A Novel Slepian-Wolf Decoding Algorithm Exploiting Geometric Regularity Constraints with Anisotropic MRF Modeling

Yongsheng Zhang[1], Hongkai Xiong[1], Chang Wen Chen[2]
[1]Department of Electronic Engineering, Shanghai Jiao Tong University, Shanghai 200240, China
[2]Department of Computer Science and Engineering, State University of New York at Buffalo, NY 14260-2000, USA

***Abstract*— This paper proposes a novel Slepian-Wolf decoding algorithm for distributed video coding by exploiting not only the statistical correlation between the side-information and source but also the spatio-temporal consistency constraint of video sequences. The proposed algorithm models the log-likelihood-ratio (LLR) information for Slepian-Wolf decoding as an anisotropic MRF model and solving the inference by iteratively performing conventional probabilistic Slepian-Wolf decoding, which imposes the global bit-wise constraint from the Slepian-Wolf encoding process, and MRF optimization with belief propagation (BP) to enforce the local geometric regularity constraint of video frames. Experimental results demonstrate a considerable performance gain beyond existing Slepian-Wolf decoding algorithms in literature.**

I. Introduction

Distributed Video Coding (DVC) acts as a new video coding paradigm motivated by the requirements of recent emerging applications, *e.g.* mobile camera phone and wireless visual sensor networks, which desire low encoding complexity due to the battery life constraint [1]. To improve the rate-distortion (RD) performance, a variety of approaches have been developed, such as constructing more accurate side information and adopting decorrelation transform [1]. In this paper, a novel Slepian-Wolf (SW) decoding algorithm for DVC applications is introduced by exploiting the spatial regularity constraint with anisotropic Markov random field (MRF) modeling.

Practical DVC schemes are generally composed of a quanitzer and an SW codec. Although SW codec schemes realized by probabilistic channel decoding achieve a very close performance to the theoretical lower bound of encoding rate for random binary sources [2], there is still a big gap between DVC and the state-of-the-art predictive video coding schemes, *e.g.*, H.264/AVC [1]. In existing SW coding schemes, the correlation between source and side-information is modeled by a statistical *i.i.d.* model, *e.g.*, Laplacian distribution, which measures only the conditional probability distribution of source given corresponding side-information. However, more sophisticated spatial regularity correlation exists between neighboring pixels in video frames.

Although de-correlation transform has been introduced into DVC schemes to eliminate the spatial redundancy of video frames [1], it only exploits spatial correlation inside the transform window and introduces excessive computation to the encoder. To exploit the spatial correlation while maintaining the low encoding complexity, *Zhang et.al.* have introduced an algorithm to exploit spatial smoothness constraint with isotropic Markov modeling [3].

In this paper, a novel SW decoding algorithm is proposed to exploit not only the statistical correlation between side-information and source but also the local geometric regularity constraint of video frames with anisotropic MRF modeling, instead of the isotropic MRF model in [3]. The anisotropic MRF model consists of two terms: a *data term* measuring the deviation of the MRF solution to the initial estimate from conventional SW probabilistic decoding, and a *geometric regularity (GR) term* enforcing the anisotropic local structure regularity constraint of natural video frames. By iteratively performing conventional SW decoding and MRF optimization, the global bit-wise constraint from the SW encoding and the local geometric regularity constraint of video frames can be both exploited in the decoding process. The advantages of this novel decoding algorithm involve two aspects: First, the low encoding complexity virtue of pixel-domain DVC scheme is preserved since the encoding process is unaltered. Second, DVC schemes with the proposed algorithm are more flexible because better RD performance can be achieved by introducing more sophisticated geometric regularity models without changing the encoder.

The rest of this paper is organized as follows. Section II presents the DVC architecture with the proposed SW decoding algorithm. Formulation and implementation details are provided in Section III and Section IV, respectively. Experimental results are shown in Section V, and section VI concludes this paper.

II. DVC Architecture with the Proposed Slepian-Wolf Decoding Algorithm

The architecture of the DVC scheme with the proposed SW decoding algorithm is shown in Fig. 1, where the encoding process is identical to conventional pixel-domain DVC (PD-DVC) schemes. A subset of frames, termed as “Key frames”, are encoded and decoded using conventional predictive encoding schemes, *e.g.* H.264/AVC, and serve as reference frames for DVC decoding. The rest frames, called “WZ frames”, are fed into the Wyner-Ziv(WZ) encoder. Pixels in a WZ frame are uniformly quantized into M bits, and all these bits are grouped to form M bit-planes according to their significance. Each bit-plane is fed into the SW encoder to generate WZ

The work was supported in part by the NSFC, under grants 60802019, 60928003, 60736043, and the Program for New Century Excellent Talents in University, under Grant NCET-09-0554.

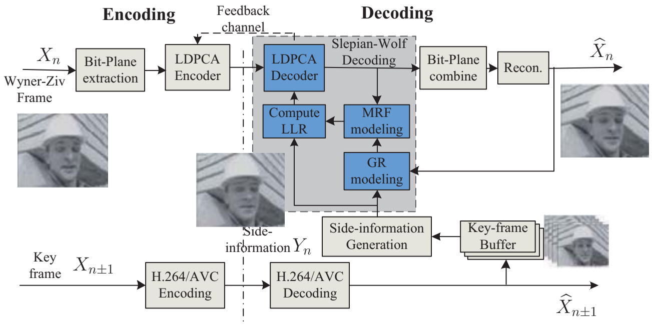


Fig. 1. Pixel domain DVC scheme with the proposed SW decoding algorithm.

bit-stream. The LDPCA approach [2] is used as conventional SW codec in this paper.

At the decoder side, side-information Y_n is firstly generated by bi-directional motion compensation with previously decoded Key frames [4]. Statistical correlation between the generated side-information and current WZ frame is then modeled to compute the log-likelihood ratio (LLR) of each bit in module "Compute LLR". The derived LLR will be fed into the SW decoder together with received WZ bit-stream to recover the original bit-plane. The decoder would request additional bits from the encoder's buffer through feedback channel (FBC) until successful decoding. Correctly decoded bit-planes are combined to determine the quantization bin $[B_L, B_U]$ of each pixel in the WZ frame. Finally, the "Recon." module would produce a best estimate of the original pixel value, based on the derived quantization bin and side-information, as the output of the DVC decoder.

In the proposed SW decoding algorithm, in addition to the LLR information, spatial geometric regularity constraint of the WZ frames is also exploited by an anisotropic MRF model in the "MRF modeling" module. The geometric regularity constraint is represented by the gradients in different directions for each pixel obtained in module "GR modeling". With MRF optimization, the LLR information from LDPCA decoding is refined and fed back to the LDPCA decoder for next iteration of decoding. In fact, it could be noticed that the proposed algorithm is applicable to all SW codec schemes with probabilistic decoding.

III. Problem Formulation

A. Motivation

In conventional SW coding, the source and side-information are typically modeled as correlated *i.i.d.* random sources. Under this model, the only constraint that can be exploited for SW decoding is the statistical correlation between corresponding source and side-information. Although the efficiency of this model has been validated on correlated *i.i.d.* random sources with capacity approaching performance [2], it is not suitable for DVC applications where remarkable spatial regularity presents between neighboring pixels in WZ frames. Therefore, more sophisticated models are desired to exploit the spatial geometric regularity constraint together with the statistical correlation to reduce the required SW bit-rate.

B. The anisotropic MRF model

In this paper, the LLR information of each bit-plane is modeled by a first-order 8-connected anisotropic MRF model. Second and higher order cliques in the MRF are ignored for simplicity. The underlying nodes consist of two types of cliques: clique $\{l_i, \widetilde{l}_i\}$ and cliques $\{l_i, l_j\}$. Clique $\{l_i, \widetilde{l}_i\}$ is called *data term*, which uses *energy function* $E_i(l_i)$ to measure the deviation between MRF solution l_i and LLR $\widetilde{l}_i$ from LDPCA decoding. Cliques $\{l_i, l_j\}$ are called *geometric regularity (GR) term*, which enforces the spatial geometric regularity consistency constraint with *energy function* $E_{i,j}(l_i, l_j)$, where $j \in \mathcal{N}(i)$ indicates the index of direct neighbors of node i. Accordingly, the *energy function* $E(\mathbf{l})$ for the anisotropic MRF model can be written as

$$\begin{aligned} E(\mathbf{l}) &= \sum_i E_i(l_i) + \lambda \sum_{\{i,j\}\in\mathcal{N}} E_{i,j}(l_i, l_j) \\ &= \sum_i |(l_i - \widetilde{l}_i)| + \lambda \sum_{\{i,j\}\in\mathcal{N}} |l_i - g(h(l_j) + \nabla_{i,j})| \end{aligned} \quad (1)$$

where parameter λ regularizes the relative ratio between the *data term* and the *GR term*. $\nabla_{i,j}$ is the target gradient between neighboring pixel pair x_i and x_j, which measures the local geometric regularity constraint of the WZ frame, $\widetilde{l}_i$ is the estimated LLR by LDPCA. Since the actual value of $\nabla_{i,j}$ is unavailable, it is approximated by $\nabla'_{u(i),u(j)}$, the gradient between pixel $x_{u(i)}$ and $x_{u(j)}$ in the reference frame by compensating relative motion $u(\cdot)$. Function $h(l_j)$ gets the corresponding side-information $\widehat{y}_j$ of pixel x_j that produces LLR l_j, while function $g(\widehat{y}_j)$ returns the LLR l_j given $\widehat{y}_j$. In this means, given LLR l_j of neighbor pixel x_j and the gradient $\nabla_{i,j}$ between pixel x_i and x_j, the LLR for pixel x_i can be approximated by $g(h(l_j) + \nabla_{i,j})$.

IV. Implementation Details

In the anisotropic MRF modeling, there are several items worth addressing: optimization algorithm for the anisotropic MRF problem, and function $h(\cdot)$ and $g(\cdot)$ to convert values between the pixel and the LLR domain.

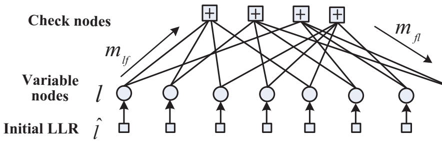
A. Belief Propagation

In LDPCA decoding process, belief propagation (BP) algorithm is used to recover each bit-plane with the received SW bit-rate by continuously propagating messages between neighboring nodes, as shown in Fig. 2. The message propagation progress is realized with the following update equation

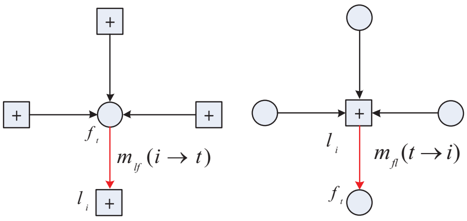
$$m_{lf}(i \to t) = \widehat{l}_i + \sum_{k\in N(i)\setminus t} m_{fl}(k \to i) \quad (2)$$

$$m_{fl}(t \to i) = \ln \frac{1 + \prod_{k\in N(t)\setminus i} \tanh(m_{lf}(k \to t)/2)}{1 - \prod_{k\in N(t)\setminus i} \tanh(m_{lf}(k \to t)/2)} \quad (3)$$

where $m_{lf}(i \to t)$ and $m_{fl}(t \to i)$ are message from variable node l_i to check node f_t and message from check node



(a) Factor graph



(b) Local message passing

Fig. 2. LDPC factor graph and local message passing.

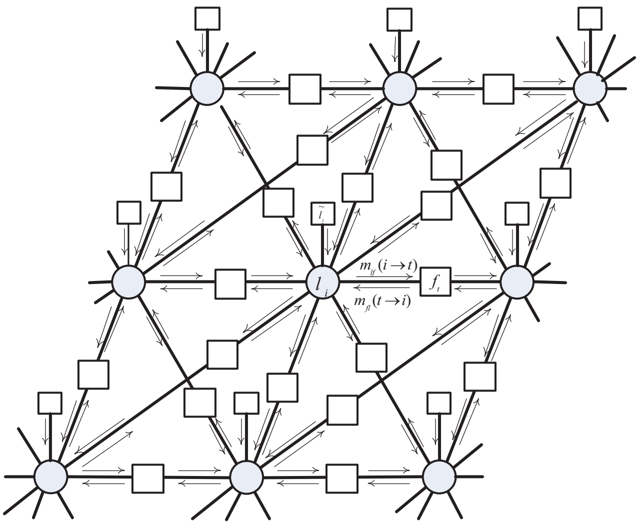


Fig. 3. Factor graph of the anisotropic MRF model, and local message pass in the network. Circles are variable nodes, gray squares are factor nodes.

f_t to variable node l_i, $N(i) \setminus t$ is the set of check nodes incident to message node i excluding f_t, $N(t) \setminus i$ is the set of message nodes incident to check node f_t excluding l_i, and function $\tanh(x) = \frac{e^x + e^{-x}}{e^x - e^{-x}}$. At beginning of message passing, $m_{lf}(i \to t)$ is initialized as $\widehat{l}_j$.

Since the anisotropic MRF optimization is also implemented in the LLR domain, we modified the message passing equation of LDPC decoding to adapt the MRF optimization problem in LLR domain. The update formula for message passing from factor nodes to variable nodes are same as that in eq. (3). However, factor nodes in the anisotropic MRF model considered in this paper have no more than two neighboring variable nodes, as shown in Fig. 3. Therefore, the message update function from check nodes to variable nodes can be accordingly simplified as

$$m_{fl}^{MRF}(t \to i) = \ln \frac{1 + \tanh(m_{lf}(k \to t)/2)}{1 - \tanh(m_{lf}(k \to t)/2)} = m_{lf}(k \to t). \tag{4}$$

That is, a factor node just forward the received message to its neighbors without processing. To model the anisotropic property of the MRF model, $m_{fl}^{MRF}(t \to i)$ is modified as follow to take into account the anisotropic property of natural images

$$m_{fl}^{MRF}(t \to i) = g(h(m_{lf}(k \to t)) + \nabla_{i,j}), \tag{5}$$

where function $g(\cdot)$ and $h(\cdot)$ are used to transit values between the pixel and the LLR domain.

B. Transition function $g(\cdot)$ and $h(\cdot)$

LLR is defined by $l = \ln(\frac{p(0|\mathcal{C})}{p(1|\mathcal{C})})$ where $p(b|\mathcal{C})$, $b \in \{0, 1\}$, is the conditional probability of bit b given constraint $\mathcal{C}$. Constraint $\mathcal{C}$ indicates available constraint at the decoder side, including quantization bin $[B_L, B_U - 1]$ determined by previously decoded bit-planes and conditional probability $f(x|y)$ which represents the statistical correlation between side-information y and source x. Generally, $f(x|y)$ is modeled by a discrete Laplacian distribution $f(x|y) = \frac{\alpha}{2} e^{-\alpha|x-y|}$ [1] because both x and $\widehat{x}$ are integers. Therefore, $p(0|\mathcal{C})$ and $p(1|\mathcal{C})$ can be obtained by summing $f(x|y)$ over region $[B_L, B_M - 1]$ and $[B_M, B_U - 1]$ respectively.

In what follows, we will derive the transition function between LLR and side-information y for LLR $l \leq 0$. For $l > 0$, an intermediate estimate $\widetilde{y}$ can be estimated by $-l$ and use $B_U - 1 - \widetilde{y} + B_L$ as the estimate of y. For $l \leq 0$, $p(0|\mathcal{C})$ and $p(1|\mathcal{C})$ can be expressed by

$$p(1|\mathcal{C}) = \sum_{x=B_M}^{B_U - 1} f(x|y) = \frac{\alpha}{2} e^{\alpha y} e^{\alpha} \frac{e^{-\alpha B_U} - e^{-\alpha B_M}}{1 - e^{\alpha}}. \tag{6}$$

$$\begin{aligned} p(0|\mathcal{C}) &= \sum_{x=B_L}^{B_M - 1} f(x|y) \\ &= \frac{\alpha}{2} \frac{e^{\alpha(B_L - y)} - 1 + e^{\alpha} e^{\alpha(y - B_M)} - e^{\alpha}}{1 - e^{\alpha}}. \end{aligned} \tag{7}$$

Let $A = e^{\alpha B_L}$, $B = 1 + e^{\alpha}$, $C = e^{\alpha} e^{-\alpha B_M}$, $D = e^{\alpha}(e^{-\alpha B_U} - e^{-\alpha B_M})$, $Y = e^{\alpha y}$, we have

$$l = g(y) = \ln \frac{p(0|\mathcal{C})}{p(1|\mathcal{C})} = \ln \left(\frac{A/Y - B + CY}{DY} \right), \tag{8}$$

Let $r = e^l$ and solve eq. 8, we have

$$\widehat{y} = h(l) = \frac{1}{\alpha} \ln(Y) = \frac{1}{\alpha} \ln \left(\frac{B + \sqrt{B^2 - 4(C - rD)A}}{2(C - rD)} \right). \tag{9}$$

With the derived function $h(\cdot)$ and $g(\cdot)$, the MRF optimization can be realized with update equations (2) and (5).

C. Target gradient $\nabla_{i,j}$

Since target gradient $\nabla_{i,j}$ is unavailable at the DVC decoder side without accessing to the original WZ frame, we resort to the temporal consistency property of video sequences to approximate $\nabla_{i,j}$ by gradient in adjacent frame after compensating the relative motion. Motion field estimation algorithm

TABLE I

BIT-RATE SAVING OF THE PROPOSED ALGORITHM.

Bit-plane	Decoding algorithms	Test sequences			
		Bus	Crew	Football	Foreman
BP1	LDPCA	14596.1	7222.5	32218.6	12526.9
	The proposed	12141.4	6501.0	24791.4	10227.2
	Bit saving(%)	16.82%	9.99%	23.05%	18.36%
BP2	LDPCA	29468.4	22781.5	25275.5	10946.9
	The proposed	26331.4	20548.1	20032.2	9159.5
	Bit saving(%)	10.65%	9.80%	20.74%	16.33%
BP3	LDPCA	41952.0	27647.0	60843.2	22217.7
	The proposed	39796.6	24839.3	53858.5	18415.4
	Bit saving(%)	5.14%	10.16%	11.48%	17.11%
BP4	LDPCA	62715.4	48731.1	91465.0	40859.2
	The proposed	62861.4	44978.6	86420.7	37838.2
	Bit saving(%)	-0.23%	7.70%	5.51%	7.39%

[5] is used to estimate this relative motion. It should be noted that although the derived motion field couldn't produce better side-information, it could avoid the block effect in the side-information generated by block-based motion compensation, and help to exploit the smoothness constraint and gradient directions.

V. EXPERIMENTS

In experiments, four CIF format ($352 \times 288@15$Hz) sequences are tested: *Bus*, *Crew*, *Football* and *Foreman*. The video frame encoding structure is "I-WZ-I-WZ-I-...", and the side-information is generated by bidirectional motion compensation [4] with searching range 16 and searching block size 8×8. LDPCA algorithm with block length 6336 and maximum iteration 200 is used as SW codec in both the proposed and conventional DVC schemes. In the proposed scheme, two BP iterations are performed as MRF optimization after each 10 iterations of LDPCA decoding to impose spatial regularity constraint to the LLR information.

Table I presents the average bit number required to encode different bit-planes. Results show that the proposed SW decoding algorithm achieves significant bit-rate saving, especially for sequences with relative simple spatial structures and less side-information quality, *e.g.* "Football" and "Foreman" sequences, because simple spatial structure improves the spatial correlation between neighboring nodes and lower side-information quality induces less reliable LLR and induces higher variance in the LLR from LDPCA iteration. In both cases, the spatial geometric regularity constraint is very helpful to constraint the LLR information and thus improves the performance of the proposed SW decoding algorithm.

Figure 4 shows the rate-distortion performance of the proposed algorithm compared with LDPCA decoding algorithm. Results show that, the bit-rate saving is equivalent to about $0.5 \sim 1$dB performance gain compared to conventional LDPCA decoding algorithm. Since the reconstruction quality of WZ decoder is only related to the decoded bit-plane and side-information quality, both schemes achieve identical reconstruction quality with the same reconstruction algorithm. To further improve the RD performance, MRF-based constraint

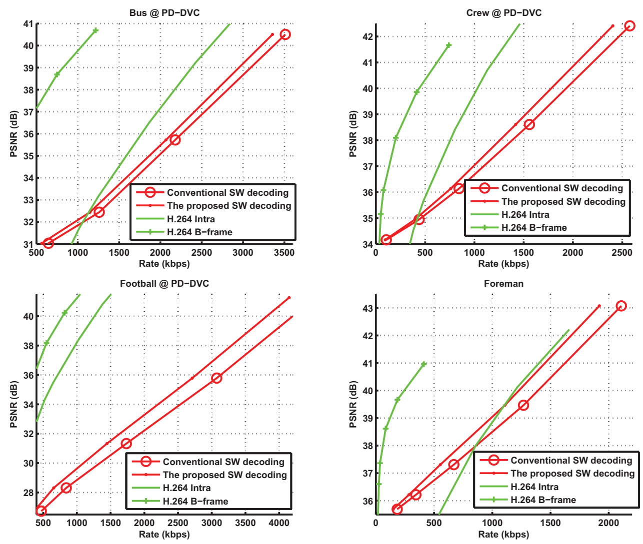


Fig. 4. The rate-distortion performance of PD-DVC schemes with the proposed Slepian-Wolf decoding algorithm.

can also be exploited in the reconstruction process to improve the decoded WZ frame quality [6].

VI. CONCLUSIONS

In this paper, we proposed a novel SW decoding algorithm for PD-DVC scheme by exploiting the spatial geometric regularity constraint of WZ frames. The proposed algorithm models the LLR of a bit-plane as an anisotropic MRF model and solving the problem by belief propagation. Experiment results show about $5\% \sim 23\%$ bit-rate saving. Since the proposed SW decoding algorithm is independent with side-information generation and WZ-frame reconstruction process, these techniques can be combined to further improve the overall performance of DVC applications.

REFERENCES

[1] B. Girod, A. Aaron, S. Rane, and D. Rebollo-Monedero, "Distributed video coding," *Proceedings of the IEEE*, vol. 93, no. 1, pp. 71-83, Jan. 2005.

[2] D. Varodayan, A. Aaron, and B. Girod, "Rate-adaptive codes for distributed source coding," *Signal Processing*, vol. 86, pp. 3123-3130, Nov. 2006.

[3] J. Zhang, H. Li, and C. W. Chen, "Distributed image coding based on integrated Markov modeling and LDPC decoding," *in Proc. International Conference on Multimedia & Expo (ICME)*, Hannover, Germany, Jun. 23-26 2008.

[4] J. Ascenso, C. Brites, and F. Pereira, "Improving frame interpolation with spatial motion smoothing for pixel domain distributed video coding," *in 5th EURASIP Conference on Speech and Image Processing, Multimedia Communications and Services*, Slovak Republic, Jul. 2005.

[5] C. Liu. "Beyond pixels: exploring new representations and applications for motion analysis," Ph.D dissertation, Computer Science and Artificial Intelligence Laboratory (CSAIL), Massachusetts Institute of Technology, Cambridge, MA, May 2009.

[6] Y. Zhang, H. Xiong, Z. He, S. Yu, C. W. Chen, "Reconstruction for distributed video coding: a context-adaptive Markov random field approach," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 21, no. 8, pp. 1100-1114, Aug. 2011.