Reduced-Resolution Intra Block Coding Mode

Sung Jae Park, Jung-hak Nam, Dong-gyu Sim, Seoung Jun Oh, and Jin Woo Hong

ABSTRACT—In this letter, a new intra-block coding mode is presented to improve the coding efficiency for band-limited signals. A band-limited block is sub-sampled, and the sub-sampled signal is coded on the basis of the conventional prediction/transform coding. The rest of the samples are reconstructed by interpolation at the decoder side without any side information. Experimental results show that the proposed algorithm achieves coding gains of 2.7% for common intermediate format (CIF), 4.29% for quarter CIF, and 6.39% for 720p60 sequences against the H.264/AVC JM10.2 reference software.

Keywords—H.264/AVC, interpolation, intra-block coding.

I. Introduction

H.264/AVC significantly outperforms the previous video coding standards with many new coding tools. Among these tools, several intra-block coding tools can particularly improve coding efficiency without exploiting temporal redundancy. Thus, intra-only coding has been applied to many high-quality applications such as studio and digital cinema applications. The H.264/AVC intra-prediction method based on several 4×4 spatial prediction modes [1] results in an increased coding gain even for image sources containing a large amount of spatial high-frequency components. In general, natural images and videos have many homogenous regions whose power spectrum is band-limited. According to the Nyquist sampling theorem, the signals that are sufficiently band-limited can be sub-sampled and perfectly reconstructed without any information loss.

In this letter, we propose a new reduced-resolution intra-block coding mode to improve the coding efficiency for band-limited blocks and to enable spatial prediction to be conducted in parallel.

II. Proposed Algorithm

Common natural and synthetic images contain many band-limited blocks, which are more easily found in larger images. Because of this band-limit property, an anti-aliasing filter is not required in our work. An 8×8 block is partitioned into four 4×4 sub-blocks, **a**, **b**, **c,** and **d** as shown in Fig. 1.

In each 8×8 block, the 4×4 pixels which are vertically and horizontally sub-sampled by a factor of 2 are predicted with 9 modes, in the same manner in H.264/AVC. The residual of the sub-sampled block is coded with the H.264/AVC integer transform/quantization and entropy coding. The rest of the sub-sampled blocks are neither coded nor transmitted to a decoder. However, they are reconstructed using multiple 4-tap interpolation filters. Note that several sub-sampling-based

a(1,1)	b(1,1)	a(1,2)	b(1,2)	a(1,3)	b(1,3)	a(1,4)	b(1,4)
c(1,1)	d(1,1)	c(1,2)	d(1,2)	c(1,3)	d(1,3)	c(1,4)	d(1,4)
a(2,1)	b(2,1)	a(2,2)	b(2,2)	a(2,3)	b(2,3)	a(2,4)	b(2,4)
c(2,1)	d(2,1)	c(2,2)	d(2,2)	c(2,3)	d(2,3)	c(2,4)	d(2,4)
a(3,1)	b(3,1)	a(3,2)	b(3,2)	a(3,3)	b(3,3)	a(3,4)	b(3,4)
c(3,1)	d(3,1)	c(3,2)	d(3,2)	c(3,3)	d(3,3)	c(3,4)	d(3,4)
a(4,1)	b(4,1)	a(4,2)	b(4,2)	a(4,3)	b(4,3)	a(4,4)	b(4,4)
c(4,1)	d(4,1)	c(4,2)	d(4,2)	c(4,3)	d(4,3)	c(4,4)	d(4,4)

Fig. 1. Four 4×4 sub-blocks in an 8×8 block.

Manuscript received Sept. 12, 2008; revised Nov. 21, 2008; accepted Dec. 19, 2008.

This work was supported by MKE/IITA, Rep. of Korea [2008-F-011-01, Development of Next-Generation DTV Core Technology] and the Research Grant of Kwangwoon University in 2008.

Sung Jae Park (phone: + 82 2 940 5102, email: sjpark@media.kw.ac.kr) and Seoung Jun Oh (email: sjoh@kw.ac.kr) are with the Department of Electronic Engineering, Kwangwoon University, Seoul, Rep. of Korea.

Jung-hak Nam (email: qejixfyza@kw.ac.kr) and Dong-gyu Sim (phone: + 82 2 940 5470, email: dgsim@kw.ac.kr) are with the Department of Computer Engineering, Kwangwoon University, Seoul, Rep. of Korea.

Jin Woo Hong (email: jwhong@etri.re.kr) is with the Broadcasting & Telecommunications Convergence Research Laboratory, ETRI, Daejeon, Rep. of Korea.

 © 2009 ETRI

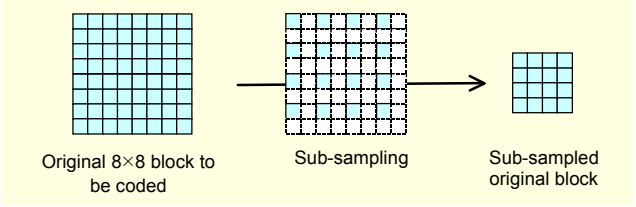

Fig. 2. Sub-block **a** to be coded from an 8×8 input block.

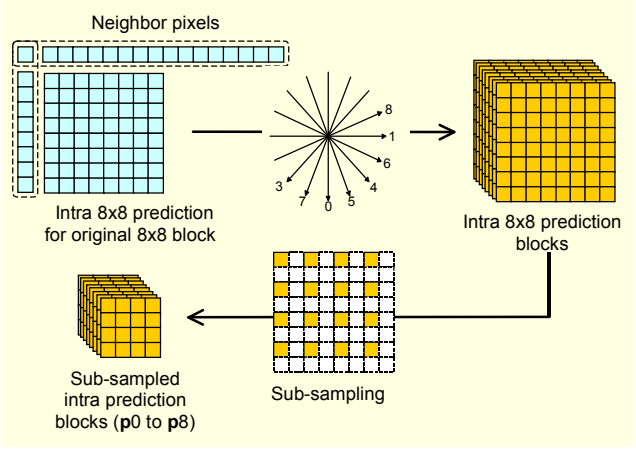

Fig. 3. Sub-sampled intra prediction blocks for the sub-block **a**.

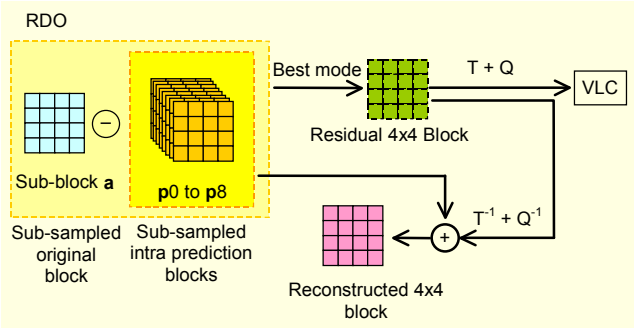

Fig. 4. Reconstruction process for sub-block **a** with nine sub-sampled intra prediction blocks.

video coding algorithms have been proposed [2]. However, they were utilized to reduce the computational complexity in mode selection. The proposed algorithm is a new video coding tool, so the proposed intra prediction should be conducted in the encoder/decoder sides.

The proposed intra-block coding mode consists of the four following steps.

Step 1. The 4×4 sub-sampled block to be coded is extracted from an original 8×8 block by sub-sampling as shown in Fig. 2.

Step 2. Nine H.264/AVC intra 8×8 prediction blocks are generated with neighboring pixels, and the nine 8×8 prediction blocks are sub-sampled to generate nine 4×4 sub-sampled intra prediction blocks (**p**0 to **p**8) as shown in Fig 3. Then, they are used for rate-distortion optimization (RDO) by comparing them with sub-block **a**.

Step 3. The 4×4 residual signals between the sub-sampled original and sub-sampled prediction blocks are coded using the H.264/AVC 4×4 integer transform and quantization. Then, the quantized coefficients are transferred to the H.264/AVC variable length coding module and the inverse quantization/ transformation module.

As shown in Fig. 4, the 4×4 reconstructed block is obtained by adding the reconstructed residual block to the prediction block. For each prediction mode, steps 2 and 3 are repeated and the best prediction mode can be chosen according to RDO.

Step 4. The rest of the sub-sampled blocks, such as **b**, **c,** and **d** in Fig. 1 are reconstructed by using only interpolation filters without any prediction. If the 8×8 original block is band-limited, the block can be perfectly reconstructed, aside from quantization error. The pixel values in sub-block **b** in Fig. 1 are interpolated with horizontal 4-tap filters $h_n(k)$. The pixel values are computed by

$$b(m,n) = \sum_{k=1}^{4} a(m,k) \cdot h_n(k) \qquad \text{for } 1 \le m,n \le 4,$$

where $b(m,n)$ denotes the (m,n)th element of sub-block **b** to be interpolated with horizontal filter $h_n(k)$, and $a(m,n)$ represents the (m,n)th element of sub-block **a**, which is the block decoded according to step 3. The coefficients of the horizontal filter $h_n(k)$ are (k=1,2,3,4). Depending on the sub-pel position, we make use of four 4-tap filters: $h_1(k)$=[16,20,-5,1]/32, $h_2(k)$=[-4, 20,20,-4]/32, $h_3(k)$=[1,-5,20,16]/32, and $h_4(k)$=[0,1,-5,36]/32. These coefficients of the proposed algorithm are derived from the 6-tap filter, [1,-5,20,20,-5,1]/6 in H.264/AVC. To reduce dependency of neighboring blocks, we use only decoded pixels in sub-block **a** in this filtering process. For example, when we interpolate the position $b(1,1)$, the left two pixels are not available so that we use (1,1) for the two pixels as follows:

$$\begin{aligned} b(1,1) &= [a(1,1)\times 1 + a(1,1)\times(-5) + a(1,1)\times 20 \\ &\quad + a(1,2)\times 20 + a(1,3)\times(-5) + a(1,4)\times 1]/32 \\ &= \big(a(1,1)\times 16 + a(1,2)\times 20 + a(1,3)\times(-5) + a(1,4)\times 1\big)/32. \end{aligned}$$

As a result, the 6-tap filter [1,-5,20,20,-5,1]/32 is simplified into the 4-tap filter [16, 20, -5, 1]/32 based on the replication of the boundary pixel. Likewise, $h_2(k)$, $h_3(k)$, and $h_4(k)$ can be derived.

In the same manner, the pixel values in sub-block **c** are interpolated by vertically applying the 4-tap interpolation filters. The pixel values of sub-block **c** are reconstructed by

$$c(m,n) = \sum_{k=1}^{4} a(k,n) \cdot h_m(k) \qquad \text{for } 1 \le m,n \le 4,$$

where $c(m,n)$ denotes the (m,n)th element of **c** to be interpolated with the vertical filters. The pixel values in sub-block **d** are interpolated with the 4-tap filer [1/4,1/4,1/4,1/4] for

four neighboring pixels as follows:

$$d(m,n) = \frac{1}{4}\sum_{k=0}^{1}\sum_{l=0}^{1} a(m+k, n+l) \qquad \text{for } 1 \le m,n \le 4,$$

where $d(m,n)$ denotes the (m,n)th element of **d**. Since these three interpolation tasks do not have any dependency, they can be concurrently executed.

III. Experimental Results and Conclusion

We implemented the proposed algorithm on the JM 10.2, released by JVT. The experimental conditions are summarized in Table 1. The Bjontegaard delta peak signal-to-noise ratio (dB) and the Bjontegaard delta bit rate (%) [3] are used to evaluate the performance of the proposed method.

Table 1. Experimental conditions and test sequences.

GOP structure			Only intra frames
Entropy			CAVLC
Profile			Baseline profile
Frame skip	Frame rate	Size	Sequence (coding frames)
0	30	CIF	Foreman (300), Mobile (300), Tempete (260)
	60	720p	BigShips (150), Crew (150), City (150), Night (150), Shuttle (300, start at 150)
1	30	QCIF	Container (150), Foreman (150), Silent (150)
	30	CIF	Paris (150)

Table 2. BDPSNR and BDBR of the proposed algorithm for several test sequences.

Size	Sequence	QP 22, 27, 32, 37		QP 32, 37, 42, 47	
		BDPSNR	BDBR	BDPSNR	BDBR
CIF	Foreman	0.045	-1.231	0.216	-6.332
	Mobile	-0.016	0.158	0.003	-0.344
	Tempete	-0.009	0.049	0.077	-2.706
	Paris	0.006	-0.116	0.054	-1.419
	Average	0.006	-0.285	0.088	-2.700
QCIF	Container	0.027	-0.352	0.008	-0.367
	Foreman	0.017	-0.511	0.213	-6.127
	Silent	0.018	-0.685	0.215	-6.386
	Average	0.021	-0.516	0.145	-4.293
720p60	Big Ships	0.066	-2.074	0.219	-6.955
	Crew	0.183	-5.584	0.282	-6.177
	City	0.041	-1.186	0.225	-6.754
	Night	0.095	-1.792	0.234	-5.131
	Shuttle	0.297	-6.888	0.350	-6.919
	Average	0.136	-3.505	0.262	-6.387

Table 3. Ratio of macroblocks which are encoded by the proposed method. (%)

Size	Sequence	QP					
		22	27	32	37	42	47
CIF	Mobile	0.52	0.86	1.67	4.87	12.65	33.97
	Tempete	2.93	2.36	4.04	12.33	27.86	32.04
QCIF	Foreman	3.29	6.84	11.27	22.48	38.67	39.98
	Silent	1.36	2.88	10.08	32.20	39.62	21.28
720p	Night	13.42	19.05	25.39	31.41	27.14	19.20
	Shuttle	24.42	27.17	23.27	16.28	9.38	4.45

To employ the proposed new mode, one flag bit is needed for a macroblock. If the proposed intra coding mode is selected as the best mode, the flag bit is set to 1. Otherwise, it is set to 0.

The quantization parameter (QP) values of the mid-bitrate range are 22, 27, 32, and 37, and those of the low-bitrate range are 32, 37, 42, and 47. We achieved average bit savings of 2.7% for common intermediate format (CIF), 4.29% for quarter CIF (QCIF), and 6.39% for 720p sequences at the low-bitrate range against H.264/AVC. At the middle bitrate range, the average bit savings are 0.29% for CIF, 0.56% for QCIF, and 3.5% for 720p sequences. They are summarized in Table 2.

Table 3 shows the ratio of macroblocks which are encoded by the proposed method. The proposed intra coding mode is designed to improve coding efficiency for band-limited blocks, so the proposed intra mode is frequently selected for the high-resolution videos and videos coded with high QP values.

In this letter, we proposed a new intra block coding mode based on reduced-resolution prediction coding. In the proposed algorithm, a sub-sampled block is coded, and the macroblock is reconstructed with up-sampling and interpolation. Experimental results show that the proposed method achieves bit saving of 6.4% for 720p sequences.

References

[1] J. Ostermann et al., "Video Coding with H.264/AVC: Tools, Performance, and Complexity," *IEEE Circuits and Systems Magazine*, vol. 4, no. 1, 2004, pp. 7-28.

[2] K.S. Lee and H.J. Kim, "An Implementation of H.264 Intra Predictor Based on Sub-sampling," *Circuits and Systems, APCCAS, IEEE Asia Pacific Conf.*, Dec. 2006, pp. 1079-1082

[3] G. Bjontegaard, "Calculation of Average PSNR Differences between RD-Curves," ITU-T Q.6/SG16 VCEQ, VCEQ-M33, Apr. 2001.