different spectral channels or with varying lighting conditions) and *geometric* (an array of sensors are placed around the scene and images are related to each other through multi-view geometry). Both types of correlation structure are nontrivial to model; relatively speaking, geometric correlation is more difficult to exploit because of the geometric uncertainty (i.e., to establish the correspondence among images). Therefore, we choose to focus on the former case and take distributed coding of multispectral images as the example.

One fundamental issue with distributed image coding is the domain to exploit the correlation structure. Previous works (e.g., Wyner-Ziv coder [10]) opt to do so in the space of quantized pixel values or transform coefficients. We propose a novel set-theoretic framework for distributed coding of multispectral images. Under this new framework, both the correlation structure and observation constraint of image source can be characterized by convex sets. It follows that joint decoding of multispectral images can be implemented via iterative projection-onto-convex-set (POCS) operations at the centralized node. The convexity of constraint sets guarantees the convergence of our iterative decoding algorithms.

We also study two protocols with distributed coding of

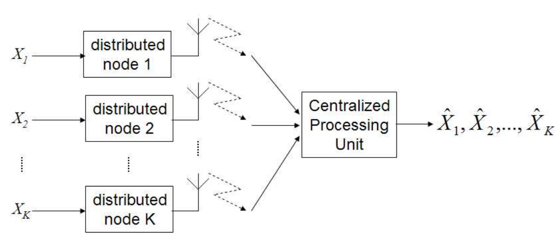

Fig. 1. Distributed source coding paradigm.

multispectral images: asymmetric and symmetric [6]. In the asymmetric protocol, one spectral channel is treated with the highest priority for it is used as the side information to aid the decoding of other bands at the base station (refer to Fig. 2). In the symmetric protocol, all spectral channels (or group of channels) are peer to each other and mutually work as the side information at the decoder. Both protocols are of interest in practical applications and it is not known which one leads to better Rate-Distortion (R-D) performance. We advocate the symmetric protocol for its simplifying mathematical analysis and our experiment results also appear to support its superiority for the chosen test images.

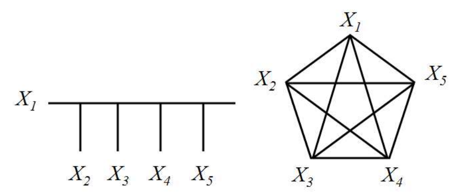

Fig. 2. Left: asymmetric case ($K = 5$); Right: symmetric case ($K = 5$).

2. PROBLEM FORMULATION

We consider lossy coding of K images: $X_1, X_2, ..., X_K$, each of which represents the acquisition of the scene at a different spectral channel. In the traditional coding paradigm, source encoder has the access to all nodes and therefore can employ a variety of mathematical tools (e.g., decorrelating transforms) to optimize the R-D performance. However, in the distributed coding paradigm, each X_k is processed by an individual node (distributed source encoder) without the knowledge of others; only the centralized decoder is allowed to collect the information from all nodes and attempt to best resolve their uncertainty in the R-D sense. Formally, we want to minimize $D = \sum_{k=1}^{K} D_k$ subject to the rate constraint (i.e., $\sum_{k=1}^{K} R_k \leq R$) and the distributivity constraint (i.e., there is no communication among the nodes).

We suggest that two issues are essential to the DSC problem: 1) the correlation structure of the source $X_1, X_2, ..., X_K$; 2) the protocol shared by distributed encoders and centralized decoder. In most previous works of DSC, the correlation structure is statistical dependency between two binary sources. Despite the mathematical elegance and tractability, such correlation model does not fit the imaging process in practical sensor networks. We believe that there are primarily two types of correlation structure in practical sensor networks: photometric and geometric. In the former case, each node acquires the scene at a different sepctral channel (or lighting condition). In the latter case, each node acquires the scene from a different position (or angle). Relatively speaking, geometric correlation structure is more difficult to exploit in DSC because it involves resolving geometric uncertainty (i.e., to estimate the correspondence among images) at the decoder; while correspondence itself is a long-standing open problem. For this reason, we limit our discussion to the former case here.

There are two generic types of protocols shared by distributed encoders and centralized decoder:

1) Asymmetric - the protocol specifies one node (w.l.g., we assume it is X_1) to be the prioritized source. Accordingly, the majority of bits will be allocated to compress X_1 to reach a relatively small distortion because the decoded X_1 will work as the side information to facilitate the decoding of other nodes for $k > 1$ (see Fig. 2). For any other node than X_1, a coarse representation is transmitted through the channel for the reason of reducing the bit rate and we can only rely on the centralized decoder to reduce its distortion by exploiting its correlation to X_1.

2) Symmetric - each node in the network is treated equally by the protocol. Consequently, at the centralized decoder, all nodes other than X_l can be used as the side information to help the decoding of X_l. For some K values, it is also possible to group some nodes and treat them as a "supernode" for the reason of simplifying the implementation. What distinguishes symmetric from asymmetric is the chicken-and-egg flavor in the decoding. That is, if $X_i \rightarrow X_j$ denotes that X_i is used as the side information to decode X_j (i.e., there is a directed edge from node X_i to X_j), we will find many closed loops in the graph shown in Fig. 2 (e.g., $X_1 \rightarrow X_2 \rightarrow X_3 \rightarrow X_4 \rightarrow X_5 \rightarrow X_1$).

3. SET THEORETIC DISTRIBUTED CODING ALGORITHMS

In this section, we present a novel set theoretic framework for distributed coding of multispectral images. The motivation behind is that both the observation constraint and the correlation structure of multispectral image source can be formulated into convex sets. Therefore, joint reconstruction of the images at the centralized node boils down to identifying a point at the intersection of the convex sets. It follows that distributed decoding can be implemented by alternatively projecting onto the defined convex sets. The key to the effectiveness of such set theoretic DSC algorithms is the appropriate definition of convex sets. We will study asymmetric and symmetric cases respectively next.

A. Asymmetric Case

In this case, we assume that X_1 is coded at the rate R_1 and a coarse representation of X_k ($k > 1$) is coded at a reduced rate $r < R_1$ by a standard wavelet codec. The key challenge is how to exploit the side information (reconstructed X_1) at the decoder to further reduce the distortion of X_k. We propose to solve such decoding with side information using the convex set theory.

The first constraint set is introduced to resolve the uncertainty of X_k by exploiting its correlation to X_1. Since they are the description of the same scene at different spectral channels, the correlation model in the redundant wavelet domain introduced in [3] can be employed. The constraint set is given by

$$S_0 = \{X_k : |W_1(m,n) - W_k(m,n)| < T, (m,n) \in \Omega_h\}. \quad (1)$$

where W denotes wavelet transform of X, Ω_h the high-frequency bands in the wavelet domain and T a positive real value. The proof of convexity of the above set is referred to [3].

The second constraint set is to enforce the consistence with the observation made with the received X_k. Depending on what kind of coarse representation is transmitted to the decoder, we can define a variety of constraint sets. For example, if the encoder chooses to code a decimated version of X_k, then the constraint set is given by

$$S_1 = \{X_k : X_k(m,n) = \bar{X}_k(m,n), \forall (m,n) \in \Omega_D\}. \quad (2)$$

where $\bar{X}(m,n)$ is the quantized sample and Ω_D is the sampling lattice corresponding to the decimation. Alternatively, if the encoder simply codes X_k by a standard wavelet coder, we can then define the constraint set to be

$$S_2 = \{X_k : |W_k(m,n) - \bar{W}_k(m,n)| < Q, \forall (m,n)\}. \quad (3)$$

where $\bar{W}$ is the quantized wavelet coefficients and Q is the stepsize of uniform quantizer. The convexity of S_1, S_2 can be easily verified [3, 5]. With more than one constraint set, we can refine the estimate of X_k by alternatively projecting it onto the defined constraint sets. The convexity of constraint sets guarantees the convergence of iterative decoding.

One tantalizing open question is the R-D optimization problem in the asymmetric case, i.e., how to minimize the total distortion $D = \sum_{k=1}^{K} D[X_k]$ under the constraint $R_1 + (K-1)r \leq R$? Since the distortion terms in the summation are not independent (e.g., $D[X_1]$ will affect $D[X_k]$ through the constraint set S_0), existing Lagrangian-multiplier based techniques are not appropriate for solving such nonlinear constrained optimization problem. An approximated solution is to empirically generate the look-up-table $D_k = f(D_1)$ and then exhaustively search the optimal point along the operational R-D curve. For simplicity, we adopt the following ad-hoc rate-allocation strategy in our implementation: $R_1 = \frac{3}{4}R, r = \frac{1}{4(K-1)}R$.

B. Symmetric Case

In the symmetric case, all nodes have the same priority. Let us consider the baseline case $K = 2$ first. Due to the symmetry, source encoder will code the coarse representations of both X_1 and X_2 and rely on the decoder to decode them in a joint fashion for the purpose of lowing the distortion. As mentioned above, we need to handle the "chicken-and-egg" problem caused by symmetric treatment of two nodes (when $K = 2$, the closed loop is $X_1 \rightarrow X_2 \rightarrow X_1$).

To overcome such difficulty, we propose to consider the following constraint set

$$S_3 = \{d : d(m,n) = \frac{1}{N} \sum_{(m',n') \in N(m,n)} d(m',n')\}. \quad (4)$$

where $d(m,n) = X_1(m,n) - X_2(m,n)$ denotes the color-difference signal and $N(m,n)$ indicates a local neighborhood of (m,n) with the size of N. Roughly speaking, the above constraint set specifies the color-difference signal can be characterized by an N-th order linear interpolative model. The convexity of S_3 can be readily justified [4] and it is symmetric about X_1, X_2 (invariant to swapping).

In the symmetric case, we also need to be more careful about the definition of observation constraint sets. For example, if the encoder chooses to code decimated versions of X_1, X_2, then the observation constraint sets can be written as

$$S_4 = \{X_k : X_k(m,n) = \bar{X}_k(m,n), \forall (m,n) \in \Omega_{Dk}; k = 1,2\}. \quad (5)$$

where $\bar{X}_{1,2}(m,n)$ still denote the quantized samples; but $\Omega_{D1,D2}$ (the sampling lattices on which decimation is performed at nodes 1 and 2 respectively) need to be complement to each other for the reason of symmetry. For example, we suggest to choose $\Omega_{1,2}$ to be the two quincunx-sampled lattices with different phases.

Due to the symmetry of nodes X_1 and X_2, it is reasonable to assume that POCS-based decoding effectively reduces the distortion of both nodes by the same factor $a > 1$. For example, if we model $X_1 \sim N(0, \sigma_1^2), X_2 \sim N(0, \sigma_2^2)$, then $D = \frac{1}{a}(\sigma_1^2 2^{-2R_1} + \sigma_2^2 2^{-2R_2})$. Then we can revoke the classical Lagrangian multiplier based technique to solve the R-D optimization problem in symmetric DSC. One way of generalizing into arbitrary $K > 2$ is to consider pairwise relationship among the nodes. However, such strategy requires the inspection of $K(K+1)/2$ constraint sets (i.e., a complete graph with K vertices and $K(K+1)/2$ edges), which increases at the order of $O(K^2)$. Instead, we propose to only consider the correlation in a cyclic order (i.e., only consider one directed and closed loop in the graph shown in Fig. 2), which will keep the size of constraint sets to be $O(K)$.

4. EXPERIMENT RESULTS

In this section, we use some preliminary experiment results to demonstrate the performance of the proposed distributed

Fig. 3. Kodak test image *sail* used in this work (dimension 1536×1024).

coding algorithms. The color image (three spectral channels - R/G/B) used in our experiments is shown in Fig. 3. We have implemented two distributed image coders (DIC): 1) Coder I - asymmetric case, where the G channel is treated as prioritized node X_1 and the constraint sets are chosen to be Eqs. (2) and (3)). The G channel is coded at full-resolution and R/B channels are coded at half-resolution by standard wavelet codec (e.g., JPEG2000) with the ad-hoc rate allocation strategy ($R_1 = \frac{3}{4}R, R_2 = R_3 = \frac{1}{8}R$). The implementation details of projection operators can be found in [3]; 2) Coder II - symmetric case, where four decimated channels, namely $R_{1,2}, G_{1,1}, G_{2,2}, B_{2,1}$ (subscripts denote the phases of decimated sampling lattices), are coded by JPEG2000 at an equal bit rate ($\frac{R}{4}$). We note that such Bayer-pattern sampling strategy is a mixture of symmetric protocol of $K = 3$ and $K = 2$ for the reason of simplifying the implementation (i.e., R/B channels are treated as the supernode and we ignore the correlation between R and B). The constraint sets are chosen to be Eqs. (4) and (5)), whose projection operators are referred to [4].

Table 1 shows the R-D performance comparison between three coders: JPEG2000 and our two DICs. It can be seen from Table I that on the average symmetric protocol achieves higher PSNR than asymmetric one. At the middle bit rate (1-2bpp), the PSNR performance of DIC-II falls behind that of JPEG2000 by less than 2dB. When compared with the dramatic coding efficiency loss in the previous DSC works, the performance of the proposed set theoretic DSCs appears to be promising.

5. REFERENCES

[1] J.D. Slepian and J.K. Wolf, "Noiseless coding of correlated information sources," *IEEE Trans. on Infor. Theory*, vol. 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 Infor. Theory*, vol. 22, pp. 1-10, Jan. 1976.

[3] B. K. Gunturk, Y. Altunbasak and R. M. Mersereau, "Color plane interpolation using alternating projections ," *IEEE Trans. on Image Proc.*, vol. 11, no. 9, pp. 997-1013, Sep. 2002.

[4] X. Li, "Demosaicing by successive approximation," to appear in *IEEE Trans. on Image Proc.*, 2004.

[5] H. Stark Ed. *Image Recovery: Theory and Application*. Academic Press, 1987.

[6] S. S. Pradhan and K. Ramchandran,"Distributed source coding: Symmetric rates and applications to sensor networks ", Proc. IEEE Data Compression Conference (DCC), March 2000.

[7] S. Pradhan, J. Kusuma, and K. Ramchandran. Distributed compression in a dense microsensor network. *IEEE Signal Proc. Magazine*, pages 51–60, March 2002.

[8] Rohit Puri and Kannan Ramchandran. "PRISM: A New Robust Video Coding Architecture Based on Distributed Compression Principles," 40th Allerton Conf. on Commun., Control and Comput., Allerton, IL, 2002.

[9] A. Aaron, E. Setton and B. Girod, "Towards practical Wyner-Ziv coding of video," *Proc. IEEE Int. Conf. on Image Proc.*, Sep. 2003.

[10] A. Aaron and B. Girod, "Compression with Side Information Using Turbo Codes," *Proc. IEEE Data Compression Conference*, April 2002.

[11] R. Wagner, R. Nowak, and R. Baraniuk. Distributed image compression for sensor networks using correspondence analysis and super-resolution. In *IEEE Int. Conf. On Image Proc.*, 2003.

[12] J. Garcia-Frias, "Compression of correlated binary sources using turbo codes," *IEEE Compression Letters*, vol. 5, no. 10, pp. 417-419, Oct. 2001.

[13] A. Liveris et al., "Compression of binary sources with side information at the decoder using LDPC codes," *IEEE Compression Letters*, vol. 6, no. 10, pp. 440-442, Oct. 2002.

Rate	channel	JPEG2000	DIC-I	DIC-II
1bpp	R	41.1	38.3	39.4
	G	41.5	39.8	39.9
	B	40.8	38.3	39.4
2bpp	R	44.5	39.7	42.3
	G	45.4	44.0	43.3
	B	44.1	39.7	42.3

Table 1. PSNR (dB) comparison among JPEG2000, DIC-I (asymmetric case) and DIC-II (symmetric case) for the *sail* image.