International Journal of Modern Engineering Research (IJMER)
www.ijmer.com Vol.2, Issue.4, July-Aug. 2012 pp-2676-2679 ISSN: 2249-6645

Transcoding Method for Regions of Interest

Soon-kak Kwon
Department of Computer Software Engineering, Dongeui University, Korea

Abstract : *When a bitstream of the higher bitrate is converted to the lower bitrate through the transcoding process, the ROI (Regions of Interest) can be considered. We classify the ROI as the groups of macro-blocks with non-zero motion vectors. For simple implementation, the motion vectors are not additionally estimated, but extracted from the coded bitstream of the higher bitrate before trascoding. Then, a proposed method assigns a specific quantization step-size differentially according to the ROI within a video. The picture quality can be increased by applying the quantization step-size as a small value relatively for the ROI compared with the non-ROI. We can find the result that the proposed method gives the improved picture quality by assigning the quantization step-size differentially.*

Keywords - *ROI, Transcoding, Picture Quality*

I. INTRODUCTION

For both the homogeneous and heterogeneous transcodings, the bitrate reduction may be needed when the original higher bitrate cannot be accepted. The bitrate reduction algorithm converts from high bitrate bitstream to lower rate bitstream. Although there are several methods for bitrate reductions [1-6], the re-quantization method is well used because of the simple implementation and a less processing time.

All of transcoding methods deal with proper allocation of bitrate or quantization step-size in order to satisfy good picture quality or high PSNR (Peak Signal to Noise Ratio) on the constraint of the limited bitrate after transcoding. However, it is insufficient to meet the need of picture quality for ROI (Regions of Interest) within video sequence.

The several methods [7-12] have been researched for ROI coding, such as a classification of the foreground and background, a detection of the human face, a determination based on user attention model. However it did not include the bitstream transcoding schemes.

We propose a new transcoding method that makes better video quality relatively in ROI. ROI is classified to the moving regions. For searching moving regions, we use the coded information within the bitstream before transcoding. If the coded information before transcoding is used to find ROI, then the extra complexity is not needed. First, the non-zero motion vectors of each macro-blocks are grouped to separate the zero motion vectors. Then the groups with large numbers of non-zero motion vectors are set to ROI, the groups with small numbers of non-zero motion vectors are set to the non-ROI.

The organization of this paper is as follows. Section II provides the conventional relationship of bitrate and quantization step-size for transcoding. In Section III, the proposed transcoding method is described for improving the subjective picture quality by properly allocating the quantization parameters dependently on the ROI. Section IV shows the simulation results of the proposed method about quantization parameter compared with the conventional one. In Section V, we summarize the main results.

II. RELATIONSHIP OF BITRATE AND QUANTIZATION STEP-SIZE FOR TRANSCODING

The previous researches [13,14] showed the relationship between bitrate and quantization step-size for transcoding, a linear relative formula as a logarithmic function between bitrate (R) and quantization step-size ($Qstep$),

$$\log Qstep = b + a \cdot \log R \qquad (1)$$

where, b and a are dependent on the video characteristics. Also a transcoding algorithm has been presented which updates the model parameters given for the previous picture or slice using an approximated relationship between bitrate and quantization step-size according to the coded picture-type. That is, if a target bitrate in bits per second, BR_t(bps) is given in the case of already an encoded bitrate BR_1(bps) in bits per second, then target quantization step-size $Qstep_t$ is obtained from following equation.

$$Qstep_t = Qstep_1 \cdot \left(\frac{BR_t}{BR_1}\right)^a \qquad (2)$$

where, the value of a is approximated by the previous coded results.

Also our previous research[12] has proposed the transcoding method to control the picture quality depended on the subjectively ROI. The subjectively ROI determined using motion vectors were classified by following four different cases in the aspect of subjective importance.

(1) Central focus video: The focus is on the central regions rather than the whole frame in the case where object movement is primarily in the central regions of a picture. The background may either be fixed or moving uniformly through camera panning. This can be considered a default type if a frame is not classified into the other three types because humans tend to concentrate their viewpoint on the central region of a video.

(2) Peripheral focus video: The region of interest is focused on the contour in the case where the scene is being zoomed-out. The representative motion vector values of the top and bottom regions have non-zero values while those of the central regions are approximately zero.

(3) Upper focus video: The region of interest is in the upper regions in the case where object movement is located in the upper regions of a picture. The representative motion vector values in the upper regions of a frame have non-zero values with zero values in the lower regions.

(4) Lower focus video: The region of interest is in the lower regions in the case where object movement is located in the lower regions of a picture. The representative motion vector values at in the lower of a picture have non-zero values with zero values in the upper regions.

After the above classification of picture for the subjectively interest regions, the value of quantization step-size was differently assigned according to the location of slice of picture.

III. Proposed Transcoding Method for ROI

We propose a new transcoding method that makes better video quality relatively in ROI. ROI is classified as the groups of macro-blocks with non-zero motion vectors. For getting the motion vectors, we use the coded information within the bitstream before transcoding. First, the non-zero motion vectors of each macro-blocks are grouped to separate the zero motion vectors. Then the groups with large numbers of non-zero motion vectors are set to ROI, the groups with small numbers of non-zero motion vectors are set to the non-ROI.

Step1) Grouping the non-zero motion vectors of macro-blocks for the bitstream before transcoding

Step2) Eliminating the group with small numbers of non-zero motion vectors. The small number of macro-blocks are not important to give the subjective picture quality. Therefore, in this paper, the group having less than 15 macro-blocks is eliminated from the group of non-zero motion vectors.

Step 3) Classifying the ROI. ROI is the non eliminated groups and the others the non ROI. The groups of large number of non-zero motion vectors is only considered to ROI.

Step 4) Reassigning the quantization step-size. For the ROI, the quantization step-size is assigned to lower value than the target quantization step-size of Eq. (2), but for the non-ROI, the higher quantization step-size is assigned.

$$Qstep = Qstep_t - \alpha \text{, for ROI} \qquad (3)$$

$$Qstep = Qstep_t + \alpha \text{, for Non-ROI} \qquad (4)$$

where, $Qstep$ is final assigned quantization step-size, and α is a constant. α can affect to the difference in the quantization step-size between the maximum and the minimum values. A large value of α makes the large difference of picture quality between ROI and non-ROI.

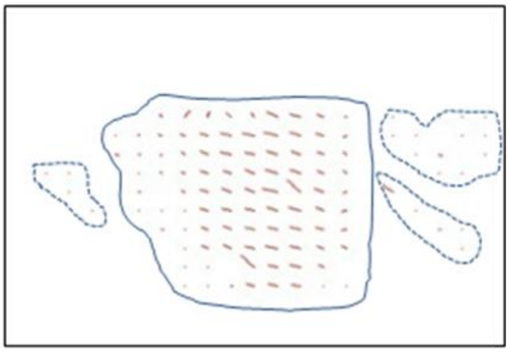

Fig 1. Grouping of macro-blocks according to non-zero motion vectors

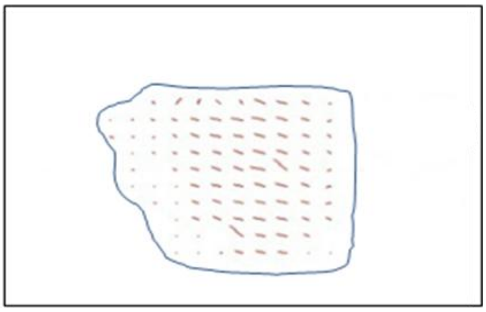

Fig 2. Eliminating of Grouped macro-blocks

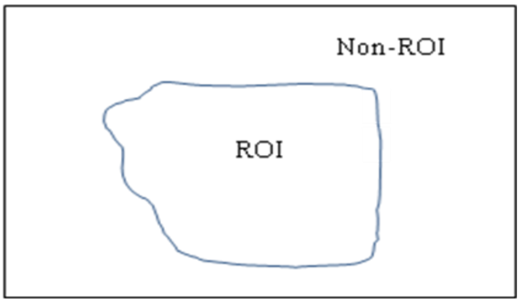

Fig 3. Classified ROI and Non-ROI

IV. Simulation Results

The JM [15-18] that is a standard coding software tool of the H.264 was used to the simulation where Foreman video with horizontal and vertical resolutions of 352 and 228 pixels was selected.

The B-picture is not included due to the use of the H.264 baseline profile, and 15 pictures are configured as one GOP in which the video applies 60 pictures. The number of slices is the same as for each picture and that are determined by 18 along the vertical direction.

The quantization step-size was differentially applied for each macro-blocks by using the value of determined in Eq. (3,4). Also, the coding was performed as an objective bitrate for the original video based on four different α of Eq. (3,4) , such as 2, 4, 6, and 8, from the classification of the difference between the maximum and the minimum values of the quantization step-size.

1) Measurement of the objective PSNR

This paper investigated the PSNR (peak signal to noise ratio) that is an objective picture quality criterion. The conventional method applied a bitstream coded with of 1.5Mbps, is transcoded to the bitrate of 0.8Mbps. Then, the quantization step-size was differentially assigned with the unit of macro-blocks using the proposed method for the Foreman video, according to the criterion of subjectively ROI.

Table 1 shows the average value of the PSNR for 60 pictures after transcoding in the Foreman video and maximum difference of quantization parameters within picture is set to 2, 4, 6 and 8. It can be seen that the values of the PSNR decreased according to the differential application of the quantization step-size among macro-blocks compared to the method that applies the same quantization step-size values for all macro-blocks within same slice. In the case of α =6, the PSNR was decreased about 0.07dB, and it was decreased about 0.28dB for the α =8.

Table 1. Average PSNR for the bitstream coded with of 1.5Mbps, is transcoded to the bitrate of 0.8Mbps.

Conventional method [dB]	Proposed method, α [dB]			
	2	4	6	8
37.83	37.81	37.76	37.70	37.55

2) Measurement of subjectively picture quality

For measuring of the subjective picture quality, we used DSCQS (Double Stimulus Continuous Quality Scale) [19]. According this method, the assessors are positioned at a distance from the monitor equal to three times the diagonal length of the monitor used to display the videos. They then observe two videos in sequence on the monitor; one is an original video and the other is a video either using the conventional method or the proposed method. The presentation order of the original and processed videos was random.

The presentation of the test material:

1 Video A (Original or Processed) 12s,
2 Gray 3s,
3 Video B (Processed or Original) 12s,
4 Gray 3s

Assessors evaluated the picture quality of both videos using an ITU-R quality scale (Excellent=5, Good=4, Fair=3, Poor=2, Bad=1) [12,20].

The final subjectively picture quality assessment score was calculated from the mean over all assessors;

$$U = \frac{1}{N}\sum_{i=1}^{N} u_i \qquad (5)$$

where is score which is determined by each assessor, and N is number of assessors.

The measurement of the subjectively picture quality was performed by distributing reconstructed videos to 10 assessors without notification of the methods and the ROI. We got the same average score of the subjectively picture quality obtained from 10 evaluators as shown in Table 2.

Table 2. Results of the subjectively picture quality for the bitstream coded with of 1.5Mbps, is transcoded to the bitrate of 0.8Mbps .

Conventional method	Proposed method, α			
	2	4	6	8
3.0	3.1	3.2	**3.4**	3.3

In the results, because the Foreman video showed no movements in the background and some movements at the central region, it represented good picture quality in the aspect of a subject manner based on the assigning of the quantization step-size with ROI of the central moving regions. The desirable difference between the maximum and the minimum values of the quantization step-size between ROI and non-ROI was 6. If the difference between the maximum and the minimum values showed more than 8, then the subjectively picture quality was getting worse.

Although the objective PSNR value that is the average value of the entire picture slightly was decreased for the differential assign of the quantization step-size within the video sequence according to ROI, the subjective picture quality that has a focus on interest points within the video sequence was increased. It means that the measurement of objective picture quality only will not fully satisfy the visual picture quality of human being. In addition, it is possible to present an improvement effect in picture quality by applying the results of the evaluation of subjective picture quality to the objective picture quality while there is no change in bitrates.

Fig .4 illustrates the reconstructed picture of Foreman video for the bitstream coded with of 1.5Mbps, is transcoded to the bitrate of 0.8Mbps with conventional method and the proposed methods of different α . When α =6, the subjective picture quality of Foreman video looked better relatively. When α =8, although the subjectively picture quality of ROI was getting best, the subjective picture quality of non-ROI was getting the coarsest picture quality level.

(a) Reconstructed frame of conventional method

(b)Reconstructed frame of proposed method, α =4

(c)Reconstructed frame of proposed method, α =6

(d) Reconstructed frame of proposed method, α =8

Fig. 4 Comparison of the four reconstructed frames of Foreman video

V. CONCLUSION

This paper classified ROI and non-ROI within the video according to motion vectors. The moving regions of non-zero motion vectors is set to focus region, then quantization step-size (parameter) is assigned to the focus regions differentially compared to the non focus regions. Proposed method obtained an improvement effect in subjective picture quality compared to the conventional method through the intensive application of the quantization step-size.

Although this study focused on the H.264 video coding, it can be directly used to other video coding standards including the HEVC[21]. In addition, it can be expected that this research will be extensively applied with other various bitstream transcoding techniques.

ACKNOWLEDGEMENTS

This work was supported by Dongeui University Foundation Grant (2010)..

REFERENCES

[1] T. Shanableh and M.Ghanbari, Heterogeneous video transcoding to lower spatio–temporal resolutions and different encoding formals, *IEEE Trans. Multimedia*, 2(2), 2000, 101-110.

[2] M. Xia, A. Vetro, B. Liu, and H. Sun, Rate-distortion optimized bit allocation for error resilient video transcoding, *Proc. IEEE Int. Symp. Circuits and Systems*, 2004, 945-948.

[3] W. Lie, M. Tsai, and T. Lin, Rate-distortion optimized DCT-domain video transcoder for bit-rate reduction of MPEG videos, *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, 5, 2004, 969-972.

[4] L. Chen, L. Cheng, and S. Yu, Accurate rate control method in transcoding," *Electron. Lett.*, 40(1), 2004, 16-18.

[5] J. Lei and N. Georganas, Accurate bit allocation and rate control for DCT domain video transcoding, *Proc. IEEE Canadian Conf. Electrical and Computer Engineering*, 2002, 968-973.

[6] H. J. Lee, T. Chiang, and Y. Q. Zhang, Scalable rate control for MPEG-4 video, *IEEE Trans. Circuit Syst. Video Technology*, 10(9), 2000, 878-894.

[7] D. Chai, K. N. Ngan and A. Bouzerdoum, Foreground / background bit allocation for region-of-interest coding, *Proc. Int. Conf. on Image Processing*, 2000, 438 - 441.

[8] M. J. Chen, M. C. Chi, C. T. Hsu, and J. W. Chen, ROI video coding based on H.263+ with robust skin-color detection technique, *IEEE Trans. Consumer Electronics*, 49, 2003, 724-730.

[9] M.-C. Chi, C.-H. Yeh, and M.-J. Chen, Robust region-of-interest determination based on user attention model through visual rhythm analysis, *IEEE Trans. Circuit Syst. Video Technology*, 19(7), 2009, 1025-1038

[10] L. Tong and K. R. Rao, Region of interest based H.263 compatible codec and its rate control for low bit rate video conferencing, Proc. *Int. Symp. Intelligent Signal Processing and Communication Systems*, 2005, 249-252.

[11] H. J. Song and C. C. J. Kuo, A region-based H.263+ codec and its rate control for low VBR video, *IEEE Trans. Multimedia*, 6(2), 2004, 489-500.

[12] S.-k. Kwon, A. Punchihewa, D. G. Bailey, S.-W. Kim, and J. Lee, Adaptive simplification of prediction modes for H.264 intra-picture coding, *IEEE Trans. Broadcasting*, 58(1), 2012, 125-129.

[13] Y.-l. Wu, S.-k. Kwon, Transcoding method of H.264 coded bitstream for interest region, *The International Journal of Virtual Reality*, 8(3), 2009, 45-50.

[14] S.-k. Kwon, S.-w. Kim, J.-h. Lee, and J.-m. Lee, An adaptive transcoding method, *International Journal of Computer Science and Network Security*, 8(12), 2008, 154-160.

[15] S. k. Kwon, A. Tamhankar and K. R. Rao, Overview of H.264/MPEG-4 part 10, *Journal of Visual Communications and Picture Representation*, 17(2), 2006, 186-216.

[16] ITU-T Rec. H.264 / ISO/IEC 11496-10, Advanced video coding, Final Committee Draft, Document JVT-F100, 2002.

[17] T. Widegand, G. Sullivan, G. Bjonttegaard and A. Luthra, Overview of the H.264/AVC video coding standard, *IEEE Trans. Circuit Syst. Video Technology*, 13(7), 2003, 560-576.

[18] H.264/AVC JM Software, http://iphome.hhi.de/suehring/tml/download.

[19] M. Bernas, Image quality evaluation, *Int. Symp. Video/Image Processing and Multimedia Communications*, 2002, 133-136.

[20] ITU-R Rec. BT. 500, Methodology for the subjective assessment of the quality of television pictures, 2002, Geneva.

[21] K. Ugur, K. Andersson, A. Fuldseth, G. Bjøntegaard, L. P. Endresen, J. Lainema, A. Hallapuro, J. Ridge, D. Rusanovskyy, C. Zhang, A. Norkin, C. Priddle, T. Rusert, J. Samuelsson, R. Sj¨oberg, and Z. Wu, High performance, low complexity video coding and the emerging HEVC standard, *IEEE Trans. Circuit Syst. Video Technology*, 20(12), 2010, 1688-1697.