International Journal of Multimedia and Ubiquitous Engineering
Vol.10, No.8 (2015), pp.221-230
http://dx.doi.org/10.14257/ijmue.2015.10.8.23

An Optimized Motion Estimation Algorithm and Application in the FRUC System

Min-Jun Deng[1,2], Ping Gan[1],Zhuo Chen[1],Xiao-Qing Shen[1], Dong-Lian Li[1], Ming-Yan Yu[1],Yu Zhang[1],Cai -Lan Zeng[1] and He Huang[3]

[1]College of Communication Engineering, Chongqing University, Chongqing, China
[2]State Key Laboratory of Coal Mine Disaster Dynamics and Control, Chongqing University, Chongqing, China
[3]Chongqing Communications Research and Design Institute, Chongqing University, Chongqing, China
E-mail: vigor_gp@163.com

Abstract

Based on the optimized three-step search algorithm. Combining threshold judgment as well as local full search, a more efficient motion estimation algorithm is proposed. The algorithm not only inherited the traditional three-step algorithm's quick speed but also kept the advantages of a relatively small amount of calculation, besides it can avoid the local optimum problem in the three-step search algorithm (TSS). In addition, the algorithm combined with the threshold judgment and local full search algorithm, so it also maintains satisfactory visual quality. Comparing the algorithm with TSS and local full search algorithm (LFS). The algorithm has great performance in search points and peak signal-to-noise ratio. Experimental results show that compared with LFS, search points drop by 34.61% ~ 54.47% .While compared with the TSS, the search points only rise by 6.15% ~ 12.21%. The average PSNR of proposed Algorithm is 0.24dB higher than LFS and 3.30dB higher than TSS.

Keywords: *Motion Estimation, Motion Vector Prediction, Frame Rate up—Conversion*

1. Introduction

The development of large screen display device forward new requirements on image processing technology. Tailing jitter and fuzzy phenomenon would appear when giant TV screen is displaying fast dynamic images. In order to solve above problems, the concept of frame rate up is presented, reducing every frame images' displaying time on the screen to solve above problems, the typical way is inserting a frame image between two source images to realize frame rate up compensation(FRUC) [1].Motion estimation(ME) and motion compensation(MC) technology is the key steps to realize FRUC. High quality motion compensation (MC) is on the premise of fast and accurate motion estimation (ME). Currently, the motion estimation algorithm based on Bi-block-matching has been widely used in frame up system., which takes current block in the interpolated frame (FI) as the mirror center and searches for symmetrically matching blocks in the previous and next reference frames (FP and FN) respectively, as shown in Figure1[2].

ISSN: 1975-0080 IJMUE

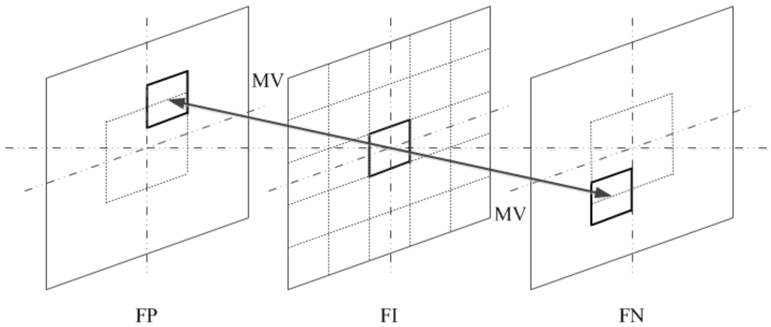


Figure 1. Bi-Directional Motion Estimation

The whole process of FRUC can be made up of the following steps:

Firstly, the interpolated frame (FI) are divided into non-overlapping and same size macro blocks .Secondly, Assuming all the pixels in one macro block have the same motion vector, bi-directional ME is performed at each interpolated block in turn to find optimal matching blocks in FP and FN. Thirdly, Making use of the motion vector derived from the second step, combining previous and next reference frames to finish the motion compensation. Finally, Repeat step two and step three, until all the macro blocks have been compensated.

Currently, there are many algorithms used in FRUC. Full search algorithm (FS) is the highest accuracy motion estimation algorithm in the frame rate up system, while FS costs a huge amount of the arithmetic operations, besides the search speed is slow, so it is difficult to meet the requirements of real-time processing of digital video. The traditional three-step search (TSS) takes the search speed and search accuracy into account ,but the traditional three-step search method is easy to fall into local optimum, so it is likely to cause motion estimation bias. The diamond search algorithm finely solved the problem of local optimum search, and the algorithm has high speed, but the search steps and search points can't be estimated. All of these restrict the algorithm's application in the FRUC system [3].

This paper proposed an algorithm that combines with the prediction of the MV files, improved three-step search and local full search

For the temporal correlation between video sequence, This paper take advantage of motion consistency between previous and following frames, optimal MVs of previous FI is used as initial motion vector of current FI .if the initial motion can meet the requirement ,the initial motion vector will be adopted in the current FI. Otherwise we will search each block's motion vector of the current FI. The search process will be executed from left to right and line by line. When we search the current block's motion vector, four adjacent macro blocks' motion vector can be used for the current block. During the search this paper assumes the current block has the same motion vector with its adjacent blocks .Combining with the SAD criterion determine whether the assumption is established. If the adjacent blocks' motion vector is suited for the current block, this paper will adopt the adjacent blocks' motion vector for the current block. Otherwise the improved three step search and local full search will be executed for the current block.

2. Motion Estimation Algorithm

2.1 Matching Criterion

This paper adopts the bidirectional motion estimation algorithm that based on the rule of SAD. Matching criterion essentially is a measure of the error, the value measured according to matching criterion is corresponding to the degree of similarity [4]. SAD and MSE criterions are widely used in frame rate up system .they are respectively defined as follows:

$$SAD(i,j)=\sum_{m=1}^{M}\sum_{n=1}^{N}\left|f_{r-1}(m-i,n-j)-f_{r+1}(m+i,n+j)\right| \quad (1)$$

$$MSE(i,j)=\frac{1}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N}\left[f_{r+1}(m-i,n-j)-f_{r-1}(m+i,n+j)\right]^2 \quad (2)$$

As shown above, where (i,j) is MV, $M\times N$ is block size, $f_{r-1}(m-u,n-v)$ and $f_{r+1}(m+u,n+v)$ denote pixel values of matching blocks in FP and FN depending on the motion vector (i,j) MSE matching result has higher accuracy, while MSE criteria need multiplication and division operation, the computational complexity is high and hardware cost is large, so this article adopts the SAD matching criteria.

2.2 Making Prediction

For the temporal correlation between video sequence, this paper takes the final MV field of previous FI as the initial MV field of current FI, just shown as the following Figure2. Then according to the SAD criterion get the SAD value of the current FI[5]. Then compare the SAD value with presented threshold .If the SAD value is smaller than the threshold value, the MV field of previous FI will be used in the current FI. Otherwise the current FI will be divided into non-overlapping and same size macro blocks, So this paper will search every block' s motion vector in the current FI.

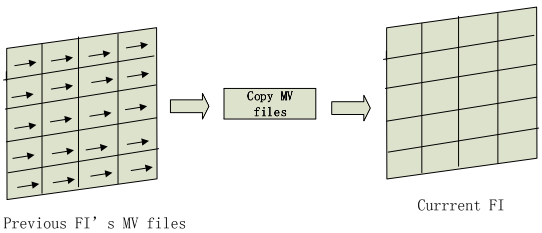


Figure 2. Copy Previous FI's MV Filed

As the moving target's size, the neighboring macro block's motion vector has strong correlation with the current block. Before searching the motion vector of the current block, this paper will make use of the adjacent block's motion vector as predictive vector [6]. The method of vector prediction is shown in Figure 3.

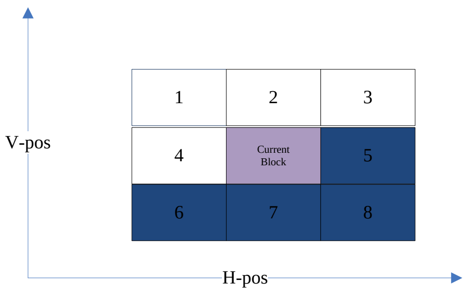


Figure 3. Motion Vector Prediction

There are eight blocks adjacent to the current block .When we search the motion vector of the current block, the motion vector of blocks (with number 1-4) had been calculated out. This paper Initialize the current block's motion vector with the blocks' (with number 1-4) motion vector . According to the SAD criterion, we will get the SAD value of the current block's initial motion vector .Comparing the SAD value with the threshold value, then judge whether the initial motion vector is suitable for the current block. Making full use of the adjacent blocks' motion vector not only greatly reduce the search time and search points, but also improve the smoothness of motion vector field.

2.3 Proposed ME Algorithm

Three step search (TSS) is very simple and robust. It searches for the best motion vectors in a course to fine search pattern .Initially we have to choose the step size, 8–norm is chosen as the initial step size. In the first step, eight blocks at a distance of step size from the centre one are picked for comparison. According to the SAD criterion, we get the SAD value of each blocks .The block with the smallest SAD value will be selected as the next centre points .In the second step, the step size will be halved ,and eight blocks at a distance of half step size from the new centre are picked up for comparison, get the new centre point. Repeat the above step, until the step size is 1-norm. The searching process is shown as the following Figure4.While the algorithm is vulnerable to the local optimum problem. For the step size of the first searching step is big, the first search perhaps can't get optimal matching point in some special case. So the subsequent searching will be end up with failure. This defect has limited the application of the algorithm [7].

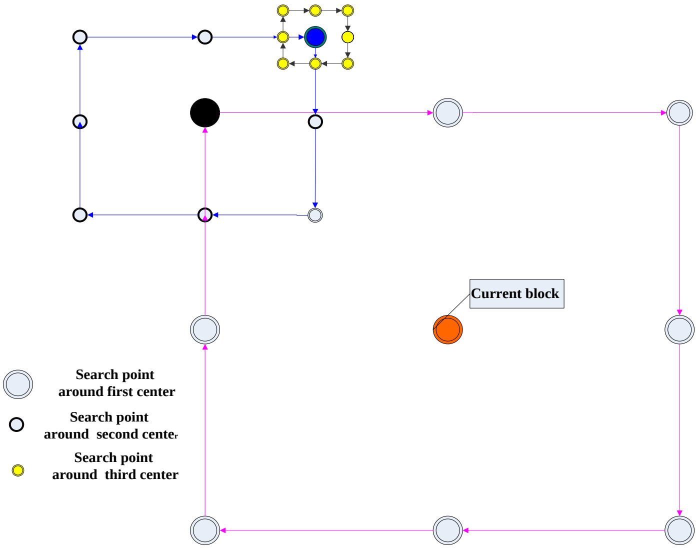


Figure 4. TSS searching Process

This paper adopts the improved three-step search (ITSS) method that proposed by reference [8]. The searching process of the improved three-step algorithm is shown in Figure5.

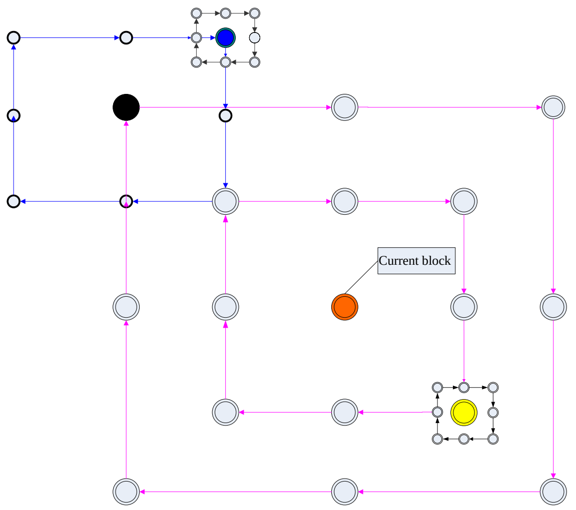


Figure 5. ITSS Searching Process

Compared with TSS, in the first step, eight blocks at a distance of half step size from the centre added to the comparison. According to the SAD values of the sixteen blocks, this paper will choose the block with smallest SAD value as the next centre .The following steps is same with the TSS. The improved three-step search method can effectively prevent the local optimum problem. However when the video sequence changes rapidly, the images compensated with the improved three-step search method is not ideal. After comprehensive analysis, because the improved three-step search method do not judge the final matching block's SAD value, some of blocks get incorrect motion vectors. According to the video sequences, this paper set corresponding threshold and compare matching block's SAD with the threshold value. If the matching block's SAD value is greater than the set threshold value. The current block's motion vectors will be considered as invalid, then the current block will be made local full search (LFS).

2.4 Steps of Algorithm

The algorithm this paper proposed is shown as follow:

Step1: Determine whether the current FI has the same vector files with the previous FI. If the SAD value of the current FI, the previous FI's MV files will be copied to the current FI. Otherwise the FI will be divided into non-overlapping and same size blocks, each blocks will be executed with step 2.

Step2: Initialize the current block's motion vector with the adjacent blocks' motion vector, judge whether the initial motion vector is suitable for the current block. According to the SAD value of each blocks, the current block will be determined executed step3, or ended the searching process in the step 2.

Step3: The block that will be searched with improved three-step search algorithm. If the matching block's SAD value is smaller than the setting threshold value, this paper will adopt the motion vector. Otherwise the current block will be into step4.

Step4: The block from step3 will be made search with local full search, then the current block get its motion vector.

3. The Algorithm Applied in the Frame Rate Up System

Frame rate up system this paper forwarded is shown in the Figure 6. Firstly, the original sequences were divided into many macro-blocks with same size and get each block's motion vectors. Secondly, smooth the motion vectors and improve the image sequences' quality. Finally, each macro-block will be divided into four small macro block with size of 8*8, and then every motion vectors will be refined into four motion vectors. Making motion compensation and get a new video sequence.

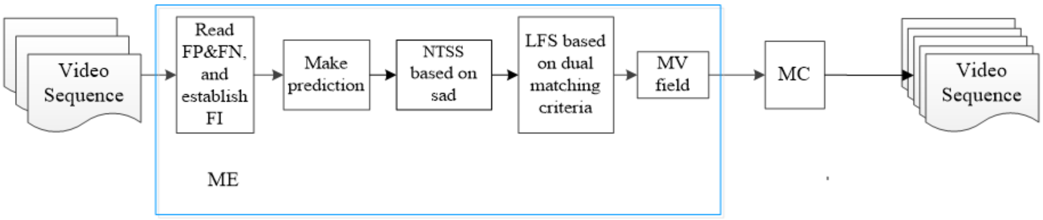


Figure 6. Frame Rate up System Structure

The frame rate up system this paper proposed is shown as the follow:

Step1: Two successive images will be extracted from the video sequences as reference frames, at the same time one blank image will be interpolated between reference frames.

Step2: The blank frame will be divided into several non-overlap macro-blocks with the same size.

Step3: Bi-directional motion estimation algorithm [9] is proposed in this paper, which takes current block in the interpolated frame (FI) as the mirror center and searches for symmetrically matching blocks in the previous and next reference frames (FP and FN) respectively ,then get each macro-blocks' motion vectors.

Step4: Motion compensator makes use of motion vectors obtained from Step3 and the reference frames to get the interpolated block.

The MC method proposed in this paper is weight-based adaptive overlapped block motion compensation (OBMC) algorithm [10]. This paper fully consider the correlation in interpolated block and its 8 neighboring blocks, the weighted sum of compensation values corresponding to MVs of 9 blocks is computed as interpolated pixel value, as shown in formula (3).

$$Fi(s)=\sum_{i=1}^{9}0.5\times w_i(s)\times\left(Fp\left(s-\overrightarrow{mv_i}\right)+Fn\left(s+\overrightarrow{mv_i}\right)\right) \tag{3}$$

In formula (3), $w_i(s)$ denotes the corresponding weighing coefficient of the macro-block whose pixel is S . mv_i denotes the macro-block's motion vector. Each pixel value in the interpolated block is derived by summating 9 MVs' compensation value which is consist of the block itself and its 8 neighboring blocks, besides each motion vectors' weigh is derived by the filtering window function.

4. Experimental Results and Discussion

In order to prove the feasibility of the proposed algorithm, objective criteria and subjective criteria are respective used to compare the proposed algorithm, local full search and the improved three-step search method. The objective criteria is consist of the average number of search points and PSNR. This paper adopts four kinds video sequence with different characteristics to verify the effectiveness of the proposed algorithm. They are foreman, mobile football and coastguard. The first 101 frames of sequences are selected, of which 50 even frames are taken as references of FIs and 51 odd frames form the test sequences are used to generate 50 FIs by FRUC system .In addition, the block size is set as 16*16 pixels, and search ranges of LFS, TSS and the proposed algorithm are all set as -7 +7.

4.1 NSP (Number of Search Points) and PSNR

Table 1. Comparison of Average NSPs per Block for Three Methods

Test sequence	*Average NSP per block*		
	LFS	*NTSS*	*Proposed method*
Foreman	*225.0*	*29.4*	*74.3*
Mobile	*225.0*	*28.6*	*65.7*
Football	*225.0*	*31.3*	*62.2*
Coastguard	*225.0*	*30.2*	*46.9*

Table 2. Comparison of Average PSNRs (dB) for Three Methods

Test sequence		LFS	NTSS	Proposed method
Foreman	PSNR	*34.1*	*32.2*	*34.7*
	Δ PSNR	*0.0*	*-1.9*	*+0.6*
Mobile	PSNR	*30.2*	*27.8*	*31.3*
	Δ PSNR	*0.0*	*-2.4*	*+1.1*
Football	PSNR	*23.5*	*22.4*	*24.3*
	Δ PSNR	*0.0*	*-1.1*	*+0.8*
Coastguard	PSNR	*26.0*	*23.3*	*26.3*
	Δ PSNR	*0.0*	*-2.7*	*+0.3*

Table 1 lists the average NSPs per block for three methods and four sequences. Table2 lists the average PSNRs of FIs generated by three methods, LFS as a reference standard. From Table1 we can see that the average NSP of LFS is the largest, the proposed algorithm in the middle and NTSS the smallest. This paper makes full use of each macro block's correlation. Making prediction of the current macro block's motion vectors, greatly reduced the search points. From the experimental data in table 1 and table 2, we can draw the conclusion that TSS has the lowest PSNR in spite of the smallest number of search points in the three methods. Average PSNR of LFS increases by 2.7dB than TSS, while LFS has the largest NSP. PSNR of FIs generated by the proposed algorithm is the highest, 0.7dB higher than LFS, which proves the effectiveness of the proposed algorithm.

4.2 Subjective Visual Quality

(a) TSS (b) LFS (c) Proposed Method

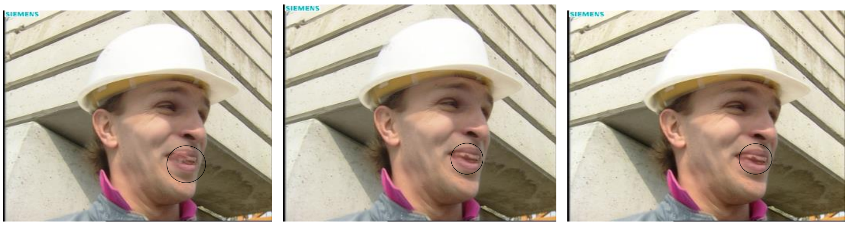
Figure 7. Comparison of FIs with Three Methods for 'Football'

(a) TSS (b) LFS (c) Proposed Method

Figure 8. Comparison of FIs with Three Methods for 'Coastguard'

(a) TSS (b) LFS (c) Proposed Method

Figure 9. Comparison of FIs with Three Methods for 'Mobil



(a) TSS (b) LFS (c) Proposed Method

Figure 10. Comparison of FIs with Three Methods for 'Foreman'

Subjective visual qualities of FIs generated by three methods are compared in Figure 7~ Figure 10 for sequences 'Football', 'Coastguard', 'Mobile' and 'Foreman'. Obviously, although some problems remain unsolved yet, the overall visual quality of proposed algorithm is the highest among three methods. Such as the number section in 'Mobile', body movement in 'Football' and human body in 'Coastguard' all have better visual effect with the proposed method [10].

5. Conclusion

This paper proposed an optimized motion estimation algorithm combined with threshold judgment, improved three-step search and local full search. The algorithm improved the image quality without significant increase in search points. The experimental results show that the interpolated video images generated by the proposed algorithm have higher PSNR and better subjective visual effect than LFS and TSS methods, applied in ME/MC-based FRUC system. The proposed algorithm is effective and simple, which also has certain practical value.

Acknowledgements

Project No.106112014CDJZR168801 supported by the Fundamental Research Funds for the Central Universities.

As time The work was supported by The Transportation Construction Projects of Ministry of Transport of the People's Republic of China (Grant No.2011318740240) &

The Visiting Scholar Foundation of State Key Laboratory of Coal Mine Disaster Dynamics and Control(Chongqing University) (Grant No.2011DA105287—FW201207).

References

[1] Yi W. C., Wen H. P., "Parametric OBMC for Pixel-Adaptive Temporal Prediction on Irregular Motion Sampling Grids," Circuits and Systems for Video Technology, IEEE Transactions on, doi: 10.1109/TCSVT.2011.215834, vol.22 no.1, January **(2012)**, pp.113,127.

[2] Zhou W. and Zhou X., "A fast hierarchical 1/4-pel fractional pixel motion estimation algorithm of H.264/AVC video coding," Industrial Electronics and Applications (ICIEA), 2013 8th IEEE Conference on, doi: 10.1109/ICIEA.2013.6566493, vol. no., 19-21 June **(2013)**, pp. 891-895.

[3] Liu Y., F. Tian and R. Liu, "A novel approach to frame rate up-conversion based on morphing," Journal of Computational Information Systems, vol. 7 no. 14, **(2011)**, pp. 5219-5226.

[4] Yuanzhouhan C., "Motion Compensated Frame Rate Up-conversion Using Soft-decision Motion Estimation and Adaptive-weighted Motion Compensated Interpolation," Journal of Computational Information Systems, vol. 9 no. 14, **(2013)**, pp. 5789-5797.

[5] Shuxiang G., Chenguang Q. and Xiufen Y., "A kind of global motion estimation algorithm based on feature matching," Mechatronics and Automation, 2009. ICMA 2009. International Conference on, doi: 10.1109/ICMA.2009.5246379, vol. no., 9-12 August. **(2009)**, pp. 107-111.

[6] Xuan J. and Chau L. P., "An efficient three-step search algorithm for block motion estimation," Multimedia, IEEE Transactions on, doi:10.1109/TMM.2004.827517, vol. 6 no. 3, June **(2004)**, pp. 435-438.

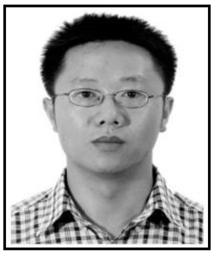
[7] Basher H. A., "Two Minimum Three Step Search algorithm for motion estimation of images from moving IR camera," Southeastcon, 2011 Proceedings of IEEE,doi:10.1109/SECON.2011.5752971, vol. no., 17-20 March **(2011)**, pp. 384-389.

[8] Pandit A. K., Verma S., Tomar G. S. and Kannoujia D., "Fast Motion Estimation Using Modified New Block Matching Algorithm: MTSSDS for Efficient Video Compression in Mobile Domain," Computational Intelligence and Communication Networks (CICN), 2011 International Conference on, doi: 10.1109/CICN.2011.77, vol. no., 7-9 October **(2011)**, pp.363-366.

[9] Tsung H. T., Hsueh Y. L., Hong G. C., "Overlapped block-based adaptive bilateral motion estimation," Circuits and Systems (APCCAS), 2010 IEEE Asia Pacific Conference on, doi: 10.1109/APCCAS.2010.5774980, vol. no., 6-9 December. **(2010)**, pp. 572-575.

[10] Ismail Y., Elgamel M. A. and Bayoumi M. A., "Fast Variable Padding Motion Estimation Using Smart Zero Motion Prejudgment Technique for Pixel and Frequency Domains," Circuits and Systems for Video Technology, IEEE Transactions on, doi: 10.1109/TCSVT.2009.2017417, vol. 19 no. 5, May **(2009)**, pp. 609-626.

Author



Min Jun, Deng, He was born in 1979, Sichuan province in China. He received his B.S. degree from Southwest University in 2001, The M.S. degree in Chongqing University of Post and Telecommunications in 2006. He is studying Ph.D. degree in Circuits and systems from Chongqing University. His research interests include Image Processing and mobile communication.