Fast Intra Prediction Algorithm for H.264/AVC Based on Quadratic and Gradient Model

A. Elyousfi, A. Tamtaoui, E. Bouyakhf

Abstract—The H.264/AVC standard uses an intra prediction, 9 directional modes for 4x4 luma blocks and 8x8 luma blocks, 4 directional modes for 16x16 macroblock and 8x8 chroma blocks, respectively. It means that, for a macroblock, it has to perform 736 different RDO calculation before a best RDO modes is determined. With this Multiple intra-mode prediction, intra coding of H.264/AVC offers a considerably higher improvement in coding efficiency compared to other compression standards, but computational complexity is increased significantly. This paper presents a fast intra prediction algorithm for H.264/AVC intra prediction based a characteristic of homogeneity information. In this study, the gradient prediction method used to predict the homogeneous area and the quadratic prediction function used to predict the nonhomogeneous area. Based on the correlation between the homogeneity and block size, the smaller block is predicted by gradient prediction and quadratic prediction, so the bigger block is predicted by gradient prediction. Experimental results are presented to show that the proposed method reduce the complexity by up to 76.07% maintaining the similar PSNR quality with about 1.94%bit rate increase in average.

Keywords—Intra prediction, H.264/AVC, video coding, encoder complexity.

I. INTRODUCTION

THE H264/AV C is newest video coding standard of the ITU-T Video Coding Experts Group and the ISO/IEC Moving Picture Experts Group [1,2,3,4]. Compared to the previous video coding standards, H.264/AVC has significantly better performance in terms of better peak signal-to-noise ratio (PSNR) and visual quality at the same bit rate [5]. This is accomplished mainly due to the consideration of variable block sizes for motion compensation, multiple reference frames, integer transform [an approximation to discrete cosine transform (DCT)], in-loop deblocking filter, context-based adaptive binary arithmetic coding (CABAC), but also due to better exploitation of the spatial correlation that may exist between adjacent Macroblocks, with the multiple intra-mode prediction in intra (P) slices.

The Intra prediction is a necessary part for any video coding standard. Including MPEG-1, MPEG-2, H.261, H.263, MPEG-4 part 2 and H.264/AVC. But the intra prediction in H.264/AVC is different from the one in the other standards. In the previous standards, the intra prediction is used in the transform domain [6,7,8]. For example, in *MPEG*-1, *MPEG-2* and *H.*261 the intra prediction is DC prediction and in *MPEG*-4 and *H.*263, additional AC prediction is used. However, in *H.*264/*AVC*, the intra prediction is conducted by using spatially neighbouring samples of a given block, which are already transmitted and decoded.

The *H*.264 video coding standard supports intra prediction for various block sizes. For coding the luma signal, one 16 x 16 macroblock may be predicted as a whole using Intra-16 x 16 modes, or the macroblock can be predicted as individual 4x4 blocks using nine Intra-4x4 modes. In the profiles that support Fidelity Range Extension (*FRExt*) tools, a macroblock may also be predicted as individual 8x8 blocks using nine intra-8x8 modes [9]. In this paper we only discuss the typical block type intra 4 x 4 and intra 16 x 16. Intra prediction for the chroma signal uses similar techniques as those for luma Intra-16x 16 predictions. *H*.264/*AVC* uses rate-distortion optimization (*RDO*) [10] technique to obtain the best result maximizing visual quality and minimizing bitrates. To choose the best macroblock mode, *H*.264 encoder calculates the RDcost (Rate distortion cost) of every possible mode and chooses the mode having the minimum value. Therefore, the computational complexity is extremely increased compared with previous standards, so it makes *H*.264/*AVC* difficult for applications with low computational capability, such as mobile devices.

To reduce this complexity, few approaches have been proposed on fast intra prediction algorithm. In (Pan et al., 2003, 2005)[11,12], it is based on the local edge information, and thus adopts the edge direction to predict the possible mode. In (Jongho Kim* and Jechang Jeong, 2005)[13], the directional masks and mode information of neighbouring blocks used to select the probable modes. In (Jun Sung Park, and Hyo Jung Song, 2006)[14], it is based on the idea that direction of a bigger block is similar to that of smaller block, the effects of fast mode decision is reduced. In this paper, we present a fast intra prediction algorithm using a gradient prediction function and a quadratic prediction function to improve the encoding speed without much sacrifice at RDO performance. Based a characteristic of homogeneous areas, the homogeneous area is predicted by the gradient prediction

function. Since, the nonhomogeneous area is predicted by the quadratic prediction function.

The rest of this paper is organized as follows: Section 2 describes the intra mode decision in *H.264/AVC*. In Section 3, we propose a fast intra prediction algorithm in detail. Section 4 gives the Experimental results to show the performance of the proposed algorithm. Finally, the paper is concluded in section 5.

II. PROCEDURE FOR PAPER SUBMISSION

The H.264 standard exploits the spatial correlation between adjacent macroblocks/blocks for Intra prediction. In JV T, the current macroblock is predicted by adjacent pixels in the upper and the left macroblocks that are decoded earlier. For the luma prediction samples, the prediction block may be formed for each 4 x 4 subblock or for a 16 x 16 macroblock. One case is selected from a total of 9 prediction modes for each 4 x 4 luma block, 4 modes for a 16 x 16 luma block and 4 modes for each chroma blocks. To take the full advantage of these modes, the H.264 encoder can select the best mode using the rate distortion optimization (RDO).

A. 4 x 4 luma intra prediction modes

In 4x4 Intra prediction modes, the values of each 4x4 block of luma samples are predicted from the neighboring pixels above or left of a 4 x 4 block, and nine different directional ways of performing the prediction can be selected by the encoder as illustrated in Fig. 1 and Table I (and including a DC prediction type numbered as mode 2, which is not shown in the figure).

TABLE I
INTRA 4 X 4 PREDICTION MODES

Num	INTRA 4 X 4 PRED MODE
0	vertical
1	horizontal
2	DC
3	Diagonal-down-left
4	Diagonal-down-right
5	Vertical-Right
6	Horizontal-down
7	Vertical-left
8	Horizontal-up

Each prediction direction corresponds to a particular set of spatially-dependent linear combinations of previously decoded samples for use as the prediction of each input sample. For the purpose of illustration, Fig.1(a) shows a 4 x 4 block of pixels a, b, c ...p, belonging to a macroblock to be coded. Pixels A, B, C ...H, and I, J, K, L, M are already decoded neighboring pixels used in computation of prediction of pixels of current 4x4 block. Fig.1(b) depicts the eight directional modes. The vertical mode extrapolates a 4x4 block vertically with 4 neighbouring pixels, A, B, C, D, whereas the horizontal member utilizes the horizontal adjacent pixels, I, J, K, L to perform the prediction. With the exception of DC, the other modes operate in a similar manner, according to their

corresponding orientations. The direction less member, DC, extrapolates all pixels as (A+B+C+D+I+J+K+L)/8[1,2,3].
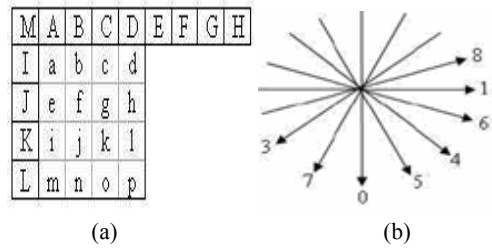

(a)   (b)
Fig. 1 prediction directions

B. 16 x 16 luma and 8 x 8 chroma intra prediction modes

The 16 x 16 luma intra prediction modes are selected in relatively homogeneous area, four prediction modes are supported as listed in Table II comprising of the dc, vertical, horizontal and plane prediction. These modes are specified similar to modes in Intra-4x4 prediction except the plane prediction. In vertical prediction, each of the 16 columns (of 16 pixels each) of current macroblock are predicted using only 1 past decoded pixel each, similar to the case of prediction of 4 pixels of column by a single decoded pixel in the case of 4 x 4 intra prediction. The horizontal prediction predicts an entire row of 16 pixels by a past decoded neighboring pixel. The process is repeated for each of the 16 rows. The dc prediction uses an average of past decoded row and column of pixels to predict all pixels of the 16 x 16 block. The planar prediction uses weighted combination of horizontal and vertical adjacent pixels [4,5]. The neighboring pixels used for prediction of 16 x 16 luminance component of current macroblock belong to neighboring decoded macroblock. For the chrominance (chroma) components, there are 4 prediction modes that are applied to the two 8 x 8 chroma blocks (U and V), which are very similar to the 16x16 luma prediction modes such as DC (Mode 0), horizontal (Mode 1), vertical (Mode 2), and plane (Mode 3). To take the full advantage of these modes, the H.264 encoders select the best mode using the rate distortion optimization (RDO) technique [10].

TABLE II
INTRA 16x16 PREDICTION MODES

Num	Intra 16x16 prediction mode
0	Vertical
1	Horizontal
2	DC
3	Plan

C. Best mode selection using rate-distortion optimization (RDO)

The RDO mode decision method finds the optimal mode in the RD sense. To be more precise, it exhaustively searches the best mode by measuring the RD cost based on the actual rate and distortion after entropy coding and reconstruction, respectively. The RDO procedure to encode one macroblock

with intra mode is described below [13]:

Step 1: Select the best intra-mode for a 4x4 luma block among 9 modes by minimizing the following equation (1)

$$J(s, c, \mathrm{mod}\, e\, / Qp, \lambda_{\mathrm{mod}\, e}) = SSD(s,\, c,\, \mathrm{mode}/Qp) +$$
$$\lambda_{\mathrm{mode}} * R(s,\, c,\, \mathrm{mode}/Qp)$$

(1)

where QP is the macroblock quantized, λ_{mode} is the Lagrangian multiplier for mode decision, **mode** indicates one of the 9 prediction modes of a 4x4 luma block. SSD is the sum of the squared differences between the original 4x4 block luminance signal denoted by s and its reconstructed signal denoted by c, and R(s, c, MODE / QP) represents the rate, i.e., the number of bits associated with choosing **mode**. It includes the bits for the intra prediction mode and the transformed coefficients for the 4x4 luminance block. Repeat this procedure for sixteen 4x4 luma blocks of a macroblock.

Step 2: As contrary to the RDO technique for intra 4x4 luma block mode decision, determine the best intra-mode for a 16x16 macroblock among 4 modes by choosing the mode that results in the minimum SATD (Sum of Absolute Transformed Difference).

Step 3: Determine the best intra-mode for a 8x8 chroma block among 4 modes by minimizing the following,

$$J_{ch}(s,\, c,\, \mathrm{mode}/Qp,\, \mathrm{,mode}) =$$
$$SSD(s,\, c,\, \mathrm{mode}/Qp)\, + \lambda_{mode} * R(s,\, c,\, \mathrm{mode}/Qp)$$

(2)

Where s represents the original U or V chrominance values of an 8x8 chroma block, R(s, c, MODE/QP), in this case includes only the bits for the transformed coefficients.

Step 4: Compare the RD cost for the two best modes, i.e., the 4x4 mode obtained from Step 1 and the 16x16 mode from Step 2, and choose the best one as the macroblock prediction mode.

H.264/AVC encodes the MB by iterating the entire luma intra prediction mode 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 N8-chr*(N4*16 + N16), where N8-chr, N4, and N16 represent the number of modes for 8x8 chroma blocks, 4x4 and 16x16 luma blocks, respectively. It means that, for an MB, it has to perform 4*(9*16 + 4) = 592 different RDO calculations before a best RDO mode is determined [11, 12]. As a result, the complexity of the encoder is extremely high. To reduce the encoding complexity with little RD performance degradation, a fast intra prediction algorithm is proposed in the next section.

III. BACKGROUND AND THE PROPOSED METHOD

One of the reasons for adopting multiple different block sizes in h.264/AVC intra coding is to represent homogeneous information when the fixed block size is used. So the multiple direction intra modes are used for each size block to reduce the prediction error. H.264/AVC standard checks all possible intra prediction modes of every block which belongs to P-frame as well as I-frame in order to achieve optimal coding efficiency. The Rate Distortion optimization (RDO) process of intra modes is quite complex. To reduce this complexity, several researches have been proposed on fast intra prediction algorithm, by reducing the number of candidate for the best intra prediction mode and saving the directional prediction mode used in this standard. In this paper, we propose a fast intra prediction algorithm to improve the encoding speed without much sacrifice at the RD performance. This algorithm is described as comprising two prediction methods: the first is the gradient prediction method it used to predict the homogeneous area, the second method is the quadratic prediction method it used to predict the nonhomogeneous area.

A. The quadratic prediction method

For the not homogenous block, the difference between pixels have large. Therefore, a good prediction could be achieved if we predict the pixels using those neighboring pixels that are in the same direction of the edge. H.264/AVC tested a number of prediction direction to selected one of direction by computing the RDO of each prediction direction, but this method not valuable of real-time application. So, our proposed based the idea that the block is represented a as same surface, when the pixel of neighboring block are the elements of this surface a values know, since, the pixels of the block to coded are the elements of the surface a values unknown. In this case, a quadratic prediction function is used to determine the intra prediction error of a variable size block in H.264.
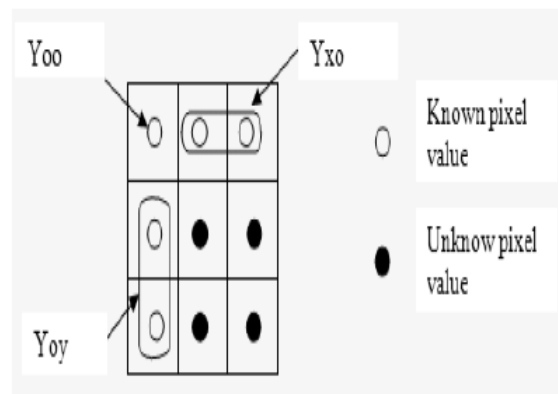

Fig. 2 Nine pixel of quadratic prediction

In H.264, the best intra prediction mode of a variable size block is determined by minimizing the RDO, which contains both the SAD and bit-rate values. Based a characteristic of nonhomogeneous area, we see that it reasonable to consider the block as same a surface. From this observation, a quadratic prediction function is used to the predictions of the pixels, which is given by (3)

$$F(x, y) = Ax^2 + Bx + Cy^2 + Dy + E \qquad (3)$$

where x and y are the offsets to the up-left pixel to the current block, while A, B, C, D and E are parameters to be determined. As shown in Fig. 2, the intra prediction area contains 9 pixels, where Y11, Y12, Y21 and Y22 are the pixels to predicted at locals coordinates (1,1), (1,2), (1,1) and (2,2) respectively. Since, Yxo and Yoy represent the pixels of top line and left column to the current block respectively, and the five corresponding matching errors, denoted as F(0,0), F(0,1), F(0,2), F(1,0), and F(2,0), are computed in the intra prediction procedure. So the prediction of these pixels are completely determined with quadratic prediction function, which can be employed to determine the five parameters A, B, C, D and E in eq (3), then based on eq (3), we have:

$$\begin{cases} F(0,0) = E \\ F(0,1) = C + D + E \\ F(0,2) = 4C + 2D + E. \\ F(1,0) = A + B + E \\ F(2,0) = 4A + 2B + E \end{cases}$$

If for simplicity, let

$$\begin{cases} I = F(0,2) - 2F(0,1) \\ J = F(0,2) - 4F(0,1) \\ K = F(2,0) - 2F(1,0) \\ L = F(2,0) - 4F(1,0) \end{cases}$$

The solution of this equation is presented as fellow:

$$\begin{cases} A = (F(2,0) - 2F(1,0) + F(0,0)) / 2 \\ B = (4F(1,0) - F(2,0) - 3F(0,0)) / 2 \\ C = (F(0,2) - 2F(0,1) + F(0,0)) / 2 \\ D = (4F(0,1) - F(0,2) - 3F(0,0)) / 2 \\ E = F(0,0) \end{cases}$$

With A, B, C, D and E computed, the current block can be predicted by (3). Those parameters are varied with a characteristic of area. So, the quadratic prediction function is valid for more case of information distributes:

1) In case of horizontal homogeneity, all pixels of each line are the same value, since we have:

$$\begin{cases} Y\ 00 = Y\ 10 = Y\ 20 \Rightarrow F(0, 0) = F(1, 0) = F(2, 0) \\ \text{and} \\ F(0, 0) = F(1, 0) = F(2, 0) \Rightarrow A = B = 0. \end{cases}$$

Then each pixels of this area is predicted by (4)

$$F(x, y) = Cy^2 + Dy + E \qquad (4)$$

From this equation, we observe that the prediction function is depend only with y, since all pixels of the same lines are predicted by the same value, because, these pixels are the same y value.

2) In case of vertical homogeneity, all pixels of each rows are the same value, since we have:

$$\begin{cases} Y\ 00 = Y\ 01 = Y\ 02 \Rightarrow F(0, 0) = F(0, 1) = F(0, 2) \\ \text{and} \\ F(0, 0) = F(0, 1) = F(0, 2) \Rightarrow C = D = 0. \end{cases}$$

Then each pixels of this area is predicted by (5)

$$F(x, y) = Ax^2 + Bx + E \qquad (5)$$

From this equation, we observe that the prediction function is depend only with x, since all pixels of the same rows are predicted by the same value, because, these pixels are the same x value.

3) In case of homogeneous area, all pixels of this area are the same value, since we have:

$$\begin{cases} Y\ 00 = Y\ 10 = Y\ 20 = Y\ 01 = Y\ 02 \Rightarrow \\ F(0, 0) = F(0, 1) = F(0, 2) = F(1, 0) = F(2, 0) \\ \text{and} \\ F(0, 0) = F(0, 1) = F(0, 2) = F(1, 0) = F(2, 0) \\ \Rightarrow A = B = C = D = 0 \end{cases}$$

Then each pixels of this area is predicted by (6)

$$F(x, y) = E \qquad (6)$$

From this equation, we observe that the prediction function is not depending with pixels location, since all pixels of this area are predicted by the same value E.

4) In case of diagonal left direction pixels, all pixel of each diagonal direction are the same value, since we have:

$$\begin{cases} F(1, 0) = F(0, 1) \Rightarrow A = C \\ \text{and} \\ F(2, 0) = F(0, 2) \Rightarrow B = D \end{cases}$$

Then each pixels of this area is predicted by (7)

$$F(x, y) = A(x^2 + y^2) + B(x + y) + E \qquad (7)$$

From this equation, we observe that the prediction function is depending with the offset to the up-left pixel to the current block and the three parameters A, B and C.

The quadratic prediction function has five parameters, we

observe that those parameters can be obtaining the positive, negative or null value, since this variety of the parameters is depend a characteristic of the area. So to ameliorate this prediction, we use the gradient prediction as described below.

B. The gradient prediction method

In case of homogeneous block, the textures in the region have similar spatial property. The pixels along the direction of local edge normally have similar values. Since, the current block has strong correlations with neighboring blocks. Thus it can be supposed that the pixels of the current block and their neighboring pixels are on a same gradient plane. With the use of the characteristic of smoothness block, we propose a special gradient prediction technique to further improve the coding efficiency with low complexity. Thus the predictions of these pixels Px,y can be depicted by (8)

$$P_{x,y} = Y_{x,0} + \delta_{x,y} \tag{8}$$

where gx,y, is the variation by pixel Yo,o and Yo,y as defined by (9)

$$\delta_{x,y} = Y_{0,y} + Y_{0,0} \tag{9}$$

It should be mentioned that this method is valid for some mode direction employed by h.264/AVC. We can simplicity by example of horizontal and vertical mode prediction. As shown in fig.3, the area can be divided in two ways. The dashed lines represent the 8x8 boundaries and the rectangular points indicate the reference pixels in the upper block and the left block. The circular points show the actual pixels in the current block. The 8x8 block area is divided into four areas by
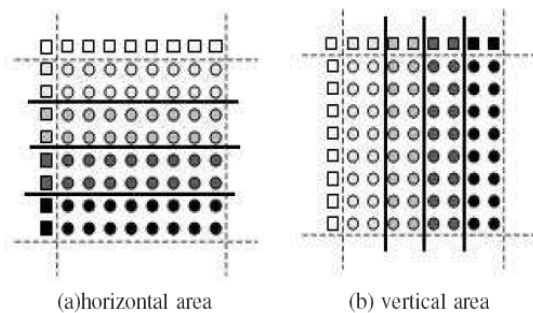

(a)horizontal area   (b) vertical area
Fig. 3 Horizontal and vertical areas

the thick line. When it is divided horizontally as shown in Fig.3(a), each area is predicted from the left block. When it is divided in vertically as shown in fig.3(b), each area is predicted from the upper block. So, our gradient prediction technique is valid in these two mode prediction, horizontal and vertical prediction mode:

1) In case of horizontal homogeneity, all pixels of each lines are the same value, since the neighboring pixel of the upper block as a same values, then we have:

$$\delta_{x,y} = Y_{x,0} - Y_{0,0} = 0 \tag{10}$$

or

$$P_{x,y} = Y_{x,0} + \delta_{x,y} \tag{11}$$

$$= Y_{x,0} + Y_{0,y} - Y_{o,o} \tag{12}$$

then

$$P_{x,y} = Y_{0,y} \tag{13}$$

From the horizontal homogeneous area, with our gradient prediction, we show that all pixels from the same line are predicted by the first pixel of this line. So the gradient prediction method and the mode horizontal used by h.264/AVC are the similar prediction.

2) In case of vertical homogeneity, all pixels of each rows are the same value, since the neighboring pixel of the left block as a same values, then we have:

$$\delta_{x,y} = Y_{0,y} - Y_{0,0} = 0 \tag{14}$$

or

$$P_{x,y} = Y_{0,0} + \delta_{x,y} \tag{15}$$

$$= Y_{x,0} + Y_{0,y} - Y_{0,0} \tag{16}$$
then

$$P_{x,y} = Y_{x,0} \tag{17}$$

From the vertical homogeneous area, with our gradient prediction, we show that all pixels from the same rows are predicted by the first pixel of these rows. So the gradient prediction method and the mode vertical used by h.264/AVC are the similar prediction.

3) In case of homogeneous area, all pixels of the area are the same value, since we have:

$$Y_{0,y} = Y_{x,0} = Y_{0,0} \tag{18}$$

or

$$P_{x,y} = Y_{x,0} + \delta_{x,y} \tag{19}$$

$$= Y_{x,0} + Y_{0,y} - Y_{0,0} \tag{20}$$

then

$$P_{x,y} = Y_{0,0} \tag{21}$$

The gradient function prediction is used to predict the homogeneous area, since this method is valid for two mode prediction used by H.264, the Horizontal and the vertical mode. So h.264 calculates the RDO for each mode contracted

a gradient method where the RDO is calculus only once.

C. Proposed luma and chroma spatial prediction

The Intra coding block size is highly dependent on the smoothness of the block. The smaller block is well suited for an MB with detailed information, while the bigger block is will suited for a smooth one. Fig.4 shows two example frames
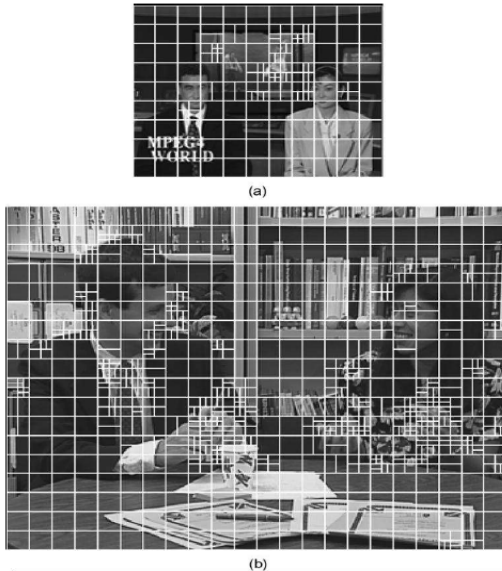

Fig. 4 Block sizes chosen after H.264 RDO. (a) News. (b) Paris. [15]

from the QCIF sequence News and the CIF sequence Paris, respectively [15].The different size of white boxes overlay on the images represent the different intra coding block size that are determined by RDO in h.264/AVC. We can see the correlation between the smoothness and the intra coded block size. From this observation, we detailed the both method which is depend with a characteristic of area.

1) Intra prediction for 4x4 luma blocks.

The block 4x4 has a smaller size and then it is beneficial for detailed areas. Since, this block can contain homogeneous

Fig. 5 4x4 intra prediction with quadratic function

information as well nonhomogeneous information. The Mode 2 (DC) of 4 x 4 blocks does not have direction and they use the mean value of adjacent block, since this mode has a higher probability to be the best prediction mode of the nine modes [14]. For 4x4 intra block, we applied the gradient prediction method, the quadratic prediction function and the mode DC. So, in case of 4x4 intra prediction with quadratic function, the block 4x4 is divided into four areas and each area is applied individually. The different areas and those parameters of quadratic prediction function are supported as shown in fig.5, where the black and white dot represent the pixels of the current and neighboring block respectively. The prediction values for pixels are calculated by quadratic prediction function. Actually, they are different on parameters of prediction quadratic function.

2) Intra prediction for 16x16 luma macroblocks.

The macroblock 16x16 is a comparatively larger area, since the pixel correlations in such area are not as high as those in 4x4 blocks. Therefore, no much gain can be achieved even if the information from up-left 16x16 blocks is also used in Intra-16x 16 quadratic predictions. In h.264/AVC, 16x16 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, there are 4 directional modes, different from the case of 4x4 luma blocks, such as horizontal, vertical, DC and plane mode. So for fast 16x16 intra prediction, we have implemented the 16x16 intra prediction algorithm used the gradient prediction method and the modes plan used by H.264/AVC.

3) Intra prediction for 8 x 8 Chroma block.

The chroma intra prediction mode of H.264 consists of four modes: DC, horizontal, vertical, and plane mode. These modes are the same as for the luma component in Intra-16x16 macroblocks. So, the statistical characteristics of luma and chroma components are very different. The variance of the chroma pixel values is much smaller than that of the luma ones. The luma pixel values gradually change according to the angle between the direction of the source light and the surface of the object. On the other hand, the chroma pixel values in a certain area change very little [16]. With the uses of the characteristic of chroma and the subject to reduce the complexity with better quality, we use only the DC mode for chroma intra prediction.

Joint model (JM) reference software provided by joint video team (JVT) performs the intra mode prediction after the inter mode prediction procedure [17]. Also, the RDO option is supported. This structure gives good results for implementation of our algorithm in JM software. The proposed algorithm has low computational complexity compared with other schemes [11,..14], because it requires low number of RDO calculation.

D. Analysis of Computational Complexity

Table III summarizes the number of candidates selected for RDO computation by the proposed method. As can be seen from Table III, the encoder with the fast intra prediction
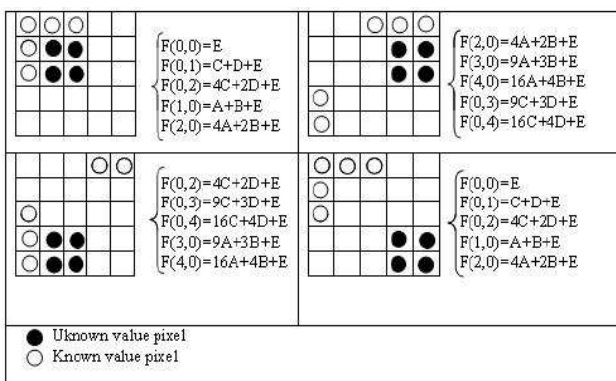
algorithm would need to perform only 1*(3*16 + 2) = 50. Thus our proposed algorithm reduces number of RDO calculation significantly compared to the 1 *(4 *16+4) = 68, 2*(4*16 + 2) = 132 and 4*(9*16 + 4) = 592 modes that are used in the current RDO calculation in H.264/AVC video coding with (J. S. Park and H. J. Song 2006), (F. Pan, X. Lin,..2005) and full search method of H.264 respectively.

TABLE III
COMPARISON OF THE NUMBER OF THE CANDIDATE MODES

	Block size	Total number of modes	Number of proposed
Luma(Y)	4x4	9	3
Luma(Y)	16x16	4	2
Chroma(U,V)	8x8	4	1

IV. EXPERIMENTAL RESULTS

To evaluate the performance of the proposed method, our proposed method was implemented into H.264/AVC reference software JM10.1 [17] and tested with various Quantization Parameters QP, The system platform is the Intel Pentium 4 Processor of speed 1.8GHz, 512MB DDR RAM, and Microsoft Windows XP. The test conditions are as follows: (1) MV search range is 16 pixels for QCIF, CIF, (2) RD optimization is enabled, (3) Reference frame number equals to 5, (4) CABAC is enabled, (5) GOP structure is IPPPP or I-frame only, (6) the number of frames in a sequence is 100. Comparisons with the case of exhaustive search were performed with respect to the change of average PSNR (4PSNR), the change of average data bits (4Bit), and the change of average encoding time (4Time), respectively.

In order to evaluate the timesaving of the fast intra prediction algorithm, the following calculation is defined to find the time differences. Let T_{JM} denote the coding time used by JM10.1 encoder and T_{FI} be the time taken by the fast intra mode decision algorithm, and is defined as:

$$\Delta Time = \frac{T_{FI} - T_{JM}}{T_{JM}} *100\% \qquad (22)$$

A group of experiments were carried out on the recommended sequences with different quantization parameters as defined bellow. The average ¢Bits is the bit rate difference expressed as a percentage between JM10.1 [17] encoder and the fast algorithm, respectively. The PSNR values of luma (Y) and chroma (U, V) is used which is based on the equations below:

$$PSNR = 10 \log_{10}(255^2 / MSE) \qquad (23)$$

where MSE is the mean square error.

The differences between PSNR and bit rate are calculated according to the numerical averages between the RD-curves derived from the JM 10.1 original encoder and the proposed fast algorithm, respectively. The detailed procedures for calculating these differences can be found in a JVT document

by Bjontegaard [18], which is recommended by the JVT Test Model Ad Hoc Group [19].

Table IV shows the simulation results of the proposed algorithm with JM10.1 for various sequences with I-only type. The quantization parameter set was chosen to be [10,14,18,...42,46]. We also show the simulation results for IPPPP type sequences in Table V. Note that in the tables, positive values mean increments and negative values mean decrements. It can be seen that the proposed algorithm

TABLE IV
SIMULATION RESULTS FOR I-FRAME-ONLY SEQUENCES

Sequences	Δ Y-PSNR (dB)	Δ UV-PSNR (dB)	Δ Bit-rate (%)	Δ Time (%)
mobile	-0,257	-0,031	+3,69	-75,09
silent	-0,169	-0,021	+4,19	-74,27
salesman	-0,232	-0,184	+3,59	-75,07
mthrdotr	-0,227	-0,068	+5,29	-75,86

TABLE V
SIMULATION RESULTS FOR IPPPP SEQUENCES

Sequences	Δ Y-PSNR (dB)	Δ UV-PSNR (dB)	Δ Bit-rate (%)	Δ Time (%)
Claire	-0,016	-0,202	+3,53	-75,61
silent	-0,065	-0,052	+1,99	-74,97
news	-0,322	-0,104	+4,62	-76,39
mthrdotr	-0,404	-0,115	+1,94	-76,07
salesman	-0,223	-0,077	+2,01	-75,63
grandma	-0,233	-0,080	+0,77	-75,69

achieves very high encoding time saving (up to about 76%) with a little loss of PSNR and increment of bitrates.

Tables VI and VII show the tabulated performance comparison of our proposed with the full search method for the image sequences described below with I-frame-only and IPPP type sequences, respectively. The quantization parameter set was chosen to be [10,14,18,...,42,46]. Experimental results of the proposed method show a significant reduction of computation in between 70.00%, and 77.78%, a slight increase in bit rate in between 0.75% and 6.63%, and similar PSNR in comparison with full search method.

TABLE VI
RESULTS ON QCIF TEST I-ONLY SEQUENCE "SILEN"

QP	Δ Y-PSNR (dB)	Δ UV-PSNR (dB)	Δ Bit-rate (%)	Δ Time (%)
10	0,350	0,120	+4,45	-70,36
14	0,480	0,110	+2,03	-77,78
18	0,130	0,065	+2,48	-77,23
22	0,110	0,035	+3,26	-76,66
26	0,090	0,055	+3,95	-76,04
30	0,080	-0,040	+5,60	-75,02
34	0,070	0,060	+6,36	-73,80
38	0,070	-0,030	+6,63	-73,05
42	0,130	-0,030	+5,14	-71,96
46	0,180	-0,130	+1,94	-70,81

TABLE VII
RESULTS ON QCIF TEST IPPP SEQUENCE "SILEN"

QP	Δ Y-PSNR (dB)	Δ UV-PSNR (dB)	Δ Bit-rate (%)	Δ Time (%)
10	-0,66	0,135	+2,36	-77,87
14	0,40	0,145	+1,40	-77,20
18	0,12	0,000	+3,31	-70,00
22	0,20	0,150	+1,95	-76,59
26	0,22	0,08	+1,765	-75,63
30	0,13	0,055	+2,45	-74,94
34	0,07	0,085	+0,20	-74,36
38	0,12	0,060	+1,35	-74,18
42	0,07	-0,080	-0,75	-73,88
46	-0,02	-0,110	+5,85	-75,10

Fig. 6,7,8 shows the RD performance and the computation time for the IPPP sequences "SILEN". It can be seen from these figures that the two RD curves, one from the original full search and the other from the proposed algorithm, are almost overlapping each other. It means that the performance of the proposed algorithm is almost similar to that of the original full-search. From Fig. 8 we can observe that the encoding time with fast intra prediction is distinctly less than that of without full search under the same test conditions.
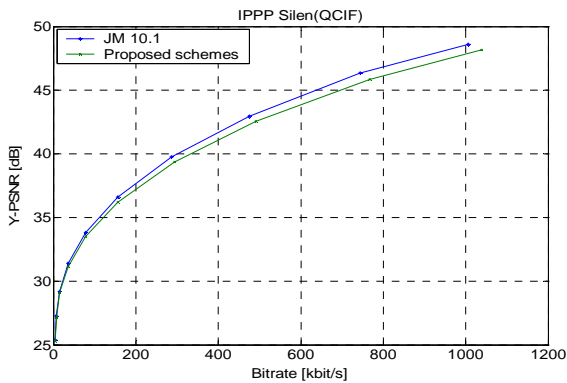

Fig. 6 Comparison of PSNR-Y and bit-rate for the IPPP sequences of silen
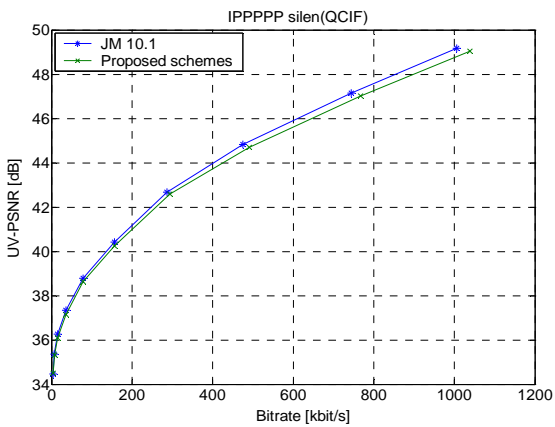

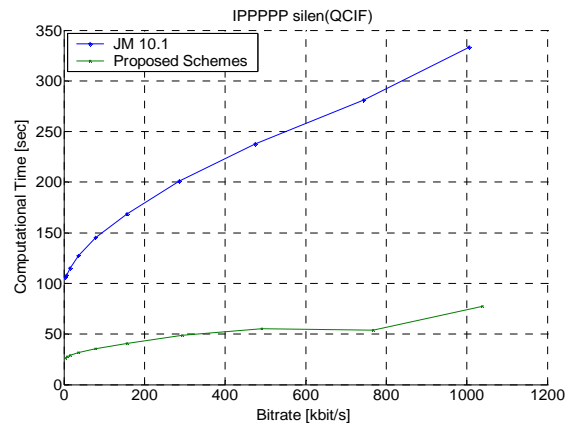Fig. 7 Comparison of PSNR-UV and bit-rate for the IPPP sequences of silen

Fig. 8 The computational time comparison of silen IPPP sequence

Fig. 9,10,11 shows the RD performance and the computation time for the I-only sequences "SILEN". It can be seen from these figures that the two RD curves, one from the original full search and the other from the proposed algorithm, are almost overlapping each other. It means that the performance of the proposed algorithm is almost similar to that of the original full-search. From Fig. 11 we can observe that the encoding time with fast intra prediction is distinctly less than that of without full search under the same test conditions.
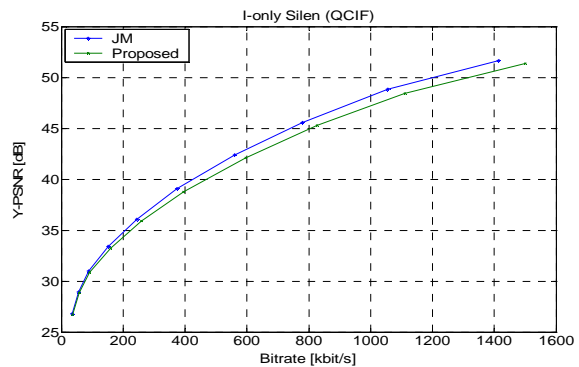

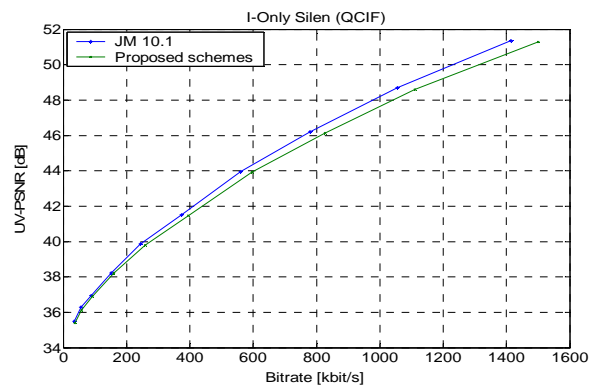Fig. 9 Comparison of PSNR-Y and bit-rate for the silen I-only sequences

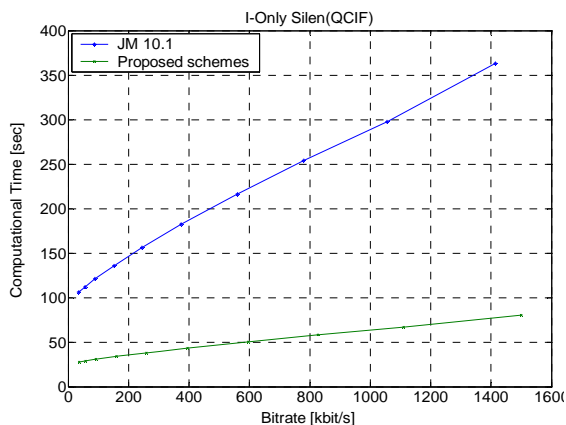Fig. 10 Comparison of PSNR-UV and bit-rate for the silen I-only sequences

Fig. 11 The computational time comparison of silen I-only sequence

V. CONCLUSION

In this paper, a fast intra prediction algorithm for H.264 video coding is proposed based on the gradient prediction method and the quadratic prediction function. With our method, the number of mode combinations for luma and chroma blocks in an MB that takes part in RDO calculation has been reduced significantly from 592 to as low as 50. From the experimental results, we can see that the proposed method can achieve a considerable reduction of computation complexity while maintaining similar bit rate and PSNR.

REFERENCES

[1] ITU-T Recommendation H.264 and ISO/IEC 14496-10 (MPEG-4) AVC, "Advanced Video Coding for Generic Audiovisual Services," (version 1: 2003, version 2: 2004) version 3: 2005

[2] T Wiegand, G Sullivan, G Bjntegaard, and A Luthra, "Overview of the H.264/AVC Video Coding Standard", IEEE transactions on circuits and systems for video technology, vol. 13, pp. 560-576, July 2003

[3] A Puri, X Chen, A Luthra, "Video coding using the H.264/MPEG-4 AVC compression standard," Signal Processing: Image Communication 19 (2004) 793-849.

[4] Iain E. G. Richardson, "H.264 and MPEG4 Video Compression: Video Coding for Next Generation Multimedia", John Wiley and Sons, 2003

[5] "Report of the formal verification tests on AVC (ISO/IEC 14 496-10 - ITU-T Rec. H.264)", MPEG2003/N6231, Dec. 2003.

[6] B. Erol, M. Gallant, G. Ct and F. Kossentini," The H.263+ Video Coding Standard: Complexity and Performance", IEEE Data Compression Conference, Snowbird, Utah, pp. 259-268, March 1998.

[7] ISO/IEC IS 13818, Information Technology - Generic coding of moving pictures and associated audio information, Part 2: Video. ISO/IEC JTC1/SC29/WG11 (2004)

[8] ITU-T and ISO/IEC JTC 1, "Generic coding of moving pictures and associated audio information - Part 2: Video," ITU-T Recommendation H.262 - ISO/IEC 13818-2 (MPEG-2), Nov. 1994

[9] G. J. Sullivan, P. Topiwala, 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.

[10] G. Sullivan and T. Wiegand, "Rate Distortion Optimization for Video Compression," IEEE Signal Processing Magazine, pp. 74-90, Nov' 98

[11] F. Pan, X. Lin, S. Rahardja, K.P. Lim, Z.G. Li, G.N. Feng, D.J. Wu, and S. Wu, "Fast mode decision for intra prediction," JVT-G013, 7th JVT Meeting, Pattaya, Thailand, Mar. 2003.

[12] 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.

[13] J. Kim, J. Jeong, "Fast intra-mode decision in H.264 video coding using simple directional masks", VCIP 2005, of proceedings of SPIE Vol. 5960, pp.1071-1079.

[14] J. S. Park, and H. J. Song, "Selective Intra Prediction Mode Decision for H.264/AVC Encoders", Transactions on Engineering, Computing and Technology Volume 13 May 2006, pp.51-55.

[15] D. Wu, F. Pan, K. P. Lim, S. Wu, Z. G. Li, X. Lin, S. Rahardja, and C. C. Ko "Fast Intermode Decision in H.264/AVC Video Coding", IEEE Transactions On Circuits And Systems For Video Technology, Vol. 15, No. 6, July 2005

[16] W. S. Kim, D. S. Kim and K. W. Kim, "Complexity Reduction of Plane mode in Chroma Intra Prediction", presented at the 5th JVT meetig(JVTE050), Geneva, CH, 9-17 October, 2002.

[17] JM Reference Software Version 10.1 http://iphome.hhi.de/suehring/tml/download/.

[18] G. Bjontegaard, Calculation of Average PSNR Differences between RDcurves, presented at the 13-th VCEG-M33 Meeting, Austin, TX, Apr. 2001.

[19] JVT Test Model Ad Hoc Group, Evaluation Sheet for Motion Estimation, Draft version 4, Feb. 2003.