IMPROVED INTRA TRANSFORM SKIP MODE IN HEVC

Liang Zhao[1], Jicheng An[2], Siwei Ma[3], Debin Zhao[1], Xiaopeng Fan[1]

[1]Department of Computer Science and Technology, Harbin Institute of Technology, Harbin, China
[2]Mediatek, Beijing, China
[3] Institute of Digital Media, Peking University, Beijing, China
{liang.zhao, dbzhao, fxp}@hit.edu.cn; jicheng.an@mediatek.com; swma@pku.edu.cn

ABSTRACT

Recent years we have witnessed a striking rise of the compound video, which is generated by computers. The High Efficiency Video Coding standard (HEVC) adopts the intra transform skip mode to improve the coding efficiency of compound video. In this paper, we improve the intra transform skip mode from two aspects. Firstly, we study the statistical character of the intra prediction residual and propose a separate coefficient scan order for the intra transform skip mode. Secondly, we analyze the probability of the intra transform skip mode for different coding unit (CU) sizes and prediction unit (PU) sizes, and then propose one limitation to reduce the applied range of intra transform skip mode. The experimental results show that the first method can achieve 1.6%, 1.1%, and 0.7% luma coding gains on average in terms of BD-rate for class F under HE10-AI, HE10-RA, and HE10-LB conditions respectively; the second method can bring 9% encoder run time reduction with almost the same BD-rate changes. The proposed methods have been partially adopted into HEVC standard.

Index Terms— HEVC, compound video, intra transform skip mode

1. INTRODUCTION

Recent years we have witnessed a striking rise in the popularity of the compound images and video such as Web pages, PDF files, slides, online games, and captured screens, which are generated by computers. For natural images and video, the existing image and video coding standards (e.g., JPEG2000 [1] and H.264/AVC [2]) have shown good coding performance. However, they are not good at compressing the compound images. Therefore, how to efficiently compress compound images has become a prevalent and critical problem.

Attempts in this field can be grouped into two categories: layer-based and block-based approaches. Most of the layer-based coding algorithms use the standard three-layer mixed raster content (MRC) representation, where one compound image is decomposed into a foreground layer, a background layer and a binary mask plane [3][4]. Block based approaches classify images into several types of blocks and use different coding methods for different types of blocks [5] [6].

Since block-based intra transform skip mode makes minor changes to the conventional hybrid video coding framework and obtains significant coding gain for compound video, HEVC adopts it to process compound video [7]. Except for adding one flag to indicate whether a 4x4 intra transform unit (TU) uses transform skip mode or not, intra transform skip mode makes no change to the prediction, de-quantization and scaling, in-loop filters and entropy coding. When transform skip mode is selected, transform is skipped from the coding structure. For compound video, intra transform skip mode is more efficient because experiments demonstrate that non-transform coding is much better than transform coding [6]. To make a tradeoff between the coding complexity and performance, intra transform skip mode is only applied to 4x4 TUs.

In this paper, we improve the intra transform skip mode from two aspects. Firstly, we study the statistical character of the intra prediction residual and propose a separate coefficient scan order for the intra transform skip mode. This method has been proposed in our proposal JCTVC-J0053 [11] and the concept of our proposal is similar to that of the proposal JCTVC-J0093 [13] in the same meeting. Secondly, we analyze the probability of the intra transform skip mode for different coding unit (CU) sizes and prediction unit (PU) sizes, and then propose one limitation to reduce the applied range of intra transform skip mode.

The rest of this paper is organized as follows. Section II presents an overview of mode dependent coefficient scan in HEVC. Section III gives a detailed description of the proposed improved intra transform skip mode. Experimental results are provided in Section IV. Section V concludes this paper.

2. MODE DEPENDENT COEFFICIENT SCAN

HEVC provides 35 intra prediction modes including 33 directional modes and 2 non-directional modes as illustrated in Fig. 1. For 4x4 TUs, there are three kinds of quantized coefficient scan orders: diagonal scan, horizontal scan and vertical scan as depicted in Fig. 2. Each coefficient scan order is assigned to the transform coefficients depending on the intra prediction modes, which is called mode dependent coefficient scan. Table I shows the mapping between intra prediction modes and coefficient scan orders.

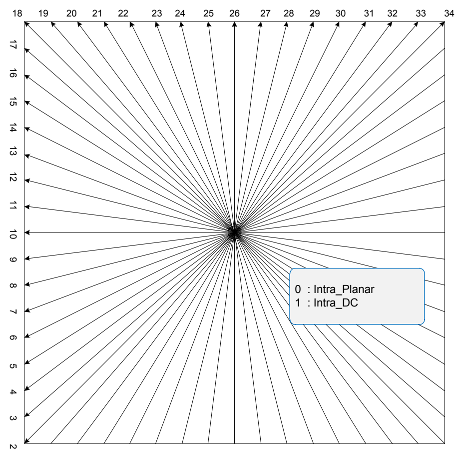

Fig. 1 Intra prediction modes in HEVC [8]

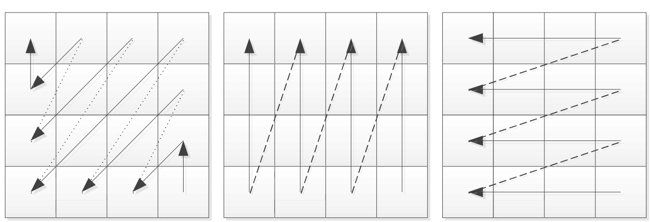

Fig. 2 Different scan orders for 4x4 intra TUs

Table I Mapping between intra prediction modes and scan orders

IntraPredMode	ScanOrder
6-14	Vertical
22-30	Horizontal
Others	Diagonal

2. IMPROVED INTRA TRANSFORM SKIP MODE

To improve the intra transform skip mode in HEVC, two methods are developed. In the first method, we design a separate coefficient scan order for the intra transform skip mode; in the second method, we propose to limit the applied range of the intra transform skip mode.

2.1 Separate Scan order for Intra Transform Skip Mode

Transform coding rearranges the energy distribution, which assembles most energy of the whole block to the top-left few coefficients [9]. It means that the amplitudes of the top-left coefficients are mostly larger than the bottom-right coefficients. However, when transform is skipped the energy distribution of quantized residual is different. It is reasonable to consider a different scan order to achieve better entropy coding.

For the intra prediction modes which are horizontal like prediction and use left side reference pixels, as illustrated in the first region of Fig. 3, the magnitudes of the prediction residues are usually larger near the right boundary and decrease from right to left as the prediction distance is shortened. That energy distribution is inversed compared to the case of transform-on, where the energy is supposed to be compacted to the left lines. Therefore, an inverse vertical scan as shown in the middle part of Fig. 4 is better for the intra transform skip mode which originally use vertical scan.

Similarly, for the intra prediction modes which are vertical like prediction and use above side reference pixels, as illustrated in the second region of Fig. 3, the magnitudes of the prediction residues are usually larger near the bottom boundary and decrease from bottom to top as the prediction distance is shortened. That energy distribution is inversed compared to the case of transform-on, where the energy is supposed to be compacted to the top rows. Therefore, an inverse horizontal scan as shown in the right part of Fig. 4 is better for the intra transform skip mode which originally uses horizontal scan.

Finally, for the remaining intra prediction modes, as illustrated in the third region of Fig. 3, the magnitudes of the prediction residues are usually larger at the bottom-right corner and decrease from bottom-right corner to top-left corner as the prediction distance is shortened. That energy distribution is inversed compared to the case of transform-on, where the energy is supposed to be compacted to the top-left corner. Therefore, we propose to use an inverse diagonal scan as shown in the left part of Fig. 4 for the intra transform skip mode which originally uses diagonal scan.

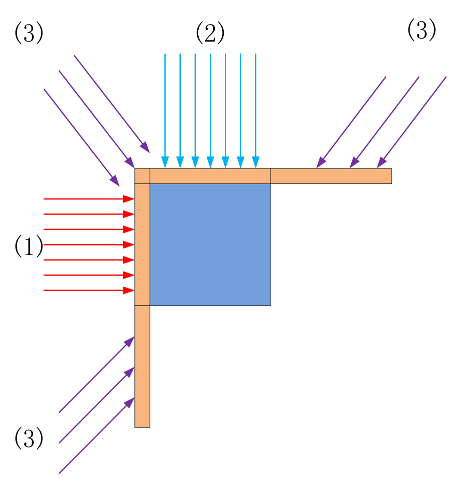

Fig. 3 Different regions for intra prediction mode

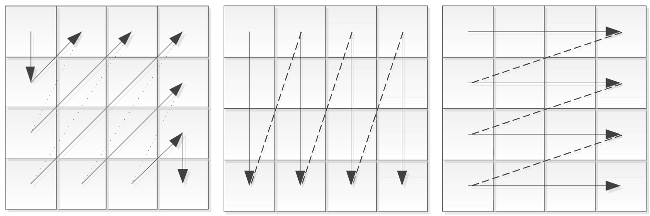

Fig. 4 Separate scan orders for intra transform skip mode

2.2 Limited Range for Intra Transform Skip Mode

In current HM, for a 4x4 intra luma or chroma TU, regardless of the size of CU and the partition mode, a flag will be transmitted to the decoder to indicate whether the transform skip mode is used or not. Therefore, for luma TU, there are three cases where the transform skip flag is transmitted as illustrated in Table II. For chroma TU, besides the above three cases, there is one additional case as illustrated in Table III.

Table II Three cases both for luma and chroma

	Luma CU	PU	TU
Case 1	16x16	2Nx2N	4x4
Case 2	8x8	2Nx2N	4x4
Case 3	8x8	NxN	4x4

Table III The fourth case for chroma

	Luma CU	PU	TU
Case 4	32x32	2Nx2N	4x4

Table IV The number of transform skip mode in different cases for luma

	Case 1	Case 2	Case 3
BasketballDrillText	1586	4078	20741
ChinaSpeed	3528	14439	161267
SlideEditing	5434	20573	305036
SlideShow	517	2232	21022

Table V The number of transform skip mode in different cases for chroma

	Case 1	Case 2	Case 3	Case 4
BasketballDrillText	181	1346	6671	0
ChinaSpeed	910	3886	36268	0
SlideEditing	1709	8526	132938	0
SlideShow	83	324	2364	0

To analyze the effectiveness of transform skip mode in different cases, we use four compound sequences provided by [10] to perform experiments. They are *BasketballDrillText*, *ChinaSpeed*, *SlideEditing* and *SlideShow* respectively. The platform is HM7.0 and the configure file is all intra configuration, where TS mode is set on.

Table IV and Table V denotes the number of case that transform skip mode is selected through the RDO process. It can be clearly seen that most of the cases appear in the third case both for luma and chroma. From our observation, we also found that most gain of the transform skip mode is obtained from the third case. Based on this observation, it is appropriate to signal the transform skip flag of a luma and chroma TU only when the third case is satisfied, which requires normative changes for the current specification.

Clearly, this limitation can also be implemented as an encoder-only method. In this way, we set the transform skip flag of the 4x4 intra luma or chroma TU to 0 when the first, second, or fourth case is satisfied. Different from the normative modification, we have to transmit this flag to the decoder although it is always 0. Both methods can speed up the encoder because of the first, second, and fourth cases are not searched.

3. EXPERIMENTAL RESULTS

To verify the performance of the improved intra transform skip mode, it was implemented on top of HM7.0. The simulation has been done according to the HE10 common test conditions [10]. For the first scheme, HE10-AI, HE10-RA, and HE10-LB conditions are conducted. For the second scheme, only HE10-AI condition is employed. Here the reference test model HM7.0 of HEVC is treated as the anchor. The “Time” in the following tables depicts both the encoder time and decoder time with proposed method compared to the anchor with default settings.

In Table VI ~ Table X, the class A, B, C, D, E and F indicate the different resolutions for the picture. The resolutions for class A, B, C, D and E are 2560x1600, 1920x1080, 832x480, 416x240, and 1280x720 respectively. Class F is the compound video sequences generated by computers. There are four sequences in Class F and the resolutions are not the same as each other. Experimental results of the first method are shown in Table VI ~ Table VIII. It can been seen that the proposed coefficient scan can provide 1.6%, 1.1%, and 0.7% average luma coding gains for class F for HE10-AI, HE10-RA, and HE10-LB, respectively. Because the transform skip mode is designed for compound video, the improvement on other sequences is marginal. It is also illustrated from the Table IX and Table X that the encoder complexity reduction of the normative and non-normative methods are both about 9% in the second method. However, without transmitting the unnecessary transform skip flag for the first, second, and fourth cases, the coding efficiency of the normative method is a little better than the non-normative method.

Table VI Results of proposed method of the first method

	All Intra HE10		
	Y	U	V
Class A	0.0%	-0.3%	-0.2%
Class B	-0.1%	-0.3%	-0.4%
Class C	-0.4%	-0.9%	-1.1%
Class D	-0.5%	-1.1%	-1.4%
Class E	-0.2%	-0.4%	-0.4%
Overall	-0.2%	-0.6%	-0.7%
Class F	-1.6%	-2.1%	-2.2%
Enc Time[%]		100%	
Dec Time[%]		99%	

Table VII Results of proposed method of the first method

	Random Access HE10		
	Y	U	V
Class A	0.0%	-0.2%	0.0%
Class B	0.0%	-0.4%	-0.5%
Class C	-0.2%	-0.7%	-1.0%
Class D	-0.2%	-0.6%	-1.3%
Class E			
Overall	-0.1%	-0.5%	-0.7%
Class F	-1.1%	-1.4%	-1.8%
Enc Time[%]		100%	
Dec Time[%]		99%	

Table VIII Results of proposed method of the first method

	Low delay B HE10		
	Y	U	V
Class A			
Class B	0.0%	0.1%	-0.1%
Class C	0.0%	-0.2%	-0.2%
Class D	0.0%	0.1%	-0.6%
Class E	-0.1%	-0.7%	-0.7%
Overall	0.0%	-0.1%	-0.4%
Class F	-0.7%	-0.7%	-0.3%
Enc Time[%]		100%	
Dec Time[%]		100%	

Table IX Results of proposed method of the second method

	All Intra HE10		
	Y	U	V
Class A	-0.01%	-0.02%	-0.01%
Class B	-0.02%	0.02%	0.02%
Class C	0.00%	-0.02%	0.00%
Class D	-0.01%	-0.01%	-0.01%
Class E	-0.02%	-0.04%	0.01%
Overall	-0.01%	-0.01%	0.01%
Class F	0.14%	-0.15%	0.15%
Enc Time[%]		91%	
Dec Time[%]		99%	

Table X Results of proposed method of the second method

	All Intra HE10		
	Y	U	V
Class A	0.00%	0.01%	0.07%
Class B	0.00%	0.07%	0.08%
Class C	0.03%	0.06%	0.09%
Class D	0.01%	0.06%	0.09%
Class E	0.01%	0.01%	0.08%
Overall	0.01%	0.04%	0.08%
Class F	0.19%	0.10%	0.36%
Enc Time[%]		91%	
Dec Time[%]		99%	

4. CONCLUSION

In this paper, we improve the transform skip mode from two aspects. Firstly, the san order of the intra transform skip mode is improved by using the separate scan order which is designed depending on the intra residual property. Secondly, the unessential cases used for intra transform skip mode is removed to relieve the encoder burden. The proposed methods have been proposed and partially adopted to HEVC standard.

ACKNOWLEDGEMENT

This work was supported in part by the National Science Foundation of China(NSFC) under grants 61272386 and 61100095, and the Program for New Century Excellent Talents in University (NCET) of China (NCET-11-0797), and the Fundamental Research Funds for the Central Universities (Grant No.HIT.BRETIII.201221)

REFERENCES

[1] D. Taubman and M. Marcellin, JPEG2000: Image Compression Fundamentals, Standards, and Practice, *Kluwer Academic Publishers*, 2001.

[2] “Draft ITU-T Recommendation and Final Draft International Standard of Joint Video Specification (ITU-T Rec. H.264/ISO/IEC 14 496-10 AVC”, JVT of ISO/IEC MPEG and ITU-T, JVTG-050, 2003.

[3] “Mixed Raster Content (MRC)”, ITU-T Recommendation T.44, Study Group-8 Contribution, 1998.

[4] P. Haffner, L. Bottou, P.G. Howard, P. Simard, Y. Bengio, Y. LeCun, “ High Quality Document Image Compression with DjVu”, *Journal of Electronic Imaging*, pp. 410-425, July 1998

[5] W. Ding, D. Liu, Y. He, F. Wu, “Block-based Fast Compression for Compound Images”, *Proc. Int. Conf. Multimedia and Expo (ICME)*, pp. 809-812, 2006.

[6] C. Lan, G. Shi, F. Wu, “Compress Compound Images in H.264/MPEG-4 AVC by Exploiting Spatial Correlation,” *IEEE Trans. Image Processing*, vol. 19, no 4, pp 946-957, 2010.

[7] C. Lan, J. Xu, G. J. Sullivan, F. Wu, “Intra Transform Skipping,” Document of Joint Collaborative Team on Video Coding, JCTVC-I0408, April, 2012.

[8] B. Bross, W. J. Han, G. J. Sullivan, J.-R. Ohm, T. Wiegand, “High Efficiency Video Coding (HEVC) Text Specification Draft 7,” Document of Joint Collaborative Team on Video Coding, JCTVC-I1003, May, 2012.

[9] P. A. Wintz, “Transform Picture Coding,” *Proc. IEEE*, vol.60, pp. 809-820, July 1972.

[10] F. Bossen, “Common Test Conditions and Software Reference Configurations”, Document of Joint Collaborative Team on Video Coding, JCTVC-I1100, May, 2012.

[11] J. An, L. Zhao, Y. Huang, S. Lei, "Residue Scan for Intra Transform Skip mode," Document of Joint Collaborative Team on Video Coding, JCTVC-J0053, July, 2012.

[12] L. Zhao, J. An, Y. Huang, S. Lei, "Simplification for Intra Transform Skip Mode," Document of Joint Collaborative Team on Video Coding, JCTVC-J0389, July, 2012.

[13] D. He, J. Wang, G. Martin-Cocher, "Rotation of Residual Block for Transform Skipping," Document of Joint Collaborative Team on Video Coding, JCTVC-J0093, July, 2012.