Wyner-Ziv Video Coding for Highway Traffic Surveillance Using LDPC Codes

Linlong Guo, Bowei Shan, and Yong Fang(✉)

School of Information Engineering, Chang'an University, Xi'an, People's Republic of China
{2018124070,bwshan,fy}@chd.edu.cn

Abstract. This paper presents a 2D Q-ary Sliding-Window Belief Propagation (2DQSWBP) algorithm to decode Low-density Parity-check (LDPC) code which is utilized to compress highway traffic surveillance video under Wyner-Ziv video framework. This framework is beneficial for camera device with limited memory and computing ability. The differences of successive frames is modeled as Truncated Discrete Laplace (TDL) distribution. The experimental result shows the 2DQSWBP outperforms the Q-ary Belief Propagation (QBP) algorithm in both Bit-Error-Rate and computing time.

Keywords: 2DQSWBP algorithm · LDPC code · QBP algorithm · Wyner-Ziv video coding

1 Introduction

Recent advances in the Fifth Generation (5G) telecommunication and Internet of Things as a Service (IoTaaS) enables real-time video surveillance of highway traffic using mobile camera device. In this scenario, the video codes at the camera usually has limited memory and computing ability, therefore a low-complexity encoder is always required.

The successive frames in a video sequence typically are very similar, and current video coding standards, such as MPEG or ITU-T H.26x, take advantage of this similarity to make compression efficient. The implementation of traditional video codes requires encoder 5 to 10 times more computational complex than the decoder, which is well suited for streaming video-on-demand system instead of surveillance system.

The Wyner-Ziv video coding [1,2] (also known as distributed video coding) is an asymmetric compression scheme which encodes individual frame independently while decodes them conditionally. The theoretical foundation of the Wyner-Ziv video coding is the Slepian-Wolf theorem [3] on distributed source coding and the Wyner-Ziv [4] theorem for lossy source coding. These theorems state an astonishing result that for two statistically dependent discrete signals,

B. Li et al. (Eds.): IoTaaS 2019, LNICST 316, pp. 639–645, 2020.
https://doi.org/10.1007/978-3-030-44751-9_54

X and Y, with entropies $H(X)$ and $H(Y)$, even if the encoders are independent, the achievable rate region for a joint decoder is

$$\begin{aligned} &R_X \geq H(X|Y), R_Y \geq H(Y|X), \\ &R_X + R_Y \geq H(X,Y). \end{aligned} \tag{1}$$

In this paper, the Wyner-Ziv video coding is implemented by Low-density Parity-check (LDPC) codes for highway traffic surveillance video compression. The difference between the successive frames are modeled as Truncated Discreate Laplacian (TDL) distribution. The 2D Q-ary Sliding-Windows Belief Propagation (2DQSWBP) algorithm is proposed to decode LDPC codes. The experiment result shows that 2DQSWBP outperforms Belief Propagation (BP) algorithm for decoding the traffic video.

2 Wyner-Ziv Video Coding

2.1 Framework

In the Wyner-Ziv video coding, the encoder is implemented by Pixel-Domain and Transform-Domain (PDTD) scheme. The framework of PDTD is illustrated in Fig. 1

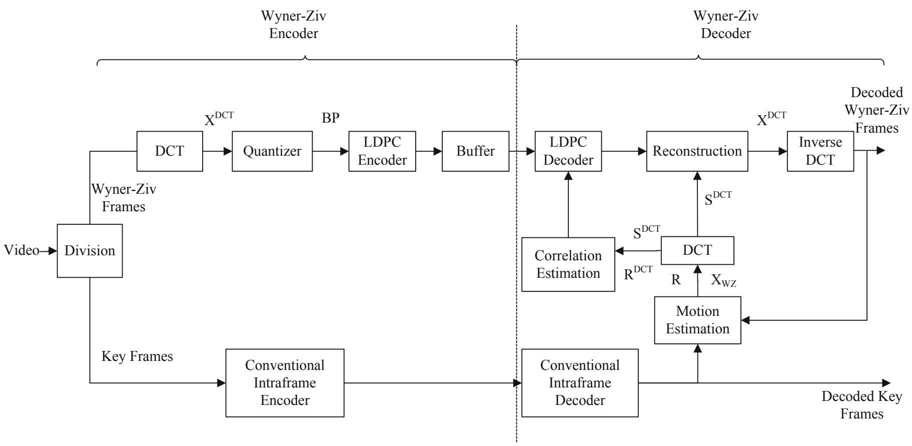


Fig. 1. Framework of Wyner-Ziv video coding.

PDTD divides video frames into two parts: the first frame is named as Key Frame, and the rest are named as Wyner-Ziv Frames. Key Frame is encoded and decoded using a conventional intraframe codec. For Wyner-Ziv Frames, A block-wise Discrete Cosine Transformation (DCT) is performed to obtain the transform coefficients X^{DCT} which are independently quantized and grouped into coefficient bands. These bands are compressed by LDPC encoder. At the decoder, the side information frame X_{WZ} and motion compensation residual frame R are generated by previously reconstructed frames with the help of motion compensation. Then a block-wise DCT is applied to them, resulting in side information S^{DCT} and residual frame R^{DCT}. The correlation between X^{DCT} and S^{DCT} is modeled as a Laplacian distribution. LDPC decoder reconstructs the coefficient bands with the corresponding side information and performs a inverse DCT to generate the reconstructed Wyner-Ziv frames.

Above coding paradigm has many advantages. First, it makes it possible to shift the computational complexity from encoder to decoder. This is beneficial for many wireless camera devices which normally have limited computing ability and battery life. Second, since there is no side information at the encoder, the frame error will not be propagated and the recovered videos appear to be more robust.

2.2 2DQSWBP Algorithm

In 2013, Fang [5] proposed a 2D SWBP algorithm for binary LDPC codes, and a Q-ary SWBP [6] in 1D form is introduced in 2019. As what we have known, the 2D Q-ary SWBP algorithm has not been presented. The video is consisted of successive 2D frames. Each frame has many pixels and each pixel is represented by a q bit value which varies from 0 to $Q = 2^q - 1$. In this paper, we will use 2DQSWBP to decode Wyner-Ziv video, and the flowchart of 2DQSWBP is illustrated in Fig. 2

For 2D paradigm, Let

$$x^{n,n} \triangleq \begin{pmatrix} x_{0,0} & \cdots & x_{0,n-1} \\ \vdots & \ddots & \vdots \\ x_{n-1,0} & \cdots & x_{n-1,n-1} \end{pmatrix}, \tag{2}$$

$y^{n,n}$ be two $n \times n$ 2D sources, where $x_{i,j} \in [0 : Q)$. Let the correlation between them be $z^{n,n} = x^{n,n} - y^{n,n}$.

$$p_{z^{n,n}}(z^{n,n}) = \prod_{i=0}^{n-1} \prod_{j=0}^{n-1} p_{z_{i,j}}(z_{i,j}), \tag{3}$$

where $p_{z_{i,j}}(z_{i,j})$ obeys Truncated Discreate Laplacian (TDL) distribution

$$p_{z_{i,j}}(a|y_{i,j}, \sigma_{i,j}^2) = \frac{\exp\left(-\sqrt{\frac{2}{\sigma_{i,j}^2}}|a|\right)}{\sum_{a'=-y_{i,j}}^{M-1-y_{i,j}} \exp\left(-\sqrt{\frac{2}{\sigma_{i,j}^2}}|a'|\right)}. \tag{4}$$

We call $\sigma_{i,j}^2$ the local nominal variance of $Z_{i,j}$ and define the global nominal variance of $z^{n,n}$ as

$$\sigma^2 \triangleq \frac{1}{n \times n} \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} \sigma_{i,j}^2. \tag{5}$$

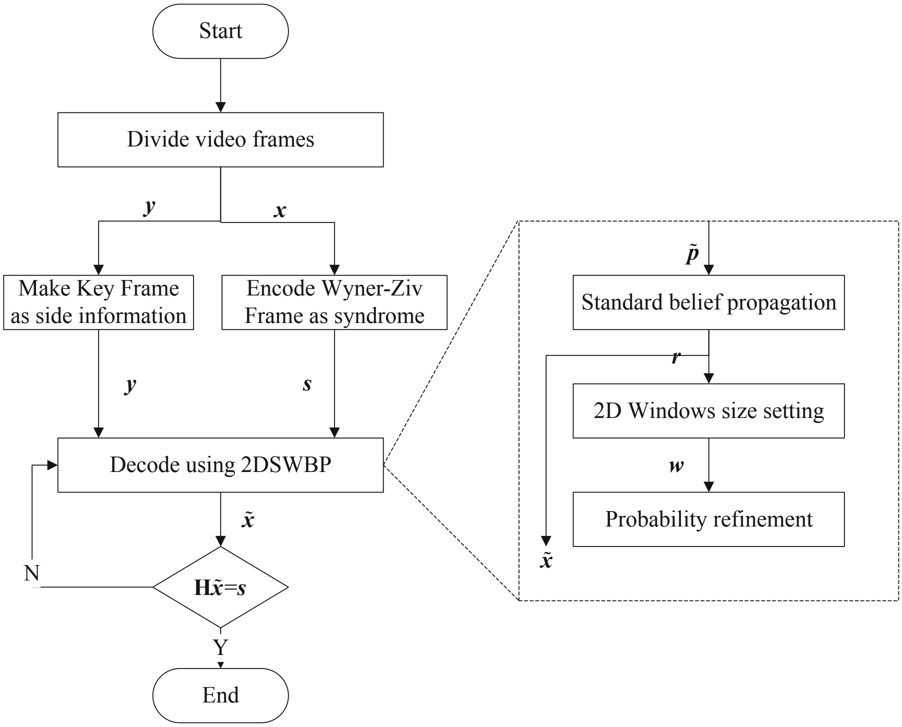


Fig. 2. Flowchart of 2DQSWBP.

At the encoder, $x^{n,n}$ is first vectorized into an Q-ary temporary vector $v^{n,n}$. Then $v^{n,n}$ is performed a matrix-vector multiplication over the finite field GF(Q) to compress source to syndrome s^m:

$$s^m = \mathbf{H}v^{n,n}, \tag{6}$$

where $\mathbf{H}$ is an $m \times (n \times n)$ sparse parity-check matrix.

At the decoder, standard Q-ary BP algorithm (QBP) utilize side information $y^{n,n}$ to recover $x^{n,n}$, which has been introduced in [6]. After each QBP iteration, the local nominal variances of symbol nodes are reestimated according to their overall probability mass functions (pmf). Let $r_{i,j}(a)$ for $a \in [0 : Q)]$ be the overall pmf of symbol node $x_{i,j}$. The Expected Squared Euclidean Distance (ESED) between each source symbol and its corresponding side information symbol is computed by

$$d_{i,j} = \sum_{a=0}^{Q-1} r_{i,j}(a)|a - y_{i,j}|^2. \tag{7}$$

The optimal half window size u^* is determined by

$$u^* = \arg\max_{u} d_{i,j}(u). \tag{8}$$

Then, $\sigma^2_{i,j}$ is computed by averaging $d_{r',c'}$ over a $(2u^*+1) \times (2u^*+1)$ 2D window around $x_{i,j}$, *i.e.*,

$$\hat{\sigma}^2_{i,j}(u^*) = \frac{\sum_{i',j' \in \mathcal{W}_{i,j} \setminus (i,j)} d_{i,j}}{|\mathcal{W}_{i,j}| - 1}, \tag{9}$$

where

$$\mathcal{W}_{i,j} \triangleq \{(i',j') : i' \in \mathcal{W}_i \cap j' \in \mathcal{W}_j\}, \tag{10}$$

where

$$\mathcal{W}_i \triangleq [\max(0, i-u) : \min(i+u, n-1)], \tag{11}$$

and $|\cdot|$ denotes set cardinality. Finally, $r_{i,j}(a)$ is updated with the refined $\sigma^2_{i,j}(u^*)$

3 Experimental Results

A highway traffic surveillance video is used to perform our tests. Each frame in the video has 128×128 pixels. We construct a regular LDPC code with codeword length 16384, and information bit number 8192. Then the rate is 1/2. Our experimental platform is configured with Intel(R) Core(TM) i7-7700 CPU 3.60 GHz(4 Cores) main frequency, 8 GB memory, and 64-bit Windows-10 operation system. All above algorithms are developed in MATLAB 2017a environment.

The aim of the experiment is to compare the performance between the 2DQSWPB and QBP algorithm. The alphabet cardinality Q is fixed to $2^8 = 256$. The source size $n \times n$ is fixed to 128×128. The LDPC code is used to compress the source. The local nominal variance of $z^{n,n}$ varies according to $\sigma^2_{i,j} = 2b^2$. Five value of b are tested, *i.e.*, $b \in \{1.2, 1.4, 1.5, 2, 2.5\}$.

Two metrics, *i.e.*, Bit-Error-Rate (BER) and running time, are used to evaluate the performance of proposed algorithms. The results are plotted in Fig. 3, from which we find that 2DQSWBP outperforms QBP with less computing time (Fig. 4).

From the results we can conclude that:

We select a frame from the highway traffic surveillance video and depict it for $b = 1.4$ in Fig. 5(a). The recover frame by 2DQSWBP is depicted in Fig. 5(b). We find 2DDSWBP can successfully recover the video and when $b > 1.4$, QBP cannot recover the video.

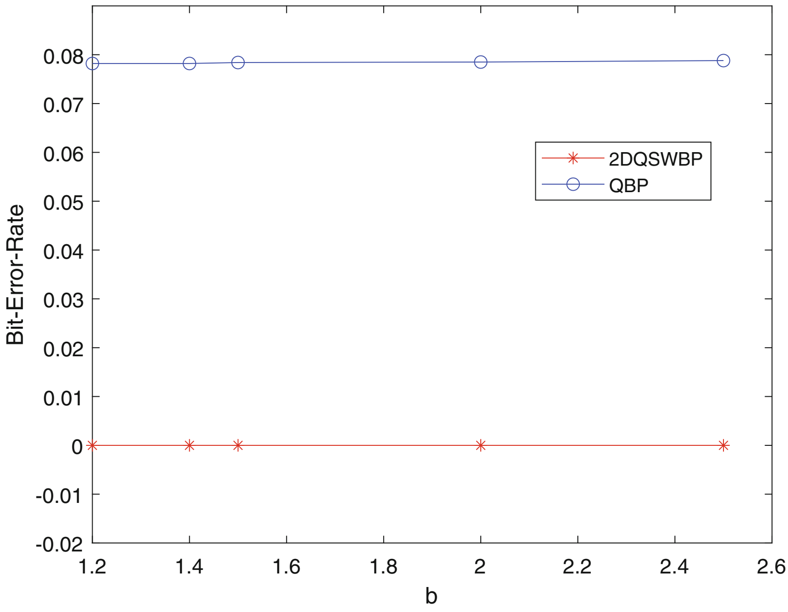


Fig. 3. Bit-Error-Rate under two algorithms.

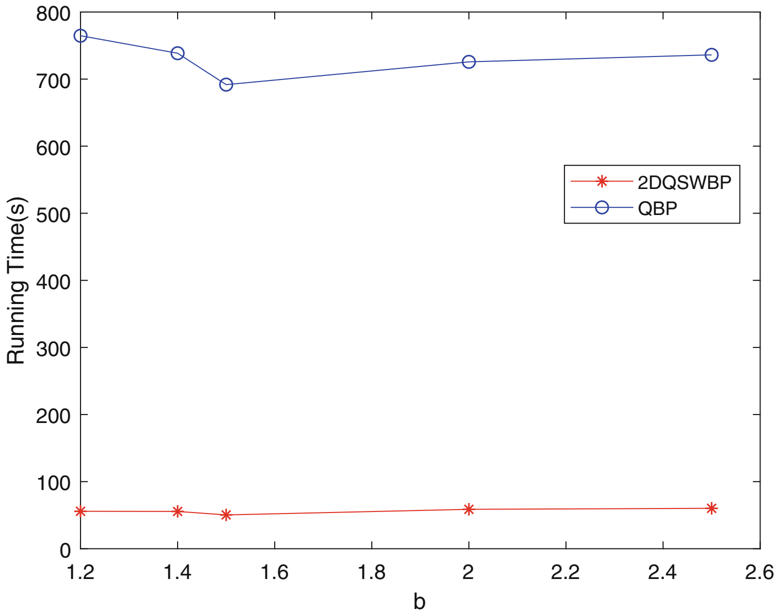
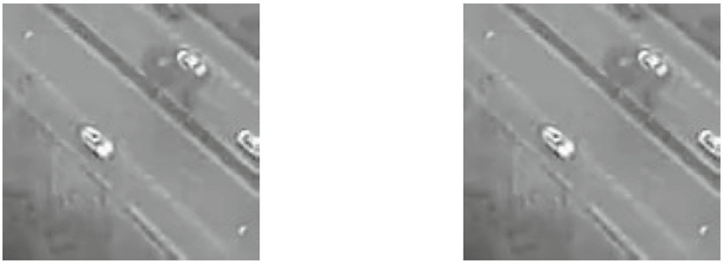


Fig. 4. Running time under two algorithms.



(a) (b)

Fig. 5. A frame from the highway traffic surveillance video (a) Orignal Frame (b) Recovered Frame by 2DQSWBP

4 Conclusion

To deal with highway traffic surveillance video, the innovation of this paper is using the LDPC code to compress video frame under Wyner-Ziv video coding framework and proposing a 2DQSWBP algorithm to decode corresponding LDPC codes. The correlation between the successive frames are modeled as TDL distribution. Experimental results show that this algorithm can successfully recover the orignal video and 2DQSWBP outperforms QBP algorithm in both BER and computing time.

References

1. Aaron, A., Zhang, R. Girod, B.: Wyner-Ziv coding of motion video. In: Conference Record of the Thirty-Sixth Asilomar Conference on Signals, Systems and Computers, Pacific Grove, CA, USA, vol. 1, pp. 240–244 (2002). https://doi.org/10.1109/ACSSC.2002.1197184
2. Girod, B., Aaron, A.M., Rane, S., Rebollo-Monedero, D.: Distributed video coding. Proc. IEEE **93**(1), 71–83 (2005). https://doi.org/10.1109/jproc.2004.839619
3. Slepian, D., Wolf, J.K.: Noiseless coding of correlated information sources. IEEE Trans. Inf. Theory **19**(4), 471–480 (1973). https://doi.org/10.1109/TIT.1973.1055037
4. Wyner, A.D., Ziv, J.: The rate-distortion function for source coding with side information at the decoder. IEEE Trans. Inf. Theory **22**(1), 1–10 (1976). https://doi.org/10.1109/TIT.1976.1055508
5. Fang, Y.: Asymmetric Slepian-Wolf coding of nonstationarily-correlated M-ary sources with sliding-window belief propagation. IEEE Trans. Commun. **61**(12), 5114–5124 (2013). https://doi.org/10.1109/TCOMM.2013.111313.130230
6. Fang, Y., Yang, Y., Shan, B., Stankovic, V.: Joint source-channel estimation via sliding-window belief propagation. IEEE Trans. Commun. (2019, submitted)