A TWO-LEVEL RATE CONTROL APPROACH FOR VIDEO TRANSCODING

Maria Pantoja and *Nam Ling*
Department of Computer Engineering
Santa Clara University
Santa Clara, California 95053, U.S.A
Email: {mpantoja, nling}@scu.edu

ABSTRACT

We present a two-level rate control approach for VC-1 to H.264 transcoding. First, a low complexity algorithm in which the key is to find the relationship between quantization parameters (QPs) in VC-1 to QPs in H.264. Second, a medium complexity algorithm in which the key is to use mean absolute differences and sum of absolute transform differences calculated in VC-1 to estimate the complexity of macroblocks in H.264 for a pixel/transform domain transcoder. The low complexity rate control tool has a limitation of only able to rate-control transcode for QP ranges from 10 to 29. To transcode the entire QP range we propose a combination of low and median complexities tools. Results show that the proposed rate control transcoding is less complex than that of a full-cascaded transcoder with regular rate control turned on, while maintaining target bit-rate and PSNR.

Index terms - video coding, visual communications, transcoding, H.264, VC-1, rate control

1. INTRODUCTION

In applications where different video formats are involved with the need to convert from one to another, transcoders are used. For an output bitstream to adapt to a reasonable channel bandwidth while maintaining visual quality, rate-control transcoding is addressed. For encoders without rate control the video sequence can be compressed with a constant quantization parameter (QP), the resulting output video frames may have similar qualities but the bit-rate varies depending on the complexity of each frame. These variations are sometimes unacceptable and most of the time constraints are imposed by the encoder buffer size and network bandwidth, which make the uses of rate control algorithm a necessity. Rate control (RC) is not part of the standards but is included as an informative part for most of them. Examples include [3, 4, 5] for H.264. VC-1, to our knowledge, has not recommended any particular rate control algorithm, but uses rate control for its reference encoder. There are several works on rate-control transcoding, which includes MPEG-2 to H.264 [7] and H.263 to H.264 [8]. Examples of more recent work include a frame layer adaptive rate control MPEG-2 to H.264 [15] and a macroblock layer rate control trascoder from MPEG-2 to AVS [16]. We have selected to work on VC-1 to H.264 rate control transcoding as both are among the latest video coding standards with wide applications, with VC-1 being less complex than H.264.

Two main classes of algorithms are used for transcoding (e.g. [9, 10]): a pixel domain method which fully decodes the input video and then re-uses the information gathered to do fast encoding; and a transform domain method which partially decodes to the transform coefficient levels and converts the coefficients from one standard to another with matrix multiplication techniques. In this paper we propose a two-level approach (a low complexity and a medium complexity) rate control for transcoding. We apply the method to the pixel domain transcoder of [9] to transcode I-frames, and the transform domain transcoder of [10] to transcode P and B frames, for VC-1 to H.264 transcoding. Figure 1 represents the transcoder used.

The main contribution of this paper is the development of a medium complexity rate control, which calculates a complexity estimator and performs texture bits allocation. The method calculates the complexity estimators in VC-1 in terms of mean of absolute differences (MAD) for I-frames and sum of absolute transform differences (SATD) for P and B frames. The method then uses them to estimate the complexity of the macroblocks (MBs) for H.264 to accelerate H.264 rate control. To improve the performance of rate control we also use VC-1 measures of texture complexity to allocate bits per frame; this information can alert ahead of time of especially complex regions of the picture, which allows generating a corrected target for texture bits in H.264. The medium complexity method combines with a low complexity method to accelerate the rate control for VC-1 to H.264 transcoding. Section 2 gives an overview of the rate control performed in VC-1 and H.264. Section 3 describes the low and medium complexities tools and how they can be used to accelerate H.264 rate control with the information obtained from VC-1. In Section 4, experimental results are shown and analyzed. Section 5 presents a conclusion.

978-1-4244-5654-3/09/$26.00 ©2009 IEEE

2. OVERVIEW OF RATE CONTROL FOR VC-1 AND H.264

VC-1 rate control in the reference decoder depends on the settings of the Hypothetical Reference Decoder (HRD) and further details can be referred to [1]. H.264 rate control is described in [3]. In this section we very briefly summarize the equations that are important for the understanding of Section 3. To compute the target bits for the current frame a fluid traffic model [6] is used and to determine the quantization step *QS*, the following second order equation can be used [11]:

$$Texture\_Bits = c_1 * \frac{MAD}{QS} + c_2 * \frac{MAD}{QS^2} \quad , \qquad (1)$$

where c_1 and c_2 are constant value coefficients that may be estimated empirically and dynamically updated; *QP* is related to *QS* as discussed in [2].

After *QP* is calculated H.264 performs rate distortion optimization (RDO) to select the MB type to be used, by applying the Lagrangian cost function [12]. The Lagrangian mode decision for an MB, S_K, proceeds by minimizing the following cost function,

$$D(S_k, I_k \mid QP) + \lambda_{MODE} R(S_k, I_k \mid QP) , \qquad (2)$$

where I_K is the set of MB types, *D* is the distortion, *R* is the rate, and $\lambda_{MODE} = 0.85 * 2^{(QP-12)/3}$.

3. TWO-LEVEL RATE CONTROL

3.1 Low Complexity Rate Control

In frame layer rate control, I, P, and B frames for VC-1 are mapped to H.264 I, P, and B frames, respectively, and we use BI frames of VC-1 to detect scene changes for H.264. If the frame in VC-1 is a BI frame we map all MBs in this frame to intra mode in H.264 and re-use the $QP_{\text{VC-1}}$, as described below. In cases where a very low target bit-rate is required and frame skip is necessary, to avoid drift error we only skip B frames.

At MB layer rate control, details for low complexity RC tool can be found in [13]; here we summarize them briefly:

a) For H.264 *QP* ranges from 10 to 29, VC-1 can achieve the same bit-rates as H.264.
b) The following equation can be used to map *QP*s in VC-1 ($QP_{\text{VC-1}}$) to *QP*s in H.264 ($QP_{\text{H.264}}$).

$$QP_{H.264} = a * QP_{VC-1}^2 + b * QP_{VC-1} + c \ , \qquad (3)$$

where *a* = -0.02, *b* = 1.10, and *c* = 9.92. *QP*s in VC-1 are signaled at frame or slice level while *QP*s in H.264 are signaled at basic unit (BU) level. *QP*s are averaged in sets of MBs to match the H.264 BU size. With the values of the $QP_{H.264}$ we can estimate the value of block mode λ_{MODE} in Equation (2). With the mapping of Equation (3) to obtain $QP_{H.264}$ we reduce the computational complexity to perform H.264 rate control in transcoding.

3.2 Medium Complexity Rate Control

By using the tool described in Section 3.1 we observe that the target bit-rates can be maintained without a significant loss in PSNR only for $QP_{H.264}$ ranges from 10 to 29. To widen the applicable range we propose to design a slightly more complex RC transcoder. There are several VC-1 variables that can be re-used to implement it.

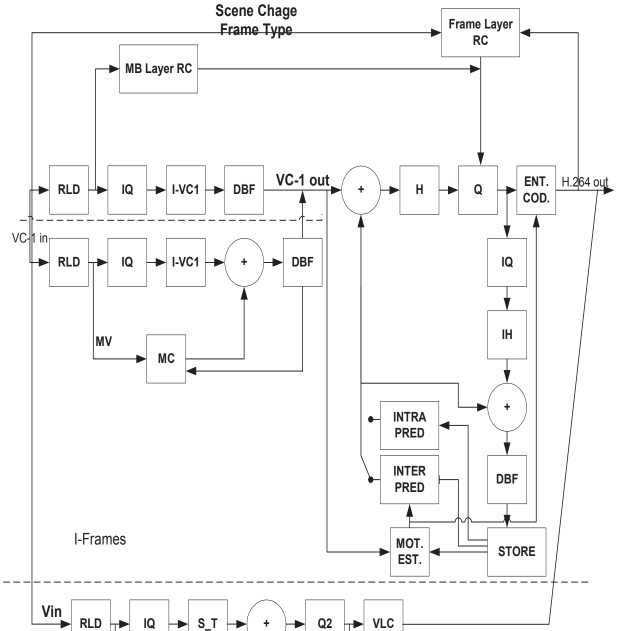

Figure 1. A VC-1 to H.264 transcoder.

3.2.1 Complexity Estimation

For the transform domain transcoder used for P and B frames, where there is no motion estimation (ME) performed [9] (due to motion vectors re-use), it is difficult to obtain a value for MB complexity estimation (MAD) [14]. Therefore to estimate the complexity for the MB we use a different complexity estimator, the sum of absolute transform differences (SATD), calculated using the sum of the absolute values of the frequency transform of the residuals. Since we transcode from VC-1 to H.264 the SATD is calculated as:

$$SATD = \sum_{i,j} \left| VC(residual)_{i,j} \right| , \qquad (4)$$

where *VC* is the VC-1 forward transform. SATD predicts visual quality more accurately than MAD or sum of absolute differences (SAD) from the standpoint of objective and subjective metrics. Also, SATD is more convenient for transform domain transcoding, which only partially decodes incoming P/B frames to the transform coefficient level.

For pixel domain transcoding for I frames, we use MAD as the incoming frames are decoded by pixel domain transcoding to the pixel level and because of its simplicity of calculation.

3.2.2 Texture Bit Allocation

In H.264, bit allocation is typically performed per group of pictures (GOP), even though the standard does not specify a GOP structure (and neither does VC-1). GOPs are usually used to simplify rate control tasks, to allow predictable insertion of intra frames where the application needs, and to perform bits per frame allocation. Bits per frame allocation in H.264 ($BitsF_{H.264}$) and per BU ($BitsBU_{H.264}$) can be dynamically adjusted based upon the value of VC-1 bits per frame ($BitsF_{VC\text{-}1}$).

With the value of the complexity estimation for VC-1 (Section 3.2.1) and the value of texture bits assigned by the rate controller for VC-1, we substitute these two values into Equation (1) to obtain the value of *QS*, reducing the complexity for H.264 rate control greatly. As mentioned in Section 3.1, once we have the value of $QP_{H.264}$ we can estimate the value of block mode (λ_{MODE}) in Equation (2).

To improve the performance of texture bit allocation we could also take into account the complexity of the frame. It is common to have fluctuations on actual texture bits with respect to the allocated texture bits. These fluctuations are usually due to the fact that *QP* changes need to be constrained or smoothed to avoid drastic changes in visual quality that would otherwise produce visible artifacts in the picture; a rate limiter is applied which typically limits changes in *QP* to no more than 2 units. VC-1 measures texture complexity to allocate bits per frame. Information provided by VC-1 can alert ahead of time of especially complex regions of the picture which allows generating a corrected target for H.264 texture bits, at both frame and BU levels.

Providing *N* frames of buffering between VC-1 and H.264 allows us to look into the future at the characteristics of the sequence on an *N*-frame sliding window ahead in time. A small *N* is required for low latency applications and is an interesting problem in minimizing buffering memory requirements (for future study). For our experiments in Section 4 we use *N* = 1. On the other hand, large *N* allows us to get benefits similar to dual-pass off-line encoder, where a sequence is encoded once and the RC parameters generated are used to re-adjust the second pass compensation for the excess or lack of bits with respect to overall sequence target.

4. EXPERIMENTS AND RESULTS

In our experiments, JM 13.1 was the H.264/AVC reference software used. To run the experiments a Dell Inspiron 300m running at 1.2 GHz and 256 MB memory was used. Several video sequences were tested; the results for "Foreman", "Claire", and "Walk", at 176×144, all with GOP of IPPP…, encoded at 30 fps, are listed below.

The details of the results for using low complexity rate control tools can be found in [13]. We can see there that there is no significant loss in PSNR when the low complexity tools are used while the bit-rates are almost the same as those of the full-cascaded transcoder with rate control turned on. However, we reduce the average time used to encode a sequence by around 57%.

The results for using the medium complexity tools are described as follows. In Table 1, we demonstrate the comparisons of the results for the video sequences for target bit-rates of 256 and 2,000 kbps. These target bit-rates were selected to test the extreme cases (QP<10 and QP>29). The comparisons are between the fully cascaded transcoder with rate control turned on for both VC-1 and H.264 and our rate control algorithm applied to the pixel/transform domain VC-1 to H.264 transcoder. From the table we see that there is no significant loss in PSNR for our cases while the bit-rates are almost the same.

Table 1. Transcoding Bit-rate and PSNR Results

Sequence	**Actual bit-rate (kbps)**		**PSNR (dB)**	
	Fully cascaded (RC turned on)	Our rate control transcoder	Fully cascaded (RC turned on)	Our rate control transcoder
Target bit-rate = 256 kbps				
Foreman	256.61	256.50	23.29	22.65
Claire	250.51	250.72	31.66	31.16
Walk	256.12	256.34	26.12	26.24
Target bit-rate = 2,000 kbps				
Foreman	1960.30	2000.21	36.12	36.34
Claire	1980.13	2000.34	42.35	42.13
Walk	2000.09	2000.17	32.6	31.8

Table 2 shows the computational complexity results for our medium complexity rate control transcoder as compared to that of the fully cascaded transcoder with rate control turned on. From the table we reduce the average time for transcoding by about 52% overall.

Table 2. Rate Control Transcoding Time Units

Sequence	**Fully cascaded (RC turned on)**	**Our rate control transcoder**
Target bit-rate = 256 kbps		
Foreman	9.4	4.8
Claire	8.7	4.0
Walk	10.1	4.6
Target bit-rate = 2,000 kbps		
Foreman	48.12	23.1
Claire	40.22	18.3
Walk	51.21	27.2

5. CONCLUSION

From the discussions and experimental results the low complexity rate control has a better computational complexity performance than that of the medium complexity rate control but can only be used for H.264 *QP* from 10 to 29. The medium complexity transcoder applies to a wider range of *QP*s but with less improvement in complexity performance. Video quality and bit-rates are maintained in both cases as compared to those of the fully cascaded transcoder with regular rate control turned on. We therefore propose to use a two-level approach - the low complexity transcoder for *QP* ranges within 10 to 29, and the medium complexity transcoder for all other QPs. Experimental results show that the proposed combination of low and medium complexities rate control for our transcoding is less complex than that for a full-cascaded transcoder with regular rate control, and yet maintaining target bit-rate and PSNR.

REFERENCES

[1] "Standard for Television: VC-1 Compressed Video Bitstream Format and Decoding Process," *SMPTE 421M-2006*.

[2] H.264/AVC Reference Software JM12.1, available at http://iphome.hhi.de/suehring/tml/.

[3] Z. G. Li, F. Pan, K. P. Lim, X. Lin, and S. Rahardja, "Adaptive Rate Control for H.264" *Proc. IEEE International Conference on Image Processing,* September 2004, vol. 2, pp.745-748.

[4] M. Jiang and N. Ling, "Low-Delay Rate Control for Real-time H.264/AVC Video Coding," *IEEE Transactions on Multimedia*, vol. 8 no. 3, pp. 467-477, June 2006.

[5] M. Jiang and N. Ling, "On Lagrange Multiplier and Quantizer Adjustment for H.264 Frame-Layer Video Rate Control," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 16, no. 5, pp. 663 – 669, May 2006.

[6] Z. G. Li, C. Zhu, N. Ling, X. K. Yang, G. N. Feng, S. Wu, and F. Pan, "A Unified Architecture for Real-Time Video-Coding Systems," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 13, no. 6, pp. 472-487, June 2003.

[7] Y.-N. Xiao, H. Lu, X. Xue, V.-A. Nguyen, and Y.-P. Tan, "Efficient Rate Control for MPEG-2 to H.264/AVC Transcoding," *Proc. IEEE International Symposium on Circuits and Systems*, Kobe, Japan, May 2005, pp. 23-26.

[8] C.-W. Ho, O. C. Au, S. H. Gary Chan, S.-K. Yip, and H. M. Wong, "Low Complexity Rate Control for Efficient H.263 to H.264/AVC Video Transcoding," *Proc. IEEE International Conference on Image Processing,* Sep. 2006, vol.3, pp.652-655.

[9] M. Pantoja, N. Ling, and W. Shang, "Coefficient Conversion for Transform Domain VC-1 to H.264 Transcoding," *Proc. 2007 IEEE Workshop on Signal Processing Systems*, Shanghai, China, Oct. 17 - 19, 2007, pp. 363 – 367.

[10] J.-B. Lee and H. Kalva, "An Efficient Algorithm for VC-1 to H.264 Video Transcoding in Progressive Compression," *Proc. IEEE International Conference on Multimedia and Expo,* July 2006, pp.53-56.

[11] T. Chiang and Y.-Q Zhang, "A New Rate Control Scheme Using Quadratic Rate Distortion Model," *IEEE Transactions on Circuits and Systems for Video Technology,* vol. 7, no. 1, pp. 246-250, February 1997.

[12] T. Wiegand, H. Schwarz, A. Loch, F. Kossentini, and G. J. Sullivan, "Rate-Constrained Coder Control and Comparison of Video Coding Standards," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 13, no. 7, pp. 688-703, July 2003.

[13] M. Pantoja and N. Ling, "Low Complexity Rate Control for VC-1 to H.264 Transcoding," *Proc. 2009 IEEE International Symposium on Circuits and Systems*, Taipei, Taiwan, May 24-27, 2009, pp. 888-891.

[14] M. Jiang, X. Yi, and N. Ling, "Improved Frame-Layer Rate Control for H.264 Using MAD Ratio," *Proc. 2004 IEEE International Symposium on Circuits and Systems*, Vancouver, Canada, May 23 - 26, 2004, pp. III-813 – III-816 (Vol. III).

[15] X. Wang, Y. Zhang, H. Li, and W. Zhu, "Adaptive Rate Control for Dynamic Bandwidth in Video Transcoding," *Proc. IEEE International Conference on Communication Systems*, Nov. 19-21, 2008, pp.773 – 777.

[16] H. Zhen, H. Ruimin, Z. Li, and Q. Liu, "Macroblock Layer Rate Control in MPEG-2 to AVS-M Transcoder for Mobile Communication," *Proc. International Conference on Wireless Communications, Networking and Mobile Computing,* Sept. 21-25, 2007, pp.2829 – 2832.