A Comparative Study on Gradient-Based Approaches for Optical Flow Estimation

Takashi FUSE*, Eihan SHIMIZU** and Morito TSUTSUMI***
* Department of Civil Engineering, University of Tokyo
** Professor, Department of Civil Engineering, University of Tokyo
*** Assistant Professor, Department of Civil Engineering, University of Tokyo
Hongo 7-3-1, Bunkyo-ku, Tokyo, 113-8656
E-mail: <fuse, shimizu, tsutsumi>@planner.t.u-tokyo.ac.jp
JAPAN

KEY WORDS: Optical Flow Analysis, Movement Detection, Object Tracking, Image Sequence, Measurement, Image Processing, Computer Vision.

ABSTRACT

The most readily available motion parameter from sequential image is optical flow. Among various optical flow estimation techniques, gradient-based approach is a common technique. This approach is based on the assumption that the brightness of a point in the image remains constant during a short time interval, while the location of that point in the image may change due to motion. This assumption leads to a single local constraint on the optical flow at a certain point in the image. It is, however, ill-posed as the constraint constitutes only one equation of two unknowns, that is, x-component and y-component of the flow vector. In order to solve this problem, various methods have been proposed. There are, however, only a few comparative studies from the viewpoint of the application to the specific and practical motion analysis. This paper reviews the gradient-based approaches theoretically and compares their performance empirically from the point of view of application to vehicle motion analysis. The basic methods of gradient-based approach are reviewed as follows: (1) Increase in the number of observation equations: (a) spatial local optimization method, (b) temporal local optimization method, (c) multispectral constraints method, (d) second order derivative method and by their combination; (2) Imposition of a condition: (a) spatial global optimization method, (b) temporal global optimization method and their combination. The result of empirical comparison shows the difficulty of estimation of precise and dense optical flow by ordinary gradient-based approaches, when sequential images are taken at an interval about 1/30 seconds. Hence, it is difficult to analyze vehicle motion by the gradient-based approaches in this case.

1. INTRODUCTION

Sequential image processing techniques have made progress for motion analysis under the improved performances of optical sensor and personal computers. Using the sequential image processing techniques, 3D reconstruction and structure from motion have been attempted, and such attempts will lead to computer vision or robot vision. For 3D reconstruction and structure from motion, stereo sequential images are employed. Stereo matching requires displacement vectors at feature points. In order to acquire the displacement vectors, the extraction techniques of the displacement vectors have been investigated. Optical flow is usually used for this purpose and gradient-based approach is a common technique for optical flow estimation. Optical flow is the distribution of apparent velocities of movements of brightness patterns in an image. And gradient-based approach is based on the assumption that the brightness of a point in the image remains constant during a short time interval, while the location of that point in the image may change due to motion. This assumption leads to a single local constraint on the optical flow at a certain point in the image (Horn and Schunck, 1981). Standard computer vision applications require precise and dense optical flow. In various techniques used for optical flow estimation, gradient-based approach is the most suitable to meet these requirements. It is, however, ill-posed as the constraint constitutes only one equation of two unknowns, that is, x-component and y-component of the flow vector. Further constraints are, therefore, necessary to solve for two unknowns. In order to solve this problem, various constraints have so far been proposed. It is said that these approaches have large influence upon the result.

On the other hand, there have been only a few comparative studies from the viewpoint of the application to the specific and practical motion analysis. This paper reviews the gradient-based approaches theoretically and compares their performance empirically from the point of view of application to vehicle motion analysis.

2. GRADIENT-BASED APPROACH

2.1 Gradient Constraint Equation

An equation that relates the change in image brightness at a point to the motion of the brightness pattern was derived by Horn and Schunck (1981). Let the image brightness at the point (x, y) in the image plane at time t be denoted by $I(x, y, t)$. Now consider what happens when the pattern moves. The brightness of a particular point in the pattern is constant, so that

$$\frac{dI}{dt} = 0. \tag{1}$$

The differentiation (1) can be expanded in a Taylor series

$$\frac{\partial I}{\partial x}\frac{dx}{dt} + \frac{\partial I}{\partial y}\frac{dy}{dt} + \frac{\partial I}{\partial t} = 0. \tag{2}$$

If we let

$$u = \frac{dx}{dt} \text{ and } v = \frac{dy}{dt}, \tag{3}$$

and then,

$$I_x u + I_y v + I_t = 0 \tag{4}$$

where we have also introduced the additional abbreviations I_x, I_y and I_t for the partial derivatives of image brightness with respect to x, y and t, respectively, that is

$$I_x = \frac{\partial I}{\partial x}, \quad I_y = \frac{\partial I}{\partial y}, \quad I_t = \frac{\partial I}{\partial t}. \tag{5}$$

This equation expresses a plane, which have normal vector $(u,v,1)$, and a measured point (I_x, I_y, I_t) is on the plane. Due to single linear equation in the two unknowns u and v, the parameter u and v, that is x-component and y-component of optical flow respectively, cannot be determined. As a consequence, the optical flow (u, v) cannot be computed locally without introducing additional constraints.

In order to solve this problem, various methods have been proposed. The basic methods of gradient-based approach are reviewed in following section.

2.2 Increase in the Number of Observation Equations

One of the approaches of solving the gradient constraint equation is increase in the number of observation equations:

(a) by the assumption that a constant velocity over each spatial neighborhood (spatial local optimization method) (Barron, Fleet and Beauchemin, 1994, Kearney, Tompson, and Boley, 1987, Lucas and Kanade, 1981);
(b) by the constant velocity over temporal neighborhood (temporal local optimization method) (Kearney, Tompson and Boley, 1987, Nomura, Miike and Koga, 1991);
(c) by use of three channels (RGB, HSI) of each pixel (multispectral constraints method) (Markandey and Flinchbaugh, 1990, Mitiche, Wang and Aggarwal, 1987, Woodham, 1990);
(d) by use of second order derivatives of each pixel (second order derivative method) (Bainbridge-Smith and Lane, 1997, Nagel, 1983, Tistarelli and Sandini, 1990, Tretiak and Pastor, 1984, Uras, Girosi et al, 1988)
(e) by their combination.

The spatial local optimization method estimates optical flow by solving a group of observation equations obtained from a small spatial neighborhood of the image as a system of linear equations. Two observation equations are sufficient to arrive at unique solution for (u, v). More than two equations may be included in the system to reduce the effects of errors in the observation equations. Let the small spatial neighborhood, Ω, be equal to $n \times n$ pixels. The observation equation at each pixel in small spatial neighborhood can be obtained. We can get n^2 observation equations as an overdetermined linear equation system:

$$\boldsymbol{G}\boldsymbol{f} = -\boldsymbol{b} \tag{6}$$

where

$$\boldsymbol{G} = \begin{bmatrix} I_{1x} & I_{1y} \\ \vdots & \vdots \\ I_{ix} & I_{iy} \\ \vdots & \vdots \\ I_{n^2x} & I_{n^2y} \end{bmatrix}, \quad \boldsymbol{f} = \begin{bmatrix} u \\ v \end{bmatrix}, \quad \boldsymbol{b} = \begin{bmatrix} I_{1t} \\ \vdots \\ I_{it} \\ \vdots \\ I_{n^2t} \end{bmatrix} \tag{7}$$

has the least squares solution

$$\hat{\boldsymbol{f}} = -\left(\boldsymbol{G}^T\boldsymbol{G}\right)^{-1}\boldsymbol{G}^T\boldsymbol{b} \tag{8}$$

provided that the inverse of $\boldsymbol{G}^T\boldsymbol{G}$ exists.

In the temporal local optimization method, small temporal local neighborhood is n frames, and then we can obtain n observation equations. Multispectral constraints method has three observation equations. These overdetermined linear equation systems have the least squares solution in the same way as spatial local optimization method. In the case of second order derivative method, three observation equations are obtained, re:

$$\boldsymbol{G} = \begin{bmatrix} I_x & I_y \\ I_{xx} & I_{xy} \\ I_{yx} & I_{yy} \end{bmatrix}, \quad \boldsymbol{b} = \begin{bmatrix} I_t \\ I_{tx} \\ I_{ty} \end{bmatrix} \tag{9}$$

2.3 Imposition of the Smoothness Condition

Another approach of solving the gradient constraint equation is imposition of condition, that is:

(a) spatial smoothness of optical flow (spatial global optimization method) (Barron, Fleet and Beauchemin, 1994, Beauchemin and Barron, 1997, Horn and Schunck, 1981, Schunck, 1984);
(b) temporal smoothness (temporal global optimization method);
(c) their combination.

One way to express the additional condition is to minimize the square of the magnitude of the gradient of the optical flow velocity:

$$\left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial u}{\partial y}\right)^2 \text{ and } \left(\frac{\partial v}{\partial x}\right)^2 + \left(\frac{\partial v}{\partial y}\right)^2. \tag{10}$$

The total error, E, to be minimized as

$$E = \sum_x \sum_y \left\{ \left(I_x u + I_y v + I_t\right)^2 + \alpha^2 \left(\left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial u}{\partial y}\right)^2 + \left(\frac{\partial v}{\partial x}\right)^2 + \left(\frac{\partial v}{\partial y}\right)^2 \right) \right\}. \tag{11}$$

The minimization is to be accomplished by finding suitable values for the optical flow velocity (u, v). Using the calculus of variation, following equations are obtained.

$$\begin{aligned} I_x^{\,2} u + I_x I_y v &= \alpha^2 \nabla^2 u - I_x I_t \\ I_x I_y u + I_x^{\,2} v &= \alpha^2 \nabla^2 v - I_y I_t \end{aligned}. \tag{12}$$

However, it would be very costly to solve these equations simultaneously by one of the methods, such as Gauss-Jordan elimination. So, these equations should be solved by iterative method that is Gauss-Seidel method (Press et al, 1988).

In the temporal optimization method, smoothness condition is expressed as

$$\left(\frac{\partial u}{\partial t}\right)^2 \text{ and } \left(\frac{\partial v}{\partial t}\right)^2. \tag{13}$$

The total error is minimized in the same way as spatial global optimization method.

3. EXPERIMENTS

3.1 Empirical Comparison among Gradient-Based Approaches

In this chapter, various gradient-based approaches described in Chapter **2** are applied to real sequential images of traffic scene. The size of the frame in the sequential images is 720 by 480 pixels. And time interval is 1/30 second. Figure 1 shows the frame of sequential images used in the optical flow estimation. The vehicles in Figure 1 move from upper right to bottom left. Velocities of the vehicles in this image were measured, and the results were about (1) 20 pixels/frame, (2) 3 pixels/frame and (3) 2 pixels/frame, respectively. These values were used as measurements for comparison among the several approaches.

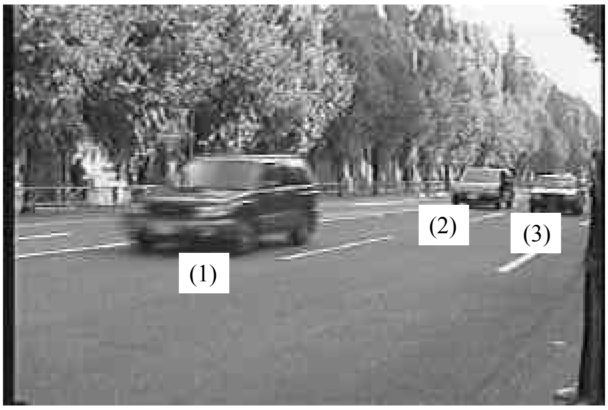

Figure 1: Image of Traffic Scene.

Figure 2 through Figure 9 show optical flow estimated by each method at a frame. Estimated optical flow is depicted as segment at an interval of 20 pixels, and the length of segments is ten times as long as estimated value.

Figure 2 shows the result which was solved by spatial local optimization method. The spatial neighborhood was defined as 5by 5 pixels. A constant optical flow over these neighborhoods was assumed. The magnitude of one of the estimated flow vectors in each vehicle was (1) 6.4 pixels/frame, (2) 2 .0 pixels/frame and (3) 1.5 pixels/frame, respectively. While there were precise flow vectors in the vehicles (2) and (3), even in the same vehicles flow vectors could not be obtained at many pixels. This problem will be described later. In the vehicle (1), though there were

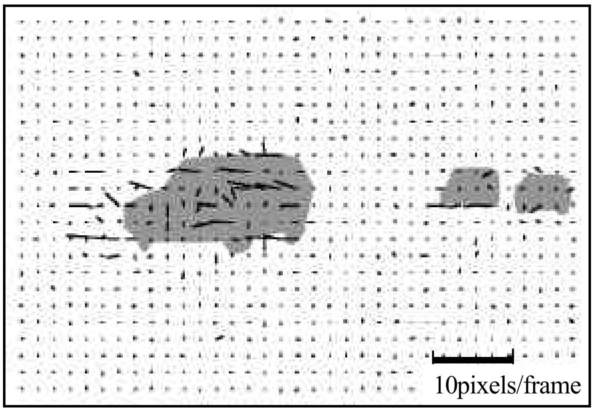

Figure 2: Spatial Local Optimization Method.

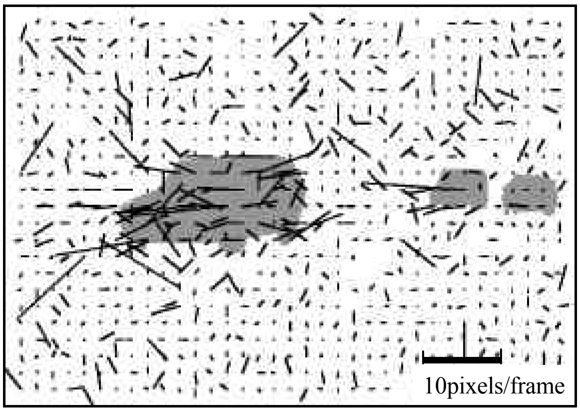

Figure 3: Temporal Local Optimization Method.

many flow vectors, inaccurate flow vectors were included.

Figure 3 shows the result which was solved by temporal local optimization method. The temporal neighborhood was defined as 3 frames. Optical flow estimated in the vehicle were much inaccurate, more over in the background, which are not moved naturally, were estimated as large magnitude. It is said that temporal local optimization method performed well, only when the velocity of object is less than 1 pixel/frame. Hence, when the time interval is about 1/30 seconds, it is not appropriate that a constant optical flow over temporal neighborhood is assumed.

The results which was solved by multispectral constrains method (Figure 4) was also worse. Because brightness of channels (RGB) in images taken in the air are similar, the solution is unstable (Figure 5).

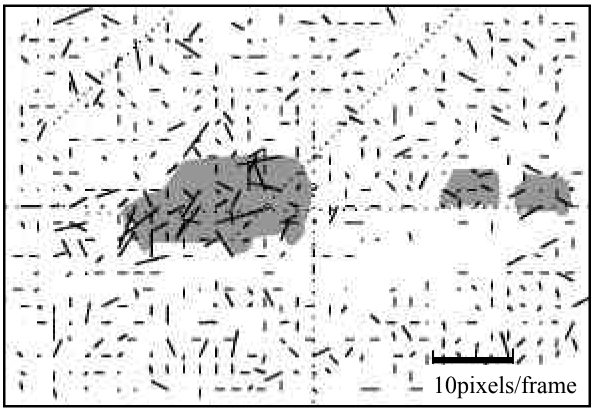

Figure 4: Multispectral Constraints Method.

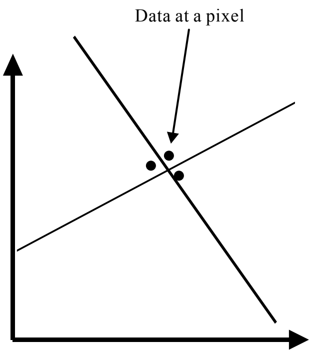

Figure 5: Unstability of the Solution by Multispectral Constraints Method.

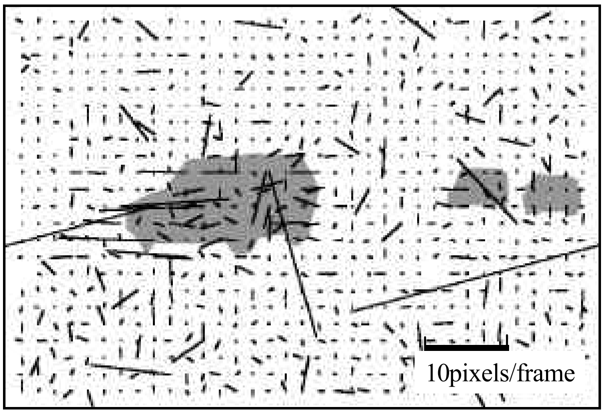

Figure 6: Second Order Derivative Method.

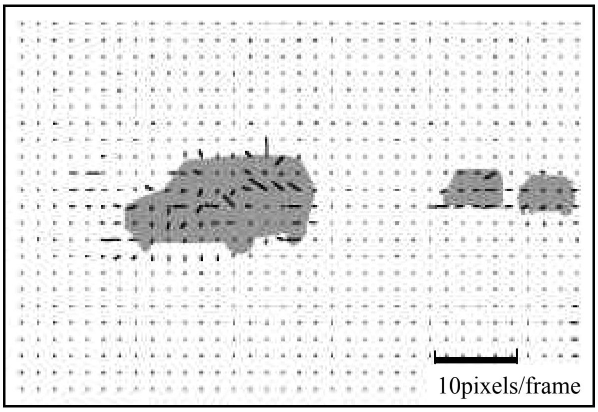

Figure 7: Spatial Global Optimization Method.

The results which was solved by second order derivative method (Figure 6) were also worse. In the second order derivative method, degree of freedom of overdetermined linear equation system is 1 as well as multispectral constraints method. Accordingly, the solution is also unstable and influence of noise become larger.

Table 1 shows averages of standard deviations of optical flow (u, v) estimated by the approaches, which solve the constraints equation by increasing in the number of observation equations. Comparing the orders of those averages, (a) spatial local optimization method gives the best estimated value in all over the image.

In global optimization method, iterative number was 100, and coefficient was defined as 100. In the result which was solved by spatial global optimization method, the optical flow tended to be small (Figure 7). The magnitude of one of the estimated flow vectors in each vehicle was (1) 5.3 pixels/frame, (2) 2.0 pixels/frame and (3) 0.9 pixels/frame, respectively. The smoothness constraint smoothes variety of direction of flow vectors, so the magnitudes of flow vectors are under estimated.

In addition to above-mentioned methods, which were applied to the real sequential images, the results by other method were worse.

Table 1: Averages of Standard Deviations of Estimated Values by Each Method

	Std. Dev. of u (pixels)	Std. Dev. of v (pixels)
(a)Spatial Local Optimization Method	0.13	0.15
(b)Temporal Local Optimization Method	1036	1014
(c)Multispectral Constraints Method	720	636
(d)Second Order Derivative Method	697	677
(a)+(b)	67	65
(a)+(c)	49	48
(a)+(d)	45	44
(b)+(c)	202	203
(b)+(d)	190	188
(c)+(d)	140	139
(a)+(b)+(c)	28	27
(a)+(b)+(d)	26	25
(a)+(c)+(d)	19	19
(b)+(c)+(d)	85	39
(a)+(b)+(c)+(d)	12	5

3.2 Improvement of Spatial Local Optimization Method

The result, which was solved by spatial local optimization method (Figure 2), was better than other gradient-based approaches. In the spatial local optimization method, there were pixels at which flow vectors could not be estimated. It is called as aperture problem (Figure 8). It shows a moving plane of constant brightness. When we see the plane through the aperture, we cannot recognize the moving of the plane.

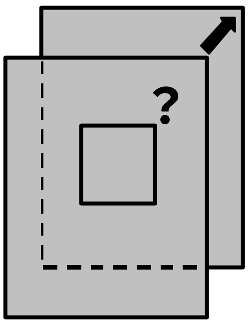

Figure 8: Aperture Problem.

Consequently, flow vectors at all feature points cannot be obtained by employing the basic methods of gradient-based approach and their combined methods. Hence, it is difficult to analyze details of vehicle motions taking into account the shape of the vehicles by the flow vectors which are solved by basic methods of gradient-based approach.

Figure 9 shows the distribution of t-value of estimated optical flow only in the region of vehicles. In these figures, color at the pixel whose t-value is more than 2.0 is black. Particularly in the (a) t-value of estimated $\hat{u}$, estimations of optical flow are significant at the pixels which locate near edges or have optical flow of small magnitudes.

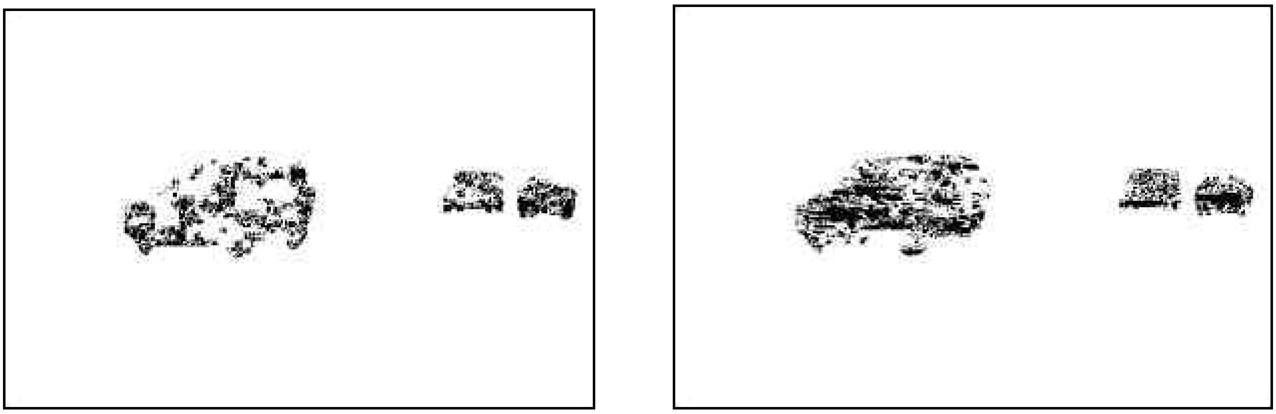
(a) T-Value of Estimated $\hat{u}$ (b) T-Value of Estimated $\hat{v}$

Figure 9: T-Value of Estimated Optical Flow in the Region of Vehicles

According to experimental results, when the velocity of a vehicle is large, the estimated value tends to be incorrect. It is said that gradient-based approaches perform well, when the velocity of object is small. In this case, resolution of the image is important. To avoid this problem, hierarchical estimation of flow vectors has been proposed (Sato and Sasaki, 1986, Tominaga et al, 1989). At first, multi-resolution images $\{I_0^1, I_1^1, \cdots I_{k-1}^1, I_k^1\}$, $\{I_0^2, I_1^2, \cdots I_{k-1}^2, I_k^2\}$ are prepared. The image I_k^1 and I_k^2 is composed 2^k by 2^k pixels. From low resolution image to high resolution image, optical flow is estimated step by step. Let $\boldsymbol{f}_k$ be optical flow in image I_k. When $\boldsymbol{f}_k$ is estimated, the I_k^2 is shifted following $\boldsymbol{f}_k$. At the time, the flow vector is

$$\boldsymbol{f} = \boldsymbol{f}_k + 2\times \boldsymbol{f}_{k-1}\,. \tag{14}$$

This process is carried out repeatedly. Finally, estimated optical flow is

$$\boldsymbol{f} = \boldsymbol{f}_k + 2\times \boldsymbol{f}_{k-1} + 4\times \boldsymbol{f}_{k-2} + \cdots + 2^k\,\boldsymbol{f}_0\,. \tag{15}$$

In addition, flow vectors were extracted near edges, and no flow vectors inside the vehicles. According to this result, the edges were defined as spatial neighborhood, and then spatial local optimization method were applied (Davis, Wu and Sun, 1983).

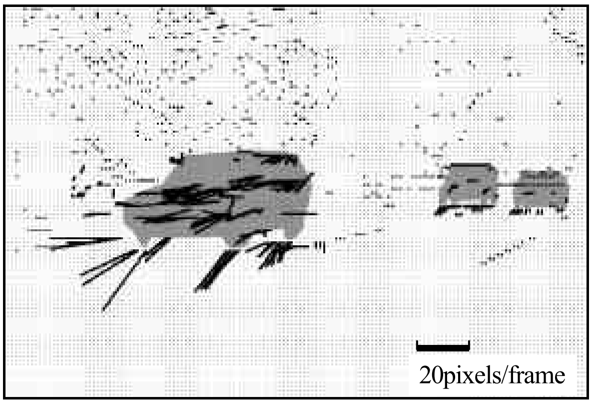

Figure 10: Hierarchical Estimation and Spatial Local Optimization Method with Neighborhood of Edge.

Figure 10 shows the result which was solved by hierarchical estimation and spatial local optimization method with neighborhood of edge. Estimated optical flow is depicted as segment at an interval of 5 pixels, and the length of segments is three times as long as estimated value. The magnitude of flow vector was improved. However, the results were not precise and dense sufficiently to be employed for 3D reconstruction and structure from motion.

4. CONCLUSION

The conclusions of this paper are as follows:

(1) Theoretical review of gradient-based approaches from the viewpoint of regularization;
(2) Empirical comparison among basic method of gradient-based approach from the viewpoint of application to vehicle motion analysis.

In this paper, the sequential image was taken at an interval about 1/30, and such a sequential image can be acquired easily. According to the empirical comparison, it is difficult to estimate precise and dense optical flow by the basic methods of gradient-based approach and their combination, when sequential images are taken at an interval about 1/30 seconds. And then, it is difficult to analyze details of vehicle motions such as 3D reconstruction by the flow vectors which are solved by the basic methods of gradient-based approach and their combination in this case. Hence, other approaches, which are different with gradient-based approach fundamentally, are required.

The future works are as follows:

(1) Application to sequential images which are taken at an shorter interval;
(2) Empirical comparison with methods of pattern matching such as Least Squares Matching.

REFERENCES

Bainbridge-Smith, A. and Lane, R.G., 1997. Determining Optical Flow Using Differential Method. Image and Vision Computing 15, pp.11-22.

Barron, J.L., Fleet, D.J. and Beauchemin, S.S., 1994. Performance of Optical Flow Techniques. International Journal of Computer Vision, 12(1), pp.43-77.

Beauchemin, S.S. and Barron, J.L., 1997. The Computation of Optical Flow. ACM Computing Surveys, 27(3), pp.433-467.

Davis, L.S., Wu, Z. and Sun, H., 1983. Contour-Based Motion Estimation. Computer Vision, Graphics and Image Processing, 23, pp.313-326.

Horn, B.K.P. and Schunck, B.G., 1981. Determining Optical flow. Artificial Intelligence, 17, pp.185-203.

Kearney, J.K., Tompson, W.B. and Boley, D.L., 1987. Optical Flow Estimation: An Error Analysis of Gradient-Based Methods with Local Optimization. IEEE Transactions on Pattern Analysis and Machine Intelligence, PAMI-9(2), pp.229-244.

Lucas, B.D. and Kanade, T., 1981. An Iterative Image Registration Technique with an Application to Stereo Vision. In: Proceedings of the Seventh International Joint Conference on Artificial Intelligence, pp.674-679.

Markandey, V. and Flinchbaugh, B.E., 1990. Multispectral Constraints for Optical Flow Computation. In: Proceedings of Third International Conference on Computer Vision, pp.38-41.

Mitiche, A., Wang, Y.F. and Aggarwal, J.K., 1987. Experiments in Computing Optical Flow with Gradient-Based, Multiconstraint Method. Pattern Recognition, 20, pp.173-179.

Nagel, H-H., 1983. Displacement Vectors Derived from Second-Order Intensity Variations in Image Sequences, Computer Vision, Graphics, and Image Processing, 21, pp.85-117.

Nomura, A., Miike, H. and Koga, K., 1991. Field Theory Approach for Determining Optical Flow. Pattern Recognition Letters, 12, pp.183-190.

Press, W.H., Flannery, B.P., Teukolsky, S.A. and Vetterling, W.T., 1988. Numerical Recipes in C. Cambridge University Press, Cambridge.

Sato, M. and Sasaki, H., 1986. Hierarchical Estimation of Displacement Vectors in Image Sequences. Transactions of the Institute of Electronics, Information and Communication Engineerings, J69 (D5), pp.771-776.

Schunck, B.G., 1984. The Motion Constraints Equation for Optical Flow. In: Proceedings Seventh International Conference on Pattern Recognition, pp.20-22.

Tistarelli, M. and Sandini, G., 1990. Estimation of Depth from Motion Using an Anthropomorphic Visual Sensor. Image and Vision Computing, 8(4), pp.271-278.

Tominaga, H., Komatsu, N., Miyashita, T. and Hanamura, T., 1989. A Motion Detection Method on Video Image by Using Hierarchical Pixels. Transactions of the Institute of Electronics, Information and Communication Engineerings, J72 (D-Ⅱ3), pp.395-403.

Tretiak, O. and Pastor, L., 1984. Velocity Estimation from Image Sequences with Second Order Differential Operators. In: Proceedings of International Conference on Pattern Recognition, pp.16-19.

Uras, S., Girosi, F., Verri, A. and Torre, V., 1988. A Computation Approach to Motion Perception. Biological Cybernetics, 60, pp.79-87.

Woodham, R.J., 1990. Multiple Light Sourse Optical Flow. In: Proceedings of Third International Conference on Computer Vision, pp.42-46.