A Fast and Adaptive Search Algorithm Based on Rood Pattern and Gradient Descent

Mu-Long Lin, Qing-Ming Yi, Min Shi

Abstract

In order to achieve the real-time property for video coding, a fast and adaptive algorithm based on starting search point prediction and early-termination strategy is proposed. It analyzes center-bias property and spatial correlation property of motion vector field, and utilizes the respective characteristics of block based gradient descent search (BBGDS) and adaptive rood pattern search (ARPS) algorithm. The proposed algorithm adaptively chooses different searching strategies according to the type of the image, makes full use of the cross image motion vector distribution characteristics and optimizes the traditional ARPS algorithm. The experimental results show that the proposed algorithm is about 2.3~9.2 times faster than Diamond Search (DS), 1.2~4.0 times than ARPS. The algorithm can meet the real-time demand without reducing the image quality.

1. Introduction

Motion estimation is a key technology of video coding. It can effectively remove redundancy between frames among image sequences so as to achieve high compression ratio. However, its high computational complexity is one of the key factors of restricting real-time compression coding of digital image. How to improve the motion estimation efficiency has been hot spot in video coding research area. Due to simplicity and efficiency, block-matching algorithm (BMA) is widely used in various video coding standards, such as MEPG-1/2/4 and H.26x. The FS algorithm is simple with highest accuracy. However, it calculates and compares each point in the block, which occupies great cost of computations. As a result, FS makes motion estimation the main bottleneck in real-time video coding applications and it is hard for hardware realization. Therefore, the researchers put forward a variety of fast algorithms, such as Three-
Step Search (TSS) [1], DS [2], Hexagon-Based Search (HEXBS) [3], BBGDS [4], ARPS [5], etc. TSS has a large initial step, which may be easy to trap the search into the local minimum and mislead the search direction. DS and HEXBS are more efficient, but they use a fixed template all the time and do not distinguish the motion types and image characteristics, which results in a waste of calculating resources. ARPS has fast search speed, but it still only uses a single template. BBGDS uses eight search directions from the center, but it does not consider the motion correlation between adjacent blocks. Based on analyzing the existing algorithms, a fast search algorithm which takes advantage of BBGDS in strong direction search and high search accuracy, and utilizes the improved ARPS method is proposed.

2. Analysis of motion vector characteristic

In order to improve search speed and image quality in the motion estimation by designing more appropriate template and strategy for search, it is necessary to investigate video sequence distribution characteristics of motion vectors. Table I shows the image vector distribution probability statistics in the first 100 frames in the different natural video sequence with FS algorithm.

| Test sequence | Distribution probability of the MVs |
|---------------|-----------------------------------|
|               | (0,0) | (±1,0) | (0,±1) | (±2,0) | (0,±2) | others |
| akiyo         | 0.9894 | 0.0106  | 0 | 0 |
| news          | 0.9531 | 0.0378  | 0.0012 | 0.0079 |
| claire        | 0.9620 | 0.0311  | 0.0014 | 0.0055 |
| foreman       | 0.4965 | 0.2801  | 0.0236 | 0.1998 |
| mobile        | 0.9579 | 0.0389  | 0.0027 | 0 |
| flower        | 0.0587 | 0.4220  | 0.2341 | 0.2852 |
| Coastguard    | 0.2973 | 0.5826  | 0.0573 | 0.0628 |
| Missamerica   | 0.6542 | 0.2463  | 0.0179 | 0.0816 |

Statistical data from Table I show that most of the video sequences contain static micro block (MB) or quasi-static MB. Average distribution probability of the vectors, whose radius in 2 around the original point is higher than 80%. They are mainly in the horizontal and vertical direction, which indicates that MVs of video sequence have center-bias property.

Furthermore, video image is a motion sequence with high spatial and temporal correlations. These properties can be used to predict the MV of the current MB. Considering real-time applications, the proposed algorithm only uses spatial correlation. However, utilizing temporal correlations for motion estimation needs recording all the MVs in the previous frames, which will increase cost of the memory and data access time. It is not convenient for hardware implementation [6]. MVFAST [7] adopts the left, top, top-right adjacent blocks of the current MB to predict the current MV, while ARPS only adopts the left block to predict the current MV. The proposed algorithm considers comprehensive equilibrium between computations and motion vector distribution characteristics, and employs left block in horizontal and top block in vertical direction. It does not consider the top-right block because of the high spatial correlation of top block.

3. Adaptive Road Pattern Search

As previous section mentions, ARPS algorithm makes full use of spatial correlation and adopts the MV of left block to predict current MV. The entire algorithm includes adaptive coarse search and local accurate search.
In the coarse search, if the MB is leftmost, a fixed-size arm length of two pixels is taken as the search step size, the others adaptively choose the search step size according to the predicted MVp. The maximum absolute value of horizontal component or vertical component is taken as step size as follow.

\[
\text{stepSize} = \max(|MVp(x)|, |MVp(y)|)
\]  

(1)

Here, MVp is prediction MV. Its corresponding point is also taken as one of the search points. If this point is overlapping with one of the vertex points in the rood shape, it needs to search 5 points, otherwise 6 points, as shown in Fig. 1. After searching these 5 or 6 points, a minimum block distortion (MBD) point is obtained. The MBD point is taken as the center point to go into the local accurate search. ARPS uses a small rood pattern in accurate search (SRPS), as shown in Fig. 2.

4. The proposed search algorithm

4.1 Overview of the proposed algorithm

The proposed algorithm makes full use of the center-bias property in horizontal and vertical direction with spatial correlation property in motion vectors. And it employs left block in horizontal and top block in vertical direction to predict reference blocks. Moreover, in order to get more precise prediction MVs, we adopt the BBGDS algorithm for the topmost blocks and leftmost blocks, and adopt an improved ARPS algorithm for other blocks. The algorithm mainly uses several techniques as follow.

1) Early termination strategy. Set a threshold for the static block to terminate early.

2) Motion type classification search. Choose the appropriate search step-size adaptively according to the motion type.

3) Optimized ARPS algorithm. Make full use of the property of the image motion vector distribution to obtain the MV quickly.

4.2 Early termination strategy

In the process of the block-matching search, there are three common matching criterions. They are Mean of the Absolute Frame Difference (MAD), Mean of the Square Error (MSE) and Sum of Absolute Difference (SAD). Here, the SAD criterion is adopted. From the following formula, we can find that it does not require multiplication, and is relatively easy to be implemented in hardware.

\[
\text{SAD}(i, j) = \sum_{m=1}^{M} \sum_{n=1}^{N} |f_{k}(m, n) - f_{k+1}(m + i, n + j)|
\]  

(2)

Here, \((i, j)\) is displacement vector, \(f_{k}\) and \(f_{k+1}\) is the pixel value of the current frame and previous frame, \(M*N\) is the block size. If \(\text{SAD}(i_0, j_0)\) is minimal, the point \((i_0, j_0)\) is the optimal match point. As previously mentioned, distribution of the MVs in video sequence has center-bias property, and the
MVs highly concentrate near the center of the search window. Static MB or quasi-static MB occupies a large proportion, so we can set a threshold $T$ to judge static MB for early termination. Check zero vector $(0,0)$ before searching $MV$, and then compare $SAD(0,0)$ with $T$. If $SAD(0,0) < T$, it means that the current block is very similar to the corresponding block in the reference frame, then we can judge the current block as a static block, set $(0,0)$ as its $MV$, and terminate search directly. Otherwise, micro-block has relative motion, it needs further search to determine the $MV$. Statistical experiment shows that more than 98 percent of static blocks $SAD$ values are less than $512[8]$, so we take 512 as the threshold $T$. Further research shows that take $T=512$ and take $T=0$ to determine static block, the PSNR value from the final compensation image is almost the same. Compared with traditional search algorithms, early termination strategy can reduce much redundant in the search matching process, and significantly increase the speed of motion estimation.

4.3 Classification of motion type

There exists strong spatial correlation in video sequence. Based on this property, it is easy to perform the fast and accurate search after predicting the $MV$ of current block. The proposed algorithm adopts the maximum value of horizontal component or vertical component from the left block and top block around the current block to be the motion vector length (MVL), as shown in the formula follows.

$$MVL = \max(|MV_x|, |MV_y|, |MV'_x|, |MV'_y|)$$ (3)

$MV_x$ is motion vector of left block, and $MV_y$ is motion vector of top block. According to the MVL of current block, we classify the motion types for the image sequences and select different strategies in accordance with the classification shown as follow.

1) If MVL<2, this block is small motion type image, which can be judged as a static block, then go to SRPS directly.

2) If 2≤MVL≤4, the current block is medium motion type, then improved ARPS is adopted.

3) If MVL>4, it shows that the current block is a large motion image, and it needs to predict the starting search point, then begins searching from this point with SRPS.

The leftmost blocks and topmost blocks do not have left or top block at the same time, so they are not necessary to be classified as the motion types. However, they are processed specially with BBGDS.

4.4 Initial search point prediction

Most of the traditional BMAs start the search from point $(0,0)$, but for large motion types, their real MVs deviate far from the zero vector. Thus, it needs iterative searches to find the best approximation $MV$. Methods that confirm the starting search point mainly include the median method, the weighted average method and the MB matching Sum of Absolute Difference (SAD) comparison method. SAD comparison method compares all the SAD values of the prediction MBs, and takes the point with minimum SAD as the initial search point. Comparing with other methods, its prediction accuracy is the highest, and the predicted initial point must be a MV of the adjacent MBs. So it is applied in this proposed algorithm. Experimental statistics show that current block has the strongest correlation with its left, top and top-right adjacent block and weak correlation with other directions. The left and top blocks are adopted in this paper. For large motion sequence, taking the corresponding point with minimum SAD from the left and top block as the initial search point can effectively avoid redundant search.

4.5 Description of the proposed algorithm

The algorithm mainly uses two structures of block gradient descent search pattern and rood shape pattern. For the leftmost blocks and topmost blocks, BBGDS or SRPS is adopted. The left block is
adopted as reference in the topmost, if MVL< 2, then use SRPS directly, else use BBGDS. Take the top block of the current block in the leftmost as reference, if MVL< 2, then use SRPS directly, otherwise use BBGDS. This can not only make full use of the advantage of high precision search in BBGDS, but also avoid redundant search for a small motion.

**Fig. 3 Flow chart of the algorithm**

For other blocks, optimized ARPS algorithm is adopted. Traditional ARPS adopts a same step-size as the rood-shaped arm length in horizontal and vertical direction. The proposed algorithm pays more attentions to the MV distribution and takes the left and top block for prediction. It considers the actual motion in horizontal and vertical direction. Hence, take

\[
\text{stepSize}(x) = \max(|MV'_x(x)|, |MV'(x)|)
\]

as the rood-shaped arm length for the horizontal direction, and take

\[
\text{stepSize}(y) = \max(|MV'_y(y)|, |MV'(y)|)
\]

as the rood-shaped arm length for vertical direction. This is conducive to find the best matching point more quickly. Flow chart of the algorithm is shown in Fig. 3.

The detailed algorithm steps are introduced as follow.

**Step 1:** Early termination for static block. Calculate the value SAD(0,0) of zero vector (0,0). If SAD(0,0)<T, then the current MV is (0,0), go to step 7, or else go to step 2.

**Step 2:** Classification of motion type. For the leftmost blocks and topmost blocks, the MVL of each block is \(\max(|MV(x)|, |MV(y)|)\) or \(\max(|MV'_x(x)|, |MV'(x)|)\). If MVL<2, go to step 6, else go to step 3. For other blocks, \(\max(|MV(x)|, |MV'(x)|, |MV'_x(x)|, |MV'_y(y)|)\).

If MVL<2, go to step 6; if 2≤MVL≤4, go to step 5; else if MVL>4, go to step 4.

**Step 3:** BBGDS algorithm, find out the MBD point, then go to step 7.

**Step 4:** Initial search point prediction. Compare the SAD value of the left block to the top block for the current block, set the point with minimum SAD value as initial search point, then go to step 6.

**Step 5:** Optimized ARPS algorithm. Start search from the original point at the corresponding block in the reference frame, set

\[
\text{stepSize}(x) = \max(|MV'_x(x)|, |MV'(x)|)
\]
as rood-shaped arm length for horizontal direction, set

\[ \text{stepSize}(y) = \max (|MV_x(y)|, |MV_y(y)|) \]

as the rood-shaped arm length for vertical direction. Find out the MBD point, then go to step 7.

Step 6: SRPS algorithm. Set the point with minimum SAD value found in last step as the center in the small rood pattern, and check if the SAD value of each point is less T or not when searching. If SAD < T, go to step 7 directly, else, continue at step 6.

Step 7: Stop the algorithm.

5. Experimental simulation and analysis

In order to prove the availability of the proposed algorithm, we use the typical sequence akiyo, news, foreman, flower and Stefan for testing. News is the typical sequence full of spatial details with video playing in the intermediate position. The sequence foreman is selected as medium motion video image. The sequence Stefan has fast motion in global and complicated details. Furthermore, the sequence flower with rich details and scene translation is used. The first 50 frames are used in the experiment. Stefan and flower is CIF image, and others are QCIF. FS, TSS, DS, BBGDS and ARPS algorithms are adopted at the same experimental situations to compare with the proposed algorithm. Experiments are on the Matlab platform, and use macro block size of 16*16 pixels and 7 as the search parameter. Two indicators are used to compare the performance of each algorithm. One is the average searching points per MB, the other is the average PSNR of compensated image. Average searching points are used for comparison of the computational complexity, and the average PSNR is adopted for comparison of search precision. Experiment data in detail are shown in Table II and Table III.

| Seq. | FS | TSS | DS |
|------|----|-----|----|
| Akiyo | 184.56 | 21.48 | 11.43 |
| News | 184.56 | 21.48 | 11.45 |
| Foreman | 184.56 | 21.54 | 13.08 |
| Flower | 204.28 | 23.32 | 15.76 |
| Stefan | 204.28 | 23.39 | 16.03 |

| Seq. | BBGDS | ARPS | Proposed |
|------|-------|------|----------|
| Akiyo | 7.85 | 4.93 | 1.24 |
| News | 7.92 | 5.00 | 1.51 |
| Foreman | 10.2 | 6.68 | 4.85 |
| Flower | 12.52 | 8.52 | 5.86 |
| Stefan | 13.45 | 8.31 | 6.89 |

| Seq. | Akiyo | News | Foreman | Flower | Stefan |
|------|-------|------|---------|--------|-------|
| FS | 44.54 | 38.51 | 33.06 | 26.34 | 25.91 |
| TSS | 44.54 | 38.49 | 32.83 | 25.81 | 25.34 |
| DS | 44.54 | 38.49 | 32.84 | 26.32 | 25.13 |
| BBGDS | 44.54 | 38.49 | 32.96 | 26.33 | 24.88 |
| ARPS | 44.54 | 38.49 | 32.67 | 26.3 | 25.56 |
| Proposed | 44.54 | 38.49 | 32.67 | 26.3 | 25.45 |
Table II shows that in the sequence with slow motion like akiyo, the optimal MV can be found out to use 11.43 points on average with DS algorithm, 4.93 points with ARPS algorithm, but only 1.24 points with the algorithm proposed. Compared with the FS algorithm, this algorithm increases 148.8 times in search speed. It also obviously speeds up in the medium motion and large motion sequence. Moreover, the algorithm not only finds out the optimized MV very fast, but also keeps high search accuracy at the same time. Its average PSNR value of compensated image falls slightly compared with FS as shown in Table III.

6. Conclusion

In order to improve the search speed and the search accuracy of motion estimation in image sequence, the distribution characteristics of image motion vector and spatial correlation is analyzed firstly. Based on the gradient descent search method with high precision characteristic and fast search speed characteristic in ARPS algorithm, a novel algorithm called fast adaptive hybrid algorithm combination these two algorithms is proposed. For the image quality virtually unchanged premises, the search speed of the proposed algorithm is the fastest comparing with ARPS and other algorithms. Experiment results show that this algorithm meets the real-time demand, and it is quick, effective and easy for implementation.

7. Acknowledgment

This paper is supported by 2010 Guangdong Province Science and Technology Project(2010B010600024), 2010 Guangzhou Science and Technology Support Project (Video format transcoding algorithm research and AVS transcoder design) and 2010 Guangdong Province Combination of Production, Studying and Researching Key Plan (2010A090200049).

References

[1] T. Koga et al., “Motion-compensated interframe coding for video conferencing,” in National Telecommunication Conference, New Orleans, Nov. 1981, G5.3.1-5.3.5
[2] 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, Feb. 2000, pp. 287-290
[3] C. Zhu, X. Lin, and L. M. Chau, “Hexagon-Based Search Pattern for Fast Block Motion Estimation,” IEEE Trans. Circuits Syst. Video Technol., vol. 12, no. 5, May. 2002, pp. 349-355
[4] L. K. Liu, E. Feig, “A block-based gradient descent search algorithm for block motion estimation in video coding,” IEEE Trans. Circuits Syst. Video Technol., Aug. 1996, vol. 6, no. 4, pp. 419-422
[5] Y. Nie, K. K. Ma, “Adaptive roof pattern search for fast block-matching motion estimation,” IEEE Trans. Image Processing, vol. 11, no. 12, Dec. 2002, pp. 1442-1448
[6] Z. H. Lin, Y. B. Zou, “A Novel Fast Motion Estimation Algorithm Based on Starting Search Point Prediction,” IEEE Conf. Measuring Technol, 2009
[7] K. K. Ma and P. I. Hosur, “Performance Report of Motion Vector Field Adaptive Search Technique (MVFAST),” in ISO/IEC JTC1/SC29/WG11 MPEG99/m5851, Noordwijkerhout, NL, Mar’00.
[8] Hosur P. I, K. K. Ma, “Motion Vector Field Adaptive Fast Motion Estimation,” International Conf. Information, Communication and Signal Processing’99, Singapore city, Singapore, 1999:7-10