Proceedings of the 13th
KACC, October 1998

An Optical Flow Based Time-to-Collision Predictor

T. Yamaguchi, H. Kashiwagi and H. Harada

Faculty of Engineering, Kumamoto University
Kurokami 2–39–1, Kumamoto 860-8555, Japan
Tel: +81-96-342-3737, Fax: +81-96-342-3729
E-mail: kuchi@mech.kumamoto-u.ac.jp

Abstract

This paper describes a new method for estimating time-to-collision which exhibits high tolerance to noise contained in camera images.

Time to collision (TTC) is one of the most important parameters available from a camera attached to a mobile machine. TTC indirectly stands for the translation speed of the camera and is usually calculated either from successive images or optical flow by using intimate relationship between TTC and flow divergence. In most cases, however, it is not easy to get accurate optical flow, which makes it difficult to calculate TTC.

In this study it is proved that if the target has a smooth surface, the average of divergence over any point-symmetric region on the image is equal to the divergence of the center of the region. It means that required divergence can be calculated by integrating optical flow vectors over a symmetric region. It is expected that in the process of the integration, accidental noise is canceled if they are independent of optical flow and the motion of the camera.

Experimental results show that TTC can be estimated regardless of the surface condition. It is also shown that influence of noise is eliminated as the area of integration increases.

Key words: navigation, time-to-collision, optical flow, intelligent sensor

1. Introduction

It is necessary for smart vehicles and robots navigating in a structurally unconstrained environment to utilize visual scene of their environment which contains no remarkable landmark. In such cases, it is important to know the direction of their own motion, and it is often required to predict a time to reach an object in front of them. Temporal images acquired by a visual sensor attached to a mobile robot contain considerable information about its own motion. In earlier researches, it was shown that this information can be extracted from a small difference between successive images[1, 2]. It was, however, difficult to extract such subtle change from a series of images in realtime.

Optical flow field taken by a moving camera is a consequence of its motion relative to the observed object. In case of navigation, most of surrounding objects are supposed to be stationary and often composed of flat surfaces, e.g., wall or floor. If the observed object is a planar wall normal to the optical axis of the camera, it is proved that time-to-collision (TTC) to the wall can be estimated from the divergence of optical flow[3, 4] as

$$T_{\mathrm{TTC}} = \frac{2}{\mathrm{div}\boldsymbol{v}}. \tag{1}$$

Note that TTC is defined as the time to reach the object if approaching velocity does not change, and it is an essential dynamic motion parameter like velocity and useful in planning navigation[5].

According to eq. (1), TTC can be calculated only from a series of motion images without knowing any other parameters including the distance to the wall and focal length of camera lens. This method has been limited to an object with a surface normal to the camera. In this paper, we try to expand this algorithm to estimate TTC to a surface with more arbitrary shape.

Optical flow analysis is also expected to play an important role in the estimation of direction of self motion[6]. Motion direction is encoded in an image as an emergent point of motion which is defined as a staring point of all flow vectors. This singular point is often called focus of expansion (FOE). We will show that to estimate TTC to a generalized surface, it is desirable to estimate the FOE at first.

We have already proposed a method of calculating optical flow vectors from video rate images[7]. But optical flow computed from images contains inevitable noise so that TTC estimated from the divergence of optical flow may become unstable. To avoid this problem, we propose a noise immune method of estimating flow divergence at a point using a number of flow vectors in the vicinity of the point.

2. Preliminaries

System of Coordinates and Relative Motion

The camera coordinate system adopted here is a model of a pin-hole camera. In this model, each object point

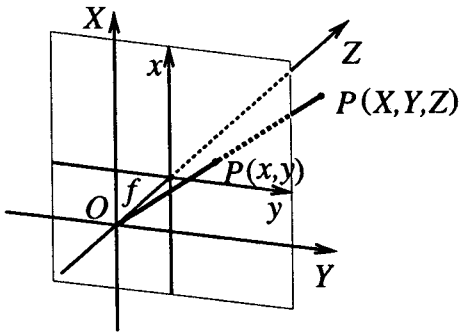


Fig. 1. System of coordinates for the modeled camera.

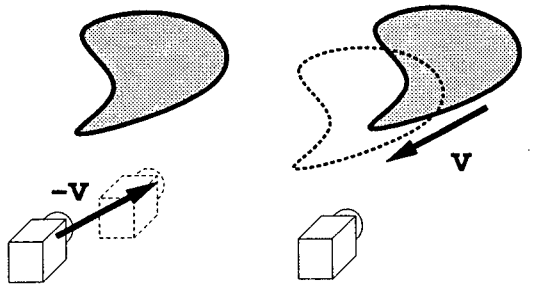


Fig. 2. Relative motion. (a) real motion. (b) apparent motion viewed from the camera.

(X, Y, Z) in space is mapped by optical system to a point (x, y) on an image plane. The pin-hole (the node point of a lens) is located at the origin of the coordinate system, and the relation between object and image points are described as

$$\begin{cases} x &= f\dfrac{X}{Z} \\ y &= f\dfrac{Y}{Z}, \end{cases} \tag{2}$$

where f is the distance between the pin-hole and the image plane (the focal length of the lens). For convenience, the image plane is modeled as located on the same side of the object to the origin (see fig. 1).

When a camera is approaching a stationary object (see fig. 2(a)), the object and its surroundings may be observed by the camera as if they were approaching the camera. It is equivalent to the relative motion, where the camera coordinate system is stationary (see fig. 2(b)). In this paper, coherent translation of all surrounding objects and camera's translation are regarded as equivalent translation, for convenience of explanation.

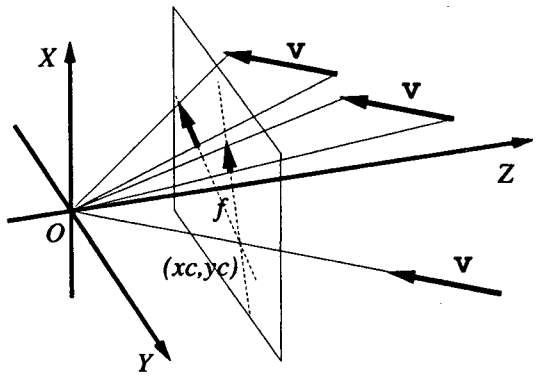


Fig. 3. Relationship between optical flow field and FOE.

Estimation of FOE

When a given point (X, Y, Z) in object space translates at a velocity $(V_X\ V_Y\ V_Z)^T$, the velocity $(v_x\ v_y)^T$ at the corresponding point (x, y) on the image plane is derived from eq. (2) as

$$\begin{cases} v_x &= \dfrac{1}{Z}(fV_X - xV_Z) \\ v_y &= \dfrac{1}{Z}(fV_Y - yV_Z). \end{cases} \tag{3}$$

It is easy to prove from eq. (3) that the time trajectories of projected points on the image caused by translation compose a set of lines sharing a single point. This point is called focus of expansion (FOE) or emergent point of motion, and is denoted as (x_c, y_c), where

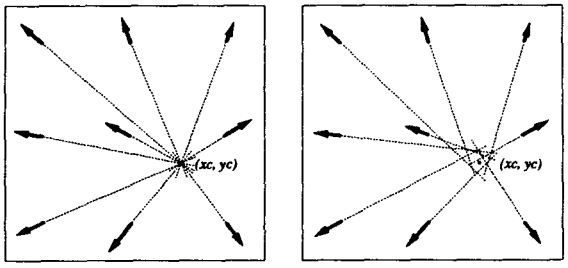
$$\begin{cases} x_c &= f\dfrac{V_X}{V_Z} \\ y_c &= f\dfrac{V_Y}{V_Z}. \end{cases} \tag{4}$$

It is important that optical flow vectors always diverge from an unique FOE at every moment regardless of distance to or shape of the object. (see fig. 3)

If the translation vector lies on a line passing the origin, the projection of this vector on the image corresponds to the FOE. Therefore, to determine the direction of camera motion, we need to know the FOE.

From the above consideration, to measure the FOE it is necessary to estimate the most shared point by extrapolated lines of velocity vectors. One crucial problem is, however, that measured optical flow vectors from actual images will be stained with noise, and this makes it difficult to estimate a single sharing point because all extrapolated lines do not always intersect one point (see fig. 4).

In this case, optimal estimate can be determined as a point which has the least sum of square distance from the



(a) (b)

Fig. 4. Schematic diagram of flow vectors and their extrapolated lines in case of ideal (a) or not (b).

lines. The extrapolated line of a velocity vector $(v_{xi}\ v_{yi})^T$ at a point (x_i, y_i) is described as

$$v_{yi}(x - x_i) = v_{xi}(y - y_i). \tag{5}$$

The distance between a point (x, y) and the line is

$$d_i = \frac{|v_{yi}(x - x_i) - v_{xi}(y - y_i)|}{\sqrt{v_{xi}^2 + v_{yi}^2}}. \tag{6}$$

In this research, we adopt an FOE estimator minimizing the square sum of distance d_i, which is described as

$$\begin{pmatrix} \hat{x}_c \\ \hat{y}_c \end{pmatrix} = \begin{pmatrix} \sum_i \frac{v_{yi}^2}{v_{xi}^2 + v_{yi}^2} & -\sum_i \frac{v_{xi}v_{yi}}{v_{xi}^2 + v_{yi}^2} \\ -\sum_i \frac{v_{xi}v_{yi}}{v_{xi}^2 + v_{yi}^2} & \sum_i \frac{v_{xi}^2}{v_{xi}^2 + v_{yi}^2} \end{pmatrix}^{-1} \begin{pmatrix} \sum_i \frac{v_{yi}^2 x_i - v_{xi}v_{yi}y_i}{v_{xi}^2 + v_{yi}^2} \\ \sum_i \frac{v_{xi}^2 y_i - v_{xi}v_{yi}x_i}{v_{xi}^2 + v_{yi}^2} \end{pmatrix}. \tag{7}$$

3. Extension of TTC Estimation

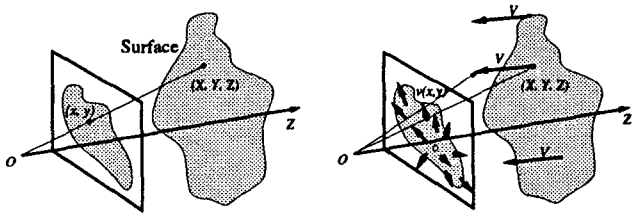
Estimation of TTC to a Curved Surface

If the surface shape of an object visible from a camera has no occlusion, the depth to the surface corresponding to any point on the image can be described as a function of image coordinates (see fig. 5), i.e.,

$$Z = Z(x, y). \tag{8}$$

When the camera translates at velocity $(V_X\ V_Y\ V_Z)^T$, the divergence of optical flow field on the image plane can be derived from eq. (3) as

$$\begin{aligned} \operatorname{div} \boldsymbol{v} &= -\frac{2V_Z}{Z} + \frac{\partial(1/Z)}{\partial x}(fV_X - xV_Z) \\ &\quad + \frac{\partial(1/Z)}{\partial y}(fV_Y - yV_Z). \end{aligned} \tag{9}$$



(a) (b)

Fig. 5. (a) surface with arbitrary shape. (b) flow vectors induced from camera motion.

From eq. (4), the second and third term of the right hand of eq. (9) at FOE (x_c, y_c) becomes zero and TTC at this point is described as

$$T_{\mathrm{TTC}} = -\frac{Z_c}{V_Z} = \frac{2}{\operatorname{div} \boldsymbol{v}_c}, \tag{10}$$

where Z_c and $\operatorname{div} \boldsymbol{v}_c$ are the depth and the divergence at FOE, respectively.

This equation implies:

- TTC at FOE is described as double of the inverse of flow divergence.
- This estimator does not depend on the shape of a surface nor the focal length of a lens and can be calculated only from the optical flow field observed on the image plane.

Divergence Estimation from Flow Vectors with Noise

To calculate divergence of a flow vector field, it is necessary to carry out a kind of differentiation of flow vector. In most cases, however, measured optical flow is apt to contain some noise caused by incompleteness of video apparatus or signal conditioner, for example. In general, random error in flow vector strongly affects the accuracy of estimated derivative. In this section, we seek for a more stable estimator of flow divergence.

It can be assumed that there exists a plane which can approximate the adjacent region on the surface around a given point (see fig. 6). This plane can be selected as a tangential plane or least-square fitted plane. If the plane is described as

$$pX + qY + Z = Z_0, \tag{11}$$

from eq. (2) we can get

$$\frac{1}{Z} = \frac{px + qy + f}{fZ_0}. \tag{12}$$

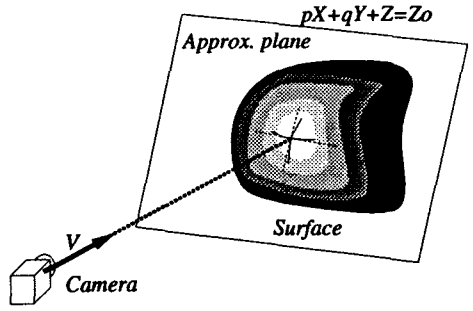


Fig. 6. Planar approximation of the object surface adjacent to the viewed point.

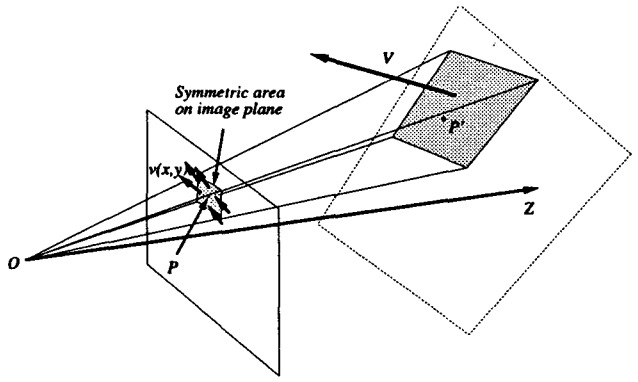


Fig. 7. Perspective projection of an inclined plane onto an image plane.

Velocity vector on the image plane is derived by substituting eq. (12) to eq. (3).

$$\begin{cases} v_x &= \frac{1}{fZ_0}(px+qy+f)(fV_X - xV_Z) \\ v_y &= \frac{1}{fZ_0}(px+qy+f)(fV_Y - yV_Z). \end{cases} \quad (13)$$

By differentiating eq. (13) by x and y respectively, flow divergence proved to be expressed as the following form,

$$\mathrm{div}\, \boldsymbol{v} = -2\frac{V_Z}{Z_0} + \frac{1}{Z_0}(pV_X + qV_Y) - 3\frac{V_Z}{fZ_0}(px+qy). \quad (14)$$

Eq. (14) indicates that the divergence of flow field is a first order function of the location (x, y). It means that the average of the divergence over a point-symmetric region S becomes the divergence at the center of the region P, i.e.,

$$\mathrm{div}\, \boldsymbol{v}(x,y) = \frac{\iint_S \mathrm{div}\, \boldsymbol{v}\ \mathrm{d}x'\mathrm{d}y'}{\iint_S \mathrm{d}x'\mathrm{d}y'}, \quad (15)$$

where the point $P(x, y)$ represents the center of region S (see fig. 7).

By using Gauss' divergence theorem, it is shown that the area integration of divergence over a region can be expressed as a line integration around the boundary loop of the region. Therefore, the following relationship is held,

$$\mathrm{div}\, \boldsymbol{v}(x,y) = \frac{\oint_{\partial S}(v_x \mathrm{d}y' - v_y \mathrm{d}x')}{\mathrm{Area}(S)}. \quad (16)$$

The integration region can be selected as whatever figure it satisfies symmetry, but in practice the region should be limited to a set of rectangles. This is because the line integration can be represented by summation of an outgoing component of each velocity, i.e. either $\pm v_x$ or $\pm v_y$,

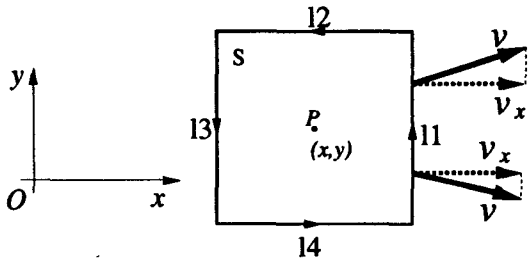


Fig. 8. Principle of divergence estimation by a line integration along a rectangle boundary.

as shown in eq. (17) (see fig. 8).

$$\oint_{\partial S}(v_x \mathrm{d}y' - v_y \mathrm{d}x') = \int_{l1} v_x \mathrm{d}y' + \int_{l2} v_y \mathrm{d}x' - \int_{l3} v_x \mathrm{d}y' - \int_{l4} v_y \mathrm{d}x'. \quad (17)$$

In fig. 9, divergence can be calculated from region S_1, S_2, S_3, and S_4. To utilize measured vectors as many as possible, we can calculate the divergence from all of the above mentioned regions as

$$\mathrm{div}\, \hat{\boldsymbol{v}}(x,y) = \frac{\sum_{\mathrm{R}} v_x + \sum_{\mathrm{U}} v_y - \sum_{\mathrm{L}} v_x - \sum_{\mathrm{D}} v_y}{\sum_i \sqrt{\mathrm{Area}(S_i)}}. \quad (18)$$

4. Experimental Results

Experimental Conditions and Optical Flow

Images used to these experiments are shown in fig. 10. These images are 256 × 256 in size and 8 bit monochrome. Image 1 to 3 are computer-synthesized motion images

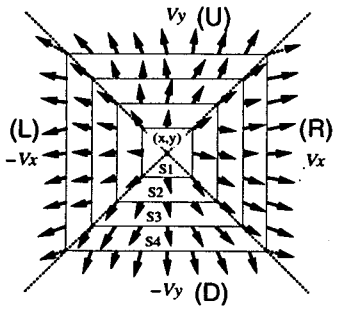


Fig. 9. A method of expanding effective integration length.

which represent inclined planes with uniform texture. Image 4 to 6 are captured by a moving CCD camera. For each image, successive three frames of images are simulated by geometric translation or captured by stop-motion camera.

Object planes in image 1 and 4 are normal to camera's optical axis. In image 2 and 5, object planes have a normal vector $(p\ q\ 1)^T = (0\ \frac{1}{\sqrt{3}}\ 1)^T$ and in image 2 and 6 it is $(p\ q\ 1)^T = (\frac{1}{\sqrt{3}}\ 1\ 1)^T$ (cf. eq. (11)).

Note that in the following experiments, optical flow vectors are calculated from the three successive frames before and after a nominal image by a method based on spatio-temporal gradient invariance[7] we proposed before.

Estimation of FOE

In this experiment, each FOE is estimated from image 1 to 6. View point or the camera center is set to translate along the optical axis so that the FOE should be at the origin.

For each image, FOE are calculated from optical flow vectors within four different regions in size (entire, 127 × 127, 32 × 32, and 16 × 16) by eq. (7). Experimental results are shown in fig. 11

It is supposed that estimation error in image 1 to 3 mainly due to limited representation of grayness (8bit), which is equivalent to signal-to-noise ratio of about 47 dB. In image 4 to 6, there is more additional noise from actual CCD camera and image capturing system, which results in larger estimation error.

Estimation of TTC

In this experiment, we tried to estimate TTC from the same images used in the previous experiment. TTC for the images utilized in this experiment is set to the time equivalent to 100 frames for image 1 to 3, and 200 frames for image 4 to 6. Experimental results are shown in fig. 12. In this figure, estimated TTC versus the size of region within which velocity vectors are taken into account is

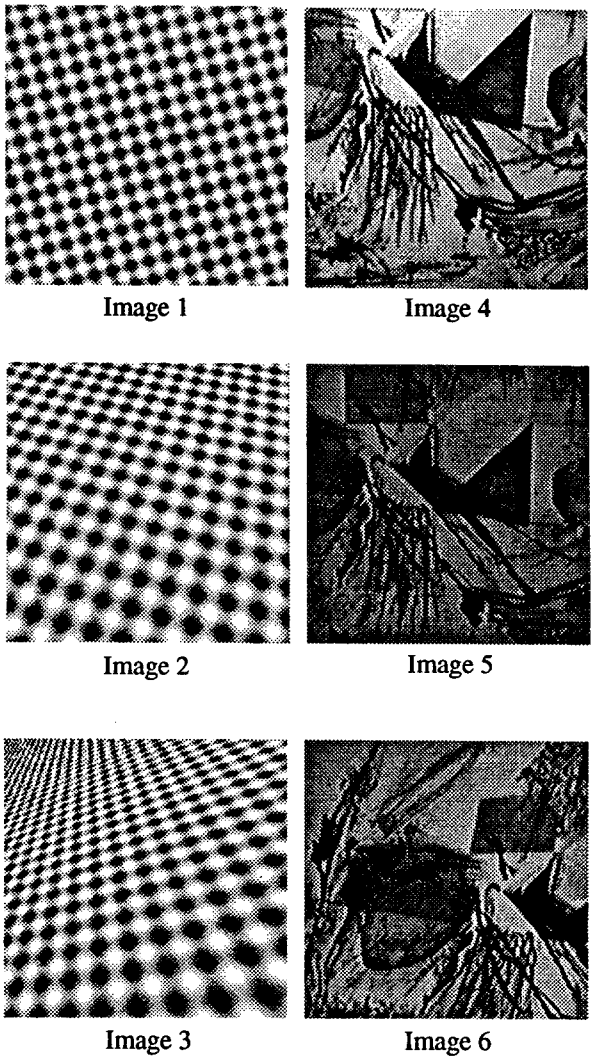


Fig. 10. Images utilized in experiments. Image 1 to 3 are synthesized and image 4 to 6 are captured.

shown. It is shown that as the size of region increase, estimated TTC tends to become stable and approach the correct value.

For synthesized images ((a),(c),(e)), evaluated TTC converges on the nominal value, while estimated TTC for captured images is more perturbed ((b),(d),(f)).

5. Discussion

Dependency of FOE error on TTC estimation

Proposed TTC estimator depends on the accuracy of FOE estimation. Therefore the effect of FOE error on TTC should be considered.

The systematic error of FOE estimation can be evaluated from eq. (9). If estimated FOE have deviation

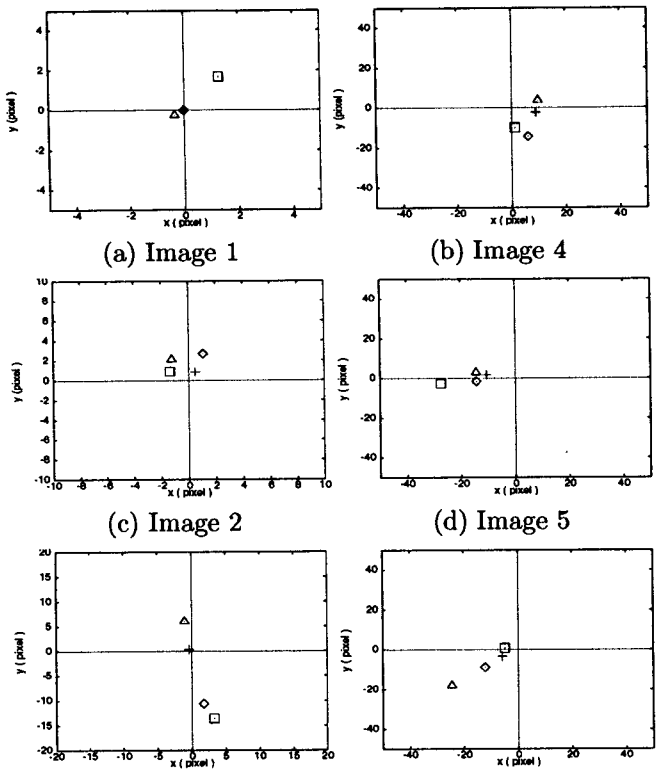


Fig. 11. Estimated FOE for different regions. △: from 16 × 16 flow vectors. □: from 32 × 32. +: from 127 × 127. ◇: from entire vectors.

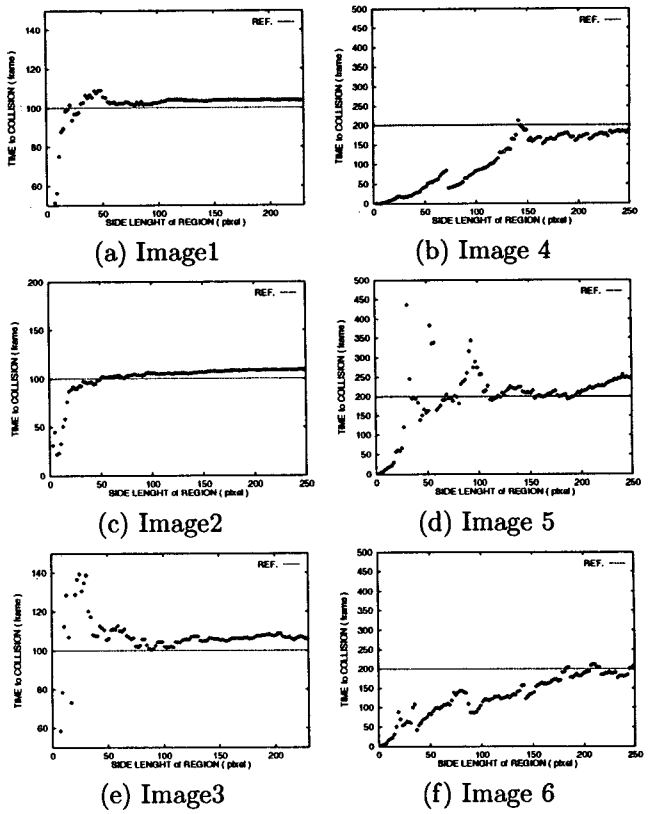


Fig. 12. Estimated TTC.

$(\Delta x, \Delta y)$ form true FOE (x_c, y_c), estimated divergence is described as

$$\operatorname{div} \boldsymbol{v} = -\frac{V_Z}{Z}\left(2 + \frac{\partial(\log Z)}{\partial x}\Delta x + \frac{\partial(\log Z)}{\partial y}\Delta y\right). \quad (19)$$

If Δx and Δy are small, eq. (19) becomes

$$\operatorname{div} \boldsymbol{v} \approx -\frac{V_Z}{Z}\left(2 + \log\frac{Z}{Z_c}\right). \quad (20)$$

Eq. (20) can give relation between optical flow divergence and TTC at a vicinal point around FOE.

6. Conclusions

In this paper, we formulated new algorithm to calculate TTC to a surface or plane which is not always normal to camera's optical axis.

Experimental results show that this method has the following advantages,

- It is possible to calculate FOE and TTC to a slanted surface.
- These estimations can be carried out from images with noise.

References

[1] K. Prazdny, "Egomotion and relative depth map from optical Flow," Biol. Cybern., **36**, 87–102, 1980

[2] K. Nakayama, "Biological image processing: a review," Vision Res., **25**, 625–660, 1985

[3] M. Subbarao, "Bounds on Time-to-Collision and Rotational Component from First-Order Derivatives of Image Flow," CVGIP, **50**, 329–341, 1990

[4] N. Ancona, T. Poggio, "Optical Flow from 1-D Correlation: Application to a Simple Time-to-Crash Detector," Int. J. of Computer Vision, **14**, 131-146, 1995

[5] J. D. McDonald, A. T. Bahill, and M. B. Friedman, "An adaptive control model for human head and eye movements while walking," IEEE trans. SMC, **13**, 167–174, 1983

[6] K. Kanatani, "Image Understanding," Morikita Publishing, 1990 (in Japanese)

[7] T. Yamaguchi and H. Yamasaki, "Velocity Based Vestibular-Visual Integration in Active Sensing System," Proc. of IEEE Conf. Multisensor Fusion and Integration for Intelligent Systems, 639–646, 1994