International Journal of Computer Applications (0975 – 8887)
Volume 37– No.6, January 2012

Fast Intra Mode Decision Algorithm for H264/AVC HD Baseline Profile Encoder

Nejmeddine Bahri, Imen Werda, Amine Samet, Mohamed Ali Ben Ayed, Nouri Masmoudi
University of Sfax
National Engineering School
Sfax, Tunisia

ABSTRACT
The high performance of H.264/AVC video encoder is accompanied with a wide computation complexity especially for high definition (HD) video sequences. One of the major H.264/AVC features to be optimized is the mode decision for both inter and intra prediction. Thus, based on high correlation observed between selected inter prediction mode and intra mode decision, a fast intra mode decision algorithm based on the best inter prediction mode for H264 high definition (HD) baseline profile encoder is proposed. The evaluation of the proposed approach was based on the rate distortion and PSNR variation, execution time and percentage of skipping intra4x4 and intra16x16. The proposed scheme is performed on 720p (1280x720) and 1080p (1920x1088) HD video sequences. Experimental results show that the proposed algorithm can save up to 60% of intra prediction computation time, 16% of skipping intra16x16 and up to 83% for intra4x4 without inducing PSNR degradation and bit-rate increase.

General Terms
Video compression techniques and signal processing

Keywords
H264/AVC, Fast intra mode decision, High definition, baseline profile.

1. INTRODUCTION
The H.264/AVC [1] is a video encoder standard that developed by Joint Video Team (JVT) of ITU-T VCEG and ISO/IEC MPEG. This standard achieves better video coding efficiency by saving up to 50% of bit-rate and maintaining the same visual quality compared to previous standards as MPEG2. This efficiency is accompanied by a high computational complexity due to many new features included in this standard such as intra and inter predictions designed to reduce spatial and temporal redundancies.
The H264/ AVC standard adopts different profiles such as [2]:

- Baseline profile: Primarily used for lower-cost applications with limited computing resources as videoconferencing and mobile applications.
- Main profile: Originally intended as the mainstream consumer profile for broadcast and storage applications.
- Extended profile: Intended as the streaming video profile, this profile has relatively high compression capability and some extra tricks for robustness to data losses and server stream switching.
- High profile: is the primary profile for broadcast and disk storage, especially for High-Definition (HD) television.

Currently, HD resolution will be more and more used in several media applications such as TV, cinema and even video conferencing. Thus, several companies such as Texas Instruments [3], EyeLytics [4] and Alma Technologies [5] work to develop their own H264 HD baseline profile encoder.
The H264/AVC modules profiling show that inter and intra predictions increase greatly the H264/AVC computation time [12]. When adopting fast motion estimation approaches [6], computational complexity of inter prediction module will be drastically reduced as presented in Fig1. This figure presents the run time percentage for H264/AVC modules when using a full search inter prediction algorithm and when using fast inter prediction algorithm. As a result, required time for checking intra prediction modes becomes relatively important compared to that of inter prediction module. Therefore, a fast intra mode decision algorithm is required.

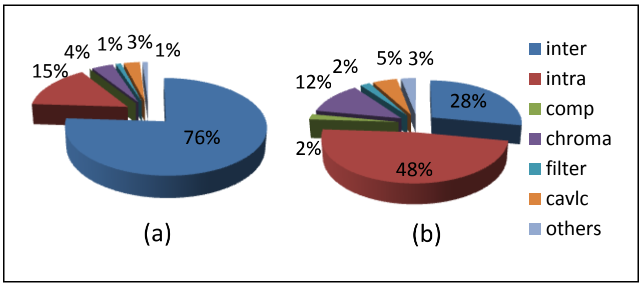

Fig1: Run-time percentage for H264/AVC modules:
(a) Profiling using full search inter prediction algorithm
(b) Profiling using fast inter prediction algorithm

In this paper, a fast intra mode decision approach is proposed, based on the correlation between the best inter prediction mode and MB homogeneity. The rest of this paper is organized as follows. After a brief overview of H.264/AVC intra/inter mode decision in Section 2, we present previous works adopted to reduce intra prediction complexity in section 3. Our proposed algorithm is described in Section 4. Experimental results are discussed in section 5. Finally, conclusion is presented in section 6.

2. INTRA/INTER MODE DECISION FOR H264/AVC
H264/AVC baseline encoder presents two frame types:
- Intra frame (I Frame) where only the intra prediction is performed to reduce the spatial redundancy.
- Inter frame (P Frame) where both intra and inter predictions are executed to eliminate respectively spatial and temporal redundancies.
For inter prediction, there are seven different block sizes (Fig2): 16×16, 8×16, 16×8, P8x8 {8×8, 8×4, 4×8 and 4×4}, while intra prediction includes two different block sizes:

❖ Intra4×4 (I4MB) performed for each 4x4 block and containing nine directional prediction modes as shown in Fig3.
❖ Intra16×16 (I16MB) applied for each 16x16 MB and including four directional prediction modes as illustrated in Fig4.

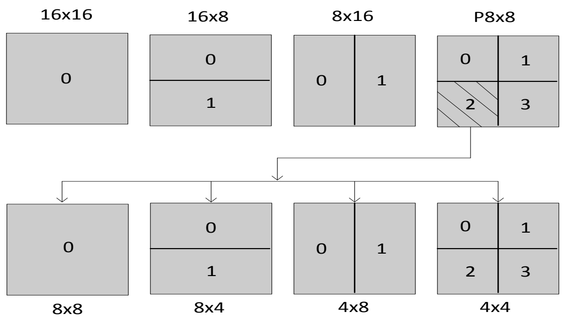

Fig2: Inter prediction modes

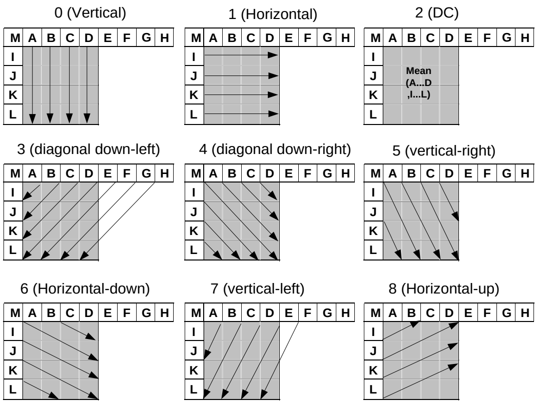

Fig3: 4x4 intra prediction modes

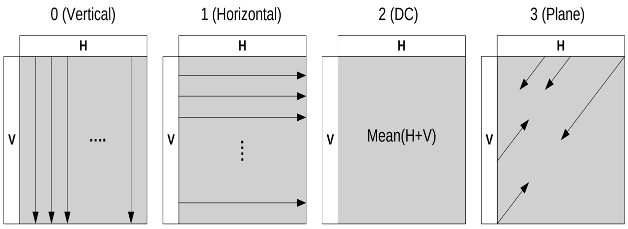

Fig4: 16x16 intra prediction modes

In order to derive the best mode between the different intra and inter prediction modes, a Rate distortion optimization (RDO) function which includes both the distortion and the rate, is defined in the H.264 reference software JM [18]. So, the encoder selects the mode that minimizes the following cost function:

$$Cost_{mode} = Distortion\ (MB) + \lambda_{mode}\ Rate\ (MB) \qquad (1)$$

The Fig5 summarizes the major steps for intra16x16, intra4x4 and inter prediction in P Frame.

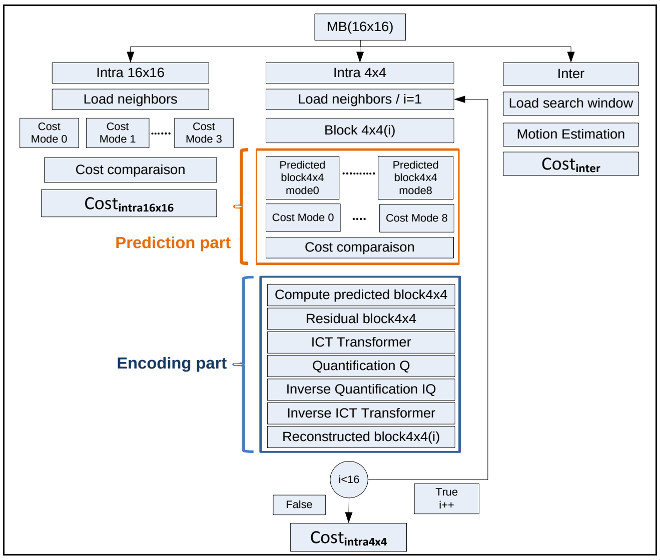

Fig5: Intra and inter mode decision steps

❖ For intra16x16, costs relative to the four intral6x16 modes are calculated and the minimum one is chosen as $Cost_{intra16x16}$.

❖ For intra4x4, the cost computation is performed on two steps:

- Prediction step: The current MB is divided into sixteen 4x4 blocks. For each one, 9 predicted blocks are generated and the best one which gives the minimum cost is selected.
- Encoding step: The residual block between the current block and its prediction is computed, transformed and quantized to generate the reconstructed block used for the next block prediction.
  The $Cost_{intra4x4}$ will be computed as the sum of the sixteen minimum costs relative to the sixteen 4x4 blocks.

❖ Regarding inter prediction, motion estimation is performed in order to calculate the costs relative to the 7 inter block sizes. The minimum cost among 16x16, 16x8, 8x16 and P8x8 costs is chosen as $Cost_{inter}$.

Finally $Cost_{intra16x16}$, $Costi_{ntra4x4}$ and $Cost_{inter}$ are compared in order to choose the best mode decision for the current MB.

3. PREVIOUS WORKS

The adoption of fast motion estimation algorithm makes intra predictions as the most demanding part in terms of computation time among all the H264/AVC modules. This complexity is due to the large number of RDO calculations performed for each MB in order to select the best intra prediction mode. Many algorithms have been proposed to reduce intra prediction complexity. Among those who worked on the optimization of the intra prediction we quote:
Pan *et al.* [7], present an intra mode decision scheme using Sobel operator to compute the direction histogram and used its dominating edge direction to reduce the number of candidate modes.
Chao-Chung *et al.* [8], propose Fast Three Step Intra Prediction Algorithm for 4x4 blocks. In fact, they only examined 6 modes instead of 9 modes in the full search algorithm.
Sourabh Rungta *et al.* [9], present a fast mode selection algorithm using texture analysis of video sequences. They computed the MB variance, horizontal and vertical texture

direction before comparing them to thresholds in order to achieve an early termination intra and inter mode decision.
Takeshi *et al.* [10], based on frequency characteristic of the block 4x4, some parameters are computed in order to classify the frequency distribution component of the input signal and used them to reduce the number of candidate modes.
Jun Sung Park *et al.* [11], propose a fast selective intra mode decision algorithm exploiting the directional correlation between 16x16 MB and 4x4 block. The result of intra16x16 is then used to reduce the number of intra 4x4 candidate modes.
Huang *et al.* [12], Based on statistic study, they found that a MB is very likely to be inter coded when the best inter mode cost is much lower than intra16x16 cost while intra modes would be selected as the best mode in the contrary case. Also the MB variance was computed to classify the texture complexity in order to take the decision of the block size (I8MB/I16MB) or (I8MB/14MB).
Lin *et al.* [13], present an H.264 intra frame codec IP design which supports HD size 1080p encoding process their design has been improved with the variable-pixel parallel architecture and the modified three-step Chao-Chung' s algorithm [8].
Golam and Wu [14], a correspondence between the neighboring pixels variance of the 4x4 current block and its best prediction mode is considered. The variance of its 12 neighboring pixels and the variance of its top 8 neighboring pixels are computed and compared them with two thresholds to restrict the number of possible candidate modes.
Jordan *et al.* [15], a criterion C and a threshold T is used to decide if the previously used coding mode of a 4× 4 block in the frame N-K will be reused for the current block of the frame N or a full intra mode-decision search is started.
Elyousfi *et al.* [16], present a fast intra prediction algorithm using a gradient and a quadratic prediction functions to improve the encoding speed. The homogeneous areas are predicted by the gradient prediction function, since non-homogeneous areas are predicted by the second method.
And finally Jeon B and J. Lee [17], perform a spatial domain analysis to decide if the intra modes should be checked or not. The average boundary error (ABE) and the average rate (AR) are computed and compared. If AR <k*ABE, there is no need to consider any intra modes for the current MB. So that, only inter mode decisions are performed. Otherwise, both inter and intra modes should be considered.
Even if some of these algorithms were able to reduce intra prediction computation time, they still have some drawbacks. First, they increase the H264/AVC complexity computation due to the pre-calculations required in computing edge direction and boundary variance. Second, the assumption of edge direction is not always true in every block; furthermore threshold technique makes proposed schemes dependent on the video sequences contents. Moreover, methods reducing the number of intra candidate modes, affect only the prediction part while intra prediction module is composed of two parts (Fig5):

- A prediction part where a cost function is performed for each intra modes.
- An encoding part used to compute residual block, and to determinate neighboring pixels for the next block.

Finally these approaches induce a bit-rate increase when the rate control option is turned off and heavily degrade PSNR quality.

4. PROPOSED ALGORITHM

To solve problems of previous algorithms, a fast intra prediction approach is presented. When adopting this approach, pre-calculations step is not required and conditional threshold used to achieve early termination mode decision is not utilized. The aim of the proposed scheme is to reduce the whole intra prediction complexity not only the prediction part as adopted by the most previous algorithms.
The best way to achieve this aim is choosing the appropriate condition to decide between performing or skipping intra4x4 and intra16x16. Our proposed approach is a result of several analyses performed on different resolution sequences: CIF and HD. These analyses show that:

❖ The inter mode is the most selected mode for P frames compared to intra4x4 and intra16x16, as shown in Fig6. This figure presents the mode decision percentages for HD video sequences with different quantification parameters (QP). As a result, when a fast intra mode decision is proposed in P slices, no important degradation in visual quality is observed.

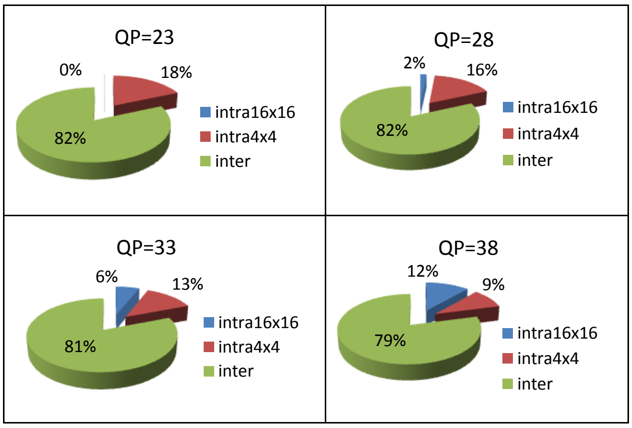

Fig6: modes decision percentages for HD 1280x720 video sequences

❖ Fig 7 illustrates the inter mode decision map in P frame.16x16, 16x8 and 8x16 inter prediction block sizes are generally used for backgrounds and stationary blocks whereas P8x8 block sizes are used in detailed and fast motion areas.

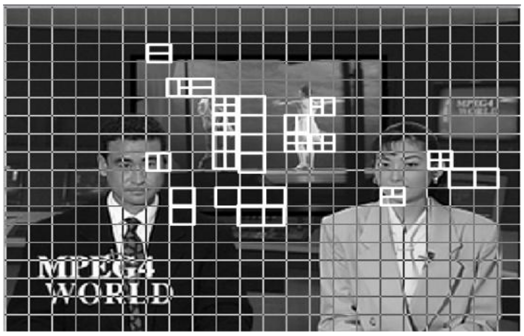

Fig7: MB inter mode decision map of News CIF sequence

❖ Intra16x16 is generally used for backgrounds, stationary and homogenous blocks characterized by a tiny luminance change; whereas intra4x4 is used for detailed areas where there is a high luminance change as shown in Fig 8 which presents the intra mode decision map in I Frame.

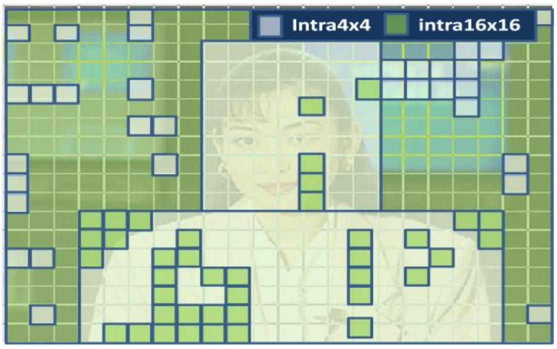

Fig8: MB intra mode decision map of Akiyo CIF sequence

Considering these observations, we can affirm that there is a high correlation between inter and intra prediction modes. In fact, we propose a fast intra mode decision based on inter prediction modes.

The proposed approach flowchart consists of the following steps, as detailed in Fig 9:

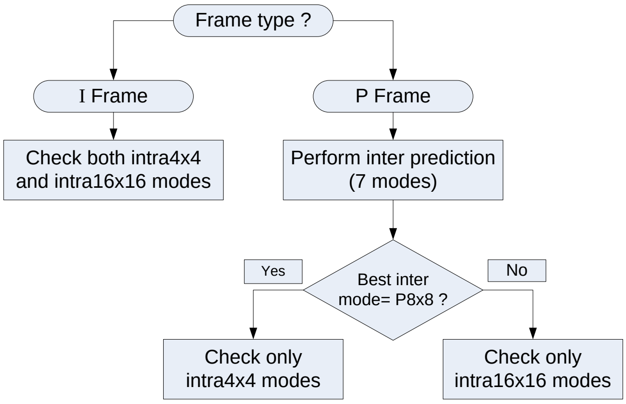

Fig9: Flowchart of the proposed algorithm

❖ For I Frames, intra mode decision was not changed: the full search intra prediction process is performed and all intra16x16 (I16MB) and intra4x4 (I4MB) modes are checked.

❖ For P Frames, the intra mode decision is related to the best inter prediction mode as follows:

- If the best inter mode is P8x8, which means that 8x8, 8x4, 4x8 or 4x4 block size is selected, intra4x4 is performed and intra16x16 is skipped because the MB is considered not homogenous and characterized by detailed texture.
- Else, intra16x16 is checked and intra4x4 is skipped because the MB is considered homogenous.

5. EXPERIMENTAL RESULTS

The proposed algorithm was implemented using the H.264 baseline profile reference software JM 17.1[18] according simulation parameters as described in Table 1.

Table 1. Simulation parameters

Intra period	8
Search range	16
QP values	20,28,38
Frames to be encoded	32
Error metric for mode decision	SAD
Entropy encoder	CAVLC
Rate control	off
Number of Reference Frames	1
Frame rate	30
Sequence sizes	720p (1280x720) and 1080p (1920x1088)

Experimental simulations are performed using HD video sequences, with two resolutions such as 720p and 1080p, with different characteristics: Crowdrun, ParkJoy, Sunflower, OldTown and BlueSky. The system platform is a laptop based on Intel Pentium4, 2 Dual cores, 2.2GHz processor speed with 2 Go as internal RAM.

For performance comparison, our proposed scheme will be evaluated based on three criteria:

❖ ΔPSNR: presents the PSNR variance between the proposed algorithm and the JM reference software.

❖ ΔBits (%): represents the percentage increase in bit-rate for the implemented scheme compared to JM reference software.

❖ ΔTime (%): is defined as the percentage of time saving of intra prediction between the proposed algorithm and JM.

The above three criteria are detailed by the following equations:

$$\Delta \text{Bits}(\%) = \frac{\text{Bitrate (proposed)} - \text{Bitrate (reference)}}{\text{Bitrate (reference)}} * 100 \qquad (2)$$

$$\Delta \text{PSNR} = \text{PSNR(proposed)} - \text{PSNR(reference)} \qquad (3)$$

$$\Delta \text{Time}(\%) = \frac{\text{Time (proposed)} - \text{Time (reference)}}{\text{Time (reference)}} * 100 \qquad (4)$$

The simulation results are presented in Table2 and Table3 which show respectively our proposed algorithm performances on 720p and 1080p HD video sequences. The three criteria are calculated for different QP values.

As presented in this table, we can observe that our proposed algorithm can save up to 60% of intra prediction computation time without inducing a PSNR loss and bit-rate increase.

Table 2: Experimental results for 720p (1280x720) HD video sequences

	QP=20			QP=28			QP=38		
Sequence	**ΔPSNR(dB)**	**ΔBits (%)**	**ΔTime (%)**	**ΔPSNR(dB)**	**ΔBits (%)**	**ΔTime (%)**	**ΔPSNR(dB)**	**ΔBits (%)**	**ΔTime (%)**
Crowdrun	0	0	-52.3	0	0	-54	0	0	-59.2
Parkjoy	0	+0.004	-54.4	-0.02	+0,005	-58.97	-0,04	+0,003	-62.5
Sunflower	-0.01	0	-56.12	0	0	-55.92	0	0	-60.4
OldTown	0	0	-60.43	0	0	-56.21	0	0	-58

Table 3: Experimental results for 1080p (1920x7088) HD video sequences

	QP=20			QP=28			QP=38		
Sequence	**ΔPSNR(dB)**	**ΔBits (%)**	**ΔTime (%)**	**ΔPSNR(dB)**	**ΔBits (%)**	**ΔTime (%)**	**ΔPSNR(dB)**	**ΔBits (%)**	**ΔTime (%)**
Crowdrun	0	+0.001	-51.45	0	0	-53.46	-0.01	0	-58.64
Parkjoy	-0.07	+0.01	-52.7	-0.03	+0.002	-54.78	-0,05	+0.006	-59.45
Bluesky	0	0	-54.86	0	+0.002	-54.23	-0.01	-0.002	-61.78
Sunflower	-0.01	+0.004	-56.26	-0.01	+0.003	-57.8	-0.02	-0.004	-62.14

Since the computational time is relative to the performances of the platform which doesn't give an accurate idea on the encoding speed and time saving, we can determine how many times intra4x4 and intra16x16 are checked when the proposed approach is applied and then deduce the percentages of skipping intra4x4 and intra16x16 with respect to the JM reference software.

These percentages are defined respectively as follows:

$$\Delta I16\% = \frac{Nb\_I16\_prop - NbI16\_ref}{Nb\_I16\_ref} x100 \qquad (5)$$

$$\Delta I4\% = \frac{Nb\_I4\_prop - NbI4\_ref}{Nb\_I4\_ref} x100 \qquad (6)$$

Nb_I16_ref and Nb_I4_ref represent respectively number of call for intra16x16 and intra4x4 modules in the JM reference software. Nb_I16_prop and Nb_I4_prop represent respectively number of call for intra16x16 and intra4x4 modules when the proposed algorithm is applied.

$$Nb\_I16\_ref = Nb\_I4\_ref = \text{Number of MBs in 1 Frame} * \text{Number of frame to be encoded.} \qquad (7)$$

i.e. for 720p (1280*720) HD sequence: Number of 16x16 MBs in 1 frame= (1280/16)*(720/16) =3600MBs. If the number of frames to be encoded=32 then, Nb_I16_ref= Nb_I4_ref=3600*32=115200.

Tables 4 and 5 show respectively the percentages of skipping intra16x16 and intra4x4 for P Frames with different QP values.

We can observe that our proposed approach can reduce by average 16.05% the number of performing intra16x16 and 83.72% for intra4x4. The gain of our proposed algorithm can yield to 90% for intra4x4 in many cases.

We notice that the percentage of skipping intra4x4 is increased especially for large QP values, contrary to intra16x16, which represents an advantage for us because intra4x4 computation time is more important than intra16x16.

Table 4: Saving use results for 720p HD video sequences in P Frames

	QP=20		QP=28		QP=38	
Sequence	**ΔI16 (%)**	**ΔI4 (%)**	**ΔI16 (%)**	**ΔI4 (%)**	**ΔI16 (%)**	**ΔI4 (%)**
Crowdrun	-50.54	-49.11	-26,73	-73,22	-11,51	-88,44
Parkjoy	-31.87	-64.06	-1,69	-98,30	-0,55	-99,20
Sunflower	-9.89	-89.34	-11,65	-86,89	-5,31	-93,86
Oldtown	-5.66	-94.33	-24,66	-73,97	-11,11	-87,72

Table 5: Saving use results for 1080p HD video sequences in P Frames

	QP=20		QP=28		QP=38	
Sequence	**ΔI16 (%)**	**ΔI4 (%)**	**ΔI16 (%)**	**ΔI4 (%)**	**ΔI16 (%)**	**ΔI4 (%)**
Crowdrun	-51.10	-53.27	-36.83	-63.16	-9.61	-90.38
Parkjoy	-35.79	-64.20	-24.07	-75.92	-4.46	-95.53
Bluesky	-11.48	-88.51	-4.66	-95.33	-0.57	-99.42
Sunflower	-9.46	-90.53	-4.5	-95.49	-1.60	-99.27

6. CONCLUSION

In this paper, based on exploiting the correlation observed between inter and intra prediction modes decision, a fast intra mode decision for H264/AVC HD baseline profile is presented. In our proposed algorithm, the complexity of intra prediction is reduced by choosing the appropriate condition to perform or skip intra4x4 and intra16x16 according to the best inter prediction mode. Our approach is evaluated based on the rate distortion and PSNR variation, execution time and percentage of skipping intra4x4 and intra16x16.

The proposed algorithm can save up to 60% of intra prediction computation time without inducing PSNR loss or bit-rate increase. In addition our approach allows saving 16.05% on average of performing intra16x16 and 83% of performing intra4x4.

7. REFERENCES

[1] Joint Video Team (JVT) of ISO/IEC MPEG & ITU-T VCEG, "Draft ITU-T Recommendation and Final Draft international Standard of Joint Video Specification (ITU-T Rec. H.264 ISO/IEC 14496-10 AVC)", JVT-G050, 2003.

[2] H264/MPEG-4 AVC Wikipedia the free encyclopedia:http://en.wikipedia.org/wiki/H.264/MPEG-4_AVC

[3] H.264 HD Baseline Profile Encoder (v2.00) on DM6446: http://www.ti.com/lit/ml/sprs527/sprs527.pdf

[4] Main / Baseline Profile HD h264 Encoder http://www.eyelytics.com/Products/H264Encoder.html

[5] H264-BP-E H.264/AVC HD & ED Video Encoder Core http://www.cast-inc.com/ip-cores/video/h264-bp-e/index.html

[6] Imen Werda, Haithem Chaouch, Amine Samet, Mohamed Ali Ben Ayed and Nouri Masmoudi, "Optimal DSP-Based Motion Estimation Tools Implementation For H.264/AVC Baseline Encoder," IJCSNS International Journal of Computer Science and Network Security, VOL.7 No.5, May 2007.

[7] 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 Transactions on Circuits and Systems for Video Technology, vol. 15, no. 7, pp. 813-822, July 2005.

[8] Chao-Chung Cheng and Tian-Sheuan Chang, "Fast Three Step Intra Prediction Algorithm for 4x4 blocks in H.264," Proc. IEEE Canadian Conference on Electrical and Computer Engineering, pp1981-1984, May 2003.

[9] Sourabh Rungta, Kshitij Verma and Anupam Shukla, "A Fast Mode Selection Algorithm Using Texture Analysis for H.264/AVC," IJCSI International of computer Sciences Issues, Vol. 7, Issue 4,No 9, July 2010

[10] Takeshi. Tsukuba, Isao. Nagayoshi, Tsuyoshi. Hanamura and Hideyoshi. Tominaga, "H.264 Fast Intra-Prediction Mode Decision Based on Frequency Characteristic," Proc. of European Signal Processing Conference (EUSIPCO), 2005.

[11] Jun Sung Park and Hyo Jung Song, "Fast selective intra mode decision H.264/AVC," IEEE Consumer Communications and Networking Conference 2006.3rd, Vol.2, pp.1068-1072 Jan. 2006.

[12] Yi-Hsin Huang, Tao-Sheng Ou, and Homer H. Chen, "Fast Decision of Block Size, Prediction Mode, and Intra Block for H.264 Intra Prediction," IEEE transactions on circuits and systems for video technology, Vol.20, No.8, august 2010.

[13] Yu-Kun Lin, Chun-Wei Ku, De-Wei Li, and Tian-Sheuan Chang, "A 140-MHz 94 K Gates HD1080p 30-Frames/s Intra-Only Profile H.264 Encoder," IEEE transactions on circuits and systems for video technology, Vol.19, No.3, March 2009.

[14] Mohammed Golam Sarwer and Q. M. Jonathan Wu, "Improved Intra Prediction of H.264/AVC," Effective Video Coding for Multimedia Applications, Sudhakar Radhakrishnan (Ed.), ISBN: 978-953-307-177-0, InTech (2011).

[15] Harald Jordan, Florian H. Seitner, Michael Bleyer and Margrit Gelautz, "Runtime-Optimised Intra-4×4 Mode-Decision for H.264/AVC Video Encoding," Proceedings of the 6th International Symposium on Image and Signal Processing and Analysis (2009)

[16] A.Elyousfi, A.Tamtaoui and E.Bouyakhf, "Fast Intra Prediction Algorithm for H.264/AVC Based on Quadratic and Gradient Model," World Academy of Science, Engineering and Technology 63, 2010.

[17] Jeon B and J. Lee, "Fast Mode Decision for H264", ISO/IEC JTC1/SC29/WG11 and ITU-T SG16, Input Document JVT-J033 (2003, December).

[18] The H.264/AVC encoder reference software JM 17.1: http://iphome.hhi.de/suehring/tml/download/old_jm/