EFFICIENT INTRA CODING STRUCTURE FOR HIGH RESOLUTION VIDEOS USING LINE-BY-LINE PREDICTION AND ADAPTIVE TRANSFORM SELECTION

Jung-Ah Choi and Yo-Sung Ho

Gwangju Institute of Science and Technology (GIST)
261 261 Cheomdan-gwagiro, Buk-gu, Gwangju 500-712 Republic of Korea

ABSTRACT

In this paper, an efficient intra coding structure for high resolution videos is proposed. First, we use a line-by-line intra 16×16 prediction to improve the prediction accuracy. Second, we modify the intra coding structure for the line-by-line prediction and an adaptive transform selection is added to skew the prediction mode. Experimental results show that the proposed method provides approximately 10.415% bit saving, compared to the current H.264/AVC FRExt high profile over several high-definition (HD) test sequences.

Index Terms— High quality video coding, H.264/AVC, intra coding, intra prediction, adaptive transform

1. INTRODUCTION

Recently, the demands for high quality video are increased and a lot of video materials are required to be distributed with high resolution. Since these high resolution videos have large data size, they require a higher compression. Therefore, a new generation of video compression technology for high resolution video is required.

To satisfy these commercial demands, the call for proposals (CfP) on video compression technology is finally issued jointly by ITU-T Video Coding Expert Group (VCEG) and ISO/IEC Moving Picture Expert Group (MPEG) on January 22, 2010 [1].

A Joint Collaborative Team (JCT) on video coding standard development will also be established accordingly between VCEG and MPEG. They decide high efficiency video coding (HEVC) will have a backward compatibility with the H.264/AVC.

H.264/AVC is the latest video coding standard established by Joint Video Team (JVT) between MPEG and VCEG. H.264/AVC shows good coding efficiency, however it cannot achieve an outstanding performance in high resolution video including high-definition (HD) videos.

H.264/AVC uses an intra prediction to reduce the spatial redundancy between spatially adjacent blocks. The intra prediction technique is recognized to be one of the main factors that contributes to the high coding performance of the conventional H.264/AVC.

Several researches have been studied on the improvements of intra coding based on H.264/AVC [2][3]. Bi-directional prediction (BIP) is proposed [2]. Authors are also designed new directional transform for intra frame. The concept of template matching is suggested [3].

In this paper, we propose a new intra coding structure that can replace the conventional H.264/AVC intra coding. We design the line-based prediction instead of the traditional block-based prediction. Also, we analysis the mode distribution after encoding the high resolution video using discrete cosine transform (DCT) and mode-dependent directional transform (MDDT). Then, we adaptively select the transform and find the efficient structure for skewing prediction mode to intra 16×16 mode.

2. MODE DEPENDENT DIRECTIONAL TRANSFORM

After intra prediction, the predicted block is subtracted from the original block and the residual block is produced. Although the residual block contains small number of data, there is still structural information left in the residual block.

To compact the residual signal energy, MDDT was proposed [5]. MDDT shows better energy compaction performance than DCT, since it is based on KLT.

MDDT consists of a series of pre-defined separable transforms. Each transform is efficient in compacting energy along one of the prediction directions, thus favoring one of the intra modes. The type of MDDT is coupled with the selected intra prediction mode, so is not explicitly signaled.

However, MDDT is not optimal, because MDDT is separable and designed based on general statistics. It may not accord with local statistics of certain video sequences.

3. PROPOSED ALGORITHM

3.1. Analysis of Mode Distribution of H.264/AVC

We encoded the several HD test sequences using the conventional H.264/AVC with DCT. In this experiment, the quantization parameter (QP) is equal to 22 and CABAC is used. The best mode distribution for each sequence is shown in Fig. 2.

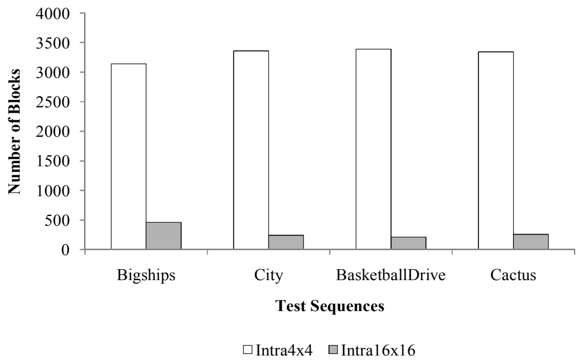


Fig. 2. Mode distribution in 720p HD sequence.

In high resolution video, most parts are coded as smaller block due to the complex texture. Choosing smaller partition sizes may result in a lower number of bits to encode the residual error at the expense of a larger number of bits to encode the mode information for each partitioned block.

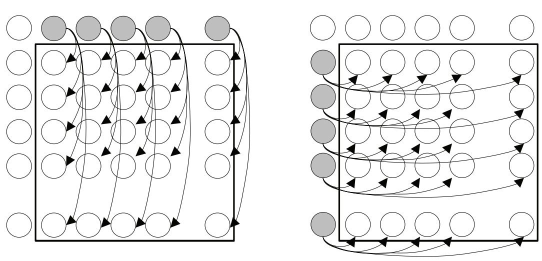
By contrast, choosing larger partition size requires a smaller number of coding bits to encode mode information, at the expense of containing a significant amount of energy in residual error in areas of high detail.

Thus, if we use more accurate intra 16×16 prediction and skew the best mode to intra 16×16 mode, we can efficiently reduce the entire coding bits.

3.2. Proposed Intra Coding Structure

3.2.1. Line-by-line (LbL) Prediction

Figure 3 shows the prediction process for the vertical mode and the horizontal mode of intra 16×16 block in H.264/AVC. In the conventional H.264/AVC, the block-based prediction is used. The prediction formulas of the vertical mode and the horizontal mode are stated in Eq. (1) and Eq. (2), respectively. Here, **pred**(x,y) is predicted luminance sample and **p**(x,y) is neighboring sample. The numbers from -1 to 15 represent the pixel position. For instance, -1 means the already encoded neighboring pixels, the shaded pixels in the Fig. 3.



(a) Vertical mode (b) Horizontal mode

Fig. 3. H.264/AVC intra 16×16 prediction.

Vertical mode:

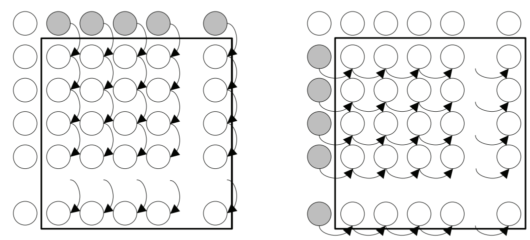
$$\mathbf{pred}(x,y) = \mathbf{p}(x,-1) \tag{1}$$

Horizontal mode:

$$\mathbf{pred}(x,y) = \mathbf{p}(-1,y) \tag{2}$$

As shown, 256 pixels within the current macroblock are predicted using 16 neighboring pixels. Compared with the number of neighboring pixels, a large number of pixels should be predicted from the neighboring pixels. Thus, we cannot expect high prediction accuracy in the block that has variation in pixel values such as the gradient.

To improve the prediction accuracy, the line-based prediction is proposed instead of the traditional block-based prediction [6]. In the vertical and horizontal mode of intra 16×16, 256 pixels in the current macroblock are predicted using 16 neighboring pixels. Generally, closer pixels give better prediction accuracy. The process of the line-by-line prediction is shown in Fig. 4. We reformulated the prediction equation as follows:



(a) Vertical mode (b) Horizontal mode

Fig. 4. Line-by-line intra 16×16 prediction.

Vertical mode:

$$\mathbf{pred}(x,0) = \mathbf{p}(x,-1) \tag{3}$$

$$\mathbf{pred}(x,y) = \mathbf{r}(x,y-1) \tag{4}$$

Horizontal mode:

$$\mathbf{pred}(0,y) = \mathbf{p}(-1,y) \tag{5}$$

$$\mathbf{pred}(x,y) = \mathbf{r}(x-1,y) \tag{6}$$

First, we make the predicted pixel values of the first line within the current macroblock using Eq. (3) or Eq. (5) according to the mode. Then, the predicted pixel values are subtracted from the original pixel values and the residual data is formed. The transform, the quantization, the inverse quantization, the inverse transform are applied one after another to make the reconstructed pixel values, **r**(x,y).

This reconstructed pixel values are used as the reference data of the next line. According to the prediction mode, we use Eq. (4) or Eq. (6) as the prediction formula. The above prediction process is continued until the last line of the macroblock.

Note that we do not need any additional syntax elements to the H.264/AVC standard for the proposed algorithm. By applying the line-by-line prediction, the prediction accuracy of intra 16×16 block is highly improved. Thus, we can efficiently save the amount of the residual data.

3.2.2. Adaptive Transform Selection

We encoded one of the high resolution test sequence, *Jets* using different transforms; MDDT and DCT. For the same PSNR and bit-rate, we compared the mode distribution of two cases and the result is shown in Fig. 5.

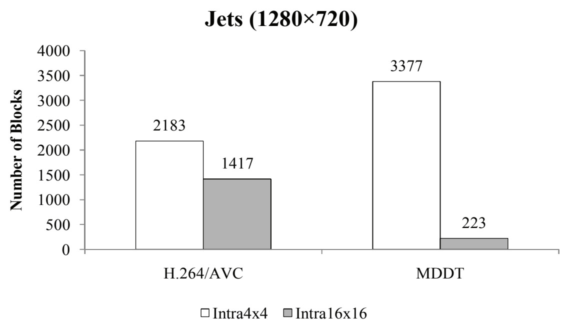


Fig. 5. Mode distribution comparison.

The mode distributions of two cases are quite different. From the result, we can observe that more intra 4×4 modes are selected when we use MDDT although two cases provide the same performance. Since intra 4×4 mode requires more coding bits for sending prediction modes, we try to give some penalty to prevent the excessive selection of the intra 4×4 mode.

MDDT shows better performance than DCT. It is shown that use of the directional transforms alone provides 5% performance gain on average [7]. Thus, if we use DCT to intra 4×4 mode and MDDT to intra 16×16 mode, more intra 16×16 modes are likely to be selected in the mode decision.

We modified the intra coding structure to give penalty to intra 4×4 mode. We adaptively use MDDT and DCT for each intra prediction mode. Figure 6 shows the block diagram of the proposed algorithm and the details are as follows:

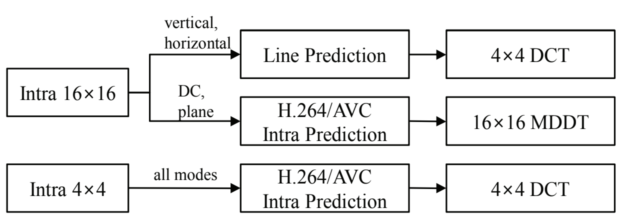


Fig. 6. Block diagram of the adaptive transform selection.

1) Intra 16×16 block (vertical and horizontal mode): we use the LbL prediction to improve the prediction accuracy. In LbL prediction, the predicted line (16 pixels) should pass to transform and quantization before the prediction of the next line. Thus, we use 4×4 DCT. We do not use 4×4 MDDT, since the residual data of the LbL prediction is quite different with the training residual data of 4×4 MDDT.
2) Intra 16×16 block (DC and plane mode): we use 16×16 MDDT to encourage the selection of intra 16×16 mode.
3) Intra 4×4 block (for all prediction modes): we use DCT to discourage the unnecessary selection of intra 4×4 mode.

4. EXPERIMENTAL RESULTS

In order to examine the performance of the proposed algorithm, we performed experiments on several test materials of the YUV 4:2:0 format with HD (1280×720) and Full HD (1920×1080) resolutions. All the tested sequences are intra only coded. We implemented the proposed algorithm on the H.264/AVC reference software version JM 12.2 [8]. Table 1 shows the encoding parameters.

Table 1. Encoding parameters

ProfileIDC	100 (high profile)
LevelIDC	51
IntraPeriod	1 (intra only)
QPISlice	22, 27, 32, 37
SymbolMode	1 (CABAC)
Transform8x8Mode	0

Our proposed method was applied to H.264/AVC intra coding by replacing the conventional intra coding. The Bjøntegaard delta peak signal-to-noise ratio (dB) and the Bjøntegaard delta bit rate (%) are used to evaluate the performance of the proposed algorithm [9].

Table 2. Performance comparison with H.264/AVC

Test Sequence	Size	Frame rate (Hz)	BD-PSNR (dB)	BD-Rate (%)
Bigships	1280 × 720	60	0.32	-6.564
City		60	0.69	-9.326
Crew		60	0.25	-8.711
Night		60	0.56	-7.736
Average			0.46	-8.084
BasketballDrive	1920 × 1080	50	2.05	-19.125
BQTerrace		60	1.02	-11.421
Cactus		50	0.84	-10.019
Average			1.30	-13.522

Table 2 provides BD-PSNR and BD-Rate when compared to H.264/AVC with DCT used for transform. From the table we observe that on average only 0.52 dB BD-PSNR or 10.415% BD-Bit rate we sacrifice using our proposed technique compared to the H.264/AVC. Through the proposed algorithm, the number of intra 16×16 modes are significantly increased using the proposed algorithm. Obviously, large blocks provide less coding bits.

Rate-distortion curves of the best case and worst case in 720p HD and 1080p Full-HD are given in Fig. 7 and Fig. 8. We evaluate the proposed method compared to H.264/AVC and MDDT. From the rate-distortion curves, we can observe that the proposed scheme is superior to both the current H.264/AVC FRExt high profile and the original MDDT.

5. CONCLUSION

In this paper, we propose an efficient H.264/AVC intra coding structure using the line-by-line prediction and the adaptive transform selection. We start the concept of most parts of the high resolution video is encoded using small blocks due to the complex texture. Then, we analyzed the mode distribution of MDDT and DCT in same quality and bit-rate. From the observations, we find that MDDT selects the excessive intra 4×4 mode. Thus, we designed a new line-by-line intra prediction and an adaptive transform selection to encourage the intra 16×16 mode rather than the intra 4×4 mode. Experimental results showed that the proposed method provides about 10.415% bit savings, compared to the current H.264/AVC FRExt high profile.

ACKNOWLEDGE

This research was supported by the MKE(The Ministry of Knowledge Economy), Korea, under the ITRC(Information Technology Research Center) support program supervised by the NIPA(National IT Industry Promotion Agency) (NIPA-2010-(C1090-1011-0003))

REFERENCES

[1] Video, Test and Requirements Subgroups, "Call for Evidence on High-Performance Video Coding (HVC)," ISO/IEC JTC1/SC29/WG11 MPEG, doc. w11113, Kyoto, April 2009,.
[2] T. Shiodera, A. Tanizawa, and T. Chujoh, "Bidirectional Intra Prediction," ITU-T SG16/Q.6 VCEG, VCEG-AE14, Marrakech, January 2007.
[3] T. Tan, C. Boon, Y. Suzuki, "Intra Prediction by Template Matching", in Proceedings ICIP 2006, October 2006.
[4] T. Wiegand, G. J. Sullivan, G. Bjøntegaard, and A. Luthra, "Overview of the H.264/AVC video coding standard," IEEE Trans. Circuits Syst. Video Technol., vol. 13, no. 7, pp. 560–576, July 2003.
[5] Y. Ye and M. Karczewicz, "Improved Intra Coding," ITU-T SG16/Q.6 VCEG, VCEG-AG11, Shenzhen, October 2007.
[6] J.A. Choi and Y.S. Ho, "Line-by-line Intra 16×16 Prediction for High-quality Video Coding," ICME 2010, to be published.
[7] Y. Ye and M. Karczewicz, "Complexity Analysis of Improved Intra Coding," ITU-T SG16/Q.6 VCEG, VCEG-AH20, Antalya, October 2007.
[8] http://iphome.hhi.de/suehring/tml/download/old_jm/jm12.2.zip, Joint Video Team, Reference Software Version 12.2.
[9] G. Bjøntegaard, "Improvements of the BD-PSNR Model", ITU-T SG16/Q.6 VCEG, VCEG-AI11, Berlin, July, 2008.

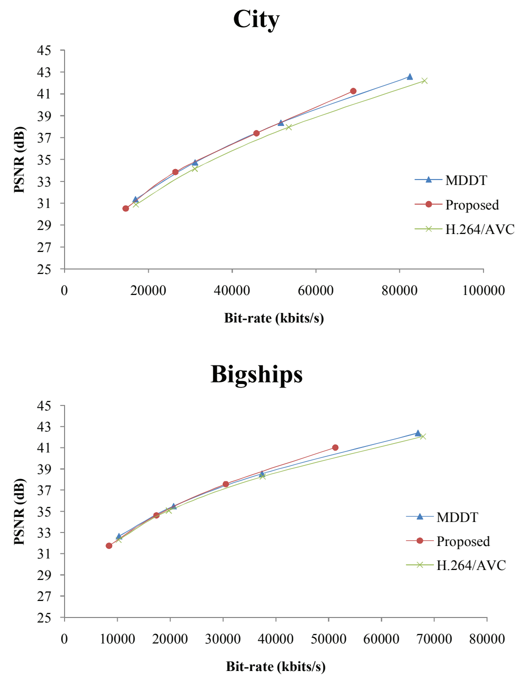


Fig. 7. Rate-distortion curves for 720p HD test sequences.

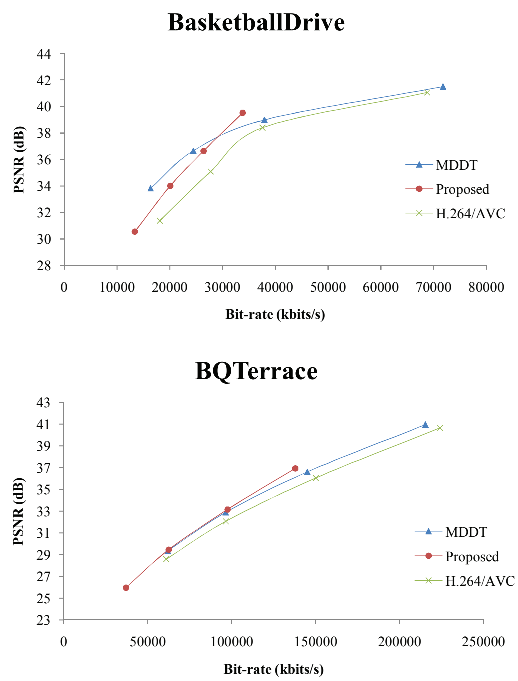


Fig. 8. Rate-distortion curves for 1080p Full-HD test sequences.