Fast Multi-Reference Frame Motion Estimation via Downhill Simplex Search

Chen-Kuo Chiang and Shang-Hong Lai

Department of Computer Science, National Tsing Hua University,
Hsinchu 300, Taiwan, R.O.C.
d948323@oz.nthu.edu.tw, lai@cs.nthu.edu.tw

ABSTRACT

Multi-reference frame motion estimation improves the accuracy of motion compensation in video compression, but it also dramatically increases computational complexity. Based on tracing motion vector trajectories, fast approximated motion estimation results can be obtained for multi-reference frames. In this paper, we extend the downhill simplex search to multiple reference frames and propose several enhanced schemes to improve its efficiency and accuracy. Experimental results show that the proposed algorithm outperforms several representative single-reference frame block matching methods.

1. INTRODUCTION

In multi-reference frame motion estimation (ME), the motion vector of one block can be predicted from many reference frames. Recently, a number of algorithms have been proposed to reduce the computational complexity. Center-biased frame selection [1] and recent-biased search [2] perform ME with predefined search patterns in 3D space. Su and Sun [3] used composed MVs to predict approximated results in multi-reference frames. In addition, a simplex minimization method [4] [5] is applied in each previous frame to form an initial simplex for searching the minimal solution of the block distortion function.

Recently, an improved downhill simplex search (DSS) algorithm [6] was proposed for single-frame motion estimation by regarding it as a function minimization problem in a finite 2D space. Experimental results show that fast block-based motion estimation can be achieved by using an efficient function minimization algorithm other than using a predefined search pattern.

In this paper, we extend the downhill simplex search to multiple reference frames and propose several improved schemes to improve the efficiency and accuracy of the motion estimation algorithm. The rest of this paper is organized as follow: Section 2 introduces the proposed downhill simplex search algorithm for multiple reference frame motion estimation. In section 3, we present some improved schemes in our algorithm, including a special rounding technique and an early-stop error function evaluation scheme. Section 4 gives experimental results and comparison of the proposed algorithm with other block matching algorithms. Finally, we conclude this paper in section 5.

2. ALGORITHM DESCRIPTION

2.1. Downhill Simplex Search

Downhill simplex search is a derivative-free multidimensional function minimization method. In the downhill simplex search, a collection of n + 1 points in n-dimensional space is called a simplex. In the iterative simplex update process, the point with the highest function value is iteratively replaced by a new point with a smaller function value until the stopping criterion is satisfied. In our previous work, MVs are searched in 2-D space by single reference frame. Therefore, three points of the simplex form a triangle for simplex minimization. In multi-frame motion estimation, four points are selected to form a tetrahedron to find motion vectors in 3-D space, as shown in Figure 1.

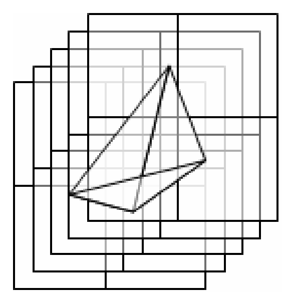

Figure 1. Multi-reference frame downhill simplex search using four points to form a tetrahedron.

2.2. Initial Simplex

Downhill simplex search can be roughly divided into two parts. In the first part, we need to initialize the simplex. It is important since we have better chance to find the correct solution very quickly when the solution is near the initial simplex or inside the simplex. A simple initialization method for downhill simplex search is to find three points around the center of the current block. This method works well for blocks with small motion vectors in 2D space. Motion vectors can also be predicted from the available motion vectors which we have estimated in the neighboring blocks at the current or previous frame. However, the overall situation including spatial and temporal displacement should be taken into consideration in multiple reference frames. And, it seems inefficient to apply downhill simplex search directly in each of the previous frames to find initial simplex.

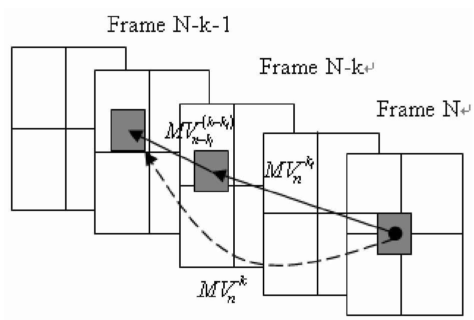

Figure 2. Tracing motion trajectories

In [3], the correlation between motion vectors of neighboring blocks can be exploited to find a better reference frame. By tracing along the motion trajectories, the motion vector in the previous frame can be composed by:

$$MV_n^{-k} = MV_n^{-k_l} + MV_{n-k_l}^{-(k-k_l)} \tag{1}$$

where MV_n^{-k} represents the motion vector of frame n referring to frame n-k shown in Figure 2. For example, MV_n^{-5} can be composed of $\mathrm{MV}_n^{-4} + \mathrm{MV}_{n-4}^{-1}$. Note that 95.16% of the composed motion vectors with MSE are below 1 pixel error. Thus, we can use the approximated motion vectors in reference frames to form the initial simplex. The way to form an initial simplex is as the following: In each frame, the most recent reference frame is first searched using downhill simplex search. Then, motion vectors in any other reference frames are composed from previous results. In the last step, four motion vectors among all candidates with the minimal block distortion values are then chosen to form an initial simplex.

2.3. Iterative Procedure

After the initial simplex is determined, the simplex is updated iteratively until the stopping criterion is satisfied. Finally, the point with the lowest function value in the simplex is the final solution. It consists of four steps: namely, reflection, expansion, contraction, and shrinkage, as shown in Figure 3. In the reflection step, we define a reflection point Pr as

$$P_r = P_{ave} + \alpha(P_{ave} - P_h)\ ,\ \alpha > 0 \tag{2}$$

$$\mathrm{P}_{ave} = \frac{1}{n+1}\sum_{i=1}^{n+1} P_i \tag{3}$$

The Block Distortion Measure (BDM) is calculated as the function value for each point. Yr is the BDM value of Pr and Pave is the average of all points of the simplex. In the expansion step, we define an expansion point Pe as

$$P_e = P_{ave} + \gamma(P_r - P_{ave}),\ \gamma \ge 1 \tag{4}$$

Ye is the BDM value of Pe. In the contraction step, we define a contraction point Pc as

$$P_c = P_{ave} + \beta(P_h - P_{ave}),\ 0 < \beta < 1 \tag{5}$$

Yc is the BDM value of Pc. In the shrinkage step, we define a contraction point P_i^{new} as

$$p_i^{new} = (p_i + p_l)/2\ , for\ i=1 \ldots n+1, i \neq l \tag{6}$$

Yl is the function value of the smallest point and Yh is the function value of the largest point. Ym refers to those points. Each iteration starts with the reflection step. The BDM values are calculated to see which point can be replaced with the point of the highest value in the simplex. Here, sum of square errors (SSE) is used as the BDM. In the shrinkage action, all other points are moved toward the lowest point. Figure 4. shows the flow chart of the iterative simplex update procedure.

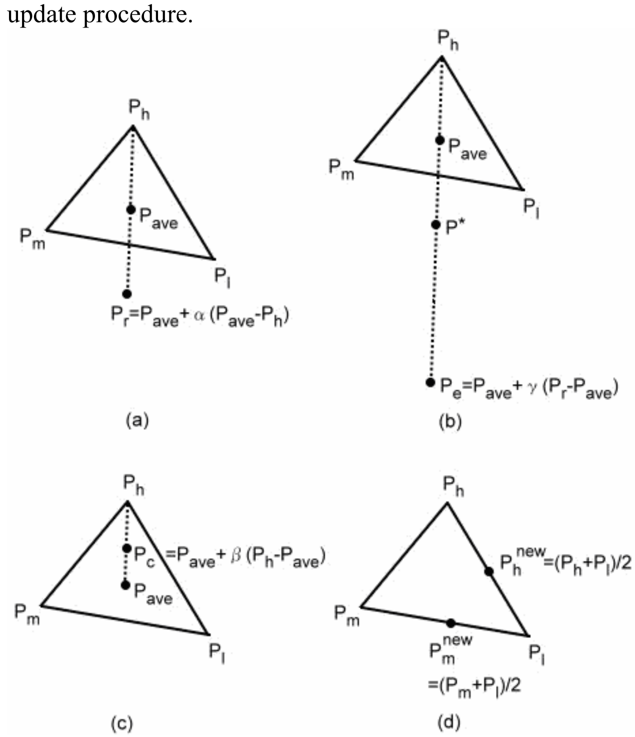

Figure 3. The four steps in the DSS iteration: (a) reflection, (b) expansion, (c) contraction, and (d) shrinkage

3. IMPROVED SCHEMES

In this work, we propose an improved downhill simplex search algorithm for multiple reference frame motion estimation. Throughout the iterative simplex update

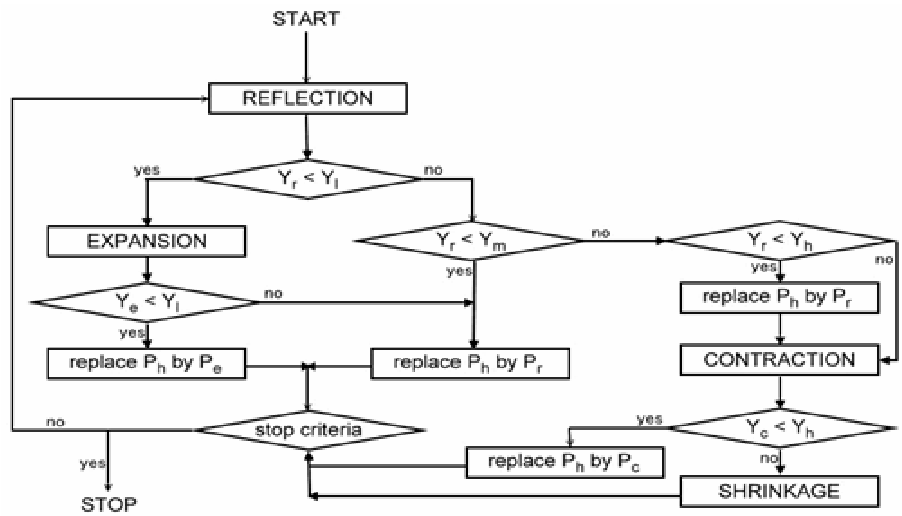

Figure 4. Flow chart of the downhill simplex search

process, several improved schemes are adopted to achieve better efficiency and compression quality.

3.1. The early-stop scheme

The early-stop scheme can be applied in the downhill simplex search. In the iterative procedure, the point with the highest SSE, namely Ph, is iteratively replaced with a better point. However, new points usually have a higher SSE than that of the current Ph. Therefore, the SSE computation can be terminated as soon as the accumulated SSE of the new location exceeds the SSE of Ph. The method can be used in the iteration steps or refinement. It helps reduce the computational load greatly. It is more significant in SSE than in the sum of absolute differences (SAD) error measure. Note that even the accumulation is aborted, the part of SSE that is already calculated is still considered as a fractional number in our search location numbers.

3.2. New location rounding scheme

The simplest rounding method is to round the search location to the nearest integer point, but this may degrade the accuracy. As shown in Figure 5, (x,y) is a point with

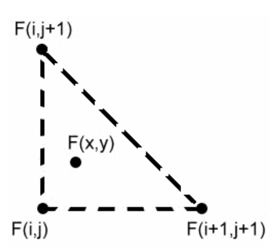

Figure 5. A new location rounding scheme.

fractional position. The coordinates (i,j), (i+1,j) and (i,j+1) are three neighboring integer points. Function F represents the SSE value of the point. At the fractional point, it requires interpolation to compare function values, taking shrinkage step for example, which takes much computational efforts for interpolation.

Instead of interpolation, we calculate F(i,j), F(i+1,j) and F(i,j+1) since the purpose of BMA is to find the most similar block and downhill simplex search uses a triangle to minimize function values in the same frame. The location with the smallest SSE value is then used to update point (x,y). In multiple reference frames, three neighboring integral points with the same temporal displacement are compared first. Then, the smallest is chosen from all candidates.

4. EXPERIMENTAL RESULTS

We compare six block matching algorithms, including full search (FS), four step search (FSS), diamond search (DS), simplex minimization search (SMS), our proposed downhill simplex search (DSS) and multi-reference full search (MR-FS) with our multi-reference frame downhill simplex search (MR-DSS) through experiments on four benchmarking video sequences (foreman, coastguard, news, and container). As shown in Table 1, QCIF sequences were chosen for testing. For different sequence length, we compute average number of search locations per frame. In each frame, motion is estimated and compensated using original frame in our proposed algorithm to avoid error propagation due to poor prediction results. Reference frame length is set to 5 frames which complies with H.264 coding standard, and $\alpha = 1, \beta = 0.5, \gamma = 2$ with the search range within ± 16 pixels in both horizontal and vertical directions. Our MR-DSS method performs a two-level downhill

Table 2. Simulation Results : PSNR(dB), Average Number of Search Locations/ Frame, Reference frame length = 5.

BMA	Foreman		Coastguard		News		Container	
	PSNR	Location	PSNR	Location	PSNR	Location	PSNR	Location
FS	32.21	82104.00	33.25	82104.00	37.64	82104.00	42.17	82104.00
FSS	31.73	1703.99	33.13	1507.57	37.61	1255.12	42.17	1219.58
DS	31.77	1595.29	33.17	1054.94	37.61	965.75	42.16	921.59
SMS	31.31	1106.33	32.53	1478.89	37.54	1045.95	42.13	1043.76
DSS	31.94	645.58	33.23	548.96	37.59	515.96	42.13	499.73
MR-FS	33.15	387195.00	33.61	387195.00	37.80	387195.00	42.42	387195.00
MR-DSS	32.48	1653.74	33.41	1317.28	37.65	1035.48	42.21	996.19

search. One is in the initial step. Each frame uses the single reference frame DSS to search the motion vector with the minimal distortion in the previous one frame. Then we compose MVs in any other reference frames as candidates of the initial simplex. Four minimal points are chosen to form an initial simplex. After that, a 3D-version DSS starts to get more accurate motion vectors. As in Table 2, our proposed MR-DSS provides significant reduction in computational complexity. Compared with MR-FS, a speed-up ratio is up to 350~390. Table 1 represents the speed-up ratio by *R = #Location in MRFS / #Location in MRDSS*. In fact, our proposed method outperforms many single reference block matching algorithms, mainly due to fast initial simplex composition. And the early-stop scheme further boosts the search speed. According to Table 2, MR-DSS is almost double the DSS search number. The result fits with the two-level method. On the other hand, MR-DSS provides higher PSNR than other methods within fewer search numbers and it achieves similar PSNR compared to MR-FS in some sequences. A special feature of the proposed method is the nearly constant average computational cost for videos of different motion types.

5. CONCLUSION

In this paper, we extend the previous single-frame downhill simplex search algorithm to multi-reference frame motion estimation. The main contribution of this work is that we proposed a very efficient multi-reference frame motion estimation algorithm with very stable computational efficiency for different types of video. According to the property of the downhill simplex search, we select tracing motion trajectories to form an initial simplex so that it reduces very high computational complexity in multiple reference frames. In addition, a special location rounding scheme and an early-stop strategy also help boost the search speed and improve the accuracy. Experimental results show the proposed multi-reference frame downhill simplex search algorithm in general provides faster motion estimation than many popular single-reference frame ME algorithms while maintaining good video compression quality.

Table 1. Computation reduction R in four sequence :
(R = #Location MRFS / #Location MRDSS)

Sequence	Resolution	Computation Reduction R
Foreman	QCIF, 320 frame	234.13
Coastguard	QCIF, 97 frame	293.94
News	QCIF, 200 frame	373.93
Container	QCIF, 180 frame	388.68

6. ACKNOWLEDGEMENT

This work was supported by MOEA research project under the grant 94-EC-17-A-01-S1-034.

7. REFERENCES

[1] C.-W. Ting, L.-M. Po and C.-H. Cheung, “Center-biased frame selection algorithms for fast multi-frame motion estimation in H.264,” IEEE Int. Conf. Neural Networks & Signal Processing, Nanjing. China, December 14-17, 2003.

[2] C. W. Ting, W. H. Lam, and L. M. Po, “Fast block-matching motion estimation by recent-biased search for multiple reference frames,” Proc. IEEE International Conference on Image Processing, Singapore , pp. 1445-1448, Oct. 2004.

[3] Y. Su and M.T. Sun, “Fast multiple reference frame motion estimation for H.264,” IEEE International Conference on Multimedia and Expo (ICME), 2004.

[4] M.E. Al-Mualla, C.N. Canagarajah, and D.R. Bull, “Simplex minimization for single- and multiple-reference motion estimation,” IEEE Trans. Circuits Syst. Video Techn, VOL. 11, NO. 12, pp. 1209-1220, Dec. 2001.

[5] M.E. Al-Mualla, C.N. Canagarajah and D. R. Bull, “Simplex minimization for multiple-reference motion estimation,” Proc. IEEE International Symposium on Circuits and Systems, Geneva, Switzerland, May 28-31, 2000.

[6] H.-C. Fei, C.-J. Chen, S.-H. Lai, “Enhanced downhill simplex search for fast video motion estimation,” Proc. Pacific Rim Conference on Multimedia, Jeju Island, Korea, pp. 512-523, 2005.