IEEE Int. Conference Neural Networks & Signal Processing
Zhenjiang, China, June 8~10, 2008

FAST INTRA MODE DECISION ALGORITHM FOR H.263 TO H.264/AVC TRANSCODING

Xuan Jing[1], Wan-Chi Siu[1], Lap-Pui Chau[2] and A.G. Constantinides[3]

[1]Department of Electronic and Information Engineering, Hong Kong Polytechnic University, Hung Hom, Hong Kong
[2]School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore, 639798
[3]Department of Electrical and Electronic Engineering, Imperial College London, London, SW7 2AZ, UK
E-mail: {enxjing, enwcsiu}@polyu.edu.hk , elpchau@ntu.edu.sg, a.constantinides@imperial.ac.uk

ABSTRACT

In this paper, we present a fast mode decision method for H.263 to H.264/AVC intraframe transcoding. The proposed algorithm is based on the observation that features of DCT coefficients extracted from H.263 coded intraframe strongly relates to the optimal mode in H.264/AVC intra mode decision. Specifically, the total number of non-zero AC coefficients of four 8x8 DCT blocks in H.263 is taken as a measurement in the intra block size decision. The objective is to reject improbable block types in earlier stages in order to achieve computation saving. In addition, a transform domain edge direction estimation is also adopted into our scheme to further speed up the intra mode prediction. Simulation results show that the proposed approach can reduce the computational complexity of intraframe transcoding by up to 65% while maintaining the rate-distortion (R-D) performance.

Key Words — Transcoding, Mode decision, H.264/AVC

1. INTRODUCTION

H.264/AVC [1] is so far the most efficient standard for video compression which achieves much higher coding efficiency than previous video coding standards. Due to its significant improvement in compression capability over H.263 [2], H.264/AVC is more suitable for use in a wide range of applications such as HDTV and videoconferencing devices. Although H.264/AVC is expected to replace the use of H.263 in the near future, it is of great importance for transcoding videos between these two formats. In this work, we address the problem of transcoding intraframes from H.263 to H.264/AVC. Compared with the baseline profile of H.263 where there is no intra prediction, the H.264/AVC enables a rich set of intra prediction modes which can eliminate the spatial redundancy considerably. Generally, nine prediction modes for 4x4 luma blocks, four modes for 16x16 luma blocks and four modes for 8x8 chroma blocks are employed in H.264/AVC. The best coding mode for each macroblock (MB) is selected using rate-distortion optimization (RDO) technique which aims at minimizing the Lagrangian cost function.

Exhaustively examining all prediction modes in the H.264/AVC re-encoding stage is extremely time consuming. Thus to design fast mode decision algorithms in transcoding becomes highly desirable for real-time low delay and low complexity applications [3,4,5]. In [6], a fast MB partition mode decision algorithm based on data mining tools for interframe transcoding from H.263 to H.264/AVC is proposed. Several fast intra mode decision methods for H.264/AVC have been proposed recently [7][9]. As in [7], Pan *et al* presented an edge detection algorithm using Sobel operator to speed up the mode decision process. Similarly, Su *et al* [9] predicted the texture directions for each block based on the transform domain processing. These techniques can be regarded as good candidates for H.263 to H.264/AVC intraframe transcoding method. However, since no explicit fast algorithm to select between I16MB and I4MB block type is available in Pan's and Su's methods, directly employing this kind of algorithms into a transcoder will not perform optimally in terms of computational complexity.

In this paper, we develop a fast intra block type prediction scheme to enhance the efficiency of the existing intra mode decision method used in the H.263 to H.264/AVC transcoder. Based on the observation we find that the number of non-zero AC coefficients of one MB extracted from H.263 input bitstream has very high correlation with its optimal intra block type in the H.264/AVC re-encoding process. Therefore we use this feature to form an efficient intra block type classification algorithm. As a result, computations for checking unlikely block types in intra mode decision can be reduced. On the other hand, we also adopt the existing directional prediction techniques to further reduce the number of candidate intra prediction modes. Experimental results reveal that by applying our proposed scheme, significant time saving for the intraframe transcoding can be achieved as compared to the traditional method. In addition, the R-D performance in terms of PSNR and the bit rate are maintained.

The rest of the paper is organized as follows. In Section 2 we briefly review the intra mode decision defined in

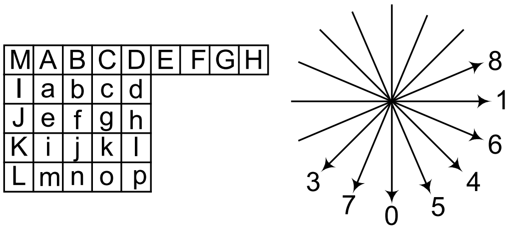
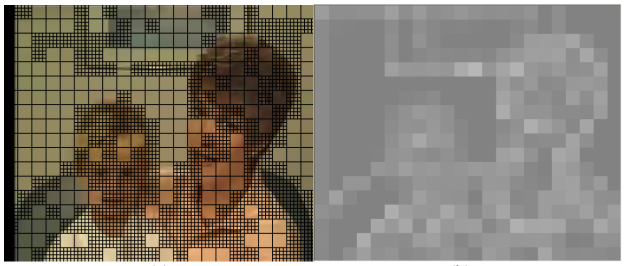


Fig. 1. Pixels used in I4MB mode decision and its prediction modes of different directions



(a) (b)

Fig. 2. Correlation between the optimal intra block type and the value of *nzcoef*. (a) Sample frame from *Mather&Daughter* with intra block type for each MB (b) Corresponding value of *nzcoef.* An offset intensity value of 128 is added to each pixel in (b) for display purpose.

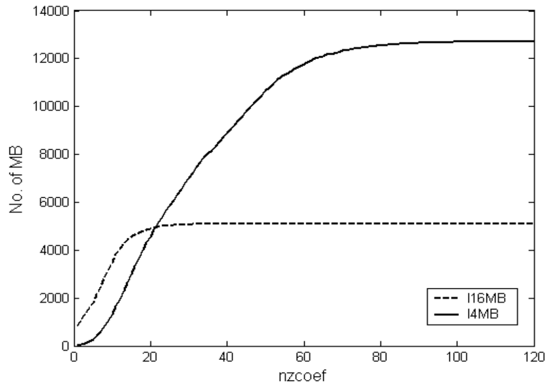


Fig. 3. Cumulative distribution of I16MB mode and I4MB mode

H.264/AVC. An efficient intra block type selection algorithm is introduced in Section 3. Section 4 describes the overall fast intra mode decision scheme for our proposed intraframe transcoder. The experimental results are demonstrated in Section 5 for comparison. Finally, we draw a conclusion in Section 6.

2. INTRA PREDICTION METHOD IN H.264/AVC

The intra coding algorithm of H.264/AVC differs from previous video coding standards in that intra prediction in spatial domain has been conducted. For each MB, there are two intra prediction block types for luma component i.e., I16MB and I4MB. Besides, I8MB is used for both chroma components. As for I16MB and I4MB, each 16x16 or 4x4 block is predicted from its spatially adjacent pixels according to the particular prediction modes. For example, Fig. 1 illustrates the pixels used in I4MB mode and its corresponding prediction directions. Totally there are nine possible prediction modes available for I4MB block type, and four modes for I16MB as well as I8MB. To select the best coding mode for one MB, normally the encoder needs to perform RDO mode decision to check every possible mode and finally choose the one with the minimum R-D cost as computed according to the following equation,

$$J = D + \lambda \cdot R \qquad (1)$$

where D and R represent the distortion and the number of bits for coding the MB respectively. λ is the Lagrangain multiplier which is defined by the standard. In general, a total of $4 \times (16 \times 9 + 4) = 592$ RDO operations have to be performed for a given MB [7]. This would be a huge computational burden for some low complexity real-time applications.

3. EFFICIENT INTRA BLOCK TYPE SELECTION

Variable block size technique plays an important role in H.264/AVC for both intraframe and interframe coding. Basically, larger block types are more suitable for coding homogeneous regions within a video frame while smaller block types are more often selected for coding non-homogeneous regions. If we can perform accurate block type classification before exhaustively examining all prediction modes for both large and small block types, unnecessary computations can be avoided without affecting the prediction performance. In order to perform such block type classification, a certain feature has to be selected as a measure. Particularly, for real-time low complexity transcoding applications, the feature selected should have the following two important properties. First, the extra computations required to calculate this measure must be as low as possible. Second, the feature should have strong correlation with the optimal block types.

For H.263 to H.264/AVC transcoding scenario, much useful information which can represent the spatial characteristics has already been encoded in the compressed bit stream. Therefore, it is beneficial to take advantage of such information in the transcoder design. In this work, we consider a very effective feature as our measure in block type classification, *nzcoef*, *i.e.*, the number of non-zero AC coefficient of a MB. Note that, in our scheme, *nzcoef* stands for the total number of non-zero AC coefficients for four 8x8 DCT blocks of a given MB in H.263 compressed data. Fig. 2 shows the correlation between the value of *nzcoef* of each MB in H.263 and its optimal block type after H.264/AVC re-encoding process. In Fig. 2(b), the gray level intensity value of each MB is set as *nzcoef.* The larger the *nzcoef* the brighter the MB will be and vice versa. By comparing Fig. 2(a) with Fig. 2(b), it can be seen that, the value of *nzcoef* has very strong correlation with the spatial complexity in a given video frame. Normally, the

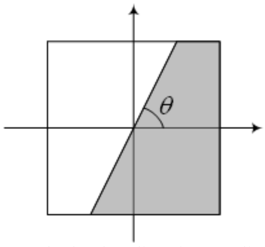


Fig. 4. Block edge direction prediction

areas coded using smaller block type I4MB in Fig. 2(a) will have bigger intensity value in Fig. 2(b). This indicates that *nzcoef* is a useful measure which contains the spatial domain information important to our intra block type classification.

To develop a systematic block type selection scheme, we have conducted extensive experiments to further differentiate both block types, I4MB and I16MB, using a set of criteria. In our experiments, a number of 17000 MBs extracted from H.263 intra-coded frames of various video contents are used to generate our rule for block type classification. After decoding from H.263 bit stream, the decoded videos are re-encoded by H.264/AVC encoder using full search intra mode decision method with different quantization parameters. Fig. 3 illustrates the cumulative distribution of I4MB mode and I16MB mode. It can be observed from Fig. 3 that, when the value of *nzcoef* is larger than a certain threshold, the probability for the MB being coded using I16MB is extremely low. On the contrary, when *nzcoef* is less than a certain value, most likely the I16MB will always be chosen as the best mode. In addition, we also find that, it is difficult to use only one threshold to conduct an accurate binary classification because there exists a range that both block types are possibly be selected as the best mode. Based on the observation, we propose to classify the possible candidate intra block type by the following rules as defined in Eq. (2).

$$\text{mode} = \begin{cases} \text{I16MB} & \text{if } nzcoef < TH1 \\ \text{I16MB \& I4MB} & \text{if } TH1 \le nzcoef < TH2 \\ \text{I4MB} & \text{if } TH2 \le nzcoef \end{cases} \quad (2)$$

In other words, after decoding the current intra frame from H.263 input bit stream, we first determine the possible intra block types to be used before performing H.264/AVC intra mode decision. Therefore, unnecessary computations for checking the improbable block types for each MB will be avoided. According to our extensive experiments, the two thresholds TH1=6 and TH2=16 are adopted in our scheme as they can provide a good balance between mode decision complexity and accuracy.

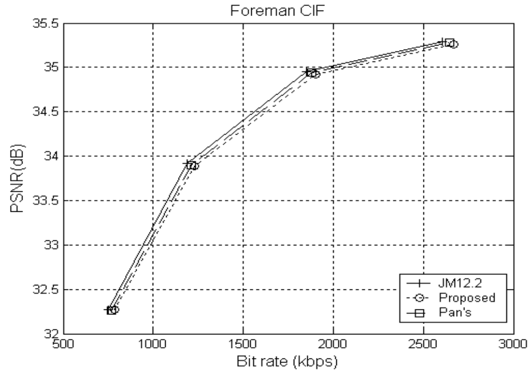
4. FAST INTRA MODE DECISION SCHEME

In this section, we present our fast prediction mode decision scheme which is performed after previous intra block type selection. As mentioned in Section 2, there are nine prediction modes for luma I4MB mode and four prediction modes for luma I16MB as well as for chroma I8MB mode. The target of our proposed method is to reduce the number of candidate prediction modes such that further time savings can be achieved. When I4MB mode is chosen, each of nine prediction modes (mode 2 is DC mode) has to be evaluated for every 4x4 block within one 16x16 MB. In [8][9], transform domain coefficients are utilized to determine the edge direction for a block. Specifically, for a spatial domain 4x4 block, the edge direction can be effectively computed based on its corresponding few transform domain AC coefficients as in Eq. (3)

$$\tan\theta = \frac{F_{0,1} + F_{0,2} + F_{0,3}}{F_{1,0} + F_{2,0} + F_{3,0}} \quad (3)$$

where θ is the angle between the edge direction and the horizontal axis as shown in Fig. 4. $F_{u,v}$ are the corresponding transform domain AC coefficients of this particular 4x4 block. In H.264/AVC, 4x4 integer transform (IT) is adopted instead of the traditional 8x8 DCT transform in H.263. After decoding one frame from the H.263 input bitstream, the proposed transcoder will perform the 4x4 integer transform on the decoded frame to obtain its IT coefficients. The core transform is shown in Eq. (4)

$$Y = H \cdot X \cdot H^T \quad (4)$$



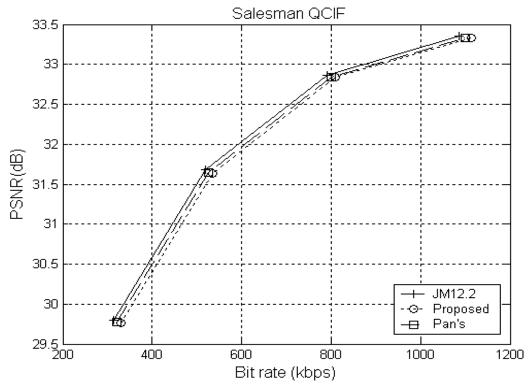


Fig. 5. R-D performance of different intra mode prediction methods

TABLE I EXPERIMENTAL RESULTS FOR DIFFERENT INTRAFRAME TRANSCODERS

	Sequence	Pan's method			Proposed method		
		Δ PSNR (dB)	Δ Bitrate (%)	Δ Time (%)	Δ PSNR (dB)	Δ Bitrate (%)	Δ Time (%)
QCIF	Salesman	-0.02	4.08	-48.32	-0.03	4.15	-56.61
	Foreman	-0.04	3.58	-49.00	-0.04	3.80	-57.43
	Carphone	-0.07	4.12	-47.69	-0.06	4.33	-55.65
CIF	Foreman	-0.03	3.91	-51.08	-0.02	4.29	-62.54
	Mother&Daughter	-0.04	3.77	-55.02	-0.05	3.84	-65.14
	Football	-0.03	3.06	-53.14	-0.03	3.46	-62.67
	Coastguard	-0.06	3.33	-50.68	-0.08	3.78	-60.25
	News	-0.05	5.11	-52.43	-0.06	5.42	-63.20

where X is the 4x4 matrix of pixel values and Y is the 4x4 matrix of resulting IT coefficients which are to be used to calculate the edge direction in Eq. (3). For each 4x4 block, the intra prediction mode in Fig. 1 which is the closest to the computed edge direction angle θ will be chosen as the candidate mode for R-D cost evaluation. In addition, the two neighboring modes as well as the DC mode will also be enabled in our scheme. Similarly, for I8MB mode to be examined for chroma components, the mode mostly close to the edge direction obtained from Eq. (3) will be enabled together with the DC mode. For luma I16MB, all modes are enabled in our proposed scheme in order to maintain a better tradeoff between the complexity and efficiency. In the following section, the efficiency of our proposed fast intra mode decision scheme will be demonstrated.

5. EXPERIMENTAL RESULTS

The proposed algorithm has been implemented to the H.264/AVC reference software JM12.2 [10] for the re-encoding stage of the transcoder. In addition, Pan's method [7] was also implemented and both fast intra mode decision algorithms were compared with the original full search method. The experiments were conducted using the first 100 frames of three QCIF (176×144) test video sequences and five CIF (352×288) sequences of various spatial characteristics. In our experiments, all frames were coded as I-frames using quantization parameters QP=24, 28, 32 and 36. For all tests, the RDO was enabled in mode decision and CABAC was adopted for entropy coding.

To evaluate the efficiency of different intra mode decision algorithms, three metrics have been selected *i.e.*, Δ PSNR, Δ Bitrate and Δ Time. Table 1 tabulates the detailed experimental results for comparison. Note that, for these results, the positive values represent increments while the negative ones represent decrements. From the table we find that both fast intra mode decision algorithms can achieve accurate mode prediction as compared to the full search method. It is observed that by using our proposed method, the average PSNR loss is less than 0.1 dB and the average bitrate increment is within 5%. These results are similar to Pan's method. For the computational complexity savings, our proposed algorithm can reduce the encoding time by up to 65% and on average 60% time saving can be achieved for all test sequences. As compared to Pan's method, more computational reductions are obtained in our algorithm. This is because our intra block type selection scheme can effectively reject those improbable modes in an earlier stage. Fig. 5 illustrates the R-D performances of different schemes for *Foreman* and *Salesman* respectively. Obviously, the difference in R-D performance between our proposed algorithm and JM12.2 is negligible.

6. CONCLUSION

In this paper, we have proposed a fast mode decision scheme for H.263 to H.264/AVC intraframe transcoding. In order to perform fast intra block type decision before exhaustively checking both I16MB mode and I4MB mode, we adopt the number of non-zero AC coefficients from H.263 input as a criterion for early termination. Furthermore, efficient transform domain edge direction prediction method has also been employed to accelerate the intra mode prediction. Experimental results show that the proposed scheme achieves about 60% computation reduction for the transcoder with negligible visual quality and bit rate degradation. Thus it is very desirable for real-time H.263 to H.264/AVC video transcoding applications.

ACKNOWLEDGEMENT

This work was supported by the Centre for Multimedia Signal Processing, Department of Electronic and Information Engineering, Hong Kong Polytechnic University and the Research Grant Council of the Hong Kong SAR Government (PolyU 5268/06E).

REFERENCES

[1] T. Wiegand, G.J. Sullivan, G. Bjontegaard and A. Luthra, "Overview of the H.264/AVC video coding standard," *IEEE Trans. Circuits Syst. Video Technol*, vol. 13, no. 7, pp. 560-576, Jul. 2003.

[2] ITU-T Recommendation H.263: *Video coding for low bit rate communication*, ITU-T SG11, Jan. 1998.

[3] A. Vetro, C. Christopoulos and H. Sun, "Video transcoding architectures and techniques: an overview," *IEEE Signal Processing Mag*., vol. 20, no. 2, pp. 18-29, Mar. 2003.

[4] K.-T. Fung and W.-C. Siu, "Low complexity H.263 to H.264 video transcoding using motion vector decomposition," in *Proc. Int Symp. Circuits Syst.*, ISCAS'05, vol. 2, pp. 908-911, May 2005.

[5] J. Xin, C.-W. Lin and M.-T. Sun, "Digital video transcoding," in *Proc. IEEE*, vol. 93, no. 1, pp. 84-97, Jan. 2005.

[6] G. Fernandez-Escribano, J. Bialkowski, H. Kalva *et al*, "H.263 to H.264 transcoding using data mining," in *Proc. Int. Conf. Image Process.*, ICIP'07, vol. 4, pp. 81-84, Sept. 2007.

[7] F. Pan, X. Lin, S. Rahardja, K.P. Lim, Z. Li, D. Wu and S. Wu, "Fast mode decision algorithm for intraprediction in H.264/AVC video coding," *IEEE Trans. Circuits Syst. Video Technol*, vol. 15, no. 7, pp. 813-822, Jul. 2005.

[8] B. Shen and I.K. Sethi, "Direct feature extraction from compressed images," in *Proc. SPIE Storage & Retrieval for Image and Video Databases* IV, vol. 2670, 1996.

[9] R. Su, G. Liu and T. Zhang, "Fast mode decision algorithm for intra prediction in H.264/AVC," in *Proc. Int. Conf. Acoust., Speech, Signal Process.*, ICASSP'06, vol. 2, pp. 921-924, May 2006.

[10] [Online] Available JM12.2 reference software : http://iphome.hhi.de/suehring/tml/download/old_jm