Intra-Prediction Mode Decision for H.264 in Two Steps

Song-Hak Ri, Joern Ostermann

Institut für Informationsverarbeitung, University of Hannover
Appelstr 9a, D-30167 Hannover, Germany

Abstract. Two fast mode decision methods for intra prediction in H.264 are proposed in this work to reduce the encoder complexity. The first method, the so-called two-level fast Intra4x4 mode decision, is based on the idea of reducing possible candidate directions of a 4x4 luma block in Intra4x4 mode decision. The second method enables an early selection of suitable 4x4 or 16x16 luma partitions of a given macroblock.

The performance of the proposed algorithms are evaluated using the H.264 reference codec JM 10.1, which is compared with exhaustive RD mode decision of H.264 in metrics such as the encoding time, the average PSNR and the coding bit-rate for test sequences. It is observed that the proposed fast algorithms can save about 52% of encoding time resulting in a minor PSNR degradation of 0.03dB lower PSNR and a 2.5% higher bit rate.

Index Terms—H.264/AVC, fast intra mode decision, video coding

1. INTRODUCTION

The emerging H.264 coding standard incorporates many state-of-the-art techniques to achieve outstanding coding performance. The enhanced intra prediction technique is one of the important factors that contribute to the success of H.264. This feature also makes H.264 Intra-frame coding better than the JPEG2000 still image compression standard [7]. To achieve high coding efficiency, the intra prediction in H.264 employs the rate-distortion optimisation (RDO) technique [7-8]. In order to perform RDO, the encoder encodes video by exhaustively searching the best mode in the RD sense among various predefined modes. As a result, the computational complexity of the H.264 encoder is dramatically increased, which makes it difficult for practical applications such as real-time video communications.

Several attempts have been made to address the complexity bottleneck of the intra mode decision [1-6]. Pan et al. [3-4] proposed a fast mode decision scheme based on pre-processing, which measures the average edge direction of a given block so as to reduce the number of probable modes to achieve complexity reduction. The overall performance is 20~30% faster than the exhaustive RDO method at the cost of 2% extra bits and 0.024dB of PSNR degradation, and 55~65% faster at the cost of 5% extra bits. T. Qian et al. [1] and J. Xin [2] exploits simplification of transformation, which achieves about 13% of time saving with 0.01~0.15dB of PSNR degradation. Another further improved approach [5] based on context-based decimation of unlikely candidates, subsampling of matching operation and bit-width truncation shows about 45% of time saving and 0.3dB PSNR degradation.

In this work, we present a simple yet effective fast intra mode decision algorithm for a H.264 codec, combining two methods. The first method, the so-called two-level fast Intra4x4 mode decision, is based on the idea of reducing possible candidate directions of a 4x4 luma block prediction in Intra4x4 mode decision. The second method enables an early selection of suitable 4x4 or 16x16 luma partitions of a given macroblock. The performance of the proposed algorithms is evaluated using the H.264 reference codec JM 10.1 with exhaustive RDO mode decision. Metrics are the encoding time, the average PSNR and the coding bit-rate for test sequences.

The rest of this paper is organised as follows. The intra mode prediction for H.264/AVC is reviewed in Section 2. The proposed mode decision algorithms are detailed in Section 3. Finally, experimental results and conclusions are given in Sections 4 and 5, respectively.

2. INTRA MODE DECISION FOR H.264

In the section, we review the H.264/AVC intra mode prediction scheme and analyse the computational complexity of the exhaustive search scheme. Intra prediction in H.264 exploits the spatial correlation between adjacent macroblocks.

Two different types of intra prediction are possible for the prediction of the luminance component Y. The first type is called Intra4x4 (I4MB) and the second one Intra16x16 (I16MB).

Using the Intra16x16 type, the macroblock, which is of the size 16 by 16 picture elements (16x16), is divided into sixteen 4x4 subblocks and a prediction for each 4x4 subblock of the luminance signal is applied individually. For prediction purposes, nine different prediction modes are supported. One mode is the DC prediction mode, where all samples of the current 4x4 subblock are predicted by the mean of a set of samples. The set of samples consists of pixels neighbouring to the left and to the top of the current block and already reconstructed at the encoder and at the decoder side.

In addition to DC-prediction mode, eight prediction modes for a specific prediction direction are supported. All possible directions are shown in Fig. 1.

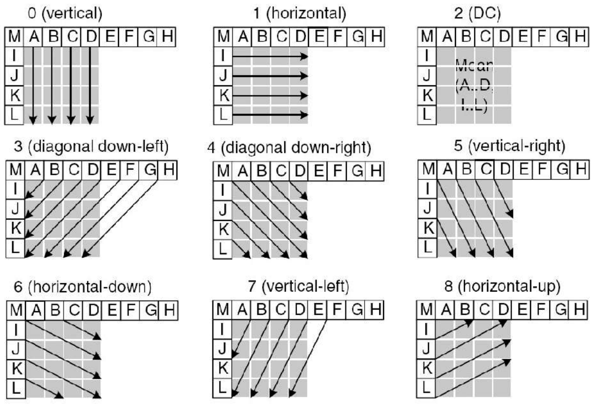

Fig. 1. Nine prediction modes of 4x4 luma components (from [8]).

Using the type Intra16x16, only one out of four prediction modes is applied to the whole macroblock. The different prediction modes supported for the type Intra16x16 are vertical prediction, horizontal prediction, DC-prediction and plane-prediction. Hereby plane-prediction uses a linear function between the neighbouring samples to the left and to the top in order to predict the current samples. This mode works very well in areas of a gently changing luminance. The other prediction modes are the same as the ones of the 4x4 prediction. The only difference is that they are applied to the whole macroblock instead of to a 4x4 subblock.

The intra prediction for the chrominance components Cb and Cr of a macroblock is similar to the Intra16x16 type for the luminance component because the chrominance signals are very smooth in most cases. It is performed always on 8x8 blocks using vertical prediction, horizontal prediction, DC prediction and plane prediction.

A H.264 encoder mainly uses a rate-distortion optimisation (RDO) to obtain the best prediction mode by selecting the smallest rate-distortion cost. The RD cost for the best intra mode is computed using the following:

$$J(s,c,MODE \mid QP) = SSD(s,c,MODE \mid QP) + \lambda_{MODE} \cdot R(s,c,MODE \mid QP) . \quad (1)$$

λ_{MODE} is Lagrangian multiplier and *SSD* is the sum of the squared differences between the original block *s* and its reconstruction *c* given as

$$SSD(s,c,MODE \mid QP) = \sum_{x=1,y=1}^{16,16} \{s_y(x,y) - c_y(x,y,MODE \mid QP)\}^2 + \sum_{x=1,y=1}^{8,8} \{s_u(x,y) - c_u(x,y,MODE \mid QP)\}^2 + \sum_{x=1,y=1}^{8,8} \{s_v(x,y) - c_v(x,y,MODE \mid QP)\}^2 \quad (2)$$

and R(s*, c, MODE* | *QP*) is the number of bits associated with the best intra mode currently selected for the macroblock and includes the bits for the macroblock header and the bits for the transformed y, u and v blocks. $c_y(x, y, MODE \mid QP)$ and $s_y(x, y)$ represent the reconstructed and original luminance values; c_u, c_v and s_u, s_v the corresponding chrominance values.

According to the exhaustive procedure [7] of intra prediction in JM 10.1, the number of mode combinations for luma and chroma blocks in a macroblock is N8x(N4x16+N16), where N8, N4 and N16 represent the number of modes for 8x8 chroma blocks, 4x4 and 16x16 luma blocks, respectively. This means that, for the intra coding of a macro block in H.264/AVC, it has to perform 592 different RDO calculations to determine the optimal RDO mode. As a result, the complexity of the encoder is extremely high. To reduce the encoding complexity with little RD performance degradation, new intra mode decision is proposed in the next section.

3. PROPOSED FAST INTRA MODE DECISION

In this Section, we propose a simple yet effective fast intra mode decision algorithm for H.264 codec, combining two methods. The first method, the so-called two-level fast Intra4x4 mode decision, is based on the idea of reducing possible candidate directions of a 4x4 luma block in Intra4x4 mode decision. The second method enables an early selection of the Intra4x4 type or Intra16x16 type prediction for a given macroblock.

3.1 Two-level Fast Intra Mode Decision for 4x4 Intra Mode

Fig. 2 gives the block diagram for the Two-level mode decision scheme applied to 4x4 luma blocks. The first-level is to select the two modes with the smallest RD cost among the five modes DC

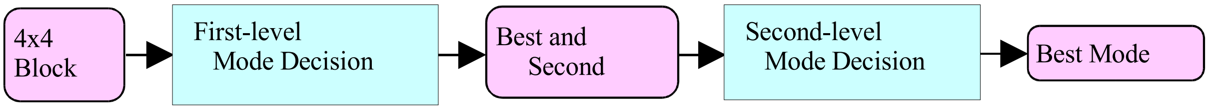

Fig. 2. Block diagram of Two-level intra mode decision.

prediction, vertical prediction, horizontal prediction, diagonal down-left and diagonal down-right prediction.

Then, for the second-level, the encoder selects the best mode among one or two modes which are selected based on the spatial relationship between the two prediction modes from the first-level mode decision. The following spatial relationships between the two best prediction modes are possible (see Fig. 3).

Case 1: The best and the second best prediction mode from the first-level mode decision are neighbours.
Case 2: The best prediction and the second best prediction mode are not neighbours.
Case 3: One of the two prediction modes is the DC prediction mode.

For case 1, the second-level mode decision computes the RD cost of the prediction mode (green-colored mode in Fig. 3) between the best mode (red-colored) and the second best mode (blue-colored) of the first-level mode decision. This cost is compared with the best mode from the first-level mode decision. For case 2, RD costs of the two modes (green-colored modes) around the best mode (red-colored mode) from the first-level mode decision are compared with the best mode of the first-level mode decision. For case 3, two modes around the mode not being the DC mode are compared with the best mode of the first-level mode decision. As a result, RD costs are computed for 6 to 7 out of 9 modes.

3.2 Early Selection of Suitable 4x4 or 16x16 Luma Partitions of a Given Macroblock

As can be seen in Section 2, the exhaustive RDO mode decision is performed by calculating RD cost for all possible prediction modes of luma and chroma. The RD cost of a given macroblock for a chroma prediction mode *uv* and a luma prediction mode *y* is denoted as *J*(*uv*, *y*). The RD cost of the best intra mode obtained by exhaustive mode decision [7] can be computed as follows:

$$\min_{uv \in UV} \min_{y \in Y} \{J(uv, y)\} \quad (3)$$

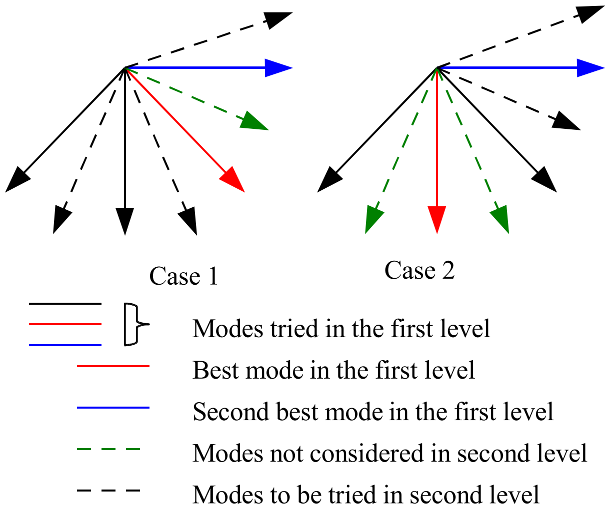

Fig. 3. Possible spatial relationships between the best mode and the second best mode in first-level mode decision.

with UV = {DC, Vertical, Horizontal, Plane} and Y = {I4MB, I16MB}. I4MB is the best combination of 4x4 modes for a macroblock and I16MB is the best 16x16 mode. The above formula implies that the encoder firstly selects the best luma mode, which is one of the two luma types, Intra4x4 and Intra16x16 mode, for a given chroma mode. This procedure is repeated for all possible chroma modes. If we can find in advance which luma type is the most suitable for luma prediction, we could save a lot of computation for intra mode decision.

From the experiment on the statistics of RD costs of Intra4x4 and Intra16x16, we observed the followings to be mostly true: If the rate-distortion cost of coding the luminance of a macroblock in I16MB type is lower than I4MB, then the overall rate distortion cost of a macroblock is lowest for I16MB independent of the selected chrominance prediction mode. In other words, if the best luma mode is 16x16 block type, then the following was most often observed:

$$J(uv, I4MB) > J(uv, I16MB), \forall uv \in UV$$

This means that we need only one luma mode decision for a first chroma mode and then we can select the best intra mode by repeating RD calculation only for each chroma mode. The proposed early selection of a suitable luma type of a given macroblock is shown in the flow chart in Fig. 4. As can be seen, in a first step, the encoder computes two RD costs for 4x4 and 16x16 typess with the chrominance mode set to DC.

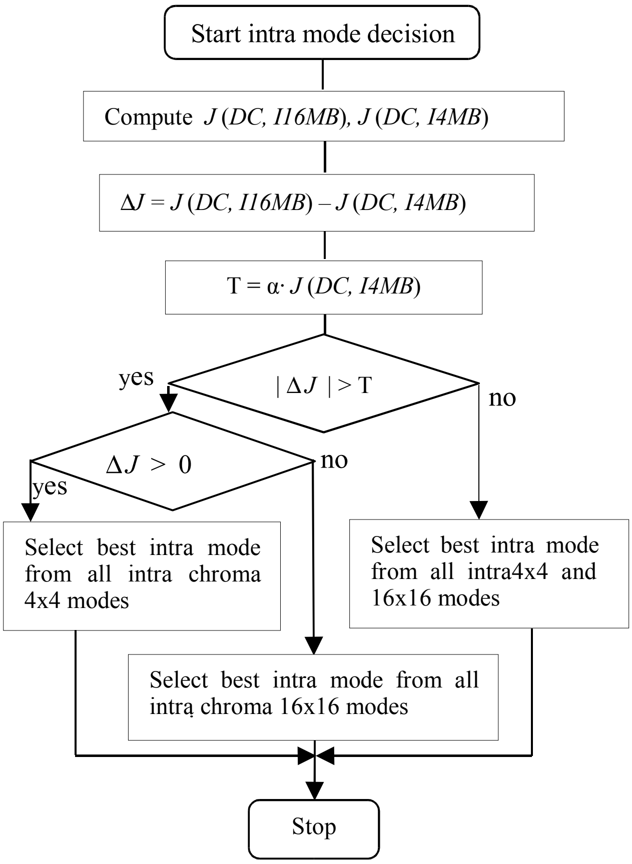

Fig. 4. Flow chart of intra mode decision by early selection of suitable partition size.

If the difference between the RD costs of two luma prediction types, Intra4x4 and Intra16x16, is relatively small, the probability to correctly select the actual best luma mode is lower due to the uncertain chroma mode. Thus, let us define a threshold T

$$T = \alpha \cdot J(DC, I4MB) \quad (4)$$

where α is a constant which is experimentally determined to be in the range of 0.03~0.05. If the following condition

$$|J(uv, I16MB) - J(uv, I4MB)| > T \quad (5)$$

holds, the encoder tries intra mode decision only for the luma type with smaller RD cost, otherwise, it tries for both luma types, Intra4x4 and Intra16x16.

4. EXPERIMENTAL RESULTS

The proposed algorithms were implemented into JM 10.1 provided by JVT for the performance evaluation. JM 10.1 was used in the baseline profile and its test conditions are as follows:

a. Frame structure is IIIII…
b. The number of frames to be encoded is 100 frames.
c. RD optimisation is enabled.
d. CABAC is enabled.
e. Video sequences are QCIF format including “foreman”, “stefan”, “mobile”, “mother and daughter”, “coastguard” and “container”.
f. The quantisation parameters (QP) are varied from 26 to 30.

The proposed algorithms are compared with the exhaustive RDO intra mode decision of H.264 in terms of decrease of execution times (TS), the decrease of YUV PSNR (ΔPSNR) and the increase of bit rate (+Bits).

In experiments about 74% of all macroblocks were coded using the preselection of 4x4 mode, 10% using the preselection of 16x16 mode. On average all prediction modes were evaluated for only 16% of the macroblocks.

When T_REF and T_IMP are total time for intra mode decision in the exhaustive method and in the proposed method, respectively, TS which indicates time saving of the improved method is denoted as

$$TS = \frac{T\_REF - T\_IMP}{T\_REF} \times 100\% . \quad (6)$$

Table 1 shows that the first method based on two-level intra mode decision can achieve 16.6% of average time savings for intra mode decision with 0.03dB of PSNR degradation and 1.33% of extra bits. But the second method (early selection of suitable luma type) shows an excellent performance compared with all previous proposals achieving about 50% of time saving with little or no PSNR degradation and few extra bits. The combinated method also shows good results with 52% of time saving and 0.03dB of PSNR degradation (see Fig. 5), while the Pan et al. [3] proposal based on pre-processing makes about 22% faster than the exhaustive RDO method at the cost of 2% extra bits. T. Qian et al. [1] and J. Xin [2] achieves about 13% of time saving with 0.01~0.15dB of PSNR degradation. Another further improved approach [5] based on context-based decimation of unlikely candidates, sub-sampling of matching operation, and bit-width truncation, shows about 45% of time saving and 0.3dB PSNR degradation.

Table 1. Comparison of proposed methods

Sequence	QP	Method1 (Two-level decision)			Method2 (Early selection of partition)			Combination			Alternative methods			
											J.Xin et al[2]		Pan et al [3]	
		ΔPSNR (dB)	+Bits (%)	TS (%)	ΔPSNR (dB)	+Bits (%)	TS (%)	ΔPSNR (dB)	+Bits (%)	TS (%)	ΔPSNR (dB)	TS (%)	ΔPSNR (dB)	TS (%)
mobile	30	0.041	0.77	15.0	0.001	0.52	40.4	0.041	1.23	46.5	0.01	11	-	-
	28	0.05	0.68	21.4	0.001	044	50.3	0.051	1.09	52.8	0.005	6	-	-
	26	0.054	0.61	15.4	0	0.43	46.4	0.055	1.00	49.8	0.005	9	-	-
stefan	30	0.059	0.63	13.2	0.002	1.18	44.6	0.059	1.73	46.2	-	-	-	-
	28	0.066	0.54	8.64	0.001	1.12	39.9	0.069	1.60	44.1	-	-	-	-
	26	0.064	0.53	17.8	0.002	1.11	42.6	0.065	1.62	46.4	-	-	-	-
foreman	30	0.014	2.99	22.1	0	0.9	57.4	0.013	3.85	59.7	0.01	14	0.032	22.1
	28	0.004	2.68	16.1	0.001	0.8	56.4	0.004	3.44	58.1	0.01	13	0.008	23.5
	26	0.005	2.43	17.7	0.002	0.78	54.9	0.007	3.16	57	0.01	12	-	-
mother & daughter	30	0.016	1.92	18.4	0.002	2.14	53.0	0.021	3.97	58.3	0.01	14	-	-
	28	0.036	2.26	19.2	0.006	1.95	47.1	0.037	4.11	53.7	0.005	12	-	-
	26	0.021	2.13	29.5	0.013	1.77	52.6	0.037	3.74	56.9	0.005	12	-	-
container	30	0.023	0.96	13.4	0.002	2.65	47.2	0.025	3.53	49.3	0.01	13	0.066	22.9
	28	0.012	1.02	10.2	0.003	2.39	42.1	0.014	3.34	45.1	0.01	13	0.02	24.1
	26	0.015	1.00	15.1	0.002	2.06	41.8	0.016	3.00	48.1	0.01	7	-	-
coastguard	30	0.017	0.98	17.8	0.001	0.47	52.9	0.017	1.40	58.4	-	-	0.001	19.4
	28	0.015	0.95	13.2	0	0.43	53.9	0.015	1.37	57.8	-	-	0.018	21.0
	26	0.021	0.83	15.1	0	0.48	47.9	0.021	1.26	52	-	-	-	-
average		0.03	1.33	16.6	0.0022	1.2	48.4	0.0315	2.47	52.2	0.01	11.3	0.024	22.2

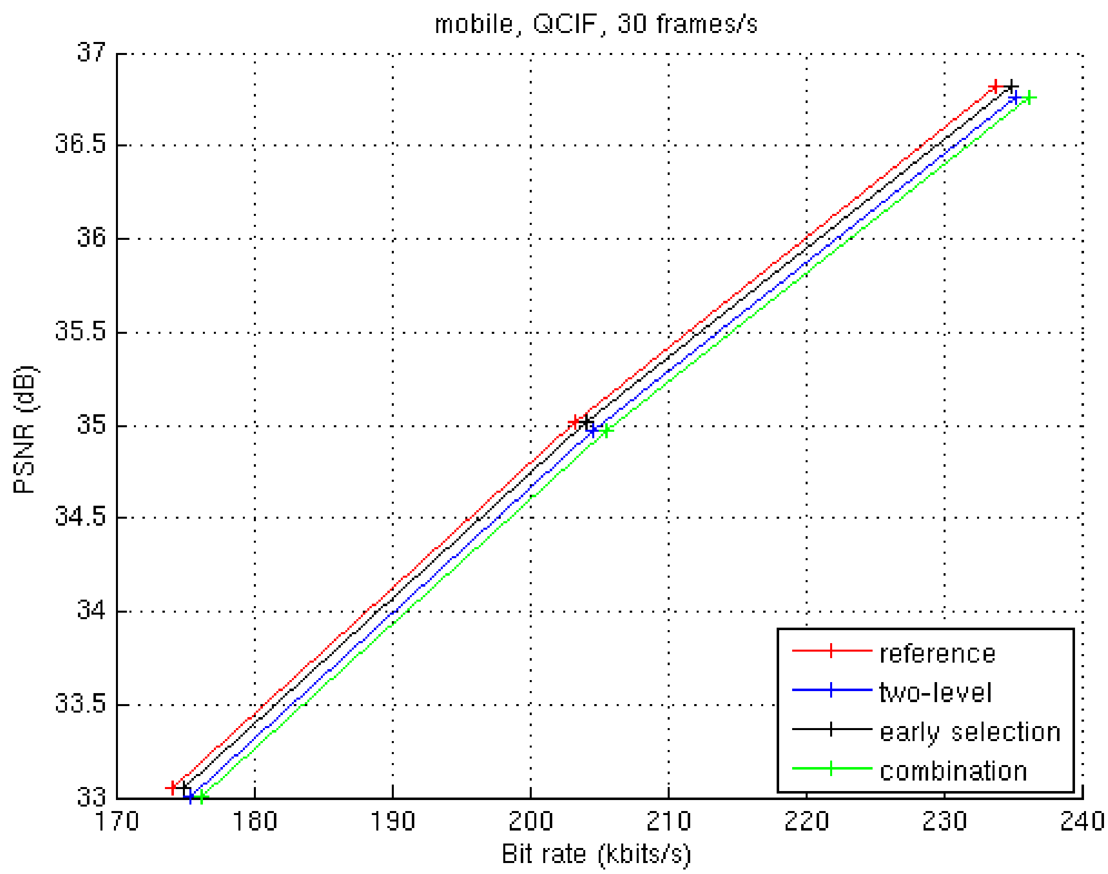

Fig. 5. Rate distortion curve of our proposals and reference method.

Table 2. Comparison of Number of RD calculation for a Macroblock

	JM 10.1	Pan et al [3, 4]	Proposal
4x4 luma block	9	4	2.12
16x16 luma block	4	2	0.84
8x8 chroma block	4	2.23	4
Number of RD calculations for a MB	592	147.18	142.88

Additionally, the average number of RD calculations for a macroblock is compared in Table 2 and the comparison of rate-distortion curves is shown in Fig.5.

5. CONCLUSIONS

In this work, we propose two methods for fast intra mode decision. The performance of the proposed algorithms are evaluated using the H.264 reference codec JM 10.1, which is compared with exhaustive RD mode decision of H.264 in metrics such as the encoding time, the average PSNR and the coding bit-rate for test sequences. It is observed that the proposed fast algorithms can save about 52% of encoding time resulting in a minor of degradation of 0.03dB lower PSNR and a 2.5% higher bit rate.

REFERENCES

[1] T. Qian, S. Jun and W. Chen, “Fast Mode Decision Scheme for Transcoding H.264 Intra Coded Frames with Spatial Resolution Reduction”, Proceedings of 2005 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), Philadelphia, Pennsylvania, March 2005.

[2] J. Xin, A. Veto, and H. Sun, “Efficient Macro block Coding-Mode Decision for H.264/AVC Video Coding”, Picture Coding Symposium 2004 (PCS 2004), San Francisco, USA, December 2004.

[3] F. Pan, X. Lin, and R. Susanto, “Fast Mode Decision for Intra Prediction”, JVT-G013, JVT 7th meeting, Pattaya, Thailand, 7-14 March, 2003.

[4] F. Pan, X. Lin, S. Rahardja, K. P. Lim, Z. G. Li, and D. Wu. S. Wu, “Fast Intra Mode Decision Algorithm for H.264/AVC Video Coding”, 2004 International Conference on Image Processing (ICIP), Singapore, 2004.

[5] P. Zhang, D. Zhao, S. Ma, and Y. Lu, “Multiple Modes Intra-Prediction in Intra Coding”, The 2004 IEEE International Conference on Multimedia and Expo (ICME 2004), Taipei, Taiwan, Jun. pp. 27-30, 2004.

[6] Yu-Wen Huang, Bing-Yu Hsieh, Tung-Chien Chen, and Liang-Gee Chen, “Analysis, Fast Algorithm, and VLSI Architecture Design for H.264/AVC Intra Frame Coder”, IEEE Trans. on Circuits and Systems for Video Technology, Vol. 15, No. 3, March, 2005, pp. 378 - 401.

[7] J. Ostermann, J. Bormans, P. List, D. Marpe, M. Narroschke, F. Pereira, T. Stockhammer, and T. Wedi, “Video Coding with H.264/AVC: Tools, Performance, and Complexity”, Circuits and Systems Magazine, IEEE, Vol. 4, No. 1, 2004, pp. 7-28.

[8] Iain E. G. Richardson, “H.264 and MPEG-4 Video Compression”, John Wiley & Sons, NJ, USA, 2003.