Proceedings of APCC2008 copyright © 2008 IEICE 08 SB 0083

Fast Multiple Reference Frame Selection Method Using Inter-mode Correlation

Chan-Young Jeong and Min-Cheol Hong, *Member, IEEE*
School of Electronic Engineering Soongsil University
Sangdo5-dong Dongjak-gu, Seoul, 156-743, Korea
E-mail : jesusjcy@vipl.ssu.ac.kr, mhong@ssu.ac.kr

***Abstract*-In this paper, we propose a fast multiple reference frame selection method for motion estimation and compensation in H.264 video coding. Using the property that an optimal reference frame has the statistical characteristics, the motion estimation of small block size is estimated from the motion estimation of the larger block size. Simulation results show that the proposal method decreased the computations about 52%. Without the sacrifice of coding performance, comparing to the H.264 multiple reference picture coding.**

Keywords : H.264/AVC, multiple reference frame, motion estimation, video coding

I. INTRODUCTION

H.264/AVC is the new video coding standard jointly developed by ITU-T VCEG and ISO/IEC MPEG [1]. In order to achieve better coding efficiency than existing standards, a number of new and advanced coding schemes have been adopted. In inter frame coding, this standard supports more flexibility in the selection of motion compensation block sizes and shapes than any previous standard, with a minimum luma motion compensation block size as small as 4×4[1] and multiple reference frames greatly improves the prediction precision[8]. The multiple reference picture coding aims to select a frame that has a minimum difference between an original input frame and the multiple decoded frames, so that the coding performance can be maximized at given encoding conditions. However, the encoding time and the computational complexity proportionally increase the number of reference frames in higher.

To reduce the computational complexity of multiple reference motion estimation, several approaches have been investigated. A fast multiple reference motion estimation algorithm is proposed by using a weighted average for motion compensation to eliminate the unnecessary motion estimation [4]. In ref [5], it composes a temporary predictive motion vector based on the motion vector map, and then refines it in a narrow search range. In ref [6], partial ME information is extracted to dynamically control the motion estimation search parameters.

In this paper, in order to reduce the number of reverence frames and find the exact match, we propose the fast multiple reference frame selection method that can efficiently determine the best reference frame from the allowed reference frame. The speed of the proposed method is over two times faster than that of the H.264 reference software JM13.2 while keeping the similar video quality and bit rate.

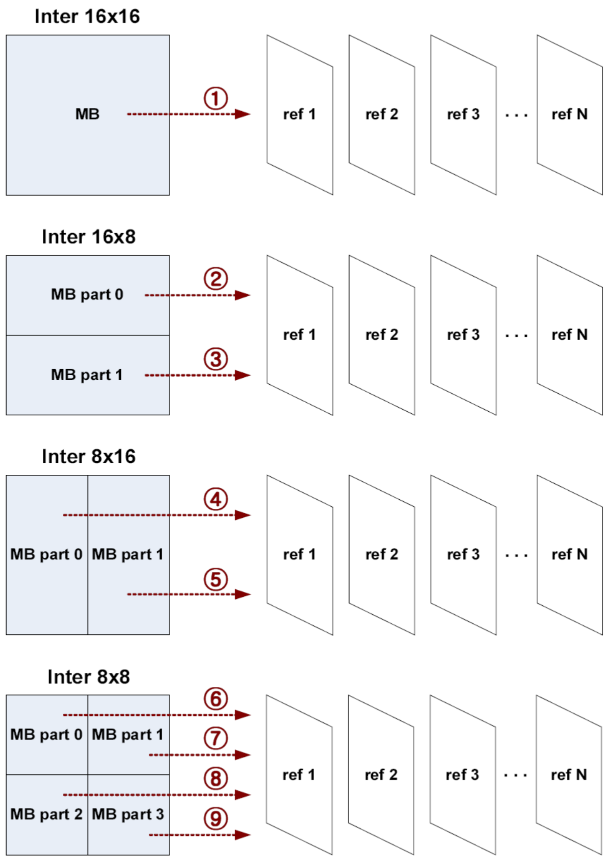

Fig. 1. Multiple reference frame motion estimation in H.264/AVC .

This paper is organized as follows. In Section II, we describe the proposed fast multiple reference frame selection method using inter-mode correlation. Simulation results are shown in Section III. Finally, conclusions are presented in Section IV.

II. PROPOSED MULTIFRAME MOTION ESTIMAION ALGORITHM

A. Statistical Characteristic between Inter-modes

Multiple reference frame motion estimation in H.264 is shown in Fig. 1. For a given inter mode, multiple reference picture selection of a macroblock aims to find a frame that is best match among the decoded reference frames. For example, motion estimation of macroblock-level such as 16x16, 16x8, 8x16 and 8x8 block sizes is performed for every reference frame. When 8x8 mode is determined as the best match mode, motion estimation for smaller block size such as 4x8, 8x4 and 4x4 block sizes is performed to obtain more accurate motion vector. It is observed that the reference frame determined in macroblock level is highly correlated to the optimal reference frame in small block size level.

In order to analyze the statistic characteristic between different modes, we have tested 5 sequences (stefan, table tennis, foreman, mobile and Claire) with various quantization index.

TABLE I
STATISTICAL CHARACTERISTIC BETWEEN INTER MODES

Sequences	Case A	Case B
Stefan	0.88	0.81
Table tennis	0.94	0.92
Foreman	0.89	0.86
Mobile	0.90	0.79
Claire	0.95	0.94
Average	0.91	0.86

Table 1 shows statistic characteristics when five reference frames are used for motion estimation. Case A represents the propability such that the choice of inter 16x16 mode in MB level results in the optimal choice of inter 16x8 or 8x16 inter-modes, rather than the choice of smaller block size inter mode. Also, Case B the probability with which inter 8x8 mode is determined, given that the condition inter 8x16 mode or 16x8 mode are known. The above results a can be described as

$$
\begin{aligned}
&Case\ A = P\left[\left(R^{m}_{16\times 8}(i) \cup R^{n}_{8\times 16}(i)\right) \middle| R_{16\times 16}(i)\right] \\
&Case\ B = P\left[R^{k}_{8\times 8}(i) \middle| \left(R^{m}_{16\times 8}(i) \cup R^{n}_{8\times 16}(i)\right)\right] \\
&\quad \text{for } i = 1, \ldots N \\
&\quad m:\ 16\times 8\ mode\ sub-MB\ index \\
&\quad n:\ 8\times 16\ mode\ sub-MB\ index \\
&\quad k:\ 8\times 8\ mode\ sub-MB\ index
\end{aligned}
$$

The selected optimal reference frame using motion estimation for the mode with arbitrary block size can be also selected in the mode with lower block size above the average 90%. Therefore, using the cost information of upper block mode, the number of reference frames that is used for motion estimation in lower block mode is restricted within a frame.

As shown in Fig. 2, If the object1 suddenly appears on ref(t-1) and not appears on ref(t-2) or ref(t-3), a lot of computation for multiple reference frames is wasted without any benefit. If we are able to predict whether current macroblock needs to multiple reference frame or not, we can adaptively decide the number of multiple reference frame and reduce the computational load of motion estimation for multiple reference frame.

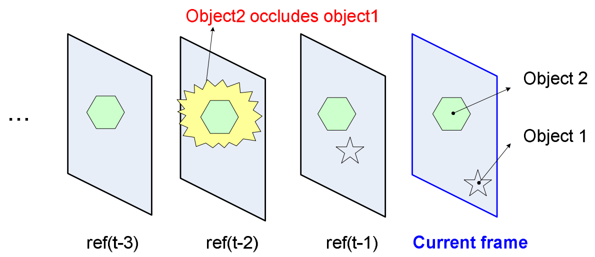

Fig. 2. Not efficient Example when using Multiple reference frame.

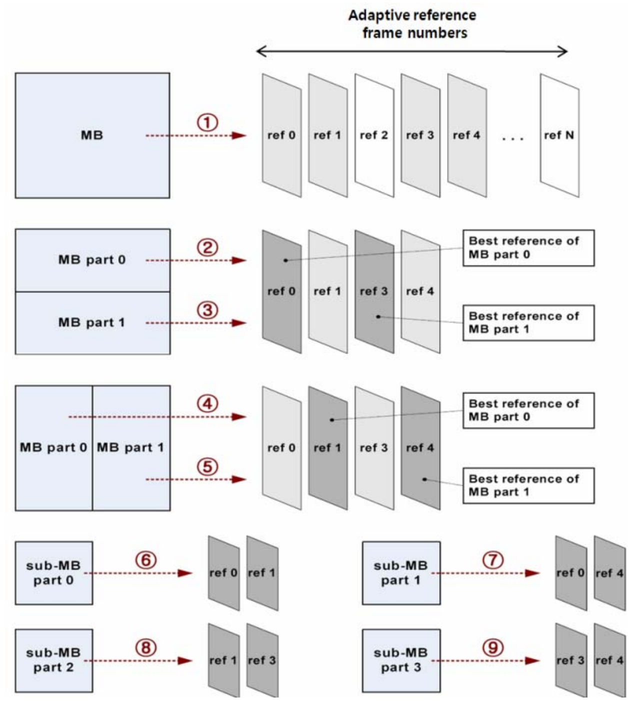

Fig. 3. Process of proposed algorithm.

B. *Fast Multiple Reference Frame Selection Method Using Correlation of Inter-modes*

Based on the statistic characteristic pointed out, we propose a fast multiple reference frame selection method. Fig. 3 shows the process of proposed algorithm.

The steps of the proposed algorithm are summarized as follows :

Step 1: If current macroblock can use four or more reference frames, go to step 2. Otherwise, go to step 3.

Step 2: Execute motion estimation of Inter16x16 mode for all available reference frames. As shown in Fig 4, if the cost of second or third reference frame is higher than the cost of first reference frame of motion estimation of Inter16x16 mode, stop searching the remaining reference frames. Otherwise, execute searching the remaining reference frames until the cost of current reference frame(t time) is higher than the cost of previous reference frame(t-1 time). Go to step 4.

Step 3: If there are available reference frames, execute motion estimation of Inter16x16 mode in order of available reference frames.

Step 4: Arrange the reference frames used in Inter16x16 mode , in order of small cost.

Step 5 : Execute motion estimation of Inter16x8 and Inter8x16 mode. If the number of arranged reference frames is more than four, only four reference frames arranged in order of small cost can be used for motion estimation for Inter16x8 and Inter8x16 mode.

Step 6 : Arrange the reference frames used in Inter16x8 mode and Inter8x16, in order of small cost, respectively. Select two optimal reference frame with small cost, respectively.

Step 7 : Execute motion estimation for Inter 8x8 mode. Only two reference frames are used for motion estimation. For example, as shown in Fig. 5, ref(0) is optimal reference frame of MB part 0 of Inter16x8 mode and ref(1) is optimal reference frame of MB part0 of Inter8x16 mode. Therefore only two reference frames, ref(0) and ref(1) are used for motion estimation of MB part 0 of Inter8x8 mode.

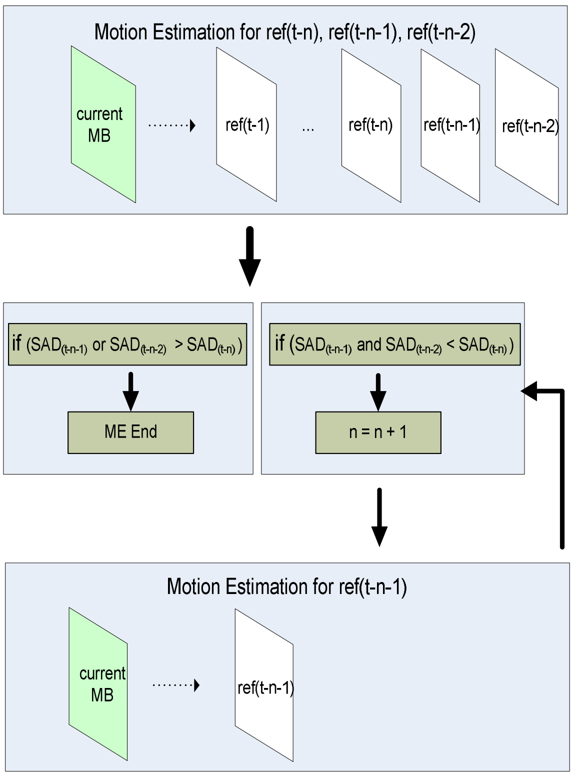

Fig. 4. Motion estimation for Inter 16x16 mode.

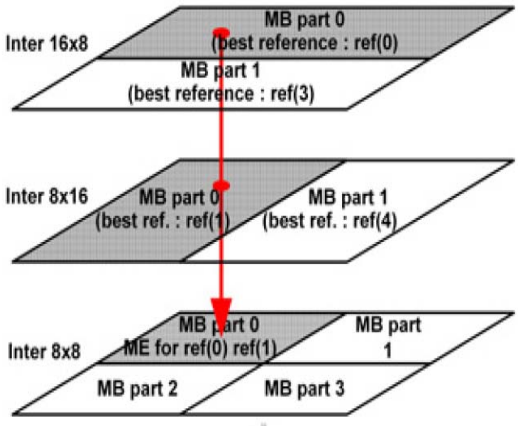

Fig. 5. Example of selection reference frame for motion estimation in Inter8x8 mode.

III. Simulation Results

The proposed algorithm is performed on H.264 reference software JM13.2. The simulation plat-form is Pentium IV 2.4GHz processor with 1G RAM. Four QCIF(176 x 144) sequences are tested to compare the performance of the proposed algorithm under different motion activities. The first two sequences contain smooth and slow motions with static background, while the other two sequences consist of zooming, panning and more vigorous motions. All sequences are coded 100 frames with IPPP structure. The motion search range is set to 32 and number of reference frames is set to 5. RD optimization is turned off. Four fixed QP values, 16, 24, 32, 40, are used in the experiments. For evaluating the performance of the algorithm, PSNR(Peak Signal to Noise Ratio) is used. For M x N dimensional 8-bit image, it is defined as

$$PSNR = 10 \log \frac{MN \times 255^2}{\left\| f - \hat{f} \right\|^2}, (2)$$

where $\|\bullet\|$ represents the Euclidean norm, and f and $\hat{f}$ denote an original frame and the reconstructed frame, respectively. The proposed algorithm is compared with MR5. Note that MR5 uses five reference frame and FS(Full Search) algorithm for motion estimation. For complexity comparisons, TMET(average of Total Motion Estimation Time) as a function of quantization index are shown in tables II, III and IV. RMP(%) means the ratio of proposal to MR5 In table II~IV.

For all the test sequences, the results of proposed algorithm have the highest speedup. Based on these results, the proposed algorithm achieves an average speedup ratio of 52%. the degradation of PSNR is only 0.00~0.15dB compared with the reference software JM13.2.

TABLE II
AVERAGE TMET , PSNR AND BITRATE FOR AKIYO SEQUENCE.

QP		16	24	32	40
TMET [sec]	MR5	778.35	701.00	540.40	381.27
	Proposal	383.17	369.85	320.41	239.16
	RMP(%)	49.23	52.76	59.29	62.73
PSNR [dB]	MR5	46.98	41.12	35.42	30.44
	Proposal	46.97	41.07	35.35	30.43
	Difference	-0.01	-0.05	-0.07	-0.01
Bitrate [kbits/s]	MR5	185.88	56.27	18.32	8.06
	Proposal	187.11	56.39	18.39	8.05
	RMP(%)	0.66	0.21	0.38	-0.12

TABLE III
AVERAGE TMET , PSNR AND BITRATE OF CONTAINER SEQUENCE.

QP		16	24	32	40
TMET [sec]	MR5	1301.43	1078.24	727.96	485.25
	Proposal	508.37	499.38	395.96	279.71
	RMP(%)	39.06	46.31	54.39	57.64
PSNR [dB]	MR5	44.89	38.70	33.51	28.39
	Proposal	44.89	38.67	33.43	28.36
	Difference	0.00	-0.03	-0.08	-0.03
Bitrate [kbits/s]	MR5	368.28	84.65	21.13	8.92
	Proposal	369.83	85.95	21.85	8.85
	RMP(%)	0.42	1.54	3.41	-0.78

TABLE IV
AVERAGE TMET , PSNR AND BITRATE OF FOREMAN SEQUENCE.

QP		16	24	32	40
TMET [sec]	MR5	1830.37	1492.45	1039.90	710.88
	Proposal	672.26	592.86	486.88	402.29
	RMP(%)	36.73	39.72	46.82	56.59
PSNR [dB]	MR5	45.23	39.07	33.76	28.56
	Proposal	45.16	38.97	33.61	28.49
	Difference	-0.07	-0.10	-0.15	-0.07
Bitrate [kbits/s]	MR5	666.79	228.01	84.80	34.46
	Proposal	687.95	233.88	84.37	34.20
	RMP(%)	3.17	2.57	-0.51	-0.75

TABLE V
AVERAGE TMET , PSNR AND BITRATE OF STEFAN SEQUENCE.

QP		16	24	32	40
TMET [sec]	MR5	2049.38	1854.48	1512.03	1137.35
	Proposal	759.84	714.70	612.74	516.96
	RMP(%)	37.08	38.54	40.52	45.45
PSNR [dB]	MR5	44.72	37.73	30.85	30.85
	Proposal	44.68	37.67	30.76	30.76
	Difference	-0.04	-0.06	-0.09	-0.09
Bitrate [kbits/s]	MR5	1560.21	641.97	198.84	63.07
	Proposal	1575.11	652.68	201.91	63.65
	RMP(%)	0.95	1.67	1.54	0.92

IV. CONCLUSION

In this paper, we propose a fast multiple reference frame selection method. Using the property that an optimal reference frame has the statistical characteristics, the motion estimation of small block size is estimated from the motion estimation of the larger block size. The experimental results show that motion estimation time(TMET) can be reduced average 52% comparing to JM13.2 with slightly reduced PSNR and increased bit-rate. But the proposed algorithm is really helpful to real time implementation of the H.264 encoder

ACKNOWLEDGMENT

This work was supported by the Seoul Future Contents Convergence (SFCC) Cluster established by Seoul R\BD Program and Soongsil University Research Fund.

REFERENCES

[1] Draft ITU-T Recommendation and Final Draft International Standard of Joint Video Specification (ITU-T Rec. H.264 | ISO/IEC 14496-10 AVC), May 2003.

[2] I. Richardson, *H.264 and MPEG-4 Video Compression*, Wiley, 2003.

[3] T. Wiegand, G. Sullivan, G. Bjontegaard and A, Luthar, "Overview of the H.264/AVC Video Coding Standard," *IEEE Trans. Circuit and Systems for video Technol.*, vol. 13, no. 9, pp. 560-576, July 2003.

[4] Yeping Su, Ming-Ting Sun, "Fast Multiple Reference Frame Motion Estimation for H.264/AVC", *IEEE Trans. Circuits Syst. Video Technol.*, vol. 16, no. 3 pp. 447-452, Mar. 2006.

[5] Sung-Eun Kim, et al., "An Efficient Scheme for Motion Estimation Using Multireference Frames in H.264/AVC," *IEEE Trans. on Multimedia*, vol. 8, no. 3, pp. 457-466, Jun. 2006.

[6] S. Saponara, et al., "Dynamic Control of Motion Estimation Search Parameters for Low Complex H.264 Video Coding," *IEEE Trans. Consumer Electronics*, vol. 52, no. 1, pp. 232-239, Feb. 2006.

[7] Joint Video Team Reference Software JM13.2[Online]. Available: http://iphome.hhi.de/suehring/tml

[8] Zhengnig Wang, Jun Yang, Qiang Peng, Changqian Zhu, "An Efficient Algorithm for Motion Estimation with Multiple Reference Frames in H.264/AVC", Image and Graphics., pp. 259-262, Aug. 2007.