A Fast Intra-Prediction Method in HEVC Using Rate-Distortion Estimation Based on Hadamard Transform

Younhee Kim, DongSan Jun, Soon-heung Jung, Jin Soo Choi, and Jinwoong Kim

A fast intra-prediction method is proposed for High Efficiency Video Coding (HEVC) using a fast intra-mode decision and fast coding unit (CU) size decision. HEVC supports very sophisticated intra modes and a recursive quadtree-based CU structure. To provide a high coding efficiency, the mode and CU size are selected in a rate-distortion optimized manner. This causes a high computational complexity in the encoder, and, for practical applications, the complexity should be significantly reduced. In this paper, among the many predefined modes, the intra-prediction mode is chosen without rate-distortion optimization processes, instead using the difference between the minimum and second minimum of the rate-distortion cost estimation based on the Hadamard transform. The experiment results show that the proposed method achieves a 49.04% reduction in the intra-prediction time and a 32.74% reduction in the total encoding time with a nearly similar coding performance to that of HEVC test model 2.1.

Keywords: HEVC, fast mode decision, encoder optimization, intra prediction, early CU partitioning, RD cost estimation.

Manuscript received Apr. 12, 2012; revised Sept. 13, 2012; accepted Oct. 12, 2012.

This research was supported by the Korea Communications Commission (KCC), Rep. of Korea, under the ETRI R&D support program supervised by the Korea Communications Agency (KCA) (KCA-2011-11921-02001).

Younhee Kim (phone: +82 42 860 5407, kimyounhee@etri.re.kr), DongSan Jun (dschun@etri.re.kr), Soon-heung Jung (zeroone@etri.re.kr), Jin Soo Choi (jschoi@etri.re.kr), and Jinwoong Kim (jwkim@etri.re.kr) are with the Broadcasting & Telecommunications Media Research Laboratory, ETRI, Daejeon, Rep. of Korea.

http://dx.doi.org/10.4218/etrij.13.0112.0223

I. Introduction

As video resolutions increase and the available bandwidth to transfer high-resolution video remains limited, a new video compression standard with a high coding performance is desirable. ISO-IEC/MPEG and ITU-T/VCEG recently formed the Joint Collaborative Team on Video Coding (JCT-VC) [1], aiming to develop the next-generation video coding standard, High Efficiency Video Coding (HEVC). As HEVC focuses on achieving a high coding efficiency [2], [3], its computational complexity dramatically increases. For use in practical applications such as high-resolution video services, HEVC requires a significant complexity reduction while maintaining a high coding performance.

For the previous coding standard, AVC/H.264, several methods [4]-[6] have been proposed to reduce the encoding complexity, which are classified into two approaches. The first approach uses fast intra-prediction methods [7]-[9]. For intra prediction, mainly fast intra-mode decision methods and fast block-size decision methods [10]-[15] have been proposed. The second approach is the use of fast inter-prediction methods [16]-[19]. For inter prediction, many fast motion search methods have been proposed, as motion estimation is the most time-consuming procedure in inter prediction.

The proposed method is relevant to the first category, a fast intra-prediction method using a fast mode decision and fast coding unit (CU) size decision. For a fast intra-mode decision, we propose a method to reduce the number of candidates. There have been several methods to reduce the number of candidates for AVC/H.264 intra prediction. An edge detection technique was employed in [15] to determine the edge direction, and, based on the direction, the number of candidate modes for the best mode was limited. Other approaches for limiting the number of candidates were based on filters [20], [21], a directional mask [22], the intensity gradient [23], and statistical properties [24]. Although these approaches reduce the encoding complexity, such preprocessing as detecting the edges or classifying their directional patterns requires an additional computational burden on the encoder. Our method limits the number of candidates without additional preprocessing by using rate-distortion cost estimation based on the Hadamard transform during the encoding processes.

The remainder of this paper is organized as follows. The HEVC test model (HM) encoding processes [25] are briefly described in section II, while the proposed method is presented in section III. Finally, experiment results and some concluding remarks are given in sections IV and V, respectively.

II. Overview of HM Encoding Processes

1. Coding Structure

HEVC has adopted three kinds of tree-structured unit representations: a CU, prediction unit (PU), and transform unit (TU). CU is the basic unit of region splitting for intra and inter predictions, PU is the basic unit of the prediction processes, and TU is the basic unit of the transform and quantization processes [25]. For each CU, the rate-distortion (RD) costs for the possible prediction mode (*Intra*, *Skip*, *Inter*) are calculated, and the mode having the minimum RD cost is selected as the best prediction mode at the current CU size. If the current CU is larger than the smallest CU (for example, 8 × 8), it is split into four equally-sized sub-CUs, as shown in Fig. 1, and an RD cost comparison between the current CU and the sum of the four sub-CUs is then conducted to determine the CU size and prediction mode. In other words, the RD costs are recursively obtained for each quadratic part, and the mode with the minimum cost is selected. For intra prediction, the CU size is the same as the PU size, except for the smallest CU, which has four additional PUs that are sized equally.

2. Intra Prediction

In HM 2.1[1)] intra prediction, 34 modes (33 angular modes and 1 direct current [DC] mode), as shown in Fig. 2 [26], are considered when selecting the best mode. Although rate-distortion optimized mode selection improves the coding efficiency, it requires heavy computational complexity. HM has already adopted a fast intra-mode decision scheme [27] in which some candidates are selected for the computational expensive rate-distortion optimization (RDO) process, and the RD cost is calculated in two ways: simplified RD cost calculation and full RD cost calculation. In the simplified RD cost calculation, the RD cost is calculated with only one fixed-sized TU, where the TU size is equal to the current PU size. The best mode is decided based on the simplified RD cost. On the other hand, in the full RD cost calculation, the RD cost is calculated using various-sized TUs, from 4 × 4 to the current PU size. The CU size is decided based on the full RD cost.

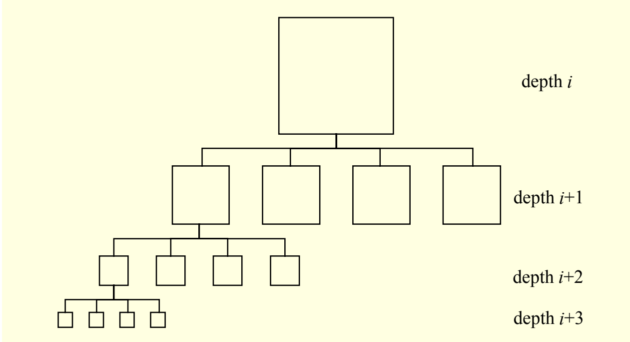

Fig. 1. Example of CU encoding structure (maximum partition depth = 4).

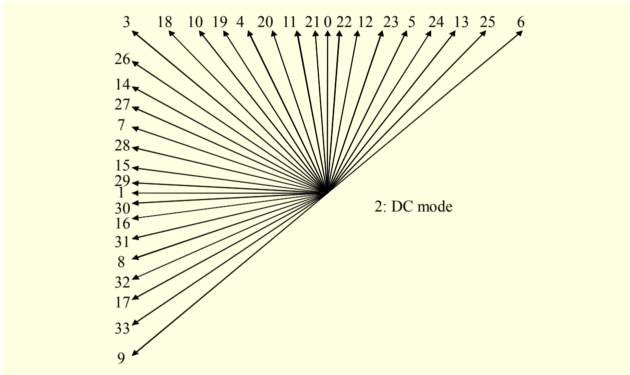

Fig. 2. Intra-prediction modes in HM.

The detailed HM intra-mode decision processes follow three steps. First, each cost based on the sum of the absolute transformed differences (SATD) between the current PU and the predictions using a Hadamard transform for all predefined modes is calculated, and *N* candidates along the ascending order of costs (denoted as HAD cost) are selected. The initial *N* is assigned differently according to the PU size: 3, 3, 3, 8, and 8 for 64 × 64, 32 × 32, 16 × 16, 8 × 8, and 4 × 4 PU sizes, respectively, as proposed in [27]. In addition to *N*, the most probable mode (MPM) [28] is added to the *N* candidates if it is not already included, and *N* is updated to *N*+1.

Second, the simplified RD costs for the selected *N* candidates are computed, and the mode having the minimum

1) Since our method is designed based on HM 2.1, HM specifically refers to HM 2.1 in this paper.

Table 1. Time spent on HM intra prediction (seconds/frame).

Q_p	N candidate selection	Simplified RD	Full RD	Luma intra prediction
22	4.27	16.98	7.56	29.07
27	4.63	13.41	5.81	24.28
32	4.13	11.05	4.91	20.43
37	4.31	9.94	3.86	18.24

simplified RD cost is selected. During this process, the RD cost is calculated using only one TU size, as previously described. In the second step, the simplified RD cost calculation module is called as many as N times.

Finally, the full RD cost with recursive TU partitioning is calculated using the mode selected from the second step. In the third step, the full RD cost calculation module is called only once for each PU.

Although the HM encoder already has a fast intra-mode decision scheme, as in the previously described three-step process, intra prediction in HM still takes a long time (for example, 23 seconds per frame for a 2,560 × 1,600 resolution sequence), as shown in Table 1, and HM intra prediction further requires a reduction in complexity.

3. Complexity Analysis of Intra prediction in HM

The profiling of HM intra prediction is conducted using the standard C clock function to find the most time-consuming module. The analysis is conducted using the *all intra high efficiency* configuration defined in the HEVC configurations [29]. The percentage of the total encoding time that the major intra-prediction modules spend encoding the People on Street (2,560 × 1,600) sequence is shown in Fig. 3, and its actual execution time is shown in Table 1. Since the results with other sequences are similar, they are not reported. As mentioned in the previous section, there are three major modules in an HM intra prediction: N candidate selection based on HAD cost, simplified RD calculation, and full RD calculation. The percentage of time spent for each module is 19.4%, 55.56%, and 22.7%, respectively. The encoder spends the most time for the simplified RD calculation because this module is called as many as times as the number of candidates, N. Our method reduces the time consumed in the simplified RD cost calculation significantly, and its simulation results are presented in section IV.

III. Proposed Fast Intra-Prediction Method

In the HM intra-mode decision, the best mode is selected

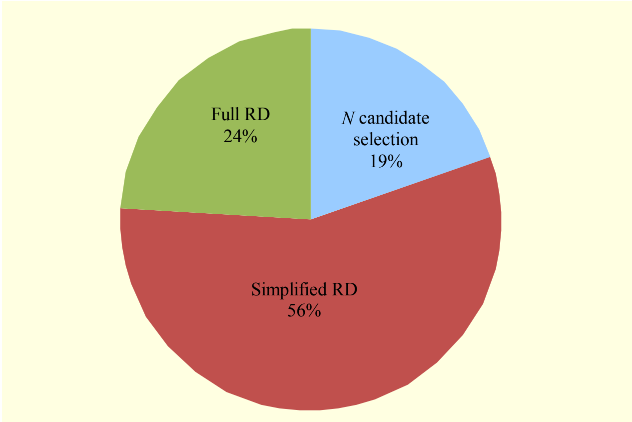

Fig. 3. Complexity analysis of three major modules of intra prediction in HM.

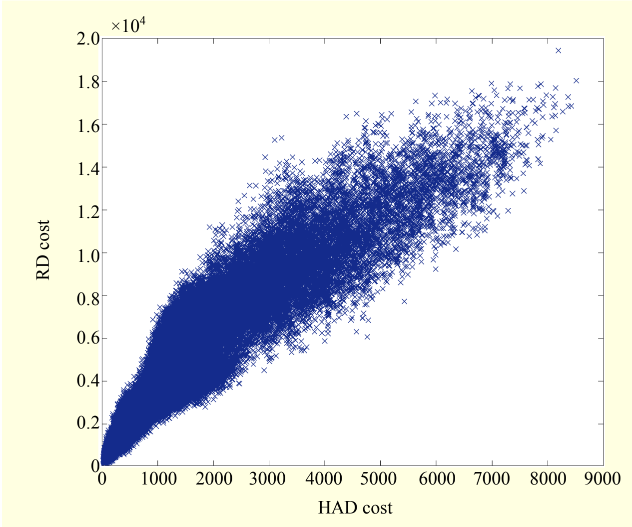

Fig. 4. Correlation between HAD and RD costs.

based on the competing RD cost, which is calculated by

$$J_m = D_m + \lambda_m(Q_p)R_m \,, \tag{1}$$

where D_m is the distortion between the original block and the reconstructed block from mode m, R_m is the estimated bits used to encode the residual with the given mode m, Q_p is a quantization parameter, and $\lambda_m(Q_p)$ is a Lagrangian multiplier given by

$$\lambda_m(Q_p) = 0.85 \times 2^{(Q_p-12)/3}. \tag{2}$$

1. Fast Mode Decision and Related Observations

We find that the HAD cost, which can be obtained using Hadamard-based SATD (distortion) and mode-bit (rate) estimation, has a high correlation with the RD cost, as shown in Fig. 4. The average correlation coefficient between the HAD

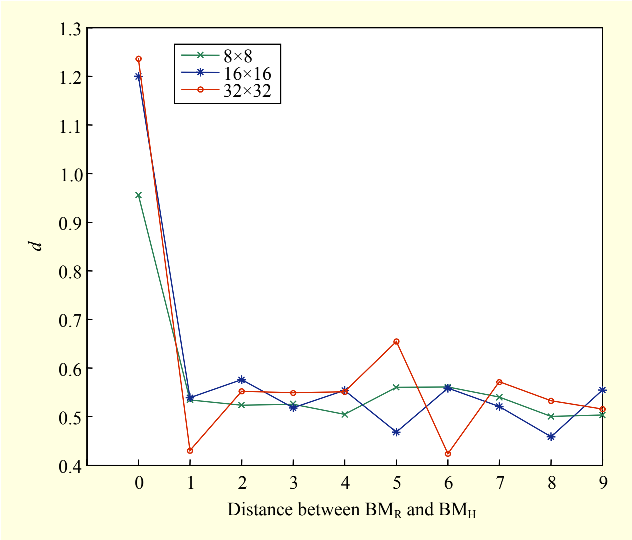

Fig. 5. Relationship between d and distance between BM$_H$ and BM$_R$. Zero distance indicates BM$_H$ hitting BM$_R$.

cost and RD cost is 0.93.

Let the mode having the minimum HAD cost be BM$_H$ and the mode having the minimum RD cost be BM$_R$. BM$_H$ may or may not be the same as BM$_R$. The probability of BM$_H$ being the same as BM$_R$ is 58% in our simulation. If we can distinguish the case of BM$_H$ hitting BM$_R$, we are able to skip the second step of the intra-mode decision, which is computing N simplified RD costs for these cases.

We also find that the PU block having a large difference in value between the minimum HAD cost and second minimum HAD cost is more likely to have BM$_H$ hitting BM$_R$. Let d be the difference between the minimum HAD cost ($C_{\text{HAD}[1]}$) and the second minimum HAD cost ($C_{\text{HAD}[2]}$), which is normalized by the PU block size as

$$d = \frac{C_{\text{HAD}[2]} - C_{\text{HAD}[1]}}{W_{\text{PU}} \times H_{\text{PU}}}, \tag{3}$$

where W_{PU} and H_{PU} are the width and height of PU, respectively.

Figure 5 shows the relationship between d and the distance between BM$_H$ and BM$_R$. We measure BM$_H$ hitting BM$_R$ based on the index distance in the candidate list. For example, the distance between $M_{\text{HAD}[i]}$ and $M_{\text{HAD}[j]}$ is $|i-j|$, where $M_{\text{HAD}[i]}$ represents the i-th mode in the candidate list sorted by the HAD cost. The zero distance of BM$_H$ from BM$_R$ in the candidate list indicates the case of BM$_H$ hitting BM$_R$. As shown in Fig. 5, the case of BM$_H$ hitting BM$_R$ has a high value of d.

To investigate this relationship further, we group PU into hitting and non-hitting groups depending on whether BM$_H$ hits BM$_R$. Table 2 presents the average value of d for the two groups and shows that the hitting group has a larger value of d than the non-hitting group.

Based on the above observations, we propose a fast mode

Table 2. Average value of d for hitting and non-hitting groups.

Group \ Size	4 × 4	8 × 8	16 × 16	32 × 32	64 × 64	Total
Hitting	1.44	0.83	1.01	1.02	1.93	1.34
Non-hitting	0.67	0.34	0.37	0.36	1.79	0.57

decision method based on d. In the proposed method, BM$_H$ is selected as the best mode m^* for the current PU without N candidate competition, and the process for N simplified RD cost calculations is hence skipped when $d \geq \delta$.

2. Computation Control Using δ

Performance and computation have a trade-off relationship, and computation control can be implemented using δ in our method. Simply, a small δ allows a fast mode decision whereas a large δ allows a small performance degradation. The performance degradation comes from an incorrect mode selection, in which the RD cost of BM$_H$ is larger than that of BM$_R$. In this section, we present the performance-computation model with respect to δ, which will help users determine the value of δ within their allowed level of performance degradation. We generate the model using statistics collected from one training sequence.

The performance degradation is estimated as

$$D_{\text{rd}}(\delta) = P_{\text{w}}(\delta) \times (RD_{\text{BM}_\text{H}} - RD_{\text{BM}_\text{R}}), \tag{4}$$

where D_{rd} is the expected RD cost penalty from an incorrect mode selection and P_{w} is the probability that PU has a larger d than δ but belongs to the non-hitting group. The RD cost for BM$_H$ is RD_{BM_H}, and the RD cost for BM$_R$ is RD_{BM_R}.

The expected time reduction can be expressed as

$$T_{\text{r}}(\delta) = P_{\text{g}}(\delta) \times T_{\text{srd}} \times (N-1), \tag{5}$$

where T_{r} is the time reduction, P_{g} is the probability that PU has a d larger than or equal to δ and belongs to the hitting group, and T_{srd} is the time required to calculate the simplified RD cost once. In addition, P_{w} and P_{g} are obtained from the training sequence with a given δ. The combined cost considering the distortion and time reduction can be expressed as

$$C(\delta) = \omega_1 T_r(\delta) + \omega_2 D_{rd}(\delta), \tag{6}$$

where ω_1 and ω_2 are the weighting factors.

Since T_{r} and D_{rd} are related to P_{w} and P_{g}, respectively, C is then related to both P_{w} and P_{g}. The statistically obtained values of T_{r} and D_{rd} for various δ show a trade-off model between the degradation and time reduction, as illustrated in Fig. 6.

To test our performance-computation trade-off model, we

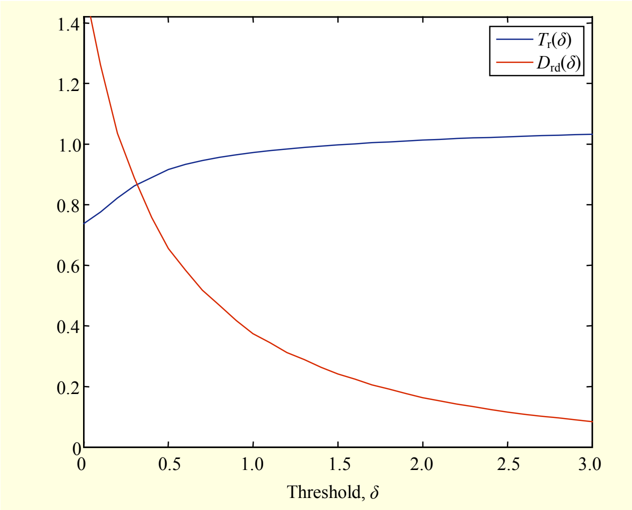

Fig. 6. Relationship between T_r and D_{rd} as function of δ.

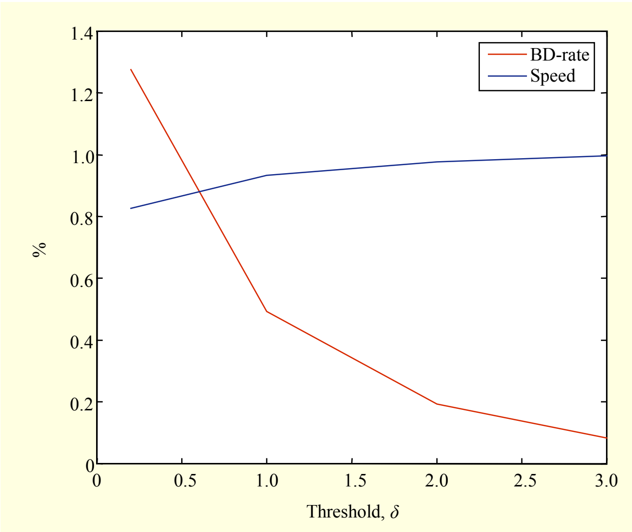

Fig. 7. Relationship between experimental BD-bitrate and encoding speed as function of δ.

encode four test sequences, which do not include the sequence used to generate the model, and measure the Bjøntegaard delta (BD) bitrate [30] and time reduction (with the HM mode decision as a reference, previously described in section II) for various values of δ, as shown in Table 3. The trade-off model between time saving and degradation for the given δ is constructed using statistics from the training sequence, and the model fits the experiment results fairly well, as shown in Figs. 6 and 7. The model is useful to estimate the time savings and distortion (BD-bitrate) with the given δ.

3. DC and MPM Modes

As previously described, in our fast mode decision, BM_H is always selected as the best mode m^* when $d \geq \delta$. When $d < \delta$,

Table 3. Bitrate and time reduction results with various values of δ.

Test sequence	δ	BD-bitrate Y (%)	BD-bitrate U (%)	BD-bitrate V (%)	Time savings (%)
Traffic (2,560×1,600)	0.2	1.5	0.3	0.5	–14.88
	1	0.5	0.1	0.1	–5.92
	2	0.2	–0.1	0.0	–2.04
	3	0.1	–0.1	0.0	–0.52
Kimono (1,920×1,080)	0.2	0.3	0.1	0.1	–17.69
	1	0.1	0.0	0.1	–5.20
	2	0.0	0.1	0.0	–1.49
	3	0.0	0.1	–0.1	–0.93
Basketball Drill (832×480)	0.2	1.2	0.0	0.0	–18.76
	1	0.4	–0.2	0.1	–7.98
	2	0.0	0.1	0.0	–3.73
	3	0.1	–0.1	–0.1	–0.47
Race Horses (832×480)	0.2	1.9	0.1	0.2	–18.84
	1	0.9	–0.1	0.0	–8.50
	2	0.4	0.2	–0.1	–3.46
	3	0.2	0.0	–0.1	–0.99

our method constructs the candidate list using only the DC mode, MPM mode, and BM_H. The best mode is selected from at most three candidates using a simplified RD cost competition, and the full RD cost is calculated using the best mode selected. The HM encoder always includes the MPM mode after selecting N candidates based on the HAD cost. Therefore, in HM, the number of candidates is at most N+1, while the number is three in our proposed fast mode decision method.

We forcibly include the DC mode as a candidate, based on our observation that the largest RD cost penalty occurs when BM_H is in DC mode. We compute the RD cost penalties as $RD_{BM_H} - RD_{BM_R}$ for each mode experimentally and find that DC mode (mode number 2 in Fig. 8) produces the largest penalty. MPM mode is also included because the best mode of the current block has a high correlation with its neighboring blocks. Hence, we propose to always include the DC mode, MPM mode, and BM_H as candidates for simplified RD cost calculation when $d < \delta$. The proposed fast intra-mode decision method always includes DC and MPM modes in addition to BM_H for the simplified RD cost calculation process if $d < \delta$.

4. Fast CU Size Decision

HM encodes a CU by splitting it into four sub-CUs

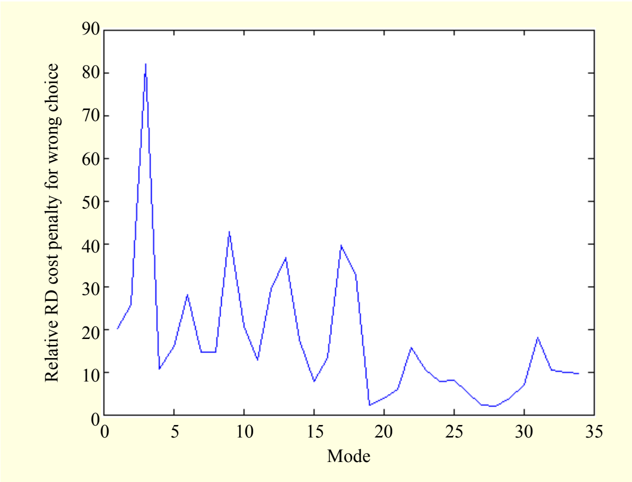

Fig. 8. Relative RD cost penalty for incorrectly selected mode.

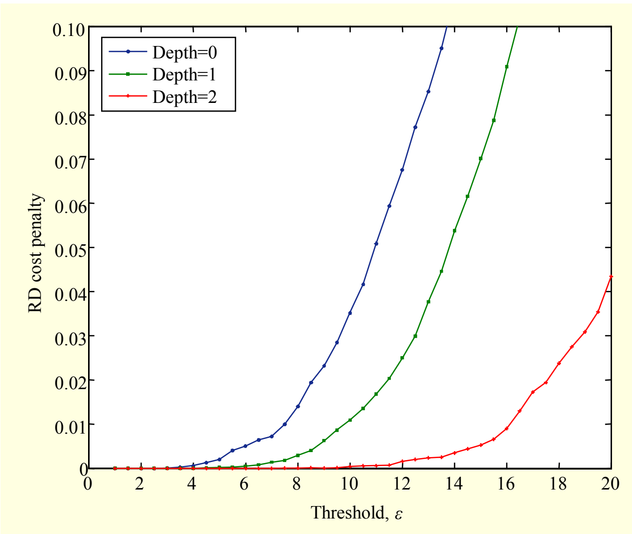

Fig. 9. RD cost penalty by fast CU-size decision with respect to ε.

recursively and finding the best mode in terms of RD cost. However, a recursive comparison results in a high encoding complexity. We observe that a CU with a small RD cost is more likely to be encoded without further partitioning or encoded with split sub-CUs for a very small gain. Based on this observation, we propose skipping further CU splitting if the full RD cost for the best mode is less than threshold ε. To determine the threshold that rarely affects the rate-distortion performance by the proposed fast CU size decision, we calculate the expected RD cost penalties according to CU size as the threshold changes, and the result is shown in Fig. 9. The experiment result shows that the normalized RD cost penalty of less than 0.001 is maintained as follows: set the threshold as 3.0 for CUs in depth 0 (64 × 64), 4.0 for CUs in depth 1 (32 × 32), and 8.0 for CUs in depth 2 (16 × 16), respectively. However, the hit rate for CUs in depth 0 with the threshold is

Table 4. Bitrate and time reduction results with respect to ε.

Test sequence	ε (depth1, depth2)	BD-bitrate Y (%)	BD-bitrate U (%)	BD-bitrate V (%)	Time savings (%)
2 Class A and 5 Class B	(3, 7)	1.2	0.0	0.1	–28.96
	(4, 8)	1.2	0.0	0.1	–30.55
	(5, 9)	1.3	0.0	0.2	–31.98
	(6, 10)	1.4	0.1	0.3	–33.51
	(7, 11)	1.5	0.2	0.5	–35.06
	(8, 12)	1.6	0.4	0.6	–36.49
	(9, 13)	1.7	0.6	0.8	–37.94

67% whereas the hit rate for the CUs in both depth 1 and depth 2 is 99%. We therefore apply the fast CU decision scheme only to CUs in depth 1 and depth 2.

Let c be the normalized full RD cost for the selected best mode m^*:

$$c = \frac{C_R^*}{W_{CU} \times H_{CU}}, \quad (7)$$

where C_R^* is the full RD cost with the chosen best mode m^*, and W_{CU} and H_{CU} are the width and height of CU, respectively. In the proposed method, further CU splitting is skipped if $c \le \varepsilon$, and ε is set as 4.0 for depth 1 CUs and 8.0 for depth 2 CUs.

We test the fast CU size decision scheme with seven test sequences (Traffic, People on Street, Kimono, Park Scene, Cactus, Basketball Drive, and BQ Terrace), and the result of the BD-bitrate and time saving with respect to ε is shown in Table 4.

5. Summary of Proposed Fast Intra-prediction Method

We propose a fast intra-prediction method using a fast mode decision and fast CU-size decision. The flow chart of the proposed fast method is presented in Fig. 10.

For a fast mode decision, BM_H is selected only based on HAD cost comparisons among all intra modes. If d is greater than or equal to δ, the best mode m^* for the current PU is chosen as BM_H. Otherwise, the simplified RD costs of BM_H, DC, and MPM modes are compared. If there is a redundant mode, the mode is removed from the RD cost calculation. The mode with the minimum RD cost is selected as m^*. The full RD cost (C_R^*) with m^* mode is then calculated by considering the quad-tree TU structure.

The second proposal is a fast CU-size decision. The full RD cost with the given best mode c is less than threshold ε, and

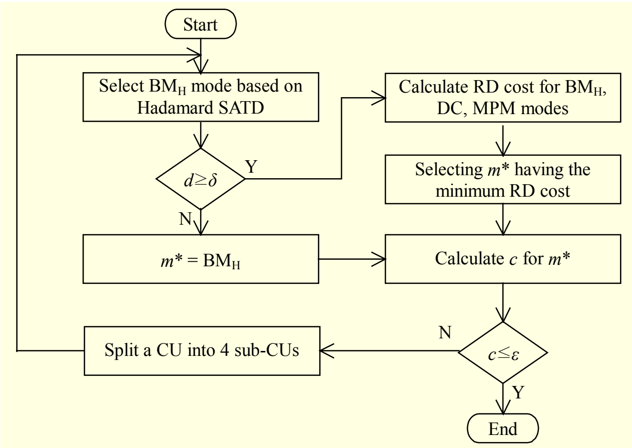

Fig. 10. Flow chart of proposed fast intra-prediction method.

Table 5. Intra-prediction time reduction (%) of proposed method.

δ	N candidate selection	Simplified RD	Full RD	Luma intra prediction
0.2	–10.37	–92.54	–7.45	–54.78
0.5	–10.06	–85.09	–6.78	–50.73
1	–8.46	–81.36	–8.29	–48.41
2	–10.36	–75.41	–13.02	–46.41
3	–4.01	–74.38	–13.95	–44.85
Avg.	–8.65	–81.76	–9.90	–49.04

further competition with split CUs to find the optimal CU size is skipped.

IV. Experiment Results

The proposed method has been implemented based on the HM 2.1 encoder, which is used as an anchor in the experiments. Our experiments are performed on the *all intra high-efficiency encoding* configuration. The conditions used in the experiments are as follows.

- System: 3.47GHz CPU with 24G RAM
- Q_P: 22, 27, 32, 37
- RDOQ: on
- ALF: on
- Entropy coding: CABAC
- Largest CU size: 64×64
- Maximum partitioning depth: 4
- Test sequences: 2 Class A (Traffic, People on Street) and 5 Class B (Kimono, Park Scene, Cactus, Basketball Drive, BQ Terrace) sequences.

Owing to the trade-off relationship between the increased speed and rate-distortion performance degradation, we evaluate the time savings of our method with respect to δ from 0.2 to 3.

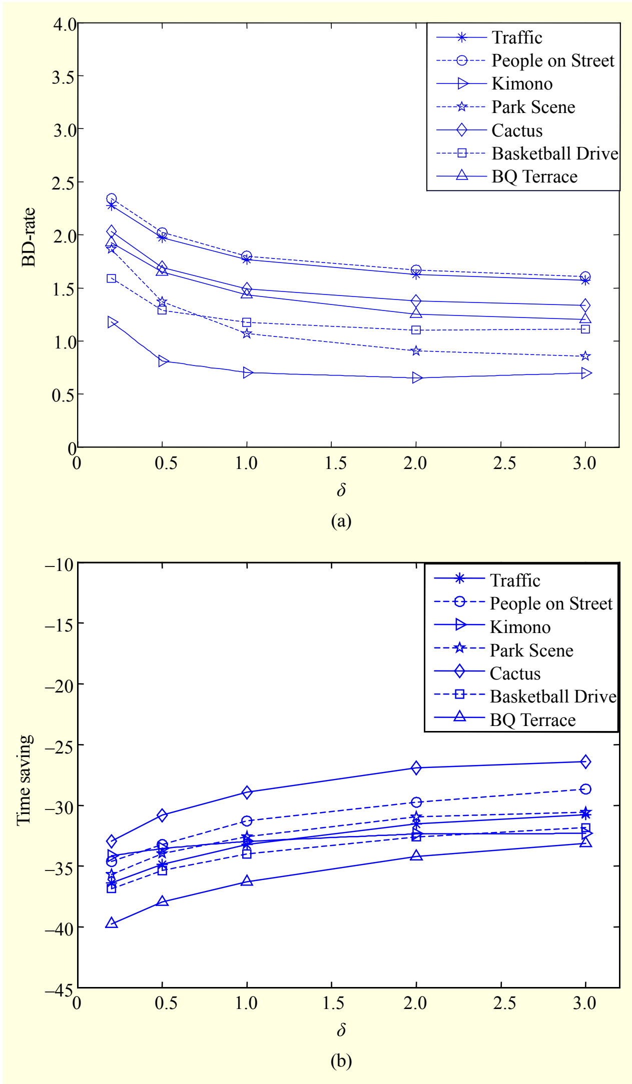

Fig. 11. Relationship between (a) BD-bitrate and δ and between (b) total encoding time savings and δ.

A faster mode decision is expected with a smaller δ, while a smaller degradation is expected with a larger δ. The results for the time reduction for the major intra-prediction processes are presented in Table 5.

As described in section II, the three major modules for HM luma intra mode decision are N candidate selections based on HAD cost, simplified RD calculation, and full RD calculation. Among the three modules, HM consumes the most time on the simplified RD calculation module, and the time is dependent on the number of N. In the proposed fast mode decision method, the average number of N is significantly reduced from 7.83 to 1.58, which results in an overall encoding time reduction. As shown in Table 5, the proposed method reduces the 81.76% time spent on the simplified RD calculation and, as a result, achieves an average of 49.04% time reduction in the luma intra-prediction process as compared to HM.

Table 6. Simulation results ($\delta = 3,\ \varepsilon = (4, 8)$).

Sequence	Q_P	ΔBR (%)	ΔPSNR (dB)	ΔTime (%)	BD rate Y	BD rate U	BD rate V
Traffic (2,560 × 1,600)	22	0.7882%	–0.1471	–42.13%	1.7	0.4	0.5
	27	–0.0123%	–0.0834	–30.92%			
	32	–0.0937%	–0.0765	–25.87%			
	37	0.0226%	–0.0615	–22.45%			
People on Street (2,560 × 1,600)	22	0.5475%	–0.1289	–42.58%	1.6	0.6	0.5
	27	0.0361%	–0.0851	–28.27%			
	32	–0.0243%	–0.0807	–24.10%			
	37	0.2487%	–0.0714	–21.40%			
Kimono (1,920 × 1,080)	22	0.2527%	–0.0203	–61.11%	0.7	0.3	0.3
	27	0.0947%	–0.0152	–35.92%			
	32	0.0841%	–0.0205	–24.77%			
	37	–0.0477%	–0.0252	–22.47%			
Park Scene (1,920 × 1,080)	22	–0.8451%	–0.1175	–37.23%	0.9	–0.3	–0.3
	27	–0.7685%	–0.0826	–30.42%			
	32	–1.0662%	–0.0691	–27.37%			
	37	–0.6734%	–0.0371	–25.74%			
Cactus (1,920 × 1,080)	22	–1.2015%	–0.1098	–32.93%	1.3	–0.3	–0.1
	27	–0.4398%	–0.0598	–26.87%			
	32	–0.2666%	–0.0582	–23.73%			
	37	0.0860%	–0.0461	–20.95%			
Basketball Drive (1,920 × 1,080)	22	–0.9975%	–0.0579	–47.59%	1.1	0.0	0.2
	27	0.1356%	–0.0183	–31.36%			
	32	0.2597%	–0.0252	–23.85%			
	37	0.2486%	–0.0335	–20.63%			
BQ Terrace (1,920 × 1,080)	22	–0.1114%	–0.1187	–38.26%	1.2	–0.4	–0.4
	27	–0.2928%	–0.0859	–34.11%			
	32	–0.5219%	–0.0926	–27.81%			
	37	–0.2063%	–0.0727	–24.45%			
Average		–0.1701%	–0.0679	–30.55%	1.2	0.0	0.1

We also evaluate the performance of our method in terms of the change in average bitrate, peak signal-to-noise ratio (PSNR), and total encoding time, which are reported based on the following formula:

$$\Delta BR\ (\%) = \frac{\text{Bitrate (proposed)} - \text{Bitrate (anchor)}}{\text{Bitrate (anchor)}} \times 100, \quad (7)$$

$$\Delta PSNR\ (\text{dB}) = Ypsnr(\text{proposed}) - Ypsnr(\text{anchor}), \quad (8)$$

$$\Delta \text{Time}\ (\%) = \frac{\text{Time (proposed)} - \text{Time (anchor)}}{\text{Time (anchor)}} \times 100. \quad (9)$$

The simulation results with δ=3.0 are shown in Table 6. Note that with the proposed method, the PSNR decrease is 0.0679 but that the bitrate decrease is 0.17%, which infers a negligible rate-distortion performance degradation of the proposed method. The rate-distortion performance is also presented using the BD-bitrate [30]. The relationship between the BD-bitrate and δ and the relationship between time saving and δ are shown in Fig. 11. In our fast encoding setting (that is, δ=0.2), a 35.46% savings in the encoding time is achieved with a BD-bitrate increase of 1.9. Based on the setting of our small performance degradation (that is, δ=3.0), a 30.55% savings in the encoding time is achieved with a BD-bitrate increase of 1.2.

We conduct an additional experiment for less performance

Table 7. Simulation results with no-simplified RD ($\delta = 3,\ \varepsilon = (4, 8)$).

Sequence	Q_P	ΔBR (%)	ΔPSNR (dB)	ΔTime (%)	BD rate Y	BD rate U	BD rate V
Traffic (2,560 × 1,600)	22	–0.3301%	–0.1058	–24.16%	1.2	0.1	0.4
	27	–0.1304%	–0.0741	–23.69%			
	32	–0.2799%	–0.0713	–23.82%			
	37	–0.1452%	–0.0506	–22.75%			
People on Street (2,560 × 1,600)	22	–0.0803%	–0.0987	–24.76%	1.3	0.4	0.4
	27	–0.0506%	–0.0797	–22.54%			
	32	–0.1931%	–0.0756	–21.52%			
	37	0.1055%	–0.0581	–19.87%			
Kimono (1,920 × 1,080)	22	–0.0343%	–0.0102	–27.61%	0.5	0.3	0.3
	27	0.0590%	–0.0140	–26.32%			
	32	0.0570%	–0.0207	–25.16%			
	37	–0.0123%	–0.0205	–24.62%			
Park Scene (1,920 × 1,080)	22	–0.8916%	–0.1034	–26.90%	0.5	–0.4	–0.3
	27	–0.8744%	–0.0762	–26.55%			
	32	–1.1520%	–0.0583	–25.20%			
	37	–0.7872%	–0.0245	–25.91%			
Cactus (1,920 × 1,080)	22	–1.0991%	–0.0966	–24.20%	1.1	–0.4	0.0
	27	–0.5383%	–0.0556	–21.55%			
	32	–0.4078%	–0.0527	–20.29%			
	37	–0.0827%	–0.0409	–19.03%			
Basketball Drive (1,920 × 1,080)	22	–0.8036%	–0.0410	–28.03%	0.9	0.1	0.2
	27	0.0420%	–0.0176	–26.14%			
	32	0.1401%	–0.0231	–23.31%			
	37	0.1316%	–0.0307	–22.02%			
BQ Terrace (1,920 × 1,080)	22	–0.2489%	–0.0997	–30.04%	0.9	–0.4	–0.6
	27	–0.4395%	–0.0834	–27.06%			
	32	–0.6748%	–0.0867	–26.21%			
	37	–0.3868%	–0.0651	–24.53%			
Average		–0.3253%	–0.0584	–24.42%	0.9	0.0	0.0

degradation, denoted by *no-simplified RD*, in which the best mode is selected only through full RD cost competition. While in HM and our proposed method the best mode is selected using the simplified RD calculation competition and the cost is updated with the full RD cost for determining the CU partitioning, in the no-simplified RD experiment, the N candidates are selected using the proposed fast mode decision scheme, and the best mode is selected based on the full RD cost. The simulation result with no-simplified RD is shown in Table 7. Compared to HM, a 24.42% time savings in the encoding time is achieved with a BD-bitrate increase of 0.9, which results in smaller performance degradation but less encoding time reduction, compared to our proposed method simulation shown in Table 6.

V. Conclusion

We proposed a fast intra-mode decision method based on the difference between the minimum and second minimum Hadamard SATD-based RD cost estimation and a fast CU-size decision method based on the RD cost of the best mode. In contrast with the previous fast intra-prediction methods, our method does not require such additional preprocessing as edge detection or masking; rather, using the fast and simple computation of the Hadamard-based SATD, the mode that yields the cost most similar to that of the RD optimized best mode is selected. Our method also chooses an early termination of CU partitioning using the RD cost computed during the encoding process. The simulation results show that our proposed method achieves an average time reduction of 49.04% in luma intra prediction and an average time reduction of 32.74% in total encoding with a small drop in rate distortion compared to the HM 2.1 encoder.

References

[1] ITU-T Q6/16 and ISO/IEC JTC1/SC29/WG11, "Joint Call for Proposals on Video Compression Technology," ITU-T SG16/Q6/VCEG-AM91, *39th VCEG*, Kyoto, Japan, Jan. 2010.

[2] S.Y. Jeong et al., "Highly Efficient Video Codec for Entertainment-Quality," *ETRI J.*, vol. 33, no. 2, Apr. 2011, pp. 145-154.

[3] H.H. Lee et al., "Enhanced Block-Based Adaptive Loop Filter with Multiple Symmetric Structures for Video Coding," *ETRI J.*, vol. 32, no. 4, Aug. 2010, pp. 626-629.

[4] J. Xin, A. Vetro, and H. Sun, "Efficient Macroblock Coding-Mode Decision for H.264/AVC Video Coding," *Proc. 24th Picture Coding Symp.*, Dec. 2004.

[5] J.C. Wang et al., "A Fast Mode Decision Algorithm and Its VLSI Design for H.264/AVC Intra-Prediction," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 17, no. 10, Oct. 2007, pp. 1414-1422.

[6] C.C. Lien and C.P. Yu, "A Fast Mode Decision Method for H.264/AVC Using the Spatial-Temporal Prediction Scheme," *Proc. 18th IEEE Int. Conf. Pattern Recog.*, vol. 4, Aug. 2006, pp. 334-337.

[7] C.C. Cheng and T.S. Chang, "Fast Three Step Intra Prediction Algorithm for 4×4 Blocks in H.264," *Proc. IEEE Int. Symp. Circuits Syst.*, vol. 2, May 2005, pp. 1509-1512.

[8] B. Meng and O.C. Au, "Fast Intra-Prediction Mode Selection for 4×4 Blocks in H.264," *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Process.*, vol. 3, Apr. 2003, pp. 389-392.

[9] Y.D. Zhang, F. Dai, and S.X. Lin, "Fast 4×4 Intra-prediction Mode Selection for H.264," *Proc. IEEE Int. Conf. Multimedia Expo*, vol. 2, 2004, pp. 1151-1154.

[10] J.F. Wang et al., "A Novel Fast Algorithm for Intra Mode Decision in H.264/AVC Encoders," *Proc. IEEE Int. Symp. Circuits Syst.*, May 2006, pp. 3498-3501.

[11] T. Tsukuba et al., "H.264 Fast Intra-Prediction Mode Decision Based on Frequency Characteristic," *Proc. European Signal Process. Conf.*, 2005.

[12] T. Hattori and K. Ichige, "Intra Prediction Mode Decision in H.264/AVC Using DCT Coefficients," *Proc. Int. Symp. Intell. Signal Process. Commun. Syst.*, Dec. 2006, pp. 135-138.

[13] M.C. Hwang et al., "Fast Intra Prediction Mode Selection Scheme Using Temporal Correlation in H.264," *Proc. IEEE TENCON*, Region 10, Nov. 2005, pp. 1-5.

[14] A. Elyousfi, A. Tamtaoui, and H. Bouyakhf, "Fast Mode Decision Algorithm for Intra Prediction in H.264/AVC Video Coding," *Int. J. Computer Sci. Netw. Security*, vol. 7, no. 1, Jan. 2007, pp. 356-364.

[15] F. Pan et al., "Fast Mode Decision Algorithm for Intraprediction in H.264/AVC Video Coding," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 15, no. 7, July 2005, pp. 813-822.

[16] D. Jun and H. Park, "An Efficient Priority-Based Reference Frame Selection Method for Fast Motion Estimation in H.264/AVC," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 20, no. 8, Aug. 2010, pp. 1156-1161.

[17] J.H. Kim et al., "An SAD-Based Selective Bi-prediction Method for Fast Motion Estimation in High Efficiency Video Coding," *ETRI J.*, vol. 34, no. 5, Oct. 2012, pp. 753-758.

[18] B.G. Kim and J.H. Kim, "Efficient Intra-mode Decision Algorithm for Inter-frames in H.264/AVC Video Coding," *IET Image Process.*, vol. 5, no. 3, Apr. 2011, pp. 286-295.

[19] X. Lu et al., "Fast Mode Decision and Motion Estimation for H.264 with a Focus on MPEG-2/H.264 Transcoding," *Proc. IEEE Int. Symp. Circuits Syst.*, vol. 2, Kobe, Japan, May 2005, pp. 1246-1249.

[20] F. Pan et al., "A Directional Field Based Fast Intra Mode Decision Algorithm for H.264 Video Coding," *Proc. IEEE Inter. Conf. Multimedia Expo*, vol. 2, June 2004, pp. 1147-1150.

[21] B. Shen and I.K. Sethi, "Direct Feature Extraction from Compressed Images," *Proc. IS&T/SPIE Conf. Storage Retrieval Image Video Databases*, vol. 2670, Jan. 1996, pp. 404-414.

[22] J.H. Kim and J.C. Jeong, "Fast Intra-Mode Decision in H.264 Video Coding Using Simple Directional Masks," *Proc. SPIE*, vol. 5960, 2005, pp. 1071-1079.

[23] A.C. Tsai et al., "Intensity Gradient Technique for Efficient Intra-Prediction in H.264/AVC," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 18, no. 5, May 2008, pp. 694-698.

[24] R. Garg, M. Jindal, and M. Chauhan, "Statistics Based Fast Intra-Mode Detection," *Proc. SPIE*, vol. 5960, 2005, pp. 2085-2091.

[25] JCT-VC, "High Efficiency Video Coding (HEVC) Test Model 2 (HM 2) Encoder Description," Document JCTVC-D502, Jan. 2011.

[26] JCT-VC, "WD2: Working Draft 2 of High-Efficiency Video Coding," Document JCTVC-D503, Jan. 2011.

[27] L. Zhao et al., "Fast Mode Decision Algorithm for Intra Prediction in HEVC," *Proc. IEEE Visual Commun. Image Process.*, Nov. 2011, pp. 1-4.

[28] J.H. Lee et al., "Intra-Mixture Prediction Mode and Enhanced Most Probable Mode Estimation for Intra Coding," *ETRI J.*, vol. 31, no. 5, Oct. 2009, pp. 610-612.

[29] F. Bossen, "Common Test Conditions and Software Reference Configurations," Document JCTVC-F900, Torino, IT, July 2011.

[30] G. Bjontegaard, “Calculation of Average PSNR Differences between RD-Curves,” Document VCEG-M33, Austin, TX, USA, Apr. 2001.

Younhee Kim received her BS and MS in computer science from Ajou University, Suwon, Rep. of Korea, in 2000 and in 2002, respectively, and her PhD in computer science from George Mason University, Fairfax, VA, USA, in 2009. She has been a senior researcher with ETRI, Daejeon, Rep. of Korea, since 2009. She is currently involved in the development of an HEVC video real-time encoding system. Her current research interests include video coding, image and video signal processing, optimizing video encoding, and information hiding in the field of multimedia communication.

DongSan Jun received his BS in electrical engineering and computer science from Pusan National University, Busan, Rep. of Korea, in 2002 and his MS and PhD in electrical engineering from the Korea Advanced Institute of Science and Technology (KAIST), Daejeon, Rep. of Korea, in 2004 and 2011, respectively. He has been a member of the engineering staff at ETRI, Daejeon, Rep. of Korea, since 2004 and an adjunct professor in the Mobile Communication and Digital Broadcasting Engineering Department at the University of Science and Technology (UST), Daejeon, Rep. of Korea, since 2011. His research interests include image computing systems, pattern recognition, video compression, and realistic broadcasting systems.

Soon-heung Jung received his BS in electronics engineering in 2001 from Pusan National University, Busan, Rep. of Korea. He received his MS in electronics engineering in 2003 from the Korea Advanced Institute of Science and Technology (KAIST), Daejeon, Rep. of Korea. From 2003 to 2005, he was a research engineer at LG Electronics, Rep. of Korea. Since 2005, he has been a senior member of the engineering staff at ETRI, and he is also working toward his PhD in electronics engineering at KAIST. His research interests are in the area of visual communication, video signal processing, video coding, and realistic broadcasting systems.

Jin Soo Choi received his BE, ME, and PhD in electronics engineering from Kyungpook National University, Rep. of Korea, in 1990, 1992, and 1996, respectively. Since 1996, he has been a principal member of the research staff at ETRI, Daejeon, Rep. of Korea. He has been involved in developing the MPEG-4 codec system, data broadcasting systems, and the 3D/UHDTV broadcasting system. His research interests include visual signal processing and interactive services in the field of digital broadcasting technology.

Jinwoong Kim received his BS and MS in electronics engineering from Seoul National University, Seoul, Rep. of Korea, in 1981 and 1983, respectively. He received his PhD in electrical engineering from Texas A&M University, College Station, TX, USA, in 1993. He has been working in ETRI since 1983 and is now a principal member of the research staff and the director of the Realistic Broadcasting Media Research Department. He has led many government-funded R&D projects on digital broadcasting technologies, including data broadcasting, viewer-customized broadcasting, and MPEG-7/MPEG-21 standard technology development. Currently, his research interests include such 3DTV technologies as stereoscopic 3DTV, multiview 3DTV, and digital holography, as well as UHDTV. He was a chair of the 3DTV project group of TTA. He was a Far-East Liaison of the 3DTV Conference in 2007 and 2008 and has been an invited speaker at a number of domestic and international workshops, including the EU-Korea Cooperation Forum workshop on ICT, the 3D Fair 2008 in Japan, and the 3DTV Conference 2010.