A SIMPLIFIED MOTION ESTIMATION USING AN APPROXIMATION FOR THE MPEG-2 REAL-TIME ENCODER

Yuzo Senda, Hidenobu Harasaki and Mitsuharu Yano

Information Technology Research Laboratories, NEC Corporation
1-1, Miyazaki 4-chome, Miyamae-ku, Kawasaki-shi, KANAGAWA 216 JAPAN
Email: senda@DSP.CL.nec.co.jp

ABSTRACT

We propose a simplified motion estimation method which provides motion vectors for all types of motion compensation used in MPEG-2. The method is a result of applying a newly introduced approximation to the canonical three-step method described in MPEG-2 Test Model. It reduces the number of necessary computations in the second and the third steps to less than **1%**, and that of data transfers to about **8%** of the canonical method. Total complexity of the proposed method is nearly that of the full-pel search motion estimation.

1. INTRODUCTION

In moving picture coding, motion compensation (MC) is essential to reduce temporal redundancy. As a most basic measure, full-pel MC has been used in ITU-T H.261 codecs. Half-pel and interpolative MC were then introduced to MPEG-1 [1]. Recently, MPEG-2 [2] adopted dual-prime prediction to adapt to an interlaced signal. These MC methods improve picture quality significantly [3]. The improvement obtained by the latter three MC methods is about 1~2dB in signal to noise ratio (SNR) at 4Mbps.

However, it is not easy to implement motion estimation (ME) which supports all types of MC used in MPEG-2. The ME requires huge numbers of computations and data transfers. The number of computations to achieve half-pel ME is about four times as many as full-pel ME. For this reason, most of conventional half-pel ME methods [4, 5] are based on a logarithmic step search which consists of two or more steps [6]. For example, the canonical method in MPEG-2 Test Model [7] consists of three steps. In the first step, a full-pel MV is calculated by the exhaustive-search method. Eight candidate half-pel MVs around the full-pel MV are evaluated in the second step. Interpolative MC or dual-prime MC is evaluated in the third step. Both in the second and in the third steps, the MC blocks corresponding to candidate MVs are generated using either spatial or temporal interpolation. The three-step ME reduces the necessary computations, but it requires much more data transfers for the second and the third steps compared with the single-step exhaustive-search method which can reduce the data transfer by using search windows [8], thus increasing complexity. As other solutions, there are some simplified methods [9, 10], but they can not be applied to interpolative and dual-prime MC because they are based on a mathematically continuous MV-MAE model. Therefore, ME has still been one of the key points of moving picture coding.

In this paper, we propose a simplified ME method using an approximation. This paper is organized as follows. Section 2 describes the principle of the approximation, and the possibility of applying the approximation to half-pel, interpolative and dual-prime MC. The process of the proposed method is also explained. Section 3 discusses the reduction of necessary computations and data transfers, and the performance of the proposed method compared with the canonical method. Finally, Section 4 concludes this paper.

2. METHODS

2.1. Approximation

A most major matching function is mean absolute error (MAE) defined as:

$$d(P) = \frac{1}{256} \sum_{i=1}^{16} \sum_{j=1}^{16} |P_{i,j} - S_{i,j}| \qquad (1)$$

where each $P_{i,j}$ is a pel of the prediction block P, each $S_{i,j}$ is a pel of the original block S, and $d(P)$ is the MAE of P.

By definition, half-pel, interpolative and dual-prime MC are average prediction. Generalized average pre-

0-7803-2431-5/95 $4.00 © 1995 IEEE

diction is written in the form:

$$P_a = \frac{1}{n}\sum_{k=1}^{n} P_k \qquad (2)$$

where P_a is the average prediction block, each P_k is the block to be referred to generate the average prediction block, and n is the number of the referred blocks.

The MAE of the average prediction is always less than or equal to the average of the referred block MAEs:

$$d(P_a) \leq \frac{1}{n}\sum_{i=1}^{n} d(P_i) \qquad (3)$$

and there is a high correlation between the left and the right terms. So, we propose the following approximation to evaluate the average prediction blocks without generating them.

$$\tilde{d}(P_a) = \psi\frac{1}{n}\sum_{i=1}^{n} d_i \qquad (4)$$

where $\tilde{d}(P)$ is the approximate MAE of P, and ψ is an approximation factor smaller than one.

2.2. Approximation of half-pel MC

Half-pel MC behaves

$$\begin{aligned} P(vx,vy) &= \frac{1}{4}[P(\lfloor vx\rfloor,\lfloor vy\rfloor) + P(\lceil vx\rceil,\lfloor vy\rfloor) + \\ &\quad P(\lfloor vx\rfloor,\lceil vy\rceil) + P(\lceil vx\rceil,\lceil vy\rceil)] \end{aligned} \qquad (5)$$

where (vx, vy) is the motion vector, P is the MC predicted block, $\lfloor x\rfloor$ is the floor function which returns the greatest integral value less than or equal to x, and $\lceil x\rceil$ is the ceil function which returns the least integral value greater than or equal to x.

As a result of applying the approximation, the approximate MAE of the half-pel MC block is calculated from the MAEs of the full-pel MC blocks:

$$\tilde{d}(P(vx,vy)) = \begin{cases} \psi_h\frac{1}{2}[d(P(\lfloor vx\rfloor,vy)) + d(P(\lceil vx\rceil,vy))] & \text{if only } vx \text{ is fractional} \\ \psi_v\frac{1}{2}[d(P(vx,\lfloor vy\rfloor)) + d(P(vx,\lceil vy\rceil))] & \text{if only } vy \text{ is fractional} \\ \psi_{hv}\frac{1}{4}[d(P(\lfloor vx\rfloor,\lfloor vy\rfloor)) + d(P(\lceil vx\rceil,\lfloor vy\rfloor)) + \\ \quad d(P(\lfloor vx\rfloor,\lceil vy\rceil)) + d(p_{mc}(\lceil vx\rceil,\lceil vy\rceil))] & \text{if } vx \text{ and } vy \text{ are fractional} \end{cases} \qquad (6)$$

where ψ_h,ψ_v and ψ_{hv} are approximation factors.

2.3. Approximations of interpolative and dual-prime MC

The behaviors of interpolative and dual-prime MC are simpler than that of half-pel MC:

$$P_i = \frac{1}{2}(P_f + P_b) \qquad (7)$$

$$P_d = \frac{1}{2}(P_o + P_e) \qquad (8)$$

where P_i is the interpolative MC block, P_f and P_b are the forward and backward MC blocks, respectively, P_d is the dual-prime MC block, P_o and P_e are the blocks predicted from odd and even fields as components of P_d, respectively. The approximate MAEs of the interpolative and dual-prime blocks are also calculated by applying the approximation:

$$\tilde{d}(P_i) = \psi_i\frac{1}{2}[d(P_f) + d(P_b)] \qquad (9)$$

$$\tilde{d}(P_d) = \psi_d\frac{1}{2}[d(P_o) + d(P_e)] \qquad (10)$$

where ψ_i and ψ_d are approximation factors.

2.4. Implementation

The proposed method is a result of applying the approximation to the canonical three-step method described in MPEG-2 Test Model. Figure 1 shows the block diagram of the proposed method. The second and the third steps of the proposed method require no picture data transfer. They only require MAE data transfers from the first step.

MPEG-2 has layered restrictions, namely profiles. Main profile (MP) supports all MC types except dual-prime MC, simple profile (SP) supports dual-prime MC instead of interpolative MC. Accordingly, the process of the proposed method for both MP and SP modes is explained step by step.

1st Step Search full-pel MVs exhaustively. Backward MVs are also searched in MP mode. Memorize all MAEs in search range to be accessed from the second and third steps.

2nd Step Refine each of the full-pel MVs to half-pel accuracy. The eight candidate half-pel MVs around the full-pel MV are evaluated by approximate MAE instead of MAE.

3rd Step In MP mode, calculate the approximate MAEs of interpolative MC from the approximate MAEs of the forward and backward half-pel MVs. In SP mode, scale the half-pel MVs to the opposite field, and modify it with differential MV (dmv).

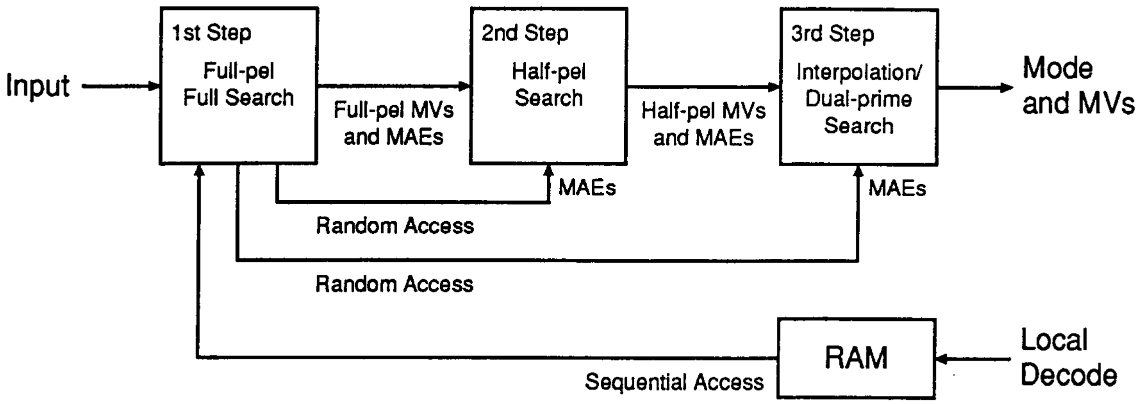

Figure 1: The block diagram of the proposed method.

Table 1: The number of necessary computations.

Mode	Method	Computations[Mop/s]	
		2nd step	3rd step
Main Profile	Test Model	1710.2	165.7
	Proposed	4.9	0.2
Simple Profile	Test Model	855.1	1722.9
	Proposed	2.4	9.7

Table 2: The number of necessary data transfers.

Mode	Method	Data transfers[Mbyte/s]	
		2nd step	3rd step
Main Profile	Test Model	84.5	48.1
	Proposed	6.5	0.0
Simple Profile	Test Model	42.2	107.8
	Proposed	3.2	8.4

Then, obtain the approximate MAEs of the scaled MV, and find the approximate MAEs of the dual-prime MC. Finally, select the best set of MC mode and MVs.

3. DISCUSSION

The followings are assumed to estimate how much the proposed method reduces the ME costs.

- Frame rate is 29.97Hz, picture size is 720x480.
- Block matching takes three operations per pel.
- Averaging takes one operation per pel.
- An MAE access equals to 2 byte data transfers.

Table 1 and 2 indicate that the numbers of the computations and data transfers of the proposed method are significantly reduced than the canonical method.

Table 3: The approximation factors used in the experiments.

Mode	Mbps	Approximation factors (×16)				
		ψ_h	ψ_v	ψ_{hv}	ψ_i	ψ_d
Main Profile	4	14	14	13	14	—
	15	13	13	12	13	—
Simple Profile	4	14	14	13	—	13
	15	13	13	12	—	12

Therefore, total complexity of the proposed method is nearly that of the full-pel ME.

However, the performance of the proposed method can be less than that of the canonical method because of the approximation. It should be made clear that its degradation is negligible.

The proposed method uses five approximation factors, namely ψ_h, ψ_v, ψ_{hv}, ψ_i, and ψ_d, each corresponding to horizontal half-pel, vertical half-pel, horizontal and vertical half-pel, interpolative, and dual-prime MC, respectively. The factors should be optimized to minimize the performance degradation, and also be simple for implementation. Thus, the factors are selected out of n/16, where n = 11 to 15. The optimum factors shown in Table 3 have been obtained from preliminary experiments.

Experiments to obtain the performance have been carried out using five sequences, each of 60 frames. Figure 2 and 3 summarize results measured in the peak-to-peak SNR. The performance differences from the canonical method are only -0.15 to -0.54 in main profile, and +0.02 to -0.26 in simple profile. Consequently, the performance of the proposed method is almost the same as the canonical method.

Implementing an improved version of the proposed method, we have developed the **real-time** MPEG-2 codec shown in Figure 4 and Table 4.

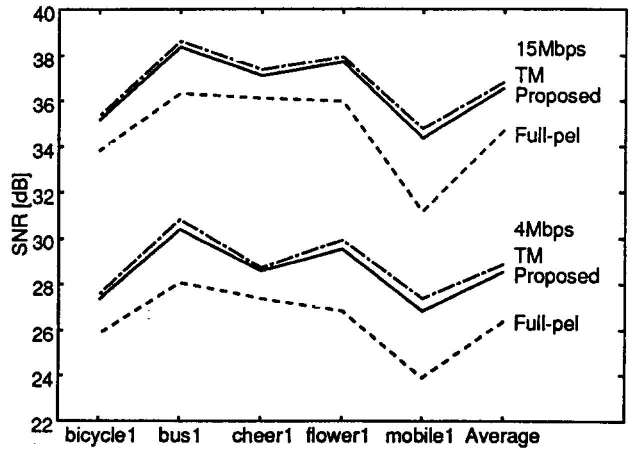

Figure 2: The performance of the proposed method compared with the canonical method in Main Profile.

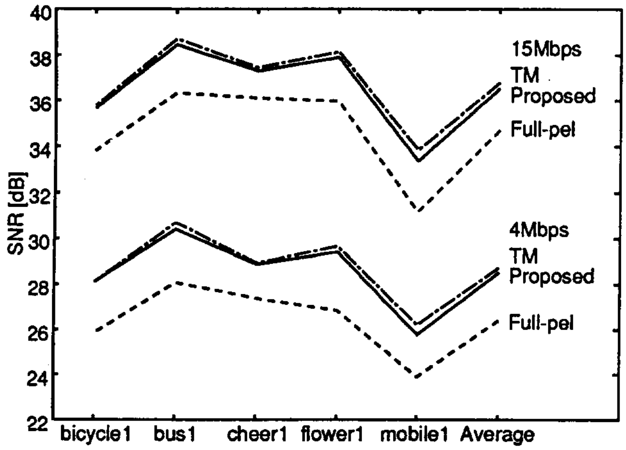

Figure 3: The performance of the proposed method compared with the canonical method in Simple Profile.

4. CONCLUSION

We have proposed a simplified ME method using an approximation. It reduces total complexity to nearly that of the full-pel ME without performance degradation.

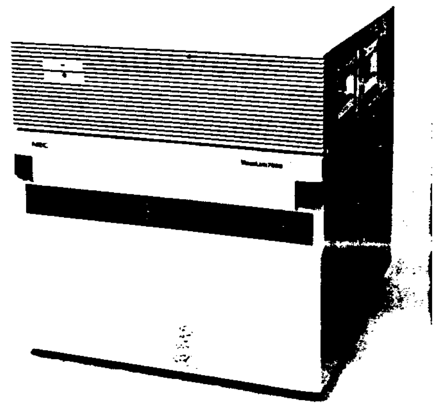

Figure 4: The photograph of VisuaLink 7000.

Table 4: The specification of VisuaLink 7000.

Video	MPEG-2 Video SP/MP@ML+4:2:2 (1) NTSC analog composite (2) D1 serial digital component
Audio	MPEG-1 Audio layer 2 (1) Analog (2) Serial digital
Mux	MPEG-2 System TS/PS
Network	(1) 6.312Mb/s I-interface (2) ATM 155.52Mb/s SDH
Size	445Wx470Hx500D

5. REFERENCES

[1] ISO-IEC/JTC1/SC29/WG11, "Coding of moving pictures and associated audio for digital storage media at up to about 1.5 Mbit/s," IS11172-2 (1992)

[2] ISO-IEC/JTC1/SC29/WG11, "Generic coding of moving pictures and associated audio," IS13818-2 (1994)

[3] B. Girod, "Motion-Compensating Prediction with Fractional-Pel Accuracy," IEEE Trans. Com., Vol.41, No.4, pp.604–611 (1993)

[4] K. M. Yang, L. Wu, I. G. Chuang, "Very High Efficiency VLSI Chip-pair for Full Search Block Matching with Fractional Precision," ICASSP89, V2.5, pp.2437–2440 (1989)

[5] S. Uramoto, A. Takabatake, M. Suzuki, H. Sakurai, M. Yoshimoto, "A Half-pel Precision Motion Estimation Processor for NTSC-resolution Video," CICC93, pp.11.2.1–11.2.4 (1993)

[6] M. Ghanbari, "The Cross Search Algorithm for Motion Estimation," IEEE Trans. COM, Vol.38, No.7, pp.950–953 (1990)

[7] ISO-IEC/JTC1/SC29/WG11, "Test Model 5 (Draft)," MPEG93/N0400 (1993)

[8] A. Artieri, F. Jutand, "A Versatile and Powerfull Chip for Real Time Motion Estimation," ICASSP89, V2.9, pp.2453–2456 (1989)

[9] J. Jeong, W. Ahn, "Subpixel-accuracy Motion Estimation Using a Model for Motion Compensated Errors," PCS93, 13.4 (1993)

[10] SGS-Thomson, "Motion Estimation Controller," STi3223, Preliminary Data Sheet, V2.0 (1993)