INTRA PREDICTION WITH ADAPTIVE CU PROCESSING ORDER IN HEVC

Amin Zheng, Oscar C. Au*, Yuan Yuan*, Haitao Yang†, Jiahao Pang*, Yonggen Ling**

*The Hong Kong University of Science and Technology, †Huawei Technologies Co., Ltd
Emails: *{amzheng, eeau, yyuanad, jpang, ylingaa}@ust.hk, †haitao.yang@huawei.com

ABSTRACT

The High Efficiency Video Coding (HEVC) utilizes Z-scan order to process coding units (CUs). For intra prediction, this order cannot fully exploit the spatial correlation between adjacent CUs. After transform and quantization, the residue still contains lots of energy along edges which consumes many bits for compression. To effectively reduce the residue energy along edges, a novel intra prediction approach is proposed, where the CU processing order is changed adaptively. Two additional orders are introduced in this paper besides traditional Z-scan order. Up to 1.9% bit saving is achieved in our experiments on HEVC test model. We also propose two fast order selection algorithms and the observed gains are obtained with 27% and 2% encoding time increase compared to HEVC, respectively.

***Index Terms*—** intra prediction, adaptive CU processing order, video coding, HEVC

1. INTRODUCTION

Intra coded pictures (I-pictures) play an important role in modern video coding systems, which are necessary to avoid error propagation and enable random access. The quality of I-pictures also affects the quality of the subsequent inter coded pictures. In traditional block-based intra coding system, a block of pixels is first predicted from its neighboring pixels in previous coded blocks, then the prediction error is followed by transform coding and entropy coding. Intra prediction exploits the spatial correlation between blocks. More accurate prediction results in less bits to code the prediction error. Therefore, it is essential to investigate effective intra prediction techniques to efficiently compress the I-pictures.

In recent video coding standards, intra prediction is mostly carried out by directional extrapolation, where the predicted pixels are generated by copying the reference pixels in previous coded blocks along an angular direction. H.264/AVC [1] utilizes 8 directions and HEVC [2] supports 33 directions. The widely used directional approach is quite an efficient intra prediction technique, however, in both H.264/AVC and HEVC, the prediction accuracy is sensitive to the main edge directions as shown in Fig.1.(a). It can be observed clearly that the edges on the right part of the building get better prediction than the edges on the left part of the building.

This work is supported in part by Hong Kong Government Innovation and Technology Fund and the State Key Laboratory on Advanced Displays and Optoelectronics Technologies (Project No: ITC-PSKL12EG02), Research Grants Council (GRF Project no. 610112), and HKUST ((HKUST Project no. FSGRF12EG01)

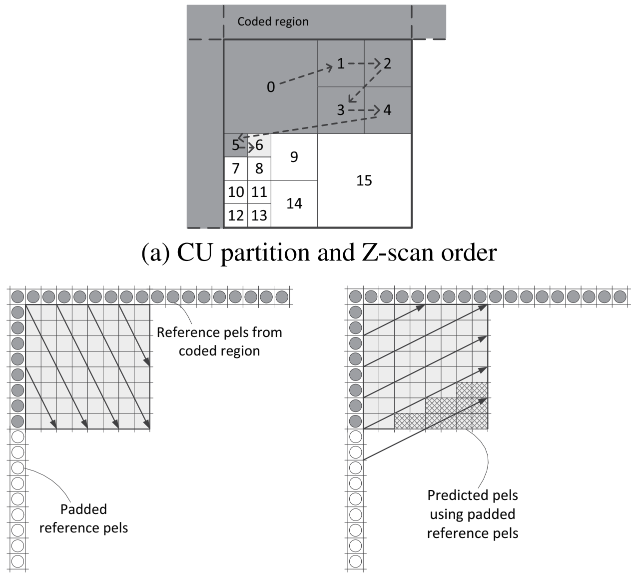
(a) HEVC (b) Proposed full RDO

Fig. 1. An example of predicted image on the first frame of *Foreman* at $QP = 32$.

Lots of new approaches have been proposed to improve the standard intra prediction methods [3][4][5]. Bi-directional intra prediction [3] combines two existing prediction modes to get a new prediction mode. Besides using reference pixels from coded blocks for prediction, [4] utilizes the pixels from current block to further reduce redundancy. By modeling image pixels with a Markov process, [5] predicts the block pixels using a weighted sum of several neighboring pixels of current block. However, these approaches haven't solved the problem that the prediction accuracy is sensitive to the edge direction.

In [6], an extra/interpolating prediction method was proposed for H.264/AVC. By changing the sub-block processing order in macroblock (MB), some sub-blocks are predicted using not only the reference pixels on the upper and/or left side but also ones on the bottom and/or right side of each sub-block. However, this method can only handle fixed size sub-block, which is not suitable for HEVC with variable size sub-block (coding unit). Different from the previous work, in this paper we propose a novel intra prediction approach for HEVC by changing the coding unit (CU) processing order of largest coding unit (LCU). Two fast order selection algorithms are employed to reduce the complexity of the encoder. The proposed scheme achieves good prediction performance for all edge directions.

The rest of the paper is organized as follows. Section 2



(a) CU partition and Z-scan order

(b) Prediction with mode 22 (c) Prediction with mode 6

Fig. 2. Examples of intra prediction in HEVC. (a) One kind of CU partition and the numbers denote the CU processing order. (b) Current CU in (a) predicted with mode 22. (c) Current CU predicted with mode 6.

reviews the intra prediction method of HEVC and analyzes its drawbacks. In Section 3, we present the adaptive CU processing order intra prediction approach and two fast order selection algorithms. Experimental results and further analysis are given in Section 4. Section 5 concludes this paper.

2. ANALYSIS OF INTRA PREDICTION IN HEVC

The main features of HEVC intra prediction which differ from previous standards are the block partitioning scheme and the intra prediction modes. In HEVC, frames are divided into largest coding units (LCUs), which can be further split in a quad-tree structure into coding units (CUs). Within one frame, all the LCUs are processed in raster scan order and within one LCU, all the CUs are processed in Z-scan order, as illustrated in Fig.2(a). The LCU in Fig.2(a) is subdivided into 16 CUs of different sizes. The flexible partitioning scheme enables the encoder to utilize the spatial correlation more effectively by choosing appropriate block size. A CU is further divided into one or four prediction units (PUs) and different PUs can have different prediction modes. HEVC supports 33 directional modes, plus DC mode and planar mode as illustrated in Fig.3. These modes are able to represent structures with various directional properties.

As CUs are processed in Z-scan order, left, below-left, above-left, above and above-right coded CUs can be used as reference to perform prediction for current CU. For example, the 8×8 PU of current processing CU denoted by 6 in Fig.2(a) is predicted as illustrated in Fig.2(b). For a $N\times N$ block, there

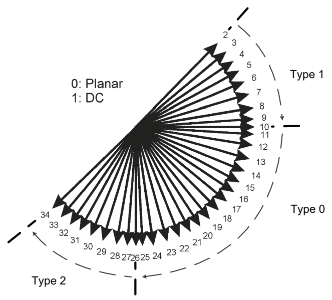


Fig. 3. 35 intra prediction modes in HEVC

are $4\times N+1$ references pixels from left and above neighboring reconstructed blocks. If neighboring blocks are unavailable, the corresponding reference pixels (padded pixels) are copied from the nearest available reference pixels. In Fig.2(b), because the below-left CU of current CU is not coded, the reference pixels shown as white circles are padded ones.

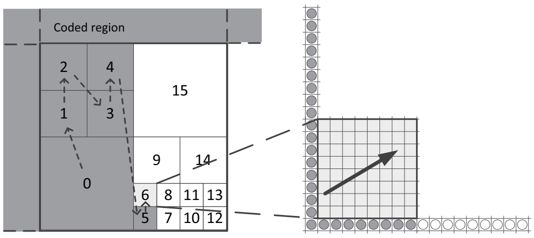
If the prediction modes are between 10 and 26 in Fig.3, no padded pixels are used for prediction as shown in Fig.2(b). Because only directly above, directly left and above-left neighboring CUs are utilized as references for prediction and these neighboring CUs are always available with Z-scan CU processing order. But when other prediction modes are employed, e.g. mode 6 in Fig.2(c), some pixels will be predicted by the padded reference pixels. Unfortunately, the padded pixels are not always same as the original reconstructed ones. Prediction by these padded pixels are less accurate than that by the reconstructed ones. That is why, in Fig.1(a), the edges on the right part of the building are predicted much better than that on the left part of the building. This problem exists in LCU with dominant down-left edges which are predicted by modes from 2 to 9 or modes from 27 to 34. Actually, it is possible to make most of the CUs to be predicted without using the padded reference pixels, by changing the CU processing order in one LCU. We propose an advanced intra prediction approach based on this idea, which will be discussed in the following section.

3. INTRA PREDICTION BY ADAPTING CU PROCESSING ORDER

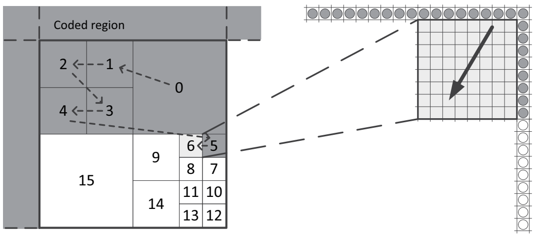
In this section, we first propose two additional CU processing orders and describe how to perform prediction in these two orders. After that, we present two fast order selection algorithms, which are based on the statistics of prediction modes and edge detection respectively.

3.1. Adaptive CU processing orders

As discussed in Section 2, down-left edges corresponding to modes from 2 to 9 and modes from 27 to 34 are not well predicted. We divide all the 33 modes into 3 categories: type 0 (modes 10 to 26), type 1 (modes 2 to 9) and type 2 (modes 27 to 34) as shown in Fig.3. If most of CUs in one LCU adopt



(a) Order 1.



(b) Order 2.

Fig. 4. Proposed two CU processing orders with corresponding prediction methods.

the prediction modes from type 1, many CUs will be predicted using inaccurate padded reference pixels in the traditional Z-scan CU processing order. To obtain more accurate reference pixels, a different CU processing order denoted by order 1 is considered especially for modes form type 1, as shown in Fig.4(a). CUs are processed first from bottom to top and then from left to right. The first processed CU is located at the bottom-left corner of current LCU. The prediction for each PU in corresponding CU is same as that in HEVC, except that the reference pixels are from the below and left neighboring blocks as illustrated in the right part of Fig.4(a). This new order makes sure that only reconstructed reference pixels are used for prediction, thus results in a more accurate prediction than traditional one. Similar to order 1, order 2 is mainly considered for those LCU whose CUs mostly utilize modes from type 2 for prediction. The processing order 2 is shown in Fig.4(b). With order 2, CUs are processed first from right to left and then from top to bottom. Reference pixels from above and right neighboring blocks are used for prediction.

Without fully compressing the current LCU, we don't know which type of modes dominans the prediction and which order should be chosen. In order to decide the best CU processing order, the Rate Distortion Optimization (RDO) method is employed. For each LCU, the cost function J is given as follows:

$$J(S) = SSD(S) + \lambda \times Rate(S) \quad S \in \{S_0, S_1, S_2\} \qquad (1)$$

where S_0, S_1 and S_2 denote order 0, order 1 and order 2, respectively. $SSD(S)$ is the sum of square difference between the original LCU and reconstructed LCU compressed in the order S, and $Rate(S)$ is the coded bits. λ is the Lagrange multiplier which is predefined. The index of the best processing order is signaled to the decoder for each LCU.

3.2. Order selection based on statistics of prediction modes

Order selection by full RDO results in that each LCU will be compressed three times, which brings considerably high computational complexity. After compressing current LCU once in the traditional Z-scan order, we can obtain many information about this LCU, such as CU partition type and intra predictions modes for each PU. According to the statistics of the prediction modes, it is not difficult to figure out what modes are preferred to predict current LCU. As mentioned in last subsection, order 1 is more suitable for modes from type 1 while order 2 is more suitable for modes from type 2. Thus, the statistics of prediction modes can facilitate the encoder to decide whether the other two orders are worthy to be tried, after order 0 is performed for current LCU.

For each mode, we calculate the total number of 4×4 blocks predicted by this mode in current LCU, after order 0 is tried. Let T_0, T_1 and T_2 be the total number of blocks predicted by modes from type 0, type 1 and type 2, respectively. If the sum of T_1 and T_2 is larger than T_0 scaled by a ratio denoted by $ratio_1$, one of the two proposed orders with larger T will be selected to compress the current LCU. The final order is chosen between order 0 and the selected one according to the RDO process. The coding process for one LCU is manifested in Algorithm 1.

Algorithm 1 Order selection based on statistic of modes

```
Compress LCU in S_0
S ← S_0
if T_1 + T_2 > ratio_1 × T_0 and T_1 + T_2 > Th_1 then
    if T_1 > T_2 then
        Compress LCU in S_1
        if J(S_1) < J(S_0) then
            S ← S_1
        end if
    else
        Compress LCU in S_2
        if J(S_2) < J(S_0) then
            S ← S_2
        end if
    end if
end if
```

In Algorithm 1, Th_1 is added to make the selection more robust in case DC or Planar mode is the dominant prediction modes. Since many LCUs are coded well in traditional CU processing order, $ratio_1$ and Th_1 ensure it is worthy to try the two different processing orders.

3.3. Order selection based on edge detection

Edge detection is used for fast intra mode decision in many works [7][8], as angular prediction modes are highly correlated with edge directions. By detecting the dominant edge

Table 1. BD-Rate results compared to HM9.2.

ΔBitrate[%]	Full RDO			Algorithm 1			Algorithm 2		
	Y	U	V	Y	U	V	Y	U	V
Class A	**-1.9**	-3.7	-3.3	**-1.4**	**-2.6**	**-2.1**	**-0.9**	**-1.8**	**-1.0**
Class B	-0.4	-1.0	-1.4	-0.2	-0.3	-0.6	0.1	0.1	-0.1
Class C	-0.6	-1.1	-1.2	-0.3	-0.4	-0.7	-0.2	-0.2	-0.4
Class D	-0.6	-1.2	-0.8	-0.2	-0.5	0.1	-0.2	-0.2	0.3
Class E	-0.8	**-4.0**	**-3.5**	-0.5	-1.2	-1.8	-0.1	-0.1	-1.0
Overall	-0.7	-1.9	-1.8	-0.4	-0.8	-0.8	-0.2	-0.3	-0.3
△enc. time	142%			27%			2%		
△dec. time	1%			1%			0		

direction in the LCU, we can estimate which modes are preferred for compressing the LCU. Then the CU processing order is selected according to the estimation.

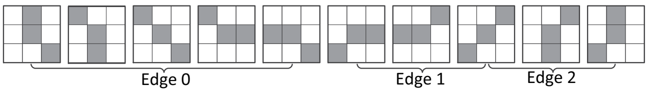


Fig. 5. Structural elements for detecting directional edge pels.

In this work, Canny edge detection [9] is utilized to extract edges in one LCU. After edge thinning, we get an one-pixel thick edge map. The structural elements in Fig.5 are used to classify each edge pixel into different categories. Edge 0, edge 1 and edge 2 are corresponding to the directional modes from type 0, type 1 and type 2, respectively. Let E_0, E_1 and E_2 be the total number of pixels belong to these three kinds of edges. Algorithm 2 manifests one LCU coding process.

Algorithm 2 Order selection based on edge detection

Edge detection for current LCU
if $E_1 + E_2 > ratio_2 \times E_0$ and $E_1 + E_2 > Th_2$ **then**
  if $E_1 > E_2$ **then**
    Compress LCU in S_1
    $S \leftarrow S_1$
  else
    Compress LCU in S_2
    $S \leftarrow S_2$
  end if
else
  Compress LCU in S_0
  $S \leftarrow S_0$
end if

where $ratio_2$ and Th_2 are similar to those in Algorithm 1. As each LCU only needs to be compressed once, the complexity of this algorithm is basically same as that of HEVC.

4. EXPERIMENTAL RESULTS

The proposed intra prediction approaches are implemented on the HEVC test model 9.2 (HM 9.2) [10]. We set $ratio_1 = ratio_2 = 1.5$, $th_1 = 40$ and $th_2 = 50$ in all of our experiments, moreover Huffman code is used to code the index of

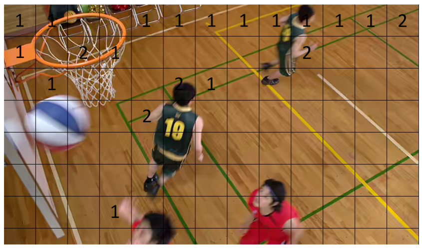


Fig. 6. CU processing orders on the first frame of *BasketballDrill* at $QP = 32$.

selected orders. The coding performance of three proposed order selection algorithms, i.e., full RDO, Algorithm 1 and Algorithm 2, was measured under common test conditions [11]. A total of 18 test sequences, which are categorized as class A to E according to resolution, were coded using intra main configuration. And four quantization parameter values are used: 22, 27, 32 and 37. The BD-Rate was measured based on [12]. The tests were performed on the following platform: Windows XP (64 bit), Intel Core i7 CPU 2.8GHz, and 16 GB RAM.

Table 1 shows the experimental results of three proposed algorithms compared to HM 9.2. The full RDO algorithm has an average of 0.7% BD-Rate(Y) reduction with 142% encoding time increase. The increase of encoding time is reduced to 27% using Algorithm 1 and 0.4% BD-Rate reduction is achieved. When the encoding time is almost same as HM9.2 using Algorithm 2, we still get an improvement of 0.2% BD-Rate saving. The BD-Rate performance of Algorithm 1 is much better than Algorithm 2, because the edge detection may be affected by noise in Algorithm 2.

There are considerable improvments for sequences containing lots of edges, such as the two sequences (*Traffic* and *PeopleOnStreet*) in Class A which have an average of 1.9%(Y), 3.7%(U) and 3.3%(V) BD-Rate reduction by the proposed full RDO algorithm. The proposed scheme achieves good prediction for all edge directions, as shown in Fig.1 that the down-left edges on the left part of the building in (b) are predicted much better than that in (a). Fig.6 shows the final selected coding orders of *BasketballDirll* by the proposed full RDO algorithm. The two proposed orders are mostly selected for those LCUs with dominant down-left edges.

5. CONCLUSION

In this paper, an improved intra prediction approach is presented. The proposed approach achieves much better prediction for edge region than HEVC by adaptively choosing CU processing order. Encouraging experimental results show that proposed approach achieves average of 0.7% bit-saving compared to HM 9.2. Two fast order selection algorithms are also proposed and the observed gains are realized with 27% and 2% encoding time increase.

6. REFERENCES

[1] *Advanced Video Coding for Generic Audiovisual Services*, ISO/IEC 14496-10, ITU-T Rec. H.264, 2005.

[2] G.J. Sullivan, J. 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.

[3] Y. Ye and M. Karczewicz, "Improved H.264 intra coding based on bi-directional intra prediction, directional transform, and adaptive coefficient scanning," in *Proc. IEEE Int. Conf. Image Process.*, Oct. 2008, pp. 2116–2119.

[4] A. Gabriellini, D. Flynn, M. Mrak and T. Davies, "Combined intra-prediction for high efficiency video coding," *IEEE Journal of Selected Topics in Signal Processing.*, vol. 5, no. 7, pp. 1282–1289, Nov. 2011.

[5] F. Kamisli, "Intra prediction based on markov process modeling of images," *IEEE Trans. Image Processing.*, vol. 22, no. 10, pp. 3916–3925, Oct. 2013.

[6] T. Shiodera, A. Tanizawa and T. Chujoh, "Block based extra/inter-polating prediction for intra coding," in *Proc. IEEE Int. Conf. Image Process.*, Sep. 2007, pp. 445–448.

[7] F. Pan, X. Lin, S. Rahardja, K.P. Lim, Z.G. Li, D.J. 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] J.F. Wang, J.C. Wang, J.T. Chen, A.C. Tsai and A. Paul, "A novel fast algorithm for intra mode decision in H.264/AVC encoders," in *Proc. IEEE Int. Symp. Circuits Syst.*, May. 2006, pp. 3498–3501.

[9] J. Canny, "A computational approach to edge detection," *IEEE Trans. Pattern Analysis and Machine Intelligence.*, vol. PAMI-8, no. 6, pp. 679–698, Nov. 1986.

[10] "HEVC Test Model," `https://hevc.hhi.fraunhofer.de/svn/svn_HEVCSoftware/tags/`.

[11] F. Bossen, *Common Test Conditions and Software Reference Configurations*, document JCTVC-H1100, JCT-VC, San Jose, CA, Feb. 2012.

[12] G. Bjontegaard, "Calculation of average PSNR differences between RD-curves," *ITU-T VCEG-M33*, pp. 1–9, Apr. 2001.