IMPROVEMENT OF H.264 SKIP MODE

Kyohyuk Lee, Woojin Han and Tammy Lee
Digital Media R&D Center, Samsung Electronics 416, Maetan3-dong, Youngtong-gu
Suwon-city, Kyounggi-do, 442-742, Korea

Keywords: Video coding, video signal processing, H.264, MPEG-4 AVC, skip mode.

Abstract: H.264 (MPEG-4 AVC) is the state of the art international video coding standard which shows better coding efficiency compared to previous standards. This contribution is on the improvement of motion derivation process of H.264 SKIP mode. H.264 exploits temporal or spatial motion field correlation to derive current motion field. Temporal or spatial direct mode macroblock for B slice and skip mode macroblock for P slice are adopted for exploitation of motion field correlation. In general, H.264 SKIP mode macroblock has great impact on coding efficiency because about 30 ~ 70% of macroblocks are set as skip mode. SKIP mode macroblock derives one motion vector for whole 16x16 macroblock region from spatial correlation. In this contribution, we improved SKIP mode motion field further instead of setting one motion vector for 16x16 macroblock region. We split 16x16 macroblock into four 8x8 sub-partitions and set each sub-partition SKIP mode motion field separately. Experimental results showed average 2.05% and up to 18.63% bit rate reduction, especially higher coding efficiency in low bit rate condition.

1 INTRODUCTION

The coding efficiency of H.264 is much superior to those of previous standards due to several new features adopted for H.264 variable motion block sizes, multiple reference pictures, intra prediction, context adaptive entropy coding and etc0. H.264 not only has new features but also has useful conventional tools for video coding such as motion compensation, texture representation by prediction itself, motion prediction, transform and etc. In view point of motion field, H.264 exploit temporal or spatial correlation for prediction of motion field to reduce required bit amounts for coding. In general, spatial correlation has better preciseness than that of temporal correlation0. Motion fields are predicted from spatially adjacent blocks and the difference between current and predicted motion is coded and transmitted to the decoder side. For specific cases H.264 exploits temporal correlation too. Temporal direct mode of H.264 exploits temporal motion correlation. Temporal direct mode derives motion field from temporally co-located macroblock in reference picture and does not transmit additional bits for additional motion field refinement. If an object has temporally uniform motion characteristics, temporal motion correlation shows more robustness than spatial motion correlation at the edge of object0. Temporal and spatial direct modes exploit temporal and spatial motion correlation respectively. H.264 can select temporal or spatial direct for coding of B slice adaptively. But in coding of P slice, H.264 can exploit spatial correlation only. Motion field of SKIP mode in P slice consists of one motion vector which is derived from spatially adjacent blocks. That is to say, one motion vector derived from spatially adjacent blocks is used for all 16x16 pixels in one macroblock.

In this paper, we improved motion field derivation process of SKIP mode in H.264 P slice and could get meaningful results in low bit rate condition.

2 SPATIAL MOTION CORRELATION IN H.264 P SLICE

H.264 SKIP mode encodes a macroblock with one bit (SKIP mode bit). If SKIP mode bit is set, the macroblock uses prediction signal as texture representation as it is. Motion field of SKIP macroblock is derived from 4 spatially adjacent blocks. All motion information in a SKIP mode

Lee K., Han W. and Lee T. (2007).
IMPROVEMENT OF H.264 SKIP MODE.
In *Proceedings of the Second International Conference on Signal Processing and Multimedia Applications*, pages 143-146
DOI: 10.5220/0002141901430146

macroblock is considered to be equal to one motion vector derived from spatially adjacent blocks. Four spatially adjacent blocks are set as the following:

- A: Left block
- B: Above block
- C: Above-right block
- D: Above-left block

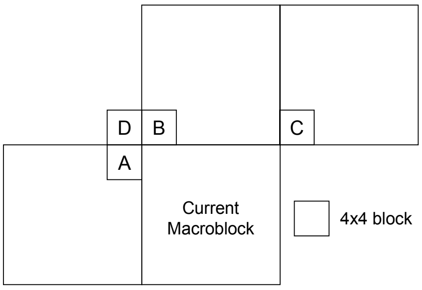

Figure 1: Spatially adjacent block position.

If above-right block is not available, above-left block is exploited instead of block C. Motion field derivation process of SKIP mode is described as follows:

- If top (A) or left block (B) is not available or has zero motion vector. Set the motion field as zero motion vectors.
- Else if one of 3 adjacent blocks has same reference index as current macroblock and the other 2 block have different reference indices from the current, set the motion field as the motion vector of the block which has same reference index.
- Else set the motion field as median of the motion vectors of A, B and C (or D).

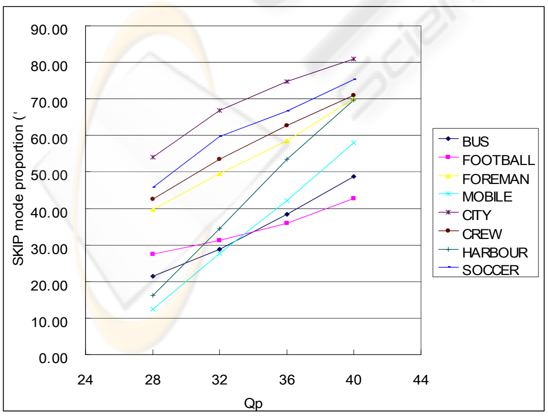

Figure 2: SKIP mode proportion in P slice.

Reference index of SKIP mode macroblock is set as that of the selected block in SKIP mode motion field derivation process.

A lot of macroblocks are set as SKIP mode in coding of P slices because SKIP mode meets the trade off between distortion and bit consumption well. Especially in low bit rate condition, proportion of SKIP mode macroblock increases because bit consumption of other macroblock modes does not meet the bit budget requirement. Figure 2 shows SKIP mode proportion in P slices (horizontal axis is quantization parameter values and vertical axis is SKIP mode proportion).

SKIP mode macroblock impacts coding efficiency not only in low bit rate condition but also in middle to high bit rate condition too. Figure 3 shows the performance graph of two cases (horizontal axis is bit rate in Kbps and vertical axis is PSNR). The first one is the performance of H.264 with exploitation of all macroblock modes. The other one is the performance of H.264 with exploitation of all macroblock modes except for SKIP mode. The first frame is coded as I slice and all the other frames are coded as P slice. Performance of coding without SKIP mode degrades dramatically in low bit rate condition.

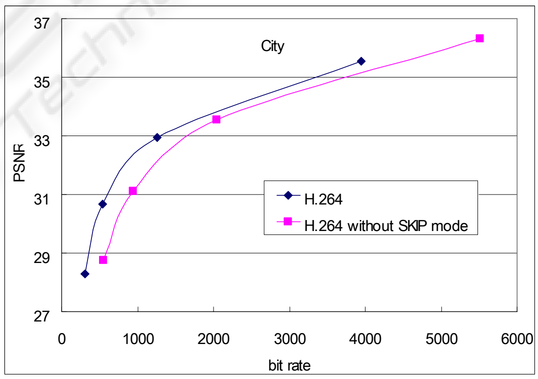

Figure 3: Performance graph with/without SKIP mode (city sequence).

3 IMPROVEMENT OF H.264 SKIP MODE

Drawback of H.264 SKIP mode is using single motion vector for setting of motion field inside the macroblock. That is to say, all 16x16 pixels inside the macroblock are set as same motion vector. In view point of motion field accuracy, the more the number of motion vector is, the more accurate the motion field is.

As we introduced in cluase 2, H.264 SKIP mode exploits four spatially adjacent blocks for derivation of motion field. When the above and left macroblock are partitioned as 16x16, motion field derivation process of H.264 SKIP mode seems to be reasonable because motion vectors of all spatially adjacent blocks can be clustered as four different motion vectors. But if the above or left macroblock has smaller motion partition than 16x16, current macroblock can exploits more than four different motion vectors. Figure 4 shows the case that above and left macroblock are splitted by 2 motion partition. In this case we have 6 different motion vectors as spatially adjacent. In extreme case, when 2 lower 8x8 blocks of above macroblock and 2 right 8x8 blocks of left macroblock are partitioned by 4x4 then current macroblock can exploit 10 different motion vectors for motion field derivation of SKIP mode. Figure D shows the case when spatial neighbor macroblock has motion partition less than 16x16.

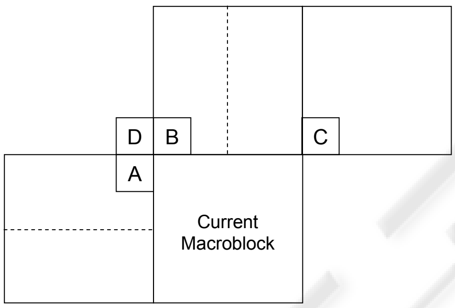

Figure 4: Small motion partition of above and left macroblock.

There are several approaches to improve motion field derivation process of H.264 SKIP mode. The first one is to change conventional derivation function (median of A, B and C or D) and the second one is to split 16x16 macroblock region into several sub partitions and to derive motion field for each partitions. And also we can combine the above two approach, changing the derivation function and splitting of macroblock region.

In this paper, we tried the second method, splitting of 16x16 macroblock region into sub-partitions and derivation of motion field for each sub-partitions. We split a macroblock into four 8x8 sub partitions and define spatially adjacent blocks for each sub-partitions respectively and derive SKIP motion field from motion vectors of previously defined spatially adjacent blocks. Processing order of four 8x8 sub partitions is equal to 8x8 mode of H.264 (zig-zag scan of four 8x8 sub partitions). And we exploit the derivation process in clause 2 without changing for derivation of motion field of each sub-partitions. Figure 5 shows spatially adjacent block position for each 8x8 sub-partitions.

All conventional SKIP modes can be replaced by new SKIP mode or new SKIP mode can be applied adaptively. We substituted new SKIP mode for all conventional SKIP mode.

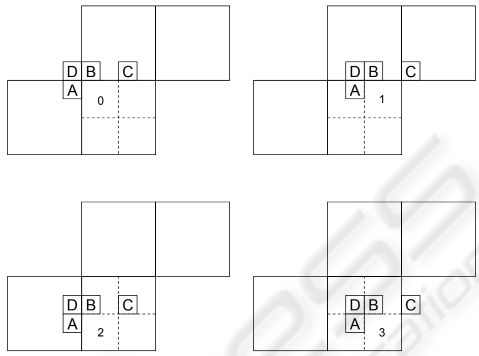

Figure 5: Spatially adjacent block position for each 8x8 sub-partitions.

4 EXPERIMENTAL RESULTS

Table 1: Experimental results.

		Qp	delta bit rate	deltal PSNR(dB)
CIF	BUS	30	-0.09%	-0.01
		36	0.49%	-0.02
		42	2.74%	0.00
	FOOTBALL	30	0.22%	-0.02
		36	0.43%	-0.01
		42	-0.39%	-0.01
	FOREMAN	30	1.10%	-0.03
		36	3.10%	-0.09
		42	6.61%	-0.11
	MOBILE	30	0.49%	-0.02
		36	1.99%	-0.04
		42	4.79%	-0.05
4CIF	CITY	30	-0.45%	0.02
		36	2.15%	0.00
		42	20.13%	-0.07
	CREW	30	0.56%	0.00
		36	1.87%	-0.01
		42	7.83%	-0.03
	HARBOUR	30	-0.35%	0.01
		36	-1.12%	0.00
		42	-0.94%	-0.16
	SOCCER	30	0.06%	0.02
		36	2.34%	0.01
		42	8.26%	-0.01

The proposed method was implemented in the reference software of H.264/AVC scalable extension which is under development in JVT as an scalable extension of H.264/AVC. H.264/AVC scalable extension consists of multi layer structure with backward compatibility to H.264/AVC in the lowest layer. Implementation was based on the lowest H.264/AVC compatible single layer condition.

Four CIF resolution sequences (bus, football, foreman and mobile) and four 4CIF resolution sequences (city, crew, harbour and soccer) were used for experiments. The first frame is coded as I slice and all the other frames are coded as P slice (no B slice). For all test sequences, 3 test points were tested.

Table 1: Experimental results shows delta bit rate and delta PSNR of exploitation of new SKIP mode compared to conventional SKIP mode. Minus values of ‘delta bit rate’ mean bit rate increase and plus values of ‘delta bit rate’ mean bit rate reduction. Minus values of ‘delta PSNR’ mean PSNR decrease and plus values of ‘delta PSNR’ mean PSNR increase.

As we can see in Table 1, when quantization parameter (Qp) goes high (low bit rate condition) we can reduce more bit amounts for coding.

In 2001, VCEG studied the relation between average PSNR differences and RD-curves0. Accoring to VCEG-M33, 0.05 dB PSNR change corresponds to 1% delta bit rate. Figure 6 shows final delta bit rate under consideratin of PSNR change. Horizontal axis represents sequence name and Qp value (i.e. BUS_30 represents BUS sequence with Qp 30). Vertical axis represents final delta bit rate according to VCEG-M33. We could get average 2.05% and up to 18.6% bit saving through proposed new SKIP mode motion derivation process.

We tried 4x4 size as sub-partition (splitting of a macroblock into 16 sub-partitions) for derivation of SKIP mode motion field, but we could not get additional improvement. If we set the sub-partition size as 4x4, above or left macroblock should be partitioned as 4x4 for additional improvement. But the proportion of 4x4 sub-partitioning in spatially adjacent macroblock is so small.

5 CONCLUSION

In this paper, we proposed a new motion field derivation process of SKIP mode. We splitted SKIP mode macroblock region into four 8x8 sub-partitions and derived SKIP mode motion field for each sub-partitions. We could get up to 18.6% bit rate reduction.

For further work, we would like to develop new motion field derivation function instead of median.

REFERENCES

Thomas Wiegand, Gary J. Sullivan, Gisle Bjontegaard and Ajay Luthra, July 2003. *Overview of the H.264/AVC Video Coding Standard*. IEEE Transactions on Circuits and System for Video Technology, vol.13, no.7 pp.560-576.

Alexis Michael Tourapis, Feng Wu and Shipeng Li, January, 2005. *Direct Mode Coding for Bipredictive Slices in the H.264 Standard*. IEEE Transactions on Circuits and System for Video Technology, vol.15, no.1 pp.119-126.

Gisle Bjontegaard, April, 2001. *Calculation of average PSNR differences between RD-curves*. Video Coding Expert Group (VCEG), Doc. VCEG-M33, Bangkok, Thailand.

Kyohyuk Lee and et al, July, 2005. *Motion prediction in temporally enhanced picture and improvement of H.264 TDM*. Joint Video Team (JVT), Doc. JVT-P075, Poznan, Poland.

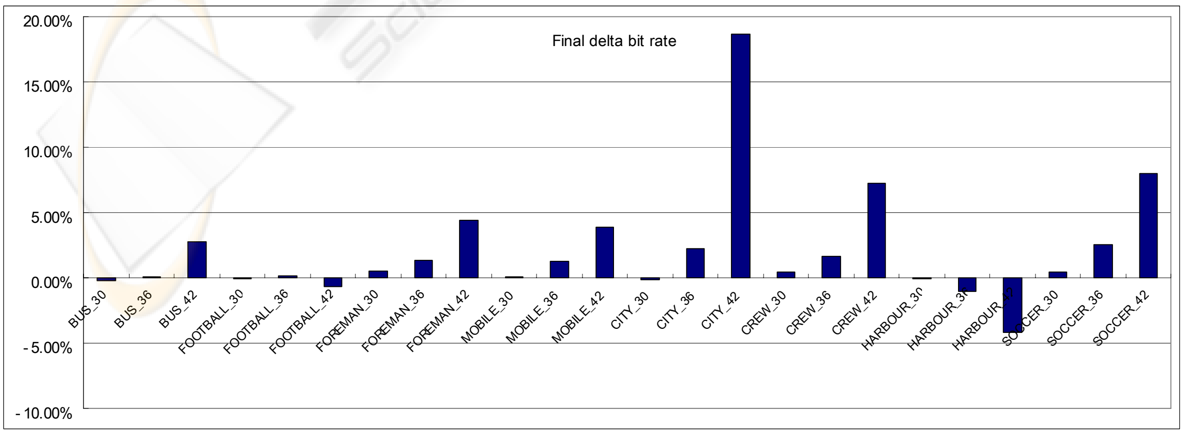

Figure 6: Final delta bit rate under consideration of PSNR change.