Adaptive Scanning for H.264/AVC Intra Coding

Yung-Lyul Lee, Ki-Hun Han, Dong-Gyu Sim, and Jeongil Seo

ABSTRACT—In this letter, an adaptive scanning that improves intra coding efficiency in the H.264/AVC standard is proposed. The proposed adaptive scanning utilizes the prediction directions (modes) that include the horizontal and vertical edge information in a block. Depending on the prediction directions, the proposed method uses three scanning methods: zigzag scanning, horizontal scanning, and vertical scanning. In the proposed method, horizontal and vertical scanning are used in vertical and horizontal prediction modes, respectively, and the normal zigzag scanning in the H.264 standard is used in all other intra prediction modes. The proposed method reduces the bit rate by approximately 2.5% compared with H.264/AVC, without the degradation of video quality.

Keywords—Video coding, intra coding, H/264, AVC, scanning, prediction.

I. Introduction

The latest video coding standard H.264/advanced video coding (AVC), which is well-known to provide high coding efficiency, was developed by the joint work of ITU-T and ISO/IEC [1]-[4]. It reduces the bit rate by approximately 30% - 50% compared with previous video coding standards such as MPEG-4 Part 2 Visual [5], [6], H.263 [7], and so on. Its high coding efficiency is made possible by new advanced coding tools such as variable block size motion estimation (ME), multiple reference frames, quarter-pixel accuracy ME, spatial prediction for intra coding [8], and so on. Usually, inter coding is superior to intra coding, but intra coding is useful for various purposes such as random access, video editing, and scene extracting. In this letter, we propose an adaptive scanning to improve intra-coding efficiency in H.264/AVC.

H.264/AVC uses zigzag scanning for compressing the quantized DCT coefficients in a 4×4 block without considering the intra prediction directions. Therefore, different scanning methods that make use of vertical, horizontal, and zigzag scanning according to the spatial prediction directions are proposed in this letter. The proposed method can be easily applied to an alternative scanning with semantic changes in the normal zigzag scanning for a frame macroblock (MB) in the H.264/AVC standard because it does not require any syntax change. In the following sections, the existing H.264/AVC intra coding is explained in brief and the proposed adaptive scanning methods are described. The experimental results of our proposed method are provided in section IV, and our concluding remarks are given in section V.

II. H.264/AVC Intra Coding

H.264/AVC has several types of intra coding modes such as intra 16×16, intra 8×8 (FRExt-only [9]), intra 4×4, and intra chroma modes. Intra 16×16, intra 8×8, and intra 4×4 modes are used to encode the luma component and those intra coding modes are performed in block-based prediction using the spatially adjacent block boundary pixels which were already decoded. The intra chroma mode is used to encode the chroma component, and the size of the prediction block in the mode depends on the image color format. For example, the intra chroma mode has an 8×8 prediction block in the 4:2:0 color format and a 16×16 prediction block in the 4:4:4 color format. In this letter, the intra chroma mode has an 8×8 prediction block because the 4:2:0 sequences are used.

To encode an MB in intra 16×16 mode in the luma component, all of the MB pixels are predicted from the block boundary pixels of the neighboring previously decoded MBs. The intra 16×16 mode has four different prediction methods

---

Manuscript received May 23, 2006, revised June 22, 2006.

Yung-Lyul Lee (phone: +82 2 3408 3753, email: yllee@sejong.ac.kr) and Ki-Hun Han (email: khan@dms.sejong.ac.kr) are with the Department of Computer Engineering, Sejong University, Seoul, Korea.

Dong-Gyu Sim (email: dgsim@kw.ac.kr) is with the Department of Computer Engineering, Kwangwoon University, Seoul, Korea.

Jeongil Seo (email: seoji@etri.re.kr) is with Radio & Broadcasting Research Division, ETRI, Daejeon, Korea.

such as vertical, horizontal, DC, and planar prediction modes. The best mode for each MB in terms of rate-distortion is selected in the encoder. For vertical and horizontal prediction, the pixels of an MB are predicted from the pixels located just above or to the left which are on the previously-decoded MB boundary, respectively. In DC prediction, the average value of the neighboring 32 pixels situated on the block boundary that is previously decoded is used as the predictor. In planar prediction, a three-parameter curve-fitting equation is used to form a prediction block having a brightness and slope in the horizontal and vertical directions that approximately matches the neighboring pixels.

The intra 4×4 mode can be alternatively selected according to the block-based rate-distortion value in the encoder. The pixels in the 4×4 block are predicted from the neighboring pixels that are above and/or left of the current block, and one prediction mode among the nine different directional prediction modes is selected. The intra chroma mode has the same prediction modes as the intra 16×16 mode in luma.

III. Proposed Adaptive Scanning in Intra Prediction

After a spatial prediction, H.264/AVC performs a 4×4 integer transform on the residual signals for energy compaction. This transform is based on the discrete cosine transform (DCT). The integer transform can be expressed in matrix form by using the separable property of the unitary transform as follows:

$$\mathrm{Y}=\left(\begin{bmatrix}1&1&1&1\\2&1&-1&-2\\1&-1&-1&1\\1&-2&2&-1\end{bmatrix}[R]\begin{bmatrix}1&2&1&1\\1&1&-1&-2\\1&-1&-1&2\\1&-2&1&-1\end{bmatrix}\right)\otimes[E],\ (1)$$

where R is 4×4 residual signals and E is the post-scaling matrix. The operation $\otimes$ means element-by-element multiplication.

The basis function of the transform is shown in Fig. 1. In Fig. 1, white positions have "+" sign and gray positions have "–" sign, in which the weight of each position is ignored. In Fig. 1, we can perceive that the edge direction in the residual block affects the distribution of transform coefficients.

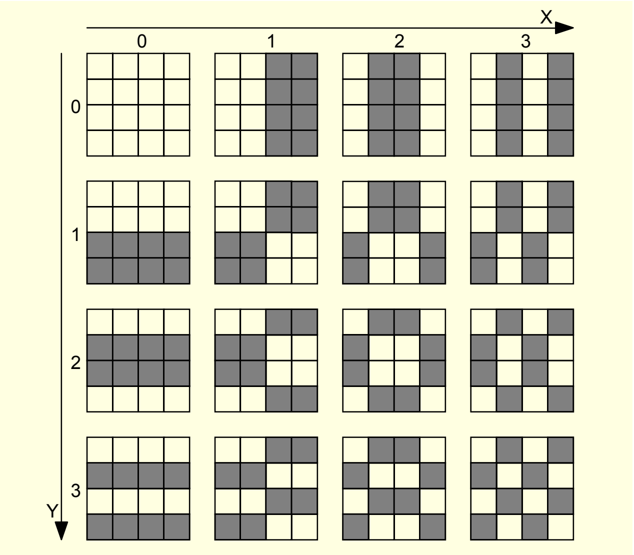


Fig. 1. The basis function of 4×4 integer transform.

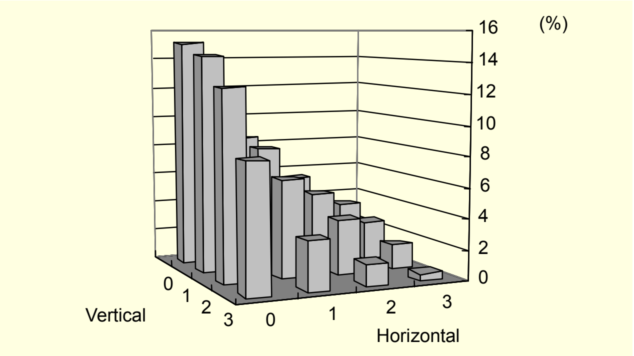


Fig. 2. Distributions of significant transformed coefficients in 4×4 block when the horizontal block prediction mode is chosen for various test sequences.

As an example, when a block contains a vertical edge, the block will be assigned the vertical prediction mode and its transformed coefficients will have a relatively large magnitude in the first row since the vertical predicted residual pixels will have a high correlation in the vertical direction (that is, 4×4 DCT coefficients of the residual pixels will be in the first row). In a similar way, when a block contains a horizontal edge, its transform coefficients will have a relatively large magnitude in the first column. As a reference, the distribution of significant (non-zero) DCT coefficients is plotted for various test sequences with the difference QP in Fig. 2 when the block is assigned the horizontal prediction mode. Analysis shows that approximately 15% of significant DCT coefficients are in the (0, 0) position, 14% are in the (0, 1) position, 12% are in the (0, 2) position, 8% are in the (0, 3) position, and so on. The first column contains about 50% of the significant DCT coefficients. Also, the first row contains about 48% of the significant coefficients when the vertical block prediction mode is chosen in our experiments. Since H.264/AVC performs spatial prediction, we can estimate the edge direction according to the intra prediction direction (mode).

Figure 3 shows the vertical and horizontal prediction in intra 4×4 mode of the H.264/AVC standard. If the prediction mode chosen is vertical, it means that each pixel value is vertically similar and the vertical edge is more probable than the horizontal edge in the block. Also, if the prediction mode chosen is horizontal, it means that each pixel value is horizontally similar

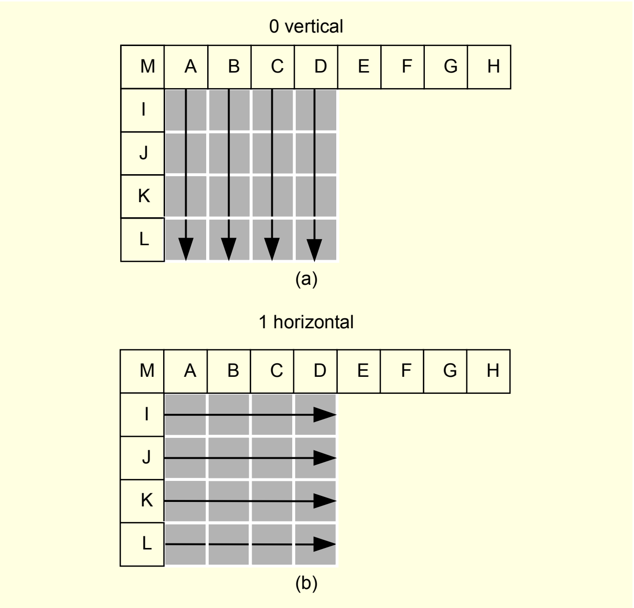
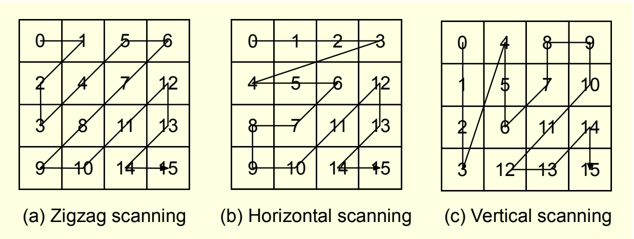


Fig. 3. Vertical and horizontal prediction in intra 4 × 4 mode.



(a) Zigzag scanning (b) Horizontal scanning (c) Vertical scanning

Fig. 4. Proposed adaptive scanning including zigzag, horizontal, and vertical scanning order, where the number in each scanning means the scanning order.

and the horizontal edge is more probable than the vertical edge in the block; therefore, the proposed method in Fig. 4 utilizes adaptive scanning according to the intra prediction direction. Figure 4(a) shows the normal zigzag scanning used in H.264. Figure 4(b) shows the horizontal scanning which is efficient in the vertical prediction mode because significant coefficients are more probable in the first row than in other rows. Figure 4(c) shows the vertical scanning that is efficient in the horizontal prediction mode because significant coefficients are more probable in the first column than in other columns. The horizontal scanning and vertical scanning in Fig. 4 are experimentally and intuitively derived from our experiments. The distribution of significant coefficients and scanning order is analyzed to improve coding efficiency in our experiments.

In the proposed method, the horizontal and vertical scanning order are used in the vertical and horizontal prediction mode, respectively, and the zigzag scanning order is used in all other prediction directions. Since both encoder and decoder perform spatial prediction, the proposed method does not require any additional information. Also, the proposed method does not require additional complexity to obtain the edge direction.

IV. Experimental Results

In this letter, we propose an adaptive scanning method to improve the intra coding efficiency of H.264/AVC. To verify the validity of the proposed method, experiments were performed on various test sequences which were recommended for the H.264/AVC experiment. We used the vertical, horizontal, and DC prediction methods for the experiments because the proposed method was applied only to vertical and horizontal prediction. All of the test sequences were coded in intra frame for four quantization values, QP, and one hundred frames were coded in every sequence. The experimental results are shown in Table 1.

In order to evaluate the performance of the proposed method, the proposed method is compared with the H.264 joint model 96 (JM96) reference codec [10].

As demonstrated in Table 1, the proposed method reduces the bit rate by approximately 2.5% on average, while similar peak signal-to-noise ratio (PSNR) values are obtained for every sequence compared with H.264/AVC. As a reference, the proposed method reduces approximately 1.7% of the bit rates when all prediction modes in the H.264 standard are used. But if more scanning methods are included in the H.264 standard according to the block prediction modes, then the gain will be increased.

Table 1. Experimental results.

Sequence	QP	H.264/AVC		Proposed method	
		PSNR (dB)	Bit rates	PSNR (dB)	Bit rates
News (QCIF)	25	39.96	1070.81	39.95	1043.16
	30	35.80	707.29	35.80	688.98
	35	31.89	450.18	31.86	435.86
	40	28.32	269.91	28.32	264.69
Container (QCIF)	25	39.37	1010.98	39.36	998.63
	30	35.54	636.77	35.55	624.41
	35	31.91	389.47	31.88	378.43
	40	28.50	234.42	28.49	229.54
Paris (CIF)	25	39.04	5433.02	39.02	5315.47
	30	34.86	3571.56	34.84	3475.70
	35	31.00	2235.27	30.98	2166.06
	40	27.45	1346.55	27.43	1305.26

V. Conclusion

In this letter, we proposed adaptive scanning according to the spatial prediction direction to improve intra coding efficiency. The proposed method does not require any additional computation or signaling bit. Experimental results show that the proposed method reduced the bit rate by approximately 2.5% on average, while the PSNR of the video sequences are maintained.

References

[1] ITU-T Recommendation H.264 and ISO/IEC 14496-10, *Advanced Video Coding for Generic Audiovisual Services*, May 2003 (and subsequent amendment [8] and corrigenda).

[2] A. Luthra, G. J. Sullivan, and T. Wiegand, "Introduction to the Special Issue on the H.264/AVC Video Coding Standard," *IEEE Trans. Circuits Syst. Video Technol.*, July 2003, pp. 557-559.

[3] 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.*, July 2003, pp. 560-576.

[4] G. J. Sullivan and T. Wiegand, "Video Compression – from Concepts to the H.264/AVC Standard," *Proc. IEEE,* Jan. 2005, pp. 18-31.

[5] ISO/IEC JTC1/SC29/WG11 N3056, *Information Technology - Coding of Audio-Visual Objects Part 2: Visual Amendment 1: Visual Extensions*, Dec. 1999.

[6] S.-M. Kim, J.-H. Park, S.-M. Park, B.-T. Koo, K.-S. Shin, K.-B. Suh, I.-K. Kim, N.-W. Eun, and K.-S. Kim, "Hardware-Software Implementation of MPEG-4 Video Codec," *ETRI J.*, vol. 25, no.6, Dec. 2003, pp. 489-502.

[7] ITU Telecom Standardization Sector, "Video Codec Test Model Near-Term, Version 10 (TMN10) Draft 1," *H.263 Ad Hoc Group*, Apr. 1998.

[8] K.-B. Suh, S.-M. Park, and H.-J. Cho, "An Efficient Hardware Architecture of Intra Prediction and TQ/IQIT Module for H.264 Encoder," *ETRI J.*, vol. 27, no. 5, Oct. 2005, pp. 511-524.

[9] G. J. Sullivan, T. McMahon, T. Wiegand, and A. Luthra (eds.), "Draft Text of H.264/AVC Fidelity Range Extensions Amendment to ITU-T Rec. H.264 | ISO/IEC 14496-10 AVC," *ISO/IEC JTC1/SC29/WG11 and ITU-T Q6/SG16 Joint Video Team Document JVT-L047*, July 2004.

[10] http://iphome.hhi.de/suehring/tml/download/old_jm/jm96.zip