# Fast Intra Coding Method of H.264 for Video Surveillance System

**Jang-Seon Ryu and Eung-Tea Kim,**

Korea Polytechnic University, Siheung-City, Kyonggi-Do, Korea

**Summary**
The H.264 standard employs several powerful coding methods to obtain high compression efficiency. Specially, to reduce the spatial redundancy more effectively, the directional intra prediction method is used. It investigates all available coding modes to decide the best one. This decision process requires very high computational time and a lot of memory access. So, it is hard to adopt it into the real-time video surveillance system, DVR(digital video recorder). To overcome this problem, we propose the efficient intra prediction mode decision method which can select the prediction block sizes between 4x4 and 16x16 sizes quickly, and then decide the directional modes. In the experimental result, the efficiency of the proposed method is shown by comparing its processing time and PSNR with those of the conventional methods. The proposed method reduces 40%~70% coding time while maintaining similar PSNR of H.264 codec.
***Key words:***
*H.264 , Spatial Encoding, Intra, Video Surveillance*

## 1. Introduction

The New Video Coding, H.264 Standard achieves much higher coding efficiency than previous video coding standards such as MPEG-1/2/4. The H.264 Video format gives perceptually equivalent video quality at significant lower bitrates than the MPEG-1/2/4 and H.263 standard by using a collection of new encoding tools. However, at a cost of a significant increase in encoding complexity so it has consumed a lot of processing time and accessing memory. [1-2]

Nevertheless, due to its high coding complexity, it is very difficult to implement a high performance real-time H.264 video coding. The spatial prediction coding method in the H.264 that had not been existed other video coding, was designed for reducing spatial redundancy by using neighboring pixel of current macroblock. It increased computation, because optimal mode was determined by Full search for all modes and choosing minimum error mode, that is, it searches intra 4x4 prediction that has 9 directional prediction modes, intra 16x16 that has 4 directional prediction modes for luminance and intra 8x8 that has 4 directional prediction modes for chrominance.

At recently, many algorithms have been proposed for improving a spatial prediction method in the H.264.[3-5] These researches contain a edge map scheme[3] which checks only selected several modes based on edge direction, EIP(Efficient Intra Prediction) algorithm based on early termination, selective computation of highly probable modes and partial computation of the cost function[4], and other fast intra-prediction mode that employs the simpler computation of the dominant edge direction within a block.[5]

However, these algorithms are decreasing complexity compared with full search, additional complex computations are required to decide mode decision and, in case of EIP, it is difficult to look for threshold to decide suitable condition.

To solve this problem, fast mode decision method was proposed in this paper which uses boundary pixels and its neighboring pixels in current macroblock.

First of all, the spatial intra prediction of H.264 was explained in section 2, and detailed algorithm of mode decision was described in section 3. In section 4, the experimental results were shown.

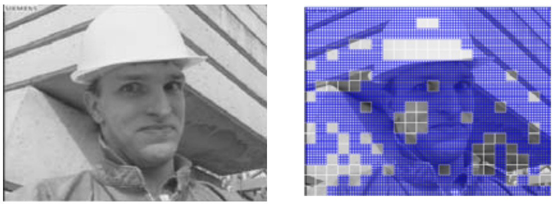
## 2. Spatial Intra Prediction of H.264

In spatial intra prediction of H.264, it has been improved by minimizing the predictive error signal through multiple directional prediction modes. Intra coding block type is highly dependent on the smoothness of the block. Intra 4x4 is well suited for a macroblock with the detailed information, while intra 16x16 is well suited for a smooth one. [6]

In the figure 1, each pane refers to a coded macroblock. The white pane is a macroblock coded as intra 16x16, while the blue one refers to intra 4x4. This shows the spatial correlation between the smoothness and the coded block type. Actually, H.264/AVC performs intra prediction by using the spatial correlation with adjacent blocks. That is, according to the given direction, the current block is predicted by using boundary pixels of upper and left blocks which are previously decoded. Then,

Manuscript received October 5, 2007
Manuscript revised October 20, 2007

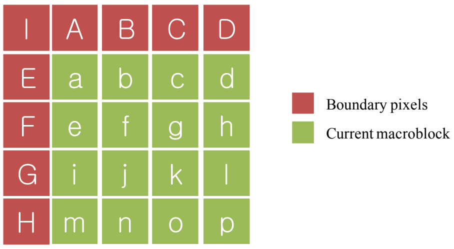
the direction which has the maximum correlation is selected. For the luminance components, intra prediction may be used for each 4x4 sub-block or 16x16 luminance blocks. For the chrominance components, there are 4 prediction modes that are applied to the two 8x8 chrominance blocks (U and V or Cb and Cr), which the resulting prediction mode for U and V components should be the same. The residual between the current macroblock and its prediction is then transformed, quantized, and entropy coded.



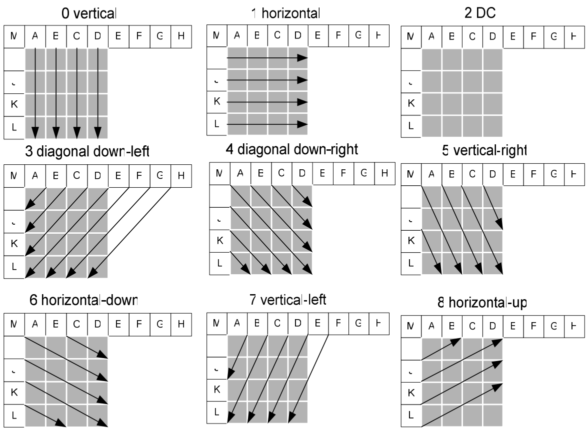
a)Original Image b) Selected Macro-Blocks

Fig 1 Block type map of the first frame of Foreman_CIF

## 2.1 .4x4 luminance prediction modes




a) Labeling prediction of samples 4x4




b) 4x4 Prediction Modes

Fig. 2 Intra Prediction for a 4x4 luminance block

Figure 2 illustrates the intra prediction for a 4x4 luminance block. In a) of this figure, a to p are the pixels to be predicted, and A to I are the neighboring pixels that are available at the time of prediction. When the intra 4x4 is used, each 4x4 block is predicted from spatially neighboring (above and left) samples. For each 4x4 block, 9 directional prediction modes are shown in the b) of figure. All the 9 modes are examined for a 4x4 block to find the optimal one with minimum cost, which is called intra 4x4 prediction routine.

## 2.2 16x16 luminance prediction

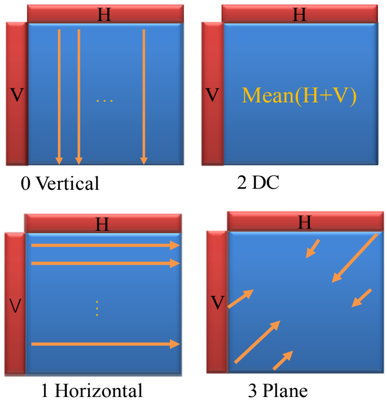



Fig. 1. Intra Prediction for a 16x16 luminance block

When using the intra 16x16 mode, the whole luminance component of a macroblock is predicted. There are four prediction modes. As shown in figure 3, Mode 0 (vertical prediction) mode 1 (horizontal prediction), mode 2 (DC prediction), and mode 3 (Plane) are specified.

## 3. Proposed Algorithm

For intra prediction in H.264, the pixel values of current block are predicted from the boundary pixels of previously decoded neighboring blocks. To reduce the computations of exhaustive search, F. Pan proposed the fast intra mode decision method, which is based on edge detection that uses sobel operation and edge direction histogram within the block. Nevertheless, the intra prediction in H.264 is carried out using not only pixels within a block but also boundary pixels of adjacent block. In addition, since Pan's method is required the preprocessing such as the edge detection and the edge direction histogram, the effect of fast mode decision is weak.

So, we proposed the efficient method which contains the fast block type decision by using the pixel correlation in the macroblock and the cost function for the mode decision.

### 3.1 Block type selection

The proposed algorithm for fast type decision is presented as follows in figure 4. Intra 16x16 block type is well suited for smooth block such as background. So, pixel value extracted from one block is similar or same to that in the predicted blocks.

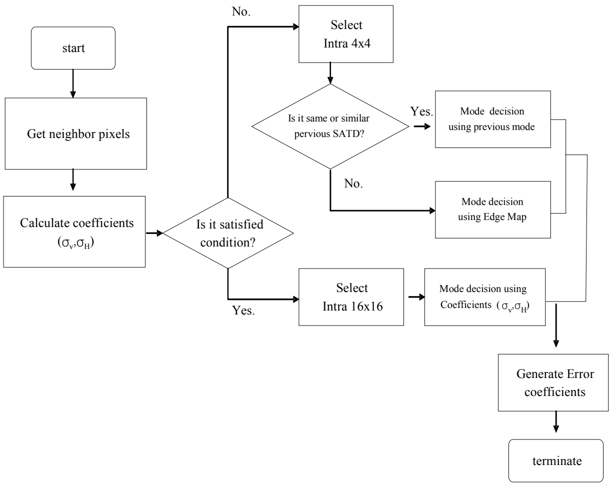



Fig. 2. Diagram of Proposed Algorithm

**Step1**: Let σ $_V$ and σ $_H$ be vertical and horizontal sum of difference between boundary pixels of current block and adjacent block, respectively. σ $_V$ and σ $_H$ are vertical and horizontal values for deciding the correlation between boundary pixels of current block and adjacent block, respectively

$$\sigma_{H_1} = \frac{\sqrt{2}}{16}\sum_{n=2}^{10}\left|(F(n,0)-F(n,-1))\right| \quad \sigma_{H_2} = \frac{\sqrt{2}}{16}\sum_{n=8}^{16}\left|(F(n,8)-F(n,-1))\right|$$
$$\sigma_{V_1} = \frac{\sqrt{2}}{16}\sum_{n=2}^{10}\left|(F(0,n)-F(-1,n))\right| \quad \sigma_{V_2} = \frac{\sqrt{2}}{16}\sum_{n=8}^{16}\left|(F(8,n)-F(-1,n))\right| \qquad \textbf{(1)}$$

where F(x,y) represents pixel value within macroblock and x and y represent index of vertical and horizontal.

In [7], intra prediction algorithm used only adjacent pixels with neighboring pixels. As a result, the performance of this method is not good when it exist the variation in the internal block. To solve this problem, we used two areas for finding internal change as shown in figure 5. The pixels in two areas are compared with the boundary pixels of the neighboring macroblock. One area is A and A` and C and C, the other area is B and B` and D and D`.To this aim, Step2 is used.

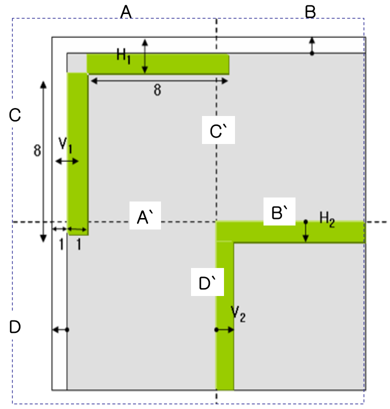



Fig. 3.Model of selection macro-block type

**Step2**: Examine whether the condition is satisfied or not.

The relation between quantization step and linear prediction error has been discussed in the reference [8]. As given in [8], the distribution of the pixel values after linear prediction in image can be modeled by a Laplacian distribution, which has a significant peak at zero.

$$P_e(e) = \frac{1}{\sqrt{2\sigma_e}}\exp\left(\frac{-\sqrt{2}\,|e|}{\sigma_e}\right) \qquad \textbf{(2)}$$

$\sigma_e$ is the standard deviation of distribution of error value.

For a zero mean Lapacian distribution, the probability value of errors will fall within ±3σ and satisfies following condition (3).

$$Qstep > 3\sigma_E \qquad \textbf{(3)}$$

where, qstep represents quantization step size.

Therefore, $\sigma_{V1}$, $\sigma_{V2}$, $\sigma_{H1}$ and $\sigma_{H2}$, given by (1) are value that is ratable the standard deviation of prediction error value calculated by current macroblock and neighboring macroblock. And these values are error values suited for distribution zero Laplacian. If (4) is satisfied, then we select intra 16x16 as coded block type for luminance.

$$\text{Qstep} > 3(\sigma_V), \;\; \& \;\; \text{Qstep} > 3(\sigma_H)$$
$$\sigma_V = \sigma_{V_1} + \sigma_{V_2} \quad \sigma_H = \sigma_{H_1} + \sigma_{H_2} \qquad \textbf{(4)}$$

If ±3σ is smaller than qstep, error value within macroblock is located near zero. Therefore, it has high correlation between boundary pixels and adjacent block and is considered to be an area that has a few changing such as background. If (4) is not satisfied, then intra 4x4 is selected as a coded block type for luminance.

## 3.2 Proposed fast mode decision

### 3.2.1 4x4 block type

In natural scene, it is observed that the current block's best mode was highly correlated with its neighboring blocks as shown in figure 6. Therefore, the prediction time can be reduced by using the correlation for current 4x4 block type and neighboring 4x4 block type.

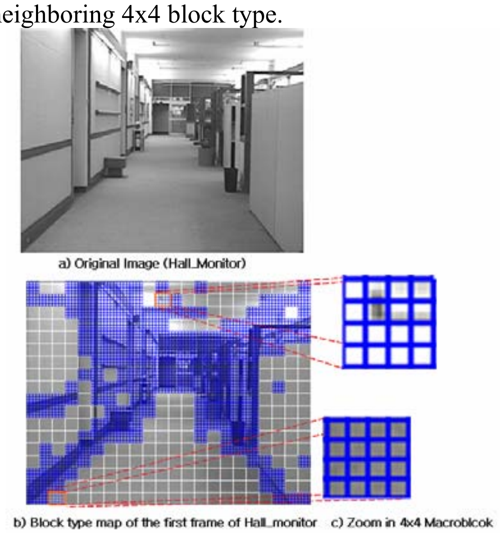



Fig. 4. Predicted sub-block by intra 4x4 in H.264

figure 7 Shows the proposed fast mode decision scheme for intra 4x4 prediction mode.

Step 1: Compute SATD according to the direction which is same to the previous prediction mode in the coding order. Eq. (5) is used to calculate SATD. It can reduce 50% complexity. [9]

Step 2: If the current SATD is not exceeding the previous value, prediction will be terminated. The previous mode is still used in the current block.

$$TR = HRH$$

$$= \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \\ 1 & -1 & 1 & -1 \end{bmatrix} \begin{bmatrix} r_{00} & r_{00} & r_{02} & r_{02} \\ r_{11} & r_{11} & r_{13} & r_{13} \\ r_{20} & r_{20} & r_{22} & r_{22} \\ r_{31} & r_{31} & r_{33} & r_{33} \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \\ 1 & -1 & 1 & -1 \end{bmatrix} \quad \mathbf{(5)}$$

$$= \begin{bmatrix} t_{r00} & t_{r01} & 0 & 0 \\ t_{r10} & t_{r11} & 0 & 0 \\ t_{r20} & t_{r21} & 0 & 0 \\ t_{r03} & t_{r31} & 0 & 0 \end{bmatrix}$$

R : residual image
H hadamard transform matrix
TR : trsted residual
SATD : $2\Sigma|tr_{ij}|$

Step 3: If condition is not satisfied, edge and edge-map are used. The direction of edge and edge-map can be found by using Eq. (6)~(9).

In that case, the candidate modes are DC mode and the mode which has the maximum amplitude in edge-map. We choose the one with the minimum SATD among two candidates.

$$\begin{aligned} dx_{i,j} &= 2f(i,j) - f(i-2, j-1) - f(i-3, j-1) \\ dy_{i,j} &= 2f(i,j) - f(i-1, j-1) - f(i-1, j) \end{aligned} \quad \mathbf{(6)}$$

where $dx_{i,j}$ and $dy_{i,j}$ represent the degree of difference in vertical and horizontal directions, respectively. Therefore, the amplitude of the edge vector can be decided by

$$Amp(\vec{D}_{i,j}) = |dx| + |dy| \quad (7)$$

The direction of the edge (in degree) is decided by the hyper-function,

$$Ang(\vec{D}_{i,j}) = \arctan\left(\frac{dy_{i,j}}{dx_{i,j}}\right), \left|Ang(\vec{D}_{i,j})\right| < 90° \quad (8)$$

Therefore the edge direction histogram of a 4 x 4 luminance block is decided as

$$\text{Histo}(k) = \sum_{Ang \in k} \text{Amp}\vec{D}, \quad k=0,1,3,4,5,6,7 \quad (9)$$

Mode0={-103.3°,-76.7°} Mode1={-13.3°,13.3°}
Mode5={-76.7°,-54.2°} Mode6={-35.8°,-13.3°}
Mode3={35.8°,54.2°} Mode4={-54.2°,-35.8°}
Mode7={54.2°,76.7°} Mode8={13.3°,35.8°}

In our method, the overall processing time is decreased because the directional prediction modes in the macroblock are very similar in the natural scene.

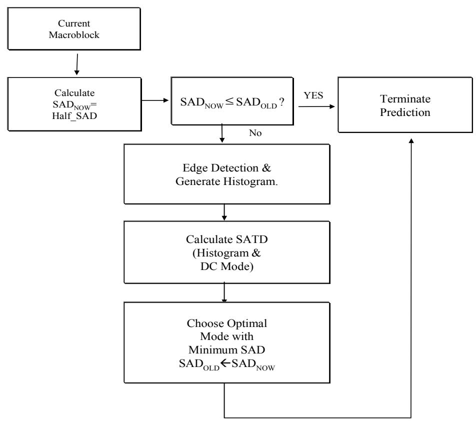



Fig. 5 Diagram of proposed Intra 4x4 Mode

### 3.2.2 16x16 macroblock

In this block, several candidate modes are generated from figure 8, in which $\sigma_H$ and $\sigma_V$ are given in Eq. (4). And then, the one with minimum cost is selected.

The overall process is as follows: if $\sigma_V$ is larger than $\sigma_H$, the candidate modes are vertical and DC mode, if $\sigma_V$ is

smaller than $\sigma_H$, the candidate modes are horizontal and DC mode, if $\sigma_V$ is equal to $\sigma_H$, the candidate modes are plan and DC mode.
Finally, determine the best mode between each candidate mode by choosing minimum SATD of the mode.

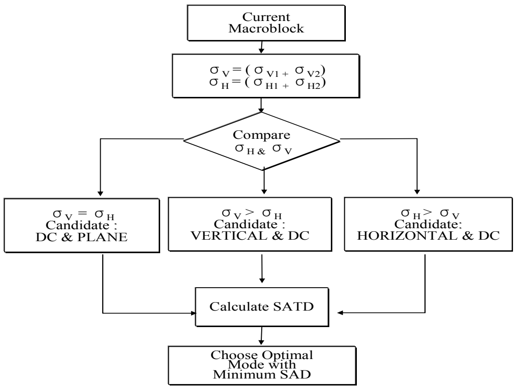


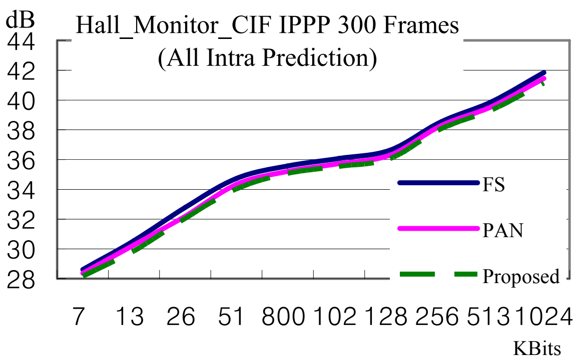

Fig. 6. Diagram of proposed 16x16 Intra modes

## 4. Experimental result

The proposed algorithm is implemented into JM10.1 provided by JVT(Joint Video Team). The test conditions are as follows;
(1) Fresh rate is 100.
(2) RD optimization is enabled (do not use inter prediction during P-frames.)
In this experiment, full-search (FS) and PAN algorithm are compared with the proposed algorithm,

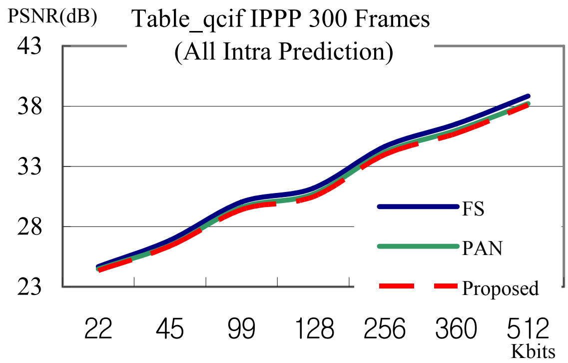
### 4.1 RD-Curve

figure 9 and 10 shows the proposed algorithm can achieve considerable computation reduction with negligible coding loss compared with FS.




| | ΔPsnr | ΔBits | ΔTime |
|---|---|---|---|
| FS | -0.594 dB | +0.06% | -82.526% |
| PAN | -0.237 dB | +0.14% | -60.006% |

Fig. 7. Tested on CIF sequences




| | ΔPsnr | ΔBits | ΔTime |
|---|---|---|---|
| FS | -0.625 dB | -0.001% | -82.451% |
| PAN | -0.171 dB | -0.082% | -64.960% |

Fig. 8. Tested on QCIF sequence.

### 4.2 Comparison Result

Table 1 shows the simulation results of the proposed algorithm. In the same bitrate, it can be seen that the proposed algorithm achieved the encoding time saving (up to about 43~70%) with negligible losses in PSNR and increments in bitrates.

| | Sequence | ΔTime(%) | ΔPSNR(dB) | ΔBits(%) |
|---|---|---|---|---|
| CIF | Tempete | -70.009 | -0.054 | +0.110 |
| | Foreman | -68.009 | -0.459 | -0.161 |
| | Stefan | -68.762 | -0.170 | -0.004 |
| QCIF | Mother and daughter | -53.529 | -0.139 | -0.133 |
| | Hall monitor | -43.007 | -0.006 | +0.025 |
| | container | -46.240 | -0.156 | -0.231 |

Table 1. Simultation result for CIF : 256 KBits , QCIF : 128 KBits , compared with PAN

## 5. Conclusion

Applying the newest video coding H.264 to the multi-channel video surveillance system, a great deal of computational time and memory is required. In particular, intra coding method, which reduces spatial redundancy, makes real-time processing difficult. To solve this problem, we proposed the efficient intra prediction mode decision method which can select the prediction block sizes between 4x4 and 16x16 sizes quickly, and then decide the directional modes.

In the simulation result, the proposed method has shown the better coding efficiency and lower complexity. The encoding time is greatly improved, and Δbitrate and ΔPSNR degradation are only about 0.1% and 0.2 dB, respectively. In conclusion, the proposed method can reduce the computational complexity without noticeable coding loss.

## 6. Reference


[1] Draft ITU-T Rec. and FDIS of Joint Video Spec. (H.264 | ISO/IEC 14496-10 AVC), JVT G050rl, Geneva23-27May2003

[2] Thomas Wiegand. "Overview of the H.264/AVC video coding standard," IEEE Trans.on Circuits and Systems for Video Technology, vol.13, no.7, pp.560-576, July2003

[3] Z. Yond-dong. D. Feng.L.Shou-xum "Fast 4x4 Intra-Predicition Mode Selection for H.264: IEEE Int. Cof. Multimedia & Export 2004

[4] Meng B Au O.C Wong C.W and Lam H.K, "Efficient Intra-Predition Algorithm in H.264," International conference on Image Processing,pp.837-840,2003

[5] Pan F.Lin X. Susanto R. Lim K.P Li Z.G Feng G.N Wu D.J and Wu S. "Fast Mode Decision Algorithm for Intra Prediction" VT G013 64th MPEG and 7th JVT meeting Pattya, Macrch 2003

[6] Iain Richardson "H.264 and MPEG-4 video comprssion" Willey 2003

[7] Jongho Kim and Jechang Jeong "Fast intra-mode decision in H.264 video coding using simple directional masks" Proc. of SPIE Bellingham, WA, 2005 Vol. 5960 1077

[8] I-Ming Pao and Ming-Ting Sun, "Modeling DCT Coefficients for Fast Video Encoding", IEEETrans. on Circuits and Systems for Video Technology, vol.9, no.4, June1999

[9] Y. W. Huang, et al,"Analysis, Fast algorithm, and VLSI architecture design for H.264/AVC intra frame coder," IEEETrans. on Circuits and Systems for Video Technology, vol.15,no.3,pp.378-501,Mar2005



### Acknowledgments

his work is financially supported by SMBA(small and medium business administration) through the project of the UniBiz (University for Inno-BIZ).


**Eung-Tae Kim** received the B.S. degree (summa cum laude) in Electronics Engineering from Inha University, Inchon, Korea, in 1991, and the M.S. degree and the Ph.D. degree in electrical & electronics engineering from Korea Advanced Institute of Science and Technology (KAIST), Taejon, Korea, in 1993 and 1999, respectively. From 1998 to 2004, he was a senior researcher at the Digital TV Lab. of LG Electronics Co. Ltd. Since 2004, he has been with the Department of Electronics Engineering at Korea Polytechnic University. His research interests include low bit-rate video transmission techniques, MPEG system and video, wireless multimedia transmission, and DTV/DMB/DVR system.

**Jang-Seon Ryu** received the B.S degree in Electronics Engineering from Korea Polytechnic University, Siheung, Korea, in 2006. Currently, he is a M. S. student in the Department of Information and Communication, Graduate school of Knowledge-Based Technology and Energy, Korea Polytechnic University. His research interests include high-performance video codec on the embedded system, multimedia system-on-chip design, digital video recorder and digital multimedia broadcasting.