

Selective Intra Prediction Mode Decision for H.264/AVC Encoders

Jun Sung Park, and Hyo Jung Song

Abstract—H.264/AVC offers a considerably higher improvement in coding efficiency compared to other compression standards such as MPEG-2, but computational complexity is increased significantly. In this paper, we propose selective mode decision schemes for fast intra prediction mode selection. The objective is to reduce the computational complexity of the H.264/AVC encoder without significant rate-distortion performance degradation. In our proposed schemes, the intra prediction complexity is reduced by limiting the luma and chroma prediction modes using the directional information of the 16×16 prediction mode. Experimental results are presented to show that the proposed schemes reduce the complexity by up to 78% maintaining the similar PSNR quality with about 1.46% bit rate increase in average.

Keywords—Video encoding, H.264, Intra prediction.

I. INTRODUCTION

THE technically aligned specification of ITU-T Recommendation H.264 and ISO/IEC MPEG4-AVC (Advanced Video coding) [1], abbreviated as H.264/MPEG4-AVC, is the state-of-the-art video coding technique to meet a wide range of applications. H.264/AVC offers a dramatic performance improvement over previous video coding standards such as H.263++ and MPEG-4. H.264/AVC provides gains in compression efficiency of up to 50% over a wide range of bit rates and video resolutions compared to previous standards [2]. Among many new features, the intra prediction technique is regarded as one of the important features that contribute to the success of H.264/AVC [3].

However, the H.264 encoder has the significant computational complexity because it selects the best coding mode by employing the rate distortion optimization (RDO) to take full advantage of the mode selection in terms of maximizing coding quality and minimizing data bits [2]. The RDO mode decision exhaustively searches every possible mode for each macroblocks (MB) to find the mode having the minimum rate-distortion cost. This optimization technique has extremely high complexity so that it is difficult to implement the encoder in real-time applications such as video telephony and video conferencing. A few attempts have been made to address this computational complexity of intra mode decision in H.264/AVC. Zhang et al. [4] proposed an intra mode decision scheme which reduces candidate modes of intra prediction with calculated local edge direction within 4×4 block. This algorithm is required additional pixel calculation to

get the parameters for edge information. They save computational time up to 70% but the bitrate is increased by 5.5%. Moreover, the performance results are varied depending on the test sequences. An irregular performance can affect the QoS in rate control of H.264/AVC. Pan et al. [5] have also proposed an intra mode decision scheme. They calculated edge direction histogram using neighboring pixels to predict the primary prediction mode. With the primary prediction mode, only small number of intra prediction modes is used for intra coding. However, they reduced the complexity of intra coding by more than 75% but computation time saving is 60% on average because additional computations for edge direction histogram are required. As the prediction error is increased, the bitrate is increased about 3.7% and PSNR loss is about 0.24dB.

In this work, we propose a simple yet effective fast mode decision algorithm for H.264 intra prediction. Based on the result of the best prediction mode of the 16×16 luma block, only a small number of the likely luma and chroma intra prediction modes are chosen for RDO calculation. The rest of this paper is organized as follows: after a brief overview of the H.264 intra mode decision in Section II, we describe the proposed selective intra mode decision in Section III. In Section IV, we provide experimental results to show the performance of the proposed scheme. Finally, we give a conclusion in section V.

II. OVERVIEW OF INTRA PREDICTION FOR H.264

The intra prediction reduces spatial redundancies by exploiting the spatial correlation between adjacent blocks in a given picture. Each picture is divided into 16×16 pixel MB and each MB is formed with luma components and chroma components. For luma components, the 16×16 pixel MB can be partitioned into sixteen 4×4 blocks. The chroma components are predicted by 8×8 blocks with a similar prediction technique as the 16×16 luma prediction. There are 9 prediction modes for 4×4 luma blocks and 4 prediction modes for 16×16 luma blocks. For the chroma components, there are 4 prediction modes that are applied to the two 8×8 chroma blocks (U and V).

A. 4×4 Luma Intra Prediction Modes

The MB is divided into sixteen 4×4 luma blocks and a prediction for each 4×4 luma block is applied individually. 4×4 Luma Intra Prediction Modes are well suited for coding of parts of a picture with a significant detail. The nine different prediction modes are supported as shown in Fig. 1, where the prediction values for pixels are calculated from the neighboring boundary pixel values. Each mode is suitable to predict directional structures in the picture at different angles (See Fig. 2). Note that DC is a special prediction mode, where the mean of the left handed and upper samples are used to predict the entire block. Normally DC prediction mode is useful for those blocks with little or no local activities.

Recently, a new profile of H.264, Fidelity Range Extensions (FRExt) was introduced. It is motivated by the rapidly growing demand for coding of higher-fidelity video materials [6]. In the FRExt amendment, an additional intermediate prediction block size of 8×8 was introduced for spatial luma prediction by extending the concepts of 4×4 intra prediction in an effort to improve coding efficiency. For the 8×8 intra prediction, 9 prediction modes are used which are the same as that of 4×4 intra prediction. However, the computational complexity of the H.264 encoder is dramatically increased according to this feature of the new extended profile.

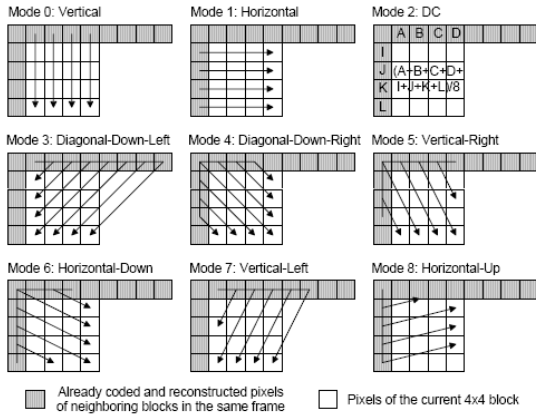


Fig. 1 Nine modes of 4×4 intra prediction in H.264/AVC

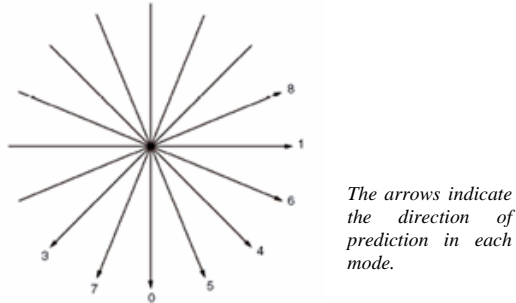


Fig. 2 Directions of 4×4 intra prediction modes in H.264/AVC

B. 16×16 Luma Intra Prediction Modes

16×16 Luma Intra Prediction Modes are more suitable for coding very smooth areas of a picture by prediction for the whole luma component of a MB. Four different prediction modes are supported: Vertical, Horizontal, DC and Plane prediction. Plane prediction mode uses a linear function between the neighboring samples to the left and to the top in order to predict the current samples.

C. 8×8 Chroma Intra Prediction Modes

The chroma intra prediction of a MB is similar to the 16×16 luma intra prediction because the chroma signals are very smooth in most cases. It is performed always on chroma blocks

using vertical prediction, horizontal prediction, DC-prediction or plane-prediction.

H.264/AVC encodes the MB by iterating all the luma intra decisions for each possible chroma intra prediction mode for the best coding efficiency. Therefore, the number of mode combinations for luma and chroma components in an MB is $C8 \times (L4 \times 16 + L16)$, where C8, L4, and L16 represent the number of modes for chroma prediction, 4×4 luma prediction and 16×16 prediction, respectively. It means that, for an MB, it has to perform $4 \times (9 \times 16 + 4) = 592$ different RDO calculations before a best RDO mode is determined [5][7]. If the 8×8 luma prediction of H.264 FRExt is included, the number of mode combinations is $C8 \times (L4 \times 16 + L8 \times 4 + L16) = 4 \times (9 \times 16 + 9 \times 4 + 4) = 736$.

III. PROPOSED FAST INTRA MODE DECISION

In this section, we propose *selective intra prediction mode decision* schemes for fast intra mode decision. The key idea of our schemes stems from the fact that the dominating direction of a bigger block is similar to that of smaller block. As in Fig. 3, the best prediction mode of 4×4 luma block within 16×16 block has the same direction as that of 16×16 luma block. The computation of the intra prediction and the chroma prediction can be reduced on the base of the overall edge information from the 16×16 intra prediction result.

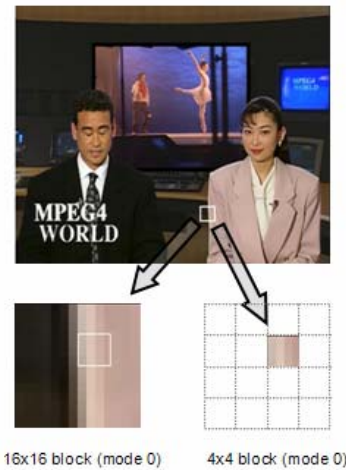


Fig. 3 Similarity in mode prediction of 16×16 block and 4×4 block

A. 4×4 Luma Intra Prediction Modes

We assort 9 modes of the 4×4 block to make four candidate groups according to the directional information of the 16×16 prediction mode. Therefore the unlikely modes are filtered out prior to the RD cost computation based on the directional correlation between 16×16 luma block and 4×4 luma block. The 16×16 intra prediction mode has 4 modes (vertical, horizontal, DC, Plane). The mode 0 (vertical) and mode 1 (horizontal) of 16×16 block have the same direction of prediction as each of mode 0 (vertical) and mode 1 (horizontal) of 4×4 block. Mode 2 (DC) of 16×16 block and mode 2 (DC) of 4×4 block do not have directions and they use the mean value of adjacent blocks. The mode 3 (plane) of 16×16 block is used for a linear function between the neighboring left and top block, so its direction of prediction is similar to the mode 3 (diagonal down-left) of the 4×4 block. Based on the above the

directional information, 9 modes of the 4×4 intra prediction are formed into each of the candidate group according to the directional information of the already calculated result of 16×16 intra prediction mode under the following rules:

- If a 16×16 prediction mode is mode 0 (vertical), candidate modes of the 4×4 block are mode 0 (vertical) and the adjacent 2 modes (mode 7, 5);
- If a 16×16 prediction mode is mode 1 (horizontal), candidate modes of the 4×4 block are mode 1 (vertical) and the adjacent 2 modes (mode 8, 6);
- If a 16×16 prediction mode is mode 2 (DC), the fundamental modes (mode 0, 1, 3, 4) are chosen as candidate modes of the 4×4 block;
- If a 16×16 prediction mode is mode 3 (plane), candidate modes of 4×4 block are mode 3 (diagonal down-left), mode 0 (vertical) and mode 1 (horizontal);
- Since DC prediction has a higher possibility to be the best prediction mode out of the 9 modes [4], each candidate group has mode 2 (DC);
- Since the mode of block U (upper) and the mode of block L (left) have a spatial-correlation with encoding block (See Fig. 4), each candidate group has the mode of block U and the mode of block L.

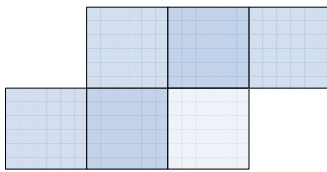


Fig. 4 Current and neighboring blocks (U: upper, L: left, C: current)

According to the criteria described above, we can determine the candidate groups as shown in Table I.

TABLE I
CANDIDATES 4×4 MODES ACCORDING TO 16×16 MODE

16×16 mode	candidate 4×4 modes
mode 0 (vertical)	mode 7, 0, 5, 2, mode of U, L
mode 1 (horizontal)	mode 8, 1, 6, 2, mode of U, L
mode 2 (DC)	mode 0, 1, 3, 4, 2, mode of U, L
mode 3 (plane)	mode 0, 1, 3, 2, mode of U, L

B. 8×8 Luma Intra Prediction Modes

Aforementioned scheme is especially effective in H.264/AVC FRExt because when the 8×8 block is applied in intra prediction, the spatial gap between bigger (16×16) and smaller (4×4) blocks is reduced.

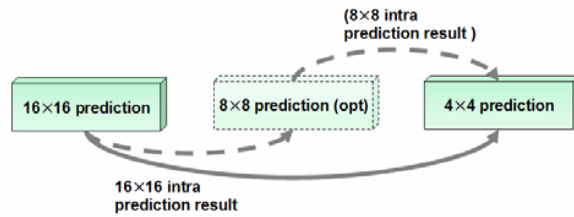


Fig. 5 The sequence of applying prediction result

Selective intra prediction mode decision scheme for 8×8 luma prediction is composed of two steps as shown below:

- Step 1: the result of 16×16 prediction mode reduces nine 8×8 intra prediction modes to candidate groups described in Table II below;
- Step 2: after step 1, the result of 8×8 prediction mode reduces nine 4×4 intra prediction modes to candidate groups described in Table III below.

TABLE II
CANDIDATES 8×8 MODES ACCORDING TO 16×16 MODE

16×16 mode	candidate 8×8 modes
mode 0 (vertical)	mode 7, 0, 5, 2, mode of U, L
mode 1 (horizontal)	mode 8, 1, 6, 2, mode of U, L
mode 2 (DC)	mode 0, 1, 3, 4, 2, mode of U, L
mode 3 (plane)	mode 0, 1, 3, 2, mode of U, L

TABLE III
CANDIDATES 4×4 MODES ACCORDING TO 8×8 MODE

8×8 mode	candidate 4×4 modes
mode 0 (vertical)	mode 7, 0, 5, 2, mode of U, L
mode 1 (horizontal)	mode 8, 1, 6, 2, mode of U, L
mode 2 (DC)	mode 0, 1, 3, 4, 2, mode of U, L
mode 3 (diag down-left)	mode 8, 3, 7, 2, mode of U, L
mode 4 (diag down-right)	mode 5, 4, 6, 2, mode of U, L
mode 5 (vert-right)	mode 0, 5, 4, 2, mode of U, L
mode 6 (hor-down)	mode 4, 6, 1, 2, mode of U, L
mode 7 (vert-left)	mode 3, 7, 0, 2, mode of U, L
mode 8 (horizontal-up)	mode 1, 8, 3, 2, mode of U, L

C. 8×8 Chroma Intra Prediction Modes

Selective intra prediction mode decision scheme for 8×8 chroma intra prediction modes is proposed for the enhancement of chroma intra prediction. Even though the luma block and the chroma block are from luminance signals and chrominance signals separately [9], they encode the same section of image, 16×16 pixel MB, and share overall directional information. We narrow down the number of chroma prediction modes from 4 modes to 2 modes, according to the best prediction mode of the 16×16 luma block.

First, we set the candidate mode of chroma prediction to DC. From the observation, DC mode has the highest probability of winning the prediction mode selection (See Fig. 6).

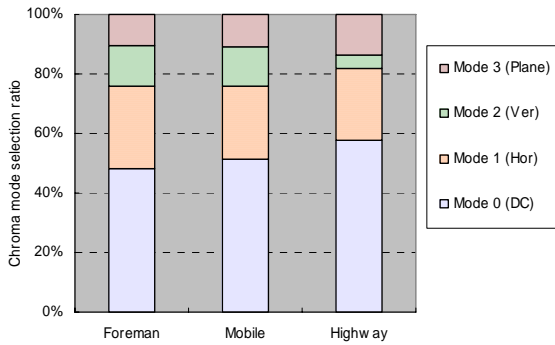


Fig. 6 Chroma mode selection ratio chart

Then, the second candidate mode of chroma prediction is set to the best mode of 16×16 luma intra prediction. If the best mode of 16×16 luma prediction is DC, then the candidate mode is DC mode only.

The following is a summary of algorithm:

Step 1: Find the best prediction mode of 16×16 luma prediction;

Step 2-a: If the best prediction mode of 16×16 luma prediction is DC mode, then the candidate mode for chroma block is DC mode only;

Step 2-b: If the best prediction mode of 16×16 luma prediction is other than DC mode, then the candidate modes for chroma prediction are DC mode and the best prediction mode of 16×16 luma prediction. For example, when the best prediction mode of 16×16 luma prediction is Vertical (mode 0), the candidate modes for chroma prediction are DC mode and Vertical (mode 2).

Fig. 7 shows the ratio that the best prediction mode of chroma block is either DC mode or the same mode as the best prediction mode of luma block. The average of hit ratio is over 85% with various sequences and it can be said that two candidate modes are efficiently enough to find the best prediction mode of chroma block.

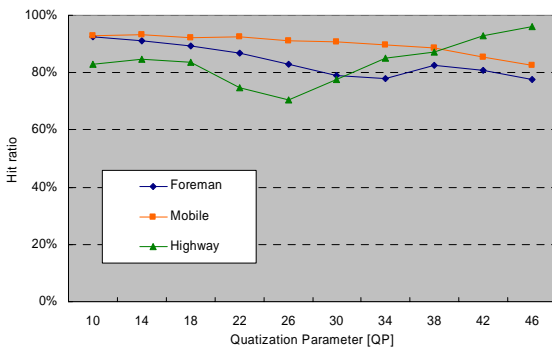


Fig. 7 The hit ratio of chroma candidate modes

D. Analysis of Computational Complexity

The number of candidate modes from *selective intra prediction mode decision* is tabulated in Table IV. With our proposed algorithm, the number of mode combinations for an MB is only $1 \times (4 \times 16 + 4) = 68$ at the best case by the section

II, whereas the current RDO calculation in H.264/AVC requires 592.

In case that the 8×8 luma block for H.264/AVC FRExt is applied, the mode combination complexity is reduced to $1 \times (4 \times 16 + 4 \times 8 + 4) = 100$. Thus our proposed algorithm reduces number of RDO calculation significantly compared to the $4 \times (9 \times 16 + 9 \times 8 + 4) = 880$ modes that are used in the current RDO calculation in H.264/AVC FRExt video coding.

TABLE IV
NUMBER OF CANDIDATE MODES

	Block size	Total number of modes	Number of candidate modes
Luma (Y)	4×4	9	4 to 7
Luma (Y)	8×8^1	9	4 to 7
Luma (Y)	16×16	4	4
Chroma (U, V)	8×8	4	1 to 2

IV. SIMULATION RESULTS

Selective intra mode decision schemes as described in Section III have been integrated with the reference software called JM [10] for a performance evaluation. The comparison results with JM encoder were examined based on the difference of the computational time, the PSNR and the bitrate for various sequences. The system platform is the Intel Pentium 4 Processor of speed 2.66GHz, 512MB DDR RAM, and Microsoft Windows XP. For each test sequence, 50 frames were encoded with I-frame coding only. The quantization parameter set was chosen to be [10, 14... 42, 46]. With this set of quantization parameters, the performance results were tabulated in Table V and VI based on the change of average luma and chroma PSNR (Y-PSNR, UV-PSNR), the change of average bitrate (Bitrate) and the change of average coding time (Time). The 8×8 block for intra prediction enabled results can be shown in Table VI. Note that in the tables, positive values mean increments and negative values mean decrements. The results with the 8×8 block enabled show that the proposed schemes reduced execution time up to 78% with only an average of 0.073 dB losses in Y-PSNR, 0.077 dB losses in UV-PSNR and 1.46% increments in bitrate.

The proposed schemes achieve faster encoding in intra prediction compared to the JM with very little RD performance degradation. Fig. 8 shows the RD performance and the computation time for the Highway (QCIF) sequence with 8×8 luma block enabled. The two RD curves, one from the JM and the other from the proposed schemes are nearly overlapping each other and it means that the selective intra prediction algorithm performs almost the same as that of the JM in terms of PSNR and data bits with saving computation time.

¹ 8×8 luma block is supported by H.264/AVC FRExt

TABLE V
EXPERIMENTAL RESULTS

Sequences	Y-PSNR (dB)	UV-PSNR (dB)	Bitrate (%)	Time (%)
Foreman (QCIF)	-0.072	-0.040	+3.38	-73.68
Mobile (QCIF)	-0.120	-0.036	+1.59	-75.38
Highway (QCIF)	-0.050	-0.082	+1.31	-73.50
Mother (QCIF)	-0.075	-0.110	+4.27	-74.11
Salesman (QCIF)	-0.081	-0.081	+2.33	-75.00
Bus (CIF)	-0.090	-0.114	+2.09	-76.85
Waterfall (CIF)	-0.090	-0.002	+1.55	-74.64
Silent (CIF)	-0.070	-0.040	+2.33	-75.04
Flower (CIF)	-0.120	-0.062	+1.59	-75.29

TABLE VI
EXPERIMENTAL RESULTS (WITH 8×8 LUMA BLOCK ENABLED)

Sequences	Y-PSNR (dB)	UV-PSNR (dB)	Bitrate (%)	Time (%)
Foreman (QCIF)	-0.060	-0.049	+1.77	-75.50
Mobile (QCIF)	-0.053	-0.035	+1.25	-75.87
Highway (QCIF)	-0.061	-0.145	+0.69	-76.00
Mother (QCIF)	-0.091	-0.110	+2.76	-75.56
Salesman (QCIF)	-0.063	-0.081	+1.45	-76.18
Bus (CIF)	-0.078	-0.132	+1.74	-78.14
Waterfall (CIF)	-0.092	-0.008	+0.90	-75.95
Silent (CIF)	-0.098	-0.058	+1.52	-76.46
Flower (CIF)	-0.069	-0.080	+1.07	-76.19

V. CONCLUSION

In this paper, a fast mode decision in H.264/AVC encoders is proposed by reducing the number of candidate modes using the directional information. For the luma intra prediction, the directional information of a bigger block is applied to smaller blocks to filter out the unlikely modes. For the chroma intra prediction, the best prediction mode of luma block is referred to choose the candidate modes of chroma block. Simulation results show that the method can achieve computation time reduction up to 78% with little loss in PSNR and bitrate.

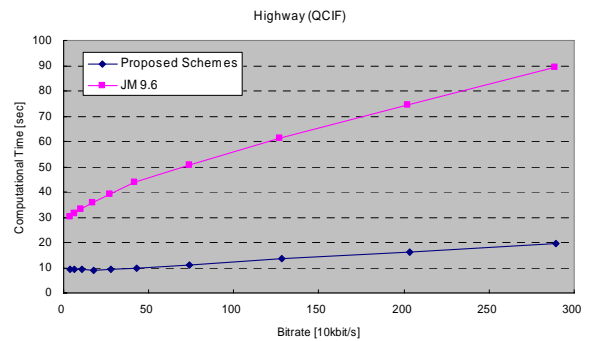
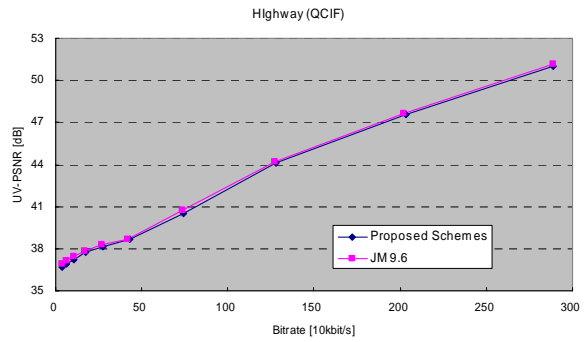
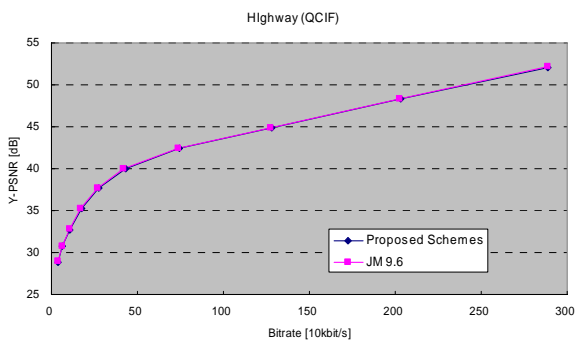


Fig. 8 The RD performance (Luma and Chroma) and computational time comparison of Highway (QCIF) sequence

REFERENCES

- [1] ITU-T Recommendation H.264 & ISO/IEC 14496-10 (MPEG-4) AVC, "Advanced Video Coding for Generic Audiovisual Services," (version 1: 2003, version 2: 2004) version 3: 2005.
- [2] S. Kwon, A. Tamhankar, K.R.Rao, "Overview of H.264 / MPEG-4 Part 10"
- [3] T. Halbach, "Performance Comparison: H.26L Intra Coding vs. JPEG2000", ISO/IEC JTC1/SC29/WG11 and ITU-T SG16 Q.6, JVT 4th meeting Klagenfurt, Austria, July 2002.
- [4] Zhang, Y, Dai, F, Lin, S, "Fast 4×4 intra-prediction mode selection for H.264", Multimedia and Expo, 2004. ICME '04. 2004 IEEE International Conference on Volume 2, 27-30 June 2004 Page(s):1151 - 1154 Vol.2
- [5] Feng Pan, Xiao Lin, Susanto Rahardja, Keng Pang Lim, Z. G. Li, Dajun Wu, Si Wu, "Fast mode decision algorithm for intraprediction in H.264/AVC video coding", Circuits and Systems for Video Technology, IEEE Transactions on Volume 15, Issue 7, pp. 813-822, July 2005.
- [6] G. J. Sullivan, P. Topiwala and A. Luthra "The H.264/AVC advanced video coding standard: Overview and introduction to the fidelity range extensions", SPIE Conf. on applications of digital image processing XXVII, vol. 5558, pp. 53-74, Aug. 2004.
- [7] Changsung Kim, Qing Li, C. C. Jay Kuo, "Fast Intra-prediction model selection for H.264 codec," SPIE International Symposium ITCOM 2003, Orlando, Florida, Sept. 7-11, 2003.
- [8] Iain E G Richardson, "H.264 and MPEG-4 Video Compression", John Wiley & Sons, pp. 180-183, 2003.
- [9] Yu-Wen Huang, Bing-Yu Hsieh, Tung-Chien Chen, and Liang-Gee Chen, Fellow, IEEE / Analysis, Fast Algorithm, and VLSI Architecture Design for H.264/AVC Intra Frame Coder / IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, VOL. 15, NO. 3, MARCH 2005.
- [10] H.264/AVC reference software JM 9.6, Jul. 2005