LETTER

Alternative Intra Prediction for Screen Content Coding in HEVC

Yangbin LIM[†], **Si-Woong LEE**[†], ***Nonmembers***, ***and*** **Haechul CHOI**[†a)], ***Member***

SUMMARY Screen content generally consists of text, images, and videos variously generated or captured by computers and other electronic devices. For the purpose of coding such screen content, we introduce alternative intra prediction (AIP) modes based on the emerging high efficiency video coding (HEVC) standard. With text and graphics, edges are much sharper and a large number of corners exist. These properties make it difficult to predict blocks using a one-directional intra prediction mode. The proposed method provides two-directional prediction by combining the existing vertical and horizontal prediction modes. Experiments show that our AIP modes provide an average BD-rate reduction of 2.8% relative to HEVC for general screen contents, and a 0.04% reduction for natural contents.

key words: *intra prediction, screen content coding, HEVC*

1. Introduction

Screen content typically includes characters, graphics, scrolling webpages, various horizontal and vertical lines, and a wide range of 3D (e.g., video game) output. The content is widely used in various applications, such as desktop sharing, video conferencing, and remote education. It is often required to compress such content with video coding solutions. Recently, the Joint Collaboration Team on Video Coding (JCT-VC) of MPEG and VCEG has developed a new video coding standard called High Efficiency Video Coding (HEVC) [1]. In this process, the team placed special emphasis on the problem of Screen Content Coding (SCC) (as reflected by several of the class F test sequences, including BasketballDrillText, ChinaSpeed, SlideEditing, and SlideShow) [2]. Currently, the team is encouraging new screen content coding algorithms under the HEVC Range Extension project [3].

The properties of screen content are quite different from those of natural content, such as images and video captured by digital cameras. Hence, legacy video codecs that target natural content, such as MPEG-2 and H.264/AVC, tend to degrade the visual quality of screen content in noticeable ways. For example, frames that include detailed text and graphics typically lose their sharp edges and high contrast under two-dimensional transforms. To address this, a transform skipping mode was adopted in HEVC [4], [5], though use of the mode is limited to a prediction unit (PU) size of 4×4. Because transform skipping can be toggled, a 4×4 intra-predicted block is coded with or without transform, as determined by rate-distortion optimization (RDO).

In addition to transform skipping mode, SCC technologies have received considerable attention from researchers. In [6], a residue scan was proposed for intra transform skip mode, enabling better adaptation to different energy distributions in the residual signal. The proposed scan order is the inverse of the transform scan order. This method achieved average BD-rate gains of 1.6% in luma coding for class F test sequences under All_Intra_Main conditions: the conditions using the main profile of all intra coding [2].

In [7], yet another scan method for coefficient coding was introduced. HEVC currently supports three types of scan methods: diagonal/up-right, horizontal, and vertical. If the intra prediction mode index is within the range of 6 to 14 (i.e. near horizontal intra prediction mode) a vertical scan is used. This is because zero coefficients tend to be successive in a vertical direction after transforming the residual that is derived from the near horizontal intra prediction mode. However, if the transform is skipped, zero coefficients of the near horizontal intra prediction modes tend to be successive in the horizontal direction. Hence, in [7], the scan type of transform skipped block is converted to horizontal for the near horizontal prediction modes, achieving an average BD-rate reduction of 0.9%.

In [8], a Base Color and Index Map (BCIM) mode was proposed for compressing screen content. Given the limited color of the text and graphics portions of screen contents, BCIM mode uses an indexed color map for color expression. In a similar manner, a palette-based coding method was also proposed [9]. The palette mode is a CU mode designed for the SCC where each CU consists of very limited distinct values. Of course, these modes can cause severe coding loss for natural content, which tends to exhibit a wide chromatic range. Furthermore, obtaining the base color values requires a complicated clustering process.

In [10], a CU-level intra motion compensation tool was proposed. This method is particularly effective for SCC because patterns (e.g. text characters, icons, lines etc.) can repeat in text and graphics regions within a picture. However, motion vector searching process also requires much computation time.

In this letter, we introduce Alternative Intra Prediction (AIP) modes. To preserve the sharp edges and corners of text and graphics of screen content, while maintaining the efficiency of natural content coding, AIP modes combine both vertical and horizontal prediction modes into

Manuscript received September 13, 2013.
Manuscript revised March 11, 2014.
[†]The authors are with the Department of Multimedia Engineering, Hanbat National University, Daejeon, Korea.
a) E-mail: choihc@hanbat.ac.kr
DOI: 10.1587/transinf.2013EDL8234

a two-directional prediction. A total of six AIP modes are suggested as replacements for the rarely used and non-consecutive HEVC intra angular modes.

2. Alternative Intra Prediction for Screen Contents

Figure 1 (a) shows a frame from the HEVC common test sequence called "SlideShow", along with detailed views of a particular region of text. Screen content generally includes a large number of such regions. In the detailed views, boxes indicate Transform Units (TUs) in which transform skipping mode was used when the sequence was coded using HM 12.0 [11]. Note that these TUs include high curvature points (i.e. "corners") defined as the junction of two or more straight edges. The high contrast found in such regions can lead to large prediction errors when neighboring pixels are referenced. Furthermore, a PU containing a corner is difficult to predict in only one direction, as the corner exhibits two or more edge directionalities.

Most PUs containing corners have horizontal and vertical edge directionalities. In Fig. 1 (b), if pixels in the upper three rows can be predicted through vertical prediction mode and pixels in the last row can be predicted through horizontal prediction mode, overall prediction performance would be improved. The proposed AIP modes provide precisely this kind of performance, by combining horizontal and vertical predictions.

Figure 2 illustrates the proposed AIP modes. For 4×4 PUs, six AIP$_i$ modes ($i = 1, \ldots, 6$) are defined according to how the PU is divided into two rectangle regions, one of which is predicted in the vertical direction and the other in the horizontal direction. The modes are meant to cover a range of typical positions for corners in a PU.

Each predicted sample is obtained from the reconstructed reference sample values $R_{x,-1}$ or $R_{-1,y}$. The 4×4 PU is denoted as Ω. Let $P_{x,y}$ designate the predicted sample value of Ω, where $(x, y) \in \Omega$. For AIP$_1$ to AIP$_3$, the predicted sample $P_{x,y}$ is defined as follows:

$$P_{x,y} = \begin{cases} R_{x,-1} & \text{if } 1 \leq y \leq i \\ R_{-1,y} & \text{otherwise} \end{cases} \quad (1)$$

Similarly, for AIP$_4$ to AIP$_6$, the predicted sample $P_{x,y}$ is obtained as follows:

$$P_{x,y} = \begin{cases} R_{-1,y} & \text{if } 1 \leq x \leq (i-3) \\ R_{x,-1} & \text{otherwise} \end{cases} \quad (2)$$

These AIP modes are composed of existing HEVC intra prediction modes 10 (horizontal) and 26 (vertical). Thus, the proposed method represents only a slight change to the HEVC standard, and minimal additions to the encoder and decoder.

HEVC now uses three syntax elements to represent the intra prediction mode, *prev_intra_luma_pred_flag*, *mpm_idx*, and *rem_intra_luma_pred_mode*. Adding the proposed AIP modes to the existing HEVC intra prediction modes would require signaling which intra prediction mode is currently selected. To minimize this burden, AIP modes could be substituted for the least used HEVC intra prediction modes.

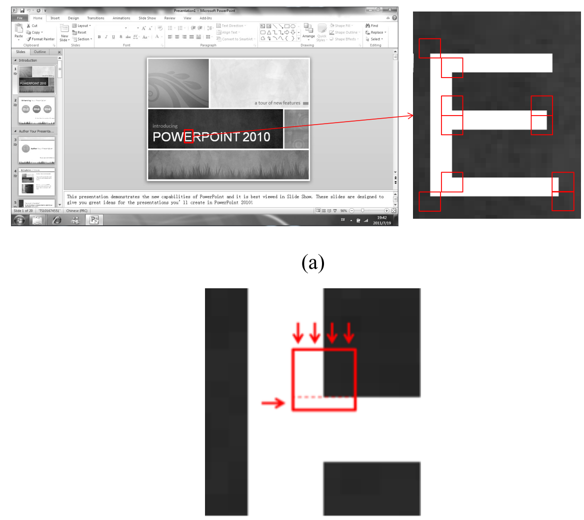

Fig. 1 (a) A frame from "SlideShow" with text detail indicating TUs coded in transform skipping mode, (b) AIP mode with vertical prediction for first three rows and horizontal prediction for last row.

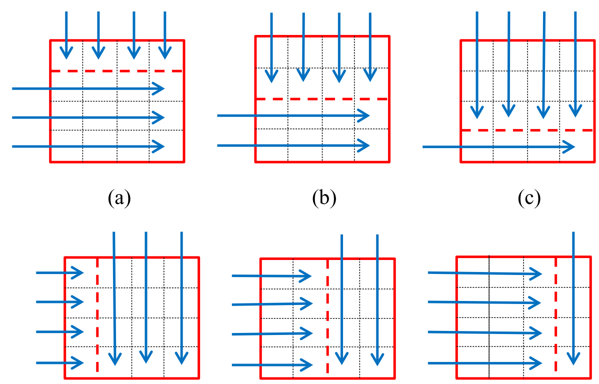

Fig. 2 Six AIP modes defined by the combination of horizontal and vertical predictions for 4×4 PU, (a) AIP$_1$, (b) AIP$_2$, (c) AIP$_3$, (d) AIP$_4$, (e) AIP$_5$, (f) AIP$_6$.

To exploit the percentages of use for the HEVC intra prediction modes, 24 test sequences from classes A to F were tested under All_Intra_Main conditions [2], using quantization parameters (QP) of 22, 27, 32, and 37. Figure 3 shows the resulting hitting counts for each HEVC intra prediction mode on 4×4 PUs. Based on these results, we suggest replacing intra prediction modes 3, 5, 13, 15, 31, and 33 (hereafter referred to as Removable Intra Prediction, or RIP, modes) with the proposed AIP modes. Since the indexes of the RIP modes are not consecutive, intra prediction modes with adjacent indexes can reduce the effects of removal. For example, indexes 2 or 4 can be used instead of index 3. As indicated in Table 1, when testing class F sequences with HEVC intra prediction modes other than the RIP modes, the

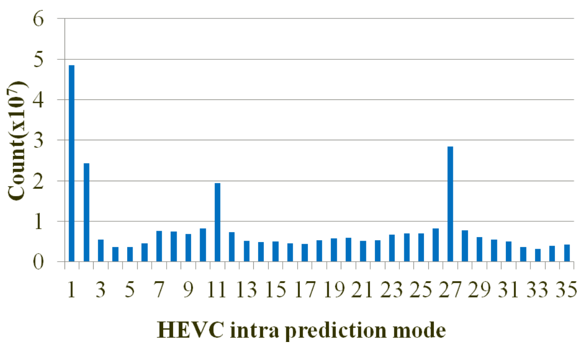

Fig. 3 Hitting counts for each HEVC intra prediction mode on 4x4 PUs when all common test sequences are coded under All_Intra_Main conditions and QP = 22, 27, 32, and 37.

Table 1 Performance of HEVC when the RIP modes are removed.

Test sequence	Resolution (pixel)	Frame rate (Hz)	Y BD-rate (%)
BasketballDrillText	832×480	50	0.2
ChinaSpeed	1042×768	30	0.1
SlideEditing	1280×720	30	0.1
SlideShow	1280×720	20	0.3
Average			0.2

average coding loss in BD-rate was only 0.2%.

For chroma intra prediction, the HEVC standard currently provides Planar, DC, horizontal, vertical, and Derived modes. In the case of Derived mode, the prediction is performed using the corresponding luma PU mode. Under the proposed method, when the chroma prediction is Derived mode and the corresponding luma mode is one of the AIP modes, chroma samples of the PU are coded using the AIP mode.

3. Experimental Results

To implement the proposed method, RIP modes were removed from HM 12.0 [11] and replaced with the proposed AIP modes. To evaluate the performance of the proposed method in comparison to HEVC, screen content sequences (class F) were tested as well as natural content sequences (classes A, B, C, D, and E). In all experiments, encoder control followed common HEVC test conditions, using the All_Intra_Main conditions [2].

All test sequences were coded for rate-distortion optimization using both the standard and modified HM 12.0. Table 2 shows the results. Note that the AIP modes were selected by an average of 3.32% versus 1.42% for RIP modes. Even when including natural contents, the AIP modes were more frequently selected than the RIP modes. Since both standard and modified HM 12.0 use the same signaling bits to indicate the prediction mode, we conclude that the reason for the higher rates of selection of AIP modes is that these modes generate fewer prediction errors.

For screen content (class F) sequences specifically, the performance of the modified HM is listed in Table 3. Note that the proposed modes show an average BD-rate reduction of 2.5% relative to standard HEVC, as well as 2.3% for chroma signals. As expected, the proposed method achieved higher coding gains for sequences with larger regions of text (SlideEditing, SlideShow, ChinaSpeed, and BasketballDrillText). The proposed method consistently outperformed HEVC over all screen content sequences. Considering that the proposed method is a normative coding tool, and not an encoding tool, the 2.5% BD-rate gain is significant.

To compare subjective quality, an original image and two reconstructed images, one coded using HEVC and other with the proposed method, are shown in Fig. 4. Note the greater clarity of edges in the image coded by the proposed method.

To test the effectiveness of AIP modes in coding natural content, BD-rate was measured for all classes of test sequences. The results, listed in Table 4, show that the AIP modes have no negative impact on natural sequences (classes A to E), and can therefore be used for both screen and natural content. Furthermore, the proposed method does not require additional computational complexity compared to HM 12.0 because AIP modes reuse prediction signals of the prediction mode 10 (horizontal) and 26 (vertical) without additional complex calculation.

Table 2 Hitting ratios for RIP modes (3, 5, 13, 15, 31, and 33) and AIP modes on 4x4 PUs.

Intra prediction mode (RIP / AIP)	HEVC / AIP (class F)	HEVC / AIP (classes A-F)
3 / AIP_1	1.44 / 3.54	1.25 / 1.78
5 / AIP_2	1.46 / 4.15	1.53 / 2.17
13 / AIP_3	1.59 / 2.88	1.63 / 1.68
15 / AIP_4	1.48 / 3.96	1.52 / 1.94
31 / AIP_5	1.25 / 3.05	1.24 / 1.55
33 / AIP_6	1.29 / 2.33	1.33 / 1.43
Average	**1.42 / 3.32**	**1.42 / 1.76**

Table 3 Performance of the AIP modes compared to current HEVC standard for screen content sequences.

Test sequence	Y BD-rate (%)	U BD-rate (%)	V BD-rate (%)
BasketballDrillText	-0.7	-0.9	-0.9
ChinaSpeed	-2.1	-1.8	-1.7
SlideEditing	-4.6	-4.4	-4.3
SlideShow	-2.8	-2.3	-2.2
Average	-2.5	-2.3	-2.3

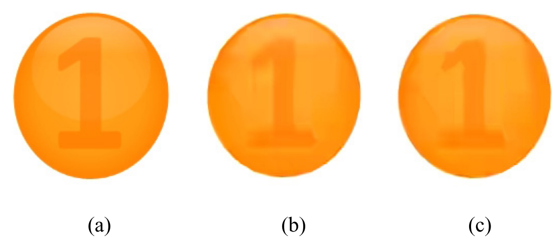

Fig. 4 A detail of the 86th frame of SlideShow under All_Intra_Main condition and QP = 37 (a) original (b) HEVC with 500016 bits (c) the proposed method with 479776 bits.

Table 4 Performance of the AIP modes in comparison to HEVC for natural (classes A to E) and screen (class F) contents.

Test sequence	Y BD-rate (%)	U BD-rate (%)	V BD-rate (%)
Class A	0.0	0.0	0.0
Class B	0.0	0.0	0.0
Class C	-0.1	-0.1	-0.1
Class D	-0.1	-0.1	-0.1
Class E	0.0	0.0	-0.1
Class F	-2.5	-2.3	-2.3
Average	-0.4	-0.4	-0.4
Enc Time [%]		100.3%	
Dec Time [%]		99.5%	

4. Conclusion

For the purpose of improving SCC efficiency and fidelity, we proposed six AIP modes that combine vertical and horizontal prediction. Specifically, we suggest that these modes replace certain, rarely used, angular modes in the current HEVC, as they provide significantly improved coding for screen content, while maintaining the quality of natural content.

Acknowledgments

This work was supported by a Basic Science Research Program through the National Research Foundation of Korea (NRF), funded by the Ministry of Education (2012R1A1A4A01014227 and 2013R1A1A1010344).

References

[1] Joint Collaborative Team on Video Coding (JCT-VC), "High Efficiency Video Coding (HEVC) text specification draft 10 (for FDIS & Consent)," JCTVC-L1003, Geneva, Jan. 2013.
[2] F. Bossen, "Common HM test conditions and software reference configurations," JCTVC-L1100, Geneva, Jan. 2013.
[3] Joint Collaborative Team on Video Coding (JCT-VC), "High Efficiency Video Coding (HEVC) range extensions text specification: Draft 4," JCTVC-N1005, Vienna, July 2013.
[4] M. Mrak and J.-Z. Xu, "Improving screen content coding in HEVC by transform skipping," Proc. European Signal Processing Conference, pp.1209–1213, Aug. 2012.
[5] C. Lan, J. Xu, G.J. Sullivan, and F. Wu, "Intra transform skipping," JCTVC-I0408, Geneva, April 2012.
[6] J. An, L. Zhao, Y.-W. Huang, and S. Lei, "Residue scan for intra transform skip mode," JCTVC-J0053, Stockholm, July 2012.
[7] T. Sugio and T. Nishi, "Modified scan method for intra transform skipping," JCTVC-J0212, Stockholm, July 2012.
[8] C. Lan, J. Xu, G.J. Sullivan, and F. Wu, "Improvements of the BCIM mode for screen content coding," JCTVC-F200, Torino, July 2011.
[9] Joint Collaborative Team on Video Coding (JCT-VC), "Non-RCE3: Modified palette mode for screen content coding," JCTVC-N0249, Vienna, July-Aug. 2013.
[10] Joint Collaborative Team on Video Coding (JCT-VC), "Non-RCE3: Intra motion compensation with 2-D MVs," JCTVC-N0256, Vienna, July-Aug. 2013.
[11] F. Bossen, D. Flynn, and K. Sühring, "High efficiency video coding test model 12 (HM 12) reference software," JCTVC-N1010, Vienna, July 2013.