Research Article

Adaptive Rate Control Algorithm for H.264/AVC Considering Scene Change

Xiao Chen and Feifei Lu

School of Electronic and Information Engineering, Nanjing University of Information Science and Technology, Nanjing 210044, China

Correspondence should be addressed to Xiao Chen; rainofsun@netease.com

Received 2 November 2012; Revised 1 January 2013; Accepted 1 January 2013

Academic Editor: Ming Li

Copyright © 2013 X. Chen and F. Lu. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Scene change in H.264 video sequences has significant impact on the video communication quality. This paper presents a novel adaptive rate control algorithm with little additional calculation for H.264/AVC based on the scene change expression. According to the frame complexity quotiety, we define a scene change factor. It is used to allocate bits for each frame adaptively. Experimental results show that it can handle the scene change effectively. Our algorithm, in comparison to the JVT-G012 algorithm, reduces rate error and improves average peak signal-noise ratio with smaller deviation. It cannot only control bit rate accurately, but also get better video quality with the lower encoder buffer fullness to improve the quality of service.

1. Introduction

The multimedia applications are becoming popular on the Internet. The quality of service (QoS) in terms of end-to-end delay guarantees to real-time applications is especially important for the new generation of Internet applications such as video on demand and other consumer services [1–7]. Elements of network performance within the scope of QoS often include availability, bandwidth, delay, and error rate.

QoS involves prioritization of network traffic. QoS can be targeted at the network interface or in terms of specific applications. In order to make the video stream well adapted for the time delay and the network sources such as the bandwidth and the buffer, especially for the low bandwidth or time-varying wireless channel, two technologies, the traffic shaping and the rate control, are developed. The traffic shaping technology belongs to the transport layer method to improve the QoS. There are two categories of approaches for guaranteeing end-to-end performances. One is bounded modeling of frames and the other stochastic modeling of frames [8–11]. For the bounded modeling based approaches, the fundamental is network calculus [11–17]. For the stochastic modeling, the fractal time series is essential [17–22]. Li et al. provides a Holder exponent to describe the fractal time series [21] and use it to investigate the scaling phenomena of

traffic data [22]. While the rate control technology belongs to the compression layer method, it compresses the original video sequences according to the needs of the application and the available bandwidth. The emerging Internet video streaming media transmission, wireless channel transmission, the MPEG-4 object encoding transmission, and transmission of the actual application require the high efficiency rate control algorithm to meet the needs of real-time video transmission.

As a new international video compression standard for IP and the wireless communication, H.264 has not only absorbed the advantages of the entire previous coding schemes, but also focuses on the current advanced coding techniques. After being promulgated in 2003 formally, H.264 has elicited a wide range of interests in industrial and academic fields [23].

The rate control is an essential part of H.264. The main task of rate control is to allocate a certain number of bits for each frame in the purpose of controlling the output rate to adapt to the current bandwidth and minimizing the image distortion. It can effectively avoid the bit error and the packet loss caused by the excessive congestion in the real-time data transmission. But in the rate control algorithms for H.264/AVC, the quantization parameter is used in both the rate control and rate distortion optimization (RDO), which leads to chicken and egg dilemma [24]. In order to solve

this dilemma, many scholars have done a lot of research [25–28]. The work in [25] solved the dilemma by enhancing the ρ domain model. The relational model between rate and quantization is advanced for the dilemma [26]. JVT-G012 algorithm proposes a linear model for MAD prediction to solve the chicken and egg dilemma [29]. This method can obtain good coding result and solve the dilemma well. When there is no scene change in video sequence, JVT-G012 algorithm has good performance. However, the video quality would have a serious decline in the situation of scene change. There are two main reasons. On the one hand, it uses a fixed length of group-of-pictures (GOP) structure, which can not effectively detect scene change in video sequence. On the other hand, it is mainly based on the linear model to determine the allocation of coding bits and quantization parameter. When a scene change happens, the predicted MAD has larger deviation leading to the serious decline in coding quality after scene change.

Aiming at the problem of scene change, there are a variety of rate control algorithms, including gray value detection algorithm [30], intramode macro block statistics algorithm [31], motion searching detection-based algorithm [32], and edge detection algorithm [33], and so forth. The differences between those methods can be summarized in two main areas—how to detect a scene change and how to deal with the scene change. Edge detection algorithm has good performance, but it uses computer image recognition technology. So the algorithm is very complicated and this greatly limited its application. Intramode macro block statistics algorithm and motion searching detection-based algorithm need to compile a code on the current frame for the second coding sequence. Gray value detection algorithm is based on absolute difference and can better reflect the degree of scene changes. But the gray absolute difference would be very large when it comes to global motion or the image is not strong within the relevant contents. So it does not reflect the true complexity of this case.

Therefore, this paper proposes a novel adaptive rate control algorithm with little additional calculation for H.264/AVC based on scene change expression. We use a scene change factor to adjust bit allocation adaptively for every frame in video sequence. Experimental results show that our method can effectively improve the video quality in the situation of scene changes.

2. Effect of Scene Change on Coding Performance

When the scene change occurs in video sequence, the temporal correlation between images disappears or diminishes, which has a great impact on internal prediction. The encoding quality will be affected and the impact depends on the scene change frame position in the GOP. There are three types of frames—I, P, and B frames. I frame uses intraframe coding, P frame uses one-way prediction interframe coding, and B frame uses bidirectional prediction interframe coding. The following analyzed the impact of scene change on the coding performance when it occurs at three different types of frames, respectively.

When scene change occurs in I frame, because I frame uses intra-frame coding without reference to other frames, the subsequent P or B frames can be normally encoding. Therefore, the scene change in I frame has no impact on coding performance at all. When the scene change occurs in B frames, this B frame and the first subsequent P or I frame are in the same scene because B frame is bidirectional prediction frame. I frame is intraprediction frame and always has a good coding results; the first P frame after scene change has a good coding effect. So the current B frame would get a better prediction. Therefore, we also do not need to do any processing. When the scene change occurs in P frame, because P frame is forward predictive coding frame, the current P frame is predicted according to the previous P or I frame. At the same time, the current P frame may also be as the prediction reference frame of the next P or B frame in the current GOP. Therefore, the scene change occurred in the P frame has great impact on the image quality of the current P frame and subsequent frames. Since the current P frame and its reference frame are in different scenes, the interprediction coding is completely invalid. Macro block must perform RDO before taking intracoding mode selection. The optimization process costs a large number of coding time. In addition, most macro blocks uses intracoding mode taking a lot of bit rate resources; buffer fullness will increase dramatically, resulting in significant decrease in the bit rate distribution and image quality of the follow-up frames. This impact is likely to continue an all the following frames in the GOP.

From the above analysis, the scene change occurred in I frame and B frame has little effect on the coding performance, but great impact on P frame. Therefore, only the scene change in P frame is necessary to be considered.

3. Proposed Adaptive Rate Control Algorithm

Detecting the scene change in video sequence is not required in scene adaptive methods. Instead, the relative change between adjacent frames is considered. It is not necessary to change the GOP structure and to determine whether scene change occurs. It avoids missing scene change and miscarriage of justice.

The JVT-G012 rate control algorithm allocates the bits for nonencoded frames on average. When a scene change happens, the predicted MAD has larger deviation leading to the serious decline in coding quality after scene change. Thus we propose a scene adaptive method to resolve the above-mentioned shortcomings. The implementation of rate control mainly includes bit allocation, calculations of quantization parameter, and buffer control. Our adaptive rate control algorithm consists of three steps. First, according to the remaining bits and scene change factor, we allocate the target bits in frame layer. Then we calculate the quantization parameter. Last, we perform RDO and update the model parameters.

3.1. Scene Change Factor. Now most researchers use MAD_{ratio} to indicate the frame complexity measurement. However,

FIGURE 1: scf change of suzie-football sequence.

MAD ratio is the ratio of the predicted MAD of current frame to the average MAD of all previously encoded P frames in the GOP; when the current frame has scene change, the predicted MAD of the current frame fails. So our frame complexity measurement is represented by frame complexity quotiety FC_j as follows [34, 35]:

$$FC_j = \mu \text{MAD}_{\text{ratio}}(i, j) + (1 - \mu) C_j, \quad (1)$$

where $\text{MAD}_{\text{ratio}}(i, j)$ is the ratio of predicted MAD of the j th frame in the i th GOP to the average MAD of all encoded P frames in the i th GOP, and j is the frame number. $C_j = \sqrt{H_j} / \sqrt{H_{j-1}}$. H_j is the average difference of gray histogram of the j th frame in the i th GOP [13]. μ is the weighting coefficient. According to the frame complexity quotiety, we propose a scene change factor scf as follows:

$$\text{scf}_j = \frac{\lambda * FC_j + FC_{j-1}}{FC_j + \lambda * FC_{j-1}}, \quad (2)$$

where λ is the weighting coefficient.

In Figure 1, the scene change factor can reflect the complexity of the image. When scene change occurs, scf increases sharply and its value is more than 2. In this figure, the values of scf which belong to football sequence are significantly greater than the values which belong to suzie sequence. We can see that the combination frame complexity measure we proposed is reasonable. When scf is bigger than 2, there should be scene change. Therefore, we can more effectively allocate the target bits according to scf.

3.2. Bit Allocation. The target bits T_j allocated for the j th frame in the i th GOP are determined by residual bits, frame

rate, target buffer size, actual buffer occupancy, and the available channel bandwidth:

$$T_j = \beta T_r + (1 - \beta) \times \left\{ \frac{u(n_{i,j})}{F_r} + \gamma [T_{bl}(n_{i,j}) - B_c(n_{i,j})] \right\}, \quad (3)$$

where F_r is the predefined frame rate, $u(n_{i,j})$ is the available channel bandwidth for the sequence, $T_{bl}(n_{i,j})$ is the target buffer level, and $B_c(n_{i,j})$ is the occupancy of virtual buffer. β is a constant and it is 0.5 when there is no B frame and 0.9 otherwise. γ is a constant and it is 0.75 when there is no B frame and 0.25 otherwise. T_r is contained by the formula

$$T_r = \begin{cases} T'_r, & \text{scf} < 0.5, \\ T'_r * 0.05, & 0.5 \leq \text{scf} < 0.9, \\ T'_r * 0.4 * \text{scf}, & 0.9 \leq \text{scf} < 1.0, \\ T'_r * 0.5 * \text{scf}, & 1.0 \leq \text{scf} < 2, \\ T'_r * 2, & \text{scf} \geq 2, \end{cases} \quad (4)$$

where $T'_r = T_r(n_{i,j})/N_r$, $T_r(n_{i,j})$ is the residual bit of all unencoded frames in the i th GOP, and N_r is the number of unencoded P frames.

When scf is much larger than 2, which indicates that great changes have taken place in the scene, macro block uses intra-frame coding mode and a large number of bits are allocated for this frame. When scf is between 0.9 and 2, scene changes not much. In order to supplement bits cost by the big scene change frame, the allocated bits for these frames are reduced lightly. When scf is between 0.5 and 0.9, the scene changes very little. The allocated bits for these frames are reduced significantly. When scf is very small, it is usually located behind a large scene change frame. Because the former one uses too many bits, the bits assigned to this frame in the algorithm are not much, so there is no adjustment to this frame.

4. Experimental Results

We have performed our proposed rate control algorithm by enhancing the JM8.6 test model software. The JVT-G012 rate control method is selected as a reference for comparison (as is implemented on reference software JM8.6) using different test sequences under various target bit rates. The combined test sequences used are in QCIF4:2:0 formats: suzie-football, foreman-mobile, bus-coastguard-news, foreman-coastguard-news, and football-foreman-mobile-suzie. In the experiments, the frame rate is set to 15 frames per second; that is, f/s; the target bit rate is various; the total number of frames is set to 100 or 150; the initial quantization parameter is set to 28; and the length of GOP is set to 100.

The experimental results are shown in Table 1. Our proposed rate control algorithm outperforms JVT-G012 under various target bit rates for different video sequences. Our proposed rate control algorithm can control the bit rates under various target bit rates for different video sequences accurately. The average error of the actual bits is reduced

TABLE 1: Performance comparisons of proposed and JVT-G012 algorithm.

Sequence	Target	JVT-G012			Proposed		
	Rate (kbits)	Bit rate (kbits)	PSNR (dB)	PSNR deviation	Bit rate (kbits)	PSNR gain (dB)	PSNR deviation
Suzie-football	32	33.44	25.72	7.32	32.85	0.15	7.16
Suzie-football	64	64.80	27.58	7.39	64.40	0.44	7.19
Foreman-mobile	32	32.33	29.25	3.86	32.15	0.29	3.64
Foreman-mobile	64	65.29	30.02	5.9	65.50	0.46	4.34
Foreman-silent-news	48	47.98	36.50	1.11	48.18	0.18	1.15
Foreman-silent-news	64	64.02	37.97	1.53	63.93	0.10	1.59
Bus-coastguard-news	32	32.00	26.43	5.57	31.86	0.12	5.48
Bus-coastguard-news	64	64.00	29.68	5.98	64.17	0.12	6.02
Football-foreman-mobile-suzie	64	64.59	27.50	4.32	63.99	0.26	4.22
Football-foreman-mobile-suzie	96	95.70	29.45	4.93	95.34	0.20	4.85
Football-foreman-mobile-suzie	128	128.01	31.02	5.39	128.10	0.08	5.56

TABLE 2: PSNR comparisons (before and after scene change).

Sequence	Rate	JVT-G012			Proposed		
		Before	After	Decrease	Before	After	Decrease
Suzie-football	32	37.47	23.58	13.89	37.23	23.89	13.34
Suzie-football	64	39.44	26.27	13.17	39.37	26.70	12.67
Foreman-mobile	32	34.58	27.06	7.52	35.18	29.12	6.06
Foreman-mobile	64	34.50	24.07	10.43	36.74	28.71	8.03

FIGURE 2: PSNR curves of suzie-football sequence comparing JVT-G012 and proposed algorithms (64 kbps).

FIGURE 3: PSNR curves of foreman-mobile sequence comparing JVT-G012 and proposed algorithms (64 kbps).

compared with that of the JVT-G012 algorithm. The more accurate bit rate avoids the bit error and the packet loss (jumped frame) caused by the excessive congestion in the real-time video transmission.

The proposed algorithm also improves the average peak signal noise-ratio (PSNR) and PSNR deviation for the sequences. In Table 1, it shows that our method achieves an average PSNR gain of about 0.22 dB with similar or lower PSNR deviation as compared to the JVT-G012 algorithm.

The maximum of PSNR deviation decrease is 26.44% compared with the original algorithm. The proposed algorithm obtains lower average PSNR deviation by 2.84% compared with the JVT-G012 algorithm. This shows the proposed algorithm can smooth the PSNR fluctuation between frames to some extent.

Table 2 shows the average PSNR comparisons of our proposed algorithm in some test sequences after the scene

FIGURE 4: PSNR of football-foreman-mobile-suzie sequence comparing JVT-G012 and proposed algorithms (64 kbps).

FIGURE 5: Buffer fullness comparison of football-foreman-mobile-suzie sequence of JVT-G012 and proposed algorithms (64 kbps).

change occurs at different bit rates. Although the PSNR declined, it is shown that the decrease is reduced compared with the JVT-G012 algorithm.

Although the video quality caused by scene change is inevitable, a slight decrease in most frames can be used to replace the serious decline in some frames of video sequence to smooth video PSNR, thereby improving video quality. Figures 2–4 show frame by frame PSNR for the sequence suzie-football, foreman-mobile, and football-foreman-mobile-suzie with the comparison of the proposed algorithm with the JVT-G012 algorithm. For example, in Figure 3, that is, the comparisons of average PSNR for football-mobile sequence, there is a dramatic decline in PSNR in the JVT-G012 algorithm from 91 to 100th frames. In our proposed algorithm, the decline is reduced, but the PSNR from the 53rd to the 89th frame has a small decrease which compensates for the dramatic decline in PSNR from the 91st to the 100th frame.

We also make the comparisons of buffer fullness for the sequence in Figure 5. It shows that our proposed algorithm has less fluctuation in buffer fullness, especially for the frames of rich details, as compared to the JVT-G012 algorithm. Therefore, our algorithm achieves much steadier buffer fullness when compared to the JVT-G012 algorithm, which avoids the potential overflow. This implies the improved quality of service.

5. Conclusion

In this paper, we propose an adaptive frame layer rate control algorithm for H.264/AVC using the scene change factor. The algorithm allocates bits in frame layer according to the scene change factor. The experimental results show that our algorithm achieves accurate rate control and a better visual quality. The average PSNR is advanced by 0.22 dB. Our algorithm can improve the smoothness of the image quality. In addition, it avoids the potential overflow because of much steadier encoder buffer fullness. So the algorithm can improve the QoS on the Internet real-time video transmission under the H.264 standard.

Acknowledgments

This work was supported by Qing Lan Project and the National Natural Science Foundation of China (10904073).

References

- [1] A. Jamalipour and P. Lorenz, "End-to-end QoS support for IP and multimedia traffic in heterogeneous mobile networks," *Computer Communications*, vol. 29, no. 6, pp. 671–682, 2006.
- [2] S. H. Kang and A. Zakhori, "Effective bandwidth based scheduling for streaming media," *IEEE Transactions on Multimedia*, vol. 7, no. 6, pp. 1139–1148, 2005.
- [3] A. Mahanti, D. L. Eager, and M. K. Vernon, "Improving multirate congestion control using a TCP Vegas throughput model," *Computer Networks*, vol. 48, no. 2, pp. 113–136, 2005.
- [4] V. Sivaraman, F. M. Chiussi, and M. Gerla, "Deterministic end-to-end delay guarantees with rate controlled EDF scheduling," *Performance Evaluation*, vol. 63, no. 4-5, pp. 509–519, 2006.
- [5] Y. Jiang, "Per-domain packet scale rate guarantee for expedited forwarding," *IEEE/ACM Transactions on Networking*, vol. 14, no. 3, pp. 630–643, 2006.
- [6] P. Balaouras and I. Stavrakakis, "A self-adjusting rate adaptation scheme with good fairness and smoothness properties," *Computer Networks*, vol. 48, no. 6, pp. 829–855, 2005.
- [7] M. Chen, J. Zhang, M. N. Murthi, and K. Premaratne, "Delay-based TCP congestion avoidance: a network calculus interpretation and performance improvements," *Computer Networks*, vol. 53, no. 9, pp. 1319–1340, 2009.
- [8] V. Firoiu, J. Y. Le Boudec, D. Towsley, and Z. L. I. Zhang, "Theories and models for internet quality of service," *Proceedings of the IEEE*, vol. 90, no. 9, pp. 1565–1591, 2002.
- [9] J. Beran, R. Sherman, M. S. Taqqu, and W. Willinger, "Long-range dependence in variable-bit-rate video traffic," *IEEE Transactions on Communications*, vol. 43, no. 2, pp. 1566–1579, 1995.

- [10] M. Li and P. Borgnat, "Foreword to the special issue on traffic modeling, its computations and applications," *Telecommunication Systems*, vol. 43, no. 3-4, pp. 145-146, 2010.
- [11] H. Michiel and K. Laevens, "Teletraffic engineering in a broadband era," *Proceedings of the IEEE*, vol. 85, no. 12, pp. 2007-2032, 1997.
- [12] J.-Y. Le Boudec and P. Thiran, *Network Calculus*, vol. 2050 of *Lecture Notes in Computer Science*, Springer, Berlin, Germany, 2001.
- [13] Y.-M. Jiang and Y. Liu, *Stochastic Network Calculus*, Springer, 2008.
- [14] M. Li and W. Zhao, *Analysis of Min-Plus Algebra*, Nova Science, 2011.
- [15] K. Pyun, J. Song, and H. K. Lee, "The service curve service discipline for the rate-controlled EDF service discipline in variable-sized packet networks," *Computer Communications*, vol. 29, no. 18, pp. 3886-3899, 2006.
- [16] Y. Jiang, Q. Yin, Y. Liu, and S. Jiang, "Fundamental calculus on generalized stochastically bounded bursty traffic for communication networks," *Computer Networks*, vol. 53, no. 12, pp. 2011-2021, 2009.
- [17] M. Li and W. Zhao, "Representation of a stochastic traffic bound," *IEEE Transactions on Parallel and Distributed Systems*, vol. 21, no. 9, pp. 1368-1372, 2010.
- [18] M. Li and W. Zhao, "On $1/f$ noise," *Mathematical Problems in Engineering*, vol. 2012, Article ID 673648, 23 pages, 2012.
- [19] M. Li, W. Zhao, and B. Chen, "Heavy-tailed prediction error: a difficulty in predicting biomedical signals of $1/f$ noise type," *Computational and Mathematical Methods in Medicine*, vol. 2012, Article ID 291510, 5 pages, 2012.
- [20] M. Li and W. Zhao, "Visiting power laws in cyber-physical networking systems," *Mathematical Problems in Engineering*, vol. 2012, Article ID 302786, 13 pages, 2012.
- [21] M. Li, Y.-Q. Chen, J. Y. Li, and W. Zhao, "Hölder scales of sea level," *Mathematical Problems in Engineering*, vol. 2012, Article ID 863707, 22 pages, 2012.
- [22] M. Li, W. Zhao, and S. Chen, "mBm-based scalings of traffic propagated in internet," *Mathematical Problems in Engineering*, Article ID 389803, 21 pages, 2011.
- [23] X. Li, A. Hutter, and A. Kaup, "Efficient one-pass frame level rate control for H.264/AVC," *Journal of Visual Communication and Image Representation*, vol. 20, no. 8, pp. 585-594, 2009.
- [24] S. Ma, W. Gao, F. Wu, and Y. Lu, "Rate control for JVT video coding scheme with HRD considerations," in *Proceedings of the International Conference on Image Processing (ICIP '03)*, vol. 3, pp. 793-796, September 2003.
- [25] I. H. Shin, Y. L. Lee, and H. Park, "Rate control using linear rate- ρ model for H.264," *Signal Processing*, vol. 19, no. 4, pp. 341-352, 2004.
- [26] S. Ma, W. Gao, and Y. Lu, "Rate-distortion analysis for H.264/AVC video coding and its application to rate control," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 15, no. 12, pp. 1533-1543, 2005.
- [27] M. Jiang, X. Yi, and N. Ling, "Improved frame-layer rate control for H.264 using mad ratio," in *Proceedings of the IEEE International Symposium on Circuits and Systems (ISCAS '04)*, vol. 3, pp. III813-III816, May 2004.
- [28] Z. G. Li, W. Gao, F. Pan et al., "Adaptive rate control for H.264," *Journal of Visual Communication and Image Representation*, vol. 17, no. 2, pp. 376-406, 2006.
- [29] Z. Li, F. Pan, and K. Lim, "Adaptive basic unit layer rate control for JVT," in *Proceedings of the 7th JVT Meeting*, Pattaya II, Thailand, March 2003.
- [30] W. A. C. Fernando, C. N. Canagarajah, and D. R. Bull, "Fade-in and fade-out detection in video sequences using histograms," in *Proceedings of the IEEE International Symposium on Circuits and Systems (ISCAS '00)*, vol. 4, pp. 709-712, May 2000.
- [31] D. Lelescu and D. Schonfeld, "Statistical sequential analysis for real-time video scene change detection on compressed multimedia bitstream," *IEEE Transactions on Multimedia*, vol. 5, no. 1, pp. 106-117, 2003.
- [32] R. Lienhart, "Reliable transition detection in videos: a survey and practitioner's guide," *International Journal of Image and Graphics*, vol. 1, pp. 469-486, 2001.
- [33] M. Sharifi, M. Fathy, and M. T. Mahmoudi, "A classified and comparative study of edge detection algorithms," in *Proceedings of the IEEE International conference on Information Technology: Coding and Computing (ITCC '02)*, pp. 117-120, 2002.
- [34] X. Chen and F. Lu, "A reformative frame layer rate control algorithm for H.264," *IEEE Transactions on Consumer Electronics*, vol. 56, no. 4, pp. 2806-2811, 2010.
- [35] X. Chen and F. Lu, "An improved rate control scheme for H.264 based on frame complexity estimation," *Journal of Convergence Information Technology*, vol. 5, no. 10, pp. 117-123, 2010.