Estimation of Moving Information for Tracking of Moving Objects

Sung-Kwan Kang, Jong-An Park*†
Electronis · Information & Communications Engineering, Chosun University
Sang-Hwa Jeong
Mechanical Engineering, Chosun University

Tracking of moving objects within video streams is a complex and time-consuming process. Large number of moving objects increases the time for computation of tracking the moving objects. Because of large computations, there are real-time processing problems in tracking of moving objects. Also, the change of environment causes errors in estimation of tracking information. In this paper, we present a new method for tracking of moving objects using optical flow motion analysis. Optical flow represents an important family of visual information processing techniques in computer vision. Segmenting an optical flow field into coherent motion groups and estimating each underlying motion are very challenging tasks when the optical flow field is projected from a scene of several moving objects independently. The problem is further complicated if the optical flow data are noisy and partially incorrect. Optical flow estimation based on regulation method is an iterative method, which is very sensitive to the noisy data. So we used the Combinatorial Hough Transform (CHT) and Voting Accumulation for finding the optimal constraint lines. To decrease the operation time, we used logical operations. Optical flow vectors of moving objects are extracted, and the moving information of objects is computed from the extracted optical flow vectors. The simulation results on the noisy test images show that the proposed method finds better flow vectors and more correctly estimates the moving information of objects in the real time video streams.

Key Words : Object Tracking, Optical Flow, Hough Transform, Voting

1. Introduction

Three-dimensional (3D) motion estimation is of relevance to many problems related to dynamic scene analysis such as 3D object reconstruction (Prazdny, 1983; Adiv, 1989), object tracking (Burt et al. 1989; Broida and Chellappa, 1989), and robot navigation (Nelson and Aloimonos, 1989; Subbarao, 1990; M. C. Han, K. S. Hong, J. K. Kim, 1996). One way to estimate the 3D motion is to evaluate its perspective projection on the image plane. This is usually called "velocity field", and represents the apparent velocity of the image pixels from one frame to the next.

One of the most notable approaches to find velocity field is based on the estimation of a measure of the change of image brightness in the frame sequence, commonly referred to as optical flow. Optical flow represents an approximation of the velocity field which is a purely geometric concept. Verri and Poggio (1989), and Nagel (1989) have analysed operating conditions for the equality of optical flow and velocity field. In many applications optical flow is a sufficient approximation of the velocity field and can be reasonably employed in its place.

† First Author
* Corresponding Author,
E-mail : japark@chosun.ac.kr
TEL : +82-62-230-7064 ; **FAX** : +82-62-232-6776
Electronics · Information & Communications Engineering, Chosun University. (Manuscript **Recevied** April 14, 2000; **Revised** December 23, 2000)

Different approaches for optical flow estimation exhibit different behavior with respect to discontinuities and for different types of motion, depending on the techniques used. In the literature two approaches for optical flow estimation can be identified: (1) regularization-based approaches and (2) multiconstraint-based approaches.

Regularization-based approaches (Horn and Schunck, 1981; Nagel, 1983; Nagel and Enkelmann, 1989; Schunck, 1989) consider velocity field estimation as an ill-posed problem. Solutions are obtained by minimizing a function where a smoothness constraint is appropriately weighted to regularize the solution. Usually, these methods lead to iterative solutions and the velocity is evaluated at every point of the image. Most of these methods yield a "dense" optical flow field, in the sense that the estimation process assigns a vector to every pixel in the image, and not only to the pixels of object boundaries. Drawbacks of these approaches are related to the fact that difficulties occur at the regions of object occlusions. Further, the depth of propagation of the field depends on the number of iterations used and on the weighting factor of the regularizing.

Multiconstraint-based approaches (Cafforio and Rocca, 1979; Campani and Verri, 1990; Nesi, Delbimbo and Sanz, 1991) to optical flow estimation are based on the principle that it is possible to define a set of constraint equations for the point under consideration. This set of equations is usually solved by using numerical methods for the inversion or pseudo-inversion of the coefficient matrix, or by using least-squares techniques. Traditional numerical methods, like the least -squares technique, are averaging methods, and are thus susceptible to errors in two important cases, namely, that of occlusion and of noise. In the case of occlusion, the two objects contribute conflicting velocities. at the border between two moving objects, and taking their average yields a less satisfactory optical flow estimation. which deviates from both. Noise also enhances the damaging effect in the case where the solution is found by averaging. The contribution to the solution from noise has the same weight as that from the object, and with a significant presence of noise, a considerable deviation from the actual optical flow can occur.

Another drawback of these optical flow based methods is that the detected motion boundary is not precise because the motion is not homogeneous near the motion boundary. Furthermore, the object contour is not determined when the motion of the object is similar to that of its neighbor. Therefore, several methods have been proposed which use not only optical flow but also other information such as color and edge (Thompson, 1980; Gemen, 1985; Black, 1992; Etoh, 1993; Nagao, 1993;).

In this paper, a method based on the multiconstraint-based approach is presented, which evaluates optical flow using optical flow constraint (OFC) equations in the neighborhood of each pixel. The solution is derived with the help of Combinatorial Hough Transform (CHT) (Leavers, Ben-Tzvi and Sandler, 1989; Ben-Tzvi and Sandler, 1990) and with the help of vote accumulation, which avoids drawbacks associated with the least-squares method. Calculation of many points along the constraint lines is also avoided by taking into consideration the transformed slope-intercept parameter domain. For reducing the operating time, the logical operation method for computing E_x, E_y, E_t (brightness gradient in the x, y and t directions respectively between two input images) is used. We also perform the extraction in the real moving area in the image for reducing the operating time of the CHT.

2. Extracting the Brightness Derivatives

2.1 Detecting the moving area

Combinatorial Hough Transform (CHT) requires a lot of computations. To reduce the computations, the moving area is detected by differentiating the two consecutive input images as:

$$D(x, y) = ABS[I_{t+1}(x, y) - I_t(x, y)] \quad (1)$$

where $I(t)$ is the first image at time t, $I(t+1)$ is

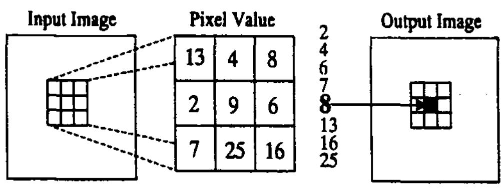


Fig. 1 Median Filter

Fig. 2 Detected moving area

the second image at time $t+1$ and $ABS[\ ]$ means the absolute value.

Median filter (Fig. 1), erosion and dilation processing are used for removing the noise. Fig. 2 shows the detected moving area.

2.2 Extracting the brightness derivatives

Assuming that the image brightness $E(x(t), y(t), t)$ is stationary with respect to time (i. e. $dE/dt=0$), the flow of its feature pattern can be modeled by a sort of continuity equation.

$$E_x u + E_y v + E_t = 0 \tag{2}$$

where the abbreviation for partial derivatives of the image brightness has been introduced, and u and v correspond to dx/dt and dy/dt respectively, and represent the components of velocity vector V of the feature pattern along the x and y directions on the image plane respectively (the optical flow components). Eq. (2) is usually called the "Optical Flow Constraint" (OFC), and estimation methods based on OFC equation are commonly referred to as gradient-based methods.

E_x, E_y, E_t are computed using Eq. (3).

$$E_x \approx \frac{1}{4}\{E_{i+1,j,k} - E_{i,j,k} + E_{i+1,j+1,k} - E_{i,j+1,k} + E_{i+1,j,k+1} - E_{i,j,k+1} + E_{i+1,j+1,k+1} - E_{i,j+1,k+1}\}$$

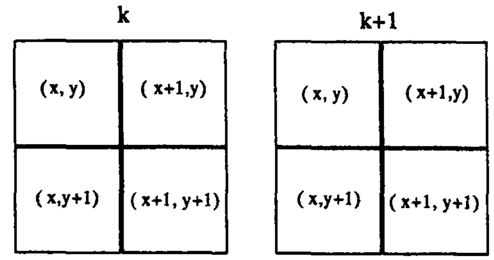


Fig. 3 Illustration of three partial derivatives of image brightness

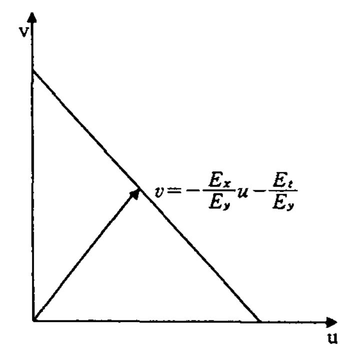


Fig. 4 Representation of a constraint line on the (u, v) plane

$$E_y \approx \frac{1}{4}\{E_{i+1,j,k} - E_{i,j,k} + E_{i+1,j+1,k} - E_{i+1,j,k} + E_{i,j+1,k+1} - E_{i,j,k+1} + E_{i+1,j+1,k+1} - E_{i+1,j,k+1}\}$$

$$E_t \approx \frac{1}{4}\{E_{i,j,k+1} - E_{i,j,k} + E_{i+1,j,k+1} - E_{i+1,j,k} + E_{i,j+1,k+1} - E_{i,j+1,k} + E_{i+1,j+1,k+1} - E_{i+1,j+1,k}\} \tag{3}$$

The derivatives of brightness are estimated from the discrete set of image brightness measurements available. It is important that the estimates of E_x, E_y and E_t be consistent. That is, they should all refer to the same point in the image at the same time. While there are many formulae for approximating differentiation, we use a matrix which gives us an estimation of E_x, E_y and E_t at one point of pixels formed by four-neighborhood measurements. The relationship in space and time between these measurements is shown in Fig. 3. Each of the estimation is the average of four first differences taken over adjacent measurements.

The OFC equation cannot provide a unique solution by itself. In fact, the OFC can be regard-

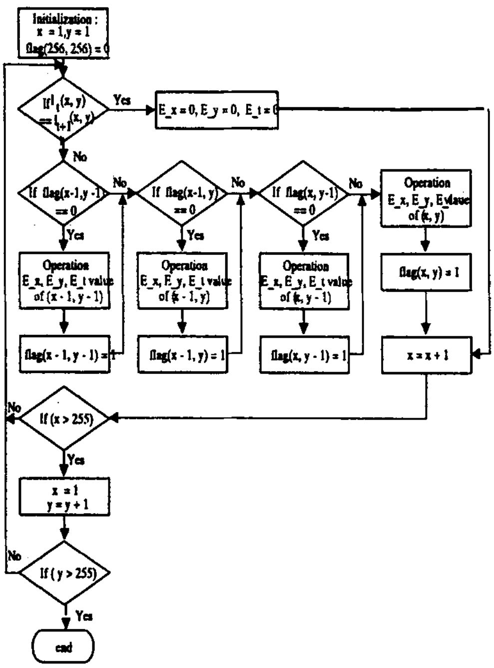
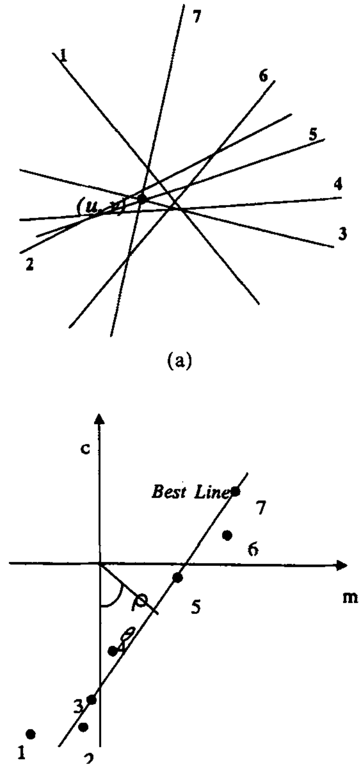


Fig. 5 Algorithm for logical comparison



(a)

(b)

Fig. 6 Constraint line parameterization (a) constraint line with the estimated optical flow vector components (u, v) in the (u, v) plane, (b) corresponding best line in the parameter plane

ed as the equation of a line in the (u, v) plane.

$$v = mu + c \tag{4}$$

where $m = -E_x/E_y$ is the slope and $c = -E_t/E_y$ is the intercept. Any point along this line is a possible solution for the optical flow estimation problem.

In this paper, we used the logical comparison method for improving the gradient operation speed. The operation algorithm for logical comparison is outlined in Fig. 5.

3. Extracting Velocity Vectors

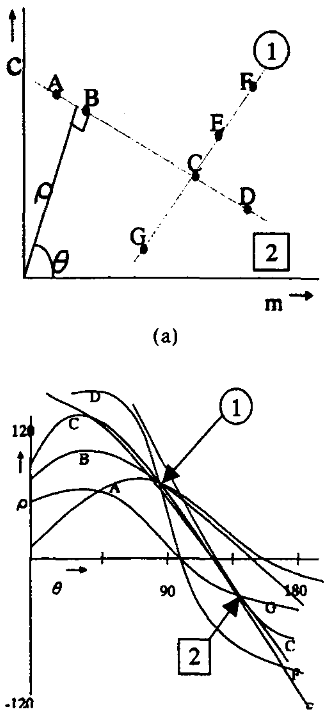
3.1 Transform to the (m, c) plane

A multiconstraint solution based on the OFC followed by the vote accumulation method identifies the most likely solution as the point (u, v) where most of the constraint lines lie in the vicinity of each pixel intersect. By means of this approach, the characteristics of each constraint line are transformed from the (u, v) plane to the slope-intercept plane (m, c), where each constraint line is represented by a point according to Eq. (4). The requirement of a common intersection point in the (u, v) plane of a set of constraint lines is equivalent to the requirement of collinearity of their corresponding points in the parameter plane (See Fig. 6). The estimation of optical flow at each pixel is thus reduced to finding the best line that matches the pattern of points corresponding to the constraint lines (around each pixel) in the parameter plane. The best line is the line on which the largest number of points reside and not the line that gives the minimum accumulative distance to the points.

In practice, we consider blocks of size 5×5, with 2 pixel overlap.

3.2 Transform to the (ρ, θ) plane and vote accumulation

Votes are cast using a Hough transform version, namely the Combinatorial Hough Trans-



(b)

Fig. 7 Comparison between (m, c) plane and (ρ, θ) plane, (a) (m. c) plane (b) Transformation into (ρ, θ) plane

form. Each couple of points (m_1, c_1) and (m_2, c_2) in the slope-intercept plane corresponding to a couple of constraint lines, adds a vote to a monodimensional accumulation histogram of the best line θ.

$$\theta = \arctan\left(-\frac{m_2 - m_1}{c_2 - c_1}\right) \tag{5}$$

$$\rho_i = m_i \cos(\theta_{\max}) + c_i \sin(\theta_{\max}) \tag{6}$$

Therefore, according to the multiconstraint approach on an $N \times N$ area, there are $(N^4 - N^2)/2$ couples of constraint equations and solutions. This could lead to an asymptotical complexity equal to I^2N^2 ($I \times I$: the size of the image). This is simplified by considering only the combinations of the constraint lines associated with the pixels in the $N \times N$ area and the constraint line of the center of the multi-point area.

Thus, $(N^2 - 1)$ pairs of equations and, hence, N^2 votes are obtained. The histogram in θ is inspected to find the most probable value, $\theta_{\max}$

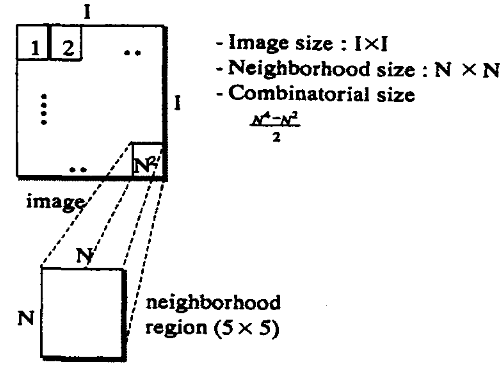
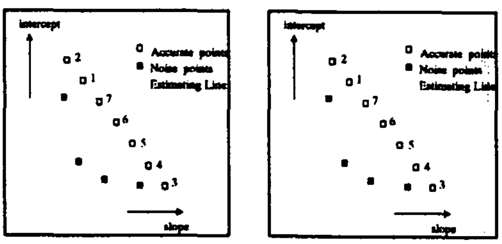


Fig. 8 Illustration of the computed complexity



(a) (b)

Fig. 9 Comparison of least-squares method and vote accumulation method, (a) vote accumulation method, (b) least-squares method

(the value corresponding to the peak of the histogram). By using that value, a second stage of N^2 votes is used to define another histogram for the other line parameter.

We computed $\rho_{\max}$ by the following procedure :

- Compute ρ_{mean}.n

$$\rho_{mean} = \frac{1}{N^2}\sum_{i=1}^{N^2}\rho_i$$

- Search ρ' within defined range in ρ set.

$$\rho' = \{i \mid \rho_{mean} \cdot \alpha < \rho_i < \rho_{mean} \cdot \beta \,;\, i = 1, \cdots, N^2\} \tag{7}$$

where α is 0.4 (minimum factor) and β is 1.5 (maximum factor) experimentally.

- $\rho_{\max}$ is the mean of ρ' within the values defined in the second process.

$$\rho_{\max} = \frac{1}{N(\rho')}\sum_{i=1}^{N(\rho')}\rho'_i \tag{8}$$

where $N(\rho')$ is the size of the set ρ'.

Figure 9 shows the advantage of vote accumulation method.

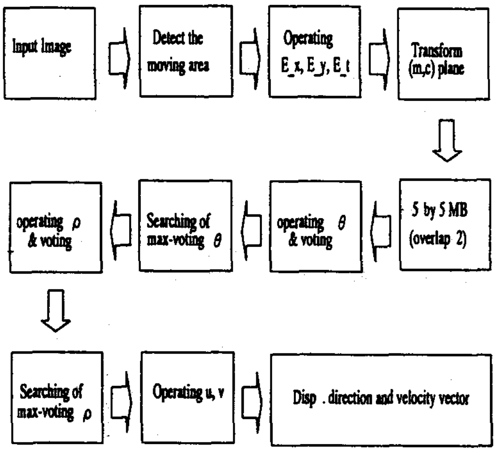


Fig. 10 Diagram of the proposed algorithm

3.3 Computing the velocity vector

The best approximation of the best line is defined by (ρ_{max}, θ_{max}). Therefore, the optical flow at each pixel is directly derivable from these line parameters.

$$u = \cot(\theta_{max}),\ v = \frac{\rho_{max}}{\sin(\theta_{max})} \tag{9}$$

3.4 Extracting the moving object

For extracting the moving area accurately, we used the mask operation between the differential image and optical flow. Then, the information of the moving object computed in section 3.3 is displayed in this area.

$$C = D \wedge O \tag{10}$$

D : differential-image region
O : optical flow region
$\wedge$: mask operation (for extracting region)

The diagram of the proposed algorithm in this paper is shown in Fig. 10.

4. Experimental Results

For the simulation of the proposed algorithm, a two-frame sequence (256×256 gray-level) as shown in Fig. 11, is used. There are several strong edges in the background of the image in the sequence. The image sequence is noisy which causes errors in the computation of optical flow vectors. The results of the proposed method are compared with the method I (Horn and Schunck, 1981) and method II (Campani and Verri, 1990).

The three methods are compared in terms of operation time and computational complexity. The simulation results are shown in Table 1.

Method I is the regularization based method for finding the optical flow vectors. Method I gives accurate results, but it is very slow due to its iterative nature as compared to the other method. For 1500 iterations, the proposed method is faster

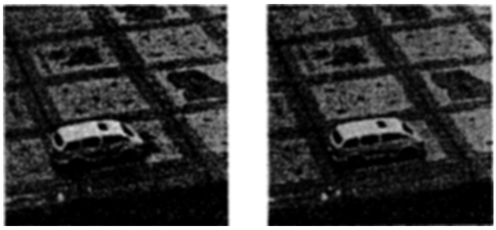
Table 1 Comparison of the operation time and complexity

Algorithm	Iteration	Complexity	Operation time
Horn & Schunck	Yes	$I_t I^2$	$61.87 \times I_t$
Campani & Verri	No	$N^2 I^2$	843.66
Proposed method	No	$N^2 I_r^2$	286.96

$I \times I$: the size of the image
$I_r \times I_r$: the size of the moving region
$N \times N$: the size of the neighborhood
I_t : the number of iterations

Table 2 Comparison of the moving information value

	Distance	Direction
Real Value	17.22	9.5
Horn & Schunck	25.02	2.3
Campani & Verri	18.44	12.5
Proposed method	16.27	10.61



(a) (b)

Fig. 11 Input Sequence, (a) First frame, (b) Second frame

(a)

(b)

(c)

Fig. 12 Comparison of the extracted optical flow, (a) Horn and Schunck's method, (b) Campani and Verri, (c) Proposed method

Fig. 13 Direction of the moving object

than that of method I. The complexity of the proposed method is also less compared to method Ⅰ.

Method Ⅱ is one of the multi-point based algorithm. But, method Ⅱ has also greater computational complexity compared to the proposed method. Also, method Ⅱ has the local minimum problem at some part of the image.

5. Conclusion

In this paper, we discussed the problem of tracking of moving objects in a video stream. There are many methods for tracking of moving objects in the literature, but we discussed the popular technique of optical flow for detection of moving objects. Optical flow represents an important family of visual information processing techniques in computer vision. Optical flow finds the velocity vectors at each pixel in the entire video scene. However, optical flow-based methods need a lot of computations, and are sensitive to noise. In this paper, we proposed a new method based on CHT and on voting for improving the accuracy and reducing the computation time. We compared the proposed method with some classical methods on the noisy test image sequence. The simulation results show that the proposed method improves the accuracy of finding the optical flow vectors, and extracts the moving information of objects more accurately. Also, the proposed method is faster than the classical methods.

Acknowledgment

This paper was supported by research funds from the Factory Automation Research Center for Parts of Vehicles(FACPOV) in Chosun University, Kwangju, Korea.

References

Adiv, G., 1989, "Inherent Ambiguities in Recovering 3-D Motion and Structure from a Noisy Field," *IEEE Trans. Pattern Analysis Mach. Intell.* 11(5), pp. 477~489.

Ben-Tzvi, D. and Sandler, M., 1990, "A Combinatorial Hough Transform," *Pattern Recognition, Lett.* 11, pp. 167~174.

Black, M. J., 1992, "Combining Intensity and Motion for Incremental Segmentation and Tracking over Long Image Sequence," *Proc. ECCV'92*, pp. 485~493.

Broida, T. J. and Chellappa, R., 1989, "Experiments and Uniqueness Results on Object Structure and Kinematics from a Sequence of Monocular Images," *Proc. IEEE Workshop on Visual Motion*, Irvine, California, U. S. A., pp. 21~30.

Burt, P. J., et al., 1989, "Object Tracking with a Moving Camera," *Proc. IEEE Workshop on Visual Motion*, Irvine, California, U. S. A., pp. 2~12.

Cafforio, C. and Rocca, F., 1979, "Tracking Moving Objects in Television Images," *Signal Process.* 1, pp. 133~140.

Campani, M. and Verri, A. 1990, "Computing Optical Flow from an Overconstrained System of Linear Algebraic Equations," *Proc. 3rd IEEE Int. Conf. on Computer Vision*, pp. 22~26.

Del Bimbo, A. and Nesi, A., 1993, "Real-Time Optical Flow Estimation," *in Proc. 1993 IEEE Systems, Man and Cybernetics Conf., (Le Touquet, France)*, pp. 17~20.

Etoh, M. et. al., 1993, "Segmentation and 2D Motion Estimate by Region Fragments," *Proc. 4th Int. Conf. Computer Vision*, pp. 192~199.

Gemen, S. et. al., 1985, "Stochastic Relaxation, Gibbs Distribution, and Bayesian Restoration of Images," *IEEE Trans PAMI.* 6(6), pp. 721~741.

Han, M. C., Hong, K. S. and Kim, J. K., 1996, "The Position Measurement of a Body Using 2D Vision Sensors," *'96 Proceedings of the KSME Spring Annual Meeting (A)*, pp. 270~275.

Horn, B. K. P. and Schunck, B. G., 1981, "Determining Optical Flow," *Artif. Intell.* 17, pp. 185~203.

Leavers, V., Ben-Tzvi, D. and Sandler, M., 1989, "A Dynamic Combinatorial Hough Transform for Straight Lines and Circles," *Proc. Alvey Vision Conf. Manchester*, U. K.

Nagao, K. et. al., 1993, "Detecting Contours in Image Sequences," *IEICE Trans. Information and System.* E76-D(10), pp. 1162~1173.

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

Nagel, H. H. 1989, "On a Constraint Equation for the Estimation of Displacement Rates in Image Sequences," *IEEE Trans. Pattern Analysis Mach. Intell.* 11(1), pp. 13~30.

Nagel, H. H. and Enkelmann, W. 1984, "Towards the Estimation of Displacement Vector Fields by 'Oriented Smoothness' Constraints," *Proc. 7th IEEE Int. Conf. on Pattern Recognition*, pp. 6~8.

Nelson, R. C. and Aloimonos, J., 1989, "Obstacle Avoidance Using Field Divergence," *IEEE Trans. Pattern Analysis Mach. Intell.* 11(10), pp. 1102~1106.

Nesi, P., Delbimbo, A. and Sanz, J. L., 1991, "Multiconstraints-Based Optical Flow Estimation and Segmentation," *Int. Workshop on Computer Architecture for Machine Perception*, Paris, pp. 419~426.

Prazdny, K., 1983, "On the Information in Optical Flow," *Computer Vision Graphics Image Process.* 23, pp. 239~259.

Schunck, B. G., 1989, "Image Flow Segmentation and Estimation by Constraints Line and Clustering," *IEEE Trans. Pattern Analysis Mach. Intell.* 11(10), pp. 1010~1027.

Subbarao, M., 1990, "Bounds on Time-to-Collision and Rotation Component from First-Order Derivatives of Image Flow," *Computer Vision Graphics Image Process.* 50, pp. 329~341.

Thompson, W. B., 1980, "Combining Motion and Contrast for Segmentation," *IEEE Trans. PAMI*, 2(6), pp. 543~549.

Verri, A. and Poggio, T., 1989, "Motion Field and Optical Flow: Qualitative Properties," *IEEE Trans. Pattern Analysis Mach. Intell*. 11(5), pp. 490~498.