Journal of Zhejiang University SCIENCE A
ISSN 1673-565X (Print); ISSN 1862-1775 (Online)
www.zju.edu.cn/jzus; www.springerlink.com
E-mail: jzus@zju.edu.cn

Dominant edge direction based fast intra mode decision in the H.264/AVC encoder*#

Byeongdu LA, Minyoung EOM, Yoonsik CHOE‡

(*Department of Electrical and Electronic Engineering, Yonsei University, Seoul 120-749, Korea*)

E-mail: {bdla6077, emy0606, yschoe}@yonsei.ac.kr

Received June 21, 2008; Revision accepted Oct. 28, 2008; Crosschecked Mar. 31, 2009

Abstract: The H.264/AVC video coding standard uses an intra prediction mode with 4×4 and 16×16 blocks for luma and 8×8 blocks for chroma. This standard uses the rate distortion optimization (RDO) method to determine the best coding mode based on the compression performance and video quality. This method offers a large improvement in coding efficiency compared to other compression standards, but the computational complexity is greater due to the various intra prediction modes. This paper proposes a fast intra mode decision algorithm for real-time encoding of H.264/AVC based on the dominant edge direction (DED). The DED is extracted using pixel value summation and subtraction in the horizontal and vertical directions. By using the DED, three modes instead of nine are chosen for RDO calculation to decide on the best mode in the 4×4 luma block. For the 16×16 luma and the 8×8 chroma, only two modes are chosen instead of four. Experimental results show that the entire encoding time saving of the proposed algorithm is about 67% compared to the full intra search method with negligible loss of quality.

Key words: H.264, Intra prediction, Mode decision, Dominant edge direction
doi:10.1631/jzus.A0820473 **Document code:** A **CLC number:** TN919.8

INTRODUCTION

In recent years, the multimedia applications environment has grown and become more complicated due to the variety of broadcasting and communication systems. Real-time encoding is one important requirement for video processing and channel coding. In response to different requirements in many environments, video compression standards such as MPEG-2, MPEG-4, H.263, and H.263+ have been created for more effective video coding, transmission, and storage. High quality and a low bit rate are the most important criteria in any video compression standard. H.264/AVC is considered to have the best performance in these two optimization dimensions compared to the other current standards. This is the latest video coding standard developed by the Joint Video Team (JVT), which is organized by the ISO Moving Pictures Experts Group (MPEG) and the ITU-T Video Coding Experts Group (VCEG) (Wiegand *et al.*, 2003). Compared to previous standards, H.264/AVC has many advanced characteristics such as 4×4 integer transformation, spatial intra prediction, quarter-pixel motion compensation, multiple reference frames, multiple block sizes for inter frame coding, and content-adaptive arithmetic coding (Richardson, 2004). With these added characteristics, H.264/AVC achieves a higher coding efficiency than other standards mainly through intra prediction and variable block size motion compensation.

To maximize the coding efficiency for intra prediction and inter prediction, H.264/AVC uses the rate distortion optimization (RDO) technique to select the best coding mode. To choose the best coding mode for a macroblock (MB), H.264/AVC calculates the rate distortion (RD) cost of every possible mode

---

‡ Corresponding author

* Project (No. IITA-2009-(C1090-0902-0011)) supported by the Ministry of Knowledge Economy of Korea under the ITRC Support Program supervised by the IITA

Expanded based on "Fast Mode Decision for Intra Prediction in H.264/AVC Encoder" by La *et al.*, which appeared in the Proceedings of the 2007 IEEE International Conference on Image Processing

and selects the one that has the minimum cost, as shown in Fig.1. The RD cost calculation is carried out through inter and intra predictions for an MB. The RD cost computation for the intra mode is conducted according to the block sizes and the edge modes for chroma and luma. The RD cost computation for motion estimation is also based on each block size and each reference frame. The computational complexity is increased by using the RDO calculation in all possible modes and reference frames of inter and intra predictions. This computational complexity is the critical issue for real-time encoding.

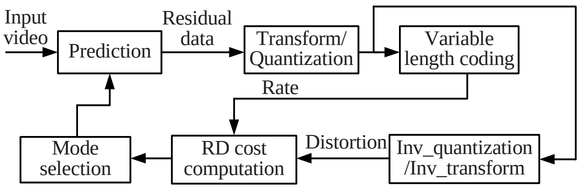

Fig.1 Rate distortion (RD) cost computation for intra prediction

To reduce this complexity, several fast algorithms for inter and intra predictions have been proposed. The approach for intra prediction is divided into two parts, one for I frames and the other for B/P frames. For B/P frames, intra prediction takes place after inter prediction in the reference software. The proportion of the resulting intra coding in the B/P frame is less than 5% of the total of encoding mode (Lee and Jeon, 2004; Xu and Lin, 2005; Gao and Lu, 2006). Therefore, several early skip algorithms for intra prediction have been proposed to determine the coding mode for the B/P frame.

The approach for the I frame is usually based on the minimization of the intra prediction candidate modes of intra prediction using preprocessing and the simplification of the intra prediction cost function. Kim *et al.*(2006) proposed an intra mode decision algorithm using directional masks and the modes of neighboring blocks to minimize the number of prediction modes. After obtaining the average and the sum of differences between the average and each pixel in a 4×4 luma block, four candidate modes are selected for RDO calculation by comparing the sum of differences with a threshold. Pan *et al.*(2005) proposed an intra prediction method using edge detection based on the Sobel operator. This algorithm calculates the edge direction histogram for luma 4×4, luma 16×16, and chroma 8×8 and accumulates the edge direction for each pixel in a block. Then the maximum accumulation direction becomes the best prediction mode. Wang J.F. *et al.*(2006) and Wang J.C. *et al.* (2007) proposed another algorithm using edge histogram descriptors. This algorithm detects the dominant edge strength (DES) using edge histogram descriptors, and a subset of the prediction modes is chosen for RDO calculation. These methods reduce the candidate modes using preprocessing as in (Zhang *et al.*, 2004; Liao *et al.*, 2005; Su *et al.*, 2006). The preprocessing for minimizing the prediction modes requires an additional computation, and the quality performance degrades due to the application of the minimized candidate modes.

Meng and Au (2003) proposed an intra mode decision algorithm that selects an optimal coding mode by computing the partial cost for down-sampled pixels, instead of for a 4×4 block, to simplify the cost function. Tseng *et al.*(2006) proposed an enhanced rate-distortion cost function, which combines the sum of absolute integer-transformed differences and a rate predictor for intra 4×4. These algorithms also add extra computation for the partial cost. In work related to the most probable mode (MPM), Storey and Nasiopoulos (2006) proposed an adaptive statistical method that increases the prediction accuracy for MPM to save bits. This algorithm generates the new MPM using the joint probability distribution of actual intra 4×4 prediction modes. Chen *et al.*(2006) proposed three conditions for skipping fewer probable candidate modes by exploiting the correlation of neighboring intra prediction modes. These three conditions come from the availability of the upper and left blocks, as well as from the half-full search and context correlation search. Rehman *et al.*(2007) proposed a fast mode decision algorithm that uses five candidate modes based on the availability of neighboring blocks. Most of these algorithms require extra computation to determine the factor required to save time or enhance performance. To meet the conditions for real-time encoding, the trade-off between the reduction of complexity and the quality is essential.

In this paper we propose a fast intra prediction mode selection algorithm based on the dominant edge direction (DED). This algorithm uses the summation and subtraction of pixel values in the horizontal and vertical directions to determine the DED, and the candidate modes for intra prediction are minimized using the DED. The number of search modes for RDO calculation is reduced from nine to three for luma 4×4 in this algorithm. For luma 16×16 and chroma 8×8, the search modes are reduced from four to two. The DED-based algorithm reduces the full intra search time to about 72% of that in the reference software. The peak signal-to-noise ratio (PSNR) degradation and the bit rate are acceptable. The edge direction calculation may be so simple that it can be implemented in VLSI hardware architecture.

The structure of the rest of this paper is as follows. In Section 2 we present an overview of intra prediction in H.264/AVC. In Section 3, we introduce our fast intra mode decision algorithm based on the DED. Performance evaluations of the algorithm are demonstrated in Section 4 and conclusions are presented in Section 5.

INTRA PREDICTION MODE DECISION IN H.264/AVC

Intra prediction modes

As mentioned previously, intra prediction uses the 4×4 and 16×16 prediction modes for luma, and the 8×8 prediction mode for chroma. The 4×4 prediction determines the prediction modes for each of the sixteen 4×4 luma blocks in an MB. The 4×4 luma block is selected in a relatively non-homogeneous area. Fig.2a shows a block with its neighboring pixels and 4×4 intra modes in nine directions. The *a*~*p* samples of a 4×4 block are predicted using reconstructed samples *A*~*M*. For example, the DC (non-edge) mode from sample *a* to sample *p* is predicted by $(A+B+C+D+I+J+K+L)/8$. The mode numbers 0, 1, 3~8 indicate the directions shown in Fig.2b. The number 2 indicates the DC prediction. For 16×16 luma prediction, the intra prediction uses the four edge directions because this block is selected in a relatively homogeneous area. The four edge direction modes in Fig.3 are numbered 0 for vertical, 1 for horizontal, 2 for DC, and 3 for the plane. The chroma intra prediction uses an 8×8 prediction mode, and the chroma U and V modes are predicted by one mode. The decision for chroma edge direction is very similar to the intra 16×16 luma prediction except that the order of the modes is different. The four edge direction modes are numbered 0 for DC, 1 for horizontal, 2 for vertical, and 3 for the plane.

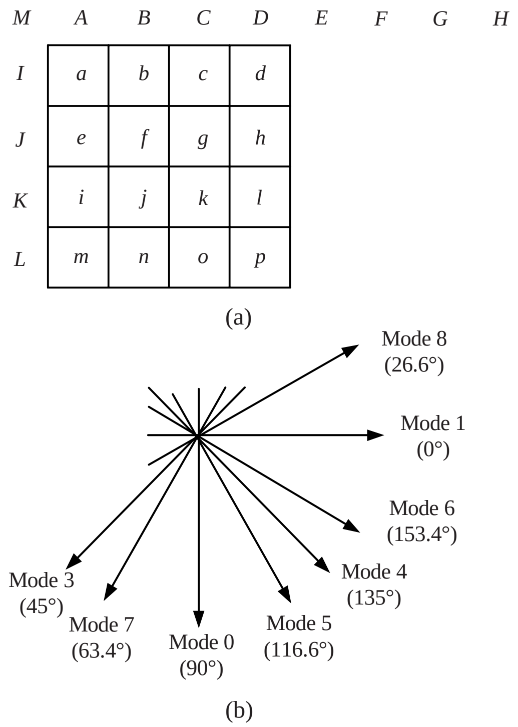

Fig.2 (a) A 4×4 block and its neighboring pixels; (b) Nine 4×4 intra-prediction modes including DC (non-edge, mode 2)

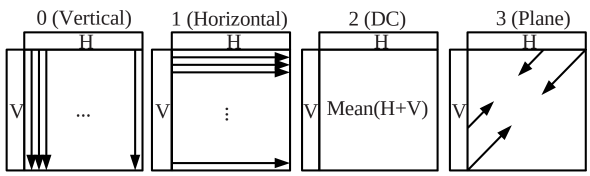

Fig.3 The 16×16 luma prediction modes

Choosing the intra prediction mode

H.264/AVC introduces the use of intra prediction modes for inter and intra frame coding. Intra prediction relies on spatial correlation with adjacent blocks, and the correlation between spatially adjacent blocks is used to encode the prediction mode for coding efficiency. The current block is predicted using boundary pixels of the upper and left blocks that were previously decoded. To determine the prediction mode for a block, the residual between the current block and the predicted block is transformed, quantized, entropy-coded, and reconstructed according to the prediction modes. The mode that has the minimum RD cost among the prediction modes is selected as the prediction mode.

The RD cost is a measure of searching for the best mode under the conditions of distortion and a constrained rate (Sullivan and Wiegand, 1998). H.264/AVC uses the RDO technique to achieve the best coding performance. To choose the best coding mode for the intra prediction, the 4×4 luma block, the 16×16 luma block, and the 8×8 chroma block are searched according to their prediction modes. When using the RDO method on the JM reference software, the best coding modes for luma and chroma are selected with the combined RD cost of the luma and chroma, whereas the best modes for luma and chroma are determined separately when using the non-RDO method. Therefore, the combined number of modes for luma and chroma components in the RDO method in an MB can be described by $n_{C8}\times(n_{L4}\times16+n_{L16})$, where n_{C8}, n_{L4}, and n_{L16} are the numbers of prediction modes for 8×8 chroma, 4×4 luma, and 16×16 luma, respectively. This means that the RDO calculation must be performed 4×(9×16+4)=592 times to select the best mode in an MB. This greatly increases the complexity and computational load of H.264/AVC. After determining the prediction mode for the coding mode, the encoder encodes the prediction mode and the residual value between the original pixel and the predicted pixels.

PROPOSED FAST MODE DECISION ALGORITHM

Intra 4×4 prediction

In H.264/AVC, the pixel values for prediction are generated from the boundary pixels of neighboring decoded blocks according to the prediction modes. In addition, the prediction mode that has the minimum RD cost of all prediction modes is chosen as the best mode after calculating the distortion and rate using encoding and decoding. The intra prediction is a process to express the current block using the reconstructed neighboring pixels. Intra prediction is based on the assumption that the neighboring pixel of the current block is correlated with the current block. This means that the edge direction of the current block is correlated with the edges of neighboring blocks. Therefore, by using the DED of the current block as a preprocessing step, the number of candidate modes for intra prediction can be minimized. To express the DED information for a 4×4 block, we define C_v and C_h as the vertical and horizontal directivity components respectively, where C_v and C_h are described in Eqs.(1) and (2), respectively. The allocated pixels according to C_v and C_h are shown in Fig.4.

a	*b*	*c*	*d*
e	*f*	*g*	*h*
i	*j*	*k*	*l*
m	*n*	*o*	*p*

$A=a+b+c+d$
$B=e+f+g+h$
$C=i+j+k+l$
$D=m+n+o+p$
$I=a+e+i+m$
$J=b+f+j+n$
$K=c+g+k+o$
$L=d+h+l+p$

Fig.4 Pixel summation in the 4×4 prediction

$$C_v=\sum_{i=0}^{3}\left(f(i,0)-f(i,3)+f(i,1)-f(i,2)\right), \quad (1)$$

$$C_h=\sum_{j=0}^{3}\left(f(0,j)-f(3,j)+f(1,j)-f(2,j)\right). \quad (2)$$

The C_v and C_h are further expressed as

$$C_v=I-L+J-K, \quad (3)$$

$$C_h=A-D+B-C, \quad (4)$$

$$C_{mul}=C_v\cdot C_h, \quad (5)$$

$$\begin{cases}\text{abs}(C_v)=|I-L+J-K|,\\ \text{abs}(C_h)=|A-D+B-C|.\end{cases} \quad (6)$$

The C_v and C_h only depend on the vertical and horizontal intensity differences, respectively. Therefore, by using the ratio between C_v and C_h, the DED is determined. If abs(C_v)>abs(C_h), abs(C_v) is divided by abs(C_h), and the DED will be the vertical direction (denoted as C_{divV}) as described in Eq.(7). Otherwise, the DED will be the horizontal direction (denoted as C_{divH}) as described in Eq.(8).

$$\text{If abs}(C_v)>\text{abs}(C_h),\ \text{then}\ C_{divV}=\frac{\text{abs}(C_v)}{\text{abs}(C_h)}; \quad (7)$$

$$\text{If abs}(C_h)>\text{abs}(C_v),\ \text{then}\ C_{divH}=\frac{\text{abs}(C_h)}{\text{abs}(C_v)}. \quad (8)$$

As described in Fig.2b, every prediction mode has its edge direction, excluding the DC mode. After deciding the DED using C_{divV} or C_{divH}, $C_v\cdot C_h$ is checked to know the DED position on the prediction modes. When the value is negative, one of DEDs 1, 4, 5, and 7 in Fig.5 will be decided as the DED. When the value is positive, one of DEDs 0, 2, 3, and 6 in Fig.5 will be decided as the DED. For example, if abs(C_v)>abs(C_h) and $C_v \cdot C_h$ is positive, C_{divV} including DEDs 0, 1, 2 and 4 will be the dominant edge direction, and one of DEDs 0 and 2 will be selected as the DED because the multiplied value is positive. When C_{divV}>4, the DED will be DED 0. The value 4 comes from the approximated value of tan 75.9°, which is almost the dividing value between prediction modes 7 and 0. The value 4 is also the dividing value of other prediction modes, such as prediction modes 0 and 5 as described in Fig.5. Therefore, one of the prediction modes is selected as the DED. Based on the DED, three prediction modes including the matched prediction mode, the one neighbouring mode on the same edge direction (horizontal or vertical) of the DED, and the DC mode (non-edge direction) are chosen for RD cost calculation to select the best mode. The 4×4 intra mode decision algorithm in Fig.6 can be determined from the relationships between the DEDs and 4×4 intra modes in Fig.5. Table 1 shows the candidate modes determined from the DED for RDO calculation.

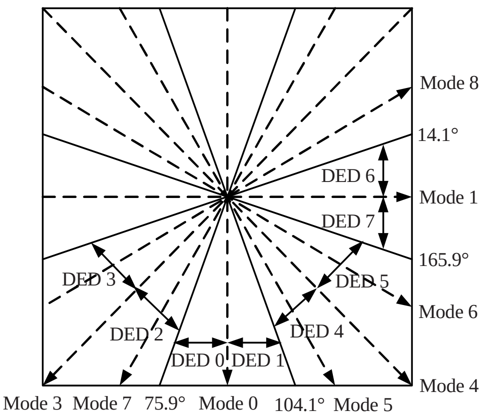

Fig.5 Relationships between the DEDs and 4×4 modes

Table 1 Candidate modes determined from the DED of the 4×4 block

DED	Candidate modes	DED	Candidate modes
0	0, 2, 7	4	2, 4, 5
1	0, 2, 5	5	2, 4, 6
2	2, 3, 7	6	1, 2, 8
3	2, 3, 8	7	1, 2, 6

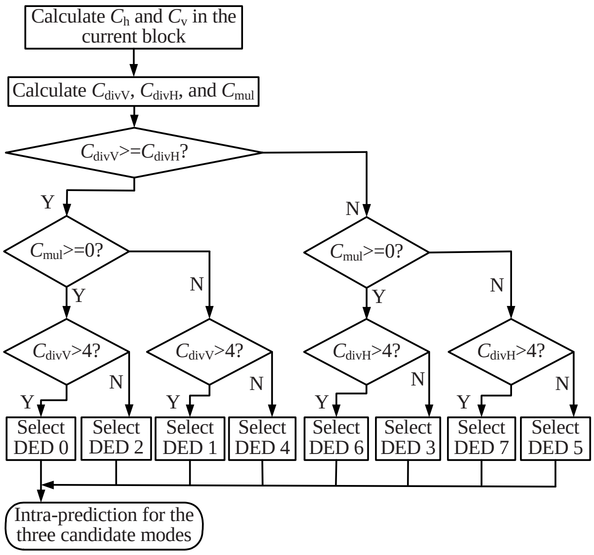

Fig.6 The proposed algorithm for the 4×4 block

Intra 16×16 prediction

For 16×16 luma, the decision method is similar to that for the 4×4 block, except for the number of prediction modes. Eqs.(1) and (2) become

$$\begin{aligned} C_v = \sum_{i=0}^{15} \big(& f(i,0) - f(i,15) + f(i,1) - f(i,14) \\ & + f(i,2) - f(i,13) + f(i,3) - f(i,12) \\ & + f(i,4) - f(i,11) + f(i,5) - f(i,10) \\ & + f(i,6) - f(i,9) + f(i,7) - f(i,8) \big), \end{aligned} \quad (9)$$

$$\begin{aligned} C_h = \sum_{j=0}^{15} \big(& f(0,j) - f(15,j) + f(1,j) - f(14,j) \\ & + f(2,j) - f(13,j) + f(3,j) - f(12,j) \\ & + f(4,j) - f(11,j) + f(5,j) - f(10,j) \\ & + f(6,j) - f(9,j) + f(7,j) - f(8,j) \big). \end{aligned} \quad (10)$$

Eqs.(9) and (10) can be described as shown in Fig.7. The 16×16 block has 256 pixels. To reduce the computational complexity, the number of pixels used to determine the DED can be minimized, because this block is selected as the best mode in a relatively homogeneous area and has only four prediction modes that are smaller than 4×4 luma. By reducing the number of pixels from 256 to 16, Eqs.(9) and (10) can be approximated as

$$C_v = \sum_{i=0}^{3} \big(f(4i,0) - f(4i,12) + f(4i,4) - f(4i,8) \big), \quad (11)$$

$$C_h = \sum_{j=0}^{3} \big(f(0,4j) - f(12,4j) + f(4,4j) - f(8,4j) \big). \quad (12)$$

a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p
q	r	s	t	u	v	w	x	y	z	a0	b0	c0	d0	e0	f0
g0	h0	i0	j0	k0	l0	m0	n0	o0	p0	q0	r0	s0	t0	u0	v0
w0	x0	y0	z0	a1	b1	c1	d1	e1	f1	g1	h1	i1	j1	k1	l1
m1	n1	o1	p1	q1	r1	s1	t1	u1	v1	w1	x1	y1	z1	a2	b2
c2	d2	e2	f2	g2	h2	i2	j2	k2	l2	m2	n2	o2	p2	q2	r2
s2	t2	u2	v2	w2	x2	y2	z2	a3	b3	c3	d3	e3	f3	g3	h3
i3	j3	k3	l3	m3	n3	o3	p3	q3	r3	s3	t3	u3	v3	w3	x3
y3	z3	a4	b4	c4	d4	e4	f4	g4	h4	i4	j4	k4	l4	m4	n4
o4	p4	q4	r4	s4	t4	u4	v4	w4	x4	y4	z4	a5	b5	c5	d5
e5	f5	g5	h5	i5	j5	k5	l5	m5	n5	o5	p5	q5	r5	s5	t5
u5	v5	w5	x5	y5	z5	a6	b6	c6	d6	e6	f6	g6	h6	I6	j6
k6	l6	m6	n6	o6	p6	q6	r6	s6	t6	u6	v6	w6	x6	y6	z6
a7	b7	c7	d7	e7	f7	g7	h7	i7	j7	k7	l7	m7	n7	o7	p7
q7	r7	s7	t7	u7	v7	w7	x7	y7	z7	a8	b8	c8	d8	e8	f8
g8	h8	i8	j8	k8	l8	m8	n8	o8	p8	q8	r8	s8	t8	u8	v8

A=a+e+i+m *I=a+m1+y3+k6*
B=m1+q1+u1+y1 *J=e+q1+c4+o6*
C=y3+c4+g4+k4 *K=i+u1+g4+s6*
D=k6+o6+s6+w6 *L=m+y1+k4+w6*

Fig.7 Pixel summation in the 16×16 prediction

The pixels related to Eqs.(11) and (12) can be indexed with the gray mark on the 16×16 luma block as shown in Fig.7. The C_v and C_h are expressed as Eqs.(3)~(6). Using the ratio of C_v and C_h in Eqs.(7) and (8), the DED will be C_{divV} or C_{divH}.

After deciding the horizontal or vertical direction of the current block, $C_v \cdot C_h$ is checked to know the DED position on the prediction modes. When the value is negative, one of DEDs 0 and 1 in Fig.8 will be decided as the DED. When the value is positive, one of DEDs 0, 1, and 3 in Fig.8 will be decided as the DED. For example, if abs(C_v)>abs(C_h) and $C_v \cdot C_h$ is positive, C_{divV} including DEDs 0 and 3 will be the dominant edge direction, and one of DEDs 0 and 3 will be selected as the DED. When C_{divV}>1.997, the DED will be DED 0. The value 1.997 comes from the value of tan 63.4°, which is almost the dividing value between prediction modes 3 and 0. The value 1.997 is also the dividing value of prediction modes 1 and 3 as described in Fig.8. Therefore, one of the prediction modes is selected as the DED. Based on the DED, two prediction modes including the matched prediction mode and the DC mode are chosen for RD cost calculation to select the best mode. The 16×16 intra mode decision algorithm in Fig.9 can be determined from the relationships between the DEDs and 16×16 intra modes in Fig.8. Table 2 shows the candidate modes determined from the DEDs for fast intra prediction.

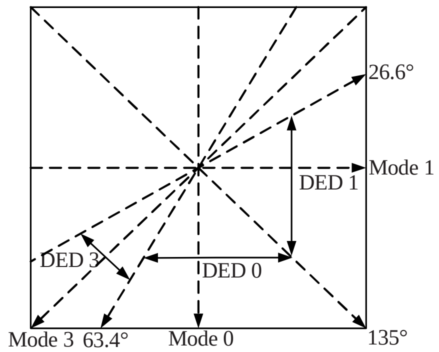

Fig.8 Relationships between the DEDs and 16×16 modes

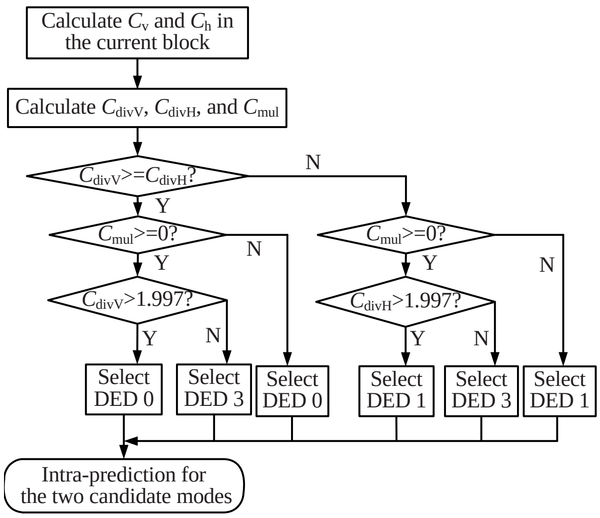

Fig.9 The proposed algorithm for the 16×16 luma block

Table 2 Candidate modes determined from the DED of the 16×16 block

DED	Candidate modes
0	0, 2
1	1, 2
3	2, 3

Intra 8×8 chroma prediction

For 8×8 chroma, the decision method is similar to that of the 4×4 block. Eqs.(1) and (2) become

$$C_v=\sum_{i=0}^{7}\big(f(i,0)-f(i,7)+f(i,1)-f(i,6)+f(i,2)-f(i,5)+f(i,3)-f(i,4)\big), \quad (13)$$

$$C_{\rm h}=\sum_{j=0}^{7}\big(f(0,j)-f(7,j)+f(1,j)-f(6,j)+f(2,j)-f(5,j)+f(3,j)-f(4,j)\big). \quad (14)$$

Eqs.(13) and (14) can be described as shown in Fig.10. The 8×8 block has 64 pixels. To reduce the computational complexity, the number of pixels used to determine the DEDs can be minimized, because this block is selected as the best mode in a relatively homogeneous area. By reducing the number of pixels from 64 to 16, Eqs.(13) and (14) can be approximated as

$$C_{\rm v}=\sum_{i=0}^{3}\big(f(2i,0)-f(2i,6)+f(2i,2)-f(2i,4)\big), \quad (15)$$

$$C_{\rm h}=\sum_{j=0}^{3}\big(f(0,2j)-f(6,2j)+f(2,2j)-f(4,2j)\big). \quad (16)$$

a	*b*	*c*	*d*	*e*	*f*	*g*	*h*
i	*j*	*k*	*l*	*m*	*n*	*o*	*p*
q	*r*	*s*	*t*	*u*	*v*	*w*	*x*
y	*z*	*a0*	*b0*	*c0*	*d0*	*e0*	*f0*
g0	*h0*	*i0*	*j0*	*k0*	*l0*	*m0*	*n0*
o0	*p0*	*q0*	*r0*	*s0*	*t0*	*u0*	*v0*
w0	*x0*	*y0*	*z0*	*a1*	*b1*	*c1*	*d1*
e1	*f1*	*g1*	*h1*	*i1*	*j1*	*k1*	*l1*

A=a+c+e+g *I=a+q+g0+w0*
B=q+s+u+w *J=c+s+i0+y0*
C=g0+i0+k0+m0 *K=e+u+k0+a1*
D=w0+y0+a1+c1 *L=g+w+m0+c1*

Fig.10 Pixel summation in the 8×8 prediction

The pixels related to Eqs.(15) and (16) can be indexed with the gray mark on the 8×8 chroma block as shown in Fig.10. The $C_{\rm v}$ and $C_{\rm h}$ are expressed as Eqs.(3)~(6). Using the ratios of $C_{\rm v}$ and $C_{\rm h}$ as given by Eqs.(7) and (8), the DEDs can be described as $C_{\rm divV}$ and $C_{\rm divH}$.

After deciding the DED using $C_{\rm divV}$ or $C_{\rm divH}$, $C_{\rm v}\cdot C_{\rm h}$ is checked to know the DED position on the prediction modes. When the value is negative, one of DEDs 1 and 2 in Fig.11 will be decided as the DED. When the value is positive, one of DEDs 1, 2, and 3 in Fig.11 will be decided as the DED. For example, if abs($C_{\rm v}$)>abs($C_{\rm h}$), and $C_{\rm v}\cdot C_{\rm h}$ is positive, $C_{\rm divV}$ including DEDs 2 and 3 will be the dominant edge direction, and one of DEDs 2 and 3 will be selected as the DED. When $C_{\rm divV}$>1.997, the DED will be DED 2. The value 1.997 comes from the value of tan63.4°, which is almost the dividing value between prediction modes 2 and 3. The value 1.997 is also the dividing value of prediction modes 1 and 3 as described in Fig.11. Therefore, one of the prediction modes is selected as the DED. Based on the DED, two prediction modes including the matched prediction mode and the DC mode are chosen for RD cost calculation to select the best mode. The 8×8 intra mode decision algorithm in Fig.12 can be determined from the relationships between the DEDs and 8×8 intra modes in Fig.11. The candidate modes for RDO calculation based on the DED are listed in Table 3.

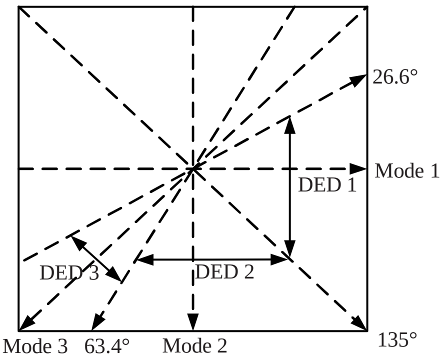

Fig.11 Relationships between the DEDs and 8×8 modes

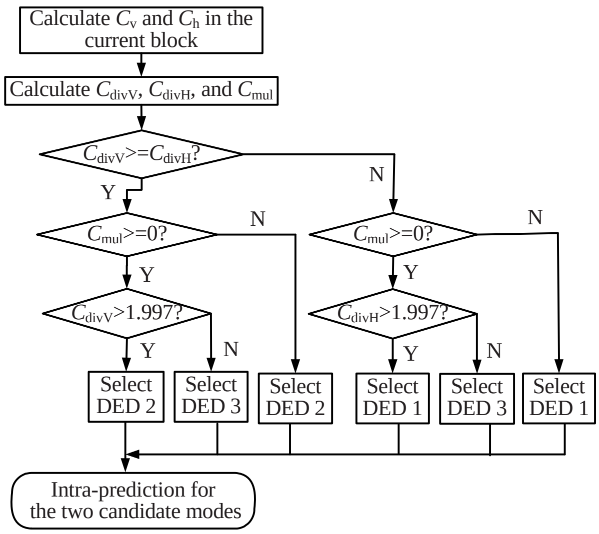

Fig.12 The proposed algorithm for the 8×8 chroma block

Table 3 Candidate modes determined from the DED of the 8×8 block

DED	Candidate modes
1	0, 1
2	0, 2
3	0, 3

Analysis of computational complexity

Several previous algorithms use the extracted dominant edge feature in a preprocessing step to minimize the prediction modes for fast mode decision. In this subsection, some of these algorithms are analyzed and compared on the basis of preprocessing and RDO calculations.

The algorithm of (Liao *et al.*, 2005) includes fast intra prediction mode selection based on edge mapping (EM) using the Sobel edge operator and an edge direction histogram in the 4×4 and 16×16 luma. Based on the edge direction histogram, a strong edge vector set is defined by comparison with the defined threshold. With the strong edge vector, small portions of the prediction modes are chosen as candidates for the best mode. Pan *et al.*(2005) proposed a fast intra prediction based on local edge information (LEI). For preprocessing, an edge map is created using the Sobel operator, and a local edge direction histogram is established for each sub-block. Based on the distribution of the edge direction, small portions of the prediction modes are searched to determine the best mode. Su *et al.*(2006) suggested a fast mode decision algorithm based on the coefficients of an integer transform (CIT) to reduce the complexity of the Sobel edge operator, as well as an adaptive threshold. According to the texture direction extracted from preprocessing, some of the prediction modes are tested to select the best mode. The algorithm of (Wang J.F. *et al.*, 2006; Wang J.C. *et al.*, 2007) includes a fast mode decision based on the DES. In the preprocessing stage, the DED is detected using five filters to reduce the computational load. Based on the DES, a subset of prediction modes is chosen for RDO calculation.

In preprocessing, in extracting the dominant feature for fast intra prediction, our algorithm is much simpler than the previously proposed algorithms, as described in Table 4, because it extracts the DED just by using the lesser computation rather than by using any complicated computation such as edge maps, transformations, or edge filters. Table 5 shows a simple comparison of the computational complexity of the RDO calculation with the other four algorithms by excluding the occurrence probabilities. Our proposed algorithm is also superior to the other algorithms with respect to the computational complexity for the RDO calculation.

Table 4 Number of operators in determining edge information for one MB

Algorithm	Number of operators			
	Addition	Subtraction	Multiplication	Division
EM (Liao *et al.*, 2005)	4480	0	0	384
LEI (Pan *et al.*, 2005)	4480	0	0	384
CIT (Su *et al.*, 2006)	1153	0	0	18
DES (Wang J.C. *et al.*, 2007)	722	0	344	72
DED (Proposed)	234	36	18	18

Table 5 Computational complexity of the RDO calculation

Algorithm	Number of search modes		
	8×8 chroma	4×4 luma	16×16 luma
JM 10.1	4	9	4
EM (Liao *et al.*, 2005)	4	3 or 5	1 or 4
LEI (Pan *et al.*, 2005)	2 or 3	4	2
CIT (Su *et al.*, 2006)	2	1 or 4	2
DES (Wang J.C. *et al.*, 2007)	2	4 or 9	2
DED (Proposed)	2	3	2

Algorithm	Number of RDO calculations	
	Minimum	Maximum
JM 10.1	4×(16×9+4)=592	
EM (Liao *et al.*, 2005)	98	376
LEI (Pan *et al.*, 2005)	132	198
CIT (Su *et al.*, 2006)	36	132
DES (Wang J.C. *et al.*, 2007)	132	292
DED (Proposed)	2×(16×3+2)=100	

PERFORMANCE EVALUATIONS

Evaluation conditions

The proposed algorithm was implemented on JM10.1 provided by the JVT. We compared the performance of the original H.264/AVC reference software with that of the previous DES-based algorithm and our proposed DED-based algorithm. We evaluated the two algorithms using common intermediate format (CIF) and quarter CIF (QCIF) sequences, and measured time savings, PSNR, and bit rate under the conditions of only intra frame (IIII) type and intra and predictive frame (IPPP) type sequences. We used the Bjontegaard delta PSNR and Bjontegaard delta bit rates (Bjontegaard, 2001) to compare the performances. The test conditions were as follows:

Test sequences (100 frames each):

QCIF: Car Phone, Container, Missa, News, Table Tennis.

CIF: Hall, Salesman, Suzie, Claire, Bream.

Intra period: 1 (IIII type), 0 (IPPP type).

RDO: on; rate control: off; frame rate: 30 frames/s.

QP: 20, 24, 28, 32, and 36; number of reference frames: 1.

Search range for motion estimation: ±16.

Hadamard transform: used; fast motion estimation: disabled.

$$\Delta PSNRY = PSNRY(proposed) - PSNRY(full\_search).$$

$$\Delta Bit = \frac{Bit\_rate(proposed) - Bit\_rate(full\_search)}{Bit\_rate(full\_search)} \times 100\%.$$

$$\Delta Time = \frac{Time(proposed) - Time(full\_search)}{Time(full\_search)} \times 100\%.$$

A negative value of $\Delta Time\_E$ and $\Delta Time\_I$ indicates the percent of time saved during the entire encoding and entire intra-prediction respectively compared to the full search by each. When the $\Delta PSNRY$ is positive, it means that the visual quality is improved.

Evaluation results

A comparison of the time saved, the PSNR, and the bit rate is shown in Tables 6 and 7 for the intra-only type. For the DED algorithm, for QCIF sequences, the entire encoding time consumption was reduced to about 67% and the PSNR decreased by −0.08 dB, whereas the bit rate increased by 5.6%; for CIF sequences, the time saving was about 72%, while the PSNR decreased by −0.04 dB and the bit rate increased by about 6.7%. Even though there is a slight bit rate increase that is worse than the DES-based algorithm on the Hall video sequence, the time saving and PSNR in the sequences such as Suzie and Claire outperform those of the DES-based algorithm while maintaining almost the same level of bit rate. Our method outperforms the DES-based algorithm in time savings for intra prediction while maintaining on average almost the same visual quality and bit rate.

Table 6 Evaluation results of the QCIF sequence (IIII type)

Sequence	$\Delta Time\_E$ (%)		$\Delta Time\_I$ (%)		$\Delta PSNRY$ (dB)		ΔBit (%)	
	DES	DED	DES	DED	DES	DED	DES	DED
Car Phone	−62.0	−67.8	−65.2	−71.8	−0.04	−0.07	3.6	6.2
Container	−60.5	−68.0	−63.6	−71.9	−0.04	−0.08	3.6	5.5
Missa	−57.6	−65.5	−60.7	−70.0	−0.04	−0.08	4.7	6.3
News	−59.7	−67.2	−62.0	−70.0	−0.05	−0.10	3.3	6.6
Table Tennis	−59.9	−67.1	−62.7	−71.2	−0.05	−0.07	2.1	3.5
Average	−59.9	−67.1	−62.8	−71.2	−0.04	−0.08	3.5	5.6

DES: proposed by Wang J.C. *et al.*(2007); DED: proposed by this paper

Table 7 Evaluation results of the CIF sequence (IIII type)

Sequence	$\Delta Time\_E$ (%)		$\Delta Time\_I$ (%)		$\Delta PSNRY$ (dB)		ΔBit (%)	
	DES	DED	DES	DED	DES	DED	DES	DED
Hall	−60.8	−67.8	−63.6	−71.6	−0.02	−0.06	5.3	8.4
Salesman	−61.9	−67.7	−64.8	−71.4	−0.04	−0.09	4.2	7.0
Suzie	−60.6	−67.0	−64.2	−73.3	−0.02	0.00	4.3	5.8
Claire	−56.4	−66.9	−59.5	−74.1	−0.03	−0.02	6.1	7.1
Bream	−52.1	−67.4	−54.7	−71.9	−0.04	−0.04	7.5	5.4
Average	−58.4	−67.4	−61.3	−72.5	−0.03	−0.04	5.5	6.7

DES: proposed by Wang J.C. *et al.*(2007); DED: proposed by this paper

The DES-based algorithm includes a preprocessing step to extract the DES using five types of edge filters, after which a subset of all intra prediction modes is processed according to the DES to select the best one. In the case of a flat video sequence such as Bream, most blocks can be chosen for searching all prediction modes by determining a non-directional edge condition defined in the DES-based algorithm. This will reduce the time savings by increasing the time for determining the DES with an additional computation while searching through all of the nine prediction modes. It is also clear that the non-directional edge condition is essential in maintaining the performance, even though it decreases the amount of time saved. Tables 8 and 9 show the simulated results for the IPPP-type video sequences. The performance and bit rate of the proposed algorithm are almost the same as those of the DED-based algorithm for the QCIF and CIF sequences.

Table 8 Evaluation results of the QCIF sequence (IPPP type)

Sequence	Δ*Time_E* (%)		Δ*Time_I* (%)		Δ*PSNRY* (dB)		Δ*Bit* (%)	
	DES	DED	DES	DED	DES	DED	DES	DED
Car Phone	−32.7	−35.1	−59.7	−64.3	0.00	0.00	0.1	0.4
Container	−33.7	−36.6	−58.0	−63.1	−0.02	−0.04	1.4	1.9
Missa	−29.4	−32.2	−55.6	−61.6	−0.04	−0.01	0.8	0.6
News	−33.1	−36.6	−58.7	−64.3	−0.03	−0.03	0.5	0.7
Table Tennis	−32.2	−34.8	−57.6	−62.7	−0.01	−0.04	0.3	0.0
Average	−32.2	−35.0	−57.9	−63.2	−0.02	−0.02	0.6	0.7

DES: proposed by Wang J.C. *et al.*(2007); DED: proposed by this paper

Table 9 Evaluation results of the CIF sequence (IPPP type)

Sequence	Δ*Time_E* (%)		Δ*Time_I* (%)		Δ*PSNRY* (dB)		Δ*Bit* (%)	
	DES	DED	DES	DED	DES	DED	DES	DED
Hall	−32.6	−35.8	−58.5	−63.5	0.00	−0.01	0.6	0.6
Salesman	−34.5	−37.2	−59.8	−63.6	−0.02	−0.03	0.8	1.4
Suzie	−31.1	−33.6	−57.8	−63.9	−0.03	−0.02	0.2	0.1
Claire	−29.5	−33.9	−54.5	−64.5	−0.03	0.01	0.1	0.7
Bream	−26.2	−31.3	−49.4	−62.5	0.01	−0.10	1.4	0.8
Average	−30.8	−34.4	−56.0	−63.6	−0.01	−0.03	0.6	0.7

DES: proposed by Wang J.C. *et al.*(2007); DED: proposed by this paper

CONCLUSION

We have proposed a fast mode decision algorithm for intra prediction in H.264/AVC. It minimizes the candidate modes based on the DED and uses pixel value summation and subtraction in the horizontal and vertical directions. The number of candidate modes for RDO calculation was reduced from 592 to 100 using DED. The evaluation results show that the proposed algorithm reduces the entire encoding time to 67% of that of the full search method in the intra-only video sequences of the reference software, with only a negligible loss of quality and increase in bit rate. By reducing the computational complexity, this algorithm can be used for real-time encoding or hardware design.

References

Bjontegaard, G., 2001. Calculation of Average PSNR Difference between RDCurve. ITU-T Q.6/16, Doc. VCEG-M33.

Chen, J.W., Chang, C.H., Lin, C.C., Ouyang, Y.H., Guo, J.I., Wang, J.S., 2006. A Condition-based Intra Prediction Algorithm for H.264/AVC. Proc. IEEE Int. Conf. on Multimedia and Expo, p.1077-1080. [doi:10.1109/ICME.2006.262721]

Gao, S., Lu, T.J., 2006. An Improved Fast Mode Decision Algorithm in H.264 for Video Communications. Int. Symp. on Systems and Control in Aerospace and Astronautics, p.57-60. [doi:10.1109/ISSCAA.2006.1627702]

Kim, J., Jeon, K., Jeong, J., 2006. H.264 Intra mode decision for reducing complexity using directional masks and neighboring modes. *LNCS*, **4319**:959-968.

Lee, J., Jeon, B., 2004. Fast Mode Decision for H.264. Proc. IEEE Int. Conf. on Multimedia and Expo, **2**:1131-1134.

[doi:10.1109/ICME.2004.1394416]

Liao, N., Quan, Z.Y., Men, A.D., 2005. Enhanced Fast Mode Decision Based on Edge Map and Motion Detail Analysis for H.264/JVT. IEEE Int. Workshop on VLSI Design and Video Technology, p.187-190. [doi:10.1109/IWVDVT.2005.1504582]

Meng, B., Au, O.C., 2003. Fast Intra-prediction Mode Selection for 4×4 Blocks in H.264. IEEE Int. Conf. on Acoustics, Speech, and Signal Processing, **3**:389-392.

Pan, F., Lin, X., Rahardja, S., Lim, K.P., Li, Z.G., Wu, D., Wu, S., 2005. Fast mode decision algorithm for intra prediction in H.264/AVC video coding. *IEEE Trans. Circuits Syst. Video Technol.*, **15**(7):813-822. [doi:10.1109/TCSVT.2005.848356]

Rehman, J., Zhang, Y., Kumarayapa, A., 2007. Fast Intra Prediction Mode Decision by Adaptively Selecting Fewer Number of Modes. Int. Conf. on Machine Learning and Cybernetics, p.2385-2389. [doi:10.1109/ICMLC.2007.4370544]

Richardson, I.E.G., 2004. H.264 and MPEG-4 Video Compression. The Robert Gordon University, Aberdeen.

Storey, D., Nasiopoulos, P., 2006. An Efficient Statistical Intra-mode Prediction Method for H.264. IEEE Int. Symp. on Signal Processing and Information Technology, p.803-806. [doi:10.1109/ISSPIT.2006.270908]

Su, R., Liu, G.Z., Zhang, T.Y., 2006. Fast Mode Decision Algorithm for Intra Prediction in H.264/AVC. Proc. IEEE Int. Conf. on Acoustics, Speech and Signal Processing, **2**:921-924. [doi:10.1109/ICASSP.2006.1660494]

Sullivan, G.J., Wiegand, T., 1998. Rate-distortion optimization for video compression. *IEEE Signal Processing Mag.*, **15**(6):74-90. [doi:10.1109/79.733497]

Tseng, C.H., Wang, H.M., Yang, J.F., 2006. Enhanced intra-4×4 mode decision for H.264/AVC coders. *IEEE Trans. Circuits Syst. Video Technol.*, **16**(8):1027-1032. [doi:10.1109/TCSVT.2006.878146]

Wang, J.C., Wang, J.F., Yang, J.F., Chen, J.T., 2007. A fast mode decision algorithm and its VLSI design for H.264/AVC intra-prediction. *IEEE Trans. Circuits Syst. Video Technol.*, **17**(10):1414-1422. [doi:10.1109/TCSVT.2007.903786]

Wang, J.F., Wang, J.C., Chen, J.T., Tsai, A.C., Paul, A., 2006. A Novel Fast Algorithm for Intra Mode Decision in H.264/AVG Encoders. IEEE Int. Symp. on Circuits and Systems, p.3498-3501.

Wiegand, T., Sullivan, G.J., Bjontegaard, G., Luthra, A., 2003. Overview of the H.264/AVC video coding standard. *IEEE Trans. Circuits Syst. Video Technol.*, **13**(7):560-576. [doi:10.1109/TCSVT.2003.815165]

Xu, L.D., Lin, X.G., 2005. Fast Mode Decision for Inter Frames in H.264/AVC. IEEE Int. Symp. on Communications and Information Technology, **1**:433-436. [doi:10.1109/ISCIT.2005.1566886]

Zhang, Y.D., Dai, F., Lin, S.X., 2004. Fast 4×4 Intra-prediction Mode Selection for H.264. IEEE Int. Conf. on Multimedia and Expo, **2**:1151-1154.