2004 IEEE International Conference on Multimedia and Expo (ICME)

Fast Priority Search Algorithm for Block Motion Estimation

Yongfang Liang[1], Ishfaq Ahmad[1], Viswanathan Swaminathan[2]

[1]*Department of Computer Science and Engineering, University of Texas at Arlington, TX, 76019*
{yliang, iahmad}@cse.uta.edu

[2]*Sun Microsystems Inc.*
Mountain View, CA 94043

Abstract

This paper proposes a median-bias fast priority search algorithm for motion estimation. The main characteristic of the proposed algorithm is that it adapts a priority median-bias search pattern enabling the identification of locations with higher probability of finding the motion vector. Moreover, based on the history of motion vectors, the algorithm includes a robust region detection technique to avoid unnecessary searches in the static background region. A minimum-distortion based analytical threshold is proposed to stop the search quickly when a "good enough" motion vector is achieved. Experimental results show that by using the proposed algorithm, the speed and accuracy of motion estimation are improved compared with the conventional approaches.

1. Introduction

Motion estimation (ME) is an essential part of video compression. Although the exhaust search can provide accurate motion vector (MV) with the minimum distortion, it is computationally intensive. Fast ME to reduce the computational complexity has been an important research issue in literatures. Research works can be roughly classified into fast matching technique (FMT) and fast search technique (FST).

FMT aims to reduce the complexity of the SAD computation, such as the successive elimination algorithm (SEA) [5], global elimination algorithm (GEA) [2] and partial SAD (pSAD) computation [3]. FST relies on searching fewer points within the search window by imposing a search pattern to the search, such as the four-step search (FSS) [8], diamond search (DS) [11] and the block-based gradient descent search (BBGDS) [6]. Recently, motion vector prediction (MVP) and fast search termination have been proposed, such as [9][10]. While the gains of MVP and fine search patterns are well documented, less attention has been given to the search order, which can potentially take advantage of FMTs. Meanwhile, the thresholds used in fast search termination to stop the search are determined empirically, which may be inaccurate and result in degraded video quality.

In this paper, relying on the statistical distribution of the motion vectors, we propose a novel median-bias fast priority search algorithm (FMPSA) for motion estimation in order to achieve high speedup and video quality. Unlike conventional approaches, we focus on the FST approach while taking the FMT into consideration. In FMPSA, the locations with higher probability of finding the MV are checked with higher priorities. Moreover, the robust static region detection technique and the minimum-distortion based analytical threshold applied in our algorithm are capable to avoid the unnecessary SAD computations and stop the search quickly.

The paper is organized as follows: Section 2 presents the proposed fast priority search ME algorithm, including the priority search pattern, the robust static region detection, and the fast termination technique. Simulation results for various sequences are given in Section 3. Finally, concluding remarks are provided in Section 4.

2. Fast median-bias priority search algorithm

2.1. Priority median-bias search

The MV of a block is highly correlated to the MVs of the adjacent blocks. This implies that the MV of the current block can be predicted from those of the spatial neighboring blocks. A good MV predictor is usually close to the optimal MV, thus we can get to the optimal MV faster. Median MVP is pointed out to be promising [9]. We will discuss and analyze the median MVP in a detailed way and propose our approach.

Let D represent the prediction error between the median predictor and the optimal MV, given by:

$$D = (mv_{median\_x} - mv_x,\ mv_{median\_y} - mv_y) \qquad (1)$$

where $(mv_{median\_x}, mv_{median\_y})$, (mv_x, mv_y) denote the median predictor and the optimal MV respectively. Experiments are carried out on typical video sequences to obtain the distributions of D. Here full search is used to obtain the MV with 16-pixel-size search region. From our results, we have:

Observation 1: The average probability of $|D|\leq 1$ is 89.27% for the tested sequences.

Table 1 gives the detailed probability distribution when $|D|\leq 1$. The first column denotes the prediction error. The i^{th} column lists the corresponding average probability for different video sequences.

Table 1: Prediction error distribution for D≤1

D	akiyo	salesman	stefan	garden	foreman	football	news	Ave. Probability
(0, 0)	0.8408	0.8516	0.4035	0.3694	0.4269	0.2720	0.8304	0.5706
(-1, 0)	0.0463	0.0610	0.2086	0.2314	0.1737	0.1837	0.0549	0.1371
(0,-1)	0.0842	0.0422	0.1314	0.1608	0.1442	0.1159	0.0523	0.1044
(1,0)	0.0016	0.0034	0.0080	0.0059	0.0111	0.0104	0.0034	0.0063
(0,1)	0.0022	0.0026	0.0049	0.0037	0.0100	0.0100	0.0029	0.0052
(1,-1)	0.0007	0.0006	0.0034	0.0049	0.0050	0.0060	0.0012	0.0031
(-1,1)	0.0007	0.0017	0.0035	0.0038	0.0050	0.0061	0.0017	0.0032
(-1,-1)	0.0163	0.0155	0.0790	0.1052	0.1103	0.0925	0.0146	0.0619
(1,1)	0.0001	0.0001	0.0008	0.0005	0.0024	0.0020	0.0003	0.0009

Observation 2: The prediction error distributes at (0, 0) with the highest probability. Meanwhile, the positions of D=(-1, 0) and D=(0, -1) also have high probability, while the probability is low for the other positions.

Observation 1 implies that by searching the small area around the median MV predictor, the optimal MV can be found with average probability 89.27%. It justifies the effectiveness of MVP: near the median predictor, we have high probability to get the optimal motion vector. Within the small area around the median predictor, we can assume the SAD surface is unimodal, thus gradient-based search pattern, such as the small diamond pattern used in [11], can be employed. According to Observation 2, different positions of the small area around the median predictor have different probabilities to be the motion vector. Observation 2 suggests that, to fully take advantage of FMT, such as the pSAD computation technique, to avoid unnecessary SAD computation, the position with higher probability should be checked with higher priority in order to stop the SAD computation earlier.

Based on the above observations, the small diamond search pattern with search priority is adopted in our algorithm. Fig. 1 illustrates the search locations and the search priorities of the search pattern. The bigger size of the point, the higher search priority this position has. At the first stage, (0,0) is checked (the biggest point), then the positions corresponding to $D=(-1,0)$ and $D=(0,-1)$ are checked secondly, followed by the other two points (Fig. 1(a)). The search pattern keeps on moving towards the direction of the current minimum (Fig. 1(b) or (c)), until the minimum SAD is found in the center or it hits the boundary of the search window. The positions closer to the median predictor are checked with higher priorities (the bigger squares).

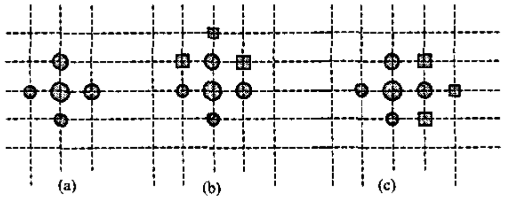

Figure 1: Priority search pattern.

2.2. Robust static region detection

From Table 1, we can also observe that for low motion sequences with static region, such as "Akiyo", "News", and "Salesman", the prediction error is 0 with very high probability. It implies that the MV is the median predictor with very high probability within the static region of the scene. Thus, if we can detect the static region, the search process can be stopped immediately without checking other positions.

We adopt motion history matrix (MHM) technique developed in our previous work [4] to perform the static region detection. The basic idea behind MHM is to keep track of the motion history of each block. The entry value of MHM shows how long this block has been without motion activity. The bigger the value, the probability is higher that it will still have no motion activity in current frame. From the value of MHM, we can predict whether the current block is a static block or not. The value of the prediction threshold is determined empirically.

With MHM, we can predict the static region blocks with high probability. However, the prediction may be incorrect due to the noise in the video sequence, which can arise from many sources, such as sensor problems in a camera, illumination variations. As a result, verification of the static block detection is necessary. In our work, the Gaussian distribution model is used to approximate the SAD distribution for a static block. Let MM and VV be the mean and variance of the SAD respectively. The confidence interval (CI) of the detection is given by:

$$CI = (MM - 2VV,\ MM + 2VV), \qquad (2)$$

For a Gussian distribution, the variable will fall within CI with probability 96%. The verification procedure is as following: when a block is predicted to be a static block, if the SAD value at (0,0) point is within CI, then it is a static block and (0,0) is chosen as the MV. Otherwise the other positions need to be checked.

2.3. Fast search termination

In a block based video encoder, after motion compensation, the residual error is first transformed into the frequency domain. After that, the transform coefficients are quantized by a predetermined quantization factor Q for further entropy compression. In SAD-optimal ME, the objective is to find the MV with the minimum SAD value. When the SAD is small enough, all the quantized transform coefficients become zeros, even a smaller SAD can be achieved, there is no difference in the dequantized transform coefficients. As a consequence, the reconstructed video quality cannot be improved and the search can be terminated immediately without loss of video quality.

The distribution of the residual error can be modeled by a Laplacian distribution with zero mean and a separable covariance [7]. Let R be the correlation coefficient matrix of the residual error of a $M \times N$ block, given by:

$$R = \begin{bmatrix} 1 & \rho & \rho^2 & & \rho^M \\ \rho & 1 & \rho & \cdots & \\ \rho^2 & \rho & 1 & & \\ & \vdots & & \ddots & \\ \rho^N & & & & 1 \end{bmatrix} \tag{3}$$

where ρ ($|\rho| < 1$) is the correlation coefficient. Typical values of ρ range from 0.6 to 0.75. Let B_{size} be the block size. After some mathematical deductions, we can estimate the probability that a specific coefficient $r(u,v)$ after quantization becomes zero by:

$$SAD < T(u,v) = \frac{B_{size} \times Q}{\sqrt{2} * n * \sqrt{[ARA^T]_{u,u}[ARA^T]_{v,v}}} \tag{4}$$

where A is the transform matrix and n is a confidence parameter. When Eq. (4) is satisfied, the probability that the quantized $r(u,v)$ becomes zero is very high. For n=1, n=2, n=3, the probability is 68%, 94%, and 99% respectively.

Specifically, the quantized DC coefficient will become zero with high probability when Eq. (5) is true.

$$SAD < T(0,0) \tag{5}$$

Since the DC coefficient dominates the transform coefficients, when the quantized DC coefficient becomes zero, we can assume all the coefficients will become zero after quantization. This assumption will not influence the reconstructed video quality apparently, because the human eyes are more sensitive to the lower frequency changes. This suggests that Eq. (5) can be used as the fast search termination criterion.

Based on information theory, when the residual error energy is small enough, there is no major difference in the reconstructed signal, which means a "good enough" reconstructed video quality is achieved, no further improvement can be obtained. Since the residual error energy can be estimated from the SAD value, it justifies that using Eq. (5) as the fast ME termination decision of a "good enough" block matching.

To evaluate the performance of the above criterion on different block sizes, we implement our proposed algorithm into the H.264/AVC video encoder [1], in which a 4x4 purely integer DCT is used and A is given by:

$$A = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 1 & -1 & -2 \\ 1 & -1 & -1 & 1 \\ 1 & -2 & 2 & -1 \end{bmatrix} \tag{6}$$

Derived from Eqs. (3)~(6), we have:

$$[ARA^T]_{u,u}[ARA^T]_{v,v} = \begin{bmatrix} 134.4150 & 76.8086 & 12.6807 & 19.2021 \\ 76.8086 & 43.8906 & 7.2461 & 10.9727 \\ 12.6807 & 7.2461 & 1.1963 & 1.8115 \\ 19.2021 & 10.9727 & 1.8115 & 2.7432 \end{bmatrix} \tag{7}$$

Hence, we have the following ME fast termination threshold for the H.264/AVC:

$$T_l = k \times B_{size} \times Q \tag{8}$$

where k is a constant.

2.4. The Proposed FMPSA

By jointly applying the above techniques, our novel FMPSA is summarized as follows. The MHM is maintained to keep track of the motion history of each block. If the corresponding MHM entry value of current block is bigger than a predetermined threshold value, the Gaussian distribution verification procedure is performed. If current block belongs to a static region, the search is stopped and the MV is (0,0). Otherwise, the media-bias priority search pattern is adopted. During the search, Eq. (5) is applied to check whether we have got a "good enough" MV in order to stop the search earlier.

In FMPSA, the static region can be detected effectively, thus the search can be quickly stopped in the early stage of the search. Moreover, unlike conventional algorithms that use a fixed search pattern without search priority, our algorithm adopts a median-bias priority search pattern. The locations with higher probability of finding the optimal MV are checked with higher priorities, thus FMTs such as pSAD computation technique can be fully taken advantage. Meanwhile, the minimum-distortion based threshold can avoid unnecessary searches without loss of video quality.

Note that in a block based video encoder, only the difference between the MV and the median MV predictor is coded into the bit stream. In FMPSA, because the positions close to the median MV predictor is always checked first, a smaller prediction error can be achieved and less bits are required to do the entropy coding In other words, the coding efficiency is improved.

3. Simulation results

In this section, computer simulations are carried out to evaluate the performance of our proposed algorithm. Conventional ME algorithms including FS, DS, FSS, and BBGDS are used in the comparison. The computer simulations are taken on typical benchmark video sequences, ranging from low motion talking head sequences to high motion sport sequences. The speedup performance and the video quality are measured by the number of checked points (CPs) and peak signal to noise ration (PSNR) respectively.

Table 2 and Table 3 summarize the comparisons. The results illustrate that for the tested sequences, the proposed FMPSA gains a great speedup with ignorable loss of video quality compared with FS. In terms of speedup, our algorithm outperforms all the other algorithms. The average CPs of our algorithm is 6.77, only 44.3%, 34.7% and 56.1% of DS, FSS and BBGDS respectively. Our algorithm is robust and efficient over a wide range of video sequences.

Table 2: PSNR comparison

PSNR	Akiyo	Coastgurad	Tennis	Stefan	Garden	Football
FS	38. 04	34. 07	35. 69	34. 10	33. 22	33. 26
DS	38. 02	34. 07	35. 66	34. 13	33. 21	33. 29
FSS	38. 00	34. 07	35. 65	34. 13	33. 20	33. 26
BBGDS	38. 01	34. 08	35. 65	34. 13	33. 21	33. 28
FMPSA	38. 00	34. 07	35. 65	34. 13	33. 21	33. 30

Table 3: Speedup comparison

CPs	Akiyo	Coastgurad	Tennis	Stefan	Garden	Football
FS	1089. 00	1089. 00	1089. 00	1089. 00	1089. 00	1089. 00
DS	13. 15	14. 43	14. 46	16. 23	15. 22	18. 18
FSS	17. 25	18. 91	18. 62	20. 27	19. 96	21. 98
BBGDS	9. 25	11. 10	10. 88	13. 18	12. 18	15. 91
FMPSA	2. 25	6. 99	5. 21	7. 93	7. 34	10. 92

4. Concluding remarks

A novel fast median-bias priority search algorithm for block motion estimation is proposed. The proposed algorithm not only reduces the computational complexity but also retains good estimation accuracy compared with some conventional algorithms. This is obtained by jointly applying different techniques. First, based on the study and analysis on the distribution of the median motion vector prediction, a median-bias priority search pattern is adopted. The locations with higher probabilities of finding the optimal motion vector are searched with higher priorities. Meanwhile, a robust static region detection approach is proposed to avoid unnecessary search within the static region. Moreover, we derive a minimum-distortion analytical threshold to stop the search when the MV is "good enough". Experimental results show that the proposed algorithm is robust and efficient. Our study can be combined with other motion estimation techniques to further improve the performance.

5. References

[1] Joint Video Team (JVT) of ISO/IEC MPEG and ITU-T VCEG, "Joint Video Specification (ITU-T Rec. H.264 | ISO/IEC 14496-10 AVC) - Joint Committee Draft," document JVT-G050r1.doc, May. 2003.

[2] Y.W. Huang, S.Y. Chien, B.Y. Hsieh, and L.G. Chen, "An efficient and low power architecture design for motion estimation using global elimination algorithm," *Proc. of 2002 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2002)*, Orlando, Florida, USA, May 2002.

[3] K. Lengwehasatit and A. Ortega "Complexity-distortion tradeoffs in vector matching based on probabilistic partial distance techniques," *Data Compression Conference DCC'99*, 1999.

[4] Yongfang Liang, Ishfaq Ahmad, Jiancong Luo, Yu Sun, "Fast Motion Estimation Using Hierarchical Motion Intensity Structure," appear in *ICME 2004*, Taipei, Taiwan, June 2004.

[5] W. Li and E. Salari. "Successive elimination algorithm for motion estimation," *IEEE Transactions On Image Processing* 4(1), pp. 105 – 107, January 1995.

[6] Lurng-Kuo Liu and Ephraim Feig, "A Block-Based Gradient Descent Search Algorithm for Block Motion Estimation in Video Coding," *IEEE Trans. Circuits Syst. Video Tech.*, vol. 6, no. 4, pp. 419-422, Aug. 1996.

[7] I-Ming Pao; Ming-Ting Sun, "Modeling DCT coefficients for fast video encoding," *IEEE Trans. Circuits Syst. Video Tech.* Volume: 9 Issue: 4, Page(s): 608 -616, Jun 1999.

[8] L. Po and W. Ma, "A Novel Four-Step Search Algorithm for Fast Block Motion Estimation," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 6, no. 3, pp. 313-317, June 1996.

[9] A. M. Tourapis, O. C. Au, and M. L. Liou, "Predictive Motion Vector Field Adaptive Search Technique (PMVFAST) - Enhancing Block Based Motion Estimation," proceedings *of Visual Communications and Image Processing 2001 (VCIP-2001)*, San Jose, CA, January 2001.

[10] W. Zheng, I. Ahmad, and M. Liu, "Adaptive Motion Search with Elastic Diamond for Mpeg-4 Video Coding," in Detection and Tracking of motion, *ICIP2001*.

[11] S. Zhu and K.K. Ma, "A New Diamond Search Algorithm for Fast Block-Matching Motion Estimation," *IEEE Trans. Image Processing*, vol. 9, NO. 2, pp. 287-290, Feb. 2000