Sparsely Sampled Field for Fast Gradient-Based Global Motion Estimation

*Yong-Ren Huang
*Dept. of Information Engineering, Shu-Te University
Yen-Chau, Kaohsiung, 82445, Taiwan
E-Mail: yrhuang@mail.stu.edu.tw

Abstract

This paper presents a approach of sparsely sampled field for gradient-based global motion estimation (GME). For decreasing the computational complexity in gradient based GME, we propose an algorithm to obtain sparsely sampled field from the original image pixels firstly. Then, the low pass filter is employed for eliminating noise of original images. Finally, we propose a one-stage gradient based GME algorithm for global motion estimation. Simulation results show the comparisons of performance between our method and others.

Keywords: sparsely sampled field, gradient-based global motion estimation**.**

1. Introduction

The global motion due to camera motion had not been considered, until the very low bit rate coding standards, such as MPEG-4, were developed. Generally, the motion in a video sequence always contains local motion of objects and global motion of camera. If we can separate global motion and local motion, it will achieve more precise estimation. In addition, the camera motion is one of the important video descriptors for content-based multimedia indexing and retrieval. The camera motion parameters can be obtained by global motion estimation. Therefore, to develop an efficient global motion estimation technique is a very important issue.

Various studies [1-16] investigated the global motion compensation technique to improve the problem due to camera motion. All these approaches use parametric motion models [11] to characterize the global motions. The most popular method for parameters estimation of global motions is the least square method [4-6][14-17]. There are two main kinds of these schemes for solving motion parameters. One is based on the motion-based algorithm [4-6][16-17].The other is pixels-based process [14-15]. The motion-based approaches are based on block-matching estimation which contains both global motion and local motion. When the motion of background is complicated, such as waves of water and leaves of trees dances in the wind, will influence the parametric estimation. And the motion field estimation results the increase of computational complexity. The pixels-based approaches are very sensitive to the intensity of pixels. So, the SIGM approach [15], estimates parameters only using selected pixels and refines the required accuracy by projection of selected pixels for redundant extraction iteratively. Although without motion estimation, the projection processes in iterative refinement still increase the computational cost.

This paper will present a fast GME algorithm from sparsely sampled field which is composed of pixels with the top gradient magnitude in each sub-block of whole image. The pixels of sparsely sampled field can be employed for the estimation of global motion parameters by iterative refinement processes. For decreasing computational complexity, we will exclude the distribution of pixels out of image range, and omit the projection and comparison from sparsely sampled pixels.

The paper is organized as follows. The principle of global motion model will present in Section 2. In Section 3, we will propose our algorithm of sparsely sampled field and the process of one-stage gradient-based GME. Finally, the simulation results and discussions will present in Section 4.

2. Problem Formulation

Let two images $I_1(x, y)$ and $I_2(x, y)$ be related to each other by perspective transform as follows.

$$\begin{aligned} x_i^{(1)} &= (a_1 x_i^{(2)} + a_2 y_i^{(2)} + a_3)/(a_7 x_i^{(2)} + a_8 y_i^{(2)} + 1) \\ y_i^{(1)} &= (a_4 x_i^{(2)} + a_5 y_i^{(2)} + a_6)/(a_7 x_i^{(2)} + a_8 y_i^{(2)} + 1) \\ m &= (a_1, a_2, a_3, a_4, a_5, a_6, a_7, a_8) \end{aligned} \tag{1}$$

The least square criterion of gradient based method estimates the parameters m by minimizing the intensity discrepancies [15].

$$m = \arg\min_m \sum_{(x_i^1, y_i^1)\in S} (I_1(x_i^{(1)}, y_i^{(1)}) - I_2(x_i^{(2)}, y_i^{(2)}))^2 \tag{2}$$

where S is the set of pixels for minimizing.

The parameters m can be formed and solved

$$m^{t+1} = m^t + (H^T H)^{-1} H^T b \tag{3}$$

where

$$H_{i,j} = \frac{\partial I_2(x_i^{(2)}, y_i^{(2)})}{\partial x_i^{(2)}} \frac{\partial x_i^{(2)}}{\partial m_j} + \frac{\partial I_2(x_i^{(2)}, y_i^{(2)})}{\partial y_i^{(2)}} \frac{\partial y_i^{(2)}}{\partial m_j}, \tag{4}$$

$$b_i = I_1(x_i^{(1)}, y_i^{(1)}) - I_2(x_i^{(2)}, y_i^{(2)})$$

In [15], $\hat{b}_i$ is developed for interpolation free and solved m.

$$\hat{b}_i = I_1(\hat{x}_i^{(1)}, \hat{y}_i^{(1)}) - I_2(x_i^{(2)}, y_i^{(2)}) - \frac{\partial I_2(x_i^{(2)}, y_i^{(2)})}{\partial x_i^{(2)}}(\hat{x}^{(1)} - \lfloor x^{(1)} \rfloor) - \frac{\partial I_2(y_i^{(2)}, y_i^{(2)})}{\partial y_i^{(2)}}(\hat{y}^{(1)} - \lfloor y^{(1)} \rfloor) \quad (5)$$

where

$$\hat{x}^{(1)} = \begin{cases} \lfloor x^{(1)} \rfloor & \text{if } x^{(1)} - \lfloor x^{(1)} \rfloor <= 0.5 \\ \lfloor x^{(1)} \rfloor + 1 & \text{if } x^{(1)} - \lfloor x^{(1)} \rfloor > 0.5 \end{cases}$$
$$\hat{y}^{(1)} = \begin{cases} \lfloor y^{(1)} \rfloor & \text{if } y^{(1)} - \lfloor y^{(1)} \rfloor <= 0.5 \\ \lfloor y^{(1)} \rfloor + 1 & \text{if } y^{(1)} - \lfloor y^{(1)} \rfloor > 0.5 \end{cases} \quad (6)$$

In real implementation, 2D coordinate shifts are

$$\varepsilon_x = x^{(1)} - \hat{x}^{(1)}$$
$$\varepsilon_y = y^{(1)} - \hat{y}^{(1)} \quad (7)$$

Let

$$\hat{\varepsilon}_x = \hat{x}^{(1)} - \lfloor x^{(1)} \rfloor$$
$$\hat{\varepsilon}_y = \hat{y}^{(1)} - \lfloor y^{(1)} \rfloor \quad (8)$$

The error of 2D coordinate shifts is

$$\Delta\varepsilon_x = \varepsilon_x - \hat{\varepsilon}_x = x^{(1)} - \lfloor x^{(1)} \rfloor$$
$$\Delta\varepsilon_y = \varepsilon_y - \hat{\varepsilon}_y = x^{(1)} - \lfloor x^{(1)} \rfloor \quad (9)$$

The accuracy of parametric estimation will be influenced by $\Delta\varepsilon_x$ and $\Delta\varepsilon_y$. So, in this paper, we modify the coordinate shift to improve the performance of parametric estimation.

3. The Proposed GME Algorithm

3.1 Sparsely Sampled Field

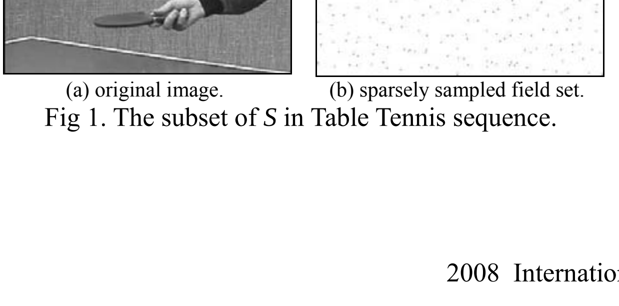
Firstly, the image size in 176×144, 352×288 or 704×480 will be divided into 8×8, 16×16 or 32×32 blocks, respectively. Let block set is $B=\{b_i\}$ where $i=1,2,..,n$. And let sparsely sampled field is composed of the pixels set $S=\{s_i\}$ where $i=1,2,..,n$.

$$s_i = \arg\max_{p_k \in b_i}(Gd(p_k)) \quad (10)$$

where

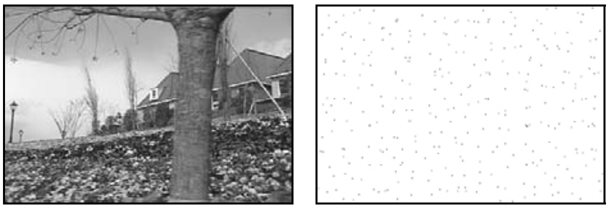
$$Gd(p_k) = |I_2(x_k, y_k) - I_2(x_k - 1, y_k)| + |I_2(x_k, y_k) - I_2(x_k, y_k - 1)| \quad (11)$$

The pixels with the highest gradient magnitude will be selected into subset S Fig 1-3 shows the sampling process.



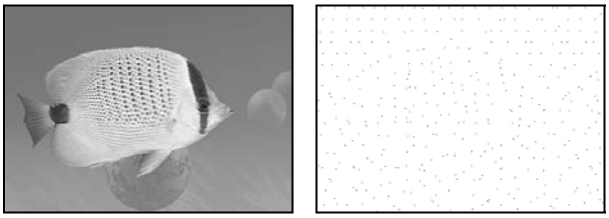
(a) original image. (b) sparsely sampled field set.

Fig 1. The subset of S in Table Tennis sequence.



(a) original image. (b) sparsely sampled field set.

Fig 2. The subset of S in Flower Garden sequence.



(a) original image. (b) sparsely sampled field set.

Fig 3. The subset of S in Bream sequence.

3.2 One Stage gradient-based GME

In our algorithm, we just estimate the parameters in original image size, not in pyramid structure. Fig 4 indicates the flowchart of one-stage gradient-based GME.

Firstly, We set the initial m=(1,0,0,0,1,0,0,0). And the input images I_1 and I_2 are processed by low pass filter.

$$\hat{I}_1 = Lowpass(I_1), \qquad \hat{I}_2 = Lowpass(I_2) \quad (12)$$

In eq.(3), $H^T H$ can be expressed by $H_{i,j}$ in eq.(4).

$$(H^T H)_{k,j} = \sum_{i \in S}\sum_{k=1}^{8}\sum_{j=1}^{8}\left(\frac{\partial \hat{I}_2(x_i^{(2)}, y_i^{(2)})}{\partial x_i^{(2)}}\frac{\partial x_i^{(2)}}{\partial m_k} + \frac{\partial I_2(x_i^{(2)}, y_i^{(2)})}{\partial y_i^{(2)}}\frac{\partial y_i^{(2)}}{\partial m_k}\right) \cdot \left(\frac{\partial \hat{I}_2(x_i^{(2)}, y_i^{(2)})}{\partial x_i^{(2)}}\frac{\partial x_i^{(2)}}{\partial m_j} + \frac{\partial \hat{I}_2(x_i^{(2)}, y_i^{(2)})}{\partial y_i^{(2)}}\frac{\partial y_i^{(2)}}{\partial m_j}\right)$$
$$= \sum_{i \in S}\sum_{k=1}^{8}\sum_{j=1}^{8} H_{i,k} \cdot H_{i,j} \quad (13)$$

For accuracy in parametric estimation, we modified Eq.(5).

$$\hat{b}_i = \hat{I}_1(\hat{x}_i^{(1)}, \hat{y}_i^{(1)}) - \hat{I}_2(x_i^{(2)}, y_i^{(2)}) - \frac{\partial \hat{I}_2(x_i^{(2)}, y_i^{(2)})}{\partial x_i^{(2)}}\varepsilon_x - \frac{\partial \hat{I}_2(y_i^{(2)}, y_i^{(2)})}{\partial y_i^{(2)}}\varepsilon_y \quad (14)$$

In our algorithm, we calculate H, $H^T H$ and $\hat{b}_i$ in the first loop and estimate parameters set m. If one of parameters is less than a predefined threshold, the estimation process stops. Otherwise, in the iterative refinement processes, if pixel of S falls out of the projection image range, we set $H_{i,j}$ to be zero and subtract the distribution of $H^T H$. As well as the pixel is extracted from S.

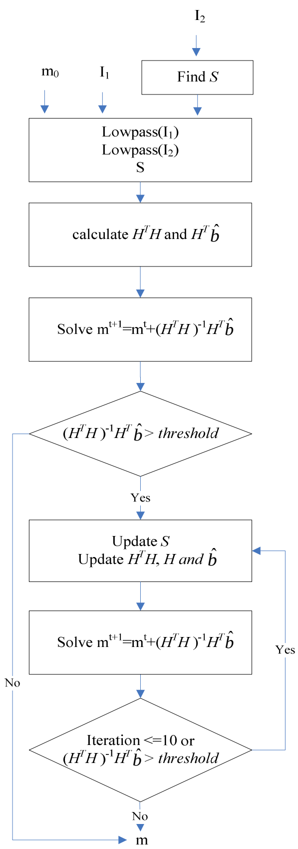


Fig 4. The flowchart of one-stage gradient-based GME.

4. Simulations and Discussions

In the simulation, the implementation circumstance is on PC with INTEL 2.66GHz CPU, memory size 512MB and MS Windows 2003 server. And the simulation software is designed by Boland Builder C++ version 6.0. The test video sequences are Foreman, Flower Garden, Table Tennis and Coast Guard. The predefined thresholds for convergence of parametric estimation are 0.01 for translation parameters a_3 and a_5 and 0.001 for others. The table 1 presents the size and mean square error (MSE) of various methods: Coarsely Sampled motion vector GM (CSGM), Recursive Least Square GM (RLS GM), Gradient Descent GM (GDGM), SIGM and proposed method. Table 2 shows the time cost of GME in various methods. The time cost calculation includes motion filed obtaining or pyramid iterative refinement for various methods.

Fig 5~8 demonstrate the comparison results of MSE by various GME methods and new-three-step MC in block size 8×8 for Foreman ,Flower Garden, Table Tennis and Coast Guard sequences in size 176×144, 352×240, 352×240 and 176×144, respectively. There are more complicated backgrounds in these sequences and the motion of background distributes disorderly. The results indicate that motion-based methods, CSGM and RLS GM, are influenced by obtained motion vector obviously. The pixel-based methods, GDGM, SIGM and our method, are more sensitive to the intensity of input data. Results indicate the influences by new appear objects obviously.

5. Conclusions

This paper presents a approach of sparsely sampled field for gradient-based global motion estimation. First, we sample the pixels of subset *S* by gradient magnitude in sub blocks. Then, one stage modified scheme is employed to estimate the global motion parameters. The proposed method needs not block-based motion estimation and reduces the computational cost of conventional SGIM by subset *S* of sparsely sampled field. Furthermore, its performance is very efficient for global motion estimation.

Acknowledge

This work is supported by the National Science Council under Grant NSC 97-2218-E-366 -002.

Reference

1. M. Hoetter. Differential Estimation of the Global Motion Parameters Zoom and Pan. Signal Processing: Image Communication 1998; 16:249-265.
2. Y. T. Tse and R. L. Baker. Global Zoom/Pan Estimation and Compensation for Video Compression. IEEE in Proc. ICASSP'91 1991; 2725-2728.
3. D. Wang and L. Wang. Global Motion Parameters Estimation Using a Fast and Robust Algorithm. IEEE Transactions on Circuits and Systems for Video Technology 1997; 7:823-826.
4. Z. Eisips and D. Malah. Global Motion Estimation for Image Sequence Coding Applications. Electrical and Electronics Engineers in ISRAEL 1991; 6-189.
5. C. T. Chu, D. Anastassiou, S. F. Chang. Hierarchical Global Motion Estimation/Compensation in Low Bitrate Video Coding. 1997 IEEE International Symposium on Circuits and Systems 1997; 1149-1152.
6. E. T. Kim and H. M. Kim. Fast and Robust Parameter Estimation Method for Global Motion Compensation in the Video Coder. IEEE Transactions on Consumer Electronics 1999; 45:76-83.
7. H. Jozawa, K. Kamikura, A. Sagata, H. Kotera and H. Watanabe. Two-Stage Motion Compensation Using Adaptive Global MC and Local Affine MC. IEEE Transactions on Circuits and Systems for video technology 1997; 7:75-84.

8. G. B. Rath and A. Makur. Iterative Least Squares and Compression Based Estimations for a Four-Parameter Linear Global Motion Model and Global Motion Compensation. IEEE Transactions on Circuits and Systems for Video Technology 1999; 9:1075-1099.
9. K. Y. Yoo, J. K. Kim. New Motion Estimation and Compensation Algorithms for Video Compression Combining Global and Local Motions. Signal Processing: Image Communication 1999; 15:201-216.
10. Smolić, T. Sikora and J. Rainer. Long-term Global Motion Estimation and its Application for Sprite Coding, Content Description, and Segmentation. IEEE Transaction on Circuits and Systems for Video Technology 1999; 9:1227-1242.
11. T. S. Huang and A. N. Netraval. Motion and Structure from Feature Correspondences: a Review. Proceedings of the IEEE 1994; 82:252-268.
12. W. G. Chen, C. G. and M. C. Lee. Repetitive and Morphological Padding for Object-based Video Coding. Image Processing 1997; 373 –376.
13. E.A. Edirisinghe, J. Jiang and C. Grecos. Shape Adaptive Padding for MPEG-4. IEEE Transactions on Consumer Electronics 2000; 46:514 –520.
14. F. Dufaux and J. Konrad, “Efficient, robust, and fast global motion estimation for video coding,” IEEE Transactions on Image Processing, 2000, 10(3): 497–501.
15. Y. Keller and A. Averbuch, Fast Gradient Methods Based on Global Motion Estimation for Video Compression, IEEE Transactions on Circuits And Systems for Video Technology, 2003, 13(4):300~309.
16. Y. Su, M. T. Sun, and V. Hsu, Global Motion Estimation from Coarsely Sampled Motion Vector Field and the Applications, IEEE Transactions on Circuits System for Video Technology, 2005, 15(2):232–242.
17. Y.R Huang, C.M. Kuo and C.L.Kuo, Efficient Global Motion Estimation Algorithm Using Recursive Least Squares, Optical Engineering 2006, 45(5): 057003-1~057003-13.

Table 1. The average of MSE in various GME methods.

Method / Sequences & size	CSGM	RLS_GM	GDGM	SIGM	Our method
Foreman 176×144 (frame 0~113)	84.1	29.8	61.8	24.6	**25.9**
Flower Garden 352×240 (frame 1~114)	238.15	197.8	234.86	277.4	**205.76**
Table Tennis 352×240 (frame 0~79)	143.4	130.0	135.4	132.9	**111.4**
Coast Guard 176×144 (frame 1~114)	43.34	41.42	56.28	38.23	**36.34**

Table 2. The average of time cost in various GME methods. (second/frame)

Method / Sequences & size	CSGM	RLS_GM	GDGM	SIGM	Our method
Foreman 176×144 (frame 0~113)	0.125	0.016	0.093	0.030	**0.002**
Flower Garden 352×240 (frame 1~114)	0.392	0.026	0.317	0.061	**0.004**
Table Tennis 352×240 (frame 0~79)	0.718	0.027	0.468	0.061	**0.003**
Coast Guard 176×144 (frame 1~114)	0.125	0.011	0.094	0.029	**0.003**

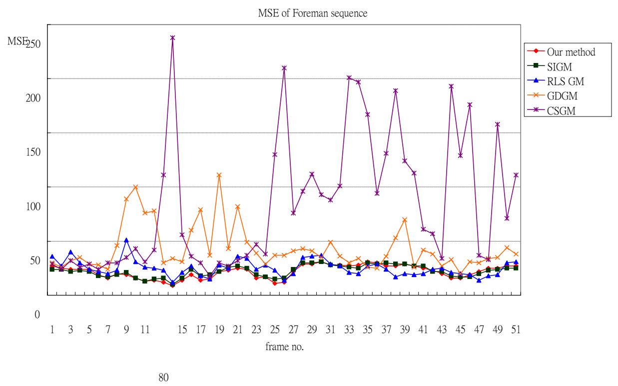


Fig 5. The mean square error of various algorithms (frame 1-50 of Foreman)

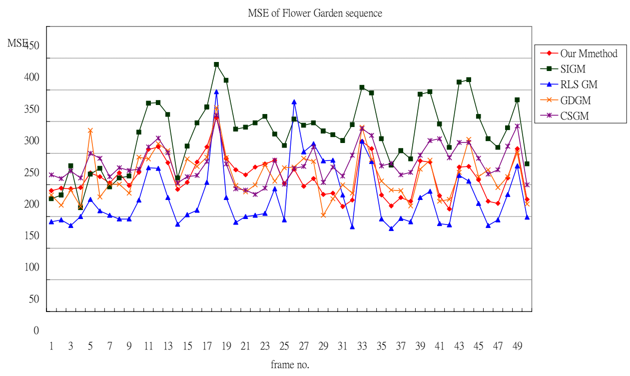


Fig 6. The mean square error of various algorithms (frame 1-50 of Flower Garden)

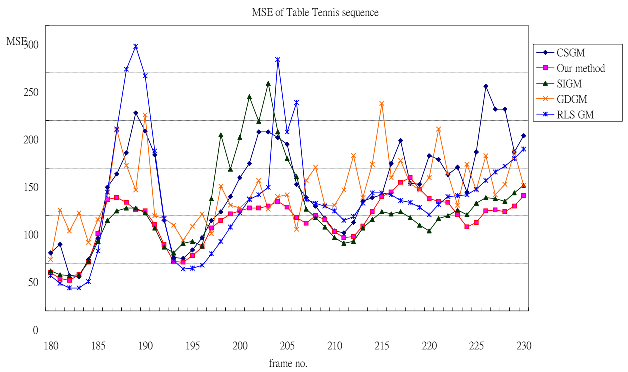


Fig 7. The mean square error of various algorithms (frame 180-230 of Table Tennis)

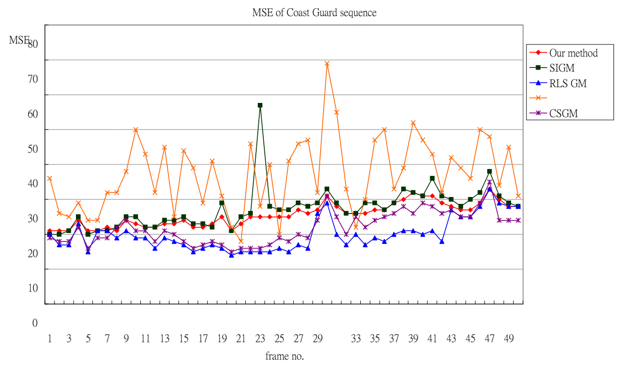


Fig 8. The mean square error of various algorithms (frame 1-50 of Coast Guard)