An Iterative Algorithm for Efficient Adaptive GOP Size in Transform Domain Wyner-Ziv Video Coding

Khanh DinhQuoc, Xiem HoangVan, and Byeungwoo Jeon

School of Electrical and Computer Engineering, Sungkyunkwan University
300 Chunchun-dong, Jangan-gu, Suwon, 440-746, Korea
{diqkhanh,xiemhoang}@gmail.com, bjeon@skku.edu

Abstract. Transform Domain Wyner-Ziv Video Coding (TDWZ) is one of the most popular paradigms of Distributed Video Coding (DVC) which supports low encoding complexity. However, there is still a gap in its coding performance compared to conventional video coding standards such as MPEG-x, or H.264/AVC. In order for TDWZ to reach comparable performance to them, a good method for deciding a proper Group of Picture (GOP) size is in great necessity. From this point of view, we propose an iterative algorithm which efficiently determines GOP size based on an intra mode decision method at frame level. This approach firstly constructs a coarse GOP size and then refines it by iterative checking for final GOP size. Experimental results show superiority of the proposed algorithm with improvement up to 2dB in term of coding efficiency.

Keywords: Distributed Video Coding, Adaptive GOP size, Intra mode decision, Hierarchical structure.

1 Introduction

In human life, visual perception plays an essential role in receiving information from the outside world. That explains why video is the most attractive form of data in multimedia system. However, uncompressed video data require extremely large bandwidth for transmission or enormous storage capacity. To solve this problem, a lot of efforts have been made to develop video coding compression techniques which help reducing the number of bits to represent a video sequence.

Since the first digital video compression standard named H.120 was published by CCITT (Consultative Committee International Telephone and Telegraph) in 1984, the former name of ITU-T (International Telecommunication Union - Telecommunication Standardization Sector), video compression techniques have undergone long and steady developments. ISO/IEC (International Organization for Standardization/International Electrotechnical Commission) published MPEG-1, MPEG-2, and MPEG-4. ITU-T has also developed H.261, H.262, H.263 independently or jointly with ISO/IEC. Most recently, under the joint effort of both ISO/IEC and ITU-T, H.264/AVC (Advanced Video Coding) was finalized in 2003, and its a few extensions, namely, professional extension, scalable extension, multiview extension follow later. Currently, HEVC (High Efficiency Video Coding) is being developed under

Y.-S. Ho (Ed.): PSIVT 2011, Part II, LNCS 7088, pp. 347–358, 2011.

JCT-VC (Joint Collaborative Team on Video Coding) with a target of compression efficiency enhancement by additional 50% compared to H.264/AVC inter coding in high profile.

However, those conventional video coding schemes require extreme amount of computational complexity at the encoder. Therefore, it is not suitable for up-link applications where low-encoding complexity is a major requirement.

With the recent explosion of hand-set devices, visual sensors and cameras, from wireless low power surveillance to mobile visual sensor network, or from video conference with mobile camera to multi-view video entertainment, a very low complexity at the encoder becomes a more essential feature due to their limited power supply and desire to have low complexity. In conventional video compression, encoder performs motion vector search, coding mode decision, rate-control, rate-distortion optimization, etc.; that means most of computing load is located at the encoder. But for the emerging applications, motion vector search, the most complexity component at encoder, is desired to be shifted to the decoder. Distributed Video Coding (DVC) scheme is one possibility to address this kind of requirement.

DVC is based on two key theorems by Slepian-Wolf [2] and Wyner-Ziv [3]. Slepian-Wolf theorem proved that two correlated sources can be encoded independently without any loss in coding efficiency if they are jointly decoded by exploiting source statistics. In 1976, Wyner-Ziv theorem further proved similar source coding method using side information for lossy compression. In 2002, two practical paradigms of DVC have been introduced by B. Girod et al. [4] and Ramchandram et al. [5]. The paradigm introduced in [4] was further developed into TDWZ (Transform Domain Wyner-Ziv Coding).

In TDWZ, Side Information (SI) generation and Channel Noise Modeling (CNM) are the two key functional components, which mostly decide its coding efficiency. Both of them depend on the distance between Wyner-Ziv (WZ) frames and key frames, that is, the GOP (Group of Picture) size. Pereira et al. [9] studied relationship between GOP size and performance of TDWZ. They showed that the rate-distortion or quality of decoded sequence relies on motion-level of sequences and different motion level should choose different GOP size. With high motion sequences such as Soccer or Coastguard, smaller GOP size gives better coding efficiency. Contrarily, for low motion sequence, larger GOP size gives better results. However when GOP size is too long, performance also decreases, for example with Hall monitor, best GOP is around four. Therefore, a proper GOP size is very important in term of coding efficiency.

Ascenso et al. [10] decided GOP size based on motion activity inside sequence. Their approach is based on four different matrices: Difference of Histogram (DH), Histogram of Difference (HD), Block Histogram Difference (BHD), and Block Variance Difference (BVD). They found that matrices DH and HD are powerful in working at frame level and detecting changes in global motion such as scene changes; matrices BHD and BVD are suitable for detecting high motion sequences, sequences with local motion in statistic background and group consecutive frames with same characteristic of motion into together. However, their method requires four complex matrices which require encoder to execute much calculation. The advantage of low complexity encoder is inflicted.

By other approach, Yaacoub et al. [11] determined a GOP size which has the minimum ratio PSNR/rate. That means they created models to estimate rate and PSNR values of both WZ and Intra frame, and set a loop for finding total PSNR and total rate to obtain the ratio of PSNR/rate. Their method requires heavy encoding complexity. Besides, using a fixed model for all sequences may not be totally accurate; for example, the result for sequence Salesman is slightly worse than fixed GOP size of 3.

In our proposed method, we determine GOP size in two steps - the first step finds a coarse GOP size by progressively looking for a frame which has higher probability of being intra-coded, then, the second step refines it based on iterative checking algorithm to decide the final GOP size. The intra mode decision at frame level compares temporal correlation and spatial correlation as described in [13]. Our simulation results show that it can adapt to various motion content of sequences and brings improvement up to 2dB in term of coding efficiency.

This paper is organized as follows. In Section 2, TDWZ with adaptive GOP size architecture is briefly introduced. Our proposed method is presented in Section 3. The proposed method is experimentally verified in Section 4; and finally, in Section 5, conclusions and some suggestions for future work are drawn.

2 Transform Domain Wyner-Ziv Coding with Adaptive GOP

Transform domain Wyner-Ziv video coding (TDWZ) is one representative DVC scheme [12] in which a video sequence is divided into key and WZ frames according to GOP size; the first and the last frames of a GOP are key frames, and all the other frames inside GOP are encoded as WZ frames. The key frames are encoded with low complexity compression techniques such as JPEG or H.264/AVC intra.

In the proposed TDWZ with adaptive GOP size (see Fig. 1), input sequence is firstly processed by GOP size controller to construct a GOP size. WZ frames are divided into blocks of 4x4 and discrete cosine transform is applied to each block to convert pixel data into frequency domain. Quantizer removes visual perception redundancy by applying large quantization step size at high frequency and small step size at low frequency.

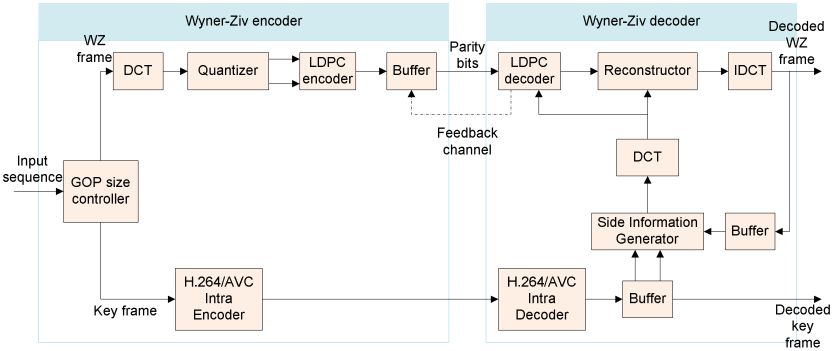


Fig. 1. Transform Domain Wyner-Ziv with GOP size controller

Then, DCT coefficients at the same frequency are grouped together and extracted into biplanes to feed into channel coding encoder.

The proposed TDWZ decoder decodes the first ($\hat{X}_t$) and the last frames ($\hat{X}_{t+GOPsize}$) of a GOP by H.264/AVC Intra decoder and store them in a buffer. Side information (SI) Generator refers to the decoded intra frames to create an error version ($Y_{t+WZindex}$) of WZ frame. Here, WZindex denotes a displaying order of a WZ frame. The procedure proposed by Ascenso et al. [6] is an efficient way to generate SI which uses bi-direction interpolation and weighted median filter. Based on backward reconstructed SI and forward reconstructed SI, a Channel Noise Model (CNM) between SI and WZ frames is made [14]. To correct error in SI, bitplanes of coefficients in frequency domain and channel model are processed at LDPC decoder to detect which bit has high probability of error. Then, requests for parity bits are sent to the encoder via a feedback channel as necessitated. Errors will be corrected by using transmitted parity bits. Under a hierarchical GOP structure, previously decoded WZ frame is used as a reference frame alternatively for a key frame which is located too far away from current WZ frame.

Although TDWZ is quite a powerful compression paradigm but it still cannot be comparable to the state-of-the-art performance of H.264/AVC. This gap must be overcome by researching more on coding efficiency improvement of TDWZ.

3 Proposed Iterative Algorithm for GOP Size Construction

The key idea of our proposed method is iteratively checking intra mode decision procedure. Firstly, we coarsely determine a GOP size as a preliminary one. Then by hierarchically checking frames inside the coarse GOP we figure out whether the preliminary GOP should be broken into smaller GOPs or not.

3.1 Intra Mode Decision

In DVC paradigm, quality of decoded WZ frame is still worse than that of decoded B frame in H.264. Sometimes it is worse than that of decoded intra frame, especially in high motion sequences such as Soccer or Stefan sequences. However, applying WZ coding is in general better in stationary frames. Therefore, there has been much effort to choose between intra coding and Wyner-Ziv coding to adapt to content of sequence [15-17]. Some researchers focused on using a threshold for mode decision, such as Belkoura et al. [15] or Do et al. [16]. However, using a same threshold, which highly depends on user's experience, for all sequences cannot give the best results always. Some others such as Ascenso et al.[17] tried to create models to estimate the rate and distortion for Wyner-Ziv mode and intra mode. Based on rate distortion model [17], they could judge which coding mode should be chosen. Nevertheless, such estimation model is not always correct for all sequences and may bring noticeable increment in encoding complexity. Xiem et al.[13] proposed a method of quite light complexity which did not rely on threshold. In this paper, we use similar idea as [13] as follows.

Denote $SAD^{T}_{t->(t-i,t+j)}$ as a difference from the current frame at time *t* to two frames at time *t* - *i* and *t* + *j*, respectively.

$$SAD^{T}_{t->(t-i,t+j)} = \sum_{r=1}^{H}\sum_{c=1}^{W}\left\{\left|F_{t-i}(r,c)-F_{t}(r,c)\right|+\left|F_{t}(r,c)-F_{t+j}(r,c)\right|\right\} \qquad (1)$$

where $F_t(r,c)$ is a pixel value at position *(r, c)* of current frame *t*; *H* and W respectively denote the height and the width of a frame. Positions of frames *t*, *t* – *i* and *t* + *j* are illustrated in Fig. 2. In general, increased motion level should lead to the difference increasing, and vice versa. Hence, $SAD^{T}_{t->(t-i,t+j)}$ is an indicator (actually in a reverse sense) for temporal correlation. In the same spirit, compute SAD^{S}_{t} as a difference among blocks inside a frame at time *t*. Similarly, SAD^{S}_{t} is a representative of spatial correlation in a reverse sense.

$$SAD^{S}_{t} = \sum_{b=1}^{B}\sum_{r=1}^{blocksize}\sum_{c=1}^{blocksize}\left|F_{t,b}(r,c)-M_{b}\right| \qquad (2)$$

where *B* denotes a total number of blocks in a frame *t*, $F_{t,b}(r,c)$ refers a pixel value at position *(r, c)* in the block *b*. In this paper, we used *blocksize*=4. M_b is a median value of three corresponding pixels of neighbor blocks: north (up), west(left) and northwest (top-left) direction. If SAD^{S}_{t} is larger than $SAD^{T}_{t->(t-i,t+j)}$, frame *t* should be encoded as WZ frame and vice versa.

The proposed GOP size determination is made in two steps (coarse construction and iterative fine determination) as follows.

3.2 Coarse GOP Size Constructor

A key point of this step is that if SAD^{S}_{t} is the smallest than $SAD^{T}_{t->(t-i,t+j)}$ values, the current frame should be coded as an intra frame. When the distances between current frame *t* to previous frame *t* - *i* and next frame *t* + *j* are larger, usually, the value of SAD^{S}_{t} is larger. We can infer that $SAD^{T}_{t->(t-i,t+j)} > SAD^{T}_{t->(t-1,t+1)}$ with every *i* > 1 and *j* > 1.

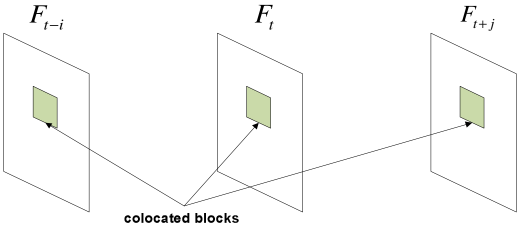


Fig. 2. Relative positions of frame *t*, *t* – *i*, and *t* + *j*

Therefore $SAD^{T}_{t->(t-i,t+j)}$ corresponds to the largest difference around time *t*. As stated above, to identify a coarse GOP size, we search for a frame which has a smaller SAD^{S}_{t} than the smallest $SAD^{T}_{t->(t-i,t+j)}$, and denote the distance between the current intra frame and the frame found as a coarse GOP size.

Besides, if GOP size is too long, the decoder must wait until finishing decoding all frames in this GOP, reorder frame index from coded index into display index, and then play entire GOP. Thus, a long GOP size take a long time to process and may undesirably affect easy video watching. In order to avoid such shortcomings, GOP size is better to be limited below some longest value, GOP_MAX. To identify the coarse GOP size, at first we compute $SAD^{T}_{t->(t-i,t+j)}$ values of three consecutive frames and SAD^{S}_{t} values of center frame. Following, a coding mode is chosen according to the two relative factors explained above. If the center frame is decided to be intra-coded, this procedure is terminated and coarse GOP size is returned. Otherwise, we pick next three consecutive frames to check again in the same way. This procedure stops when center frame is assigned either as intra frame or GOP size reaches value of GOP_MAX. Fig. 4 summaries the coarse GOP size construction.

3.3 Iterative Finer-Size of GOP Determination Algorithm

When the previous key frame is too far away from the next key frame, $SAD^{T}_{t->(t-i,t+j)}$ of current frame with respect to the previous key frame and the next key frame is larger. Thus, wrong motion vectors occur more frequently and degree of accuracy of bi-direction interpolation critically degrades. Consequently, quality of SI will be degraded then much more parity bits need to be sent to decoder to correct the worse SI. Therefore, overall TDWZ performance will be seriously affected. We can overcome this cascading problem by checking which mode is applied to the center frame by an intra mode decision procedure. In another word, we have two inequalities:

$$SAD^{T}_{t->(previous\_key,next\_key)} > SAD^{T}_{t->(t-1,t+1)} \quad \text{and} \quad SAD^{T}_{t->(t-1,t+1)} < SAD^{S}_{t}$$

That means we cannot make a conclusion about relationship between $SAD^{T}_{t->(previous\_key,next\_key)}$ and SAD^{S}_{t}. Therefore, for the next processed frame, we must check whether it should be encoded as intra mode or not by comparing the two SAD values at current frame. If $SAD^{T}_{t->(previous\_key,next\_key)} > SAD^{S}_{t}$ the current frame is determined as an intra frame and we break the coarse GOP into two consecutive GOPs.

Besides, using a hierarchical GOP structure makes overall quality of decoded sequence better [8]. Inside an arbitrary GOP, the first decoded frame should be located in the center of GOP. Quality of decoded frame is increased by receiving parity bits and correcting errors in SI. Moreover, distance between the center frame and both key frames is smaller than distance apart themselves. By taking the analyzed advantages, we can improve the quality of latterly generated SIs which will increase overall GOP coding performance. We use this structure to order all frames fed into intra mode decision to guarantee that intra-coded frame (if it exists) will be as useful as possible.

Center frame of the coarse GOP in a hierarchical order is checked whether it should be intra-coded or WZ-coded. A coarse GOP is broken into two GOPs when center frame is decided as intra-coded frame; the iteration will continue applying into two halves of GOP with respecting them as two coarse GOPs. This process continues until all frames inside coarse GOP are assigned coding mode. For example, with a coarse GOP size of 8 as shown in Fig. 3 seven iterations are performed in order to mark from 1 to 7. Based on the assigned modes – Intra or WZ - for each frame in a coarse GOP, we separate it into smaller GOPs. Flow chart diagram of iterative finer-size algorithm is shown in Fig. 5.

In summary, the proposed iterative algorithm for adaptive GOP size works as follow. Procedure at section 3.2 finds a coarse GOP size. Final mode for a specific frame is determined by hierarchical checking of intra mode decision. At last, GOP size is determined by breaking the coarse GOP if necessary.

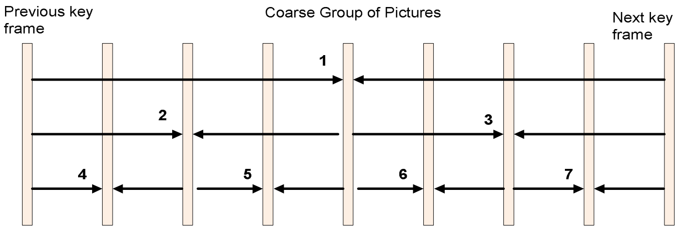
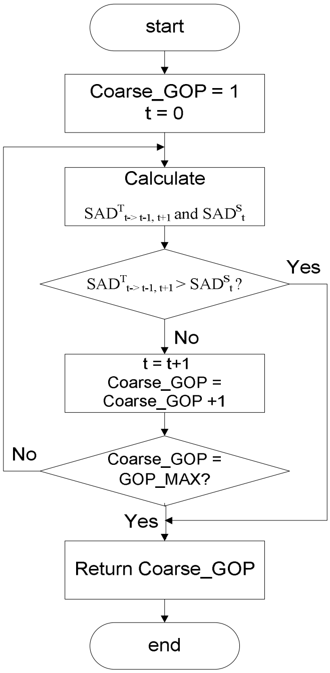


Fig. 3. Hierarchical structure for checking and encoding coarse GOP



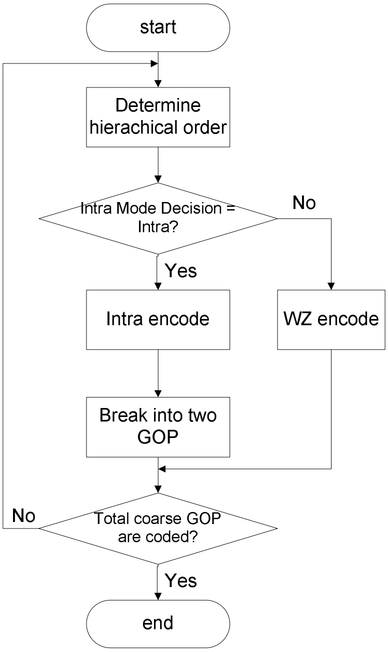


Fig. 4. Coarse GOP size determination

Fig. 5. Finer GOP determination

4 Experimental Results

Test condition

Our simulation is performed with three test sequences: Foreman, Hall monitor, and Soccer. The spatial and temporal resolutions are QCIF and 15 Hz, respectively. The number of frames is 149 frames for Foreman and Soccer, and 165 for Hall monitor. We compare coding performance of the proposed method ("**adaptive_GOP_size**") against following three methods.

- **SKKU DVC codec** which is developed by Digital Media Lab, Sungkyunkwan University [18]. GOP size of 2 is used for simulation.
- **DISCOVER codec** which was project result of EU IST FET program (Information Society Technologies – Future Emerging Technologies) [19]. It is chosen because it is the most well-known codec of Distributed Video Coding. GOP size of 2 is also used.
- **Intra Mode Decision** in [13] with GOP size of 2 and 4 is chosen to compare.

In addition, for displaying rate distortion performance, we set up quantization steps for Intra frames Q_p and quantization matrices for WZ frames Q_m. Our simulation is performed at four points of Q_p and Q_m pairs as shown in Table 1.

Table 1. Rate Distortion point for simulation

Foreman		Hall Monitor		Soccer	
Q_m	Q_p	Q_m	Q_p	Q_m	Q_p
1	40	1	40	1	40
4	34	4	34	4	34
7	29	7	29	7	29
8	25	8	25	8	25

Complexity

The proposed method ("**adaptive_GOP_size**") did not perform any motion search and compensation at the encoder to generate low quality Side Information. Also, it does not require any model for rate and distortion calculation. As described, our scheme assigned which frames should be intra-coded by just calculating difference from key frames and inner frames in GOP, and then based on it, an iterative algorithm fixes a final GOP size.

Consequently, there is not much calculation executed; thus encoding complexity is still kept very low. Fig. 6 illustrates that estimation of GOP size just occupies 4% in total encoding time. Such time usage for assigning GOP does not change much along with Qm increase, although encoding time is a variant function of Qm. This means that percentage of estimating GOP size even gets smaller when Qm gets closer to its maximum value.

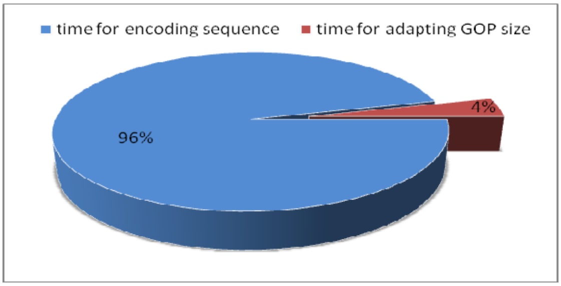
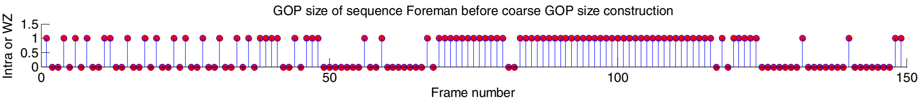


Fig. 6. Estimating GOP size time versus total encoding time (QM = 0, Hallmonitor)

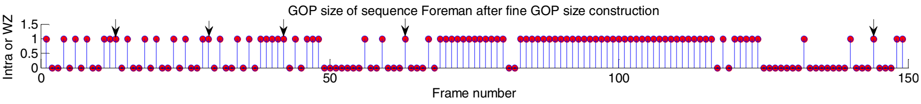
Coding efficiency performance

The proposed method includes two steps to assign the final GOP size. For the sequence Foreman, GOP sizes after the first step – coarse GOP size construction – are shown in Fig. 7 (a). There is local motion at the first half of the sequence so coarse GOP sizes are around two or three. But around the 100th frame, high global motion occurs hence GOP size becomes equal to one. In rest part of the sequence which has very slow motion, longer GOP sizes are chosen. After that, these GOP sizes were refined and results were depicted in Fig. 7 (b) which makes one observe that GOPs are broken into two or more smaller GOPs (see the down arrows which point exactly to the frames changing from WZ frames into Intra frames.) After the second steps, GOP sizes are significantly changed.

For evaluating efficiency of the proposed method, we compared the rate distortion (RD) performance in Fig. 8, 9, and 10. It is clear that RD performance is improved remarkably with about over 2 dB increment at sequence Hallmonitor or Soccer, and about 0.9 dB increment at sequence Foreman compared to SKKU results without adaptive GOP size [18]. However, when motion level increases, the gap between the simulation results and [13] decreases. In case of low motion level sequence, ratio of WZ frames should increase but it is limited because of fixed GOP size. This situation changes when motion level raises, most odd frames are encoded as Intra frames and it approaches the best GOP size. Indeed, our proposed method has shown that, the best



(a)



(b)

Fig. 7. Effect of the proposed iterative fine-size determination algorithm

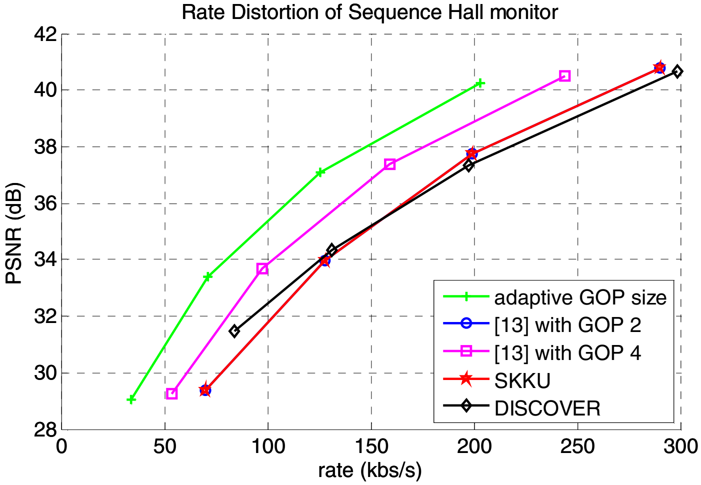


Fig. 8. Rate distortion performance of sequence Hall monitor

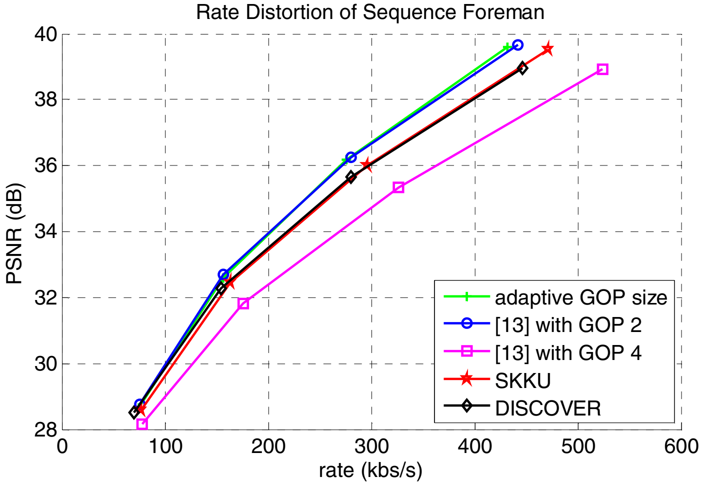


Fig. 9. Rate distortion performance of sequence Foreman

mode for sequence Soccer is almost Intra encoded. Simulation results also proved the stated analysis, and our proposed method is better than [13] by amounting to 2.2 dB in case of GOP equal to 2, and 1.2 dB with GOP equal to 4 when we examine the sequence Hallmonitor. In case of the sequence Soccer, the increment is about 0.2 dB against [13].

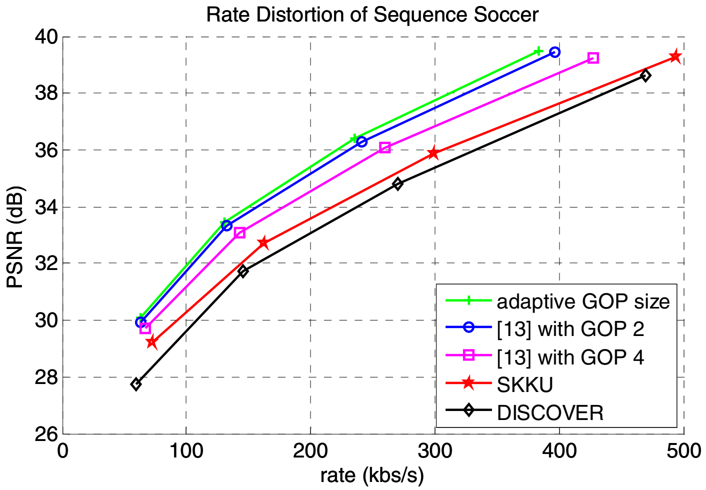


Fig. 10. Rate distortion performance of sequence Soccer

5 Conclusions and Future Works

This paper proposes an iterative algorithm for constructing GOP size in transform domain Wyner-Ziv video coding. This novel method is developed based on intra frame mode selection at frame level and iterative technique. Hierarchical structure is used for better results. Experimental results showed the superiority of the proposed method compared to previous works both in coding efficiency with improvement up to 2dB and low encoding complexity.

Because the simulation results still cannot reach the performance of H.264/AVC, there are much work remained to be done to reduce this gap. In the future works, an improved method for intra frame mode selection will be investigated to increase the accuracy of GOP size construction.

Acknowledgments. This work was supported by the IT R&D program of MKE/KEIT. [KI002142, Development of Ultra Low Complexity Video Coding Technique for Next-generation Mobile Video Service.

References

1. Wiegand, T., Sullivan, G.J., Bjontegaard, G., Luthra, A.: Overview of the H.264/AVC video coding standard. IEEE Trans. on Circuits and Systems for Video Technology 13, 560–576 (2003)
2. Slepian, D., Wolf, J.K.: Noiseless coding of correlated information sources. IEEE Trans. on Inform. Theory IT-19, 471–480 (1973)
3. Wyner, D., Ziv, J.: The rate-distortion function for source coding with side information at the decoder. IEEE Trans. on Inform. Theory IT-22, 1–10 (1976)

4. Aaron, A., Zhang, R., Girod, B.: Wyner-Ziv Coding of Motion Video. In: Asilomar Conference on Signals, Systems and Computers, Pacific Grove, CA, USA (2002)
5. Puri, R., Ramchandran, K.: PRISM: A New Robust Video Coding Architecture Based on Distributed Compression Principles. In: 40th Allerton Conference on Communication, Control and Computing, Allerton, USA (2002)
6. Ascenso, J., Brites, C., Pereira, F.: Motion Compensated Refinement for Low Complexity Pixel Based Distributed Video Coding. In: 4th Conference on Advanced Video and Signal Based Surveillance AVSS, Italy, (2005)
7. Brites, C., Ascenso, J., Pereira, F.: Studying Temporal Correlation Noise Modeling for Pixel Based Wyner-Ziv Video Coding. In: International Conference on Image Processing ICIP, USA (2006)
8. Ascenso, J., Pereira, F.: Hierarchical motion estimation for side information creation in Wyner-Ziv video coding. In: Proceedings of the 2nd International Conference on Ubiquitous Information Management and Communication (2008)
9. Pereira, F., Ascenso, J., Brites, C.: Studying the GOP Size Impact on the Performance of a Feedback Channel-Based Wyner-Ziv Video Codec. In: Mery, D., Rueda, L. (eds.) PSIVT 2007. LNCS, vol. 4872, pp. 801–815. Springer, Heidelberg (2007)
10. Ascenso, J., Brites, C., Pereira, F.: Content Adaptive Wyner-ZIV Video Coding Driven by Motion Activity. In: IEEE International Conference on Image Processing, Atlanta, GA (2006)
11. Yaacoub, C., Farah, J., Pesquet-Popescu, B.: New Adaptive Algorithms for GOP Size Control with Return Channel Suppression in Wyner-Ziv Video Coding. In: 16th IEEE International Conference on Image Processing (2009)
12. Aaron, A., Rane, S., Setton, E., Griod, B.: Transform-domain Wyner-Ziv Codec for Video. In: Proc. SPIE Visual Communications and Image Processing, San Jose, USA (2004)
13. Xiem, H.V., Park, J., Jeon, B.: Flexible Complexity Control based on Intra Frame Mode Decision in Distributed Video Coding. In: Proc. of IEEE Broadband Multimedia Systems and Broadcasting, Germany (2011)
14. Brites, C., Pereira, F.: Correlation Noise Modeling for Efficient Pixel and Transform Domain Wyner–Ziv Video Coding. IEEE Trans. on Circuits and Systems for Video Technology 18, 1177–1190 (2008)
15. Belkoura, Z., Sikora, T.: Towards rate-decoder complexity optimisation in turbo-coder based distributed video coding. In: Picture Coding Symposium, Beijing, China (2006)
16. Do, T., Shim, H.J., Jeon, B.: Motion linearity based skip decision for Wyner-Ziv coding. In: 2nd IEEE International Conference on Computer Science and Information Technology, China, (2009)
17. Ascenso, J., Pereira, F.: Low complexity intra mode selection for efficient distributed video coding. In: IEEE International Conference on Multimedia and Expo, USA (2009)
18. Park, J., Jeon, B., Wang, D., Vincent, A.: Wyner-Ziv video coding with region adaptive quantization and progressive channel noise modeling. In: Proc. of IEEE Broadband Multimedia Systems and Broadcasting, pp. 1–6 (2009)
19. DISCOVER codec, http://www.img.lx.it.pt/~discover/rd_qcif_15_gop2.html