Research Article

Basic Unit Layer Rate Control Algorithm for H.264 Based on Human Visual System

Xiao Chen^{1,2} and Haiying Liu²

¹ Jiangsu Key Laboratory of Meteorological observation and Information Processing, Nanjing University of Information Science and Technology, Nanjing 210044, China

² School of Electronic and Information Engineering, Nanjing University of Information Science and Technology, Nanjing 210044, China

Correspondence should be addressed to Xiao Chen; chenxiao@nuist.edu.cn

Received 26 May 2013; Accepted 31 July 2013

Academic Editor: Zhongmei Zhou

Copyright © 2013 X. Chen and H. Liu. 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.

In the process of the video coding, special attention should be paid to the subjective quality of the image. In the JVT-G012 algorithm for H.264, the influence of the human visual characteristic in basic unit layer rate control was not taken into account. This paper takes the influence of the human visual characteristic into the full consideration and offers ways to improve the subjective quality of the image. The visual characteristic factor, which is constituted by the motion feature and edge feature, is used to reasonably allocate the target bits, and then its quantization parameter is adjusted by encoded frame information. The experimental results show that, in comparison to the original algorithm, the proposed algorithm can not only control the bit rate more accurately but also make the peak signal to noise ratio (PSNR) stable, so as to improve the stationarity of the video image. The subjective quality of the reconstructed video is more satisfying.

1. Introduction

With the rapid development and popularization of the internet, electronic devices gradually become an indispensable part of our daily life, such as online broadcasting, online advertising, e-commerce, VOD, distance education, telemedicine, real-time video conference, smart phones, 3D video [1], VCD, DVD, HDTV, and streaming of multimedia video [2]. However, the real-time data transmission and storage of multimedia data have become more difficult due to the limited communication bandwidth, especially in the video communication. Its high capacity of data has much difficulty in the process of transmission and storage. Thus with the limits of the bandwidth and the low storage capacity, the video coding aiming at using the least bits to represent image is very important.

The rate control plays an important role in the process of video coding. Since 70% of the information is obtained from eyes and the eyes is the final receiver of video information, it is vital important to take full advantage of the human visual characteristics to get a higher subjective quality of images.

According to the special structure of the human eyes, some scholars put forward some relevant algorithms [3–10]. An adaptive bit allocation method is presented in [3], based on the space and time perception functions. The work in [4, 5] presents a method based on the region of interest in the human eyes. A novel rate control algorithm is presented in [6], based on the visual perception characteristics. The work in [7] proposes a new digital video watermark method based on the human visual system (HVS). The work in [8] proposes a method based on the region of interest, aiming at distribution of the target bits. The work in [9] presents a video quality evaluation method based on the region of interest. The work in [10] presents an algorithm to distribute the target bits of the basic unit layer, which analyze the motion information and texture features.

Although the JVT-G012 is by now the most acceptable rate control algorithm, it still has shortcomings. The work in [11] proposes a joint rate-distortion optimization for the H.264 rate control algorithm with a novel distortion prediction equation, which avoids linear regression employed in other distortion predictors and can considerably speed up

rate estimation. Multiple quantization parameters determination algorithm based on the statistics of the deviation measure is proposed in [12], which can achieve accurately QP. The work in [13] proposes a rate control technique for H.264/AVC using subjective quality of video. The work in [14] presents a complexity coefficient to combine the target bits. This paper presents a reformative basic unit layer rate control algorithm based on the HVS. The HVS was not taken into account in the JVT-G012 algorithm. Since eyes are the final receiver of the video information, so it is vital important to take the HVS into account in the video coding process. The HVS is very sensitive to the brim part and motion part. However each pixel of JVT-G012 algorithm is treated equally in the basic unit layer. Though the work in [10] takes advantage of the HVS, it is not comprehensive. In this paper the visual characteristic factor is used to improve the rate control in the basic unit layer.

2. JVT-G012 Algorithm

The JVT-G012 basic unit layer rate control mainly consists of three steps. First, it predicts the MAD of the current basic unit in the current frame. Then, it computes the target bit for the current basic unit. Last, it calculates the quantization parameter of the current basic unit and performs RDO.

2.1. Predict the MAD of the Current Basic Unit in the Current Frame. Consider

$$\text{MAD}_{cb} = a_1 * \text{MAD}_{pb} + a_2, \quad (1)$$

where MAD_{cb} is the predicted MAD of the current basic unit in the current frame. MAD_{pb} is the actual MAD of the co-located basic unit in the previous frame. a_1 and a_2 are two coefficients of the predictive model, whose initial values are 1 and 0, respectively; after finishing encoding every basic unit, the coefficients a_1 and a_2 are updated.

2.2. Compute the Target Bit for the Current Basic Unit. The target bit for the current basic unit $\tilde{b}_l$ is

$$\tilde{b}_l = T_r \times \frac{\text{MAD}_{cb}^2}{\sum_{k=1}^{N_{\text{unit}}} \text{MAD}_{cb}^2(k)}, \quad (2)$$

where T_r and N_{unit} are the number of remaining bits for the all uncoded basic units in the current frame and the number of uncoded basic units, respectively. MAD_{cb} is the predicted MAD of the current basic unit in (1).

2.3. Compute the Quantization Parameter of the Current Basic Unit and Perform RDO. Consider

$$\tilde{b}_l = \text{MAD}_{cb} * \left(\frac{X_1}{Q} + \frac{X_2}{Q^2} \right), \quad (3)$$

where Q is the quantization parameter of the current basic unit. $\tilde{b}_l$ is the target bit of the current basic unit in (2). MAD_{cb} is the predicted MAD of current basic unit in (1). X_1 and X_2 are the first-order and the second-order model parameters of the quadratic rate-distortion model, respectively. They are updated in the process of encoding.

FIGURE 1: PSNR curves of the highway sequence comparison for the JVT-G012 and the proposed algorithms.

FIGURE 2: PSNR curves of the akiyo sequence comparison for the JVT-G012 and the proposed algorithms.

3. Improved Rate Control Algorithm

The reformative basic unit layer rate control algorithm mainly consists of two steps. First, it allocates the target bits based on the HVS. Then, it adjusts the quantization parameter and performs RDO.

3.1. Compute the Target Bit of the Current Basic Unit Based on the HVS. This paper takes the HVS perception mechanism into account because the human eyes are extremely sensitive to the edge part and the motion part of images. So the

FIGURE 3: Comparison of the subjective quality in the highway sequence for the JVT-G012 and the proposed algorithms.

proposed algorithm assigns fewer bits to the unimportant region and assigns more bits to the region of interest in the process of allocating bits. It can achieve the goal of improving the overall video quality. This paper adjusts (2) with the visual characteristic factor $V(i, j)$:

$$\tilde{b}_l = T_r \times V(i, j) \times \frac{\text{MAD}_{cb}^2}{\sum_{k=1}^{N_{unit}} \text{MAD}_{cb}^2(k)}, \quad (4)$$

$$V(i, j) = \text{motion}(i, j) + \text{edge}(i, j),$$

where the motion vision characteristics and the edge vision characteristics are denoted by $\text{motion}(i, j)$ and $\text{edge}(i, j)$, respectively.

3.1.1. Motion Characteristics. In the real scene, there exist two major motion scenes as following.

The whole scene changes a little while only parts of the objects move or change. At this time, human eyes are concerned much more about the moving and changing objects. That is, when MV_{avg} is less than 4.5,

$$\text{motion}(i, j) = \begin{cases} 1 & MV < 1 \\ 2 - \frac{(6.5 - MV)}{8} & 1 < MV \leq 6.5 \\ 3 & MV > 6.5. \end{cases} \quad (5)$$

When the whole scene moves fast, MV_{avg} is more than 4.5, there are two subordinate situations in this case.

When there are many fast moving macroblocks, human eyes pay more attention to those objects that move little. This time MV is more than 2.5; $\text{motion}(i, j)$ can be expressed as

$$\text{motion}(i, j) = \begin{cases} 1 + \frac{(6.5 - MV)}{10} & 2.5 < MV \leq 6.5 \\ 3 & MV > 6.5. \end{cases} \quad (6)$$

When most of the objects move inconspicuously while only some of them move fast in the scene, human eyes pay more attention to the fast-moving part. This time, MV is less than 2.5; $\text{motion}(i, j)$ can be expressed as

$$\text{motion}(i, j) = \begin{cases} 1 & MV \leq 1.5 \\ 2 - \frac{(3.5 - MV)}{10} & 1.5 < MV < 2.5, \end{cases} \quad (7)$$

where the magnitude of motion vector for the j th basic unit in the i th frame and the magnitude of average motion vector in the remaining basic units of current frame are denoted by $MV(i, j)$ and MV_{avg} , respectively. One has $MV = \sqrt{MV_x^2 + MV_y^2}$, where MV_x represents the magnitude of the macroblock motion vector in horizontal direction and MV_y represents the magnitude of the macroblock motion vector in vertical direction.

3.1.2. Edge Characteristics. This paper describes the edge characteristics of images with the variance because there is high variance in the edge area of images:

$$\text{edge}(i, j) = \begin{cases} 1 - \frac{D(i, j)}{\max(D)} & D(i, j) < \text{medium}(D) \\ \frac{D(i, j)}{\max(D)} & \text{other,} \end{cases} \quad (8)$$

where $D(i, j)$ represents the variance for the j th basic unit in the i th frame. $\text{Medium}(D)$ and $\max(D)$ represent the mid-value and the maximum value of the variance, respectively.

3.2. Adjust the Quantization Parameter and Perform RDO. To consider the feedback information of the encoded frames, this paper adapts the quantization parameter adjustment coefficient. This paper uses η to adjust the quantization parameter, which is defined as the ratio of texture bits to the header bits:

$$QP_i(j) = \begin{cases} QP'_i(j) - 1 & 1.0 < \eta \leq 2.0 \\ QP'_i(j) + 1 & 0.6 < \eta \leq 1.0 \\ QP'_i(j) + 2 & \eta < 0.6 \\ Q_c & \text{other,} \end{cases} \quad (9)$$

where $QP'_i(j)$ is the average value of quantization parameters for all basic units in the previous frame. Q_c is the quantitative parameter in the JVT-G012 algorithm.

After the adjustment, the algorithm took into account the encoded frame information. The proposed algorithm achieves a good rate control. The algorithm performs RDO and updates the model parameters.

TABLE 1: Bit rate comparison for the proposed algorithm and the JVT-G012 algorithm.

Test sequences	Output bit rate (kbps)	
	JVT-G012	The proposed algorithm
Highway	32.41	32.06
Mother-daughter	32.10	32.04
Foreman	32.10	32.06
Claire	32.15	32.06
Hall	32.14	32.07
Silent	32.09	32.06
Akiyo	32.08	32.04
News	32.04	32.04
Carphone	32.13	32.06

TABLE 2: PSNR comparison for the proposed algorithm and the JVT-G012 algorithm.

Test sequences	PSNR (dB)		
	JVT-G012	The proposed algorithm	Gain
Highway	34.60	35.34	0.74
Mother-daughter	34.79	34.90	0.11
Foreman	28.38	28.49	0.11
Claire	39.36	39.56	0.20
Hall	34.05	34.26	0.21
Silent	31.03	31.16	0.13
Akiyo	37.93	38.16	0.23
News	31.29	31.50	0.21
Carphone	31.56	31.70	0.14

4. Simulation Results and Discussion

In order to validate the effectiveness of our algorithm in this paper, this paper has implemented the proposed rate control algorithm by the JM10.1 test model software. Also the proposed algorithm is compared with the JVT-G012 algorithm and the Zheng algorithm in [10]. The tested sequences are in QCIF4:2:0 formats: highway, mother-daughter, foreman, claire, hall, silent, akiyo, news, and carphone. In the experiments, all sequences are coded as the IPPP structure, the frame rate is set to 30 frames per second, the total number of frames is set to 200, the length of GOP is set to 40, and the target rate is 32 kbps.

Tables 1 and 2 show the comparison of the bit rate and the PSNR. As summarized in Tables 1 and 2, the proposed algorithm can control more accurately the bit rates than the JVT-G012 algorithm and obtain much better PSNR for the different video sequences. In particular for the test sequence highway, the proposed algorithm achieves a PSNR gain of 0.74 dB.

TABLE 3: Bit rate comparison for the proposed algorithm and the Zheng algorithm.

Test sequences	Output bit rate (kbps)	
	Zheng algorithm	The proposed algorithm
Highway	31.94	32.06
Mother-daughter	31.95	32.04
Foreman	31.93	32.06
Claire	32.03	32.06
Hall	32.04	32.07
Silent	32.06	32.06
Akiyo	32.01	32.04
News	32.04	32.04
Carphone	32.04	32.06

TABLE 4: PSNR comparison for the proposed algorithm and the Zheng algorithm.

Test sequences	PSNR (dB)		
	Zheng algorithm	The proposed algorithm	Gain
Highway	35.23	35.34	0.11
Mother-daughter	34.75	34.90	0.15
Foreman	28.35	28.49	0.14
Claire	39.49	39.56	0.07
Hall	34.00	34.26	0.26
Silent	31.07	31.16	0.09
Akiyo	38.04	38.16	0.12
News	31.40	31.50	0.10
Carphone	31.59	31.70	0.11

Tables 3 and 4 show the comparison of the bit rate and the PSNR. The proposed algorithm gets much better PSNR than the Zheng algorithm for different video sequences and also controls accurately the bit rate. The proposed algorithm can improve the average PSNR for all test sequences, while the Zheng algorithm can improve PSNR for a part of test sequences. For example, Zheng algorithm gets less PSNR than the JVT-G012 algorithm for test sequences: mother-daughter, foreman, and hall.

Figures 1 and 2 show that the PSNR curve is flatter than the one obtained from the JVT-G012 algorithm. The proposed algorithm suppresses the sharp drop of the PSNR and improves the stability of the picture quality.

Figure 3 shows obviously that the highway keeps the edge part of images to which human eyes are sensitive, while the one obtained from the JVT-G012 algorithm distorts and influences the subjective feeling of eyes.

Figures 4, 5, and 6 are the comparison of the subjective quality. For the test sequences claire, mother-daughter, and silent, their facial and body parts have a dramatic decline

FIGURE 4: Comparison of the subjective quality in the claire sequence for the JVT-G012 and the proposed algorithms.

FIGURE 5: Comparison of the subjective quality in the mother-daughter sequence for the JVT-G012 and the proposed algorithms.

FIGURE 6: Comparison of the subjective quality in the silent sequence for the JVT-G012 and the proposed algorithms.

in image quality in the JVT-G012 algorithm. However, these features have better visual quality in the proposed algorithm.

Figure 7 is the comparison of subjective quality for the test sequence carphone. It is found that the sensitive regions of eyes in the image, such as the face, clothing, and edge parts, have become blurred in the JVT-G012. However, these features have better visual quality in the proposed algorithm.

5. Conclusions

This paper proposes a reformative basic unit layer rate control algorithm by using the visual characteristic factor and the adjusted quantization parameter. The proposed algorithm allocates bits based on the HVS in basic unit layer and adjusts the quantization parameter with texture bits and the header

FIGURE 7: Comparison of the subjective quality in the carphone sequence for the JVT-G012 and the proposed algorithms.

bits of the current basic unit. The experimental results show that the proposed algorithm can control the bit rate more accurately and have a much better visual quality. Compared with the JVT-G012 algorithm, the PSNR can be improved by 0.2–0.7 dB. What is more, the PSNR has little fluctuation and the video image will become more stable. Compared with [10] algorithm, the PSNR for different test sequences has been improved obviously in this paper. The [10] algorithm has great effect on particular tests, while the proposed algorithm in this paper has universal applicability. In addition, the algorithm in this paper has great effect at low bit rates.

Acknowledgments

This work was supported by the Qing Lan Project and the Priority Academic Program Development of Jiangsu Higher Education Institutions.

References

- [1] C. T. E. R. Hewage and M. G. Martini, "Reduced-reference quality assessment for 3D video compression and transmission," *IEEE Transactions on Consumer Electronics*, vol. 57, no. 3, pp. 1185–1193, 2011.
- [2] Z. Chen, C. Lin, and X. Wei, "Enabling on-demand internet video streaming services to multi-terminal users in large scale," *IEEE Transactions on Consumer Electronics*, vol. 55, no. 4, pp. 1988–1996, 2009.
- [3] M. Hrarti, H. Saadane, M. Larabi, A. Tamtaoui, and D. Aboutajdine, "A macroblock-based perceptually adaptive bit allocation for H264 rate control," in *Proceedings of the 5th International Symposium on I/V Communications and Mobile Networks (ISIVC '10)*, pp. 1–4, October 2010.
- [4] M. Wang, T. Zhang, C. Liu, and S. Goto, "Region-of-interest based dynamical parameter allocation for H.264/AVC encoder," in *Proceedings of the Picture Coding Symposium (PCS '09)*, pp. 1–4, May 2009.
- [5] Y. Liu, Z. G. Li, and Y. C. Soh, "Region-of-interest based resource allocation for conversational video communication of H.264/AVC," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 18, no. 1, pp. 134–139, 2008.
- [6] R. Ruolin, H. Ruimin, and L. Zhongming, "A novel rate control algorithm of video coding based on visual perceptual characteristic," in *Proceedings of the 6th International Conference on Computer Science and Education (ICCSE '11)*, pp. 843–846, August 2011.
- [7] L. Liao, X. Zheng, Y. Zhao, and G. Liu, "A new digital video watermark algorithm based on the HVS," in *Proceedings of the International Conference on Internet Computing and Information Services (ICICIS '11)*, pp. 446–448, September 2011.
- [8] Y. Liu, Z. G. Li, and Y. C. Soh, "Region-of-interest based resource allocation for conversational video communication of H.264/AVC," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 18, no. 1, pp. 134–139, 2008.
- [9] G. Q. Lu and J. L. Li, "Video quality evaluation method based on the visual region of interest," *Computer Engineering*, vol. 35, no. 10, pp. 217–219, 2011.
- [10] Q. Zheng, M. Yu, Z. Peng, F. Shao, F. Li, and G. Jiang, "Human visual system-based rate control algorithm for H.264/AVC," *Guangdianzi Jiguang/Journal of Optoelectronics Laser*, vol. 22, no. 3, pp. 440–445, 2011.
- [11] F. Chen and Y. Hsu, "Rate-distortion optimization of H.264/AVC rate control with novel distortion prediction equation," *IEEE Transactions on Consumer Electronics*, vol. 57, no. 3, pp. 1264–1270, 2011.
- [12] J. Li and E. Abdel-Raheem, "Efficient rate control for H.264/AVC intra frame," *IEEE Transactions on Consumer Electronics*, vol. 56, no. 2, pp. 1043–1048, 2010.
- [13] S. L. P. Yasakethu, W. A. C. Fernando, S. Adedoyin, and A. Kondoz, "A rate control technique for off line H.264/AVC video coding using subjective quality of video," *IEEE Transactions on Consumer Electronics*, vol. 54, no. 3, pp. 1465–1472, 2008.
- [14] 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.