ENHANCED INTRA PREDICTION FOR INTER PICTURES

Kui Fan[1], *Ronggang Wang*[1], *Ge Li*[1], *Wen Gao*[1], *Yufan Deng*[2], *Shensian SYU*[2], *Ming-Jong Jou*[2]

[1]School of Electronic and Computer Engineering, Shenzhen Graduate School, Peking University
[2]Shenzhen China Star Optoelectronics Technology Co., Ltd., Guangdong, China
kuifan@pku.edu.cn, {rgwang, gli}@pkusz.edu.cn, wgao@pku.edu.cn, {dengyufan, Valexsyu, mingjongjou}@tcl.com

ABSTRACT

This paper presents a novel intra prediction method for inter pictures (i.e. P pictures and B pictures), denominated enhanced intra prediction (EIP). The traditional intra prediction only uses the reconstructed pixels to the left and above to derive intra prediction blocks. While the proposed method combines the left-above and right-below pixels to strengthen the prediction efficiency of the intra blocks in inter pictures. For accessing the pixels below and to the right of intra coding units (CUs), the encoding and decoding structures are adjusted to guarantee that all the inter CUs are reconstructed before all the intra CUs. With more available reference samples around the intra CUs, EIP achieves better prediction results. The proposed method is implemented on top of the H.265/HEVC reference software (HM-16.12), and the experimental results show that approximate 0.5% BD-rate reduction is achieved under Random Access (RA) and Low Delay P (LP) configurations.

Index Terms— Intra prediction, HEVC, video coding

1. INTRODUCTION

The state-of-the-art High Efficiency Video Coding (HEVC) standard is finalized in 2013, and is capable of achieving approximately 50% bit rate saving relative to the previous H.264/AVC with the same video quality [1][2]. Although HEVC supports a total of 35 intra prediction modes, including 33 angular prediction modes, DC, and Planar, the efficiency of the intra coding of HEVC is only promoted by 22.3% on average according to the test results in [3]. The improvement of intra coding in HEVC is considerably inferior to that of inter coding in HEVC. The further improvement of intra coding is desirable.

In the modern image and video coding standards, intra prediction and inter prediction are essential to remove the redundant information. Inter prediction makes full use of the correlation between the current frame and reference frames to significantly eliminate the temporal redundancy. While intra prediction utilizes the neighboring reconstructed pixels to dwindle the spatial redundancy. Since the temporal correlation between different frames are much stronger than the spatial correlation between adjacent pixels, so the intra prediction is inefficient relative to inter prediction. In addition, the spatial correlation declines as the distance between the reference sample and the intra predicted pixel increases. Therefore, large intra coding units (CUs) are only used in homogenous areas, while small intra CUs are inclined to be selected in areas with complex textures for finer prediction.

Considering the disadvantages of the current intra prediction, some methods are proposed to improve the intra coding performance. In [4], the accuracy of intra prediction is promoted by shortening the distances between the predicted pixels and the reference pixels. A new intra prediction method based on linear prediction and sparse representation concepts is presented in [5]. In [6], right and/or bottom reconstructed pixels can be referenced for intra prediction by flexibly determining the coding orders of CUs.

Besides, the Joint Video Exploration Team (JVET) of MPEG and ITU-T VCEG is established to explore compression technology designs for future video coding standard. Some potential intra methods have been adopted into the Joint Exploration Test Model (JEM) [7] developed by JVET. The directional intra modes are extended from 33 in HEVC to 65 for finer directional matching in natural images. Cross-component prediction (CCP), which is supported in the Range Extensions (RExt) of the HEVC standard [8][9], has also been applied to exploit the correlation between the chroma samples and the corresponding reconstructed luma samples in JEM. Adaptive reference sample smoothing (ARSS) is a new reference sample filtering mechanism involving two optional low pass filters. In addition to these methods, multiple reference lines or tiers are introduced in [10] and [11] to provide more reference choices for intra prediction, and are currently under the exploration stage.

Abovementioned intra tools improve the intra prediction efficiency by either involving more prediction reference information or introducing more prediction options to be evaluated. However, these methods are only dedicated to improve the coding efficiency of all intra (AI)

978-1-5090-6067-2/17/$31.00 ©2017 IEEE

constraint, and are not capable of significantly improving the coding performance of P pictures or B pictures. In this paper, the proposed enhanced intra prediction (EIP) method for HEVC aims to upgrade the coding efficiency of intra blocks in inter pictures by taking the features of intra CUs in inter pictures into account. The regular intra prediction only utilizes the reference samples to the left of and above the current CU to derive prediction blocks. In EIP, in addition to the regular intra reference samples, we consider the right-below reconstructed pixels as the additional reference samples. With more reference information, the proposed method provides more possibilities to improve the efficiency of intra prediction in inter pictures.

This paper is organized as follows. Section 2 describes the proposed EIP method. Experimental results are presented in Section 3, while Section 4 concludes the paper.

2. ENHANCED INTRA PREDICTION

As mentioned above, more reference choices for intra prediction can effectively improve the performance of intra coding. We attempt to exploit more available reference samples for intra CUs in inter pictures.

Let's consider an example of distribution with both intra and inter CUs in HEVC, as show in Fig. 1. For clarity, the CUs above and to the left of the current CU are defined as left-above CUs, and the CUs below and to the right of the current CU are named right-below CUs. The pixels immediately above and to the left of the current CU are expressed as left-above pixels, and the pixels immediately below and to the right of the current CU are expressed as right-below pixels. In the intra prediction of HEVC, the reference samples of intra CU X in Fig. 1 are only derived from left-above reconstructed pixels. Right-below reconstructed pixels are unavailable. Because in HEVC, coding tree units (CTUs) are encoded and decoded in the raster scan order, and CUs inside CTU must conform to the Z-scan order, the right-below CUs cannot be reconstructed before X. Intuitively, if right-below pixels can be referenced, the intra prediction efficiency would be improved. Inspired by this idea, we try to exploit the right-below reconstructed pixels as additional reference samples.

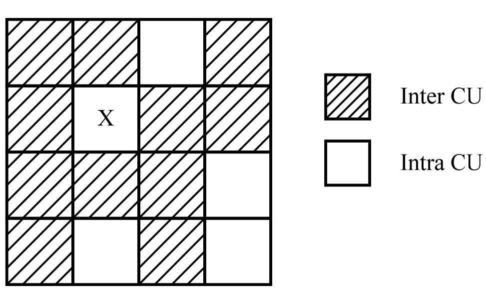


Fig. 1. An example of distribution of intra and inter CUs.

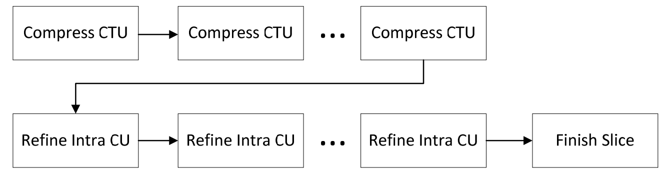


Fig. 2. New introduced encoding order of intra and inter CUs.

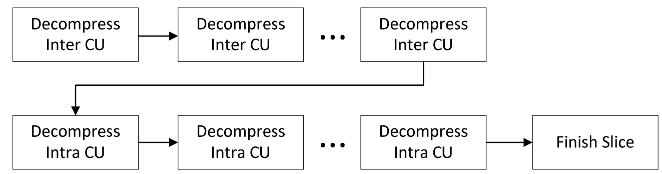


Fig. 3. New introduced decoding order of intra and inter CUs.

Only if the right-below CUs have been reconstructed, the current intra CU can refer to the right-below pixels for intra prediction. So the raster scan and Z-scan order must be broken up. In addition, intra prediction strongly relies on the previously reconstructed pixels. Therefore, the encoding or decoding of right-below intra CUs cannot precede the current CU. However, the orders of inter CUs in one frame are flexible, because inter prediction only copies blocks from previously coded frames. So if the right-below CUs are inter coded, we can access their reconstructed samples by adjusting the scan orders and performing intra encoding or decoding after all inter CUs.

First, we need refine intra CUs after regular CTU compression in encoder, and break up the scan orders in decoder. Fig. 2 and Fig. 3 demonstrate the flow chats of Slice encoding and decoding after adjusting CU orders. In encoder, a 2-pass encoding process is used. As shown in Fig. 2, the first row is the regular CTU compression in HEVC, and the second row illustrates the refinement for intra CUs. In the first round of compression, all syntax elements except for intra prediction information are determined. In the second round, only intra CUs are refined and only intra prediction information is updated. In decoder, the intra and inter CUs are completely decoupled. All inter CUs are decoded before all intra CUs. Moreover, the bitstream structure should also be changed. In encoder, CU syntax elements (including split flag, prediction mode, etc.) and inter prediction syntax elements of all inter CUs in a Slice are coded to bitstream first, followed by intra prediction syntax elements. Accordingly, the decoder can correctly decode and reconstruct all inter CUs beforehand, and then decode intra CUs and fill reconstructed intra blocks in the blank positions.

After reforming the orders of intra and inter CUs in both encoder and decoder, the right-below reconstructed pixels for the current intra CU are available if the right-below CUs are inter-predicted. The right-below reconstructed pixels can be considered as additional reference samples, which can provide more information for intra prediction. Note that only if the right-below CUs are inter coded, the right-below pixels can be referenced by the current intra block.

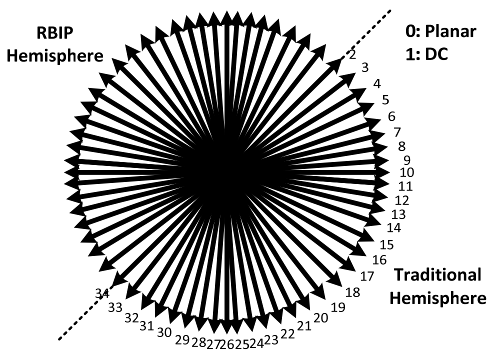
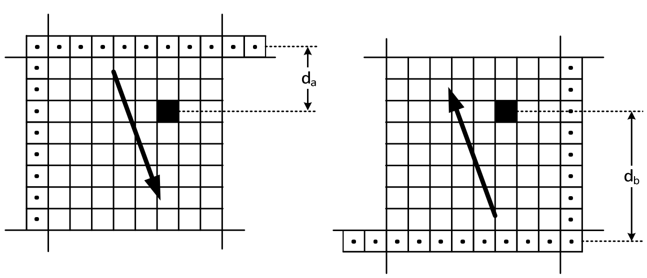


Fig. 4. Regular intra prediction directions and RBIP directions.

With the right-below reconstructed pixels available, the intra blocks can be predicted from the right-below reference samples. The intra prediction from right-below reference samples is named right-below intra prediction (RBIP). The RBIP method is in the same way as the regular intra prediction in HEVC. However, the RBIP directions and the angular modes in HEVC are centrally symmetric, as shown in Fig. 4. The traditional hemisphere includes all allowed intra prediction modes in HEVC, and RBIP hemisphere is generated by inversely extending the HEVC intra directions.



(a) regular intra prediction (b) RBIP

Fig. 5. An example of the regular intra prediction and EIP prediction.

With the additional reference samples, we propose the enhanced intra prediction (EIP) for inter pictures. In the EIP method we combine the regular intra prediction and newly introduced RBIP together to strengthen intra prediction. Fig. 5 illustrates the regular intra prediction and RBIP prediction. In Fig. 5, dots indicate reference samples, and arrows represent the intra prediction directions for the regular intra prediction and RBIP. After obtaining these two prediction blocks with the same intra mode index but opposite prediction directions, we merge these two blocks by weighted average to generate the final prediction block.

The weighting factors in combining regular intra prediction blocks and RBIP blocks are critical. We propose directionally adaptive linear weighting factors for EIP, which can be expressed by equation (1).

$$P_{comb}(x,y) = (d_b \cdot P_a(x, y) + d_a \cdot P_b(x, y) + (1 \ll (shift-1))) \gg shift \quad (1)$$

where P_a and P_b are the regular intra prediction block and EIP block, respectively. d_a and d_b are the distances from the regular prediction samples and EIP samples to the corresponding main reference borders. Shift is the normalization factor, and shift = $\log_2(d_a + d_b)$. P_{comb} is the combined prediction block by the directionally adaptive weighting. x and y denote the positions in the prediction blocks.

In the directionally adaptive linear weighting, vertically linear weighting is used for vertical intra modes, and horizontally linear weighting is used for horizontal intra modes. In HEVC, horizontal modes include 2~17 intra modes, and vertical modes includes 18~34 intra modes. For horizontal modes, the regular intra prediction blocks are mainly generated from left reference samples, and RBIP blocks are primarily predicted from right reference samples. Therefore, left and right reference samples are the main reference borders in the regular intra prediction and RBIP, respectively. Likewise, the reference samples above and below constitute to the main reference borders for vertical modes in the regular intra prediction and RBIP, respectively.

Fig. 5 illustrates two kinds of intra predictions in the proposed method, and specifies the distances in the directionally adaptive linear weighting. In Fig. 5, the intra modes represented by arrows are 22 (vertical intra mode). The pixels above and below are the main reference borders for the regular intra prediction and RBIP, respectively.

The weighting factors in equation (1) are inversely proportional to the distances to the corresponding main reference borders. In intra prediction, a shorter distance to reference samples implies higher correlation between the reference samples and predicted pixels, and consequently leads to more accurate prediction. So a shorter distance to the main reference border gives rise to a larger weighting factor. On the contrary, a longer distance between reference samples and predicted pixels results in low prediction precision. So the contribution of the predicted pixels remote from the corresponding reference samples should be attenuated in the combination of two prediction blocks. The directionally adaptive weighting factors promote the prediction efficiency of EIP, and show the best average results in our experiments compared with original linear weighting factors, bilinear weighting factors and exponential attenuation weighting factors.

In EIP, DC and planar prediction modes are also renewed. The original DC mode only averages the left-above reference samples. For the DC mode of EIP, all available reference samples around the intra block to be predicted are averaged to get the DC prediction value. In the original planar mode, the top-right reference sample is used to replace all the right reference samples, and the bottom-left reference sample is used as the bottom reference samples. In the planar mode of EIP, the right-below CUs may be available, so the genuine right-below reference pixels can be directly used. By involving right-below available reference samples, DC and planar modes become more consistent with the prediction results to be expected.

Due to the diversity of natural video content, the combination of left-above and right-below reference samples is not always superior to the original intra prediction, which only uses left-above reference samples. So when the right-below CUs are available, 1-bit flag is used to indicate whether the intra prediction block is derived from left-above pixels or the combination of left-above and right-below reference pixels. This flag is determined by RDO in encoder if the right-below CUs are available. The flag of chroma component directly inherits that of the corresponding luma component for saving the bit overhead.

3. EXPERIMENTAL RESULTS

In order to verify the performance of EIP, the proposed method is implemented on top of the HEVC reference software (HM-16.12). The experiments follow the HEVC common test condition [12], and five natural video classes (A to E) are tested. Class A to E represent 2560x1600, 1080p, WVGA, QWVGA, and 720p sequences, respectively. Bjøntegaard delta bitrate (BD-rate) with piece-wise cubic interpolation method is used to assess the objective coding performance of EIP. In addition, the encoding time and decoding time relative to the original HM-16.12 are given to evaluate the complexity of EIP.

We provide the experimental results for three common inter configurations, including Random Access (RA), Low Delay B (LB), Low Delay P (LP). Table 1, Table 2, and Table 3 show the results under RA, LB, and LP configurations, respectively. As can be seen from Table 1, the BD-rate gain of the proposed method for RA configuration is 0.49% on average. With respect to LB configuration, an average of 0.28% BD-rate reduction can be observed from Table 2. As shown in Table 3, EIP can averagely achieve 0.49% BD-rate gain for LP configuration.

The proposed method presents different performance under different configurations. By comparing the experimental results, it can be found the coding performance of EIP under RA and LP configurations is superior to that under LB configuration. Since different configurations are defined for distinct application scenarios, the coding results

Table 1. Results of EIP under RA configuration in HM-16.12

Class	Sequence	BD-rate (%)	Enc_T (%)	Dec_T (%)
Class A	Traffic	-0.43	102	121
	PeopleOnStreet	-0.86	106	106
	Nebuta	0.00	115	101
	SteamLocomotive	-0.36	111	100
Class B	Kimono	-0.29	105	119
	ParkScene	-0.28	104	113
	Cactus	-0.42	105	114
	BasketballDrive	-0.71	109	109
	BQTerrace	-0.21	107	110
Class C	BasketballDrill	-1.04	108	114
	BQMall	-0.58	111	96
	PartyScene	-0.42	111	100
	RaceHorsesC	-0.44	113	102
Class D	BasketballPass	-0.76	108	100
	BQSquare	-0.19	106	106
	BlowingBubbles	-0.53	107	97
	RaceHorses	-0.61	109	127
Class E	FourPeople	-0.59	100	109
	Johnny	-0.51	102	116
	KristenAndSara	-0.58	104	98
	Average	**-0.49**	**107**	**108**

Table 2. Results of EIP under LB configuration in HM-16.12

Class	Sequence	BD-rate (%)	Enc_T (%)	Dec_T (%)
Class A	Traffic	-0.11	105	88
	PeopleOnStreet	-0.63	109	117
	Nebuta	0.04	112	100
	SteamLocomotive	-0.08	109	101
Class B	Kimono	-0.04	104	121
	ParkScene	-0.19	100	103
	Cactus	-0.38	106	107
	BasketballDrive	-0.66	105	111
	BQTerrace	0.03	107	107
Class C	BasketballDrill	-0.88	107	105
	BQMall	-0.36	104	101
	PartyScene	-0.15	106	110
	RaceHorsesC	-0.26	109	96
Class D	BasketballPass	-0.58	106	118
	BQSquare	0.16	103	103
	BlowingBubbles	-0.28	105	104
	RaceHorses	-0.39	104	111
Class E	FourPeople	-0.40	104	104
	Johnny	-0.30	101	100
	KristenAndSara	-0.21	102	108
	Average	**-0.28**	**105**	**106**

Table 3. Results of EIP under LP configuration in HM-16.12

Class	Sequence	BD-rate (%)	Enc_T (%)	Dec_T (%)
Class A	Traffic	-0.31	105	106
	PeopleOnStreet	-0.98	109	134
	Nebuta	-0.05	122	103
	SteamLocomotive	-0.44	114	102
Class B	Kimono	-0.51	109	112
	ParkScene	-0.26	110	113
	Cactus	-0.60	108	118
	BasketballDrive	-1.05	111	108
	BQTerrace	-0.03	109	103
Class C	BasketballDrill	-1.42	111	93
	BQMall	-0.48	107	99
	PartyScene	-0.22	110	98
	RaceHorsesC	-0.51	109	102
Class D	BasketballPass	-0.82	113	96
	BQSquare	-0.01	108	119
	BlowingBubbles	-0.20	108	94
	RaceHorses	-0.40	109	100
Class E	FourPeople	-0.39	105	103
	Johnny	-0.48	102	112
	KristenAndSara	-0.67	105	108
	Average	**-0.49**	**109**	**106**

of different configurations show different characteristics. On the one hand, GOP size significantly influences the coding results. Large GOP size leads to long temporal distances between the frame to be coded and reference pictures. The efficiency of inter prediction is degraded as the distance to the reference picture increases. On the other hand, more inter prediction directions can provide better inter prediction performance. Based on above analysis, it can be inferred that inter prediction in RA and LP configurations is not very efficient to some degree. Because a large GOP size (16) is adopted for RA configuration, and LP configuration prohibits the bidirectional prediction. Consequently, EIP has more opportunities to improve the intra prediction and achieves much coding gain under these two configurations. Conversely, LB configuration supports all inter prediction directions and its GOP size (4) is very short. So the inter prediction is precise and the performance of EIP under LB configuration is not as good as that under RA and LP configurations.

In addition, it can be observed the proposed method tends to yield more outstanding performance in the sequences with intensive motions, such as PeopleOnStreet, BasketballDrive, BasketballDrill, and BasketballPass. As well known, the efficiency of inter prediction is degraded with intensive or complex motions. Intra prediction has more advantages in these areas, and better performance can be achieved by benefiting from EIP.

From Table 1 to Table 3, it can be seen that the average encoding time and decoding time only increase by less than 10%. For encoder, even though the proposed method uses a 2-pass encoding style, the encoding time only increases a little. The primary reason is that the percentage of intra CUs in inter pictures is very low. In our statistical analysis, the percentages of intra CUs are less than 10% for most of sequences in these three configurations. So the refinement is only performed for a small amount of CUs. For decoder, even though EIP is a little more complicated than the regular intra prediction of HEVC, the decoding time only ascends slightly, also because the intra CUs in inter pictures are scarce.

4. CONCLUSION

In the paper, the enhanced intra prediction (EIP) for inter pictures is presented. In addition to the left-above reference pixels, right-below reconstructed pixels can be accessed in EIP by adjusting the encoding and decoding structures. EIP combines all available reference pixels by weighted averaging the regular intra prediction block and the right-below intra prediction block. Experimental results show the proposed EIP method is able to improve the intra coding performance in all inter configurations, especially for sequences with intensive motions. Furthermore, through the proposed method, we open up a new way for intra prediction. Future exploration can be extended from this method.

5. REFERENCES

[1] G. J. Sullivan, J. R. Ohm, W. J. Han and T. Wiegand, “Overview of the High Efficiency Video Coding (HEVC) Standard,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 22, no. 12, pp. 1649-1668, Dec. 2012.

[2] J. R. Ohm, G. J. Sullivan, H. Schwarz, T. K. Tan and T. Wiegand, “Comparison of the Coding Efficiency of Video Coding Standards—Including High Efficiency Video Coding (HEVC),” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 22, no. 12, pp. 1669-1684, Dec. 2012.

[3] J. Lainema, F. Bossen, W. J. Han, J. Min and K. Ugur, “Intra Coding of the HEVC Standard,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 22, no. 12, pp. 1792-1801, Dec. 2012.

[4] X. Cao, C. Lai, Y. Wang and Y. He, "Short distance intra coding scheme for HEVC," *Pict. Coding Symp. (PCS)*, Krakow, 2012, pp. 501-504.

[5] L. F. R. Lucas, N. M. M. Rodrigues, C. L. Pagliari, E. A. B. da Silva and S. M. M. de Faria, "Sparse least-squares prediction for intra image coding," *2015 IEEE Internat. Conf. on Image Processing (ICIP)*, Quebec City, QC, 2015, pp. 1115-1119.

[6] L. Guo, D. Liu, L. Li and F. Wu, "Hierarchical quadtree-based flexible block ordering in HEVC intra coding," *2016 Visual Communications and Image Processing (VCIP)*, Chengdu, 2016, pp. 1-4.

[7] J. Chen, E. Alshina, G. J. Sullivan, J.-R. Ohm, J. Boyce, *Algorithm Description of Joint Exploration Test Model 4*, ITU-T/ISO/IEC Joint Video Exploration Team (JVET) document JVET-D1001, Oct. 2016.

[8] D. Flynn, D. Marpe, M. Naccari, T. Nguyen, C. Rosewarne, K. Sharman, J. Sole, J. Xu, “Overview of the Range Extensions for the HEVC Standard: Tools, Profiles, and Performance,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 26, no. 1, pp. 4-19, Jan. 2016.

[9] G. J. Sullivan, J. M. Boyce, Y. Chen, J. R. Ohm, C. A. Segall and A. Vetro, “Standardized Extensions of High Efficiency Video Coding (HEVC),” *IEEE Jrnl. of Sel. Topics in Signal Processing*, vol. 7, no. 6, pp. 1001-1016, Dec. 2013.

[10] Y.-J. Chang, P.-H. Lin, C.-L. Lin, J.-S. Tu, C.-C. Lin, *Arbitrary Reference Tier for Intra Directional Modes*, ITU-T/ISO/IEC Joint Video Exploration Team (JVET) document JVET-C0043, May. 2016.

[11] J. Li, B. Li, J. Xu, R. Xiong, G. J. Sullivan, *Multiple Line-Based Intra Prediction*, ITU-T/ISO/IEC Joint Video Exploration Team (JVET) document JVET-C0071, May. 2016.

[12] F. Bossen, *Common test conditions and software reference configurations*, ITU-T/ISO/IEC Joint Collaborative Team on Video Coding (JCT-VC) document JCTVC-L1100, Jan. 2013.

[13] G. Bjøntegaard, *Calculation of Average PSNR Differences Between RD-Curves*, Video Coding Experts Group (VCEG) document VCEG-M33, Apr. 2001.
[14] G. Bjøntegaard, *Improvement of BD-PSNR Model*, Video Coding Experts Group (VCEG) document VCEG-AI11, Jul. 2008.