Enhanced Inter Mode Decision Based on Contextual Prediction for P-Slices in H.264/AVC Video Coding

Byung-Gyu Kim and Suk-Kyu Song

We propose a fast macroblock mode prediction and decision algorithm based on contextual information for P-slices in the H.264/AVC video standard, in which the mode prediction part is composed of intra and inter modes. There are nine 4×4 and four 16×16 modes in the intra mode prediction, and seven block types exist for the best coding gain based on rate-distortion optimization. This scheme gives rise to exhaustive computations (search) in the coding procedure. To overcome this problem, a fast inter mode prediction scheme is applied that uses contextual mode information for P-slices. We verify the performance of the proposed scheme through a comparative analysis of experimental results. The suggested mode search procedure increased more than 57% in speed compared to a full mode search and more than 20% compared to the other methods.

Keywords: H.264/AVC, variable block mode, inter mode, intra mode, contextual mode information.

Manuscript received Sept. 28, 2005; revised Mar. 13, 2006.

Byung-Gyu Kim (phone: + 82 42 860 5766, email: bgkim@etri.re.kr) is with Embedded S/W Research Division, ETRI, Daejeon, Korea.

Suk-Kyu Song (email: sukkyu.song@samsung.com) is with System LSI Division, Samsung Electronics, Gyeonggi-do, Korea.

I. Introduction

Developments in video coding techniques have accelerated over the last several years. H.264/AVC video coding is the newest standard defined by the Joint Video Team (JVT) [1], [2] in which various techniques have been adopted to obtain a high coding efficiency compared to previous standards.

Among the new techniques of H.264/AVC video coding, motion estimation has been introduced to improve the previous video standards such as H.261 and H.263 [1], [2], [3]. Motion estimation for inter prediction is generally performed only on a 16×16 macroblock (MB). Then, each 16×16 MB is assigned one motion vector. This method can lead to minimum block distortion.

Variable block sizes for inter mode prediction maximizes the coding efficiency based on rate-distortion (RD) optimization in the H.264/AVC coding standard [2]. The block sizes are 16×16, 16×8, 8×16, 8×8, 8×4, 4×8, and 4×4. In addition, intra mode prediction (nine 4×4 and four 16×16 modes) follows inter mode prediction to find the best residual image [1], [2], [4]. For any block mode, the motion vector estimation for a minimum residual image is also performed by full search [3] or fast motion search methods such as a directional diamond search [5] and fast-adaptive rood pattern search [6]. These structures give rise to a complex encoder to be applied in real-time applications. Therefore, a fast mode decision scheme that can reduce the complexity of the H.264/AVC video encoder is needed.

Many kinds of fast inter mode decision schemes have been reported that use the following RD optimization [7]-[13]:

$$J_{RD} = SSD_{Mode} + \lambda \cdot \{R(header) + R(residual)\}, \qquad (1)$$

where J_{RD} is a bitrate-distortion value as a cost function, SSD_{Mode} is the sum of the squared differences for the given mode, λ denotes the Lagrangian multiplier, $R(x)$ is a bit amount for coding x, *header* provides header information, and *residual* is the residual data of the given MB.

Kim and others [7] proposed an algorithm using the property of an all-zero coefficients block that is produced by quantification and coefficient thresholding to effectively eliminate unnecessary inter modes. In this method, there are several threshold values that should be predefined. Also, transform coefficients are needed to make a fast decision. Jing and Chau [12] developed a fast inter mode decision scheme by using both the frame difference and the MB difference. This is a useful scheme, since only frame difference and MB difference images are required to assign a mode.

A fast algorithm using a ratio measurement of the RD costs of two classes defined as Class16 and Class8 has been proposed by Lee and Jeon [8]. In [11], Wu and others make use of the spatial homogeneity of a video object's textures and the temporal stationary characteristics inherent in video sequences. However, the method suffers from a drawback to obtain an edge image for the texture information and a difference image for the temporal stationarity characteristics.

Zhu and others [9], [10] proposed another approach for a fast inter mode decision that uses a pre-encoding, down-sampled image space. After obtaining candidate block modes, a refinement search is performed to find the best mode in the original image space. To detect an early SKIP block, Crecos and Yang [13] also used the prediction and thresholding scheme. In [14], a transcoding algorithm for MPEG-2 video to H.264 video with a fast motion search has been proposed by Lu and others. Their fast motion search is composed of a fast inter mode decision for the B- and P-frames, and fast intra prediction for intra coding. In Lu's algorithm, the suggested inter mode prediction for SKIP, 16×16, and DIRECT modes used neighboring MBs of the target MB in the current frame.

We propose a fast inter mode decision scheme using a contextual square mode (16×16 or 8×8 block mode) classification. Contextual mode information is used for an inter mode decision for categorization of the current MB as a member or non-member of the $P16\times16$ sub-block type.

Section II describes our proposed algorithm with observations in detail. We verify the performance of the proposed algorithm in section III, and conclusions are presented in section IV.

II. The Proposed Inter mode Decision Algorithm

Compression performance is improved along with an increase in the computational cost. A better strategy to achieve a fast mode search is to use a prediction technique and mode property information.

1. Observation

In many signal processing applications, the characteristics of distribution give good information to process their system. If any probabilistic property exists in the result of the mode decision of P-slices, we can employ it to evaluate the fast version of the mode decision. Thus, we observe the probability of each inter mode type through several experiments for P-slices.

Figure 1 shows examples of the probabilities for the finally determined inter mode. As shown in the figure, most of the inter modes are concentrated on the $P16\times16$ block. The remains are assigned to the $P16\times8$, $P8\times16$, or $P8\times8$ sub-block modes.

To utilize this property, we designed the inter mode decision into the classification problem of two square mode sets, that is, $P16\times16$ sub-mode ($P16\times16$, $P16\times8$, $P8\times16$) and $P8\times8$ sub-mode ($P8\times8$, $P8\times4$, $P4\times8$, $P4\times4$).

2. Square Mode Set Classification for the Inter Mode Decision

The variable block modes for the inter-frame prediction consist of the two square mode sets, $P16\times16$ sub-mode ($P16\times16$, $P16\times8$, $P8\times16$) and $P8\times8$ sub-mode ($P8\times8$, $P8\times4$, $P4\times8$, $P4\times4$). In general, all seven different modes are checked to determine the best inter mode in the full-mode decision scheme. However, this is an extensive computation.

As described in the previous subsection, we observed that $P16\times16$ sub-mode ($P16\times16$, $P16\times8$, and $P8\times16$) takes a large portion of the decided inter mode. In fact, it is rare that the $P8\times8$ sub-mode occurs, as shown in Fig. 1. Thus, it is helpful to determine if the given MB is in the $P16\times16$ sub-mode.

To achieve a fast version of the mode decision, we devised a mode classification method that reduces the number of mode searches. Based on contextual mode information, we determine whether the mode goes to $P16\times16$ sub-mode or $P8\times8$ sub-mode.

To obtain contextual mode information, both temporal and spatial contexts are used, as shown in Fig. 2. For temporal mode information, the corresponding MB and surrounding MBs of the current MB to be processed are used to obtain information for the mode prediction. We also use mode information from available neighboring MBs of the current MB in the current frame k, as depicted to the right in Fig. 2.

With this contextual mode information, we define a score criterion for classification into the square mode sets as

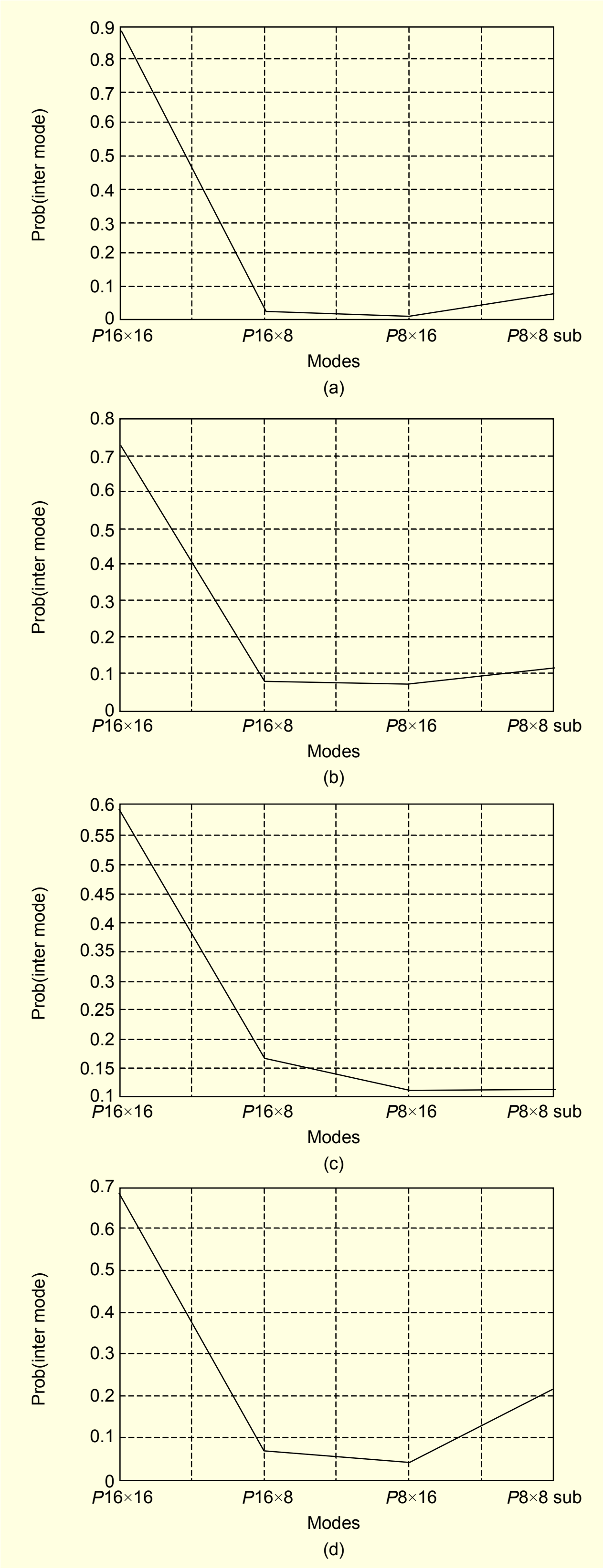

Fig. 1. Examples of the inter-mode probabilities of the (a) Mobile sequence, (b) Foreman sequence, (c) News sequence, and (d) Coastguard sequence.

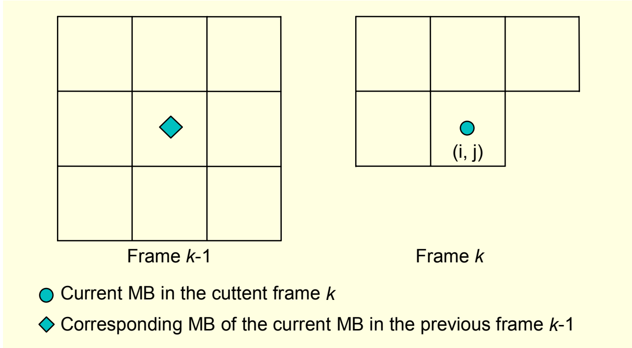

Fig. 2. Contextual maps for inter mode prediction.

$$S(i,j) = \text{Temporal info}(i,j) + \text{Spatial info}(i,j)$$
$$= \sum_{x=i-1}^{i+1} \sum_{y=j-1}^{j+1} Bin_{xy}(M_{xy}^{k-1}) + \sum_{(x,y)\in\{(i-1,j-1)\}} Bin_{xy}(M_{xy}^{k-1}), \quad (2)$$

where $Bin_{xy}(M_{xy}^{k})$ is set to 1 if M_{xy}^{k} is included in the $P16\times16$ sub-mode. Otherwise, it returns zero value. Also, M_{xy}^{k} is a mode of the MB at (x, y) at frame k. According to this score $S(i, j)$, the current MB of the P-slice is checked on the $P16\times16$ sub-mode if $S(i, j)$ is greater than zero. Otherwise, it is necessary to check on the $P8\times8$ sub-mode for the current MB.

For the proposed inter mode prediction, we guess that the mode number of the search can be decreased since most of the inter modes are included in the $P16\times16$ mode. If the current MB is selected as a $P16\times16$ sub-mode, then we need to check only three modes ($P16\times16$, $P16\times8$, $P8\times16$). Otherwise, four modes ($P8\times8$, $P8\times4$, $P4\times8$, $P4\times4$) must be checked. A motion vector search and a frame difference image are not required for the search path determination [7]-[9], [12].

In order to verify our method, the conditional probability, $\text{Prob}(M_{ij}^{k} \in P16\times16 \text{ sub-mode} \mid S(i, j) \neq 0)$ was computed through analyses of various sequences (CIF size). Table 1 shows the results. We can deduce that the proposed scheme for inter mode prediction can be applied to speed-up the mode decision with little degradation in image quality.

To show the accuracy of the proposed scheme, it may be better to use $\text{Prob}(S(i, j) = 0 \mid M_{ij}^{k} \in P8\times8 \text{ sub-mode})$.

Table 1. The conditional probability for square mode set classification.

Sequences	Prob ($M_{ij}^{k} \in P16\times16$ sub-mode | S(i, j)≠0)
Paris	97.6 (%)
News	97.2 (%)
Mobile	99.2 (%)
Hall monitor	98.9 (%)
Foreman	97.2 (%)

Through experiment, we checked that this conditional probability has a similar property to the conditional probability of Prob($M_{ij}^{k} \in P16\times16$ sub-mode | $S(i, j) \neq 0$), which is shown in Table 1.

Figure 3 shows the proposed fast mode decision algorithm that can guarantee a reliable image quality. For an input MB of the P-slice, the overall scheme for a fast inter mode decision (motion estimation) based on the suggested techniques can be summarized as follows.

Step 1. Inter mode prediction and search

By using the proposed contextual mode classification, the prediction is categorized into the two square mode sets of the $P16\times16$ and $P8\times8$ sub-modes. Possible modes are then checked to obtain the best inter mode for each square set. For example, three modes ($P16\times16$, $P16\times8$, $P8\times16$) may be possible if the current MB is classified into the $P16\times16$ sub-mode.

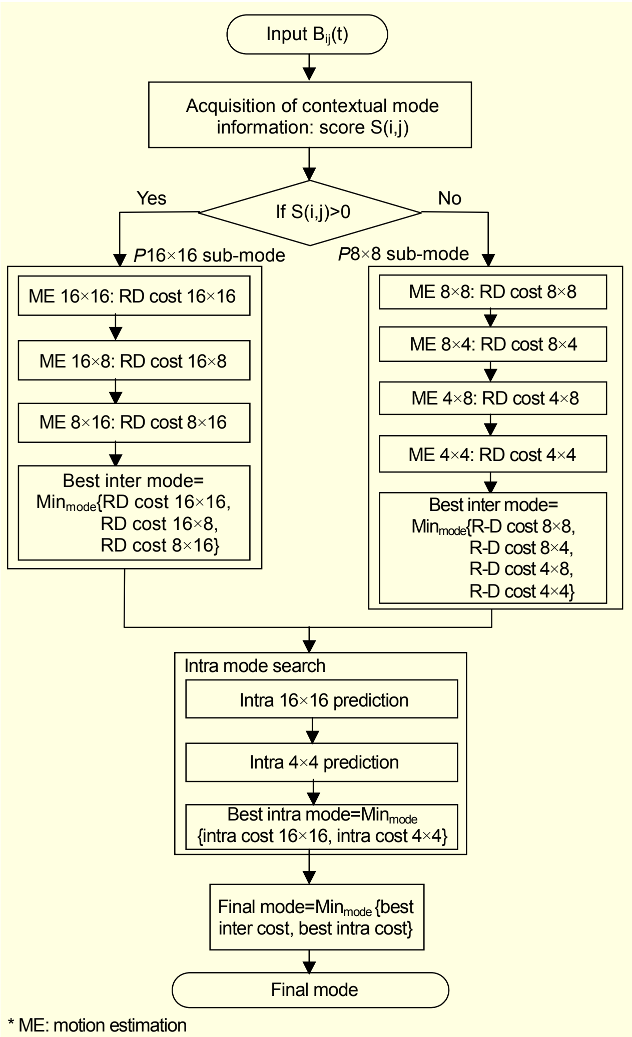

Fig. 3. Proposed scheme for fast mode decision.

Step 2. Intra mode search

An intra mode prediction and search is performed on all MBs.

Step 3. Determination of the best mode

Between the best inter mode and the best intra mode, a mode that has the lower RD cost is selected as the final mode for the given MB.

Our algorithm gives a reduction of the inter mode search. Also, since we do not use the motion vector search or transform for initially choosing a routine for the mode search, as is required in other methods [7]-[12], but rather use contextual mode information in the P-slices, the mode search procedure speed can be increased.

To analyze the estimated complexity (the number of RD mode searches) in our algorithm, we model our algorithm by using each search path as follows:

$$S_{Avg} = p_{\mathrm{I}} \cdot (\text{No. of modes for } P16\times16 \text{ sub-mode path}) + p_{\mathrm{II}} \cdot (\text{No. of modes for } P8\times8 \text{ sub-mode path}), \quad (3)$$

where p_{I} is is the posterior probability that the given MB is determined as a mode of the $P16\times16$ sub-mode, and p_{II} is the posterior probability of the $P8\times8$ sub-mode. The number of modes for a $P16\times16$ sub-mode path is 3 ($P16\times16$, $P16\times8$, $P8\times16$) and the number of modes for a $P8\times8$ sub-mode path is 4 ($P8\times8$, $P8\times4$, $P4\times8$, $P4\times4$). Therefore, if $p_{\mathrm{I}} = 0.9$ and $p_{\mathrm{II}} = 0.1$, we can typically compute the average complexity (the number of mode searches) as 3.1 modes ($\Delta S = 54.29$ %) per MB.

III. Results and Discussion

To verify the performance of the proposed fast mode decision algorithm, various MPEG standard sequences were used with CIF and QCIF sizes. Analyses were performed with encoding frames = 100, RD optimization enabled, quantization parameter (QP) = 28 or 32, a sequence type of IPPP, using CAVLC, search range of MV = ±16, and the number of the reference frames = 1.

Joint model (JM) reference software by JVT was used as a reference code [15] for evaluation of the encoding performance. However, this has a time-exhaustive structure for encoding and decoding an H.264/AVC bitstream.

Several open software packages for the H.264/AVC encoder have been developed to support real-time applications. We used a x264 encoder complying with the H.264/AVC video coding standard to check the proposed algorithm [16]. This video encoder is fast compared with the JM reference software because of the modulization and fast motion estimation scheme. We therefore implemented the proposed techniques in the x264 encoder. All algorithms were also run on a hardware platform of a Pentium-4 PC with 3.4 GHz CPU and 1 GB of RAM.

At the beginning of *P* frames, the decided mode information is not available to us. Thus, we need to implement an initialization process for the modes in the first P-frame. To do this, we utilize the full mode search (FMS) for an initial P-frame since the accuracy of its mode decision plays an important role to be used for mode prediction in the next P-frame. Otherwise, another fast mode decision algorithm can be implemented for the first P-frame if its mode decision accuracy is sufficiently guaranteed.

We defined several measurements for evaluating the encoding performance, including average $\Delta PSNR$, average $\Delta Bits$, a time-saving factor ΔS, and consumed-time saved for a mode decision, ΔT. The average $\Delta PSNR$ is the difference between the average *PSNR* of the proposed method and the corresponding value of another method. As performance improves, this criterion becomes more positive. Also, ΔS is defined as

$$\Delta S = \frac{(No.of\ RD\,search_{ref}) - (No.of\ RD\,search_{proposed})}{No.of\ RD\,search_{ref}}. \quad (4)$$

If this value increases, the performance speed is increased. The average ΔS is the bit-rate difference between the methods. Performance improves with a less negative value. Finally, the time for a mode decision ΔT is defined as

$$\Delta T = \frac{Consumed\ T_{ref} - Consumed\ T_{proposed}}{Consumed\ T_{ref}}, \quad (5)$$

under the condition that the FMS is optimum. As this value increases, the performance speed is increased.

To determine the effect of the variation of *QP* on the inter mode decision, distributions of the determined modes were generated by using the Paris and Stefan sequences according to various *QP* values, as illustrated in Figs. 4 and 5. From these results, we can observe that most of the inter modes are concentrated on a *P*16×16 sub-block mode. Sub-block modes are maintained while the *QP* value changes. This indicates that the suggested mode decision scheme is valid even though the applied *QP* value varies.

Tables 2, 3, 4, and 5 show results for *QP*=28 and *QP*=32. As we can see from these results, there is little loss of quality in all tested sequences. For the case of *QP*=28, the $\Delta PSNR$ (dB) and $\Delta Bits$ are negligible while having significant increase of speed for the mode search. Our approach can obtain a gain of over 57% in terms of the time-saving factor.

When *QP*=32, the average PSNR decreased slightly due to the increase of the *QP*. However, we can see that the proposed algorithm gets a similar degree of the $\Delta PSNR$ (dB) and $\Delta Bits$. In aspect of the time-saving factor ΔS, our algorithm is still 57% faster than the full search for mode selection as illustrated in Table 5.

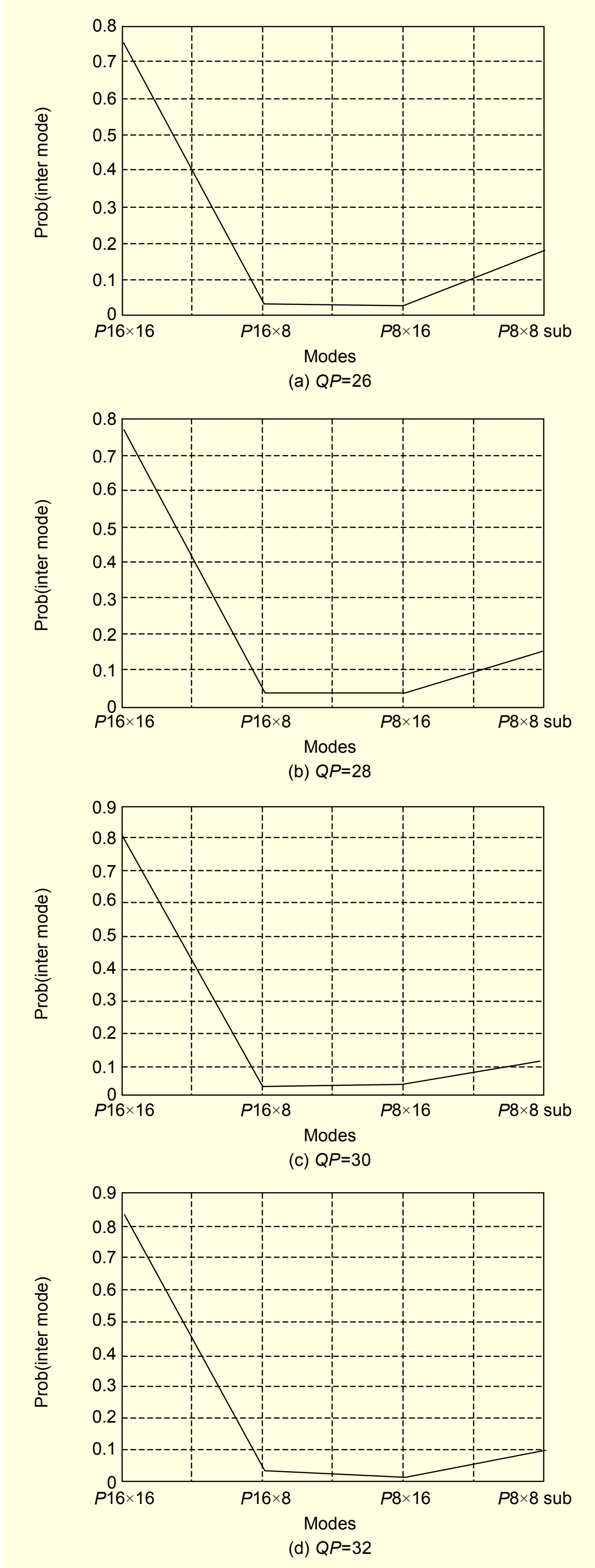

Fig. 4. The inter mode probabilities of Paris (CIF) sequence as *QP*.

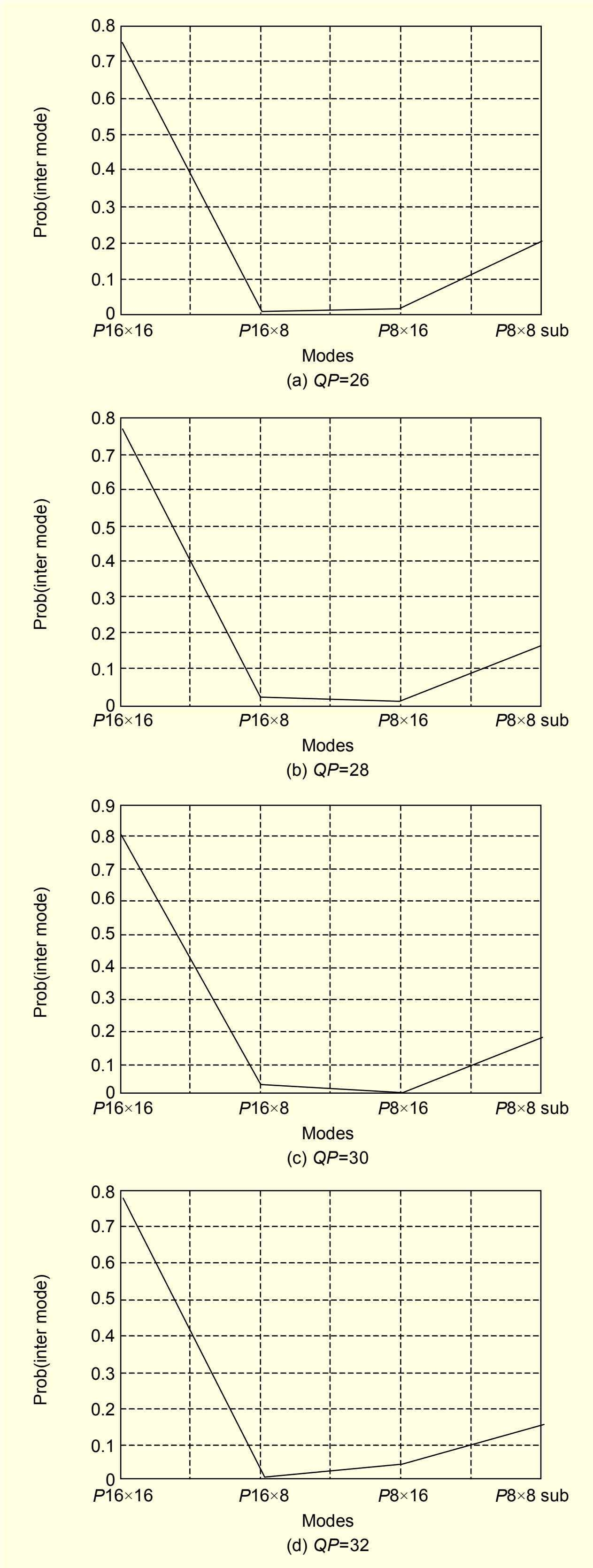

Fig. 5. The inter mode probabilities of Stefan (QCIF) sequence as *QP*.

Table 2. Performance of the proposed inter mode decision for *QP* = 28: Original *PSNR*, *Bits*, *S* and *T* (FMS: full mode search).

Contents		*PSNR*	*Bits* (kbps)	*S* (No. of R-D search)	*T* (ms)
Foreman (QCIF)	FMS	38.167	219.52	7.0000	1797
	Proposed	38.140	221.36	3.0051	1066
Paris (CIF)	FMS	38.437	1451.34	7.0000	6412
	Proposed	38.400	1460.77	3.0035	4116
Mobile (CIF)	FMS	35.678	2895.88	7.0000	7841
	Proposed	35.654	2900.80	3.0089	4865
News (CIF)	FMS	40.899	700.81	7.0000	6102
	Proposed	40.876	707.32	3.0029	3972
Stefan (QCIF)	FMS	36.250	599.86	7.0000	1857
	Proposed	36.210	604.71	3.0113	1049

Table 3. Performance of the proposed inter mode decision for *QP* = 32: Original *PSNR*, *Bits*, *S* and *T* (FMS: full mode search).

Contents		*PSNR*	*Bits* (kbps)	*S* (No. of R-D search)	*T* (ms)
Foreman (QCIF)	FMS	35.931	176.66	7.0000	1764
	Proposed	35.898	221.36	3.0051	1048
Paris (CIF)	FMS	35.493	982.44	7.0000	6374
	Proposed	35.467	990.08	3.0035	4159
Mobile (CIF)	FMS	32.700	1799.68	7.0000	7660
	Proposed	32.689	1804.87	3.0089	4843
News (CIF)	FMS	38.338	484.23	7.0000	6076
	Proposed	38.325	489.06	3.0029	4011
Stefan (QCIF)	FMS	33.337	387.82	7.0000	1791
	Proposed	33.302	390.46	3.0070	1048

Table 4. Performance of the proposed inter mode decision for *QP* = 28: Δ*PSNR*, Δ*Bits*, Δ*S*, and Δ*T*.

Sequences	Δ*PSNR* (dB)	Δ*Bits* (%)	Δ*S* (%)	Δ*T* (%)
Foreman (QCIF)	-0.026	0.84	57.07	42.45
Paris (CIF)	-0.037	0.65	57.10	35.82
Mobile (CIF)	-0.024	0.17	57.01	37.95
News (CIF)	-0.023	0.93	57.10	34.80
Stefan (QCIF)	-0.040	0.80	56.98	41.08

Table 5. Performance of the proposed inter mode decision for *QP* = 32: Δ*PSNR*, Δ*Bits*, Δ*S*, and Δ*T*.

Sequences	Δ*PSNR* (dB)	Δ*Bits* (%)	Δ*S* (%)	Δ*T* (%)
Foreman (QCIF)	-0.033	0.59	57.05	40.58
Paris (CIF)	-0.026	0.77	57.10	34.76
Mobile (CIF)	-0.010	0.28	57.04	36.77
News (CIF)	-0.013	0.99	57.09	33.98
Stefan (QCIF)	-0.035	0.68	56.04	40.20

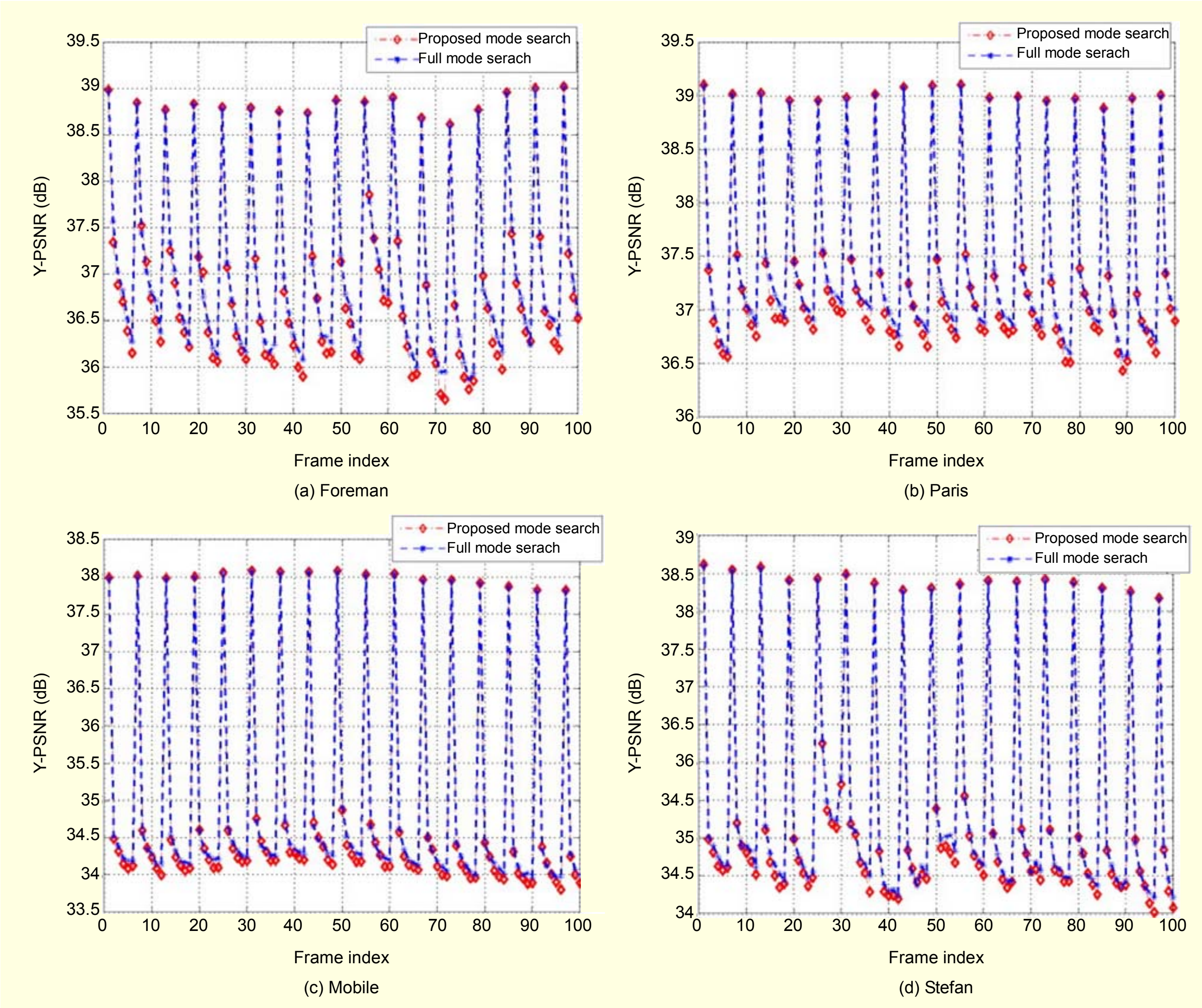

Fig. 6. PSNR values as frames pass (QP=28).

For a comparison of ΔT, we determined the time required for the overall mode decision procedure, including intra prediction. The same intra prediction was used in all analyses. The results show that our algorithm reduces the required time more than 33% for the mode decision procedure. These values are less than the ΔS values, which only consider the mode search number. Therefore, this is caused by the time consumed for both inter mode and intra mode decisions.

To determine the performance of the proposed algorithm as frames elapse, we performed *PSNR* analyses and determined the variation in the number of MBs ($P16\times16$ sub-mode and $P8\times8$ sub-mode) with time. For a more obvious illustration, we used a group of picture (GOP) structure of IPPP. In general, the *PSNR* decreases slowly as a frame moves away from the I-frame in each GOP. Figure 6 shows PSNR (Y component) variation as frames pass when QP = 28. We compared our algorithm with the FMS method. The results show that the tendency of the *PSNR* of the proposed scheme is to keep up with the FMS with a reliable error bound.

In Fig. 7, the number of MBs varies ($P16\times16$ sub-mode and $P8\times8$ sub-mode) as frames pass. The number of MBs was plotted only for P-frames. Sequences with high motion occur in the $P8\times8$ sub-mode MB. From these results, we can observe that MBs of the $P8\times8$ sub-mode exist, although they are very small. From these results, we can infer that the proposed mode decision algorithm provides a reliable performance and increases the speed of the motion search procedure.

For an objective comparison, we used Jing's [12] and Jeon's methods [8], which are both fast inter mode decision schemes. The comparative results for several sequences are shown in Table 6. For the Foreman sequence with the size of QCIF, the

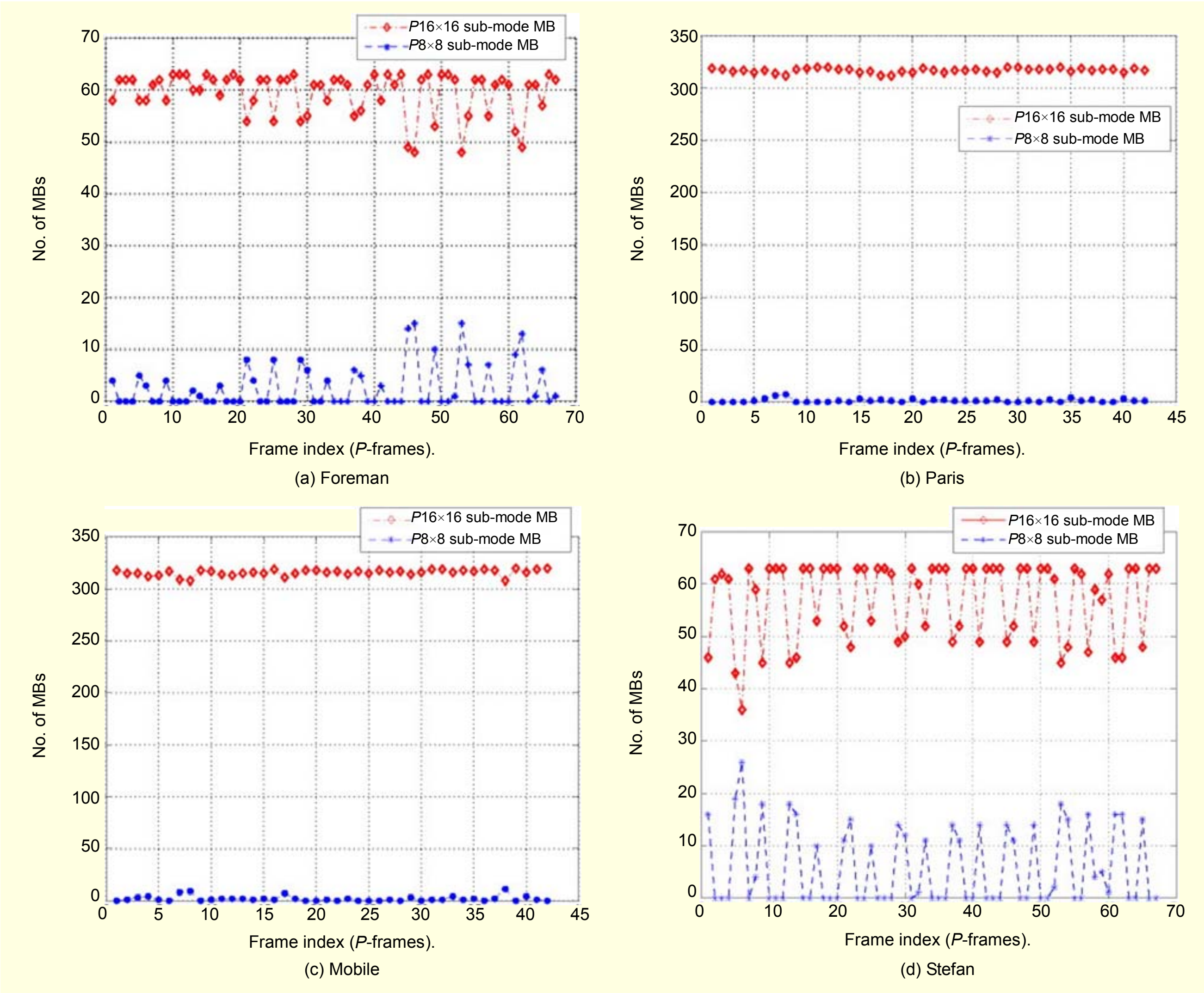

Fig. 7. The number of MBs (*P*16×16 sub-mode and *P*8×8 sub-mode) as the frame passes (*QP*=28).

proposed inter mode decision technique gives better quality based on Δ*PSNR*, Δ*S* and Δ*T*.

A similar performance was achieved for the Paris sequence. Both the speed and quality are improved with a smaller increment in bitrate. The Mobile & Calendar and News sequences show that the devised algorithm yields improved quality with smaller Δ*Bits*. Also, the mode search procedure was speeded-up more than 57% compared to a FMS and more than 20% compared to the other two methods.

To keep up generality, we implemented our fast mode decision algorithm on the JM 9.6 reference encoder. Table 7 presents the results of the applied techniques for fast mode decision based on the FMS. As shown in Table 7, the proposed fast mode decision scheme also increased the speed of the mode decision by 15 to 20% compared with the other two methods based on the time-saving factor Δ*S*. Also, our scheme is superior to the other two methods in terms of the image

Table 6. Performance comparison of the proposed inter mode decision (*QP*=28).

Sequences		Δ*PSNR* (dB)	Δ*Bits* (%)	Δ*S* (%)
Foreman (QCIF)	Jeon's	-0.06	1.08	14.07
	Jing's	-0.05	0.20	29.00
	Proposed	-0.026	0.84	57.05
Paris (CIF)	Jeon's	-0.07	1.44	57.10
	Jing's	-0.07	1.50	14.71
	Proposed	-0.037	0.65	33.00
Mobile (CIF)	Jeon's	-0.03	0.79	13.35
	Jing's	-0.05	0.25	33.10
	Proposed	-0.024	0.17	57.01
News (CIF)	Jeon's	-0.06	1.11	14.92
	Jing's	-0.06	1.40	32.20
	Proposed	-0.023	0.93	57.10

Table 7. Performance comparison of the proposed inter mode decision on the JM 9.6 reference encoder (QP=28).

Sequences		$\Delta PSNR$ (dB)	$\Delta Bits$ (%)	ΔS (%)
Foreman (QCIF)	Jeon's	-0.06	1.08	14.07
	Jing's	-0.05	0.20	29.00
	Proposed	-0.034	0.82	45.83
Paris (CIF)	Jeon's	-0.07	1.44	57.10
	Jing's	-0.07	1.50	14.71
	Proposed	-0.033	0.69	47.63
Mobile (CIF)	Jeon's	-0.03	0.79	13.35
	Jing's	-0.05	0.25	33.10
	Proposed	-0.076	0.45	54.74
News (CIF)	Jeon's	-0.06	1.11	14.92
	Jing's	-0.06	1.40	32.20
	Proposed	-0.024	0.56	45.90

Fig. 8. Decoded images for visual comparison of the Foreman sequence (QCIF). (a)-(c): frame 10, (d)-(f): frame 25, and (g)-(i): frame 50.

quality ($\Delta PSNR$) while keeping up the increase of the bitrates to less than 1.0%. In the Mobile & Calendar, there is a little loss of $\Delta PSNR$. However, the speed of the proposed mode decision algorithm can be faster than the other two methods by at least 20%, as we can see from the result.

To inspect visual quality, the decoded frames are shown in Figs. 8 and 9. In Fig. 8, there is little degradation in visual quality. Actually, a loss of $PSNR$ is at most about 0.03 (dB). As we can see from the decoded results, this almost cannot be recognized in all decoded samples.

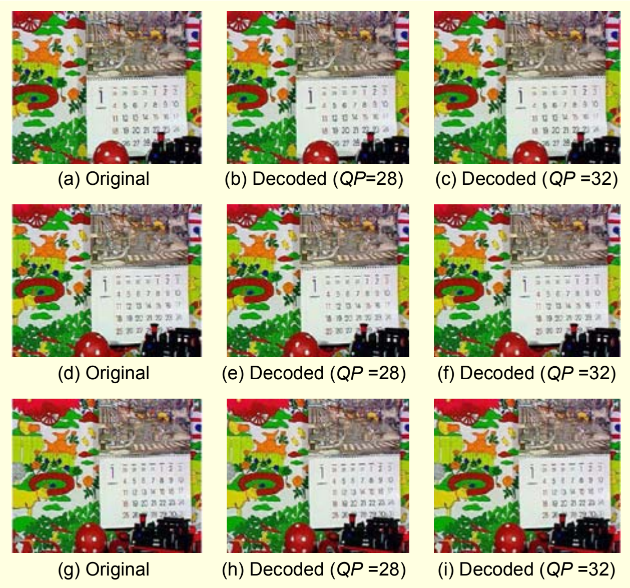

Fig. 9. Decoded images for visual comparison of the Mobile & Calendar sequence (CIF). (a)-(c): frame 10, (d)-(f): frame 25, and (g)-(i): frame 50.

In the Mobile & Calendar sequence, similar results were obtained. In the case of QP=32, there is little loss of image quality. In particular, if we observe carefully we can see that the detail quality of the surface of the red ball with white spots in the decoded image is somewhat less owing to the increased quantization parameter QP. From these results, we can summarize that the suggested algorithm is very effective in speeding-up the mode search procedure in H.264/AVC video coding.

IV. Conclusion

We have introduced a fast inter mode decision scheme that uses a contextual square mode (16×16 or 8×8 sub-mode sets) classification module. For a fast inter mode decision, contextual mode information and the probability characteristic of the inter mode were used to decrease the number of mode searches.

Combining this technique with the H.264/AVC video encoder allows the motion estimation procedure to be speeded-up with little loss of quality. We verified the enhanced performance of the proposed algorithm through comparative analyses.

Acknowledgement

The authors wish to thank the reviewers for their valuable comments and Dr. Cassell for helping to check the writing of this paper in English.

References

[1] ISO/IEC 14496-10, *Information Technology–Coding of Audio-Visual Objects–Part 10: Advanced Video Coding*, Dec. 2003.

[2] T. Wiegand, G. J. Sullivan, G. Bjontegard, and A. Luthra, "Overview of the H.264/AVC Video Coding Standard," *IEEE Trans. Circuit Syst. Video Technol.*, vol. 13, July 2003, pp. 560-576.

[3] June-Young Chang, Won-Jong Kim, Young-Hwan Bae, Jin Ho Han, Han-Jin Cho, and Hee-Bum Jung, "Performance Analysis for MPEG-4 Video Codec Based on On-Chip Network," *ETRI Journal*, vol.27, no.5, Oct. 2005, pp.497-503.

[4] Ki-Bum Suh, Seong-Mo Park, and Han-Jin Cho, "An Efficient Hardware Architecture of Intra Prediction and TQ/IQIT Module for H.264 Encoder," *ETRI Journal*, vol. 27, no. 5, Oct. 2005, pp. 511-524.

[5] H. Jia and L. Zhang, "Directional Diamond Search Pattern for Fast Block Motion Estimation," *Electronics Letters*, vol. 39, no. 22, 2003, pp. 1581 - 1583.

[6] B.-G. Kim, S.-T. Kim, S.-K. Song, and P.-S. Mah, "Fast-Adaptive Rood Pattern Search for Block Motion Estimation," *Electronics Letters*, vol. 41, no. 16, Aug. 4, 2005, pp. 25-26.

[7] Y.-H. Kim, J.-W. Yoo, S.-W. Lee, J. Shin, J. Paik, and H.-K. Jung, "Adaptive Mode Decision for H.264 Encoder," *Electronics Letters*, vol. 40, no. 19, Sept. 2004, pp. 1172-1173.

[8] J. Lee and B. W. Jeon, "Fast Mode Decision for H.264 with Variable Motion Block Sizes," *Proc. Int'l Symp. Computer and Information Sciences,* 2003, pp. 723-730.

[9] D. Zhu, Q. Dai, and R. Ding, "Fast Inter Prediction Mode Decision for H.264," *Proc. IEEE Int'l Conf. Multimedia and Expo (ICME)*, vol. 1, 2004, pp. 1123-1126.

[10] C.-H. Kuo, M. Shen, and C.-C. J. Kuo, "Fast Inter-Prediction Mode Decision and Motion Search for H.264," *Proc. IEEE Int'l Conf. Multimedia Expo (ICME)*, 2004, pp. 663-666.

[11] D. Wu, S. Wu, K. P. Lim, F. Pan, Z. G. Li, and X. Lin, "Block Inter Mode Decision for Fast Encoding of H.264," *Proc. IEEE Int'l Conf. Acoustics, Speech, and Signal Proc.(ICASSP)*, vol. 3, 2004, pp. 181-184.

[12] X. Jing and L.-P. Chau, "Fast Approach for H.264 Inter Mode Decision," *Electronics Letters*, vol. 40, no. 17, Sept. 2004, pp. 1050-1052.

[13] C. Crecos and M. Y. Yang, "Fast Inter Mode Prediction for *P* Slices in the H264 Video Coding Standard," *IEEE Trans. Broadcasting*, vol. 51, no. 2, June 2005, pp. 256 - 263.

[14] X. Lu, A. M. Tourapis, P. Yin, and J. Boyce, "Fast Mode Decision and Motion Estimation for H.264 with a Focus on MPEG-2/H.264 Transcoding," *Proc. IEEE Int'l Symp. Circuits and Systems (ISCAS)*, vol. 2, May 2005, pp. 1246 - 1249.

[15] Joint Model (JM) - H.264/AVC Reference Software, http://iphome.hhi.de/suehring/tml/download/

[16] Videolan, x264 - a free h264/avc encoder, http://developers.videolan.org/x264.html.

Byung-Gyu Kim received his BS degree from Pusan National University, Korea, in 1996 and MS degree from Korea Advanced Institute of Science and Technology (KAIST) in 1998. In 2004, he received the PhD degree in the Department of Electrical Engineering and Computer Science from Korea Advanced Institute of Science and Technology (KAIST). In March 2004, he joined in the Real-Time Multimedia Research Team at the Electronics and Telecommunications Research Institute (ETRI), Korea, where he is currently a Senior Researcher. His research interests image segmentation for the content-based image coding, real-time multimedia communication and intelligent information system for image signal processing. He is a member of IEEE, IEEE Computer and Communication Societies, and KMMS (Korea Multimedia Society). Also, he is a member of IEICE.

Suk-Kyu Song received his BS degree in electronic engineering from Pusan University of Foreign Studies, Korea, in 2001 and MS degree in electronics engineering from Kyungpook National University, Korea, in 2003. From 2003 to 2004, he worked as a Research Engineer at LG Electronics. From 2004 to April 2006, he worked in the Real-Time Multimedia Research Team at the Electronics and Telecommunications Research Institute (ETRI), Korea. From May 2006, he joined in the Samsung Electronics, Korea, where he is currently a senior research engineer. His research interests include image compression and its applications, SOC, real-time multimedia communication and intelligent information system for image signal processing.