CODING MODE DECISION FOR HIGH QUALITY MPEG-2 TO H.264 TRANSCODING

H. Kato, A. Yoneyama, Y. Takishima

KDDI R&D Laboratories Inc.
Multimedia Communications Laboratory
2-1-15 Ohara, Fujimino-shi, Saitama, Japan

Y. Kaji

Tokyo University of Science
Department of Engineering
1-3 Kagurazaka, Shinjuku-ku, Tokyo, Japan

ABSTRACT

This paper proposes a high quality coding mode decision method for MPEG-2 to H.264 transcoding. Here, we evaluate the coding modes of MPEG-2 stream in order to specify whether they should be inherited or changed in H.264 coding since many new coding schemes such as intra prediction have been introduced in H.264. By adaptively determining suitable intra/inter modes according to DCT coefficients, motion vectors (MVs), and neighboring macroblock (MB) modes of input MPEG-2 stream, high quality transcoding has been realized. In the experiment, we observed that the proposed method can improve a peak signal to noise ratio (PSNR) up to 0.29 dB at the cost of 22% more processing time compared with the conventional transcoding method.

Index Terms— Video codecs, Transform coding, Image converters, Image edge analysis, Discrete cosine transforms

1. INTRODUCTION

H.264 has been widely used in a variety of video applications including IPTV, iPod/PSP, and mobile phones due to its high coding efficiency. Although newly available source contents can be directly encoded in H.264, legacy contents which are already encoded in MPEG-2 used for such applications as TV broadcasting and DVD should be converted to H.264 in order to effectively make use of bandwidth and/or storage space. In order to realize fast transcoding, coding parameters such as MVs and coding modes of input coded stream are often inherited to the encoder. Although this approach can be applied when the input and output coding formats are similar, effective inheritance would become difficult when their coding schemes are different. As for H.264, when compared with MPEG-2, it has introduced many new coding modes such as 8x16/8x8 motion compensation and intra prediction in order to realize higher coding efficiency. Therefore, effective inheritance of MPEG-2 coding parameters to H.264 has become a key issue to realize a fast and high quality transcoding.

In this paper, we propose a fast transcoding method from MPEG-2 to H.264 while maintaining conversion quality by evaluating coding parameters such as DCT coefficients and coding modes of MPEG-2 and by estimating the appropriate intra/inter coding modes. In the following sections, firstly we describe related works of MPEG-2 to H.264 transcoding and then explain the proposed transcoding algorithm followed by experimental results.

2. RELATED WORK

Generally, transcoders can be classified into two types, one that operates in the coded domain and the other in the pixel domain.

A transcoder operating in the coded domain has a probability to significantly reduce computational load. [1] proposed a MPEG-2 to H.264 transcoding based on the error drift compensation in the coded domain. Although it can achieve fast and high quality transcoding reusing all the coding parameters such as MVs and macroblock coding modes, further improvement would be obtained if MV refinement and appropriate coding modes change are performed. However, H.264 transcoding in the coded domain still remains a challenging task because H.264 significantly differs from MPEG-1/2/4. For example, H.264 employs new coding techniques such as an in-loop filter and intra prediction that are not used in MPEG-1/2/4. The in-loop filter graduates pixels at block boundaries in order to reduce block noise. Intra prediction can reduce bit rate by using spatial correlation from the surrounding pixels. Since both techniques are processed in the pixel domain, transcoding of such processes in the coded domain may require larger calculation load rather than transcoding in the pixel domain.

On the other hand, the processing speed of pixel domain transcoding which involves decoding and re-coding can be accelerated by using input coded data as picture analysis tool or by re-using input coding modes in the encoder [2–4]. For example, [2] proposed the intra mode decision for MPEG-2 to H.264 transcoding. It computes the edge angle in a block from the MPEG-2 DCT coefficients and limits the prediction mode, which applies rate distortion optimization (RDO) according to the direction of the edge angle, which can reduce over 30% in the processing time. However, this supports only intra mode decision for a transcoder. As for the coding mode reuse, [4] proposes to reuse the MB types in MPEG-2 in the

Table 1. Features of a MB that the coding mode should be changed

MPEG-2 inter to H.264 intra	· MB that was failed in ME · MB with high spatial correlation vertically or horizontally · MB that has changed shape and color in the reference frame
MPEG-2 intra to H.264 inter	· MB without texture · MB whose luminance is flat

H.264 encoder, and about 60% of the transcoding time has been reduced. In general, these fast algorithms often employ direct inheritance of MPEG-2 coding modes in order to realize a fast processing, therefore improvement of transcoding quality is limited since newly introduced coding schemes to H.264 will not be used. For example, inter frame prediction in H.264 is improved by introducing 8x8/8x16/16x8 motion compensation block sizes to conventional 16x16 block which is uniquely used in MPEG-2. In the next section, we describe an efficient coding mode decision algorithm to solve the above problem.

3. PROPOSED MODE DECISION METHOD USING CODED DATA IN MPEG-2

Since H.264 includes many coding schemes within the intra and inter coding modes, detailed investigation of spatial / temporal correlation is necessary to make full use of its coding efficiency potential. The preliminary experiment has shown that an inappropriate mode decision such as direct inheritance of MPEG-2 coding modes often results in transcoding degradation. Therefore, intra coding and inter coding in MPEG-2 may be changed to inter coding and intra coding in H.264 according to the correlation. Table 1 shows the summary of cases where such change should be considered. For example, when the MVs of input MPEG-2 are diverse due to flat regions and therefore reliability of motion estimation is low, then coding mode should be changed to intra coding since the cost of inter coding including MV information would be larger than that of intra coding. Therefore, we estimate MB types of H.264 from the MPEG-2 coded data. Fig. 1 shows the flowchart of the proposed mode decision for an MPEG-2 to H.264 transcoder to evaluate the features in Table 1. In the following subsections, we describe decision algorithm for each coding mode.

3.1. MV check for the inter mode of MPEG-2

Fig. 2 shows the MVs of a MB that changed from the MPEG-2 inter mode to the H.264 intra mode using baseband conversion. It can be seen that MV field is diverse especially in football field where motion in relatively flat region is involved. Therefore, we use the correlation of MV to evaluate

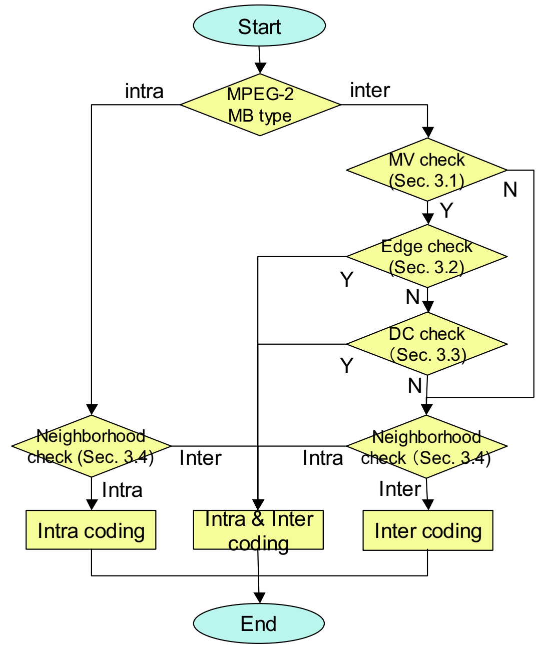

Fig. 1. Proposed mode decision for an MPEG-2 to H.264 transcoder

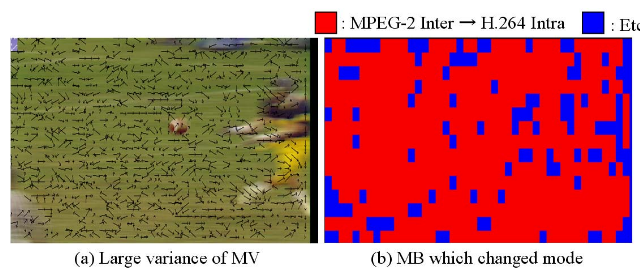

Fig. 2. Relationship between MV and a MB whose optimal mode changed from MPEG-2 inter to H.264 intra

the reliability of motion estimation (ME) since it differs from the neighboring MV when ME is inappropriate.

$$\sigma^2_{MV_X} > TH_{MV} \,||\, \sigma^2_{MV_Y} > TH_{MV} \quad (1)$$

where σ^2_{MV} is the variance of MV in neighboring blocks, and is calculated using both horizontal vertical components. TH_{MV} expresses the threshold. When the variance of MV in Eq. (1) is large, it is necessary to evaluate other features for the intra mode of H.264. Otherwise, it decides that ME is reliable and only inter coding is applied.

3.2. Edge check for the inter mode of MPEG-2

Fig. 3 shows the DCT coefficients of a MB that were changed from the MPEG-2 inter mode to the H.264 intra mode by tandem baseband transcoding. As can be seen in the figure, coding mode changes are often occurred in relatively large

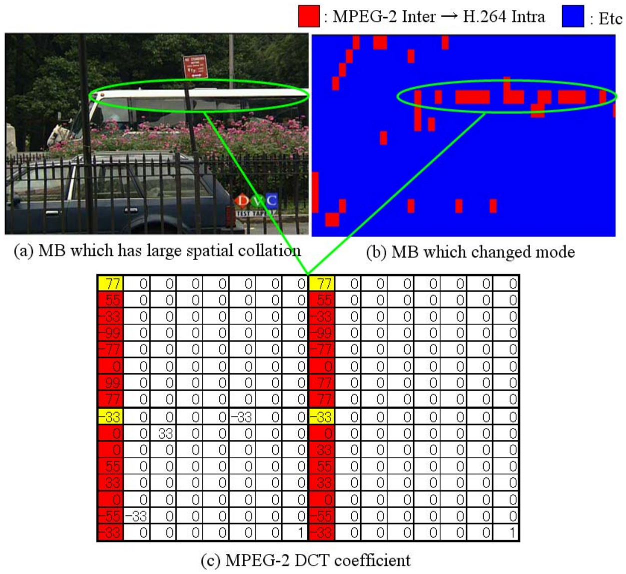

Fig. 3. DCT coefficients in a MB whose mode changed from MPEG inter to H.264 intra

edge regions where cost of intra coding becomes comparable with that of inter coding. Here, we check edge components classifying a MB as one of four types, using the DCT coefficient of MPEG-2 to evaluate the edge feature in the MB. The four types are NoEdge, VertEdge, HorEdge and VertEdge+HorEdge. The probable existence of an edge is determined by Eqs. (2) and (3).

$$Sum_{AC} = \sum_{v=0}^{7}\sum_{u=0}^{7} \left|F^i_{Y8x8}(u,v)\right| - F^i_{Y8x8}(0,0) \quad (2)$$

$$Sum_{AC} < TH_{NoEdge} \quad (3)$$

If the summation of AC Sum_{AC} in Eq. (2) is smaller than the threshold TH_{NoEdge}, a MB will be classified as NoEdge. Otherwise the direction of the edge is evaluated by Eqs. (4) and (5). A MB is classified as VertEdge when the horizontal partial sum is larger than the threshold $TH_{VertEdge}$. Alternatively, a MB is classified as HorEdge when the vertical partial sum is larger than the threshold $TH_{HorEdge}$.

$$\sum_{v=1}^{7} \left|F^i_{Y8x8}(0,v)\right| / Sum_{AC} > TH_{VertEdge} \quad (4)$$

$$\sum_{u=1}^{7} \left|F^i_{Y8x8}(u,0)\right| / Sum_{AC} > TH_{HorEdge} \quad (5)$$

where F^i_{Y8x8} is the 8x8 DCT coefficient of the luminance component in MPEG-2. Suffix i expresses the position of the individual 8x8 DCT coefficient in a MB as shown in Fig. 4.

When the determination for a MB is the combination (edge pattern) defined in Fig. 5, it is suitable for the intra prediction

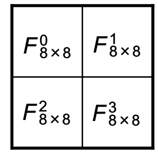

Fig. 4. MPEG-2 macro block

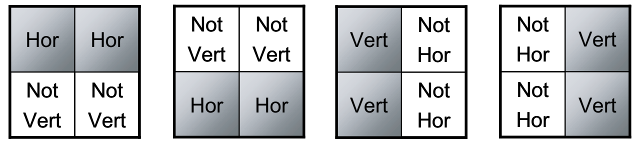

Fig. 5. Edge pattern

of H.264. Therefore, even if it is in the inter mode of MPEG-2, both intra and inter coding will be applied in order to evaluate by RDO process in H.264 since opposite mode may be effective.

3.3. DC check for the inter mode of MPEG-2

We also evaluate non-edge regions in order to investigate whether MPEG-2 inter coding should be changed to H.264 intra. Here, we investigate DC components of MPEG-2 DCT coefficients to check if DC prediction of H.264 work effectively. Otherwise, it is desirable to inherit inter since the intra coding does not work more effective than the inter coding.

The residual of prediction is estimated by Eqs. (6) and (7) using the DCT coefficient. If the required conditions are fulfilled, the coding mode will be determined by RDO. Otherwise, only inter coding is applied.

$$\sum_{i=0}^{3} \left|F^i_{Y8x8}(0,0)\right| > TH_{DC1} \quad (6)$$

$$F_{U8x8}(0,0) > TH_{DC2} \;||\; F_{V8x8}(0,0) > TH_{DC2} \quad (7)$$

where F^i_{U8x8} and F^i_{V8x8} are 8x8 DCT coefficients of the chrominance in MPEG-2. TH_{DC1} and TH_{DC2} express the threshold for luminance and chrominance, respectively.

3.4. Neighborhood check for the inter and intra mode of MPEG-2

In order to investigate spatial correlation, we also check coding modes of neighboring MBs. It is applied to both intra and inter mode. The spatial correlation is estimated by majority rule of coding modes in the neighboring blocks. When a MB has a high spatial correlation as shown by the coding modes in H.264 at left, top and top-right neighbor blocks, the coding mode of MPEG-2 can be inherited in H.264 since the bit rate can be reduced effectively. Otherwise, the coding mode will be determined by RDO.

Table 2. Simulation results

Sequence	Method	ΔTime* (%)	BDPSNR (dB)	BDBR (%)
Bus	Conventional	66.0	-0.10	3.11
	Proposed	57.5	-0.04	1.14
Flower	Conventional	67.8	-0.02	0.59
	Proposed	65.8	-0.01	0.27
Football	Conventional	59.8	-0.36	11.75
	Proposed	37.7	-0.07	2.32
Mobile	Conventional	68.3	-0.02	0.49
	Proposed	66.6	-0.01	0.37

(* Lager ΔTime means faster processing.)

4. SIMULATION RESULTS

In order to verify the effectiveness of the proposed method, experiments were conducted to compare the processing time and PSNR against the original uncompressed sequences among the proposed method, conventional method [4] and JM10.0 [5]. As input sequences, *Football* (6 Mbps), *Bus* (8 Mbps), *Flower* (10 Mbps) and *Mobile* (12 Mbps) were used. The coding parameters of the input for TM5 were set to MPEG-2 (720 * 480, M=1, N=15, 150 frames), and the output condition was set to H.264 (QP=28, 30, 32, 34). The H.264 encoder employed a fixed QP in order to eliminate the influence of rate control. The thresholds were set based on our preliminary experiment. Processing time and PSNR comparisons are shown in Table 2. For the evaluation in different QPs, we adopted BDPSNR and BDBR in [6], which represent average difference of PSNR and bit rate between two methods, respectively.

As shown in Table 2, the proposed method reduced processing time by 37.7% to 66.6% compared with JM. The degradation of the PSNR was less than 0.07 dB. We also confirmed that the proposed method has no subjective degradation.

When compared with the conventional method, it can be seen that although the improvement of PSNR in the proposed method is very limited in *Flower* and *Mobile* sequences, larger quality improvement (up to 0.29dB (=-0.07-(-0.36))) is observed in *Bus* and *Football* sequences. As for the bit rate, it is equivalent to 9.43% (=11.75-2.32) reduction. This improvement confirms that the effective MPEG-2 coding mode investigation and H.264 coding mode estimation has been performed. One of the examples can be seen in Fig. 2 where MPEG-2 inter mode is changed to H.264 intra mode in *Football* sequence.

Fig. 6 shows the PSNR difference of every frame in *Football* (QP=34) when compared with the conventional method. As can be seen in the figure, much more stable transcoding can be obtained in the proposed method since appropriate mode decision leads to avoid coding mode mismatch between MPEG-2 and H.264.

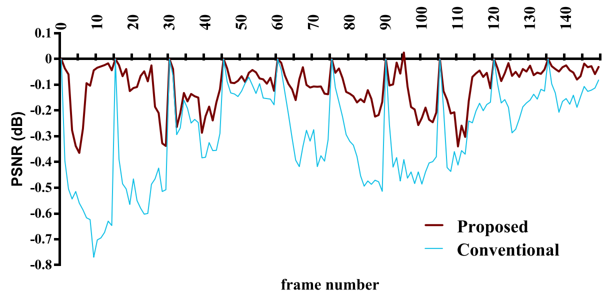

Fig. 6. Comparison of difference PSNR (*Football*, QP=34)

5. CONCLUSION

This paper proposes a coding mode decision for high quality MPEG-2 to H.264 transcoder in the pixel domain. We examined the novel coding mode decision method that evaluates the effectiveness of the coding mode in H.264 using MPEG-2 coded data. Simulation results showed that BDPSNR improved up to 0.29 dB at the cost of 22% more processing time compared with the conventional method.

6. ACKNOWLEDGMENTS

The authors would like to thank Dr. S. Akiba, Dr. Y. Nakajima and Mr. H. Yanagihara for their continuous support.

7. REFERENCES

[1] T. Qian and et al., "Transform Domain Transcoding from MPEG-2 to H.264 with Interpolation Drift-Error Compensation," *IEEE Trans. CSVT*, vol. 16, no. 4, pp. 523–534, 2006.

[2] H. Kalva and et al., "Exploiting the Directional Features in MPEG-2 for H.264 Intra Transcoding," *IEEE Trans. CE*, vol. 52, no. 2, pp. 706–711, 2006.

[3] G. F. Escribano and et al., "Speeding-up the Macroblock Partition Mode Decision in MPEG-2/H.264 Transcoding," in *IEEE ICIP*, 2006, pp. 869–872.

[4] D. Kim and et al., "A Fast MPEG2-to-H.264 Transcoding Algorithm in Spatial Domain," in *IWAIT*, 2006, pp. 695–700.

[5] K. Suehring, "H.264 / AVC reference software," http://bs.hhi.de/ suehring/tml.

[6] G. Bjontegaard, "Calculation of average PSNR differences between RD-curves," Tech. Rep. VCEG-M33, ITU-T SG16, 2001.