Distributed Video Coding Based on Constrained Rate Adaptive Low Density Parity Check Codes

Rong Ke Liu*[a], Guo Gang Hua[b], Chang Wen Chen[b]
[a]Dept. of EE, Beijing University of Aeronautics and Astronautics, Beijing, China 100083;
[ab]Dept. of ECE, Florida Institute of Technology, Melbourne, FL ,USA 32901

ABSTRACT

Distributed video coding is an emerging area of research for digital video driven partly by the widespread use of video acquisition devices by the consumers. In this paper, we present a distributed video coding scheme based on zero motion identification at the decoder and constrained rate adaptive low density parity check (LDPC) codes. Zero-motion-block identification mechanism is introduced at the decoder, which takes the characters of video sequence into account. The constrained error control decoder can use the bits in the zero motion blocks as a constraint to achieve a better decoding performance and further improve the overall video compression efficiency. It is only at the decoder side that the proposed scheme exploits temporal and spatial redundancy without introducing any additional processing at the encoder side, which keeps the complexity of the encoding as low as possible with certain compression efficiency. As a powerful alternative to Turbo codes, LDPC codes have been applied to our scheme. Since video data are highly non-ergodic, we use rate-adaptive LDPC codes to fit this variation of the achievable compression rate in our scheme. But one most basic difference between LDPC codes in our scheme and conventional channel coding is that in our scheme, we can make sure some bits are known. Those bits can work as constraints to the decoder and improve the decoding performance. We propose a constrained LDPC decoder not only to improve the decoder efficiency but also to speed the convergence of the iterative decoding. Simulation demonstrates that the scheme has significant improvement in the performances. In addition, the proposed constrained LDPC decoder may benefit other application.

Keywords: Distributed source coding, video coding, low-density-parity-check (LDPC) codes

1. INTRODUCTION

Current video coding paradigm, represented by the standardization efforts of ITU-T H.26X and ISO/IEC MPEG-X, lies on hybrid transformation and inter frame predictive coding. In this coding framework, the popular hybrid motion estimation, motion compensation, variable length coding and DCT transform are used to exploit the spatial and temporal redundancy existing in a video sequence at the encoder side. Because some encoding specific components, such as motion estimation, have computationally intensive operations even when efficient fast motion search is employed, the encoder is 5 to 10 times more complex than the decoder [1]. Such architecture is suitable for some applications, such as TV broadcast or video-on-demand, where video decoders have been sitting at the consumer side and the system has only a few encoders and numerous decoders, i.e. one-to-multiple topologies. The sheer volume of the decoder demands that the decoder be as simple as possible, often at the expense of encoder complexity.

However contemporary emerging mobile digital video applications, such as wireless low-power surveillance, multimedia sensor networks, wireless PC cameras and mobile camera phones, demand that the encoder be as simple as possible. For some applications, these battery-powered mobile portable devices usually have limited processing power and memory. In other applications, there are a great number of encoders and only one decoder. Therefore, it is desirable to have a low complexity video encoder to meet the resource limitations in mobile video with a coding efficiency similar to that of traditional video schemes, i.e. the shift of complexity from the encoder to the decoder should not compromise the coding efficiency [2].

*This work was supported by China Scholarship Council (CSC). The email and phone correspondences for the authors: Rong Ke Liu: rongke.liu@ieee.org ; phone 1 321 674 7055; Guo Gang Hua: ghua@fit.edu , phone 1 321 674 7556; Chang Wen Chen: cchen@fit.edu ; phone 1 321 674 8769

Visual Communications and Image Processing 2007, edited by Chang Wen Chen, Dan Schonfeld, Jiebo Luo,
Proc. of SPIE-IS&T Electronic Imaging, SPIE Vol. 6508, 65080R, © 2007 SPIE-IS&T · 0277-786X/07/$18

Recently, distributed video coding schemes have been proposed to provide potential reverse in computational complexity for decoder and encoder [3-5]. In these schemes much of codec complexity is moved from the encoder to the decoder. The theoretical background of these schemes is based on Slepian-Wolf and Wyner-Ziv distributed source coding theories which were established more than 30 years ago [6-7]. The distributed video coding schemes proposed recently have a general architecture: intra encoding and inter decoding. At the encoder, error control codes are applied to each intra frame and generate syndrome bits. To achieve compression effect, usually only part of the syndrome bits (punctured from the original syndrome bits) are sent to the decoder. The decoder uses the correlation between the video frames to construct an estimation of the current frame; such estimation can be viewed as a noisy version of the original frame. Then, the decoder combines the syndrome bits received and the estimation of the current frame to reconstruct the current frame. From such a process, we see that the performance of distributed video coding depends on two aspects: one is the accuracy of the estimation of current frame, the other is the decoding bit error rate performance of the error control codes. If the estimation is accurate enough, fewer number of syndrome bits would be needed to correct the error and we can use a higher coding rate error control code. Similarly, if the decoding bit error rate performance of the error control codes is good enough, a higher coding rate error control code can be applied to generate fewer syndrome bits adequate to correct the estimation errors.

There are various studies [8-10] to address those two facts to achieve better rate-distortion results. In [9], hash codewords of the current frame are generated and sent to aid the decoder to estimate more accurately the motion. The hash codewords are coarsely quantized version of a down-sampled 8 by 8 image block. To save the bit rate, the distance between two hash codeword of co-located blocks on previous and current frame is calculated to decide whether or not to send the codewords. In [10], the authors use highly compressed version of each frame as reference to perform motion estimation at the decoder in order to build more accurate estimation Although some cost for compression and transmission of those low quality frames are added to bit streams, the overall bit rate can be saved because more syndrome bits will be saved by accurate estimation. As to the second fact, trellis code was first used in the PRISM [3]. More efficient Turbo code and LDPC code were then applied in [8, 9, 11]. In [10], product-accumulate (PA) code was used due to its good performance with high code rate [12].

All of the proposed studies just applied the distributed source coding to video coding application, they separate the two facts mentioned above and improve the compression performance at the cost of encoder complexity. In this paper, we take the characters of video sequence and distributed source coding into account; propose a new scheme to integrate the frame estimation and the design of LDPC codes to improve the overall coding performance. In proposed scheme, zero motion block identification can be introduced at the decoder, which exploits the temporal correlation between adjacent frames. The decoder can use the unchanged bits in the zero motion blocks as a constraint to achieve a better decoding performance and further improve the overall video compression efficiency. The proposed scheme addresses temporal and spatial redundancy only at the decoder side without introducing any additional processing at the encoder side, thus keeping the complexity of the encoding as low as possible. Our experimental results show that the proposed constrained LDPC is even able to outperform the conventional algorithm. The rest of the paper is organized as follows. In section II, we describe the architecture of the propose scheme based on the zero motion block identification and constrained rate adaptive LDPC codes. In section III, we review the rate-adaptive LDPC encoder and propose the constrained LDPC decoder for the proposed scheme. The simulation results are given in section IV, and followed by the conclusion and discussion in section V.

2. ARCHITECTURE OF THE PROPOSED DISTIRBUTED VIDEO CODING

Fig. 1 gives the architecture of the proposed distributed video coding scheme. The general architecture of this scheme is similar to the one proposed by Aaron et al: both Slepian-wolf codec and a frame prediction module. At the encoder side, two sets of bit streams are generated: one from the low complexity video frame estimation and another one from the syndrome bits. At the decoding side, these two sets of bitstreams are combined to reconstruct the high quality video frames. The Slepian-Wolf encoder includes a rate adaptive LDPC encoder and a buffer and it produces a sequence of parity bits (syndrome bits) associated to each biplane. The parity bits generated by the LDPC encoder are then stored in the buffer, and transmitted upon request by the decoder while the systematic bits are discarded. The feedback channel is necessary to adapt to the changing statistics between the side information and the frame to be decoded, i.e. to the quality (or accuracy) of the frame reconstruction process.

There are however some major differences between the proposed solution and the one proposed by Aaron et al, namely in the Slepian-wolf Codec and zero motional block identification. At the decoder, for each reconstructed frame of the input video, zero motion vector blocks are identified. This can be done by calculating the adjacent frame's $SAD(0)$:

$$SAD(0) = \sum_{x=1}^{M}\sum_{y=1}^{N}\left|p(x,y)-n(x,y)\right| \tag{1}$$

Where p and n are previous and next to current frame, respectively. If the $SAD(0)$ is smaller than a pre-set threshold, we name the block as zero motion one. When the temporal distance between successive key frames is two, p is previous key frame and n is next key frame. When the temporal distance between successive key frames is more than 2, p is current frame's previous frame and n is current frame's next key frame. If the scene is too smooth, the temporal distance between p and n may be more than two.

There are two benefits from zero motion identification. In one side, identifying zero motion blocks may improve the accuracy of frame estimation in decoder. Since some blocks remain unchanged, when the decoder reconstruct the current frame, it can just use the corresponding blocks in reference frame to substitute the current block, and the estimation is completely correct. In the other side, if the syndrome bits generated by feeding data from all blocks are sent to decoder, the decoder may use the zero motion block identification information as constraint to perform constrained decoding, which will improve the decoding performance and the overall compression performance.

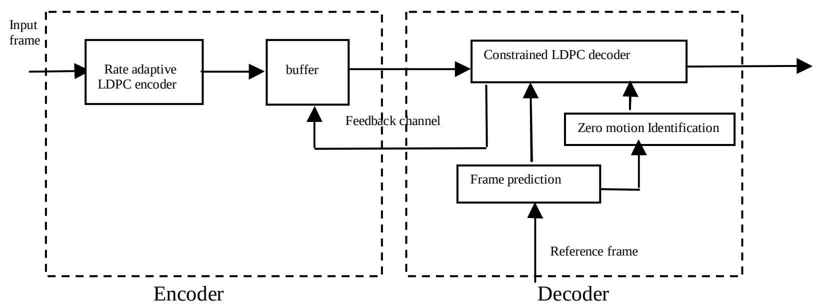


Figure. 1 Architecture of proposed distributed video source codec

Since video data are highly non-ergodic, the achievable compression ratio varies and cannot be foretold by the encoder. In this situation, a rate-adaptive scheme is suitable. Although punctured turbo codes and punctured LDPC codes were used to implement rate-adaptive source coding with and without decoder side information, LDPC-based rate-adaptive distributed source codes [14] were designed, which have better performance than alternative codes of linear complexity encoding and decoding for asymmetric DSC.

In this scheme, the zero motion blocks information is obtained by reference frame analysis, which will improve the compression efficiency of the encoder. The details of rate adaptive LDPC codes for distributed source coding and constrained LDPC decoder will be discussed in the following section.

3. LDPC CODES FOR DISTRIBUTED VIDEO CODING

In this section, we present a very brief overview of the general concept of LDPC codes and the rate adaptive ideas contained in [14], where rate adaptive LDPC codes were proposed, and focus on the constrained LDPC decoder for distributed video coding. Some modifications with respect to conventional decode algorithm are necessary in order to achieve good performance.

3.1 Review of LDPC Codes

LDPC codes are linear block codes. Instead of being described by a generator matrix G as usual, it is described by a parity check matrix H. The low density of LDPC codes lies in that its parity check matrix has sparsity of "1", namely the number of "1" in each row and column is small compared to the block length. A regular (n,k) LDPC code is a linear block code, whose parity-check matrix H contains w_c "1" in each column and w_r "1" in each row, where $w_r = w_c \cdot n/(n-k)$. In order to ensure good performance, w_c is expected to be at least 3. An LDPC code is irregular if its H is of low density but the number of "1" in each column or row is not the same. In common cases, irregular LDPC codes are superior to regular LDPC codes under the condition of a very long frame.

LDPC codes can be decoded by using a probability propagation algorithm known as the sum-product or belief propagation (BP) algorithm, which is represented by a Factor Graph that contains two types of nodes: the "variable nodes" and the "check nodes" . Each variable node corresponds to a column of a parity-check matrix, which also corresponds to a bit in the codeword. Each check node corresponds to a row of a parity-check matrix, which represents a parity-check equation. An edge between a bit node and a check node exists if and only if the bit participates in the parity-check equation.

The process of decoding is briefly described here: 1) Initialization: pass initial message f_j^a to variable Q_{ij}^a which is related with channel; 2) Calculating iteratively the message from variable Q_{ij}^a to check R_{ij}^a (Horizontal step); 3) Calculating iteratively the message from check R_{ij}^a to variable Q_{ij}^a (Vertical step); 4) Tentative decoding: Make a hard decision of variable message and obtain the tentative decoder bits $\hat{x}$, check $Hx^T = 0$ if yes, then the decoding algorithm halts. Otherwise, the algorithm repeats from the horizontal step. A failure is declared if some maximum number of iterations (e.g., 100) occurs without a valid decoding.

It is remarkably noted that when the communication system based on LDPC codes works under the different channels, the only difference to make is to modify the initialization message (likelihood ratio) in the first step of decoder procedure.

For AWGAN:

$$\mathrm{L}(c_i) = \ln\left(\frac{\Pr(c_i = 0|y_i)}{\Pr(c_i = 1|y_i)}\right) = 2y_i/\sigma^2 \tag{2}$$

Where σ^2 denotes the variance of the addictive noise and y_i indicates the soft output of the Gaussian Channel.

For binary systematical channel (BSC):

$$\mathrm{L}(c_i) = \ln\left(\frac{\Pr(c_i = 0|y_i)}{\Pr(c_i = 1|y_i)}\right) = (-1)^{y_i}\ln((1-p)/p) \tag{3}$$

Where p denotes cross probability of BSC.

For Binary Erasure Channel (BEC)

$$\mathrm{L}(c_i) = \ln\left(\frac{\Pr(c_i = 0|y_i)}{\Pr(c_i = 1|y_i)}\right) = \begin{cases} +\infty & y_i = 0 \\ 0 & y_i = E \\ -\infty & y_i = 1 \end{cases} \tag{4}$$

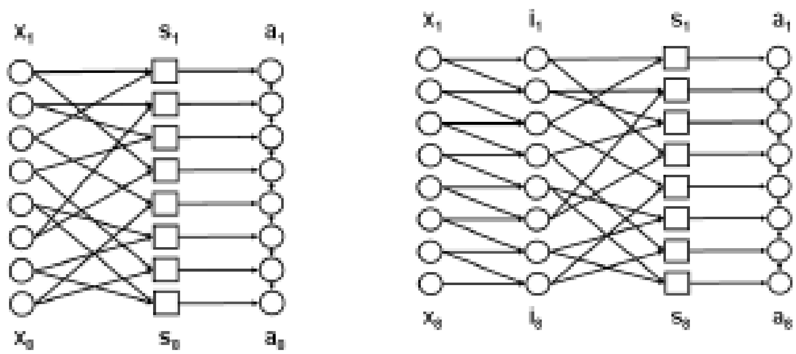
Wherein, E means the bit is eassured by BEC.

Due to assumptions that all of the received bits are completely random and independent, they are treated equally without any difference. However each bit probably has different prior knowledge in real system.

3.2 Rate-adaptive LDPC encoder for DSC

The key idea to use LDPC codes for asymmetric compression of source was introduced in [10]. It is based on dividing sequence X from source into blocks of length *N*, and using an LDPC code of rate $R_C = 1 - L/N$ to calculate the corresponding L checks, which form the compressed sequence. The resulting compression rate for source is, therefore, $R_1 = L/N = 1\text{-}R_C$.

LDPC codes have been successfully used in fixed-rated distributed source coding [11]. Considering that video data are highly non-ergodic, the achievable compression ration should be adjustable. A simple way to use such a code as part of a rate adaptive scheme would be to transmit the syndrome bits in stages and allow decoding after receipt of each increment of the syndrome. However, the performance of the high compression codes so derived is very poor because their graphs contain unconnected or singly-connected source nodes; these structural features impede the transfer of information via the LDPC iterative decoding algorithm. Two classes of rate adaptive distributed source codes, both based on LDPC, are originally proposed by Girod et al [14]. The proposed LDPCA encoder consists of an LDPC syndrome-former concatenated with an accumulator. An example is shown in Figure 2(a). These syndrome bits are in turn accumulated modulo 2, producing the accumulated syndrome. The encoder buffers the accumulated syndrome and transmits it incrementally to the decoder. The LDPCA decoder handles rate-adaptively by modifying its decoding graph each time it receive an additional increment of the accumulated syndrome. If the accumulated syndrome subset were transmitted instead of the syndrome subset, the decoding graph would maintain the degree of source nodes, otherwise the graph would be severely degraded and unsuitable for iterative decoding. Another class of rate adaptive LDPC, SLDPCA, is the concatenation of a consecutive summer with an LDPCA encoder, as figure 2(b). The definition of all above symbols is the same as [14].



(a) (b)

Fig. 2. Rate-adaptive LDPC encoder[13] (a) LDPCA (b) SLDPCA

3.3 Constrained LDPC decoder for DSC

In distributed source coding, error control coding methods are used to construct the practical distributed source coding/decoding. But there are some basic differences between distributed source coding and error control coding, since distributed source coding is a source coding other than channel coding. One of the basic differences is that in error control coding, all bits have the chance to be corrupted by the channel noise; while in distributed source coding, only the systematic bits may be “corrupted”, but the syndrome bits remain unchanged for sure. And in some cases, we know some of systematic bits are unchanged, too.

In the proposed scheme in section II, we say that in distributed video coding, if zero motion blocks are identified and informed to decoder, the decoder will know that besides the parity bits, and some of the systematic bits (variable bits) remain unchanged. As a whole, the constrained bits are all parity bits (syndrome bits) and some of systematic bits (variable bits). To make good use of these constrain in the case, some modifications of conventional sum-product algorithm (SPA) [13] are indispensable. The constraint from the syndrome information was taken into account in [11]. The message from the check nodes (syndrome bits) was modified.

As to another constraint from system bits that the variable node c_i was unchanged, the modification should be done to the initial message of variable nodes. The initial should be determined as follows:

To the Probability-domain SPA:

$$\Pr(c_i = b|y_i) = \begin{cases} 1 & \text{when } y_i = b \\ 0 & \text{when } y_i = b^c \end{cases} \tag{5}$$

Where b is 0 or 1, and b^c denotes its complement.

To Log-domain SPA:

$$\mathrm{L}(c_i) = \ln\left(\frac{\Pr(c_i = 0|y_i)}{\Pr(c_i = 1|y_i)}\right) = \begin{cases} +\infty & y_i = 0 \\ -\infty & y_i = 1 \end{cases} \tag{6}$$

4. PRELIMINARY EXPERIMENTAL RESULTS

In this section, we present the numeric results on how many blocks is zero motion and the decoding performance of our proposed scheme based on constrained LDPC decoder for distributed video coding.

4.1 Zero motion blocks

We use Car phone and Foreman sequences in our experiment. The frame rate is 30 fps. We use the block in the reference frame to substitute current block, and calculate the PSNR of the substitution (the block PSNR, not the PSNR of the whole frame). We test the percentage of the blocks which are zero motion when the thresholds are 30dB, 35dB and 40dB respectively. Table I gives the results when the block size is 4 by 4, and Table II gives the results when the block size is 8 by 8. From the results, we can find that a significant part of blocks actually can be zero motion in video sequence. It shows that the zero-motion-block-identification is valuable.

TABLE I Percentage of zero motion block (block size 4X4)

	30dB	35dB	40dB
Foreman	56.84%	42.87%	28.27%
Carphone	66.17%	55.04%	43.42%

TABLE II Percentage of zero motion block (block size 8X8)

	30dB	35dB	40dB
Foreman	49.32%	34.37%	20.31%
Carphone	60.30%	48.35%	37.13%

4.2 Performances of the schemes based on constrained LDPC decoder

In section II, we proposed a scheme to make use of the zero motion blocks. This scheme can improve the decoding performances of the LDPC codes. Since video data are highly non-ergodic, the achievable compression ratio varies. In this situation, a rate-adaptive LDPC code is an attractive solution [13]. The rate-adaptive codes have been demonstrated to be superior to linear encoding and decoding complexity alternatives for asymmetric distributed source coding. LDPCA is able to perform within 10% of the Slepian-Wolf bound in the moderate and high rate regimes.

To test the performance, 10000 randomly generated binary sequences are used. Each sequence has 6336 bits, but some of them are known unchanged in both encoder and decoder by zero motion information. The number of original bits participating in the LDPC code is fixed to N=6336. The used LDPC codes are specified in [14]. The regular LDPCA codes have a degree distribution given by $\delta_3 = 1$, where δ_r is the proportion of nodes of degree r. The irregular LDPCA codes have the following degree distribution selected

$$\delta_2 = 0.316, \delta_3 = 0.415, \delta_7 = 0.128, \delta_8 = 0.069, \delta_{19} = 0.020, \delta_{21} = 0.052$$

Figure 3 shows simulation results for regular LDPCA, irregular LDPCA and proposed constrained LDPC codes of length 6336 for 100 iterations in the decoder. The constrained LDPC code was supposed to remain 50% unchanged bits.

From Figure 3 we can see that proposed scheme has a significant improvement over the conventional SPA. It is impossible to achieve the compression if known bits are not exploited in the decoder. In the distributed source coding based on channel coding, the encoder transmits a short syndrome based on an aggressive code and the decoder attempts decoding. In the event that decoding is successful, the decoder signals this fact to the encoder, which then continues with the next block of source data. However, if the decoder fails, the encoder augments the short syndrome with additional transmitted bits, creating a longer syndrome based on a less aggressive code. The process loops until the syndrome is sufficient for successful decoding. In the proposed scheme, the known bits bring a stronger constraint and speed the iteration converge towards correct codes in the decoding process, thus the shorter syndrome is sufficient for successful decoding.

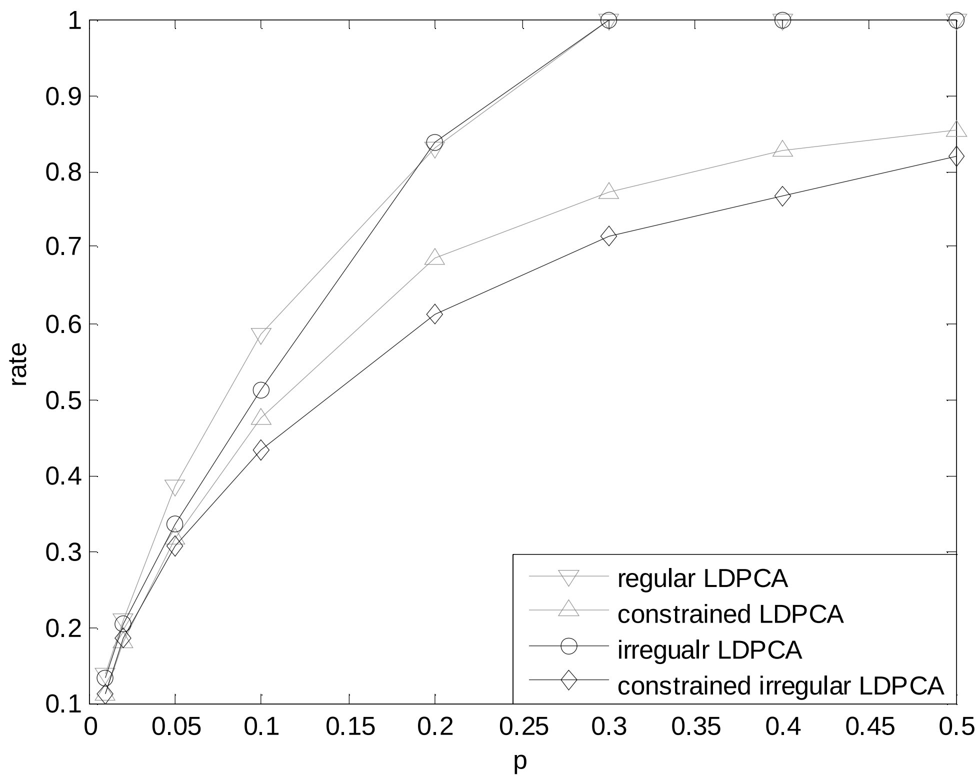


Figure 3 Performance of constrained LDPC decoder comparing with conventional SPA

Figure 4 presents the compression rate as a function of the percent of unchanged bits. The crossover probability p is 0.1, and other parameters are as above. As we can see, the compression rate is increased with the increase of the percent of unchanged bits. The reason is that when the known bits increase, the known message is increased and the decoder needs less syndrome bits.

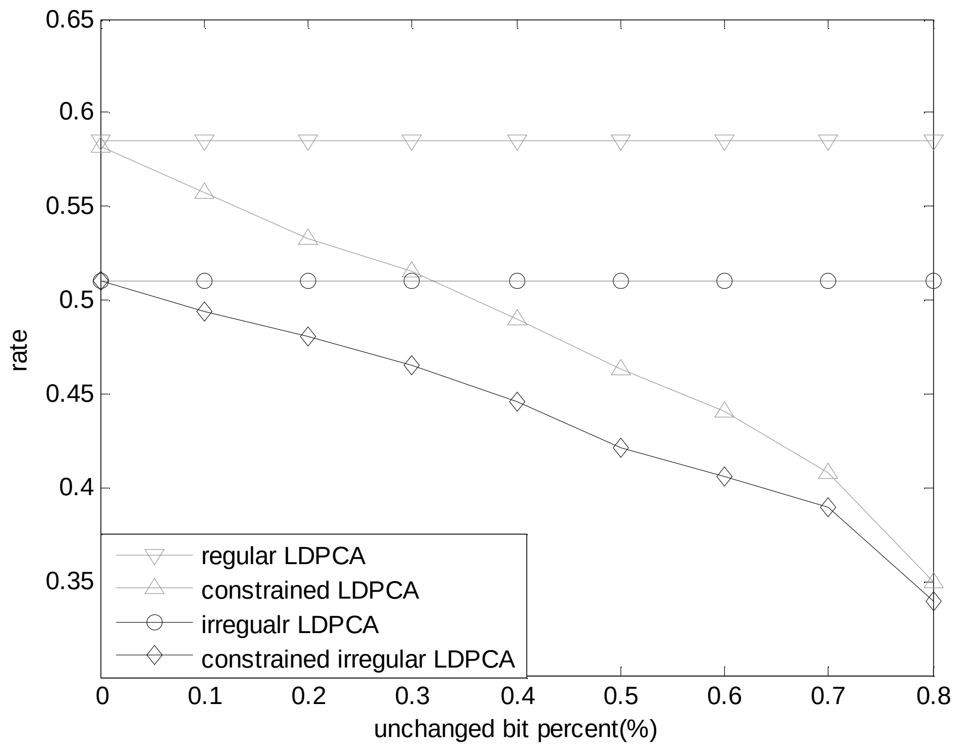


Figure 4 Compression rate performances versus the percentage of unchanged bits

If we could identify zero blocks at the encode side, we may skip the zero motion block with a few bits identifying the location of zero motion block [15]. Moreover we notice that when the percent of unchanged bits is very high, the zero motion block skip mechanism at the encoder side will have better overall performance than zero-motion block transmission because the contribution from syndrome bits is more powerful than that from known bits in this situation. Combining Figure 3 and Figure 4, we found the zero block skip is better than the zero block transmission when the percent of unchanged bits is over 50% ($p = 0.1$). But zero block skip proposed by [15] has better performance at the expense of encoder computational complexity.

5. CONCLUSION AND DISCUSS

In this paper, we take the characters of video sequence and distributed source coding into account, and propose a new scheme to integrate the frame estimation and the error control decoder to improve the overall performance. We introduce zero motion block identification at the decode side without any modification of the encoder, so keep the encoder complexity as possible as low. We also proposed a constrained LDPC decoder for distributed video coding to make good use of constraint bits from zero motion block, which may improve the error control ability and speed the coverage of iterative decoding. The simulation demonstrates that a significant part of blocks actually can be zero motion in video sequence, and the scheme has significant improvement in the overall coding performances. In addition, the constrained LDPC decoder may benefit other application where some bits have different prior knowledge from the others.

REFERENCES

1. Y. Liu and S. Oraintara, “Complexity comparison of fast block-matching motion estimation algorithms,” in Proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing, May 2004, pp. 17-21.
2. C.W. Chen, “Mobile Digital Video: Challenges and Opportunities”, in Proc. of PCS, Beijing, May,2006
3. R Puri and K. Ramchandran, “PRISM: A new robust video coding architecture based on distributed compression princples,” in Proc. of Conference on Communication, Control and Computing, Allerton, IL, October 2002.
4. A. Aaron and B. Girod, “Towards practical wyner ziv coding of video,” in Proc. of IEEE International Conference on Image Processing, Barcelona, Spain, September 2003.
5. B. Girod, A. Aaron, S. Rane and D. Rebollo-Monedero , "Distributed video coding," Proceedings of the IEEE, Special Issue on Video Coding and Delivery, vol. 93, no. 1, pp. 71-83, Jan. 2005.
6. D. Slepian and J. K. Wolf, “Noiseless coding of correlated information sources,” IEEE Transaction on Information Theory, vol. 19, pp. 471-490, July 1973.
7. D. Wyner and J. Ziv, “The rate-distortion function for source coding with side information at the decoder,” IEEE Transaction on Information Theory, vol. 22, pp. 1-10, January 1976.
8. A. Aaron, S. Rane, E. Setton, and B. Girod, “Transform domain wyner-ziv codec for video,” in Proc. SPIE Visual Communications and Image Processing, San Jose, CA, January 2004.
9. A. Aaron, S. Rane, and B. Girod, “Wyner-ziv video coding with hash-based motion compensation at the receiver,” in Proc. IEEE International Conference on Image Processing,” Singapore, October 2004.
10. M. Wu, A. Vetro, J.S. Yedidia, H. Sun, and C.W. Chen, “A study of encoding and decoding techniques for syndrome based video coding,” in IEEE International Symposium on Circuits and Systems, Kobe, Japan, May 2005, vol. 4, pp. 3527-3530.
11. A.D. Liveris, Z. Xiong, and C. Georghiades, “Compression of binary sources with side information at the decoder using ldpc codes,” IEEE Communications Letters, vol. 6, pp. 440.442, Oct 2002.
12. J. Li, K.R. Narayanan, and C.N. Georghiades, “Product accumulate codes: a class of codes with near-capacity performance and low decoding complexity,” IEEE Transactions on Information Theory, vol. 50, pp. 31.46, January 2004.
13. W. E. Ryan, "An Introduction to LDPC Codes," in CRC Handbook for Coding and Signal Processing for Recording Systems (B. Vasic, ed.) CRC Press, 2004
14. D. Varodayan, A. Aaron, and B. Girod, “Rate-Adaptive Distributed Source Coding using Low-Density Parity-Check Codes,” Proc. Asilomar Conf. Signals, Syst., Comput., Pacific Grove, CA, Nov. 2005.
15. G. G. Hua and C.W. Chen, “low punctured turbo codes and zero motion skip encoding strategy for distributed video coding”, Global Com 2006, San Francisco, CA, Nov.2006 (accepted)