An Early SKIP Mode Decision Method using PSNR Prediction in H.264/AVC

Sung-Jae Park[1], Chae-Bong Sohn[2], Ho-Chong Park[3] and Seoung-Jun Oh[4]
Department of Electronic Engineering, Kwnagwoon University
Wolgye 1-dong ,Nowon-gu, Seoul 139-701, Korea
Tel: +82-2-940-5102, Fax: +82-2-943-5381
E-mail : [1]sjpark@media.kw.ac.kr, {[2]cbsohn,[3]hcpark, [4]sjoh}@kw.ac.kr

Abstract: H.264/AVC shows high coding efficiency more than previous video coding standard by using new coding tools. However, the encoder complexity greatly increases due to these coding tools. In this paper, we propose early SKIP mode decision method to reduce the computational complexity. Simulation results show that the proposed method could reduce encoding time of the overall sequences by 33% on average than JM 10.2 without noticeable degradation of coding efficiency.

1. Introduction

The video coding standard H.264/AVC was developed by Joint Video Team (JVT) of ISO/IEC MPEG and ITU-T VCEQ [1]. H.264/AVC shows high coding efficiency more than previous video coding standard by using new coding tools. Compared with the previous coding standards, H.264/AVC has some different features such as a 4x4 integer transform, multiple reference frames, various block-based motion estimation (ME) and rate-distortion optimization (RDO) for deciding an optimal macroblock mode. Specially, variable block-based motion estimation and RDO are very important coding tools in H.264/AVC. These coding tools have high coding efficiency. These coding tools, however, greatly increase the encoder complexity. In this paper, we present an early SKIP mode decision method using PSRN prediction for reducing the computational complexity.

This paper is organized as follows. Section 2 presents an overview of mode decision. Section 3 describes the proposed method and section 4 shows the performances of proposed method. Section 5 is the conclusion of the paper.

2. Overview of mode decision

2. 1 Variable block size motion estimation

ME is an important coding tool in H.264/AVC It finds the best matching block from previous block to reduce the temporal redundancy. For ME, H.264/AVC supports seven inter modes of different block sizes, including 16x16, 16x8, 8x16, 8x8, 8x4, 4x8 and 4x4. A macroblock was divided into three different block sizes, 16x16, 16x8 or 8x16. P8x8 was divided four sub-block sizes such as 8x8, 8x4, 4x8 or 4x4. The macroblock partitions as shown in Figure 1.

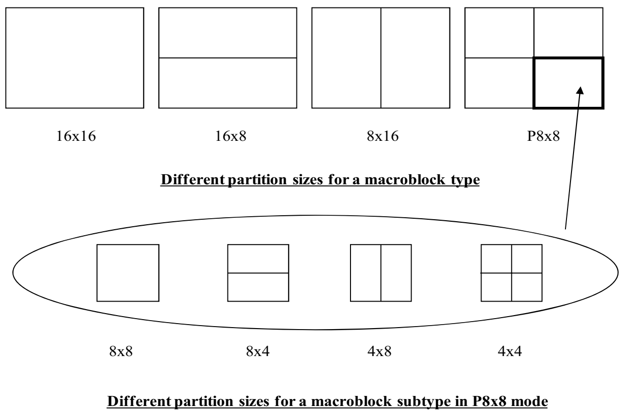


Figure 1. Variable block sizes in H.264/AVC

2. 2 Rate-distortion optimization

The H.264/AVC encoder performs ME of the variable block sizes. Then, the encoder decides an optimal macroblock mode with the least rate-distortion cost (RDcost) using lagrangian multiplier.
The procedure can be defined as follow:

$$J(s,c,MODE\,|\,QP,\lambda_{MODE}) = SSD(s,c,MODE\,|\,QP)+\lambda_{MODE}\cdot R(s,c,MODE\,|\,QP) \quad (1)$$

where s and c are the original video signals and the reconstructed video signals, respectively. QP is the quantization parameter. $MODE$ indicates an intermode in H.264/AVC and λ_{MODE} is lagrangian multiplier. R is the number of bits with QP and chosen $MODE$. SSD is the sum of squared differences between s and c. It is given by the following equation.

$$\begin{aligned} &SSD(s,c,MODE\,|\,QP) \\ &=\sum_{x=1}^{16}\sum_{y=1}^{16}(s_Y[x,y]-c_Y[x,y,MODE\,|\,QP])^2 \\ &+\sum_{x=1}^{8}\sum_{y=1}^{8}(s_U[x,y]-c_U[x,y,MODE\,|\,QP])^2 \\ &+\sum_{x=1}^{8}\sum_{y=1}^{8}(s_V[x,y]-c_V[x,y,MODE\,|\,QP])^2 \end{aligned} \quad (2)$$

where c_Y and s_Y represent the reconstructed and original luminance component and c_U, c_V, s_U and s_V the corresponding chrominance components

3. An early SKIP mode decision method

Table 1 indicates the probability of SKIP mode which was decided with the best macroblock mode. In this result, we observe that about 55% of the macroblocks encoded with SKIP mode for sequences.

Table 1. Probability of SKIP macroblock in P slices

Sequence	Probability	Sequence	Probability
Akiyo	86%	Foreman	16%
Carphone	26%	News	63%
Claire	73%	Paris	47%
Container	70%	Silent	55%
Average			**55%**

If we can decide in advance whether the mode of a macroblock is skip or not, the wasteful process of computing the RDcost can be omitted so that a huge amount of computation can be saved. A previous work for an early SKIP mode decision method (P-ESDM) was researched by Jeon [2]. According to this work, a block can be predicted as having SKIP mode in a P slice when the following set of four conditions is satisfied:

1) The best motion compensated block is MB16x16.
2) The reference frame is the one closest in the reference frame memory to the current frame.
3) The motion vector is the same as its predicted motion vector (PMV).
4) Its transform coefficients are all quantized to zero.

P-ESDM always conducts ME with 16x16 block size (MB16x16) and calculates the RDcost of MB16x16. However, if the number of reference frames was increased, it leads to a complexity increase in conducting MB16x16, as shown in Figure 2. Therefore, we need an efficient SKIP mode decision method without conducting MB16x16.

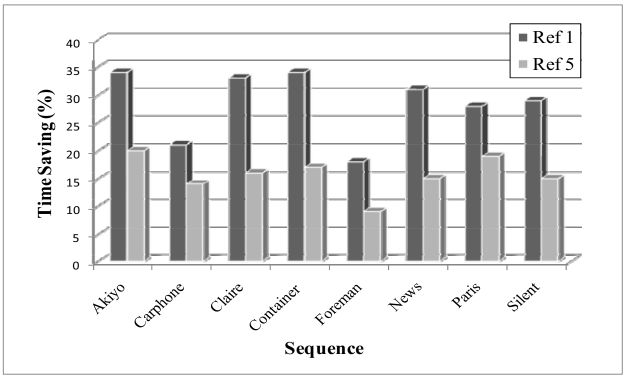


Figure 2. Time saving ratio for P-ESDM when using one reference frame and five reference frame.

When SKIP mode is selected, the encoder sends just a skip indicator without transmitting any data. The RDcost of SKIP mode calculates only distortion (*SSD*) without bitrates. If we predict the appropriate coding distortion according to *QP*, we can decide SKIP mode earlier.

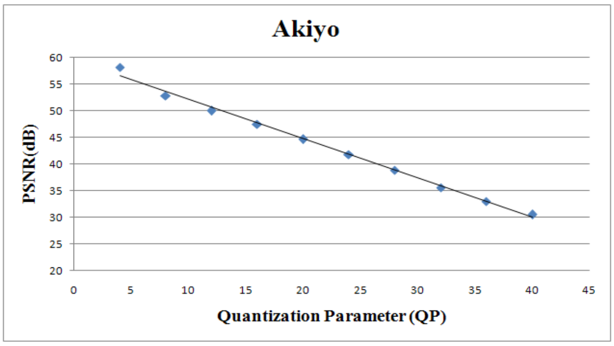


Figure 3. Average PSNR for akiyo sequence (IPPP 150 frames, 15f/s, *QP* range 4 to 48 with stepsize of 4)

Figure 3 shows the linear relation between the PSNR and the *QP*. It can be presented as Eq.(3).

$$\text{PSNR} = 10\log\left(\frac{255^2}{\text{MSE}}\right) = l \times QP + b \qquad (3)$$

where l and b are constants. Supposing the MSE is used as the distortion measure for a frame distortion, we can derive the relation between the distortion D and QP as Eq.(4).

$$D = \text{MSE} = \frac{255^2}{10^{((l\times QP+b)/10)}} \qquad (4)$$

Since statistics have shown that, when *QP* increases by 1, and PSNR usually loses 0.6~0.75 dB, we assume l=-0.68. However, b is related to image characteristic. In order to get the value of b, we use I frame's PSNR.

SKIP mode is related to the macroblock of the previous reference frame. Therefore, we must also consider macroblock's correlation denoted by r in Eq.(5).

$$m_x = \frac{1}{N\times M}\sum_{i=0}^{N-1}\sum_{j=0}^{M-1} x_{2i,2j} \qquad m_y = \frac{1}{N\times M}\sum_{i=0}^{N-1}\sum_{j=0}^{M-1} y_{2i,2j}$$
$$r = \frac{\sum_{i=0}^{N-1}\sum_{j=0}^{M-1}[(x_{2i,2j}-m_x)\times(y_{2i,2j}-m_y)]}{\sqrt{\sum_{i=0}^{N-1}\sum_{j=0}^{M-1}(x_{2i,2j}-m_x)^2}\times\sqrt{\sum_{i=0}^{N-1}\sum_{j=0}^{M-1}(y_{2i,2j}-m_y)^2}} \qquad (5)$$

Calculating all macroblock correlation leads to complexity increase. For reducing the complexity, we use the reduced size of 8x8 instead of 16x16 to compute r. In Eq.(5), x denotes a current macroblock, y is the macroblock indicated by PMV of x in a reference frame, and a pair of $(2i,2j)$ represent the position on each macroblock, where N=M=8. A value of predicted encoding distortion (*PED*) is define by Eq.(6) in this paper.

$$PED = D \times 256 \times r \quad (6)$$

The proposed method can be described as follows:

Step 1: Conduct SKIP mode and calculate RDcost of SKIP mode. Then, calculate the residual transform coding to check the coded block pattern (CBP) information.

Step 2: If CBP is zero, calculate Eq.(5) and go to step 3. Otherwise, go to *step 4*.

Step 3: If SKIP mode's luma distortion is smaller than *PED,* decide that the best macroblock mode is SKIP mode and go to *step 5*. Otherwise, go to *step 4*.

Step 4: Conduct RDO motion estimation with all inter modes and calculate RDcost for all inter modes. Choose the best mode which minimizes RDcost.

Step 5: Go to a next macroblock.

We measure the performance using the H.264 scheme as a reference. Table 2 shows statistical data of SKIP mode using the proposed method. D indicates the ratio of the correctly detecting SKIP mode. E indicates the error ratio of saying "SKIP" even though it is a non SKIP mode. The proposed method shows that the average error ratio is only 0.03 in case of SKIP mode.

Table 2. Statistical data of SKIP mode using the proposed method

Sequence	Akiyo	Carphone	Claire	Container
D	0.59	0.50	0.59	0.44
E	0.02	0.06	0.03	0.04

Sequence	Paris	News	Silent	Foreman
D	0.27	0.50	0.33	0.33
E	0.01	0.03	0.03	0.04

4. Experimental Results

We have implemented the proposed method on the JM 10.2[7]. Experimental conditions are presented as shown in Table 1. Bjontegaard delta peak signal-to-noise ratio (BDPSNR) and Bjontegaard delta bit rate (BDBR) [6] are used to measure performance of the proposed method. The average time saving (TS) in process is calculated as following Eq.(7).

$$\Delta TS(\%) = \frac{T_{\mathrm{ref}} - T_{\mathrm{proposed}}}{T_{\mathrm{ref}}} \times 100 \quad (7)$$

where T_{ref} and T_{proposed} are the encoding times of JM 10.2 and modified JM according to proposed method, respectively.

Table 3. Experimental conditions

Profile	Baseline
QP	28,32,36,40
Reference frames	5
Search Range	-32~+32
Entropy	CAVLC
Hz	15
Coding Frames	150
GOP	IPPP
SIZE	QCIF

Table 4. Performance of the P-ESDM [2]

Sequence	ΔBDPSNR (dB)	ΔBDBR (%)	ΔTS (%)
Akiyo	0.03	-0.61	17.81
Carphone	0.02	-0.50	7.96
Claire	0	-0.07	15.45
Container	0.04	-1.02	16.64
Foreman	0	-0.02	6.92
News	0.01	-0.20	17.07
Paris	-0.02	0.33	13.05
Silent	0.01	-0.25	14.09
Average	0.01	-0.29	13.63

Table 5. Performance of the proposed method

Sequence	ΔBDPSNR (dB)	ΔBDBR (%)	ΔTS (%)
Akiyo	-0.02	0.68	52.5
Carphone	0.03	-0.76	21.01
Claire	-0.03	-0.07	49.93
Container	0.02	-0.64	36.8
Foreman	-0.01	0.35	14.7
News	-0.03	0.68	37.82
Paris	-0.02	0.34	30.31
Silent	-0.01	0.28	24.24
Average	-0.01	0.13	32.66

Tables 4 and 5 show the performances of the P-ESDM and the proposed method, respectively.

Experimental results show that the proposed method reduces encoding time of overall sequences by 33%, while the P-ESDM achieves only 14% with the average time saving when the encoder uses five references. The proposed method runs over twice as fast as the P-ESDM on an average. Specially, the proposed method is about three times faster than the P-ESDM for akiyo and claire sequences, as shown in Figure 4. However, PSNR degradation is 0.01dB and bitrate increases of 0.13% on an average. It is very negligible.

Figure 5 shows comparing RD-curves with JM 10.2, P-ESDM and the proposed method. It can be seen that the PSNR of the proposed method can achieve almost the same result as H.264/AVC.

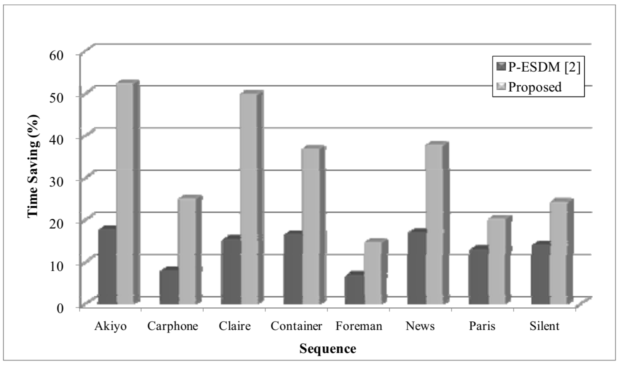
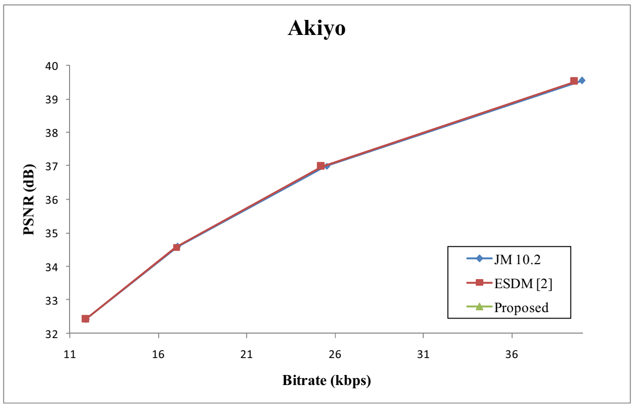
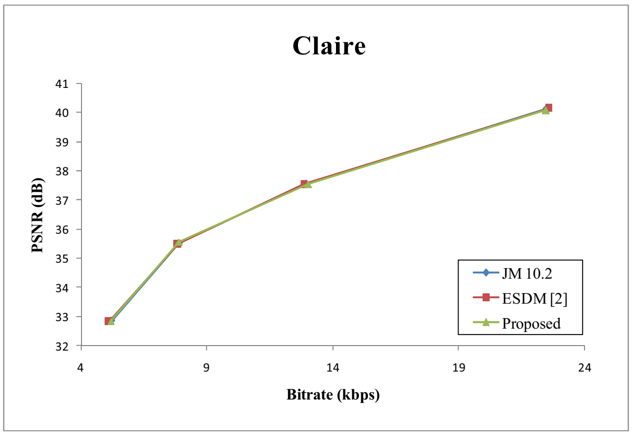


Figure 4. The average time saving comparison for sequences



(a) RD-curve for "Akiyo" sequence



(b) RD-curve for "Clarie" sequence

Figure 5. RD-curves comparisons among JM 10.2, P-ESDM and the proposed method

5. Conclusion

In this work, we proposed an early SKIP mode decision method. The proposed method used the PSRN prediction with the relation of between the coding *QP* and the reconstructed image's PSNR. For RD performances are very close to the standard, the proposed method affords significant time saving as compared to H.264/AVC. Our method is very relevant to low complexity coding system.

Acknowledgements

This work was supported by grant from Research Grant of Kwangwoon University, grant No.60012007179 and Electronics and Telecommunications Research Institute, grant No.90262007023 in 2007 .

References

[1] Ostermann. J, Bormans. J, List. P, Marpe. D, Narroschke. M, Pereira. F, Stockhammer. T, Wedi. T, "Video coding with H.264/AVC: tools, performance, and complexity," *IEEE Circuits and Systems Magazine*, vol. 4, no. 1, pp. 7-28, 2004.

[2] Choi. I, Lee. J and Jeon. B, "Fast Coding Mode Selection With Rate-Distortion Optimization for MPEG-4 Part-10 AVC/H.264", *IEEE Trans. Circuit and System for Video Technology*, vol. 16, pp. 1557-1561, Dec. 2006.

[3] K. Takagi, Y. Takishima, and Y. Nakajima, "A study on rate distortion optimization scheme for JVT coder," *in Proc. SPIE*, vol. 5150, pp. 914–923, 2003.

[4] Ma S, Wen Gao, Yan Lu, "Rate-Distortion Analysis for H.264/AVC Video Coding and its Application to Rate Control," *IEEE Trans. Circuit and System for Video Technology* , vol. 15, issue 12, pp. 1533 - 1544, Dec. 2005.

[5] Grecos. C and Yang. M, "Fast Mode Prediction for the Baseline and Main Profiles in the H.264 Video Coding Standard," *IEEE Trans. Multimedia*, vol. 8, issue 6, pp. 1125 - 1134 Dec. 2006.

[6] G. Bjonteggard, "Calculation of Average PSNR Differences between RD-curves", Doc. VCEG-M33, Apr. 2001.

[7] H.264/AVC Reference Software Model (JM 10.2):http://iphome.hi.de/uehring/ml/download/old_jm/