Residual DPCM in HEVC Transform Skip Mode for Screen Content Coding

Chan-Hee Han, Si-Woong Lee, and Haechul Choi*

Department of Multimedia Engineering, Hanbat National Engineering / Daejeon, South Korea
{chani, swlee69, choihc}@hanbat.ac.kr

***** Corresponding Author: Haechul Choi

Received August 20, 2016; Revised September 18, 2016; Accepted September 21, 2016; Published October 30, 2016

** Short Paper*

***Abstract*:** High Efficiency Video Coding (HEVC) adopts intra transform skip mode, in which a residual block is directly quantized in the pixel domain without transforming the block into the frequency domain. Intra transform skip mode provides a significant coding gain for screen content. However, when intra-prediction errors are not transformed, the errors are often correlated along the intra-prediction direction. This paper introduces a residual differential pulse code modulation (DPCM) method for the intra-predicted and transform-skipped blocks to remove redundancy. The proposed method performs pixel-by-pixel residual prediction along the intra-prediction direction to reduce the dynamic range of intra-prediction errors. Experimental results show that the transform skip mode's Bjøntegaard delta rate (BD-rate) is improved by 12.8% for vertically intra-predicted blocks. Overall, the proposed method shows an average 1.2% reduction in BD-rate, relative to HEVC, with negligible computational complexity.

***Keywords*:** HEVC, DPCM, Screen content, Intra-prediction

1. Introduction

Screen content consists of text, images, and videos variously generated, rendered, or captured by computers, cameras, and other electronic devices. The content is widely used in various applications, such as desktop sharing, video conferencing, and remote education. It is often necessary to compress such content with video coding solutions. Recently, screen content coding (SCC) has been actively studied. The Joint Collaborative Team on Video Coding (JCT-VC) of the Moving Picture Experts Group (MPEG) and Video Coding Experts Group (VCEG) has developed a new-generation video coding standard called High Efficiency Video Coding (HEVC) [1]. Of importance in the standardization of HEVC, SCC was issued, and screen content has been included as common test sequences [2]. The Joint Video Exploration Team (JVET) on future video coding that VCEG and MPEG founded in October 2015 has also been actively studying SCC.

When coding screen content, we may face different characteristics of computer-generated content, such as text and graphics, compared to the previous naturally recorded content captured by cameras and other devices. For example, with text and graphics in screen content, edges are much sharper and the contrast is obvious [3]. In legacy video coding standards, such as MPEG-2 and H.264/AVC, the prediction errors, or so-called residues, are transformed to concentrate their energy. It is efficient in cases where pixel values gradually vary in the spatial domain. In contrast, the transform in SCC may not be efficient because the sharp and/or high-contrast regions generate a large number of high-frequency components. Thus, transform skip mode was adopted in HEVC [4, 5]. In this mode, the transform is bypassed, which improves coding efficiency for specific video content, such as computer-generated graphics. HEVC restricts the use of this mode when the block size is equal to 4×4. Transform skip mode is switchable. The 4×4 intra-predicted blocks can thereby be coded with or without the transform operation, which is determined in the sense of rate-distortion optimization (RDO).

In the intra angular modes of HEVC, aside from DC and planar modes, one decoded sample (or one interpolated sample using adjacent decoded samples) is commonly used as a predicted sample for all pixels on the line. This takes place along the intra-prediction direction associated with the selected angular mode. In RDO, a block can be coded as a particular intra angular mode, which implies that original pixels have similar luminance (or chrominance) values along the selected intra-prediction direction. Thus, the residues can also be correlated along the selected intra-prediction direction because one identical prediction sample is commonly subtracted from similar original pixel values. Moreover, in the HVEC transform skip mode, the residues are not transformed but are scaled to match the dynamic range of the transform-skipped signal to that of the transformed signal. Because screen content generally has strong edges, as described above, this scaling may dramatically amplify the magnitude of the residues. The characteristics of transform-skipped blocks can be very different from those of transformed blocks. However, HEVC applies the quantization and entropy coding processes to the transform-skipped blocks in the same way as the transformed blocks.

This paper introduces an improved transform skip mode using residual differential pulse code modulation (DPCM). In the proposed method, the residues within the intra-coded and transform-skipped block are predicted from the nearest (or interpolated/extrapolated) residue along the selected intra-prediction direction. The residual DPCM can reduce the dynamic range of the residues, by which a smaller number of bits is potentially assigned to the residues in the entropy-coding process.

2. Residual DPCM for the Intra Transform Skip Mode

As described above, the intra-predicted and transform-skipped block may have a high spatial redundancy along the selected intra-prediction direction. To reduce redundancy, for the intra-predicted and transform-skipped blocks, the proposed method performs pixel-by-pixel residual prediction by referencing the spatially nearest (or interpolated/extrapolated) residue along the intra-prediction direction.

The 4 × 4 intra-predicted and transform-skipped block is denoted as $\mathbf{\Omega}$. Let $r_{x,y}$ designate the residual values of block $\mathbf{\Omega}$, where $(x,y) \in \mathbf{\Omega}$. As in the intra-predicted sample derivation process of HEVC [6], the predicted residue $\hat{r}_{x,y}$ is obtained through linear interpolation that utilizes the two neighboring residues and the selected intra-prediction direction:

$$\hat{r}_{x,y} = \left(1 - w_\theta\right) r_{i,y-1} + w_\theta r_{i+1,y-1} \qquad (1)$$

where w_θ is the weighting between the two reference residues corresponding to the projected sub-pixel location in between $r_{i,y-1}$ and $r_{i+1,y-1}$. The reference residue index i and the weighting parameter w_θ are calculated based on selected intra-prediction direction θ as follows:

$$i = \lfloor x + tan\theta \rfloor,\ w_\theta = tan\theta \qquad (2)$$

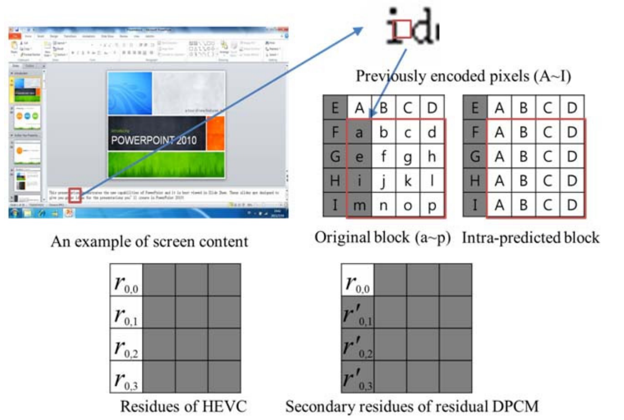

Fig. 1. An example of screen content and its residues generated by HEVC and the residual DPCM, respectively, when intra-vertical prediction mode is applied.

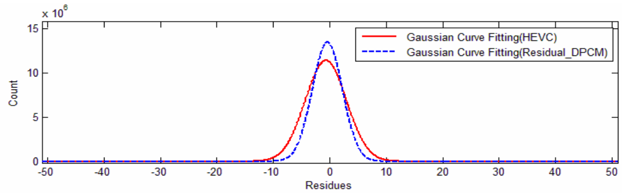

Fig. 2. Gaussian distribution curves of residues generated by HEVC and the residual DPCM, respectively.

where $\lfloor x \rfloor$ is the largest integer less than x. Describing the tangent of the intra-prediction direction in units of 1/32 samples, and having a value from -32 to +32, Eqs. (1) and (2) can be alternatively represented with the projection displacement in a way similar to that of Lainema et al. [6]. When the projection point is an integer sample (i.e., when w_θ is equal to 0 or 1), the process is even simpler, and consists of only copying an integer reference residue from the reference row.

Eqs. (1) and (2) are used for vertical-prediction directionalities (intra-prediction modes 18 to 34) when the reference row above the residue at (x,y) is used to derive the predicted residue. The residual prediction from the left reference column (intra-prediction modes 2 to 17) is identically derived by swapping the x and y coordinates in Eqs. (1) and (2).

In HEVC, residue $r_{x,y}$ is coded, whereas in the proposed method, the difference between residue $r_{x,y}$ and predicted residue $\hat{r}_{x,y}$ is coded. To clearly address the difference resulting from the residual prediction, in this paper, this difference is named *secondary residue* $r'_{x,y}$.

Fig. 1 shows a typical example of screen content and its residual pattern for intra-vertical prediction mode. In the figure, gray and white pixels represent small and high magnitudes of the residual value, respectively. For intra-vertical prediction mode, the predicted residue of $r_{x,y}$ is $r_{x,y-1}$ ($0 \le x < 4$ and $0 < y < 4$). For example, after applying the residual DPCM, the left-most column of the secondary residual block is { $r_{0,0}$, $r'_{0,1}$ (= $r_{0,0} - r_{0,1}$), $r'_{0,2}$ (= $r_{0,1} - r_{0,2}$),

Table 1. Net performance of the residual DPCM for the vertically intra-predicted and transform-skipped blocks.

Sequences	Y BD-rate (%)	U BD-rate (%)	V BD-rate (%)	Encoding time (%)	Decoding time (%)
BasketballDrillText	-8.2	-10.4	-10.3	99	96
ChinaSpeed	-12.2	-13.2	-13.3	100	96
SlideEditing	-13.1	-15.2	-14.7	100	94
SlideShow	-17.8	-18.9	-19.0	100	98
Average	**-12.8**	**-14.4**	**-14.3**	**100**	**96**

$r'_{0,3}$ (= $r_{0,2} - r_{0,3}$) }. As shown in the figure, the secondary residues resulting from the residual DPCM have only one high-magnitude (white) residue value, whereas the residues of HEVC have four high-magnitude residue values.

In practice, the residual DPCM does not always reduce the magnitude of the residues because of their randomness. However, if the residual DPCM reduces the magnitudes of the residues more frequently than it augments them, an overall coding gain can be achieved. To explore magnitude changes by the residual DPCM, the SlideShow sequence, which is one of the HEVC common test sequences [2], was verified. Fig. 2 shows the Gaussian fitting results for the residues of HEVC and the secondary residues of the residual DPCM. In the experiment, only 4 × 4 intra-vertical prediction mode is allowed, and the transform is always bypassed. As shown in the figure, the secondary residues have a higher probability for zero and a smaller variation than the HEVC residues. Thus, the secondary residues can be coded with fewer bits in the entropy coding.

The proposed residual DPCM method can be implemented in different ways depending on what the reference sample is used for in the residual prediction. One use for the reference sample is as the original residue that is not yet quantized. The original residue is directly used for a reference like Eq. (1). This is very simple, and makes it possible to perform block-based processing, which provides benefits in hardware implementation. However, a mismatch may occur between the residual prediction loops at the encoder and decoder, which results in coding loss.

The other use for the reference sample in residual prediction is the reconstructed residue, $\tilde{r}_{x,y}$, which is obtained as follows:

$$\tilde{r}_{x,y} = \hat{r}^{rec}_{x,y} + QS(r_{x,y} - \hat{r}^{rec}_{x,y}) \quad (3)$$

$$\hat{r}^{rec}_{x,y} = \left(1 - w_y\right) r^{rec}_{i,y-1} + w_y r^{rec}_{i+1,y-1} \quad (4)$$

where $QS(\cdot)$ represents the pixel-by-pixel quantization and scaling processes, and $\hat{r}^{rec}_{x,y}$ is the predicted residue derived by using the previously reconstructed residues. This scheme achieves better coding efficiency because a mismatch does not occur at all. However, it requires pixel-by-pixel processing for quantization and scaling. The pixel-by-pixel processing is strictly causal, and it may be problematic in hardware implementation. The two residual prediction schemes using the original and reconstructed residues will be evaluated in the next section.

HEVC supports 35 intra-prediction modes. Most of them require interpolation/extrapolation processes to obtain predicted samples. In the residual DPCM method, the processes are also required in order to obtain the predicted residues, as in (1) and (4). To reduce excessive computational costs associated with these processes, we apply the residual DPCM to only the five intra-prediction modes: intra-prediction modes 2 (right-up diagonal), 10 (horizontal), 18 (right-down diagonal), 26 (vertical), and 34 (left-down diagonal), in which the predicted residue is obtained by only copying one integer reference residue from the reference row (or column) as follows:

$$\hat{r}_{x,y} = r_{x+\tan\theta, y-1} \quad (5)$$

where tanθ is equal to 0, 1, or -1 according to the five intra-prediction modes.

In HEVC, the five intra-prediction modes are selected on an average of 41% for the HEVC common test sequences [2].

3. Experimental Results

The proposed residual DPCM method was implemented on HEVC Test Model 16.6 (HM 16.6) [7]. Among the common test sequences for HEVC [2], the class F sequences consist of BasketballDrillText, ChinaSpeed, SlideEditing, and SlideShow, all of which are screen content. These sequences were tested to evaluate the proposed method.

First, we evaluated the net performance of the residual DPCM, independent of other HEVC coding tools. For this net performance evaluation, all blocks in the test sequences were coded in 4 × 4 intra-vertical prediction mode, and transforms were always skipped, as in the experiments illustrated in Fig. 2. As shown in Table 1, applying residual DPCM for all blocks achieved an average 12.8% Bjøntegaard delta rate (BD-rate) reduction, compared to not applying residual DPCM. This experimental result reveals that the proposed method significantly improves the coding efficiency of transform skip mode for the vertically intra-predicted blocks.

To evaluate the overall performance of the proposed method in the sense of RDO, residual DPCM was compared with HEVC for the HEVC Main_AI configuration, which is one of the HEVC common test conditions [2]. When a block is coded as one of the five

Table 2. Coding efficiency and computational complexity for reference-sample generation schemes of residual DPCM.

Sequences	Residual DPCM using original residues					Residual DPCM using reconstructed residues				
	Y BD-rate (%)	U BD-rate (%)	V BD-rate (%)	Encoding time (%)	Decoding time (%)	Y BD-rate (%)	U BD-rate (%)	V BD-rate (%)	Encoding time (%)	Decoding time (%)
BasketballDrillText	-0.5	-0.7	-0.6	99	101	-0.5	-0.8	-0.7	101	101
ChinaSpeed	-0.7	-1.3	-1.2	100	101	-0.7	-1.5	-1.5	103	100
SlideEditing	-1.1	-2.6	-2.5	101	100	-1.7	-3.2	-3.1	103	100
SlideShow	-1.9	-1.8	-1.9	99	102	-2.0	-2.0	-2.0	103	102
Average	**-1.0**	**-1.6**	**-1.6**	**100**	**101**	**-1.2**	**-1.9**	**-1.8**	**102**	**101**

intra-prediction modes, and transform skip mode is applied for RDO, the proposed method always applied residual DPCM to the block. Experimental results are listed in Table 2 for the two reference-sample generation methods using the original residues and the reconstructed residues. In the BD-rate comparison, the anchor was the bitstream generated by using the HEVC Main_AI configuration. As shown in the table, the proposed method achieved a 1.2% BD-rate reduction with a negligible complexity increase, and it consistently outperformed HEVC over all test sequences. Note that the number of blocks coded as one of the five intra-prediction modes (with transform skip mode applied for RDO) was only 8.2% of the total number of blocks. Considering that the proposed method affected only 8.2% of the blocks, the 1.2% BD-rate gain can be considered relatively significant.

As shown in the table, the proposed method achieved more coding gains for the SlideEditing and SlideShow sequences than it did for the BasketballDrillText sequence. This is because SlideEditing and SlideShow contain wider regions of text/graphics, in which the proposed method can provide better performance. Consequently, the experimental results showed that the proposed residual DPCM can provide high coding efficiency for text/graphic regions in screen content.

4. Conclusion

For screen content coding, to reduce the spatial redundancy in intra-prediction errors, this paper presented a residual DPCM method based on HEVC. The proposed method predicts residues on a pixel-by-pixel basis along the intra-prediction direction in intra transform skip mode. Experimental results showed that the transform skip mode was improved by 12.8%, in terms of BD-rate, for vertically intra-predicted blocks. Overall, the proposed method showed an average 1.2% reduction in BD-rate relative to HEVC with negligible computational complexity. The proposed method, therefore, would be well suited to enhancing screen content coding, particularly as the distribution of screen content is becoming increasingly widespread.

Acknowledgement

This work was supported by an Institute for Information & communications Technology Promotion (IITP) grant funded by the Korea government (MSIP) (B0126-15-1013, development of generation, and consumption of jigsaw-liked ultra-wide viewing spacial media).

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] C. Lan, J. Xu, F. Wu, and G. J. Sullivan, "Screen content coding," JCTVC-B084, Geneva, July 2010.

[4] C. Lan, J. Xu, G. J. Sullivan, and F. Wu, "Intra transform skipping," JCTVC-I0408, Geneva, April, 2012.

[5] P. Onno, E. François, L. Zhao, J. An, Y.-W. Huang, S. Lei, and D. Zhao, "'Combination of J0171 and J0389 for the non-normative encoder selection of the intra transform skip," JCTVC-J0572, Stockholm , July 2012.

[6] J. Lainema, F. Bossen, W.-J. Han, J. Min, and K. Ugur, "Intra coding of the HEVC standard," IEEE Trans. Circuits and Systems for Video Technology, vol. 22, issue 12, pp. 1792-1801, Dec. 2012. Article (CrossRef Link)

[7] High Efficiency Video Coding Test Model Software 16.6, Article (CrossRef Link).