New Motion Estimation Algorithm and Its Block-Matching Criteria Using Low-Resolution Quantization

Seongsoo Lee and Soo-Ik Chae

School of Electrical Engineering, Seoul National University
San 56-1, Shinlim-dong, Kwanak-gu, Seoul, 151-742, KOREA
Phone: +82-2-880-1815, Fax: +82-2-882-4656, E-Mail: chae@sdgroup.snu.ac.kr

Abstract - We propose a new motion estimation algorithm and its block-matching criteria using low-resolution quantization. The proposed algorithm reduces both the huge computational cost of the full search algorithm and the performance degradation of the fast algorithms by matching the low-resolution images. Two search steps called the low-resolution search and the full-resolution search are employed. Simulation results show that the PSNR of the proposed algorithm is superior to those of the 4:1 alternate subsampling algorithm with less computational cost. Its computational cost is 1/38.1 of the full search algorithm.

I. Introduction

The block-matching motion estimation algorithm, in which the current frame is partitioned into equal-sized reference blocks and the best matching block of each reference block is searched in the previous frame, is widely used in image compression because of its simplicity and high efficiency. The full search algorithm [1] is optimum in the compression ratio, but the huge computational cost makes it impractical when the search window is large. To overcome this problem, several fast algorithms [2]-[4] are proposed, but they suffer from poor performance due to the local minima.

In this paper, we explain new motion estimation algorithm and block-matching criteria using low-resolution quantization. The reference block is compared at every search position to avoid a local minimum, and the low-resolution quantized images are matched to reduce the computational cost. Two-step search is employed to reduce the performance degradation due to matching low-resolution quantized image.

II. THE PROPOSED ALGORITHM

The block-matching criterion plays an important role in the motion estimation algorithm, and occupies the majority of the computational cost. The SAD is widely used in most conventional motion estimation algorithm because of its efficiency and simplicity, but it is impractical for VLSI implementation as the search window is larger. Therefore, the need for new low-computational-cost block-matching criterion is emerging.

The computational cost of the block-matching criterion can be lowered if low bit-resolution image is used, because the operation of the low bit-resolution requires less hardware (i.e. less computational cost) than that of full bit-resolution [5]. In this paper, two block-matching criteria called the different pixel count (DPC) and the pseudo-SAD (PSAD) are proposed, and will be described in Section II-A.

The low bit-resolution block-matching criterion is effective in the computational cost, but suffers from the performance degradation because the reduction in bit-resolution means the loss of pixel information. Indeed, the low bit-resolution block-matching criterion is only the “approximation” of the conventional one, and only indicates that the position has “high possibility” to be a correct motion vector. Consequently, the proposed algorithm employs two steps called low-resolution search (LRS) and full-resolution search (FRS), as shown in Fig. 1. The LRS calculates the DPC or PSAD from the low-resolution image and finds a set of candidate positions for motion vector (CMV set). The FRS calculates the SAD from the full-resolution image (i.e. original image) at each CMV position, and finds the motion vector. The pixel value should be quantized to minimize the quantization error, and this will be described in Section II-B.

Huge computational cost of the conventional motion estimation algorithm is mainly due to the large search window. To further reduce this computational cost, only part of the search window is selected as the candidate area for search window (CSW set) and the motion estimation is performed in each CSW area in the proposed algorithm. The selection of the CSW is described in Section II-C.

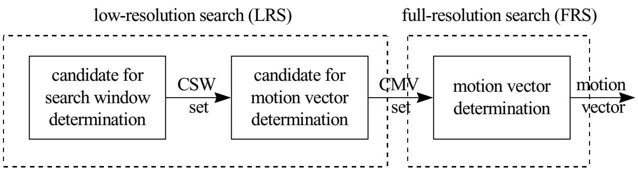

Fig. 1. Block diagram of the proposed algorithm.

A. Block-matching criterion of the low-resolution image

The block-matching criterion of the low-resolution image should satisfy three conditions: Low computational cost, Good approximation of the SAD, and Simple VLSI implementation. We propose two hardware-oriented block-matching criteria: the DPC and the PSAD.

The DPC is the number of pixels whose low-resolution quantized code, and is defined in (1). The DPC calculation can be implemented by simple logic gates instead of 8-bit adders in SAD calculation. Therefore, the computational cost is greatly reduced.

$$DPC(u,v)=\sum_{i=1}^{N}\sum_{j=1}^{N}\bar{\delta}\left[\hat{f}_k(i,j),\hat{f}_{k-1}(i+u,j+v)\right] \qquad (1)$$

where $\hat{f}_k(x,y)$ is the quantized code at the position (x,y) in k-th frame, and $\bar{\delta}(x,y)=1$ if $x \neq y$, 0 otherwise.

The PSAD is the sum of absolute difference of the reconstructed pixel value, and is defined in (2). It is the approximation of the SAD in the low-resolution. In 2-bit resolution, the PSAD calculation can be implemented by simple logic gates like the case of the DPC.

$$PSAD(u,v)=\sum_{i=1}^{N}\sum_{j=1}^{N}\left|\tilde{f}_k(i,j)-\tilde{f}_{k-1}(i+u,j+v)\right| \qquad (2)$$

where $\tilde{f}_k(x,y)$ is the reconstructed pixel value at the position (x,y) in k-th frame.

The PSAD is more accurate than the DPC as the block-matching criterion because the PSAD is more analogous to the SAD than the DPC. But in the viewpoint of computational cost, the DPC is better than the PSAD because the DPC is the N^2 accumulation of 1-bit value and the PSAD is the N^2 accumulation of k-bit value when the k-bit low-resolution image is used. The designer should choose either the DPC or the PSAD considering the trade-off between the performance and the computational cost.

B. Generation of the low-resolution image

Fig. 2 illustrates the generation of the low-resolution image. To preserve the dynamic range of the pixel value maximally, each pixel value is subtracted by the block mean, and then quantized. To reduce the computational cost to calculate the block mean, only the mean of the reference block is used for the low-resolution image of both the reference block and the search window. Simulation results indicate that 2-bit resolution is optimal in both performance and computational cost.

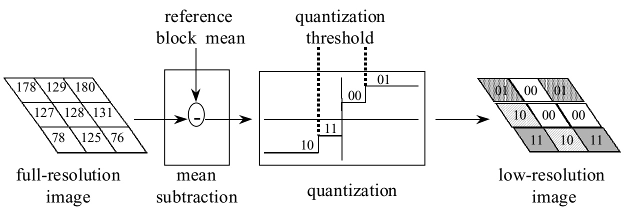

Fig. 2. Generation of the low resolution image.

The performance degradation of the proposed algorithm is mainly due to the loss of pixel information. Therefore, the quantization thresholds of each reference block are adaptively determined to minimize the average quantization error, and they can be calculated from the non-linear quantization theory [6]. For 2-bit quantization, the three quantization thresholds t_1, t_2, t_3 are given in (3).

$$t_3=-t_1=\frac{3}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\left|f_k(i,j)-\overline{f_k}\right| \;,\; t_2=0 \qquad (3)$$

where $f_k(x,y)$ is the pixel value at the position (x,y) in k-th frame, and $\overline{f_k}$ is the mean of $f_k(x,y)$.

C. Selection of the candidate area for search window

When the picture size becomes larger, the maximum value of the motion vector also becomes larger. Therefore, the search window should also become larger, which results in the increase of the computational cost. To overcome this problem, the search window is partitioned into many small area, and the candidate areas for search window (CSW set) is selected using low-resolution block-matching criterion in the proposed algorithm. Then, the motion estimation is performed only in each CSW area.

Selection of the CSW set need not be very accurate, because we have to know only the area which the motion vector is in, not the exact position of the motion vector. In the proposed algorithm, both the reference block and the search window are 4:1 averaging subsampled, and then are compared. Thus, the computational cost of the CSW selection is negligibly low.

Fig. 3 illustrates the selection of the CSW set when the search window is 64x64. The search window is partitioned into 21 16x16 areas, and the motion estimation low-resolution block-matching criterion. In the proposed algorithm, the PSAD is used in the CSW selection. The minimum PSAD score within the 16x16 area is determined as the PSAD score of the area, and the K areas which have the minimum PSAD scores are determined as the CSW set.

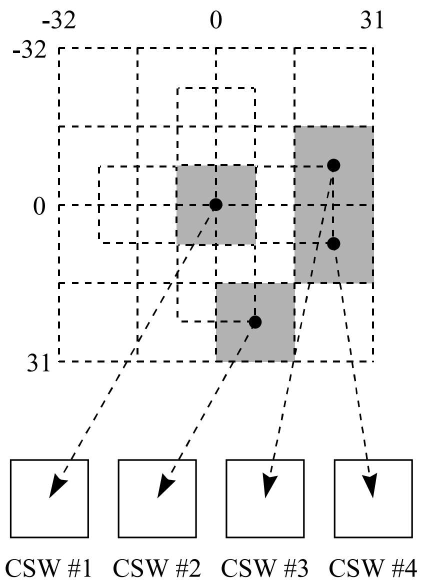

Fig. 3. The selection of the CSW.

Besides 16 non-overlapping 16x16 areas, 5 additional areas are employed. This comes from the fact that most of the movement in the real world are horizontal only or vertical only.

D. The proposed algorithm

The outline of the proposed algorithm is as follows:

(1) For each N×N reference block in the current frame, the mean of the reference block is calculated and the quantization threshold is determined.
(2) Pixel values of both the reference block and the search window are 4:1 averaging subsampled.
(3) The subsampled pixels are subtracted by the block mean, and the quantized in 2-bit resolution to generate a low-resolution images.
(4) The M×M search window is partitioned into 21 M/2 ×M/2 areas, and the PSAD score is calculated in each area. The 4 areas with minimum PSAD are determined as the CSW set.
(5) For 4 CSW areas, the pixels in the reference block and the search window are subtracted by the block mean, and the quantized in 2-bit resolution to generate a low-resolution images. This step is essentially same as (3), except that in (5) unsubsampled pixels are used.
(6) For 4 CSW areas, the DPC is calculated for every search position in the CSW area. 2 search positions with minimum DPC are found in every row of search position for the purpose of pipelining, and determined as the CMV sets.
(7) The SAD at each CMV position is calculated with full-resolution images. The candidate with minimum SAD is determined as the motion vector of a given reference block. To reduce the computational cost, 4:1 alternate subsampling [4] is used.

III. SIMULATION RESULTS

In this paper, the computational cost is defined as (total cycle per reference block) x (total gate count of all PE's). This computational cost proportional to the rough estimation of the PE hardware amount when the throughput of the all algorithm is the same. The peak signal-to-noise ratio (PSNR) is used as a performance measure. In the computational cost of the proposed algorithm, the gate counts of the quantizer, block mean calculation, and the quantization threshold determination are included.

Computer simulations were carried out on 40 frames of "football", "flower garden", "table tennis", and "popple" sequences, whose picture formats are CCIR601 (704 x 480 pixels, 8 bit/pixel, 30 frame/sec). The reference block size is 16 x 16, and the search window size is 64 x 64 (the range of motion vectors is (-32,-32)~(31,31)).

TABLE I

COMPUTATIONAL COSTS AND SIMULATION RESULTS.

Algorithm	cycle per reference block	total gate count of all PE's	computational cost (gate • cycle / block)	average PSNR (dB)			
				football	flower garden	table tennis	popple
FS [1]	1,024	400,384	4.10×10^8	25.11	27.03	28.26	30.01
4:1AS [4]	1,028	100,096	1.03×10^8	24.65	26.96	28.04	29.77
1DFS [2]	1,920	12,512	2.40×10^7	23.67	26.66	27.29	29.76
LRQME [5]	1,024	32,416	3.31×10^7	24.99	26.97	28.23	29.92
Proposed	1,024	10,504	1.08×10^7	24.87	26.95	28.15	29.81

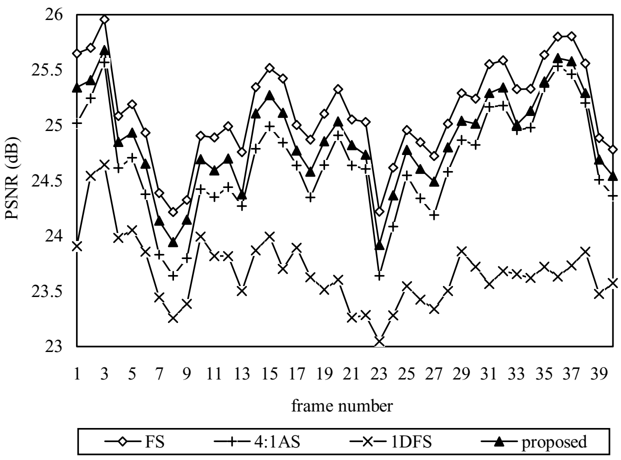

Fig. 4. PSNR for the "football" sequences.

The computational cost and the simulation results of the full search (FS), the one-dimensional full search (1DFS), the 4:1 alternate subsampling (4:1AS), the low-resolution quantization motion estimation algorithm (LRQME), and the proposed algorithm are summarized in Table I and Fig. 4. The results show that the PSNR degradation of the proposed algorithm with respect to the full search algorithm is less than 0.24dB with 1/38.1 computational cost. Furthermore, the PSNR of the proposed algorithm is superior to those of 4:1AS and 1DFS, with less computational cost.

IV. CONCLUSION

In this paper, a new motion estimation algorithm using low-resolution quantized image is proposed. Two new block-matching criteria called the different pixel count and the pseudo-SAD, are also proposed. To improve the performance, the proposed algorithm employs two search steps: the low-resolution search that generates candidates for search window and candidates for motion vectors, and the full-resolution search that finds the best motion vector in the candidates from the low-resolution search. The PSNR difference between the proposed algorithm and the full search is less than 0.24 dB, while the computational cost is reduced to 1/38.1. The PSNR of the proposed algorithm is superior to those of 4:1 alternate subsampling algorithm and the one-dimensional full search algorithm with less computational cost.

REFERENCES

[1] J. R. Jain and A. K. Jain, "Displacement measurement and its application in interframe image coding," *IEEE Trans. Commun.*, vol. COM-29, no. 12, pp. 1799-1808, Dec. 1981.

[2] M. J. Chen, L. G. Chen, and T. D. Chiueh, "One-dimensional full search motion estimation algorithm for video coding," *IEEE Trans. Circuits Syst.*, vol. 4, no. 5, pp. 504-509, Oct. 1994.

[3] T. Koga, "Motion compensated interframe coding for video-conferencing," *Proc. Nat. Telecommun. Conf.*, pp. G5.3.1-G5.3.5, 1981.

[4] B. Liu and A. Zaccarin, "New fast algorithms for the estimation of block motion vectors," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 3, no. 2, pp. 148-157, Apr. 1993.

[5] S. Lee and S.-I. Chae, "Motion estimation algorithm exploiting low-resolution quantization," *IEE Electronic Letters*, vol. 32, no. 7, pp. 647-648, Mar. 1996.

[6] N. Jayant and P. Noll, "Digital coding of waveforms," *Prentice-Hall*, pp. 131, 1984.