# A FAST INTRA MODE SELECTION METHOD FOR H.264 HIGH PROFILE

*Yue Yu and Limin Wang*

Advanced Technology, Connected Home Solutions, Motorola Inc.
6450 Sequence Dr., San Diego, CA 92121

## ABSTRACT

This paper proposes a fast intra mode selection method for high profile of H.264/MPEG-4 AVC. H.264 high profile supports three intra modes: intra_4x4, intra_8x8 and intra_16x16. There are nine possible prediction directions for intra_4x4 and intra_8x8, and four for intra_16x16. For intra_4x4, the proposed method checks all the nine possible prediction directions and determines a best direction per 4x4 block within a macroblock (MB). For intra_8x8, it only checks the prediction directions selected for the four 4x4 blocks of an 8x8 block within a MB plus the most probable mode (MPM). The saving in selecting a prediction direction for intra_8x8 can range from 45% to 100%. Further, if the four 8x8 blocks within a MB have different prediction directions, the MB is considered inhomogeneous. Tests over ten SD sequences indicated about up to 15% inhomogeneous MBs. For those inhomogeneous MBs, the prediction direction checking for intra_16x16 is skipped. For homogeneous MBs, the four possible prediction directions for intra_16x16 will be further evaluated. Simulation results showed that the proposed method does not bring any performance degradation with these complexities reduction.

***Index Terms—*** H.264/MPEG-4 AVC, mode selection, video coding, intra, high profile

## 1. INTRODUCTION

H.264/MPEG-4 AVC [1] is the latest international video coding standard from MPEG and ITU-T Video Coding Experts Group. This standard demonstrates significant performance improvement as compared with previous video coding standards, such as MPEG-1, MPEG-2, MPEG-4 part 2 etc. H.264 defines several profiles and levels for different applications [1]. In high profile, H.264 supports three intra modes at macroblock (MB) layer. They are intra_4x4, intra_8x8, and intra_16x16. Intra_4x4 and intra_8x8 have nine possible prediction directions, and intra_16x16 has four. These three intra modes along with their possible prediction directions bring more flexibility to improve the encoding performance, but at the expense of the complexity in selecting the final intra mode per MB.

An optimal mode selection method should check all possible prediction directions per mode and determine a best mode per MB. One example is JVT reference code. However, such an exhausted implementation may not be practical for a real time H.264 encoder. Recent research has focused on fast intra mode selection for both main profile [3]-[6] and high profile [7].

In this paper, we propose a fast intra mode selection method for H.264 high profile. In the Section 2, we briefly explain the intra prediction modes for H.264 high profile, and the RD-based mode decision method implemented in JVT reference code. In section 3, we describe the proposed method in reducing the complexity of intra mode decision for intra_8x8 and intra_16x16. In section 4, we present the simulation results.

## 2. THE INTRA PREDICTION MODES FOR H.264 HIGH PROFILE

Intra coding tends to reduce the spatial redundancy. There are three intra prediction modes at MB level for H.264 high profile. They are intra_4x4, intra_8x8 and intr_16x16, respectively. For intra_4x4, a MB is divided into sixteen 4x4 blocks, as shown in the Figure 1. Each 4x4 block can take one of nine possible prediction directions. Similarly to intra_4x4, for intra_8x8, a MB is divided into four 8x8 blocks, and each 8x8 block can also take one of nine possible prediction directions. For intra_16x16 mode, there are four possible prediction directions.

Ideally, given a block (4x4, 8x8, or 16x16) per intra mode (intra_4x4, intra_8x8, or intra_16x16), all the possible prediction directions should be checked for the block, and a best prediction direction is then determined for the block based upon a cost function. The cost function can be RD-based or non-RD based. The RD-based cost function in coding a block is defined as.

$$J = D + \lambda \times R \qquad (1)$$

where $D$ is the coding distortion of the block, $R$ is the number bits required for encoding this block and $\lambda$ is the Lagrangian parameter. The RD-based cost function requires the implementation of the complete encoding process per prediction direction, which can be very time consuming. The use of non-RD cost functions can significantly reduce the process in selecting a best prediction direction per block per intra mode [3]-[6].

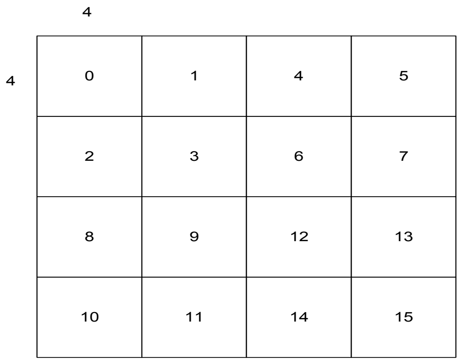

Figure 1. The layout of intra_4x4

In general, the smaller the block size is, the better the prediction can be. For example, 4x4 block often has smaller prediction residual than 8x8 block and 16x16 block. However, the prediction direction per block needs to be encoded, which requires additional bits. Intra_4x4 has sixteen blocks per MB while intra_8x8 has only four blocks per MB. Clearly, intra_4x4 needs more bits to indicate the prediction directions than intra_8x8 and intra_16x16 per MB. There is a tradeoff between prediction residual and overhead bits for prediction directions.

## 3. THE PROPOSED METHOD

There are two level decisions for intra mode coding of H.264 high profile. At MB level, there are three intra modes. They are intra_4x4, intra_8x8, and intra_16x16. At each intra mode level, there are a number of possible prediction directions. For intra_4x4, the prediction directions include vertical (0), horizontal (1), DC (2), diagonal_down_left (3), diagonal_down_right (4), vertical_right (5), horizaontal_down (6), vertical_left (7) and horizontal_up (8). Intra_8x8 has the same prediction directions. Intra_16x16 has four prediction directions, and they are vertical (0), horizontal (1), DC (2) and plane (3). A H.264 encoder needs to find a best prediction direction among the possible ones pre intra mode and a final intra mode among intra_4x4, intra_8x8 and intra_16x16.

The proposed fast method starts with intra_4x4. For each 4x4 block, all the nine possible prediction directions are checked, and a best one is selected based upon a cost function. It then moves on to check intra_8x8. intra_8x8_and intra_4x4 are very similar. It has nine possible prediction directions identical to intra_4x4. The main difference is that the block size in intra_8x8 is 8x8 while in intra_4x4 is 4x4. An 8x8 block in intra_8x8 covers four 4x4 blocks in intra_4x4. The prediction directions selected for the four 4x4 blocks certainly provide an indication of prediction direction for the 8x8 block. In other words, the prediction directions of the four 4x4 blocks should be highly correlated with the prediction direction for the 8x8 block. As a result, the four prediction directions of the four 4x4 blocks constitute a good set of candidates for the 8x8 block.

For intra_4x4 and intra_8x8, H.264 introduces a concept, named the most probable mode (MPM), which is the minimum of coding mode index values of left and upper neighboring blocks, if they are available. If the left or the upper neighboring block is not available, the corresponding prediction direction is simply set to DC [1]. If the final prediction direction for a block is the same as its MPM, only one bit is required for MPM. Otherwise, four bits are needed to indicate which of other eight prediction directions is used for the block. The prediction directions of neighboring blocks can be very correlated. In addition, three bits can also play an important role in the mode selection formula (1). Hence, the chance of MPM being selected is much higher than other modes. Therefore, MPM is also included as a candidate for intra_8x8 mode. A similar work uses DC, instead of MPM, as an additional candidate [7]. For a given block, its MPM is often a better predictor for the block’s prediction direction than DC due to the spatial correlation among the neighboring blocks. In addition, MPM requires only one bit for its prediction direction.

If a MB is considered homogeneous, a larger block size intra mode is likely selected. For example, if an 8x8 block is homogeneous, it is likely that intra_8x8, instead of intra_4x4, will be selected. Even further, if a MB locates in a flat area, it is likely predicted well in intra_16x16 In general, if prediction with a bigger block size results in a similar prediction residual as a smaller block size, the prediction with bigger block size is usually selected. This is because of the facts, 1) similar residual will consume similar bits, and 2) bigger block size intra mode requires less bits for indicating the prediction direction per MB. For example, Intra_4x4 requires many more bits in dictating the prediction directions of its sixteen 4x4 blocks than intra_16x16. On the other hand, if a MB is not homogeneous, a small block size intra mode is often selected since a bigger block size intra mode often give a poor prediction that requires more bits to be coded.

If a MB can be detected inhomogeneous in advance, big block size intra mode may be skipped to save the computation. The key is how to evaluate if a current MB is inhomogeneous. The proposed method checks the three intra modes per MB in order of intra_4x4, intra_8x8 and intra_16x16. The prediction directions selected for smaller blocks within a bigger block may indicate if the bigger block is homogeneous or not. For example, if all the sixteen 4x4 blocks or all the four 8x8 blocks within a MB have the same prediction direction, the MB can be considered homogeneous, and a homogeneous MB can be predicted well in intra_16x16. In contrast, if the sixteen 4x4 blocks or the four 8x8 blocks within a MB have very random prediction directions, the MB is inhomogeneous, and it cannot be predicted well in intra_16x16. In this case, intra_16x16 should be skipped.

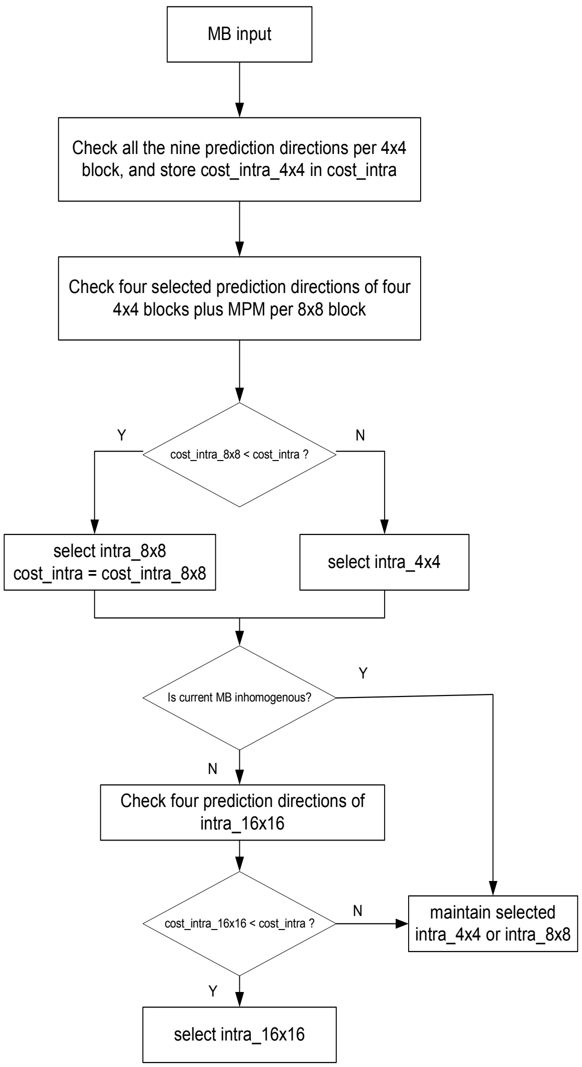


Figure 2. The flow chart of the proposed method

Figure 2 shows the flow chart of the proposed method. At MB level, for a current MB, the proposed method will first check intra_4x4 mode, then intra_8x8 mode, and finally, if it is necessary, intra_16x16 mode. For intra_4x4, for each 4x4 block of the current MB, all the nine possible prediction directions will be checked, and a best prediction direction will then be decided based upon a cost function. The sum of the costs of all the sixteen 4x4 blocks for the current MB is stored in a variable named cost_intra. For intra_8x8 mode, for each 8x8 block of the current MB, the four prediction directions selected for the four 4x4 blocks within the 8x8 block plus MPM will be further checked, and a best prediction direction is determined based upon a cost function. Since these five prediction directions may or may not be the same, at most, only five prediction directions will be checked per 8x8 block in intra_8x8 mode. If all the five prediction directions are the same, the same prediction direction will be the one for the 8x8 block. The sum of the costs of the four 8x8 blocks within the current MB is then compared with the variable cost_intra. If the intra_8x8 cost is smaller than intra_cost, intra_8x8 mode is selected and the value of cost_intra is replaced by the intra_8x8 cost. Otherwise, intra_4x4 mode is selected and the value of cost_intra remains unchanged. The four prediction directions selected for the four 8x8 blocks within the current MB is further checked. If they are different, the current MB is considered inhomogeneous, and intra_16x16 mode checking will be skipped. Otherwise, the four prediction directions for intra_16x16 will be checked, and a best prediction direction is determined for 16x16 block. The intra_16x16 cost is then compared with the variable cost_intra. If the intra_16x16 cost is smaller than cost_intra, intra_16x16 mode will be finally selected for the current MB. Otherwise, the intra mode, either intra_4x4 or intra_8x8, determined in the previous stage will remain for the current MB. Note that the cost function used in selecting the prediction direction can be either RD-based or non-RD based.

## 4. SIMULATION RESULTS

In evaluating the performance of the proposed fast intra mode selection method, simulations were conducted under the test conditions, as shown in the Table 1.

Table 1. The test Conditions

| | |
|---|---|
| Frame type | All I frames |
| No. of frames | 148 |
| Picture structure | Frame coding |
| Allowed modes | Intra_16x16, Intra_8x8 and Intra_4x4 |
| QP | Four QPs, from 28 to 40, depending on different sequences |
| Entropy code | CABAC |

Ten test sequences are used in the simulations. They are Basketball, Bus, Canoa, Coastguard, CheerLeader, F1Car, Football, Mobile, Rugby and Tempete. The ten sequences cover a great variety of motions, colors, sharpness, brightness and textures. All these ten sequences have a spatial resolution of 720x480 pixels interlaced at 30 frames/sec.

For comparison purpose, as a benchmark, the exhausted checking method is also implemented for the ten sequences. The exhausted method checks all the possible prediction directions per block (4x4, 8x8, or 16x16) per intra mode (intra_4x4, intra_8x8, or intra_16x16).

Table 2 shows the coding performance of the proposed method as compared to the benchmark method in terms of the percentage of bit rate increase, or AVSNR [2]. As seen, the increase in bit rate for most of sequences is very small and the biggest one is only 0.57%. Such a small variation in bit rate can usually be ignored in the most applications. In other words, the proposed method can achieve almost the same encoding performance as the benchmark exhausted method.

Table 2. The increase in bit rate

| Sequences | Increase bit rate (%) |
|---|---|
| Basketball | 0.33 |
| Bus | 0.09 |
| Canoa | -0.06 |
| Cheerleader | 0.57 |
| Coastguard | 0.12 |
| F1Car | 0.14 |
| Football | 0.09 |
| Mobile | 0.17 |
| Rugby | 0.22 |
| Tempete | 0.35 |

Table 3 shows, as an example, the percentages of intra_16x16, intra_8x8 and intra_4x4 modes selected using the benchmark method for Basketball.

Table 3. The % of three intra modes selected using the benchmark method for Basketball

| | QP=32 | QP=34 | QP=36 | QP=40 |
|---|---|---|---|---|
| Intra_16x16 | 19.33 | 23.27 | 27.83 | 37.22 |
| Intra_8x8 | 39.29 | 34.92 | 31.15 | 26.74 |
| Intra_4x4 | 41.38 | 41.81 | 41.02 | 36.04 |

Since the proposed method still needs to check all the possible prediction directions for intra_4x4, there is no saving for intra_4x4 mode.

For intra_8x8 mode, only the prediction directions selected for the four 4x4 blocks per 8x8 block plus the MPM will be checked. The worst case is none of the four prediction directions are the same. Five possible prediction directions will be checked per 8x8 block. The best case is all the four prediction directions and the MPM are the same. There is no need to further check any prediction direction per 8x8 block. The saving for intra_8x8 mode can range from 45% to 100%.

The second row in table 3 shows the percentages of MBs choosing intra_16x16 mode as the best intra mode for a given quantization parameter QP1, QP2, QP3, or QP4. Ideally, the percentage of MBs that skip intra_16x16 mode checking should be smaller than those numbers. Table 4 shows the percentage of MBs that skip intra_16x16 modes using the proposed method. For different sequences at different quantization levels, the percentage varies significantly, ranging from 2.6% to 15%.

Table 4. The % of MBs skipping intra_16x16 mode

| | QP1 | QP2 | QP3 | QP4 |
|---|---|---|---|---|
| Basketball | 12.72 | 12.16 | 11.19 | 9.06 |
| Bus | 5.98 | 5.34 | 4.72 | 3.91 |
| Canoa | 4.09 | 3.96 | 3.84 | 3.62 |
| Cheerleader | 15.17 | 14.18 | 13.24 | 10.95 |
| Coastguard | 5.08 | 4.63 | 4.03 | 3.49 |
| F1car | 4.22 | 3.73 | 3.28 | 2.61 |
| Football | 4.85 | 4.30 | 3.81 | 3.18 |
| Mobile | 12.49 | 11.68 | 10.78 | 9.89 |
| Rugby | 7.32 | 6.62 | 5.83 | 4.38 |
| Tempete | 13.10 | 12.52 | 11.91 | 11.28 |

## 5. CONCLUSIONS

This paper proposed a fast intra mode selection method for H.264 high profile. The proposed method significantly reduces the process in selecting the prediction direction per intra mode without degrading the coding performance. For intra_8x8, the saving in selecting a best prediction direction per 8x8 block can range from 45% to 100%. Further, if a MB is considered inhomogeneous, the checking for intra_16x16 mode is skipped, resulting in a saving of up to 15%. In this research, a strict measure is applied in evaluating if a MB is homogeneous. A loose measure should result in an additional saving.

## 6. REFERENCES

[1] “Text of ISO/IEC 14496 10 Advanced Video Coding 3rd Edition,”Joint Video Team (JVT) of ISO/IEC MPEG & ITU-T VCEG, July 2004, Redmond, WA, USA.
[2] G. Bjontegaard, "Calculation of average PSNR differences between RD-curves", VCEG-M33, Apr. 2001.
[3] C. Tseng, H. Wang and J. Yang, “Enhanced intra-4x4 mode decision for H.264/AVC coders,” IEEE Trans. on CSVT, vol. 16, no. 8, Aug. 2006, pp. 1027-1032.
[4] Y. Tu, J. Yang and M. Sun, “An efficient criterion for mode decision in H.264/AVC,” International Conference on Multimedia and Expo, July, 2006, Toronto, pp. 1705-1708.
[5] Song-Hak Ri, Joern Sstermann, “Intra-Prediction mode decision for H.264 in Two steps,” Picture Coding Symposium, April, 2006, Beijing.
[6] F. Pan, X. Lin, S. Rahardja, K. P. Lim, Z. G. Li, D. Wu and S. Wu, “Fast mode decision algorithm for intra prediction in H.264/AVC video coding,” IEEE Trans. on CSVT, vol. 15, No. 7, July 2005, pp. 813-822.
[7] Yeping Su, Ming-Ting Sun and Kun-Wei Lin, “Encoder optimization for H.264/AVC fidelity range extensions,” Visual Comm. and Image Processing, 2005, Beijing, pp. 2067-2075.