Signal Processing: *Image Communication* 23 (2008) 571–580

Contents lists available at ScienceDirect

Signal Processing: *Image Communication*

journal homepage: www.elsevier.com/locate/image

Fast sum of absolute transformed difference based 4 × 4 intra-mode decision of H.264/AVC video coding standard

Mohammed Golam Sarwer [a,b,*], Lai-Man Po [a], Q.M. Jonathan Wu [b]

[a] *Department of Electronic Engineering, City University of Hong Kong, 83 Tat Chee Avenue, Kowloon, Hong Kong*
[b] *Department of Electrical and Computer Engineering, University of Windsor, 401 Sunset Avenue, Windsor, ON, Canada N9B3P4*

ARTICLE INFO

Article history:
Received 10 February 2008
Received in revised form
1 May 2008
Accepted 5 May 2008

Keywords:
H.264/AVC
SATD
Video coding
Intra-mode
Rate-distortion optimization (RDO)

ABSTRACT

One of the new features in the H.264/AVC encoder is the use of Lagrangian rate-distortion optimization (RDO) method during mode decision at the macroblock level. The RDO technique has been employed in H.264/AVC for intra-prediction mode selection to achieve better coding efficiency. But the computational complexity of mode decision algorithm is extremely high. To reduce the complexity of mode decision, we propose an efficient and fast 4 × 4 intra-prediction mode selection scheme. The proposed method reduces the candidate of the prediction modes based on the correlation between neighboring blocks and the sum of absolute transformed difference (SATD) between the original block and the intra-predicted block. Firstly, the rank of each mode is obtained based on the SATD value. Then, the candidate modes are further reduced by using the combination of rank and most probable mode. The proposed method reduces the number of candidate mode to either one or two. Simulation results demonstrate that the proposed mode decision method reduces about 91% of mode decision time and 70% of total encoding time of intra-coding with ignorable degradation of coding performance.

© 2008 Elsevier B.V. All rights reserved.

1. Introduction

H.264/AVC is the newest international video coding standard of the Joint Video Team (JVT), formed by ISO/IEC MPEG and ITU-T VCEG [6]. H.264/AVC has achieved a significant improvement in coding efficiency compared to previous standards. Among many new features in H.264/AVC, inter- and intra-prediction modes are greatly enriched using variable block sizes and various directional prediction, respectively [4,9]. H.264/AVC uses rate-distortion optimization (RDO) technique to get the best coding result in terms of maximizing coding quality and minimizing bit rates. This means that the encoder has to code the video by exhaustively trying all the mode combinations, including the different intra- and inter-prediction modes. Therefore, the complexity and computation load of video coding in H.264/AVC increase drastically. To reduce the complexity, a number of researches have been made to explore fast algorithms in motion estimation [2,4,17]. Several fast-mode decision approaches proposed in [3,5,8,10,11,16] focused on how to eliminate unnecessary modes. To find the best interpolation direction as the best intra-mode, the information of the edge map of the whole frame is proposed in [11]. The pixels along the direction of local edges are similar values. In order to obtain the edge information in the neighborhood of the intra-block to be predicted, the edge map of the video picture is generated by using Sobel edge operators. Each pixel in the video picture will then be associated with an element in the edge map, which is the edge vector containing its edge direction and amplitude. A local edge direction histogram is then established for each block. Based on the distribution of the edge direction histogram, only small numbers of prediction modes are chosen for RDO calculation.

* Corresponding author at: Department of Electrical and Computer Engineering, University of Windsor, 401 Sunset Avenue, Windsor, ON, Canada N9B3P4. Tel.: +1 519 253 3000.
E-mail addresses: sarwer@uwindsor.ca (M. Golam Sarwer), eelmpo@cityu.edu.hk (L.-M. Po), jwu@uwindsor.ca (Q.M. Jonathan Wu).

0923-5965/$ - see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.image.2008.05.002

The histogram cell with the maximum amplitude, and its two adjacent cells, plus DC mode are chosen to take part in RDO calculation. But this algorithm evaluates every pixel in the whole frame and leads to high computation complexity. In [13], the candidate modes are reduced based on the sum of absolute transformed difference (SATD) of different modes. However, remaining modes should perform the RDO method. To reduce the computational complexity further, in this paper, we propose a fast-mode decision scheme for intra-4×4 prediction in H.264/AVC. Natural video sequences are highly spatially correlated. Hence the probability of mode of upper or left block to be the best mode of current block is high. Based on the combination of SATD and most probable mode, the proposed method reduces the number of candidate modes from nine to one or two.

The remainder of this paper is organized as follows. Section 2 provides the intra-mode search procedure for H.264/AVC. In Section 3, we describe the reduction of candidate mode based on SATD. Observations and the proposed algorithm are introduced in Section 4. The experimental results are presented in Section 5. Finally, Section 6 concludes the paper.

2. Overview of 4×4 intra-prediction

2.1. Intra-prediction

H.264 offers a rich set of prediction patterns for intra-prediction, i.e. nine prediction modes for 4×4 luma blocks and four prediction modes for 16×16 luma blocks. The latest H.264/AVC standard also defines 8×8 block and also has nine prediction modes, which are the same as those modes used in 4×4 block. In the case of 4×4 luminance block, the prediction block is defined using neighboring pixels of reconstructed blocks. The prediction block is calculated based on the samples labeled A–M as shown in Fig. 1. For example, mode 2 is called DC prediction in which all pixels (labeled *a*–*p*) are predicted by $(A+B+C+D+I+J+K+L)/8$. Mode 0 specifies the vertical prediction mode in which pixels (labeled *a*, *e*, *i*, and *m*) are predicted from *A*, the pixels (labeled *b*, *f*, *j*, and *n*) are predicted from *B*, and so on. The remaining modes are defined similarly according to the different directions as shown in Fig. 1.

2.2. Cost function for intra-4×4 mode decision of H.264/AVC

To take the full advantages of all modes, the H.264/AVC encoder can determine the mode that meets the best RD tradeoff using RDO mode decision scheme. The best mode is the one having minimum rate-distortion (RD) cost and this cost is expressed as

$$J_{RD} = SSD + \lambda R \tag{1}$$

where, SSD is the sum of squared difference between the original block and the reconstructed block. *R* is the true bits needed to encode the block and λ is an exponential function [14,15] of the quantization parameter (QP). In order to compute the RD cost for each mode, the same operation of forward and inverse transform/quantization and variable length coding is repetitively performed. All of these processing explains the high complexity of RD cost calculation. To reduce the complexity, H.264/AVC reference software [7] provides an SATD-based cost function:

$$J_{SATD} = SATD + \lambda_1 \times 4P \tag{2}$$

where SATD is the sum of absolute transformed difference between the original block and the predicted block. *P* is 0 for the most probable mode and 1 for other modes.

2.3. Most probable mode

The choice of intra-prediction mode for each 4×4 luma block must be signaled to the decoder and this could potentially require a large number of bits. For this reason, the best mode is not directly encoded into the compressed bit stream. Intra-modes for neighboring 4×4 blocks are highly correlated. For the current block, a mode is predicted based on the modes of upper and left blocks. This mode is defined as the most probable mode. The most probable mode is inferred according to the following rules [12]. If the left neighboring block or the up neighboring block is unavailable, the most probable mode is set to 2(DC). Otherwise, the most probable mode is set to the minimum of the prediction mode of left neighboring block and up neighboring block.

3. Reducing the number of candidate modes

In general, exhaustive cost calculation for all possible modes is required to find an RD optimal mode. However in

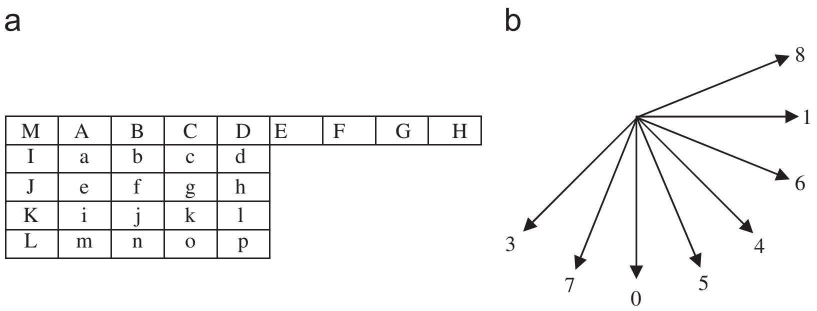

Fig. 1. Labelling and direction of intra-prediction modes.

some cases, we can ignore some modes that are almost not used because they have a very little effect on encoding efficiency. A good prediction should also produce a small value of SATD [13]. The basic idea [13] is to rank all the candidate modes using easy-to-compute SATD costs, and then evaluate Lagrange RD costs only for the few best modes decided by ranking. However, correlation between neighboring blocks is not considered in [13]. When the prediction modes of the top and left block are known, we have a probability list of the mode to be chosen for the present block. Table 1 shows the percentage of mode distribution with QP = 28. The mode that has the smallest SATD value has rank 1 and the one that has highest SATD value has rank 9. Rank of mode is obtained when the best mode is not the same as the most probable mode. From Table 1, we observed that about 96% (on average) modes are either the most probable mode, mode with rank 1, or mode with rank 2. Hence elimination of modes associated with rank 3–9 has a very little effect on encoding efficiency. However, the remaining modes should perform RDO.

4. Proposed intra-mode detection algorithm

Based on the most probable mode and rank, three different cases are described as follows.

Case 1: most probable mode is the mode associated with rank 1

If the most probable mode and the mode with rank 1 are the same, there is high chance for the mode associated with rank 1 to be the best intra-mode. To verify the above observation, extensive experiments have been conducted on different video sequences and at different QPs to find out the statistics of coding modes in test video sequences. One way is to examine the conditional probability of nine different modes with different ranks given that the most probable mode is the same as the mode with rank 1. This is called conditional probability because the condition "*most probable mode is the same as the mode with rank 1*" is given. In Fig. 2, *X*-axis represents the rank of mode and *Y*-axis shows their corresponding conditional probability. It is shown that conditional probability of the event "*best mode equal to mode with rank 1*" is about 98%. Therefore, it can be concluded that if the most probable mode is the

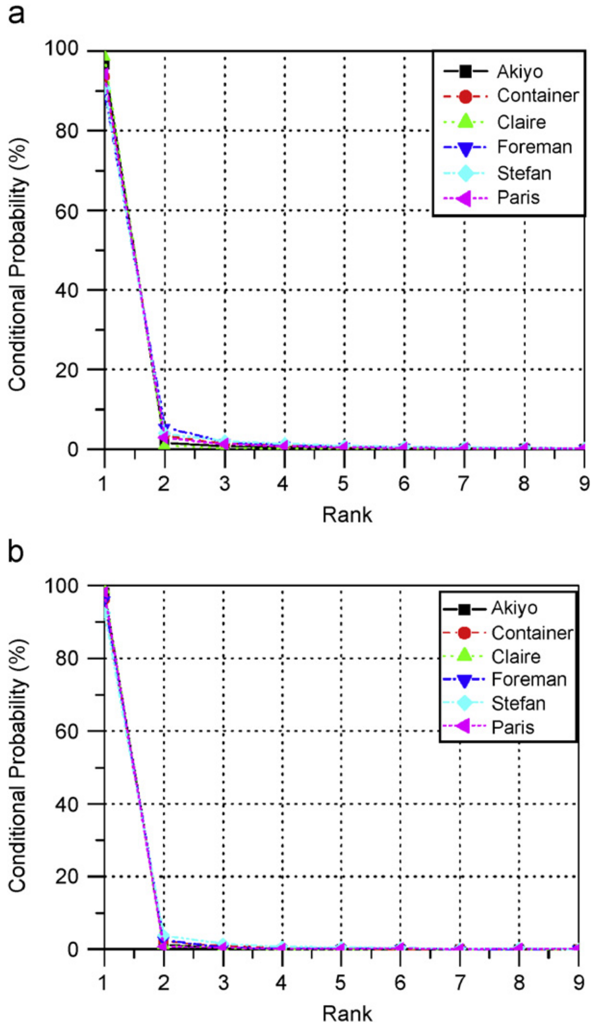

Fig. 2. Conditional probability of different video sequences based on case 1: (a) QP = 24 and (b) QP = 32.

Table 1
Percentage of mode distribution

Sequences	Most probable mode	Rank								
		1	2	3	4	5	6	7	8	9
Akiyo (QCIF)	73.03	20.14	4.09	1.51	0.60	0.29	0.16	0.10	0.05	0.02
Foreman (QCIF)	58.68	28.70	7.00	2.67	1.36	0.67	0.38	0.32	0.13	0.09
Stefan (QCIF)	56.21	24.91	8.31	4.27	2.51	1.65	1.04	0.60	0.36	0.15
Carphone (QCIF)	68.16	22.46	5.35	1.94	0.89	0.50	0.28	0.22	0.12	0.09
Claire (QCIF)	80.97	13.80	3.18	0.98	0.45	0.24	0.13	0.12	0.08	0.04
Container (QCIF)	79.80	13.04	3.65	1.53	0.78	0.50	0.31	0.19	0.12	0.09
Mobile (QCIF)	37.32	36.38	12.35	5.66	3.15	2.00	1.25	0.89	0.61	0.38
Salesman (QCIF)	53.33	31.98	7.67	3.19	1.55	0.99	0.56	0.41	0.22	0.10
Paris (CIF)	67.27	21.33	6.01	2.42	1.20	0.72	0.43	0.30	0.20	0.12
Tennis (CIF)	64.72	20.34	6.50	3.16	1.94	1.28	0.91	0.57	0.37	0.21
Tempete (CIF)	48.22	28.73	10.47	4.93	2.78	1.68	1.14	0.86	0.62	0.57

mode with rank 1 then best intra-mode is the mode associated with rank 1.

Case 2: most probable mode is the mode associated with rank 2

In this case most of the modes are those associated with either 1 or 2. Fig. 3 shows the conditional probability of nine different modes with different rank. Here the condition is the most probable mode is same as the mode with rank 2. From Fig. 3, it is shown that only a small number of modes are within rank 3–9. Hence we can eliminate the modes associated with rank 3–9. Only two modes are selected as candidate modes.

To reduce further computation, we have used thresholding technique based on relative SATD values between modes with rank 1 and 2. It is reasonable to say that if the difference between the SATD value of mode with rank 1 and that of most probable mode (rank 2) is relatively small, the probability of the most probable mode (rank 2) to be the RDO mode is higher. The relative SATD value between the most probable mode (rank 2) and the mode with rank 1 is calculated as follows:

$$\mathrm{RSATD}_{\mathrm{rank2}\rightarrow\mathrm{rank1}} = \frac{\mathrm{SATD}_{\mathrm{rank2}} - \mathrm{SATD}_{\mathrm{rank1}}}{\mathrm{SATD}_{\mathrm{rank2}}} \times 100\% \qquad (3)$$

where $\mathrm{SATD}_{\mathrm{rank2}}$ and $\mathrm{SATD}_{\mathrm{rank1}}$ are the SATD value of mode with rank 2 and mode with rank 1, respectively. In this case, the algorithm of the proposed mode decision scheme is given as follows:

if ($\mathrm{RSATD}_{\mathrm{rank2}\rightarrow\mathrm{rank1}} < T_1$)
  Best mode = Mode with rank2;
  else Best mode = Mode with rank1;

Now the factor is how to find out the threshold (T_1) value. In order to justify the proposed algorithm and to find out the threshold value, we have done some experiments with different video sequences at different QP values and

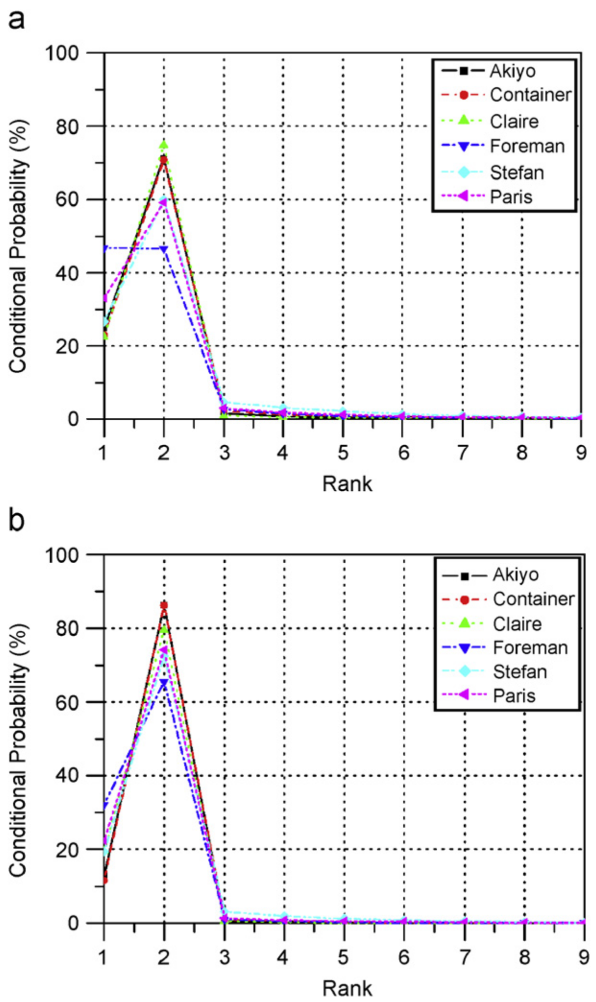

Fig. 3. Conditional probability of different video sequences based on case 2: (a) QP = 24 and (b) QP = 32.

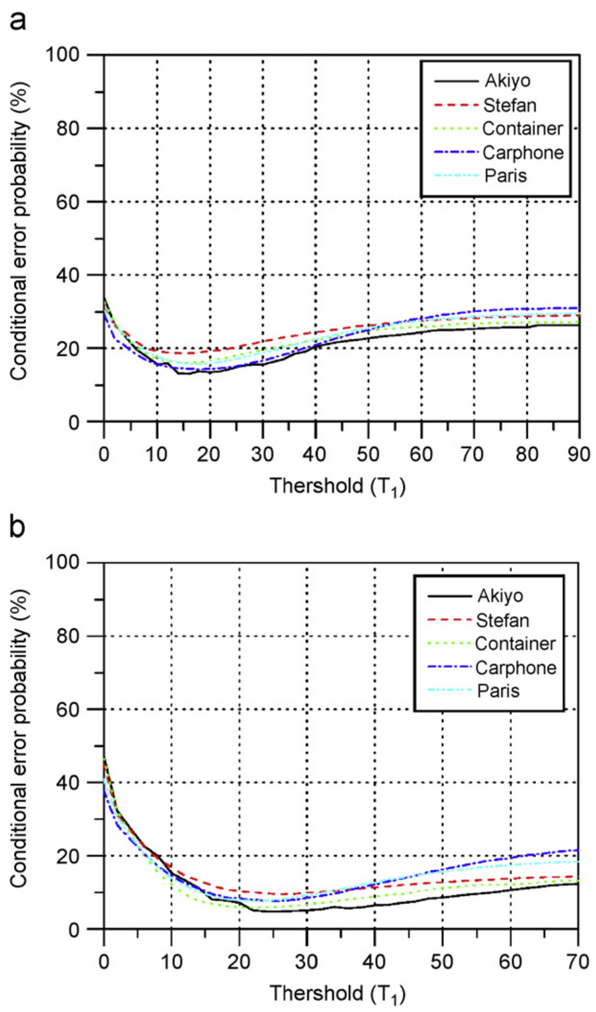

Fig. 4. Variation of conditional error probability with threshold value T_1 for case 2: (a) QP = 10 and (b) QP = 40.

observed the accuracy of mode decision with different values of threshold. By varying the threshold from 0 to 100, we have calculated the probability that a wrong mode is chosen for the current block. This probability is called as the conditional error probability. Here first condition is the most probable mode is same as the mode with rank 2 and another condition is the value of threshold is given. Fig. 4 shows variation of the probability of wrong mode decision with threshold value T_1 for difference video sequences at different QP factors. It is shown that for all of the sequences, minimum error is found with threshold value of around 15 and 20 at QP values of 10 and 40, respectively. From this observation, we found that threshold T_1 does not significantly vary with the type of video sequence and QPs. From similar experiments of different QP factors, it is shown that threshold values for minimum error are within the range 15–25 for all QP factors, and within this range, variation of conditional error probability is very low. Hence the threshold value T_1 of this algorithm is any value within the range 15–25.

Case 3: Most probable mode is the mode associated with rank 3–9

Fig. 5 shows the conditional probability of different modes with the condition "*Most probable mode is the mode associated with rank 3–9*". In this figure, the most probable mode is represented as rank 0. It is clear that most of the modes are within rank 0–2. Hence modes with rank 3–9 do not significantly influence the RD performance. In order to further reduce the complexity, we have used early termination algorithms. If SATD of rank 2 is much larger than SATD of rank 1, then probability of the event "*Mode with rank 1 is the best mode*" is higher. The relative SATD between mode with rank 1 and 2 is calculated in (3). It is reasonable to say that if SATD of the mode with rank 1, 2, and the most probable mode are more similar, the chance of the most probable mode to be the best mode is higher. To measure similarity, we have used the absolute deviation D of SATD values, which is defines as

$$D = \sum_{\text{rank}=0}^{2} |\text{SATD}_{\text{rank}} - \mu| \tag{4}$$

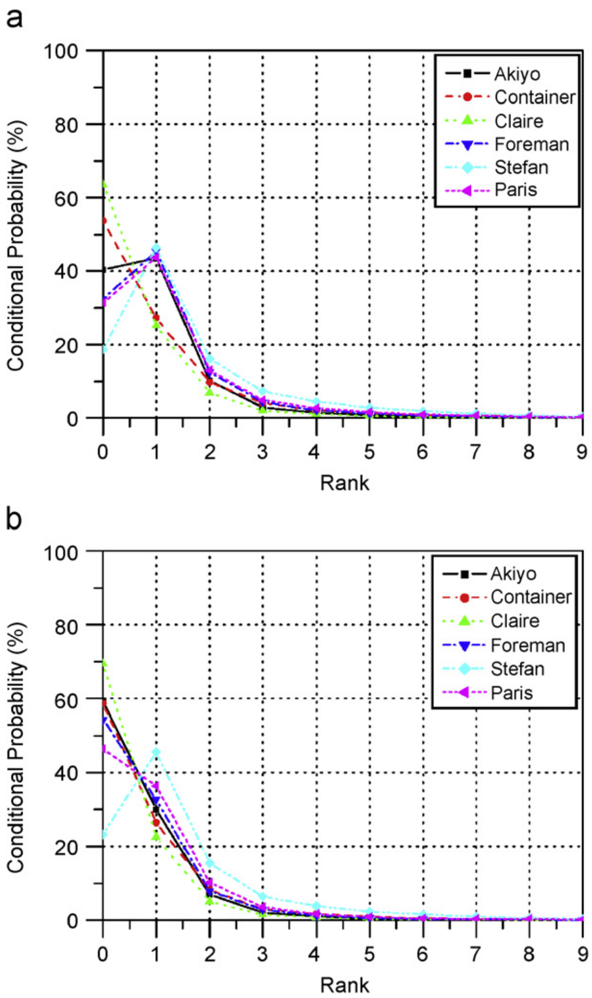

Fig. 5. Conditional probability of different sequences for case 3: (a) QP = 28 and (b) QP = 36.

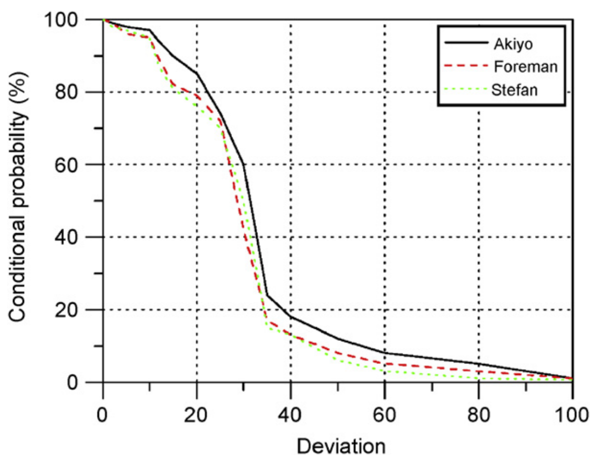

Fig. 6. Variation of conditional probability of the event "*Most probable mode is the best mode*" with deviation.

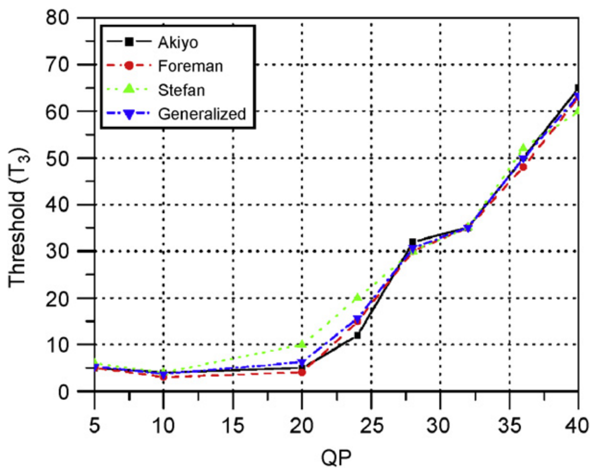

Fig. 7. Threshold T_3 of different quantization parameters.

where μ is the mean SATD of mode with rank 1, 2, and the most probable mode. Rank 0 represents the most probable mode.

To justify the above statement, variation of conditional probability of the event "*Most probable mode is the best mode*" with deviation is plotted in Fig. 6. From Fig. 6, it is shown that higher the value of deviation, the conditional probability is lower. The conditional probability is more than 80% within the deviation value 0–20. Algorithm of mode decision for case 3 is summarized as follows:

if ($RSATD_{rank2\rightarrow rank1} > T_2$)
  Best mode = Mode with rank1;
else if($D < T_3$) Best mode = Mode probable mode;
else Best mode = Mode which has minimum RD
        cost between mode with rank1 and 2;

In order to set the threshold T_2 and T_3, we have done several experiments for three different types of video

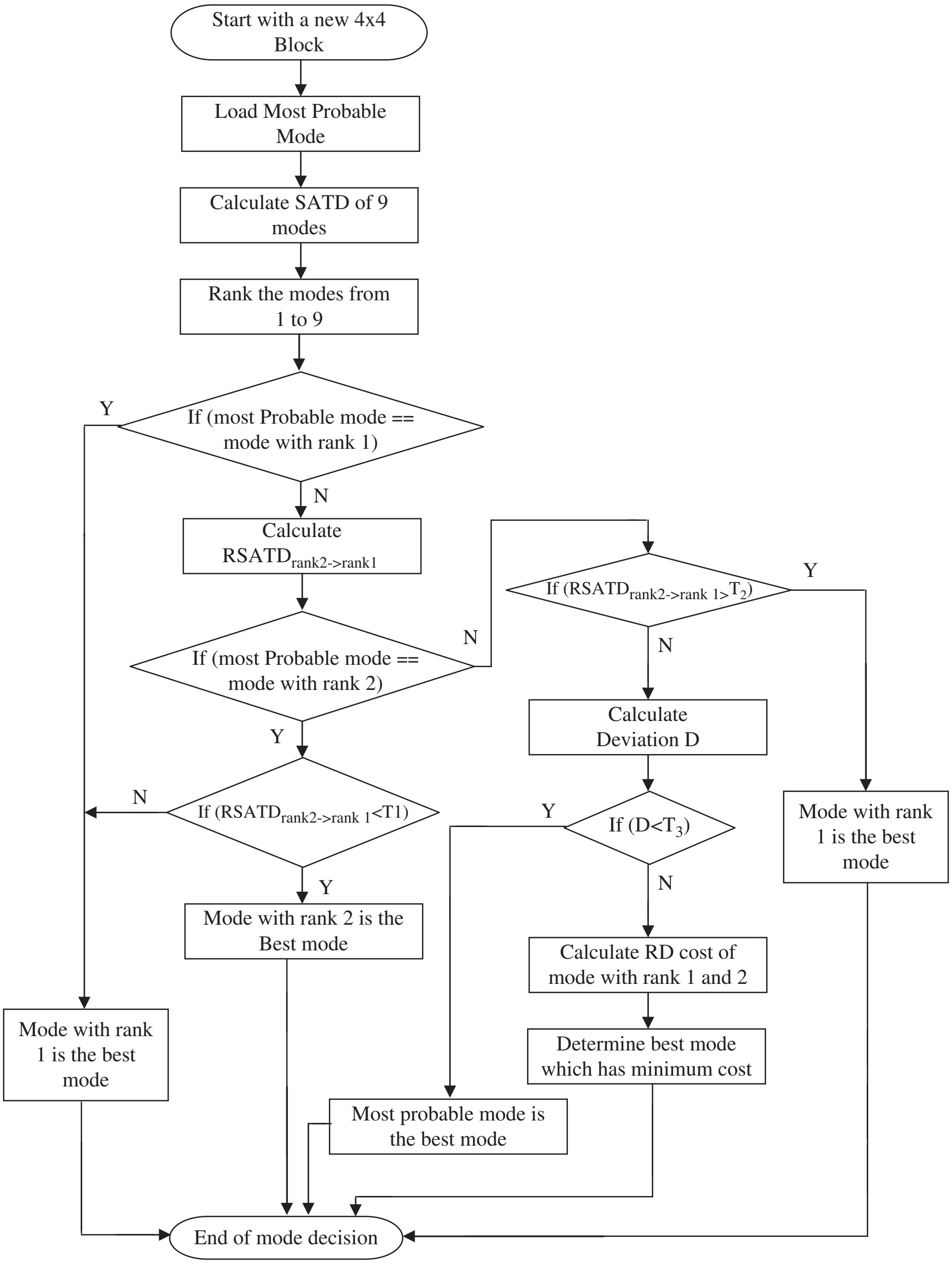

Fig. 8. Flow diagram of the proposed mode decision algorithm.

sequences (akiyo, foreman, and stefan) with QCIF format at different QP values. Akiyo, foreman, and stefan represent the sequence with low, medium, and high motion, respectively. By changing the threshold, we have observed the RD performance. We observed that T_2 is quietly independent on both of the type of video sequence and QP values and better RD performance was found at $T_2=35$. Threshold T_3 is also independent of the type of video sequence but depends on the QP values. Fig. 7 shows the variation of selected threshold T_3 with QP values. A generalized threshold curve is thus formed by averaging the threshold values of all three sequences for each QP. By using the polynomial fitting technique, the generalized threshold value T_3 is approximated as follows:

$$T_3 = 5.41 - 1.2\text{QP} + 0.06\text{QP}^2 \tag{5}$$

The flow diagram of overall mode decision is shown in Fig. 8. In this algorithm, RD calculation is required only in a part of case 3.

5. Simulation results

To verify the performance of proposed mode decision algorithm, different types of video sequences with QCIF and CIF format were used. The test parameters are: CABAC is enabled; number of frames is 100; Hadamard transform is used; frame rate is 30; threshold T_1 is 17; QPs are 28, 32, 36, and 40. JM 9.6 reference software [7] was used for evaluation of the encoding performance. The results are performed on a Pentium IV 2.8-GHz personal computer with 1-GB random access memory (RAM) and Microsoft Windows XP as the operating system. The comparison results are produced based on the average reduction of mode decision time ($\Delta T_m\%$), average reduction of total encoding time ($\Delta T_e\%$), the average PSNR differences (ΔPSNR), and the average bit rate differences ($\Delta R\%$). PSNR and bit rate differences are calculated according to the numerical averages between RD curves [1]. $\Delta T_e\%$ and $\Delta T_m\%$ are defined as

$$\Delta T_e\% = \frac{T_{e(org)} - T_{e(pro)}}{T_{e(org)}} \times 100\% \tag{6}$$

$$\Delta T_m\% = \frac{T_{m(org)} - T_{m(pro)}}{T_{m(org)}} \times 100\% \tag{7}$$

where $T_{m(org)}$ and $T_{m(pro)}$ denote the 4×4 intra-mode decision time of the JM 9.6 encoder and encoder with the proposed technique, respectively. $T_{e(org)}$ and $T_{e(pro)}$ are the total encoding time of original encoder and encoder with the proposed method, respectively.

5.1. Experiments with all intra-frame sequences

In this experiment, all the frames are intra-coded. The main profile is used for encoding. Only 4×4 intra-prediction method is used. Encoding was performed with two different cost functions ($J_{RD} = \text{SSD}+\lambda R$ and $J_{SATD} = \text{SATD}+4\lambda_1 P$) of different video sequences. The performance comparison between the original encoder, and the proposed method and J_{SATD} are tabulated in Table 2. In the case of J_{SATD} the average PSNR reduction is about 0.25 dB and average bit rate increment is about 7%, whereas in our proposed method, the average PSNR reduction is about 0.07 dB and average bit rate increment is about 2.24%. PSNR decrement and bit rate increment of *stefan* and *mobile* are a little bit larger than other video sequences. This is understandable because as shown in Table 1, percentages of mode within rank 3–9 of these sequences are a little bit larger as compared to other sequences. The proposed algorithm saves about 91% of

Table 3
Percentage of evaluating two modes

Sequence	Percentage (%)
Akiyo (QCIF)	13.92
Silent (QCIF)	21.56
Foreman (QCIF)	23.71
Carphone (QCIF)	17.87
Stefan (QCIF)	34.38
Mobile (QCIF)	40.60
Bus (CIF)	22.45
News (CIF)	13.02
Silent (CIF)	20.32

Table 2
Experimental results for all I frames

Sequence	J_{SATD}				Proposed			
	PSNR reduction ΔPSNR (dB)	Bit rate increment ΔR (%)	ΔT_e (%)	ΔT_m (%)	PSNR reduction ΔPSNR (dB)	Bit rate increment ΔR (%)	ΔT_e (%)	ΔT_m (%)
Akiyo (QCIF)	0.25	6.27	71.86	96.26	0.06	1.62	69.69	93.07
Silent (QCIF)	0.24	8.29	74.45	97.25	0.07	2.48	69.90	90.75
Foreman (QCIF)	0.21	6.36	74.54	96.11	0.06	1.93	72.09	92.30
Carphone (QCIF)	0.21	5.67	72.33	92.79	0.09	2.57	71.88	90.26
Stefan (QCIF)	0.35	8.22	75.56	96.54	0.11	2.74	70.28	89.59
Mobile (QCIF)	0.34	7.74	78.52	96.25	0.13	2.88	71.97	89.59
Bus (CIF)	0.25	7.03	72.42	96.44	0.08	2.42	71.11	91.95
News (CIF)	0.22	5.74	71.29	96.19	0.03	0.86	69.71	93.16
Silent (CIF)	0.18	7.36	70.12	95.15	0.08	3.06	69.51	92.58
Average	0.25	6.96	73.45	95.88	0.07	2.24	70.68	91.47

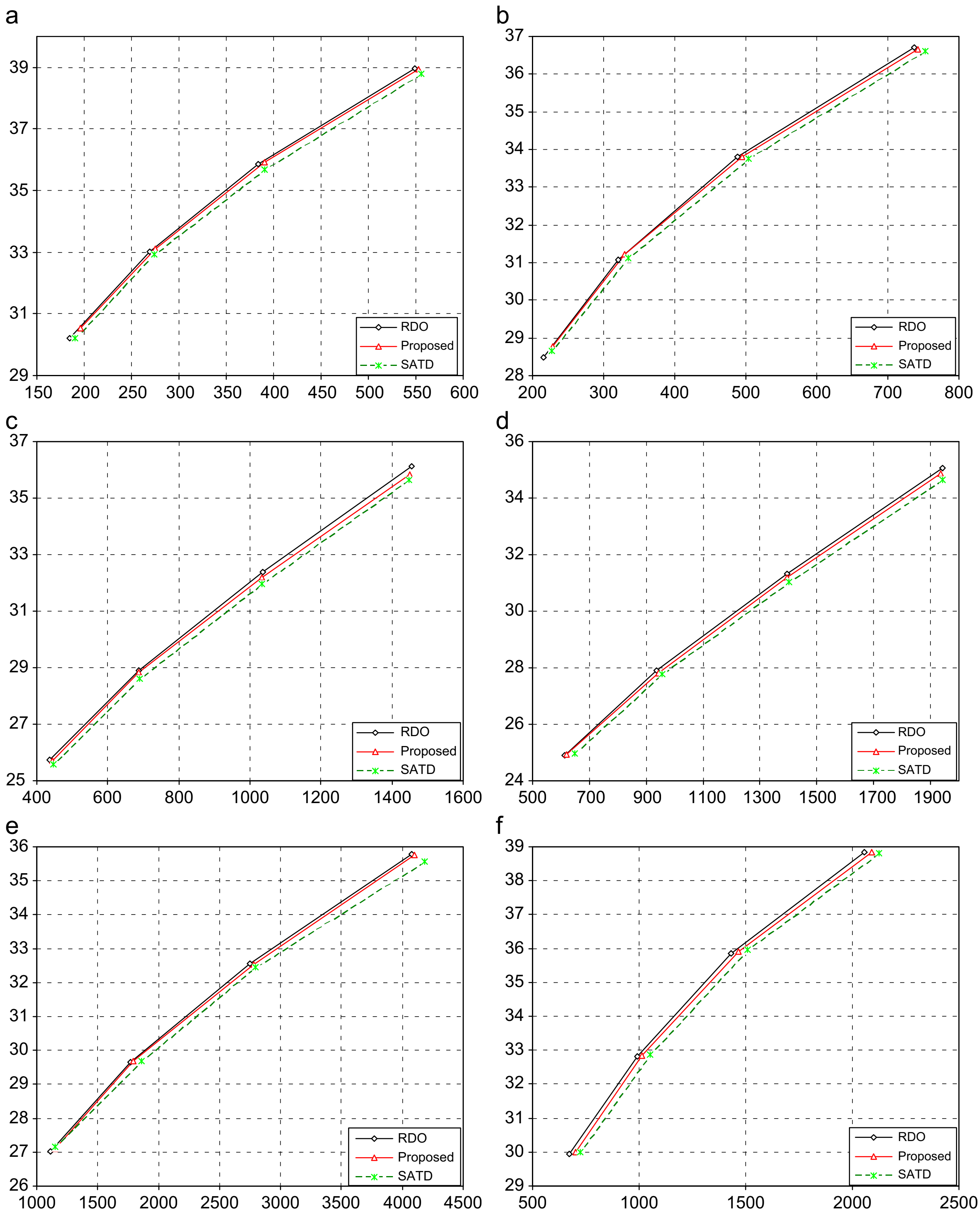

Fig. 9. RD performance of the proposed method when all frames are I frames (*X*-axis: bit rate in kbps; *Y*-axis: PSNR in dB) : (a) akiyo-QCIF (b) foreman-QCIF (c) stefan-QCIF (d) mobile-QCIF (e) bus-CIF and (f) news-CIF.

mode decision time and 70% of total encoding time as compared to J_{RD}. In the proposed method RD cost calculation is not required during case 1, case 2, and part of case3. Hence, in these conditions, computation of DCT, quantization, entropy coding, inverse quantization, and inverse DCT has been saved. RD cost calculation of two modes (mode with rank 1 and 2) is still necessary if relative SATD is smaller than a threshold T_2 and deviation is larger than the threshold T_3. But probability of evaluating two modes is small, especially for low and medium detail sequences, which are given in Table 3. Hence the complexity of proposed method is quite similar with J_{SATD} . As shown in Table 2, for complex sequences (*stefan*, *mobile*), the proposed algorithm reduces about 89% of mode decision time whereas J_{SATD} reduces about 96% of mode decision time. However, RD performance of the proposed method is much better as compared to J_{SATD}. The RD curves are given in Fig. 9. RD curve of the proposed method is much closer to RD optimized curve and better than J_{SATD} for all type of sequences.

5.2. Experiments with IPPPP sequences

In H.264/AVC, macroblocks in P frames also use intra-coding as the possible coding modes. Thus great time saving is expected by using fast intra-coding algorithm for these types of sequences. To evaluate performance of the proposed method during inter-frame coding, several test video sequences are used. In this experiment, the period of I-frame is 5 and rate control is enabled. CABAC is used as the entropy coding method in the case of main profile. Eight different types of sequences are encoded with lower and higher bit rate. Additionally, fast motion estimation method is also used. The experimental results are presented in Table 4. It can be seen that the proposed approach has reduced the total encoding time by about 40% and 45% during encoding of baseline and main profile, respectively. The PSNR reduction and bit rate increments are very negligible.

5.3. Comparison with other methods

In this experiment, the proposed method is compared with two different mode decision methods introduced in [11,13]. PSNR and bit rate differences are calculated according to numerical differences between RD curves [1]. Complexity reduction is calculated as

$$\text{Complexity reduction (\%)} = \frac{T_{ref} - T_{proposed}}{T_{ref}} \times 100\% \quad (8)$$

where T_{ref} is the mode decision time of the method presented in [11,13] and $T_{proposed}$ is the mode decision time of the proposed method. Positive values for PSNR and bit rate differences indicate increments, and negative values indicate decrements. The percentage of complexity reduction is obtained by taking average of complexity reduction of six different QPs (20, 24, 28, 32, 36, and 40). From the comparison results in Table 5, it is shown that our proposed method reduces the bit rate of about 0.66% and increases the PSNR of about 0.03 dB on average. We can see that there is a slight coding improvement relative to [11] for low and medium motion sequences. However, for high motion sequences such as stefan and mobile, RD performance of the proposed method is a little bit lower than that of fast-mode decision method in [11]. But this performance degradation does not affect the visual quality significantly. For all types of sequences, the proposed algorithm achieves a better complexity saving, which is about 70%. Experimental results of the proposed method compared to [13] are presented in Table 6. From Table 6, we can observe that as compared to [13], the proposed mode decision method reduces about 70% of mode

Table 4
Experimental results for IPPPP sequences

Sequence	Target rate (kbps)	Baseline profile			Main profile		
		PSNR reduction Δ PSNR (dB)	Bit rate increment ΔR (%)	ΔT_e (%)	PSNR reduction Δ PSNR (dB)	Bit rate increment ΔR (%)	ΔT_e (%)
Akiyo (QCIF)	30	0.01	0.19	40.09	0.01	0.26	46.05
	200	0.01	0.15	43.27	0.01	0.19	51.38
Foreman (QCIF)	30	0.03	0.61	37.44	0.02	0.03	44.94
	200	0.04	0.13	40.15	0.03	0.20	48.33
Stefan (QCIF)	30	0.02	0.26	37.17	0.02	0.06	42.92
	200	0.08	0.02	39.85	0.07	0.25	45.96
Silent (QCIF)	30	0.06	0.52	37.36	0.04	0.49	44.54
	200	0.04	0.14	40.98	0.05	0.77	48.22
Carphone (QCIF)	30	0.08	0.36	39.22	0.07	0.76	45.86
	200	0.09	0.33	42.10	0.06	0.15	48.99
Tempete (CIF)	100	0.04	0.16	38.82	0.01	0.21	44.98
	3000	0.03	0.26	43.07	0.03	0.01	49.85
Bus (CIF)	100	0.04	0.83	33.54	0.04	0.67	34.22
	3000	0.03	0.25	38.86	0.03	0.19	38.05
Flower (CIF)	100	0.02	0.16	35.13	0.04	0.43	34.44
	3000	0.06	0.28	39.03	0.04	0.61	39.08
Average		0.04	0.29	39.13	0.03	0.33	44.23

Table 5
Rate-distortion and complexity performance as compared to fast mode decision method [11]

Sequence (QCIF)	PSNR differences	Bit rate differences (%)	Complexity reduction (%)
Akiyo	0.04	−1.7	74.54
Silent	0.10	−2.3	77.32
Foreman	0.05	−0.4	70.72
Carphone	0.02	−2.4	73.64
Stefan	−0.04	1.5	65.24
Mobile	−0.01	1.3	61.65
Average	0.03	−0.66	70.50

Table 6
Rate-distortion and complexity performance as compared to fast mode decision method [13]

Sequence (QCIF)	PSNR differences	Bit rate differences (%)	Complexity reduction (%)
Akiyo	0	0.10	78.31
Silent	0	0.15	63.75
Foreman	−0.01	0.34	73.57
Carphone	0	0.17	75.03
Stefan	−0.02	0.76	64.63
Mobile	−0.02	0.40	59.16
Average	−0.008	0.32	69.08

decision time of intra-coding with ignorable degradation of coding performance.

6. Conclusions

In this paper, an efficient SATD-based intra-mode decision algorithm is proposed for H.264/AVC video coding standard. This algorithm is motivated by the fact that there is strong correlation between RDO cost and SATD. Also natural video sequences are highly spatially correlated. Hence the probability for mode of upper or left block to be the best mode of current block is high. Based on the combination of rank and most probable mode, the proposed method reduces the number of candidate modes from 9 to 1 or 2. The experimental results verified that the proposed technique is suitable for intra-mode decision of H.264/AVC. For RD performance very close to the standard, the proposed scheme affords significant time saving as compared to original H.264/AVC. The proposed technique reduces mode decision time by 91% (on average) and total encoding time by 70% during intra-frame coding of H.264/AVC. During inter-frame coding the proposed method saves about 45% of total encoding time without significant performance degradation. The proposed scheme is very relevant to low-complexity coding system.

Acknowledgment

The work described in this paper was substantially supported by a grant from Hong Kong SAR Government with CERG Project 9041251.

References

[1] G. Bjontegaard, Calculation of average PSNR differences between RD-curves, presented at the 13th VCEG-M33 Meeting, Austin, TX, April 2001.
[2] Z. Chen, P. Zhou, Y. He, Fast integer pel and fractional pel motion estimation for JVT, Joint Video Team (JVT) Docs, JVT-F017, December 2002.
[3] Y. Cheng, Z.Y. Wang, K. Dai, J.J. Guo, Analysis of inter frame coding without intra modes in H.264/AVC, in: Proceedings of the Seventh Eurographics Symposium on Multimedia, Nanjing, China, October 2004, pp. 77–86.
[4] H.-Y. Cheong, A.M. Tourapis, P. Topiwala, Fast motion estimation within JVT Codec, in: JVT-E023, Proceedings of the Fifth Meeting, Geneva, Switzerland, October 9–17, 2002.
[5] C. Grecos, M. Yang, Fast mode prediction for the baseline and main profiles in the H.264 video coding standard, IEEE Trans. Multimedia 8 (6) (December 2006) 1125–1134.
[6] International Standard, Information Technology—coding of audio-visual objects—part 10: advanced video coding, Final Draft International Standard, ISO/IEC FDIS 14496-10, December 2003.
[7] Joint Video Team, Joint Video Team (JVT) reference software version 9.6, 〈www.iphome.hhi.de/suehring/tml/download/old_jm/〉.
[8] C. Kim, H. Shih, C. Kuo, Feature-based intra prediction mode decision for H.264, in: Proceedings of ICIP 2004, October 2004, Singapore, pp. 769–772.
[9] X. Li, G. Wu, Fast integer pixel motion estimation, in: JVT-F011, Proceedings of the Sixth Meeting, Awaji Island, Japan, December 5–13, 2002.
[10] B. Meng, O. Au, C. Wang, H. Lam, Efficient intra-prediction mode selection for 4 × 4 blocks in H.264, in: Proceedings of ICME 2003, Baltimore, USA, pp. III521–III524.
[11] F. Pan, X. Lin, S. Rahardja, K.P. Lim, Z.G. Li, D. Wu, S. Wu, Fast mode decision algorithm for intra-prediction in H.264/AVC video coding, IEEE Trans. Circuits Syst. Video Tech. 15 (7) (July 2005) 813–822.
[12] I.E.G. Richardson, H.264 and MPEG-4 Video Compression—Video Coding for Next Generation Multimedia, Wiley, New York, 2003, pp. 180–181.
[13] Y. Su, J. Xin, A. Vetro, H. Sun, Efficient MPEG-2 to H.264/AVC intra transcoding in transform-domain, ISCAS 2 (2005) 1234–1237.
[14] G.J. Sullivan, T. Wiegand, Rate-distortion optimization for video compression, IEEE Signal Process. Mag. 15 (November 1998) 74–90.
[15] T. Weigand, H. Schwarz, A. Joch, F. Kossentini, G. Sullivan, Rate-constrained coder control and comparison of video coding standards, IEEE Trans. Circuits Syst. Video Tech. 13 (7) (July 2003) 688–703.
[16] C.-L. Yang, L.M. Po, W.-H. Lam, A fast H.264 intra prediction algorithm using macroblock properties, ICIP-2004, pp. 461–464.
[17] L. Yang, K. Yu, J. Li, S. Li, An effective variable block size early termination algorithm for H.264 video coding, IEEE Trans. Circuits Syst. Video Tech. 15 (6) (June 2005) 784–788.