The Spatio-Temporal Optimization to Determine Optical Flow with Combination of Local and Global Approach

L. ZHANG, H. MIIKE and K. KURIYAMA

Graduate School of Science and Engineering, Yamaguchi University, 2557 Tokiwadai, Ube 755, Japan

(Received November 2, 1998; Accepted December 21, 1998)

Keywords: Optical Flow, Global Optimization, Local Optimization, Spatio-Temporal Optimization, Gradient-Based Technique

Abstract. The gradient-based technique is one of the popular methods to determine optical flow (apparent two-dimensional velocity field). It can be classified into two broad categories: the global optimization and the local optimization. Horn and Schunck proposed the spatial global optimization to constrain the estimated velocity field. In this study, a new approach is proposed to extract the reliable optical flow fields. This is the gradient-based spatio-temporal optimization with combination of local and global approach. Main point is the introduction of a spatio-temporal measure (locally averaged) of the error function in the conservation equation. Experiments have been made to confirm the performance of the proposed method and to clarify the difference of characteristics between them.

1. Introduction

Determining optical flow (HORN, 1986; SINGH, 1991; JAHNE, 1995) is one of the most important problems in the processing of image sequence. A number of different approaches to determine optical flow have been proposed including gradient-based, correlation-based, energy-based and phase-based methods. A recent survey is due to BARRON *et al.* (1994), where the different approaches were compared on a series of synthetic and real images. In the gradient-based method, it can be classified into two broad categories: the global optimization (HORN and SCHUNCK, 1981; NAGEL and ENKELMANN, 1986; CHAUDHURY and MEHROTRA, 1995; BLACK and ANANDAN, 1996) and the local optimization (KEARNEY *et al.*, 1987; VERRI *et al.*, 1990; NOMURA *et al.*, 1991; ZHANG *et al.*, in press).

HORN and SCHUNCK (1981) are the pioneers of using the gradient-based global optimization. They use the spatial global optimization to constrain the estimated velocity field under the assumption that neighboring points on the objects have similar velocities and the velocity field of the brightness patterns in the image varies smoothly almost everywhere. The approach can be viewed as "*The spatial global optimization*". CHAUDHURY and MEHROTRA (1995) proposed the modified Horn and Schunck method. They modified the Horn and Schunck approach to include a temporal smoothness. We can view the modified Horn and Schunck approach proposed by Chaudhury and Mehrotra as "*The spatio-temporal global optimization*". BLACK *et al.* (1996) proposed the robust estimation of multiple motions approach. Instead of the traditional least-squares error function, they used an error function (for example, Lorentizian function) which is robust for handling the problem of discontinuities occurring at motion boundaries. They also described a regularization technique, which uses a robust version of the standard optical flow constraint equation and a robust first-order smoothness term. They considered the local average smoothness of velocity in a spatial neighborhood (e.g. a 3×3 neighborhood). Based on Horn and Schunck approach, in this report, we proposed a new approach to extract the reliable optical flow fields. This is the gradient-based spatio-temporal optimization with combination of local and global approach. The performance of the proposed method is confirmed by use of synthetic and real image sequences.

2. The Spatial Global Optimization

Let the image brightness function at the point (x, y) in the image plane at time t be $E(x, y, t)$. Assuming that the brightness of a particular point in the image plane is constant, we have the well-known conservation equation (HORN and SCHUNCK, 1981):

$$E_x u + E_y v + E_t = 0, \tag{1}$$

where the E_x, E_y, and E_t denote the partial derivatives of image brightness with respect to x, y and t, respectively, the u and v denote the two unknown components of velocity vector (u, v). The conservation equation (1) provides one linear equation in the variables u and v. As a consequence, the optical flow cannot be determined at a point in the image independently of neighboring points without introducing additional constraints.

Horn and Schunck introduce the spatial global smoothness constraint as the additional constraint. They assume that neighboring points on the objects have similar velocities and the velocity field of the brightness patterns in the image varies smoothly almost everywhere. The estimation of optical flow is obtained by minimizing

$$\iint_S \left(E_b + \alpha^2 E_c\right) dxdy, \tag{2}$$

where α^2 denotes a suitable weighting factor, E_b denotes a local measure of errors in the conservation equation at a pixel site,

$$E_b = \left(E_x u + E_y v + E_t\right)^2, \tag{3}$$

and E_c denotes the measure of smoothness in the velocity field (see Fig. 1),

$$E_c = \left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial u}{\partial y}\right)^2 + \left(\frac{\partial v}{\partial x}\right)^2 + \left(\frac{\partial v}{\partial y}\right)^2. \tag{4}$$

Horn and Schunck used the calculus of variation to minimize the total error given by Eq. (2), obtaining

$$\begin{aligned} \alpha^2 \nabla^2 u &= \left(E_x u + E_y v + E_t\right) E_x, \\ \alpha^2 \nabla^2 v &= \left(E_x u + E_y v + E_t\right) E_y, \end{aligned} \tag{5}$$

where $\nabla^2 = \partial^2/(\partial x^2) + \partial^2/(\partial y^2)$ is the Laplacian. Using the approximation to the Laplacian (HORN and SCHUNCK, 1981), $\nabla^2 u = \bar{u} - u$ and $\nabla^2 v = \bar{v} - v$, from Eq. (5) we can obtain

$$\begin{aligned} \left(\alpha^2 + E_x^2 + E_y^2\right)(u - \bar{u}) &= -E_x\left(E_x \bar{u} + E_y \bar{v} + E_t\right), \\ \left(\alpha^2 + E_x^2 + E_y^2\right)(v - \bar{v}) &= -E_y\left(E_x \bar{u} + E_y \bar{v} + E_t\right), \end{aligned} \tag{6}$$

where the local average $\bar{u}$ and $\bar{v}$ are defined as follows (see Fig. 11 in Appendix B)

$$\begin{aligned} \bar{u}_{i,j,t} = & \frac{1}{6}\left\{u_{i-1,j,t} + u_{i,j+1,t} + u_{i+1,j,t} + u_{i,j-1,t}\right\} \\ & + \frac{1}{12}\left\{u_{i-1,j-1,t} + u_{i-1,j+1,t} + u_{i+1,j+1,t} + u_{i+1,j-1,t}\right\}, \end{aligned}$$

$$\begin{aligned} \bar{v}_{i,j,t} = & \frac{1}{6}\left\{v_{i-1,j,t} + v_{i,j+1,t} + v_{i+1,j,t} + v_{i,j-1,t}\right\} \\ & + \frac{1}{12}\left\{v_{i-1,j-1,t} + v_{i-1,j+1,t} + v_{i+1,j+1,t} + v_{i+1,j-1,t}\right\}. \end{aligned}$$

3. Proposed Method

3.1. The spatio-temporal measure of error

In Eq. (3), Horn and Schunck used E_b denoting the local measure of errors in the conservation equation at a pixel site. Since Eq. (2) is estimated from discrete images, the estimation of optical flow will be inaccurate due to noise in the imaging process and sampling measurement error. We introduce a spatio-temporal measure to reduce the error. Considering a small spatio-temporal volume $\delta v = \delta x \cdot \delta y \cdot \delta t$ around a considerable pixel. An averaged error function is defined as

$$\hat{E}_b = \iiint_{\delta v} \left(E_x u + E_y v + E_t \right)^2 dxdydt. \tag{7}$$

A local constancy of the optical flow (u, v) is assumed to evaluate the error function in the small spatio-temporal volume $\delta v = \delta x \cdot \delta y \cdot \delta t$. Generally, the conventional local optimization approach is designed to minimize the error function directly (with the linear least-squares method or other minimization method, (e.g., KEARNEY *et al.*, 1987; NOMURA *et al.*, 1991; ZHANG *et al.*, in press)). Here, we try to combine the local optimization and global optimization approach, assuming that the optical flow (u, v) is constant in a small spatio-temporal volume $\delta v = \delta x \cdot \delta y \cdot \delta t$, and the neighboring points (among small spatio-temporal volumes δv) on the objects have similar velocities and the velocity field varies smoothly.

3.2. The spatio-temporal measure of smoothness

Horn and Schunck use the *spatial* measure of smoothness in the velocity field (see Fig. 1) to minimize the square of the magnitude of the gradient of the optical flow in the velocity plane (Eq. (4)). Considering the spatio-temporal continuity of the velocity fields in the image sequence, we introduce a concept of *spatio-temporal smoothness* (see Fig. 2). The spatio-temporal measure of smoothness is defined as (CHAUDHURY and MEHROTRA, 1995)

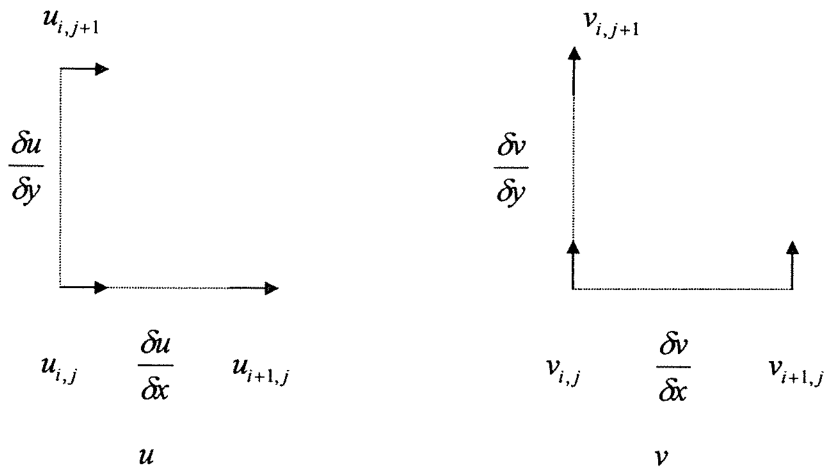


Fig. 1. The spatial measure of smoothness in the velocity field.

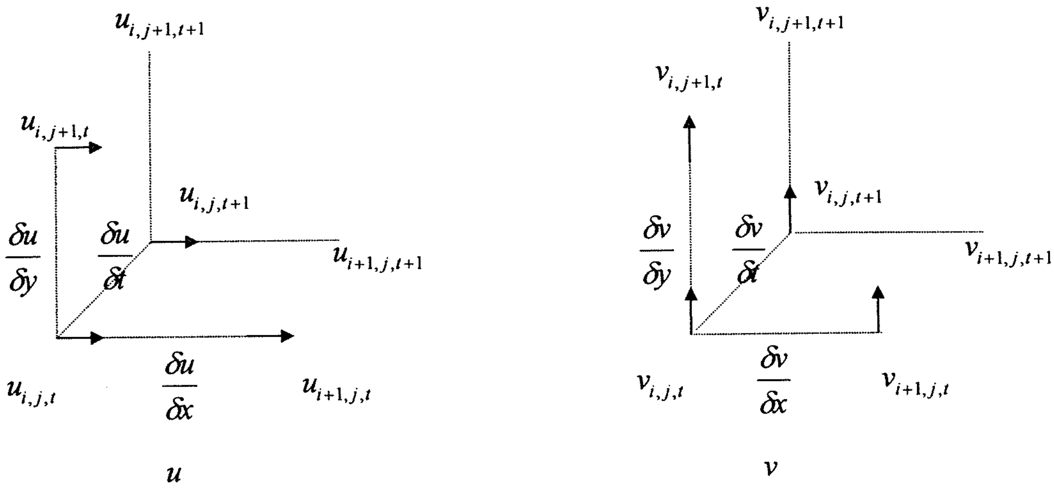


Fig. 2. The spatial-temporal measure of smoothness in the velocity fields.

$$\hat{E}_c = \left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial u}{\partial y}\right)^2 + \left(\frac{\partial u}{\partial t}\right)^2 + \left(\frac{\partial v}{\partial x}\right)^2 + \left(\frac{\partial v}{\partial y}\right)^2 + \left(\frac{\partial v}{\partial t}\right)^2. \tag{8}$$

3.3. Determining optical flow with spatio-temporal optimization

The solution of optical flow is obtained by minimization of the sum of the spatio-temporal error function $\hat{E}_b$ (Eq. (7)) and the spatio-temporal measure of smoothness $\hat{E}_c$ (Eq. (8)). Let the total errors to be minimized be

$$\iiint_v \left(\hat{E}_b + \alpha^2 \hat{E}_c\right) dxdydt, \tag{9}$$

where α^2 denotes a suitable weighting factor. Using the calculus of variation (see Appendix A) we obtain

$$\begin{aligned} \alpha^2 \nabla^2_{xyt} u &= \sum_{\delta v} E_x^2 u + \sum_{\delta v} E_x E_y v + \sum_{\delta v} E_x E_t, \\ \alpha^2 \nabla^2_{xyt} v &= \sum_{\delta v} E_y E_x u + \sum_{\delta v} E_y^2 v + \sum_{\delta v} E_y E_t, \end{aligned} \tag{10}$$

where

$$\nabla_{xyt}^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial t^2} \tag{11}$$

is a Laplacian operator. Using the approximation to the Laplacian (see Appendix B), $\nabla_{xyt}{}^2 u = \bar{u} - u$ and $\nabla_{xyt}{}^2 v = \bar{v} - v$, from Eq. (10) we obtain (see Appendix A)

$$\left(\alpha^2 + \sum_{\delta v} E_x^2 + \sum_{\delta v} E_y^2\right)(u - \bar{u}) = -\left(\sum_{\delta v} E_x^2 \bar{u} + \sum_{\delta v} E_x E_y \bar{v} + \sum_{\delta v} E_x E_t\right),$$

$$\left(\alpha^2 + \sum_{\delta v} E_x^2 + \sum_{\delta v} E_y^2\right)(v - \bar{v}) = -\left(\sum_{\delta v} E_y E_x \bar{u} + \sum_{\delta v} E_y^2 \bar{v} + \sum_{\delta v} E_y E_t\right), \tag{12}$$

where the local average $\bar{u}$ and $\bar{v}$ are defined as a weighted sum of neighboring points in $3\times3\times3$ volume (see Fig. 12 in Appendix B). When we compare between the conventional method (the global optimization method by Horn and Schunck, Eq. (6)) and the proposed method (Eq. (12)) the mainly improved points are:

1) We introduce the averaged error function $\hat{E}_b$ (Eq. (7)) instead of E_b (Eq. (3)) to estimate the errors for the rate of change of image brightness. Equation (7) assumes that the optical flow (u, v) is constant in a small spatio-temporal volume $\delta v = \delta x \cdot \delta y \cdot \delta t$. We evaluate the errors in the conservation equation in a small spatio-temporal volume δv but not at a pixel site (x, y, t). This means that we combine the global and local constraints to estimate optical flow fields.

2) We use the spatio-temporal measure of smoothness $\hat{E}_c$ (Eq. (8), Fig. 2) instead of the spatial measure of smoothness E_c (Eq. (4), Fig. 1) in the velocity fields. Horn and Schunck consider one velocity plane (Fig. 1) under the assumption that neighboring points on the objects have similar velocities and the velocity field of the brightness patterns in the image varies smoothly almost everywhere. We consider the plural velocity planes (Fig. 2) and introduce the spatio-temporal measure of smoothness in the velocity fields. We assume that neighboring points (among small spatio-temporal volumes δv) on the objects have similar velocities and the velocity field (on plural velocity planes) of the brightness patterns in the image varies smoothly almost everywhere.

3) We use the spatio-temporal global optimization (Eq. (9)) instead of the spatial global optimization (Eq. (2)) to determine optical flow. We minimized the total errors (Eq. (9)) on plural image planes but not on one image plane s (see Eq. (2)). It is expected to obtain high resolution and high reliability of optical flow field. The performance of the proposed method is confirmed by the experiments below.

4. Experiments

In this section, we try to apply the proposed method to determine optical flow fields. We compare the proposed method with the conventional gradient-based methods (spatial global optimization, HORN and SCHUNCK, 1981 and spatio-temporal global optimization, CHAUDHURY and MEHROTRA, 1995) and discuss the performance of the proposed method. The experimental data includes two synthetic image sequences and one real image sequence.

4.1. Experimental images analysis

4.1.1 Yosemite sequence (synthetic data)*

The *Yosemite sequence* has a resolution of 316 × 252 pixels and 15 frames. The brightness is quantified into 256 steps. The *Yosemite sequence* is a complex test scene. In the scene, the cloud has a translational motion with a speed of 2 pixels/frame, while speed in the lower left is about 4–5 pixels/frame. However, the brightness of cloud changes with respect to time and space. The landscape (mountains, valley, etc.) moves against depth direction. Then, motion field expands. Namely, the motion field has divergence characteristics. This sequence is challenging because of the range of velocities, occluding edges between the mountains and at the horizon, divergence and non-uniform illumination. Figure 3 shows the 1st and 10th frames of the synthetic image sequence. The theoretical optical flow field is shown in Fig. 4.

(a) (b)

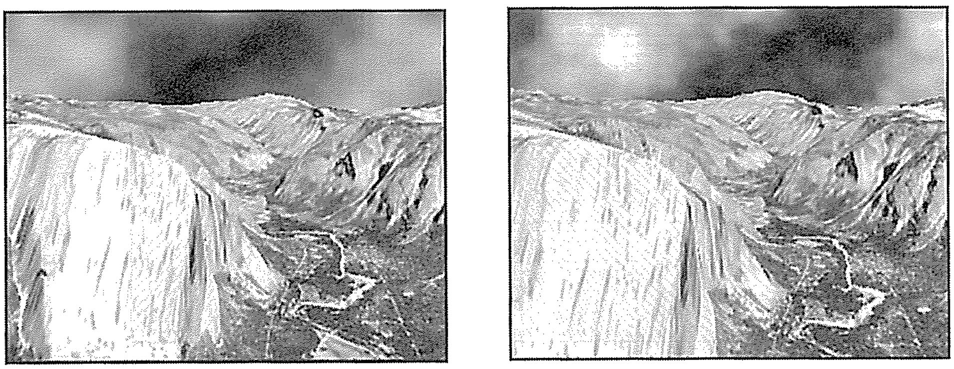


Fig. 3. The *Yosemite sequence*: (a) 1st frame and (b) 10th frame.

---

*The image sequence of *Yosemite* is obtained from the ftp-site of ftp.csd.uwo.ca.

4.1.2 Diverging Tree sequence (synthetic data)*

In the *Diverging Tree sequence*, the camera moves along its line of sight, the focus of expansion is at the center of the image, and the image speeds vary from 1.29 pixels/frame on the left side to 1.86 pixels/frame on the right. The size of the *Diverging Tree sequence* is 150 × 150 pixels and 40 frames. The brightness is quantified into 256 steps. Figure 5 shows the 5th and 25th frames of the synthetic image sequence. The theoretical optical flow field is shown in Fig. 6.

4.1.3 Toy car sequence (real data)

The usefulness of the proposed method is also tested with real image sequence. We took sequential images of toy car motions on floor through a TV camera with sampling frequency of 30 Hz. The size of the *Toy car sequence* is 236 × 110 pixels and 40 frames. Brightness is quantified into 256 steps. The toy car moves from lower left to upper right. Figure 7 shows the 10th and 25th frames of the *Toy car sequence*.

4.2. Experimental results

In this section, we report the quantitative performance of the conventional gradient-based methods (spatial global optimization, HORN and SCHUNCK, 1981 and spatio-temporal global optimization, CHAUDHURY and MEHROTRA, 1995) and the proposed method on the synthetic image sequences. We also show the optical flow fields produced by the

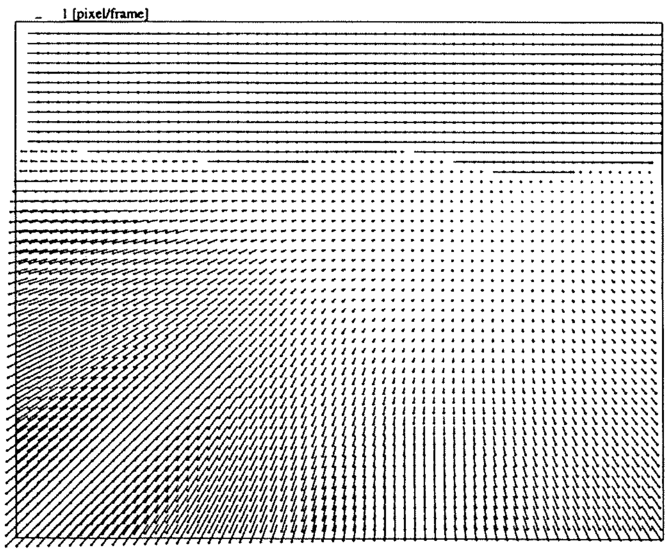


Fig. 4. Theoretical optical flow field of *Yosemite sequence*.

*The image sequence of *Diverging Tree* is obtained from the ftp-site of ftp.csd.uwo.ca.

(a) (b)

Fig. 5. The *Diverging Tree sequence*: (a) 5th frame and (b) 25th frame.

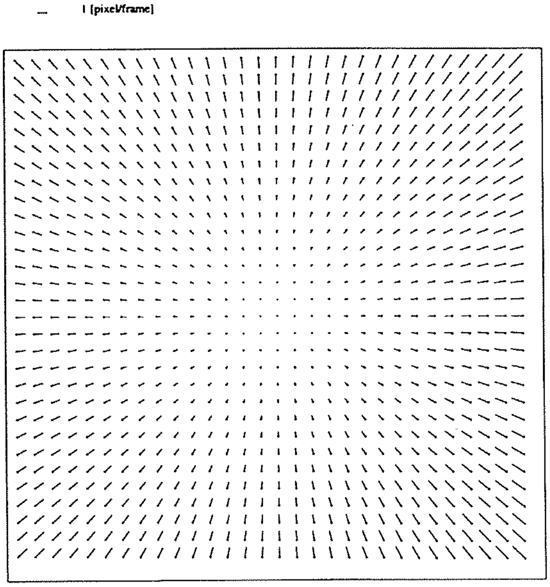
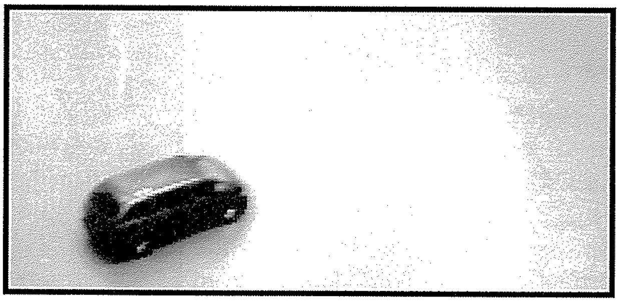


Fig. 6. Theoretical optical flow field of *Diverging Tree sequence*.

methods on the real image sequence. We try to determine optical flow fields by use of the following three methods. First, we use the conventional gradient-based method (HORN and SCHUNCK, 1981, 1986).

Spatial Global Optimization (SGO): HORN and SCHUNCK (1981, 1986) combined the basic constraint equation with a spatial global smoothness term to constrain the

(a)



(b)

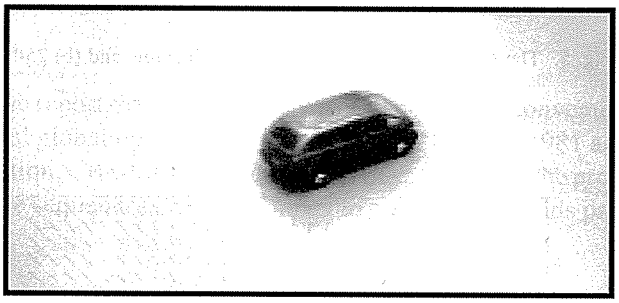


Fig. 7. The *Toy car sequence*: (a) 10th frame and (b) 25th frame.

(a)

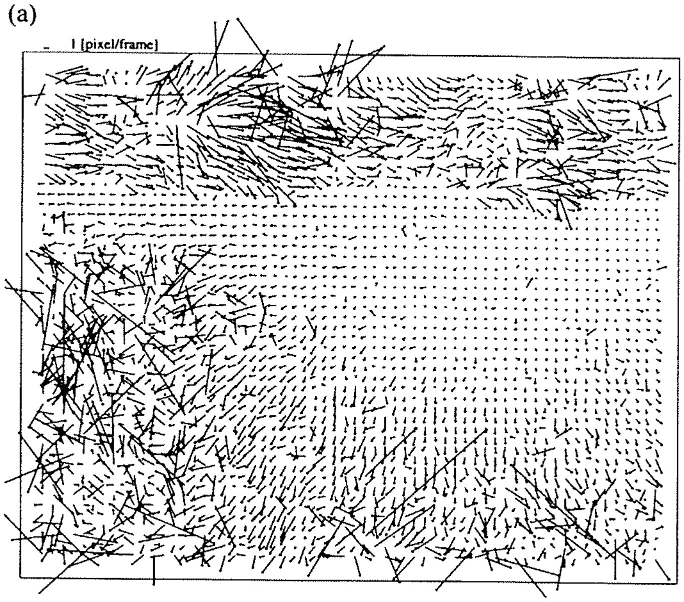
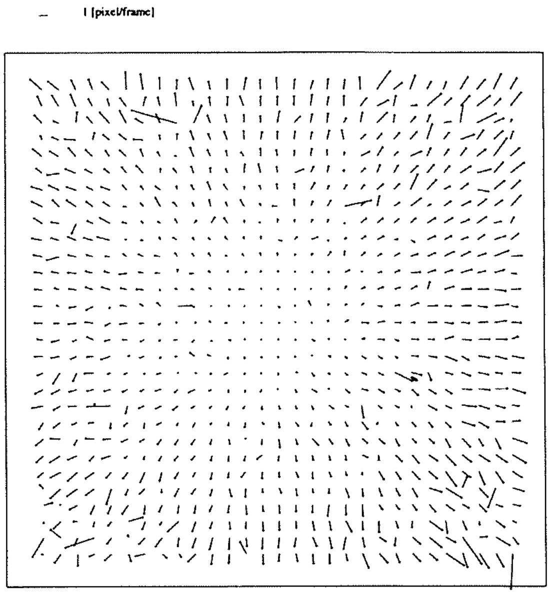


Fig. 8. The optical flow fields determined by SGO: Horn and Schunck. (a) *Yosemite sequence*, (b) *Diverging Tree sequence* and (c) *Toy car sequence*.

(b)



(c)

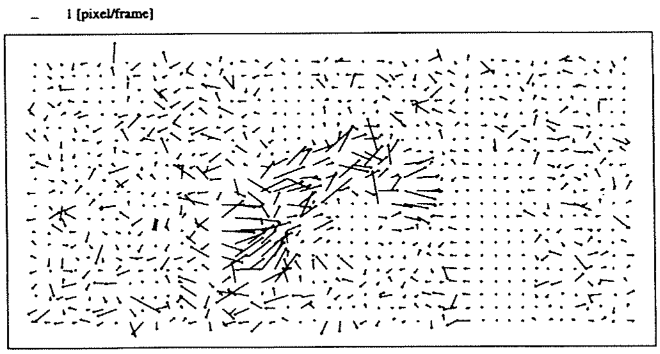


Fig. 8. (continued).

estimated velocity field $\vec{v} = (u, v)$ (Eq. (2)). We used $\alpha = 0.5$ and 100 iterations as suggested by BARRON *et al.* (1994) in testing below. Two consecutive images are used to compute the partial derivatives using a $2 \times 2 \times 2$ spatio-temporal neighborhood (HORN and SCHUNCK, 1981). The results of optical flow fields obtained by the method are shown in Figs. 8(a)–(c).

Second, Spatio-Temporal Global Optimization (STGO): CHAUDHURY and MEHROTRA (1995) proposed the modified Horn and Schunck approach. They modified the Horn and Schunck approach to include a temporal smoothness. Then, the approach can process multiframe. In the experiment we test the modified Horn and Schunck approach using $\alpha = 0.5$ and 100 iterations. Five frames of images are used to compute three velocity planes. The partial derivatives are estimated by using a $3 \times 3 \times 3$ spatio-temporal neighborhood. The results of optical flow fields obtained by the modified Horn and Schunck approach are shown in Figs. 9(a)–(c).

Third, we use the proposed Spatio-Temporal Optimization with combination of Local and Global approach (STOLG): We combined the spatio-temporal error function with the spatio-temporal measure of smoothness in the velocity fields (Eqs. (7) and (8)) to constrain the estimated velocity field $\vec{v} = (u, v)$ (Eq. (9)). Since we use the spatio-temporal optimization, the proposed approach can process not only two successive frames (as Horn and Schunck did) but also the multiframe. We use five frames of images to compute three velocity planes. The partial derivatives are estimated by using a $3 \times 3 \times 3$ spatio-temporal neighborhood. We used $\alpha = 0.5$ and 100 iterations as the same condition before. The results of optical flow fields obtained by the proposed method are shown in Figs. 10(a)–(c).

The optical flow fields obtained by the conventional gradient-based method (Horn and Schunck, see Fig. 8) and the modified Horn and Schunck approach (see Fig. 9) have

(a)

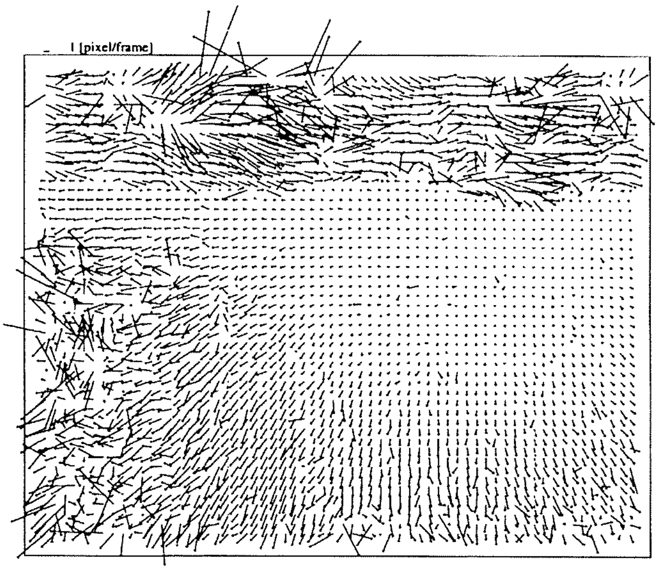
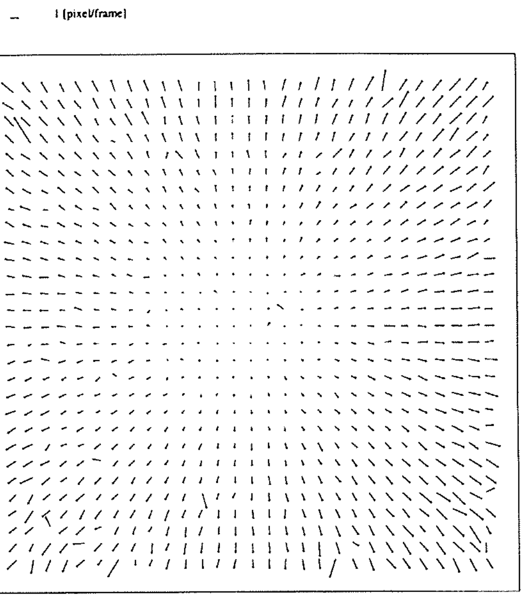


Fig. 9. The optical flow fields determined by STGO: modified Horn and Schunck (CHAUDHURY and MEHROTRA, 1995). (a) *Yosemite sequence*, (b) *Diverging Tree sequence* and (c) *Toy car sequence*.

(b)



(c)

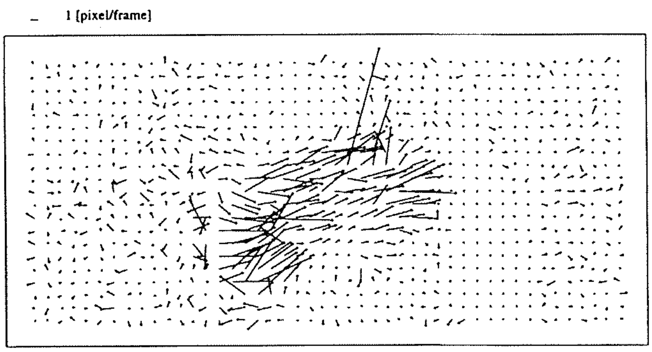
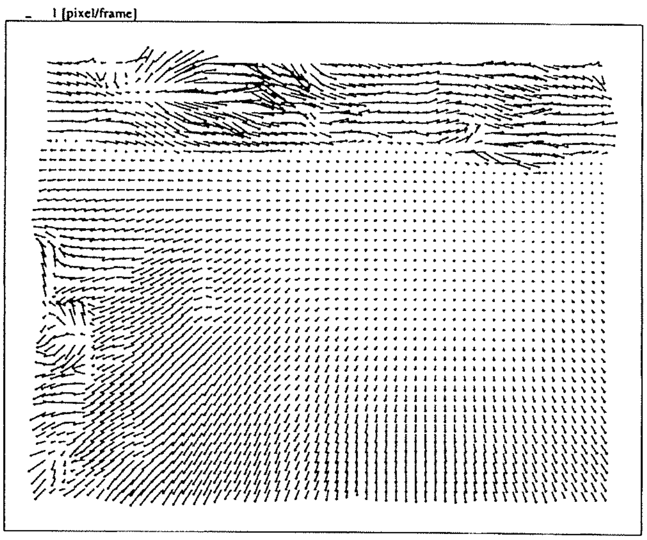


Fig. 9. (continued).

serious errors. When we apply the proposed spatio-temporal optimization under the assumption of Eqs. (7) and (8), optical flow fields are improved apparently (see Fig. 10). For more detailed and quantitative evaluation see Table 1 in Subsection 4.3. However, for *Yosemite sequence*, reduction of the error at foreground mountain surface and at the cloud position are not satisfactory, because there are the aperture problem and the spatio-

(a)



(b)

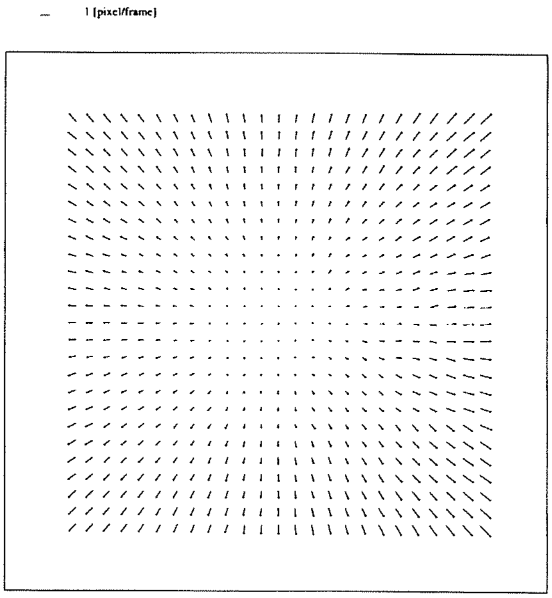


Fig. 10. The optical flow fields determined by STOLG: Proposed method. (a) *Yosemite sequence*, (b) *Diverging Tree sequence* and (c) *Toy car sequence*.

(c)

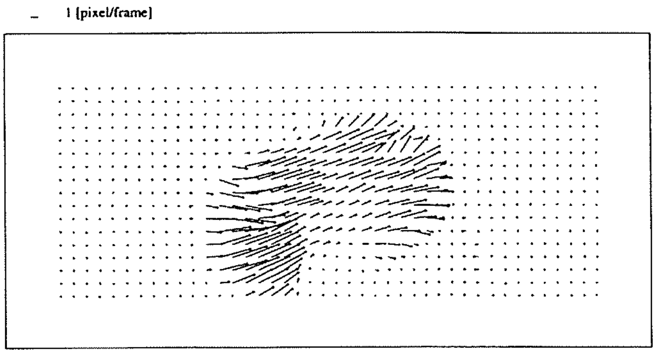


Fig. 10. (continued).

temporal non-uniform illumination. The texture of foreground mountain surface has only pinstriped, it is easy to encounter the aperture problem. When we try to determine motion fields by the gradient-based method, we have to consider the aperture problem. In order to overcome the shortage, it seems effective to introduce a spatial presmoothing (e.g. Gaussian spatial presmoothing) based on the image sequence (ZHANG *et al.*, in press) or the hierarchical approach (GLAZER, 1983). For the spatio-temporal non-uniform illumination, in our latest report (ZHANG *et al.*, in press), we introduced the pixel-based temporal filtering and the extended constraint equation with spatio-temporal local optimization to cope with the problem. The further research considering the spatio-temporal non-uniform illumination with spatio-temporal optimization combining local and global approach is expected.

4.3. *Error measurement*

We tested the proposed algorithm on the synthetic and real images. For the experiment on the synthetic images where the correct motion fields are known, we can use the angular measure of error used by BARRON *et al.* (1994) to evaluate the results. They measure the error between the correct velocity $\vec{v}_c = (u, v)$ and the estimate $\vec{v}_e = (\hat{u}, \hat{v})$ as the angle between the unit vectors in 3D space,

$$\vec{v}_3 \equiv \frac{(u, v, 1)}{\sqrt{u^2 + v^2 + 1}}. \tag{13}$$

The angular error between the correct vector $\vec{v}_{3c}$ and the estimate $\vec{v}_{3e}$ is

$$\psi_E = \arccos\left(\vec{v}_{3c} \cdot \vec{v}_{3e}\right). \tag{14}$$

Table 1. Summary of the *Diverging Tree* and *Yosemite sequence* optical flow fields results.

Algorithm	*Diverging Tree*		*Yosemite*	
	Ave. error	Std. dev	Ave. error	Std. dev
Horn and Schunck	12.02	11.73	31.69	31.18
Horn and Schunck (modified)	5.20	4.92	18.41	23.37
Proposed method	2.41	1.39	9.41	14.81

The error comparison of the proposed method with the conventional gradient-based method (HORN and SCHUNCK, 1981) and the modified Horn and Schunck approach (CHAUDHURY and MEHROTRA, 1995). The compared methods provide 100% density of optical flow vector. The preprocessing of image sequence is not introduced.

The error comparison of the proposed method with the conventional gradient-based method (HORN and SCHUNCK, 1981) and the modified Horn and Schunck approach (CHAUDHURY and MEHROTRA, 1995) is summarized in Table 1, which lists the average and standard deviation of the angular error. The proposed method records better performance.

5. Conclusions

In this paper, a new approach is proposed to extract the reliable optical flow field, this is the gradient-based spatio-temporal optimization with combination of local and global approach. We combined the spatio-temporal measure of error with the spatio-temporal measure of smoothness in the velocity fields (Eqs. (7) and (8)) to constrain the estimated velocity field $\vec{v} = (u, v)$ (Eq. (9)). The mainly improved points based on Horn and Schunck method are:

1) We use the spatio-temporal error function $\hat{E}_b$ (Eq. (7)) instead of E_b (Eq. (3)) to estimate the errors for the rate of change of image brightness. We evaluate the errors in the conservation equation in a small spatio-temporal volume δv but not at a pixel site (x, y, t). This means that we combine the global and local constraints to estimate optical flow fields.

2) We use the spatio-temporal measure of smoothness $\hat{E}_c$ (Eq. (8), Fig. 2) instead of the spatial measure of smoothness E_c (Eq. (4), Fig. 1) in the velocity fields. We consider the plural velocity planes (Fig. 2) but not one velocity plane (Fig. 1). We introduce the spatio-temporal measure of smoothness in the velocity fields, assuming that the optical flow (u, v) is constant in a small spatio-temporal volume $\delta v = \delta x \cdot \delta y \cdot \delta t$ and neighboring points (among small spatio-temporal volumes δv) on the objects have similar velocities and the velocity field (on plural velocity planes) of the brightness patterns in the image varies smoothly almost everywhere.

3) We use the spatio-temporal global optimization (Eq. (9)) instead of the spatial global optimization (Eq. (2)) to determine optical flow. We minimized the total errors (Eq. (9)) in plural image planes v but not in one image plane s (see Eq. (2)).

Since we adopted the spatio-temporal measure of error (locally averaged) and spatio-temporal measure of smoothness with spatio-temporal optimization, we obtained high resolution and high reliability of the determined optical flow fields compared to the conventional gradient-based global optimization methods. The performance of the proposed method was confirmed by the analysis of two synthetic image sequences and one real image sequence. Further investigation considering the spatio-temporal non-uniform illumination with spatio-temporal optimization combining local and global approach is expected.

The authors wish to thank Dr. Nomura, A. for his critical comments and helpful suggestions. The real image data *Toy car sequence* was taken with the help of Dr. Sakurai, T., who is now with Hokkaido University, we would like to acknowledge that. This work was partly supported by The Sasakawa Scientific Research Grant from The Japan Science Society. This work was partly supported by Creative and Fundamental R&D Program for SMEs of Japan Small Business Corporation which is consigned by New Energy and Industrial Technology Development Organization (NEDO).

Appendix A: The Spatio-Temporal Optimization with Combination Local and Global Approach

The total errors to be minimized can be written as (Eq. (9))

$$\iiint_v \left(\hat{E}_b + \alpha^2 \hat{E}_c\right) dxdydt, \tag{A1}$$

where

$$\hat{E}_b = \iiint_{\delta v} \left(E_x u + E_y v + E_t\right)^2 dxdydt, \tag{A2}$$

$$\hat{E}_c = \left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial u}{\partial y}\right)^2 + \left(\frac{\partial u}{\partial t}\right)^2 + \left(\frac{\partial v}{\partial x}\right)^2 + \left(\frac{\partial v}{\partial y}\right)^2 + \left(\frac{\partial v}{\partial t}\right)^2. \tag{A3}$$

Denoting

$$F\left(x, y, t, u, v, u_x, u_y, u_t, v_x, v_y, v_t\right) = \hat{E}_b + \alpha^2 \hat{E}_c. \tag{A4}$$

The general formulation of the variation integral for the optical flow field reads as

$$\iiint_v F\left(x, y, t, u, v, u_x, u_y, u_t, v_x, v_y, v_t\right)dxdydt \to \text{minimum.} \tag{A5}$$

The corresponding Euler-Lagrange equations are (HORN and SCHUNCK, 1986; JAHNE, 1995):

$$\frac{\partial F}{\partial u} - \frac{\partial}{\partial x}\frac{\partial F}{\partial u_x} - \frac{\partial}{\partial y}\frac{\partial F}{\partial u_y} - \frac{\partial}{\partial t}\frac{\partial F}{\partial u_t} = 0,$$

$$\frac{\partial F}{\partial v} - \frac{\partial}{\partial x}\frac{\partial F}{\partial v_x} - \frac{\partial}{\partial y}\frac{\partial F}{\partial v_y} - \frac{\partial}{\partial t}\frac{\partial F}{\partial v_t} = 0. \tag{A6}$$

The discrete representation of Eq. (A2) can be written as

$$\hat{E}_b = \sum_{\delta v}\left(E_x u + E_y v + E_t\right)^2. \tag{A7}$$

Inserting the Lagrange function (Eq. (A4)) into the Euler-Lagrange differential equation (Eq. (A6)) yields the following differential equation:

$$\sum_{\delta v}\left(E_x u + E_y v + E_t\right)\cdot E_x - \alpha^2\left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial t^2}\right) = 0,$$

$$\sum_{\delta v}\left(E_x u + E_y v + E_t\right)\cdot E_y - \alpha^2\left(\frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2} + \frac{\partial^2 v}{\partial t^2}\right) = 0. \tag{A8}$$

Denoting a Laplacian operator:

$$\nabla^2_{xyt} = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial t^2}.$$

Equation (A8) can be rewritten as

$$\alpha^2\nabla^2_{xyt}u = \sum_{\delta v} E_x^2 u + \sum_{\delta v} E_x E_y v + \sum_{\delta v} E_x E_t,$$

$$\alpha^2\nabla^2_{xyt}v = \sum_{\delta v} E_y E_x u + \sum_{\delta v} E_y^2 v + \sum_{\delta v} E_y E_t. \tag{A9}$$

Using the approximation to the Laplacian (see Appendix B), $\nabla_{xyt}{}^2 u = \overline{u} - u$ and $\nabla_{xyt}{}^2 v = \overline{v} - v$, Eq. (A9) can be written as

$$\left(\alpha^2 + \sum_{\delta v} E_x^2\right)u + \left(\sum_{\delta v} E_x E_y\right)v = \alpha^2\overline{u} - \sum_{\delta v} E_x E_t,$$

$$\left(\sum_{\delta v} E_y E_x\right)u + \left(\alpha^2 + \sum_{\delta v} E_y^2\right)v = \alpha^2\overline{v} - \sum_{\delta v} E_y E_t. \tag{A10}$$

Equation (A10) provides two linear equations in the variables u and v. Solving Eq. (A10) we obtain

$$\Delta = \alpha^2\left(\alpha^2 + \sum_{\delta v} E_x^2 + \sum_{\delta v} E_y^2\right) + \left(\sum_{\delta v} E_x^2\right)\left(\sum_{\delta v} E_y^2\right) - \left(\sum_{\delta v} E_x E_y\right)^2, \tag{A11}$$

$$u = \frac{1}{\Delta}\left\{\alpha^2\left[\left(\alpha^2 + \sum_{\delta v} E_y^2\right)\overline{u} - \left(\sum_{\delta v} E_x E_y\right)\overline{v} - \sum_{\delta v} E_x E_t\right] + \left(\sum_{\delta v} E_x E_y\right)\left(\sum_{\delta v} E_y E_t\right) - \left(\sum_{\delta v} E_x E_t\right)\left(\sum_{\delta v} E_y^2\right)\right\}, \tag{A12}$$

$$v = \frac{1}{\Delta}\left\{\alpha^2\left[-\left(\sum_{\delta v} E_y E_x\right)\overline{u} + \left(\alpha^2 + \sum_{\delta v} E_x^2\right)\overline{v} - \sum_{\delta v} E_y E_t\right] + \left(\sum_{\delta v} E_y E_x\right)\left(\sum_{\delta v} E_x E_t\right) - \left(\sum_{\delta v} E_y E_t\right)\left(\sum_{\delta v} E_x^2\right)\right\}. \tag{A13}$$

From Eqs. (A11)–(A13) we can obtain the solution of optical flow (u, v). When assuming

$$\left(\sum_{\delta v} E_x^2\right)\left(\sum_{\delta v} E_y^2\right) - \left(\sum_{\delta v} E_x E_y\right)^2 = 0,$$

$$\left(\sum_{\delta v} E_x E_y\right)\left(\sum_{\delta v} E_y E_t\right) - \left(\sum_{\delta v} E_x E_t\right)\left(\sum_{\delta v} E_y^2\right) = 0$$

and

$$\left(\sum_{\delta v} E_y E_x\right)\left(\sum_{\delta v} E_x E_t\right) - \left(\sum_{\delta v} E_y E_t\right)\left(\sum_{\delta v} E_x^2\right) = 0 \tag{A14}$$

in a small volume $\delta v = \delta x \cdot \delta y \cdot \delta t$, we get the simple solution:

$$\begin{aligned} \left(\alpha^2 + \sum_{\delta v} E_x^2 + \sum_{\delta v} E_y^2\right)(u - \bar{u}) &= -\left(\sum_{\delta v} E_x^2 \bar{u} + \sum_{\delta v} E_x E_y \bar{v} + \sum_{\delta v} E_x E_t\right), \\ \left(\alpha^2 + \sum_{\delta v} E_x^2 + \sum_{\delta v} E_y^2\right)(v - \bar{v}) &= -\left(\sum_{\delta v} E_y E_x \bar{u} + \sum_{\delta v} E_y^2 \bar{v} + \sum_{\delta v} E_y E_t\right). \end{aligned} \tag{A15}$$

Appendix B: The Approximation to the Laplacian

In HORN and SCHUNCK method (1981), Laplacian is denoted as

1/12	1/6	1/12
1/6	-1	1/6
1/12	1/6	1/12

Fig. 11. The Laplacian is estimated by 9 neighboring points (3 × 3) in one velocity plane.

1/56	2/56	1/56
2/56	4/56	2/56
1/56	2/56	1/56

2/56	4/56	2/56
4/56	-1	4/56
2/56	4/56	2/56

1/56	2/56	1/56
2/56	4/56	2/56
1/56	2/56	1/56

Fig. 12. The Laplacian is estimated by 27 neighboring points ($3 \times 3 \times 3$) in three velocity planes.

$$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}.$$

They use 9 neighboring points (3×3) to approximate the Laplacians of u and v in one velocity plane (Fig. 11). In the proposed spatio-temporal optimization combining local and global approach, however, we need to approximate the Laplacian:

$$\nabla^2_{xyt} = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial t^2}.$$

We use 27 neighboring points ($3 \times 3 \times 3$) to approximate the Laplacians of u and v in three velocity planes (Fig. 12).

REFERENCES

BARRON, J. L., FLEET, D. J. and BEAUCHEMIN, S. S. (1994) Systems and experiment—Performance of optical flow techniques, *Intern. J. Comput. Vis.*, **12**(1), 43–77.

BLACK, M. J. and ANANDAN, P. (1996) The robust estimation of multiple motions: Parametric and piecewise-smooth flow fields, *Comput. Vision and Image Understanding*, **63**(1), 75–104.

CHAUDHURY, K. and MEHROTRA, R. (1995) A trajectory-based computational model for optical flow estimation, *IEEE Trans. on Robotics and Automation*, **11**(5), 733–741.

GLAZER, F., REYNOLDS, G. and ANANDAN, P. (1983) Scene matching by hierarchial correlation, *IEEE Computer Society*, 432–441.

HORN, B. K. P. (1986) *Robot Vision*, MIT Press, Cambridge, MA.

HORN, B. K. P. and SCHUNCK, B. G. (1981) Determining optical flow, *Artificial Intell.*, **17**, 185–203.

JAHNE, B. (1995) *Digital Image Processing*. Springer-Verlag, Berlin.

KEARNEY, J. K., THOMPSON, W. B. and BOLEY, D. L (1987) Optical flow estimation: an error analysis of gradient-based methods with local optimization, *IEEE Trans. Pattern Anal. Machine Intell.*, **9**, 229–244.

NAGEL, H. H. and ENKELMANN, W. (1986) An investigation of smoothness constraints for the estimation of displacement vector fields from image sequences, *IEEE Trans. Patt. Anal. Mach. Intell.*, **8**, 565–593.

NOMURA, A., MIIKE, H. and KOGA, K. (1991) Field theory approach for determining optical flow, *Pattern Recog. Lett.*, **12**, 183–190.

SINGH, A. (1991) *Optical Flow Computation: A Unified Perspective*, IEEE Computer Society Press, Los Alamitos, California.

VERRI, A., GIROSI. F. and TORE, V. (1990) Differential techniques for optical flow, *J. Opt. Soc Am.*, **A7**, 912–922.

ZHANG, L., SAKURAI, T. and MIIKE, H. (in press) Detection of motion fields under spatio-temporal non-uniform illumination, *Image and Vision Computing*.