DOI: 10.21917/ijivp.2014.0125

ADAPTIVE GOP STRUCTURE TO H.264/AVC BASED ON SCENE CHANGE

S. Sowmyayani[1] and P. Arockia Jansi Rani[2]
Department of Computer Science and Engineering, Manonmaniam Sundaranar University, India
E-mail: [1]sowmyayani@gmail.com, [2]jansi_msu@yahoo.co.in

Abstract

This paper proposes an adaptive GOP structure with the new logic of frame comparison in H.264/AVC to achieve better quality and reduce bit rate. Initially Group of Pictures (GOP) is set to a fixed size. Frames are compared within that GOP using correlation. According to the correlation, GOP is changed within that fixed size. So, there will be no GOP size greater than that fixed size. This method does not calculate any threshold. Hence the time needed to calculate global or local threshold is eliminated. It is integrated with conventional video codec H.264/AVC. This method is compared with H.264/AVC of fixed GOP structure of sizes 4, 8, 12, 16, 32 and GOP structure with the length of entire video. The proposed method achieved gain in bit rate from 0.49% to 69.75% and PSNR gain from 2.5% to 0.3%.

Keywords:
H.264/AVC, GOP, Correlation

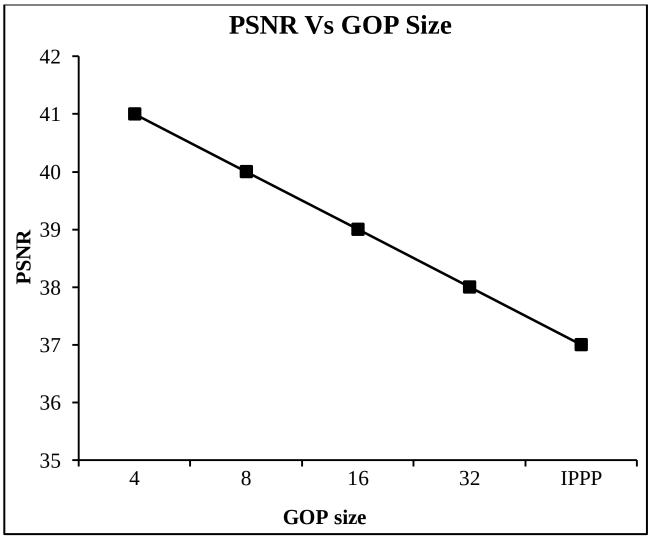
1. INTRODUCTION

In the multimedia technology, digital video is used widespread. To satisfy the needs of the user in terms of transmission speed, quality and available resources, the efficiency of video coding became very important task. The most popular video coding standard is H.264/AVC [1], [2] which is widely used in digital broadcasting. In H.264/AVC, there are several features which satisfy the need of the user. Features include motion estimation, motion prediction, multiple reference frames, variable block sizes etc. [3].

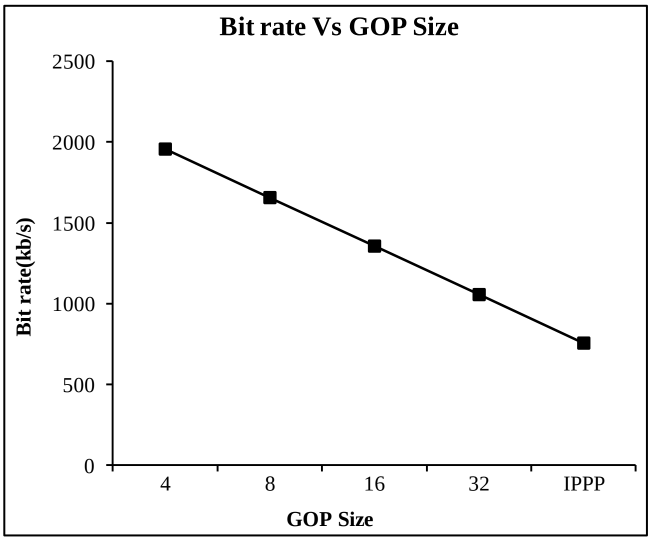
Video sequences have two types of redundancy namely spatial and temporal redundancy [4]. Encoder should eliminate those redundant parts in the video to achieve highest bit rate. For spatial redundancy, Intra prediction coding is used. For temporal redundancy, Inter prediction coding is used conventionally. For Inter Prediction Coding technique, inter frames are compared to avoid redundancy between frames. Hence in H.264/AVC, video sequence is divided into Group of Pictures (GOP). If GOP size is 4, arrangement of pictures will be “IBBPBBP”. In the GOP structure, three different picture types I-, P- and B- frames are used. I-frames use intra prediction. P-frame is predictive coded picture with inter prediction from previous I or P picture. B-frame is bidirectional coded picture with inter prediction from previous I- or P- frame and future P-frame. Since I and P picture are coded from other reference frames, they are called core pictures.

In H.264 video codec, the arrangement of picture types is arranged in periodic form. Video sequence is divided into fixed group of pictures (GOP). There are several methods to decide a GOP structure [5, 6]. The optimal number and positions of P frames in a GOP from possible candidates are searched by solving minimization problem with Legrange multiplier method [5]. This method is optimal, but calculation cost is high. The method in [6] derives a coding gain which is a function with an inter-frame correlation coefficient. This method requires pre-analysis to calculate an inter-frame correlation and it also has to search the best GOP.

Fixed GOP structures prevent encoders to improve coding efficiency for B frames, since P frames are not chosen according to the motion. In video sequence, scene changes from frame to frame. If GOP is fixed, it leads to poor performance. If two frames changed and frames are coded using inter frames, it will become inefficient. It is necessary to adapt the size of GOP structure.



(a)



(b)

Fig.1(a). Comparison of GOP size and PSNR, (b). Comparison of GOP size and Bit rate

Adaptive GOP structure is implemented in H.264/AVC [7]. Each frame is compared with its motion compensated frame which takes more time. And, only I and p frames are used in [7] which leads to higher bit rate. B frames are used to improve coding efficiency of temporal redundancy. If GOP size is large, quality of the video after decoding will be low and bit rate will be reduced. If GOP size is small, quality of video increases and bit rate also increases. The relationship between GOP size, PSNR and GOP size, bit rate is shown in Fig.1. I and P pictures are used for motion compensation for B frames. B frames should be decoded with high quality. If P pictures are selected according to the motion in the video, blocks for B frames are matched with most approximate blocks from P frames. Hence, if the GOP is adaptive to motion in the scene, coding efficiency will be improved.

This paper proposes a new logic for comparison of frames. P-frames are identified based on change in scene. Scene changes are calculated using Pearson Correlation Coefficient. Unlike conventional frame comparison, no consecutive frames are compared. This method is integrated with H.264/AVC. Finally results are compared with fixed size of GOP in H.264/AVC.

The rest of the paper is organized as follows: Section 2 describes overall architecture of the system. Section 3 illustrates the algorithm of frame comparison for adaptive GOP structure. Section 4 tells how H.264 is integrated with the system. Section 5 demonstrates the experimental results of SVC followed by conclusion in section 6.

2. SYSTEM ARCHITECTURE

The overall System architecture is shown in Fig.2. The input video sequence is given to encoder for encoding I^{th} Frame. GOP Size is determined by the Adaptive GOP algorithm from the input sequence during encoding process. Adaptive GOP algorithm sends side information to the decoder which contains GOP size for the entire video sequence. H.264 Decoder uses this side information to decode the video sequence.

3. ADAPTIVE GOP STRUCTURE

Scene changes from frame to frame in a video sequence. But all the changes are not visually identified. The logic of frame comparison is illustrated in Fig.3. In this paper, no two consecutive frames are compared. Initially, GOP size is fixed to 7. This is because, if the GOP size is set to adaptive, there is a chance to get larger size of GOP which leads to lower PSNR. This method will have adaptive GOP size but not greater than 7.

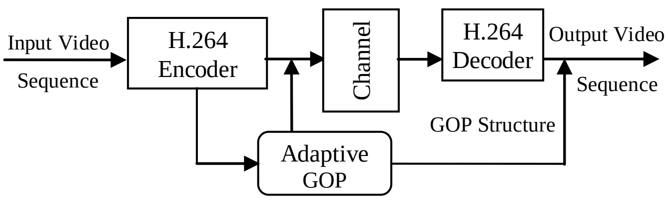


Fig.2. System Architecture

For evaluating the similarity of frame and its prediction, Pearson Correlation Coefficient is chosen. Pearson correlation coefficient is widely used to measure the similarity of two frames for cut detection [8]. The value of Pearson Correlation Coefficient can fall between 0(no correlation) and 1(perfect correlation). Correlations above 0.80 are considered as really high and lowest values will be determined as cuts [8]. The Pearson correlation coefficient for 2D signals like video sequences is expressed as follows [8].

$$\mathrm{PCC} = \frac{\sum_{i=1}^{M}\sum_{j=1}^{N}\left(f(i,j)-f^{m}\right)\left(f_p(i,j)-f_p^{m}\right)}{\sqrt{\sum_{i=1}^{M}\sum_{j=1}^{N}\left(f(i,j)-f^{m}\right)^2\left(f_p(i,j)-f_p^{m}\right)^2}} \quad (1)$$

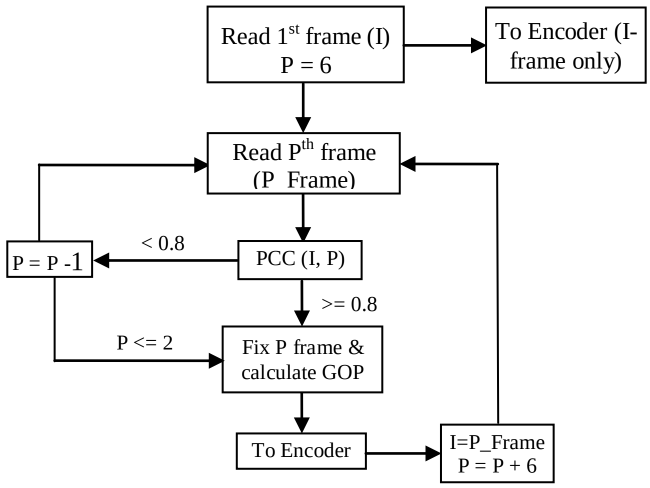


Fig.3. Frame comparison process

In the initial GOP, 1^{st} frame is compared with last frame in GOP. If the correlation between them is high, last frame is selected as P frame and GOP is set to 7. If the correlation is low, then the previous frame in the GOP is compared with 1^{st} frame. This process is repeated till we reach the frame next to 1^{st} frame in the GOP. After setting GOP size, it is given to encoder. Next 7 frames are processed again starting with the one selected in the previous process. This process is repeated till end of the video sequence is reached. Fig.4 illustrates the process by which the GOP is fixed adaptively. At each step, the correlation factor given in Eq.(1) is computed.

4. INTEGRATION WITH H.264

This method is integrated with H.264/AVC JM 11.0 reference software five conditions of H.264/AVC [9]. The process of H.264 video codec is explained in [1]. Following conditions of H.264/AVC encoder were performed for each simulation:

- QP (quantization parameter): 24
- Total number of references: 2
- Search range: 16
- Entropy coding method: CABAC

Adaptive GOP size is stored during encoding. During decoding, GOP size is retrieved for further processing.

5. EXPERIMENTAL RESULTS

Adaptive GOP structure results are compared with fixed GOP. Results are displayed in Fig.5. It is Suzie sequence which has 2 scenes: Suzie having receiver near her ear and put off the phone. In this sequence, scene change happens when the receiver is moved away. Fig.5 shows how fixed GOP and adaptive GOP algorithms work.

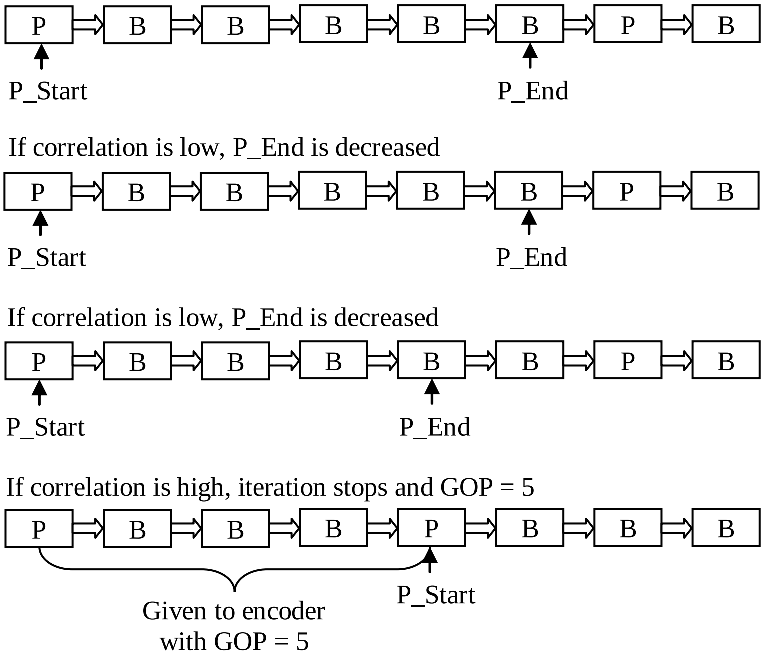


Fig.4(a)

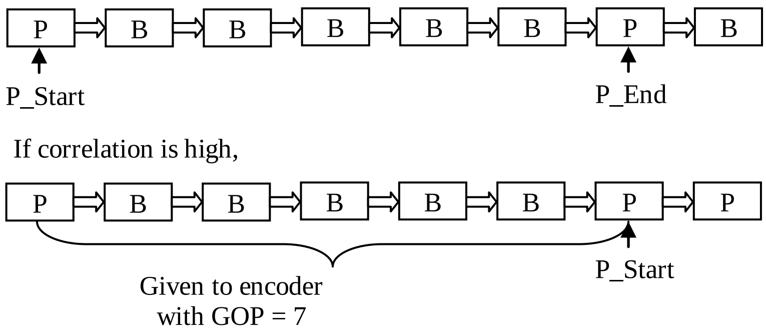


Fig.4(b)

Fig.4. Logic of frame comparison: (a). GOP less than 7, (b). GOP equals 7

Video sequence	**Fixed GOP structure**	**Adaptive GOP structure**
	I	I
	B	B
	B	B
	P	B
	B	P
	B	B
	P	P
	B	B
	B	P
	P	B

Fig.5. Results of Fixed GOP structure and Adaptive GOP structure

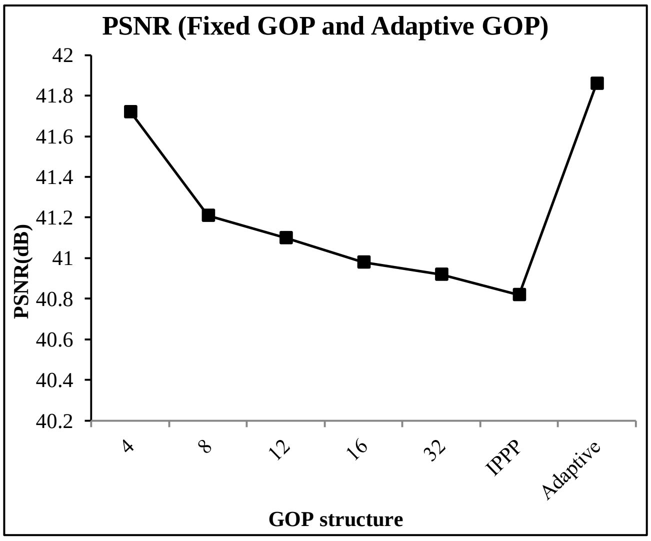
PSNR gain and Bit rate gain are compared with the situation, when whole sequence is coded with only one I frame and the rest are P frames, (IPPP...). With effort to provide similar comparison as in [10], adaptive GOP was additionally compared with fixed GOP structures of size 4, 8, 12, 16 and 32. From the observation, PSNR for fixed GOP of size 4 is high of all fixed size. Proposed method's PSNR is higher than fixed size GOP of size 4. Bit rate is less in IPPP GOP structure when compared with all other fixed GOP structure. Proposed method achieves higher bit rate than that. The results obtained during simulations are shown in Table.1.

The highest PSNR and the smallest bit rate were achieved by proposed method of adaptive GOP structure. The improvement is more significant for bit rate. Time taken for fixing adaptive GOP size is similar to fixed GOP size.

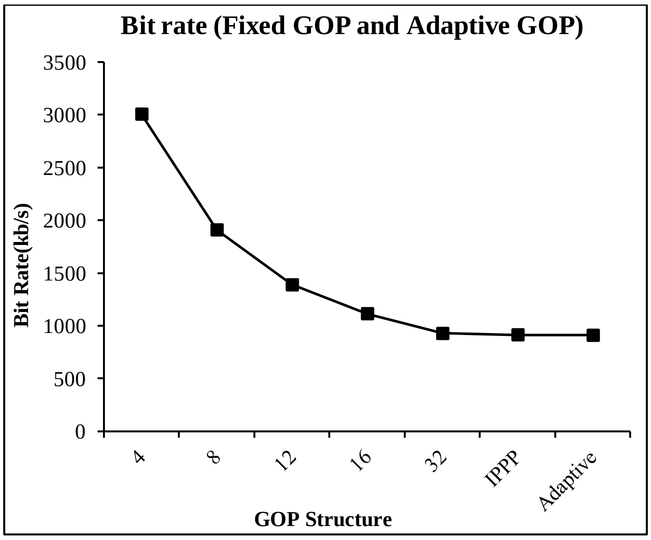
The Fig.6 shows the results obtained by fixed GOP structure and Adaptive GOP structure in charts. From the figure, it is evident that PSNR is very high when GOP size is 4. But the proposed method yields higher PSNR than that. And, bit rate is very well reduced in case of 'IPPP' GOP structure. The proposed method achieves still lesser bit rate.

Table.1. PSNR and Bit rate achieved by fixed and adaptive GOP structure

GOP	PSNR(dB)	Bit rate(kb/s)
4	41.72	3006.87
8	41.21	1909.18
12	41.1	1388.23
16	40.98	1114.35
32	40.92	927.08
IPPP	40.82	914.14
Adaptive	41.86	909.6



(a)



(b)

Fig.6. Charts comparing (a). PSNR for fixed GOP structure and Adaptive GOP structure, (b). Bit rate for fixed GOP structure and Adaptive GOP structure

The Table.2 shows gain in PSNR and bit rate of Adaptive GOP structure in comparison to fixed GOP. The gain is calculated from the PSNR and bit rate values for fixed and adaptive GOP sizes in Table.1. Proposed method achieves PSNR gain from 2.5% to 0.3% and reduction in bit rate from 0.49% to 69.75%

Table.2. PSNR and Bit rate Gain achieved by proposed method

GOP	PSNR Gain (%)	Bit rate Gain (%)
4	0.33557	69.75
8	1.577287	52.36
12	1.849148	34.48
16	2.147389	18.37
32	2.297165	1.89
IPPP	2.547771	0.49

6. CONCLUSION

In this paper, an adaptive GOP structure is used to improve coding efficiency of H.264/AVC while maintaining the quality of the video sequence. The novelty of this method is in using a method of frames comparison and in adapting GOP structure to the video characteristics. There are many advantages in this method. First, there is no consecutive frames comparison which leads to reduction in time. There is no pre-processing such as threshold calculation. Third, this method avoids GOP of larger length (more than 7), which achieves better quality. Fourth, it uses B-frames which are not used in [7]. This achieves higher bit rate.

Experimental results prove the efficiency of this method in comparison with fixed GOP structures of sizes 4, 8, 12, 16, 32 and length of the whole video sequence. Bit rate reduction is obtained by adaptive GOP differs from 0.49% to 69.75% while providing PSNR gain from 2.5% to 0.3%.

In further research, the influence of various quantization parameters can be examined.

REFERENCES

[1] T. Wiegand, G. J. Sullivan, G. Bjontegard and A. Luthra, “Overview of the H.264/AVC video coding standard”, *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 13, No. 7, pp. 560-576, 2003.

[2] JVT Editors, “Draft ITUT Recommendation and final draft international standard of joint video specification (ITU-T Rec.H.264 |ISO/IEC 14496-10 AVC)”, JVT-G050r1, 2003.

[3] T. Sikora, “Trends and perspectives in image and video coding”, *Proceedings of IEEE*, Vol. 93, No. 1, pp. 6-17, 2005.

[4] G. J. Sullivan and T. Wiegand, “Video compression - from concepts to the H.264/AVC standard”, *Proceedings of the IEEE*, Vol. 93, pp. 18-31, 2005.

[5] Jungwoo Lee and B.W. Dickinson, “Rate-Distortion Optimized Frame Type Selection for MPEG Encoding”, *IEEE Transactions on Circuits and Systems for Video Technology*, Vo1. 7, No. 3, pp. 501-510, 1997.

[6] Jiro Katto and M. Ohta, “Mathematical analysis of MPEG compression capability and its application to rate control”, *Proceedings of International Conference on Image Processing*, Vol. 2, pp. 555-558, 1995.

[7] Lenka Krulikovská, Jaroslav Polec and Michal Martinovič, “Adaptive Group of Pictures Structure Based On the Positions of Video Cuts”, *World Academy of Science, Engineering and Technology International Journal of*

Computer, Information, Systems and Control Engineering, Vol. 7, No. 7, pp. 319-322, 2013.

[8] Lenka Krulikovsk´a and Jaroslav Polec, "GOP Structure Adaptable to the Location of Shot Cuts", *International Journal of Electronics and Telecommunications*, Vol. 58, No. 2, pp. 129-134, 2012.

[9] Joint Video Team (JVT) of ISO/IEC MPEG & ITU-T VCEG (ISO/IEC JTC1/SC29/WG11 and ITU-T SG16 Q.6), "H.264/14496-10 AVC Reference Software Manual", 2009.

[10] B. Zatt, M. Porto, J. Scharcanski and S. Bampi, "Gop structure adaptive to the video content for efficient H.264/AVC encoding", *Proceedings of 17th IEEE International Conference on Image Processing*, pp. 3053-3056, 2010.