Applied Mathematics & Information Sciences
*An International Journal*

http://dx.doi.org/10.12785/amis/070447

# Distributed Video Coding Based on Multiple-source Correlation Model

*Linbo Qing*[1,*], *Xiaohai He*[1], *Xianfeng Ou*[1] *and Rui Lv*[2]

[1]School of Electronics and Information Engineering, Sichuan University, Chengdu 610065, China
[2]Chengdu Research Center, Huawei Technologies Co. Ltd., Chengdu 610065, China

Received: 13 Nov. 2012, Revised: 12 Dec. 2012, Accepted: 28 Dec. 2012
Published online: 1 Jul. 2013

**Abstract:** Distributed Video Coding (DVC) is a new video coding paradigm which has the ability to shift coding complexity from encoder to decoder. State-of-art DVC schemes are mainly based on two-source correlation model and only one correlation channel between side information and wyner-ziv frame is considered. In this paper, multiple-source correlation model is introduced into a famous DVC scheme, namely Transform Domain Wyner-Ziv video coder (TDWZ). A soft-input computing algorithm based on multiple-source correlation model is firstly proposed. The correlations among multiple side information are separately exploited. Then a new reconstruction function based on multiple source correlation model is also proposed to improve reconstructed image's quality. Experimental results show that the proposed algorithms can effectively improve the coding performance of systems.

**Keywords:** distributed video coding; multiple-sources correlation; virtual noise channel; transform domain wyner-ziv coder

## 1. Introduction

Distributed Video Coding (DVC) is a new video coding paradigm, which has the ability to shift encoding complexity from encoder to decoder, without sacrificing coding efficiency. Besides DVC's complexity shifting property, it removes the prediction loop in the encoder and is more robust to transmission error than traditional video coding paradigm. DVC is proposed mainly based on two major information theory results: the Slepian-Wolf and Wyner-Ziv theorems [1,2]. Slepian and Wolf have proved that the separate-encoding and joint-decoding of two correlated sources can achieve the same rate as joint-encoding and joint-decoding scheme. Based on two-source Slepian-Wolf theorem, the distributed coding of multiple correlated sources are discussed, similar result has been achieved [3,4]. Liveris [5] has proposed distributed multiple source coding paradigm based on LDPC. Zhu [6] has discussed distributed source coding problem of multiple sources, Distributed Joint Source-Channel Coding (DJSCC) scheme has been proposed and its theoretical limits has been deduced.

In video coding scenario, most of the pioneer works on DVC field are mainly based on DISCOVER DVC codec [7] where two-source Slepian-Wolf coding model is considered. In order to improve the coding efficiency of DVC codec, multiple side information has been considered in existing works. Fernando [8] has proposed multiple side information for Distributed Video Coding. They have generated two streams of side information and adaptively used one of them according to motion intensity. However their system is still based on two-source correlation model. Huang [9] has proposed a distributed video coding scheme based on both Motion Interpolated Side Information (MISI) and Motion Extrapolated Side Information (MESI). They have proposed a weighted soft input combination algorithm, to adaptively using virtual noise statistics for MESI and MISI. The correlation between MISI and MESI has not been considered. Ascenso [10] has also proposed DVC codec with multiple side information. In their work, 5 side information have been generated using different algorithms, then augmented LDPC Graph for simultaneously exploit the multiple virtual channel between DVC coded frame and each side information. However neither has considered the correlations among side information, which is very important when utilizing multiple sources Slepian-Wolf theorem.

The objective of this paper is to improve the coding efficiency by introduce multiple side information into

* Corresponding author e-mail: qing_lb@scu.edu.cn

DVC system, and to exploit multiple-source correlation models among source and side information in distributed way. The correlation statistics among multiple sources, including DVC coded frames, side information, are exploited to simultaneously improve soft-input computation of channel decoder and reconstruction module.

The paper is organized as follows. Section 2 reviews the theorems of distributed source coding based on multiple sources. The proposed distributed video coding scheme based on multi-source correlation statistics are discussed in Section 3. In this Section, the soft-input computing algorithms based on multi-source correlation model is firstly proposed. Then the weighted side information generation algorithm based on multi-source correlation model and modified reconstruction module are presented. Experimental results are given in Section 4 and conclusion is made in Section 5.

## 2. Distributed Source Coding of Multiple Sources

Consider $n$ correlated sources of data $X_i \in GF(p), i \in \{1,2,...,n\}$ which are separately encoded and jointly decoded, with code rate of $R_i, i \in \{1,2,...,n\}$. According to Wolf's result, the achievable rate region is:

$$\sum_{i \in S} R_i \geq H(X(S)|X(S^C)) \tag{1}$$

where $S \subseteq \{1,2,...,n\}$, $S^C$ is the complementary set of $S$, $X(M) = \{X_i, i \in M\}$.

Consider three memoryless correlated binary sources $X_1$, $X_2$ and $X_3$, where the virtual channel between them can be modeled as a Binary Symmetric Channel (BSC) cross-over probability $p_{ij}$, where

$$p_{ij} = p[X_i \neq X_j|X_j] \tag{2}$$

Suppose $X_1$ is coded independently, then according to Slepian-wolf theorem, the achievable rate for $X_2$ and $X_3$ can be given by

$$\begin{array}{l} R_2 \geq H(X_2|X_3,X_1) \\ R_3 \geq H(X_3|X_2,X_1) \\ R_2 + R_3 \geq H(X_2,X_3|X_1) \end{array} \tag{3}$$

Given $X_2$ and $X_3$ and their correlation, the cross-over probability of $X_3$ can be acquired by [6]:

$$\begin{array}{l} p[X_3 \neq X_2|X_2 = X_1, X_1] = \frac{(p_{23}+p_{13}-p_{12})}{(1-p_{12})} \\ p[X_3 \neq X_2|X_2 \neq X_1, X_1] = \frac{(p_{23}-p_{13}+p_{12})}{2p_{12}} \end{array} \tag{4}$$

## 3. Distributed Video Coding Based on Multiplesource Channel Model

### 3.1. *State-of art DISCOVER codec*

Our proposed distributed video coding is mainly based on most famous DVC framework, i.e., the DISCOVER Transform Domain Wyner-Ziv (TDWZ) video codec [7], shown in Figure 1. At the encoder, Video frames are first split into key frames and so-called Wyner-Ziv (WZ) frames, usually one WZ frame after one key frame. Key frames are intra-frame coded with a conventional video coder. A WZ frame is firstly partitioned into non-overlapped 4*4 blocks and DCT transformed. Then the quantized transform coefficients are extracted into bitplanes and LDPCA coded. For each bitplane, the syndrome is stored in a buffer and partially transmitted to decoder upon request.

At the decoder, the key frames are decoded and used to generate SI by frame interpolation for the WZ frames between two key frames. SI is then DCT transformed and extracted into bitplanes. Together with the estimated virtual channel model between WZ frame and its SI, the soft-input probability $P$ for each bitplane is estimated. Together with the soft-input $P$ and partially received syndromes, the LDPC decoding is implemented. If the LDPC decoding is not successful, more syndromes are required through feedback channel. This process is repeated until successfully decoding. After all the transform coefficients' reconstruction, the final decoding of WZ frame is accomplished under the help of SI, using reconstruction function. By exploiting the correlation between the SI and the original WZ frame, the DVC coder only requests part of the syndromes bits for decoding, thus achieves data compression.

A more effective generation of soft-input information for LDPC decoder will resulting less required syndromes. A more sophisticated reconstruction module will leading to better decoding frame quality. These two factors will thus increase the coding efficiency of the video coder. A detailed coding process can be found in [7].

### 3.2. *Proposed Soft-input Computing Algorithm based on Multiple Correlation Model*

Before soft-input computing, side information of current WZ frame has been generated. One of the most famous ways is to using Motion Compensated Interpolation (MCI). As shown in Figure 2. The forward and backward motion vector $MV_f$ and $MV_b$ are firstly estimated. Then forward motion compensated frame $SI_f$ and backward motion compensated frame $SI_b$ are generated. Side information SI is finally generated by weighted average of $SI_f$ and $SI_b$. Thus when GOP=2, $SI = 0.5*(SI_b + SI_f)$. At the process of soft-input computing, the virtual correlation channel between WZ frame and $SI$ should be

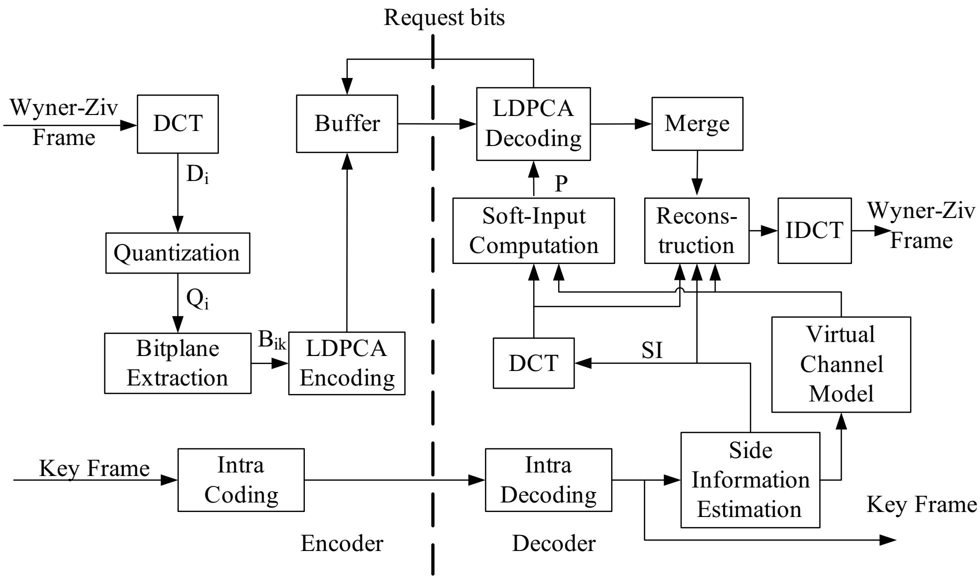


**Figure 1:** Transform domain Wyner-Ziv video coding.

firstly estimated. Several correlation noise modeling algorithm have been given in [11]. Then the soft-input of each bitplanes can be estimated. Apparently this kind of wyner-ziv video coding is mainly based on two-source DSC model.

Let us model WZ frame, $SI_f$ and $SI_b$ as three correlated sources, then multiple source correlation model should be exploited here. Since $SI_f$ and $SI_b$ are estimated using MCI, the correlations among WZ frame, $SI_f$ and $SI_b$ can modeled as Laplacian distribution pair wisely. The correlations between WZ frame, $SI_f$ and $SI_b$ can be estimated using algorithms in [11]. The correlation between $SI_f$ and $SI_b$ can be acquired straight forwardly since they are both available at decoder.

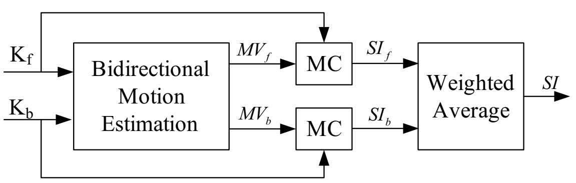


**Figure 2:** Motion Compensated Interpolation.

Let $X_i^j, SI_{bi}^j, SI_{fi}^j, i \in \{0,1,...,15\}, j \in \{0,1,...,N-1\}$ be the $j$th bitplane of the $i$th quantized coefficient for WZ frame, $SI_f$ and $SI_b$ respectively, where $N$ is the quantization bit number. The cross-probabilities among $X_i^j$, $SI_{bi}^j$, $SI_{fi}^j$ are defined as follows:

$$\begin{aligned} p[X_i^j \neq SI_{bi}^j | SI_{bi}^j] &= p_{xb} \\ p[X_i^j \neq SI_{fi}^j | SI_{fi}^j] &= p_{xf} \\ p[SI_{bi}^j \neq SI_{fi}^j | SI_{fi}^j] &= p_{bf} \end{aligned} \quad (5)$$

Given the correlation distribution among $X_i^j, SI_{bi}^j, SI_{fi}^j$, the cross-probabilities in (5) can be easily derived. According to (4) the soft-input can then computed by following formulas regarding the equality of $SI_{bi}^j$ and $SI_{fi}^j$ :

$$\begin{aligned} p[X_i^j \neq SI_{fi}^j | SI_{fi}^j = SI_{bi}^j, SI_{bi}^j] &= \frac{(p_{xf}+p_{xb}-p_{bf})}{(1-p_{bf})} \\ p[X_i^j \neq SI_{fi}^j | SI_{fi}^j \neq SI_{bi}^j, SI_{bi}^j] &= \frac{(p_{xf}-p_{xb}+p_{bf})}{2p_{bf}} \end{aligned} \quad (6)$$

## 3.3. Proposed Reconstruction Function based on Multiple Correlation Model

Since quantization has been introduce into DVC to increase coding efficiency, the channel decoded quantized symbols can only restrict the output into a specific range, i.e. $[q_n, q_{n+1})$, where the upper $N$ bitplanes of $q_n$ and $q_{n+1}$ is the same with channel decoded out, lower bitplanes are all '0's and '1's respectively. Then the basic reconstruction function is given by [7]:

$$\hat{X} = \begin{cases} q_i, & SI < q_i \\ SI, & q_i \le SI < q_{i+1} \\ q_{i+1}, & SI \ge q_{i+1} \end{cases} \quad (7)$$

Good quality of SI should firstly guaranteed for better reconstruction. As described in Section 3B, there are two side information available, $SI_f$ and $SI_b$. Due to dynamic property of motion field, the effectiveness of $SI_f$ and $SI_b$ are not the same. One the other hand, since the virtual noise is actually the difference between side information and WZ frame, so its distribution reflects effectiveness of side information. So we proposed a weighted side information algorithm for reconstruction module in DVC codec.

In the first step, the $SI_f$ and $SI_b$ are restricted into decoded quantized bin $[q_n, q_{n+1})$ by (7), to generate $\overline{SI_f}$ and$\overline{SI_b}$. Let $f_{X|SI_i}(X), i \in \{f, b\}$ represent the virtual noise distribution between WZ frames, $SI_f$ and $SI_b$, the refined $SI$ can be computed considering virtual noise distribution by

$$SI = (w_f * \overline{SI_f} + w_b * \overline{SI_b})/(w_f + w_b) \tag{8}$$

where

$$w_i = \int_{q_i}^{q_{i+1}} f_{X|SI_i}(X), i \in \{f, b\} \tag{9}$$

According to minimum distortion quantization theorem [12, 13, 14, 15], minimum quantization noise can be achieved when de-quantization value is at the center of the distribution. So in our proposed reconstruction function, the center of the noise distribution has been considered, and (7) is improved as follows:

$$\hat{X} = \begin{cases} q_i + \frac{\int_{q_i}^{q_{i+1}} (X - q_i) f_{X|SI}(X)}{\int_{q_i}^{q_{i+1}} f_{X|SI}(X)}, & SI < q_i \\ SI + \frac{\int_{q_i}^{q_{i+1}} X f_{X|SI}(X)}{\int_{q_i}^{q_{i+1}} f_{X|SI}(X)}, & q_i \le SI < q_{i+1} \\ q_{i+1} + \frac{\int_{q_i}^{q_{i+1}} (X - q_{i+1}) f_{X|SI}(X)}{\int_{q_i}^{q_{i+1}} f_{X|SI}(X)}, & SI \ge q_{i+1} \end{cases} \tag{10}$$

where $f_{X|SI}(X)$ is the distribution of the correlation noise which can be estimated using existing way[11].

## 4. Experimental Results

In order to evaluate the effectiveness of proposed algorithms, the wyner-ziv video coder has been implemented with proposed 'soft-input computation' and 'reconstruction' algorithms. In order to make fair comparisons, the test conditions used in our experiments are the DISCOVER project test conditions [7], which are widely used in literatures. The sample sequences used are *Foreman* and *Coastguard* with the first 200 frames coded at 30 frames per second; the GOP size is set to 2. Four quantization matrixes have been chosen to acquire different Rate Distortion (RD) point. Through all the tests, only the luminance component is coded for simplicity, and the rate-distortion performance of WZ frames at 15 frame per second is evaluated. The rate is measured by bit per second (bps), and the distortion is measured by Peak Signal-to-Noise Ratio (PSNR). For fair comparison, two experiments have been made for each sequence: (1) TDWZ codec following DISCOVER framework [7]; (2) TDWZ codec with proposed algorithms. The resulting RD performances are given in Figure 3 and Figure 4. In proposed TDWZ codec, the virtual noise distribution between WZ frame, $SI_f$ and $SI_b$ are estimated using offline algorithm in pixel level from [11].

As shown in Figure 3 and Figure 4, the performance of proposed TDWZ Codec is much better than DISCOVER TDWZ codec. For both test sequences, 0.5dB˜1.5dB gains has been achieved in proposed system. Since the quantization matrixs are kept the same for each experiment, the adjacent RD point of each experiment in Figure 3 and Figure 4. can represent the comparison between DISCOVER codec and proposed codec. More specifically, the gains for each quantization matrix are mainly due to two factors: (1) Less rate; (2) Better PSNR. Note that on one hand, the proposed algorithm for soft-input computing in Section 3B has only decreased the coding rate. On the other hand, the proposed algorithm for reconstruction in Section 3C has only increase image quality. So the experiments can prove that both of the proposed algorithms are effective.

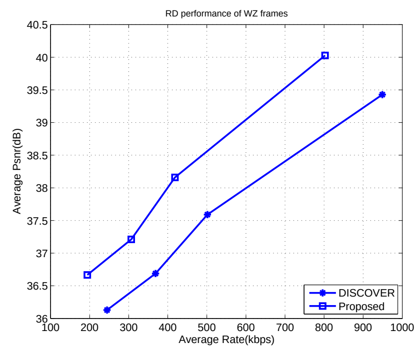


**Figure 3:** Foreman.

## 5. Conclusion

In this paper, the multiple-source correlation model has been considered in distributed video coding system, with the objective to improve the RD performance, without any change in encoder. The multiple-source correlation model is firstly considered for better soft-input, trying to reduce the number of syndrome for an aimed quality. Then each correlation model is utilized again for better reconstruction. Compared with state-of-art distributed video coding system based on two-source correlation

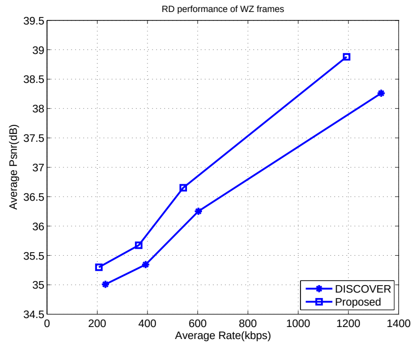


**Figure 4:** Coastguard.

model, the proposed algorithms can significantly improve the RD performance of the system.

## Acknowledgement

This project is supported by Doctor Station Foundation of National Ministry of Education of China (20110181120009), National Science Foundation of China (NSFC 61201388), and National Science Foundation of China (NSFC,61071161).

The authors are grateful to the anonymous referee for a careful checking of the details and for helpful comments that improved this paper.

## References

[1] D. Slepian and J. Wolf, IEEE Transactions on Information Theory **19**, 471-480 (1973).

[2] A. Wyner and J. Ziv, IEEE Transactions on Information Theory **22**, 1-10 (1976).

[3] J. K. Wolf, Proceedings of the 5th Colloquium Microwave Communication, 287-295 (1973).

[4] T. M. Cover, J. A. Thomas, Elements of Information Theory. StateplaceNew York: John Wiley & Sons, Inc., 184-188 (1991).

[5] A. Liveris, C. Lan, K. Narayanan, et al, Proceedings of the International Symposium on Turbo Codes and Related Topics, 63-66 (2003).

[6] X. Zhu, L. Zhang, Y. Liu, Proceedings of the 4th International Conference on Communications and Networking, Aug. 26-28; Beijing, China (2009).

[7] X. Artigas, J. Ascenso, M. Dalai, et al, Proceedings of Picture Coding Symposium, Lisbon, Portugal (2007).

[8] W.A.C. Fernando, H.K. Arachchi, W.A.R.J. Weerakkody, Electronics Letters **42**, 1447-1449 (2006).

[9] X. Huang, C. Brites, J. Ascenso, F. Pereira, S. Forchhammer, Proceedings of the 27th conference on Picture Coding Symposium, May 6-8; Illinois, USA(2009).

[10] J. Ascenso, C.Brites, F. Pereira, Proceeding of IEEE 13th International workshop on Multimedia Signal ProcessingOct. 17-19; Hangzhou, China(2011).

[11] C. Brites, F. Pereira, IEEE Trans. on Circuits Syst. Video Technol **18**, 1177-1190 (2008).

[12] S. Lloyd, IEEE Transactions on Information Theory **28**, 129-137 (1982).

[13] J. Max, IEEE Transactions on Information Theory **6**, 7-12 (1960).

[14] Y. Zhang,Y.-F. Pu,J.-R. Hu and J.-L. Zhou, Applied Math & Information Sci. **6**, 299-306(2012).

[15] Z.Ji, W.Huang, G.Kenyon, Proceedings of IEEE International Joint Conference on Neural Networks. 2844-2851 (2011).

**Linbo Qing** received his Ph.D. degree in communication and information system from the Sichuan University, Chengdu, China, in 2008. From Feb. to Aug. 2011, he was a visiting researcher at the Agder Mobility Laboratory, University of Agder, Norway where he worked on data aggregation and security in mobile communications. From Sep. 2011 to Feb. 2012, he was a visiting researcher at the Internet Media Group, Politecnico di Torino, Turin, Italy, where he worked on distributed video coding and transmission. He is currently an assistant professor in the College of Electronics and Information Engineering at the Sichuan University. He co-authored over 30 scientific publications in image processing related topics. His main research interests include image processing, video coding & transmission, and information theory.

**Xiaohai He** received his Ph.D. degree in Biomedical Engineering from the Sichuan University, Chengdu, China. He is a currently a professor in College of Electronics and Information Engineering at the Sichuan University. He is an Editor of both the Journal of Information and Electronic Engineering and the Journal of Data Acquisition & Processing. His research interests include image processing, pattern recognition, and image communication. Dr. He is a senior member of the Chinese Institute of Electronics.

**Xianfeng Ou** received the B.S. degree in Electronic Information Science and Technology and M.S. degree in Communication and Information System from Xinjiang University, Xinjiang, China, in 2006 and 2009, respectively. He is currently working toward the Ph.D. degree in the College of Electronics and Information Engineering of the Sichuan University, Chengdu, China. His research interests include image and video coding technologies.

**Rui Lv** He received Ph.D degree in communication and information system from the Sichuan University, Chengdu, China, in 2009. He currently work in Chengdu Research Center, Huawei Technologies Co. Ltd. His current research interests are video processing and compression, video communication, wireless networking.