Rate Control Algorithm for MPEG-2 to H.264/AVC Transcoding

Gao Chen, Shouxun Lin, and Yongdong Zhang

Institute of Computing Technology, Chinese Academy of Sciences
Beijing, 100080, P.R. China
`{chengao,sxlin,zhyd}@ict.ac.cn`

Abstract. There is strong need for research in transcoding technologies to enable smooth displacement from MPEG-2 to H.264/AVC since H.264/AVC has been standardized as international standard. In this paper, a novel rate control algorithm for MPEG-2 to H.264/AVC transcoding, which adopting a new block activity measurement, is proposed. Specifically, the standard deviation of the residual error is introduced into the quadratic rate distortion (R-D) model adopted in JVT-G012 instead of the mean of absolute difference (MAD) to measure macroblock (MB) complexity. Meanwhile, based on the fact that the mean square of AC coefficients in an 8×8 DCT block is equal to the variance of an 8×8 block before DCT, we derive a close-form formulation to calculate the variance of a residual MB using the DCT coefficients rather the pixel values. Obviously, this rate control method can be used for MPEG-2 to H.264/AVC transcoder in both pixel domain and transform domain. Experiments show that our proposed algorithm can meet the target bit-rate accurately and achieves a better performance than the JVT-G012.

1 Introduction

H.264/AVC is the latest international video coding standard, developed and standardized collaboratively by ISO/IEC and ITU-T as International Standard 14496-10 (MPEG-4 part 10) Advanced Video Coding (AVC) or as Recommendation H.264 [1]. H.264/AVC achieves high coding efficiency by adopting a variety of state-of-the-art tools and is expected to replace the existing standards such as H.263 and MPEG-1/2/4. Given its outstanding coding efficiency, H.264/AVC is expected to have a wide range of applications, including mobile broadcasting and storage. However, MPEG-2 video has been widely used in many existing systems, such as digital TV, DVD, and HDTV applications etc. To solve the standard incompatibility problems for Universal Multimedia Access (UMA) [2], there is a big demand for converting video in the MPEG-2 format to the one in H.264/AVC format.

Several issues on rate control for H.264/AVC transcoding have been addressed recently in [3]-[5]. An algorithm of adopting the rate control model TM5 in MPEG-2 to compute the values of quantization parameters (QP) for I and B frames based on the side information from the pre-coded MPEG-2 video is presented in [3]. In [4], a fast macroblock (MB) mode decision approach has been proposed to reduce the

J. Martí et al. (Eds.): IbPRIA 2007, Part II, LNCS 4478, pp. 72–79, 2007.

complexity of Rate Distortion Optimization (RDO) and an improved rate control method depending on statistics of input MPEG-2, which is effective in transcoding the input steams into low bit rate streams, has been proposed. In [5], an idea that we should reuse information extracted from the input MPEG-2 video stream as efficiently as possible is proposed. The experiment results demonstrate that the proposed rate control algorithm is very efficient. However, all of the aforementioned works mainly focus on the MPEG-2 to H.264/AVC transcoder in pixel domain and can not be used in the transform domain simultaneously. Recently, transcoding MPEG-2 into H.264/AVC in transform domain has been an actively studied topic in academia and industry community. In [6], an efficient method has been proposed to convert DCT coefficients to H.264/AVC integer transform coefficients completely in the transform domain. A transform domain MPEG-2 to H.264/AVC intra video transcoder is proposed in [7] and the proposed transcoder is equivalent to the conventional one in pixel domain in terms of functionality and achieves complexity saving more than 20%. Specially, a comprehensive solution to transcode MPEG-2 into H.264/AVC in transform domain is proposed in [8]. To perform the rate control for transform domain MPEG-2 to H.264 transcoding, we propose a novel rate control algorithm which can be used in both pixel domain and transform domain. To our best knowledge, no works in this respect has been reported before in this literature.

The rest of the paper is organized as follows. In Section 2 we propose a new quadratic rate distortion (R-D) model. In Section 3 we present a way to calculate the variance of MB using only the DCT coefficients. Section 4 describes our proposed rate control method for MPEG-2 to H.264 transcoding in summary. Experimental results will be presented in Section 5, and conclusion is shown in Section 6.

2 New Rate Distortion Model

According to the rate control algorithm based on the R-D theory, the quantization step size of a MB is selected according to its activity, which is usually measured by variance, mean of absolute differences (MAD), sum of absolute differences (SAD), etc. H.264/AVC rate control proposal JVT-G012 adopts the MAD as the MB activity [9]. At the same time, the R-D model adopted in JVT-G012 is the well known quadratic R-D model, which is shown in equation (1).

$$T - H = c_1 \frac{MAD^2}{QP^2} + c_2 \frac{MAD}{QP} \tag{1}$$

Where c_1 and c_2 are the model parameter.

Replacing MAD with standard deviation to measure the MB activity, we can get a new R-D model. The new R-D model is formulated in the equation as follows:

$$T - H = c_1 \frac{\sigma^2}{QP^2} + c_2 \frac{\sigma}{QP} \tag{2}$$

Where σ represents the standard deviation of the residue error.

In order to improve the accuracy of the new R-D model, we verify this new R-D model in the context of encoding system. That is, we first implement the new R-D model in the H.264 encoder of the H.264/AVC reference software JM 8.2 [10]. Then, we compare the encoding results with the one using JVT-G012. Sequences with different amounts of motion and spatial details are used in our experiments. Due to the limit of pages, only the results of six sequences are provided there and the results of other sequences are similar. As shown in Fig. 1(a) and Fig. 1(b), we can see that that using standard deviation in quadratic R-D model can obtain the same or better coding efficiency compared to using MAD, which prove the accuracy of our proposed new R-D model.

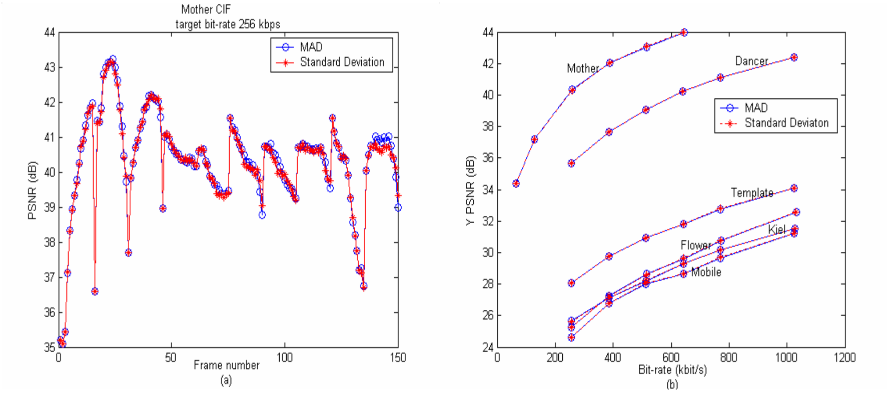

Fig. 1. (a) The PSNR (dB) of Mother (CIF) obtained by quadratic R-D model using MAD and standard deviation. (b) R-D curves of six different sequences with quadratic R-D model using MAD and standard deviation.

3 Variance Calculation in Transform Domain

Due to the reconstruction of picture in pixel domain is not needed in the context of transform domain transcoder, the MAD of residue error can not be achieved in the process of transcoding. So, the R-D model in equation (1) can not be used in transform domain transcoder. In the following, we derive a close-form formulation to calculate the variance of a MB only using the DCT coefficients. So, we can use that the R-D model in equation (2) in transform domain and pixel domain transcoder simultaneously.

In what follows, we describe the process of how to calculate the variance of a MB in transform domain. The 8×8 two dimension Discrete Cosine Transform (DCT) [6] is given by

$$F(u,v)=\frac{1}{4}C(u)C(v)\sum_{x=0}^{7}\sum_{y=0}^{7}f(x,y)\cos\left(\frac{(2x+1)u\pi}{16}\right)\cos\left(\frac{(2y+1)v\pi}{16}\right) \tag{3}$$

Where u, v, x, y =0, 1, 2, ……, 7, $\begin{cases} C(u),C(v)=\sqrt{1/2}, & u,v=0\,. \\ C(u),C(v)=1, & u,v=other \end{cases}$

For an 8×8 DCT block, we define $\overline{AC^2}$ as the mean square of AC coefficients in DCT, which is showed in equation (4).

$$\overline{AC^{2}} = \frac{1}{8\times 8}\left(\sum_{u=0}^{7}\sum_{v=0}^{7} F^{2}(u,v) - F^{2}(0,0)\right) \tag{4}$$

Where $F(0,0)$ is the DC coefficient of this 8×8 DCT block.

According to the Parseval`s theorem [11], we have:

$$\sum_{u=0}^{7}\sum_{v=0}^{7} F^{2}(u,v) = \sum_{x=0}^{7}\sum_{y=0}^{7} f^{2}(x,y) \tag{5}$$

Furthermore, let $\overline{f(x,y)}$ be the mean pixel value of an 8×8 block before DCT, then $F(0,0) = 8\times\overline{f(x,y)}$.

As described in [12], we can compute $\overline{AC^{2}}$ as follows.

$$\begin{aligned}\overline{AC^{2}} &= \frac{1}{8\times 8}\left(\sum_{u=0}^{7}\sum_{v=0}^{7} F^{2}(u,v) - F^{2}(0,0)\right)\\ &= \frac{1}{8\times 8}\left(\sum_{x=0}^{7}\sum_{y=0}^{7} f^{2}(x,y) - \left(8\times\overline{f(x,y)}\right)^{2}\right)\\ &= \frac{1}{8\times 8}\sum_{x=0}^{7}\sum_{y=0}^{7} f^{2}(x,y) - \overline{f(x,y)}^{2}\\ &= \frac{1}{8\times 8}\sum_{x=0}^{7}\sum_{y=0}^{7}\left(f(x,y) - \overline{f(x,y)}\right)^{2} = \sigma^{2}\end{aligned} \tag{6}$$

Where σ^{2} is the variance of an 8×8 block before DCT.

The above equations show that the mean pixel value and the variance of an 8×8 block can be computed directly using its corresponding DCT coefficients. That is:

$$\begin{aligned}\overline{f(x,y)} &= F(0,0)/8\\ \sigma^{2} &= \overline{AC^{2}}\end{aligned} \tag{7}$$

It is well known that in typically block-based video coding standard, the block size used for transform is corresponded to the dimension of the transform. Such as in MPEG-2, 8×8 block is used for transform corresponding to 8×8 DCT and Inverse DCT operations. However, the basic unit of rate control in video encoder is 16×16 MB usually. We need to deduce an approach to compute the variance of a 16×16 MB using the DCT coefficients of 8×8 blocks. Because that only the DCT coefficients of 8×8 DCT transform exist in MPEG-2 video stream.

Let b_i, $\overline{f_i(x,y)}$ *and* σ_i^{2}, $i = 1,2,3,4$ denote the four 8×8 blocks of a 16×16 MB, the mean values and the variances of the four 8×8 blocks, respectively. From aforementioned conclusion, we have:

$$\begin{aligned}\overline{f_i(x,y)} &= F_i(0,0)/8\\ \sigma_i^{2} &= \overline{AC_i^{2}}\end{aligned} \tag{8}$$

Fig. 2 shows the demonstration of a MB containing four 8×8 blocks.

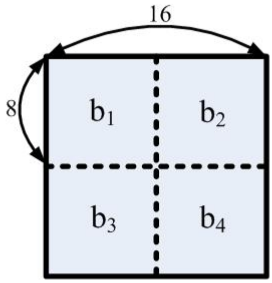

Fig. 2. Four 8×8 blocks in a MB

Let $\overline{f_{MB}(x,y)}$ and σ_{MB}^2 denote the mean value and the variance of a MB, respectively. Firstly, we can compute $\overline{f_{MB}(x,y)}$ and σ_{MB}^2 in pixel domain as follows.

$$\begin{aligned}\overline{f_{MB}(x,y)} &= \frac{1}{16\times16}\sum_{x=0}^{15}\sum_{y=0}^{15} f(x,y)\\ \sigma_{MB}^2 &= \frac{1}{16\times16}\sum_{x=0}^{15}\sum_{y=0}^{15}\left(f(x,y)-\overline{f_{MB}(x,y)}\right)^2\\ &= \frac{1}{16\times16}\sum_{x=0}^{15}\sum_{y=0}^{15} f^2(x,y)-\overline{f_{MB}(x,y)}^2\end{aligned}\tag{9}$$

Secondly, we can use $\overline{f_i(x,y)}\, and\ \sigma_i^2,\ i=1,2,3,4$ to compute σ_{MB}^2 as follows:

$$\begin{aligned}\overline{f_{MB}(x,y)} &= \frac{1}{16\times16}\sum_{x=0}^{15}\sum_{y=0}^{15} f(x,y)\\ &= \frac{1}{4}\left(\overline{f_1(x,y)}+\overline{f_2(x,y)}+\overline{f_3(x,y)}+\overline{f_4(x,y)}\right)\end{aligned}\tag{10}$$

$$\begin{aligned}&\frac{1}{16\times16}\sum_{x=0}^{15}\sum_{y=0}^{15} f^2(x,y)\\ &= \frac{1}{4}\Bigg(\frac{1}{8\times8}\sum_{x=0}^{7}\sum_{y=0}^{7} f^2(x,y)-\overline{f_1(x,y)}^2+\frac{1}{8\times8}\sum_{x=7}^{15}\sum_{y=0}^{7} f^2(x,y)-\overline{f_2(x,y)}^2\\ &\qquad +\frac{1}{8\times8}\sum_{x=0}^{7}\sum_{y=7}^{15} f^2(x,y)-\overline{f_3(x,y)}^2+\frac{1}{8\times8}\sum_{x=7}^{15}\sum_{y=7}^{15} f^2(x,y)-\overline{f_4(x,y)}^2\Bigg)\\ &\quad +\frac{1}{4}\left(\overline{f_1(x,y)}^2+\overline{f_2(x,y)}^2+\overline{f_3(x,y)}^2+\overline{f_4(x,y)}^2\right)\\ &= \frac{1}{4}\left(\sigma_1^2+\sigma_2^2+\sigma_3^2+\sigma_4^2\right)+\frac{1}{4}\left(\overline{f_1(x,y)}^2+\overline{f_2(x,y)}^2+\overline{f_3(x,y)}^2+\overline{f_4(x,y)}^2\right)\end{aligned}\tag{11}$$

Substituting equation (8), equation (10) and equation (11) into (9), we have:

$$\begin{aligned}\sigma_{MB}^2 &= \frac{1}{4}\left(\overline{AC_1^2}+\overline{AC_2^2}+\overline{AC_3^2}+\overline{AC_4^2}\right)\\ &\quad +\frac{1}{16\times16}\Big\{F^2{}_1(0,0)+F^2{}_2(0,0)+F^2{}_3(0,0)+F^2{}_4(0,0)\\ &\qquad -\left(\left(F_1(0,0)+F_2(0,0)+F_3(0,0)+F_4(0,0)\right)/2\right)^2\Big\}\end{aligned}\tag{12}$$

Where the $\overline{AC_i^2}$ and $F_i(0,0)$, i = 1,2,3,4 denote mean square of AC coefficients in DCT and the DC coefficients of the four 8×8 blocks, respectively.

4 Rate Control Algorithm

From equation (12), we can say that the variance of MB can be calculated directly using the DCT coefficients of 8×8 blocks in the case of transform domain transcoding. Combing the R-D model in equation (2), we propose a novel rate control for transform domain MPEG-2 to H.264 transcoder. Because of that most parts of our proposed algorithm inherits JVT-G012, we only present the different part there. The full procedure can refer to [13] for details.

Step 1: Using the equation (12) to calculate MPEG-2 MB activity.

Step 2: To solve the well known chicken-and-egg problem in the context of transcoding, the final standard deviation for current MB is adjusted with that gotten in Step 1 as follows:

$$\sigma_{final} = \alpha \times \sigma_{mpeg-2} + (1-\alpha) \times \sigma_{pred} \tag{13}$$

Where σ_{mpeg-2} is the standard deviation obtained from the incoming MPEG-2 DCT coefficients using the equation (12), and the σ_{pred} is the one predicted with the linear model using the actual standard deviation of encoded MB in the same spatial position of the previous frame. The constant α serves as weighting factor and its typical value is 0.5 in our experiments.

Step 3: Adopting the new R-D model in equation (2) to calculate the QP.

5 Experimental Results

Our proposed rate control method is implemented in our MPEG-2 to H.264 transcoder to verify its performance. Our MPEG-2 to H.264 transcoder utilizes a decoder provided by the MPEG Software Simulation Group [14] to decode the incoming MPEG-2 test video streams into images in pixel domain and cascades an encoder based on the reference software H.264/AVC JM 8.2 (hereafter referred to as the JM 8.2) [10] to compress the images into H.264 format bit stream with the same coding structure and resolution. In our experiments, for each test sequence, the first 150 frames are firstly encoded to MPEG-2 streams at bit-rate of 1 or 2 Mbps and a frame rate of 30 fps with the structure of group of picture (GOP) as the first frame is I frame and 14 P frames are followed (i.e., IPP……PPP).

The average peak signal-to-noise ratio (PSNR) and the actual bit-rate obtained for transcoding the pre-coded Dancer and Kiel sequences to six different target bit-rates are show in Table 1. The results show that the proposed method can provide a better performance than JVT-G012 in terms of both average PSNR and achieved bit-rate.

Table 1. PSNR and actual bit-rate botained for the Dancer and Kiel sequences at six different target bit-rates

Target bit-rate (kbps)	Dancer				Kiel			
	My proposed method		JVT-G012		My proposed method		JVT-G012	
	Actual bit-rate (kbps)	PSNR (dB)	Actual bit-rate (kbps)	PSNR (dB)	Actual bit-rate (kbps)	PSNR (dB)	Actual bit-rate (kbps)	PSNR (dB)
256	256.67	35.38	256.69	35.36	256.69	24.82	256.57	24.81
384	385.03	37.08	385.03	37.05	385.85	25.79	384.79	25.79
512	512.95	38.15	513.08	38.12	513.25	26.48	513.23	26.47
640	640.86	38.92	641.04	38.90	641.59	26.82	641.75	26.82
768	769.39	39.42	769.00	39.41	769.78	27.35	770.07	27.34
1024	1024.87	40.05	1025.22	40.04	1025.60	28.00	1025.89	28.00

Fig. 3 (a) shows the frame-to-frame PSNR results of the Dancer sequence obtained by the proposed and JVT-G012 rate control methods. No surprisingly, the fluctuation of PSNR obtained by the proposed method is less than that of the JVT-G012 method.

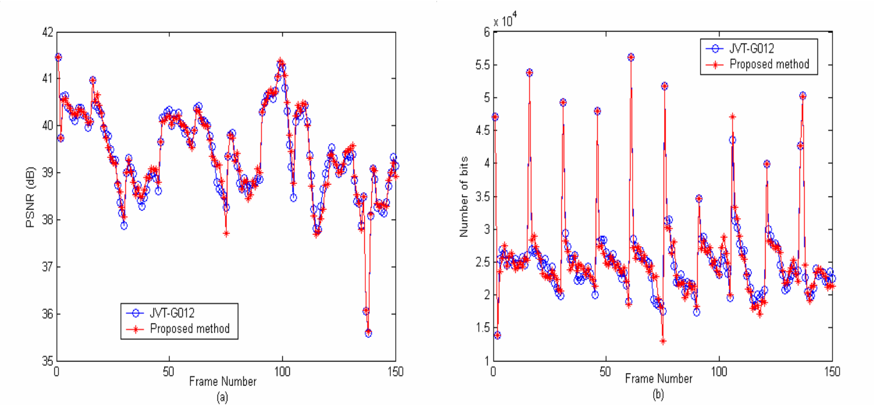

Fig. 3. (a) The PSNR (dB) of the Dancer sequence obtained by the proposed method and JVT-G012 methods. (b) The number of actual coding bits obtained by using the QP determined by the proposed method and JVT-G012.

Fig. 3 (b) shows the distribution of the number of bits over the entire sequence when the FOREMAN sequence was transcoded at the target bit-rate 768 Kbps by using the proposed rate control method and JVT-G012. It can be seen that the fluctuation of bits of the transcoded video obtained by the proposed method is a little better than that of JVT-G012.

6 Conclusion

In this paper, a new rate control algorithm, which adopting standard deviation as MB activity measurement is presented. The variance can be directly calculated in transform domain makes it suitable for the transform domain transcoder where the MAD of residue error can not obtained. The experimental results show the accuracy of the model and the effectiveness of the proposed rate control algorithm. So we can say that this algorithm will be popularized in DCT based video transcoding. In the further, we will focus on improving the efficiency of our method by reusing the motion information in MPEG-2 inputting bit-stream.

Acknowledgments. This work is supported by national nature science foundation of China under grant number 60302028.

References

1. Wiegand, T., Sullivan, G.: Draft lTU-T recommendation and final draft international standard of joint video specification (ITUT Rec. H.264 – ISO/IEC 14496-10 AVC)JVT-G050, Pattaya, Thailand (March 2003)
2. Mohan, R., Smith, J.R., Li, C.S.: Adapting Multimedia Internet Content for Universal Access. IEEE Transactions on Multimedia 1, 104–114 (1999)
3. Xiao, Y.-N., Lu, H., Xue, X., Nguyen, V.-A., Tan, Y.-P.: Efficient Rate Control For Mpeg-2 To H.264/AVC Transcoding, IEEE International Symposium on Circuits and Systems (ISCAS), pp. 1238–1241 (May 2005)
4. Zhang, P., Huang, Q.-M., Gao, W.: Key Techniques of Bit Rate Reduction for H.264 Streams, Advances in Multimedia Information Processing - PCM 2004: 5th Pacific Rim Conference on Multimedia, Tokyo, Japan, pp. 985–992 (November 30–December 3 2004)
5. Yang, J., Dai, Q., Xu, W., Ding, R.: A Rate Control Algorithm for MPEG-2 to H.264 Real-time Transcoding. In: Proc. Visual Communication and Image Processing VCIP-2005, pp.1995-2003, Beijing, China (2005)
6. Xin, J., Vetro, A., Sun, H.F.: Converting DCT Coefficients to H.264/AVC Transform coefficients. Pacific-rim Conference on Multimedia (PCM) 2, 939–946 (2004)
7. Su, Y., Xin, J., Vetro, A., Sun, H.: Efficient MPEG-2 to H.264/AVC intra transcoding in transform-domain. In: Proceedings of the IEEE International Symposium on Circuits and Systems (ISCAS ’05), vol. 2, pp. 1234–1237, Kobe, Japan (May 2005)
8. Qian, T., Sun, J., Li, D., Yang, X., Wang, J.: Transform domain transcoding from MPEG-2 to H.264 with interpolation drift-error compensation. IEEE Transaction on Circuits and Systems for Video Technology 16(4), 523–534 (2006)
9. Chiang, T., Zhang, Y.: A New Rate Control Scheme Using Quadratic Rate Distortion Model. IEEE Trans. Circuits Syst. Video Technol. 7, 287–311 (1997)
10. H.264/AVC reference software JM8.2, available online at http://bs.hhi.de/~suehring/tml/download
11. Oppenheim, A.V., Schafer, R.W.: Discrete-Time Signal Processing, 2nd (edn.). p. 60. Prentice Hall, UpperSaddle River, NJ (1999)
12. Bo, Z., Cheng-Ke, W.: A Novel Rate Control Algorithm in Video Coding for Low-delay Communications. Chinese Journal of Computers 28(1), 53–59 (2005)
13. Zh. Li, F., Pan, K., Lim, G., Feng, X.: Lin, Rahardja, S.: Adaptive Basic Unit Layer Rate Control for JVT, JVT-G012, 7th Meeting: Pattaya II, Thailand (March 2003)
14. MPEG-2 video encodec/decodec v12, available online at http://www.mpeg.org/MPEG/MSSG