Proceedings of the 17th World Congress
The International Federation of Automatic Control
Seoul, Korea, July 6-11, 2008

Moving Object Detection for Active Camera based on Optical Flow Distortion

Masaaki Shibata*. Yuichiro Yasuda. Masahide Ito*****

**Dept. of Electrical and Mechanical Engineering, Seikei University,*
Tokyo, Japan (E-mail: shibam@ st.seikei.ac.jp).
***Dept. of Electrical Engineering and Electronics, Seikei University,*
Tokyo, Japan.
****Dept. of Electrical and Mechanical Engineering, Seikei University,*
Tokyo, Japan.

Abstract: The paper describes a novel method for moving object detection for an active camera using optical flows. In the images captured with an active camera, all of static and moving objects in the workspace are moving so that the camera moves around. In our approach, the analysis of optical flows is adopted for moving objects detection. The optical flow is resolved with using image information and camera motion in our method. The camera motion is controlled with robust controller, so that the information from the camera motion contributes well precise optical flow generation. The optical flows on moving objects are distorted to differ from the theoretical one because of the object motion. Since the precise optical flow is obtained in our method, the distortions are extracted, so that the moving object is detected. The validity of our method is confirmed in several physical experiments, and then good performance is achieved. The limitations of ability and performance of our method are also examined in the experiments.

1. INTRODUCTION

A way of moving object detection for an active camera system is proposed in this paper. The optical flow generated in the sequential images taken by a moving camera is used for detecting the object motion in the images. The detection is executed with analysis of the flow distortion, so that the moving object is distinguished from the static objects.

In recent years, the industrial robots have been developed which achieve highly precise and versatile performance with using several sensing devices. Especially, the visual device, i.e. CCD or CMOS camera, PSD, and so on, contribute the robot to achieving multipurpose functions. The images captured with the visual device enable to obtain the environment information around the system. One of the purposes of using captured images taken by the camera is to recognize the object in image processing. Moving object detection is also a kind of object recognition.

The active vision system is composed of the visual device and actuating systems, in which the camera can move to change its position and orientation by itself. The moving camera what is called the active camera allows obtaining enough environmental information much more than the statically fixed camera.

For the moving object detection, many approaches have been studied; e.g. background subtraction method, inter frame difference method and the technique using the optical flow, which are well known as valid approaches in the image processing (Shio *et al*., 1991, Yachida *et al*., 1981). Any method based on such approaches is very effective to detect the moving object when the camera is statically fixed in the workspace.

In this paper, we adopt the method based on the optical flow. The optical flow is a vector which represents the object velocity in the images. The flow is usually generated with using sequentially captured images. In case that, however, the moving camera is introduced to the system, it is also difficult to detect the moving object because both the static and moving objects in the images due to camera motion. Although no flow is generated on static objects fixed in the workspace when the camera is statically settled, the flow is generated on both the static and moving objects when the camera is moving.

It has been appointed that the optical flow generation requires substantial calculation cost (Adiv, 1985, Chen *et al*., 1993). And also, it is difficult to obtain the accuracy flows sufficiently. However, the flow is generated with using image information and camera motion in our approach, while only the image information is used for the generation in conventional approaches. When the camera motion is controlled in the robust control system, precise optical flow is estimated with low calculation cost.

The optical flow generated in our method is theoretically resolved in specific motion of the camera if the target object is statically fixed in the workspace. On the other hand, the flow on a moving object is distorted in comparing to the theoretical one. This distortion is referred to distinguish the moving object in our approach.

978-1-1234-7890-2/08/$20.00 © 2008 IFAC

10.3182/20080706-5-KR-1001.1164

2. OPTICAL FLOW GENERATION

The optical flow is a vector representing the velocity of a point in an image. The flow consists of two parameters since it is defined in two dimensional image space; one is the intensity, that is, the length of the vector, the other is the orientation. In order to resolve the vector, two equations are required at least, as constraints for the parameters. In our approach, the constraints are derived with the intensity of luminance in the image and the motion of an active camera so as to resolve the respective optical flow at each point.

2.1 Constraint derived from captured image

The first constraint for optical flow generation is derived from image information. The approach mentioned below is well known as the gradient method for the flow generation (Bruhn *et al*., 2005).

When a watched point is projected onto a point pixel at the position (x, y) in an image captured with the camera, let the intensity of luminance of the pixel be $I(x,y,t)$ at the time t. When the respective point moves to $(x+\Delta x, y+\Delta y)$ during an instant Δt, equation (1) is obtained if the intensity is assumed not to change during a small instant.

$$I(x,y,t) = I(x+\Delta x, y+\Delta y, t+\Delta t) \tag{1}$$

The Taylor expansion of (1) derives the following equation, where the second order term of displacements Δx, Δy, Δt are omitted so as to be small.

$$I(x,y,t) \approx I(x,y,t) + \frac{\partial I}{\partial x}\Delta x + \frac{\partial I}{\partial y}\Delta y + \frac{\partial I}{\partial t}\Delta t \tag{2}$$

Dividing both sides of (2) with Δt and minimizing it to $\Delta t \to 0$, equation (2) is rewritten into (3).

$$\frac{\partial I}{\partial x}\frac{dx}{dt} + \frac{\partial I}{\partial y}\frac{dy}{dt} + \frac{\partial I}{\partial t} = 0 \tag{3}$$

The velocity vector in the image space can be defined as $(u,v) \equiv \left(\frac{dx}{dt}, \frac{dy}{dt}\right)$, and then the spatial gradients of the intensity of a pixel along x-axis and y-axis can be defined as $I_x = \partial I / \partial x$, $I_y = \partial I / \partial y$, respectively, while the temporal gradients of the intensity as $I_t = \partial I / \partial t$. Since the velocity vector (u,v) is the optical flow, actually, the flow satisfies the following equation.

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

Equation (4) represents the first constraint of the optical flow.

2.2 Constraint derived from camera motion

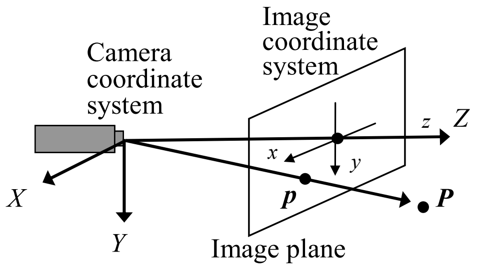


Fig.1. Camera coordinate system and image plane

The second constraint for optical flow generation is derived from active camera motion, attached on the robot tip. The camera motion is precisely controlled with robust motion controller, whose details are mentioned afterward.

At first, the relationship is made clear between coordinate systems of the camera and the image plane. Fig. 1 illustrates the camera coordinate system and image coordinate system. Note that the origin of camera coordinate system is defined to the tip of CCD camera attached on the robot, the Z-axis along the optical axis of the camera, the X-axis in right direction horizontally, and the Y-axis vertically downward. And also, note that the origin of image coordinate system is set on the virtual image plane at the intersection of the camera axis, the x- and y-axes parallel to the X- and Y-axes of the camera coordinate, respectively. An arbitrary point $\boldsymbol{P}$ in 3-dimensional workspace is projected onto the image plane at the point $\boldsymbol{p}$.

When the active camera moves around, the camera coordinate system moves together with same displacements. When the point $\boldsymbol{P}$ is statically fixed in 3-dimensional workspace and its position is represented as $\boldsymbol{X}$ in the camera coordinate system, the component values of $\boldsymbol{X}$ vary with the camera motion in translation and /or rotation. Assuming that $\dot{\boldsymbol{t}}$ is the translate velocity and $\boldsymbol{\omega}$ is the angular velocity of the moving active camera in the camera coordinate system, the point $\boldsymbol{X}$ translates with the velocity $-\dot{\boldsymbol{t}}$ toward the opposite direction and rotates with the angular velocity $-\boldsymbol{\omega}$ toward the inverse direction in the camera coordinate system. This relation satisfies the following equation.

$$\dot{\boldsymbol{X}} = -\dot{\boldsymbol{t}} - \boldsymbol{\omega} \times \boldsymbol{X} \tag{5}$$

, where '×' means the outer product for vectors.(15)

On the other hand, when the position of the point $\boldsymbol{p}$ on the image plane is represented as $\boldsymbol{x} = [x \;\; y \;\; 1]^T$, while the position of the point $\boldsymbol{P}$ is as $\boldsymbol{X} = [X \;\; Y \;\; Z]^T$ in the camera coordinate system, the relation between them is obtained as follows.

$$\boldsymbol{x} = \frac{\boldsymbol{X}}{Z} \cdot f \tag{6}$$

, where f is the focal length of the camera.

In considering the temporal differential of (6), the velocity vector in the image coordinate system is derived as follows.

$$\dot{\boldsymbol{x}} = \frac{f}{Z}\cdot\left(\dot{\boldsymbol{X}} - \boldsymbol{X}\cdot\frac{\dot{\boldsymbol{X}}\cdot\hat{\boldsymbol{Z}}}{Z}\right) \tag{7}$$

, where $\hat{\boldsymbol{Z}}$ is the unit vector of the optical axis of the camera represented in the camera coordinate system.

Substituting $\dot{\boldsymbol{X}}$ in (5) into (7) and using (6), the velocity vector $\dot{\boldsymbol{x}}$ in the image is obtained as follows.

$$\dot{\boldsymbol{x}} = \frac{1}{Z}\left(f\cdot\left(-\dot{\boldsymbol{t}} - \boldsymbol{\omega}\times\boldsymbol{X}\right) - \boldsymbol{x}\cdot\left(-\dot{\boldsymbol{t}} - \boldsymbol{\omega}\times\boldsymbol{X}\right)\cdot\hat{\boldsymbol{Z}}\right) \tag{8}$$

When the velocity vectors of the active camera are represented as $\dot{\boldsymbol{t}} = \left[\dot{t}_x \;\; \dot{t}_y \;\; \dot{t}_z\right]^T$ and $\boldsymbol{\omega} = \left[\alpha \;\; \beta \;\; \gamma\right]^T$, respectively, the optical flow $\left[u \;\; v\right]^T$ is obtained from (8) since $u = \dot{x}$ and $v = \dot{y}$.

$$\begin{cases} u = \dfrac{1}{Z}\left\{f\cdot\left(-\dot{t}_x - \beta Z + \dfrac{y}{f}\gamma Z\right) - x\cdot\left(-\dot{t}_z - \dfrac{y}{f}\alpha Z + \dfrac{x}{f}\beta Z\right)\right\} \\ v = \dfrac{1}{Z}\left\{f\cdot\left(-\dot{t}_y - \dfrac{x}{f}\gamma Z + \alpha Z\right) - y\cdot\left(-\dot{t}_z - \dfrac{y}{f}\alpha Z + \dfrac{x}{f}\beta Z\right)\right\} \end{cases} \tag{9}$$

The simultaneous equation (9) derives the following equation so as to eliminate the parameter *Z*, which is the depth to the point ***P*** in 3D camera coordinate system.

$$u\cdot\left(-f\cdot\dot{t}_y + y\cdot\dot{t}_z\right) - v\cdot\left(-f\cdot\dot{t}_x + x\cdot\dot{t}_z\right) = C \tag{10}$$

$$\begin{aligned} ;\quad C = {} & \dot{t}_x\cdot\left\{\left(f^2 + y^2\right)\cdot\alpha - x\,y\cdot\beta - f\,x\cdot\gamma\right\} \\ & + \dot{t}_y\cdot\left\{-x\,y\cdot\alpha - \left(f^2 + y^2\right)\cdot\beta - f\,y\cdot\gamma\right\} \\ & + \dot{t}_z\cdot\left\{-f\,x\cdot\alpha - f\,y\cdot\beta + \left(f^2 + y^2\right)\cdot\gamma\right\} \end{aligned}$$

Equation (10) represents the constraint for the optical flow. The equation consists of the point position (*x*, *y*) in the image and the camera motion which is represented with the translational and rotational velocities of the camera. Note that the equation does not include the target position in 3D camera coordinate system.

Consequently, the second constraint of the optical flow is obtained, which is derived from camera motion, while the first constraint in (4) is derived from image information.

2.3 *Optical flow derived from image information and camera motion*

Two constraints shown in (4) and (10) are unified into a simultaneous equation, which generates the optical flow. The equation is obtained with vector representation.

$$\boldsymbol{F}\cdot\boldsymbol{u} = \boldsymbol{g} \tag{11}$$

$$;\quad \boldsymbol{F} = \begin{bmatrix} I_x & I_y \\ -f\cdot\dot{t}_y + y\cdot\dot{t}_z & -f\cdot\dot{t}_x + x\cdot\dot{t}_z \end{bmatrix},\quad \boldsymbol{g} = \begin{bmatrix} -I_t \\ C \end{bmatrix}$$

When neither I_x nor I_y is 0, that is, the flow must exist, the inverse matrix of the regular matrix *F* can be derived.

$$\boldsymbol{u} = \boldsymbol{F}^{-1}\cdot\boldsymbol{g} \tag{12}$$

Consequently, the optical flow is estimated in (12).

3. ALGORITHM FOR MOVING OBJECT DETECTION

The moving object is detected with using the optical flow. The optical flow on the moving object is to be distinguished from one on the static object. But, in the captured images taken with the active camera, both static and moving objects generate optical flows, so that it is difficult to distinguish them with a simple approach.

In our approach, the optical flow assumed for static object is theoretically solved, and then the estimated optical flow, obtained in (12), is compared with the theoretical one. The flow on the moving object is detected in the comparison since the theoretical flow is solved where the object is assumed to be a static one.

The direction of an optical flow can be solved theoretically if a certain motion of the camera is specified. When the camera moves only in translation without pan motion, the right hand of (10) becomes 0 since the parameter *C* in (10) consists of pan motion. Therefore, the following equation is obtained from (10).

$$u\cdot\left(-f\cdot\dot{t}_y + y\cdot\dot{t}_z\right) - v\cdot\left(-f\cdot\dot{t}_x + x\cdot\dot{t}_z\right) = 0 \tag{13}$$

The direction of the optical flow ϕ is represented from (13), as follows.

$$\phi = \tan^{-1}\left(\frac{v}{u}\right) = \tan^{-1}\left(\frac{-f\cdot\dot{t}_y + y\cdot\dot{t}_z}{-f\cdot\dot{t}_x + x\cdot\dot{t}_z}\right) \tag{14}$$

The right hand of the above equation consists of parameters of the camera velocity in translation, the point position on the image and the focal length of the camera. If only the camera motion is obtained, the direction of the flow is estimated in (14).

When the unit vector $\bar{\boldsymbol{u}} = \left[\cos\phi \;\; \sin\phi\right]^T$ represents the direction of the theoretical optical flow, the angle ψ between the theoretical one $\bar{\boldsymbol{u}}$ and the estimated $\boldsymbol{u}$ is measured with the inner product of them.

$$\psi = \arccos\left(\frac{\boldsymbol{u}\cdot\bar{\boldsymbol{u}}}{\|\boldsymbol{u}\|\|\bar{\boldsymbol{u}}\|}\right) \tag{15}$$

If the angle ψ is not nearly 0, the concerned object might be moving. When such flows exist much in some region in the image, it is detected that the region contains the moving object.

4. EXPERIMENTS

4.1 Experimental System

The motion of the active camera is controlled in the robot system. Fig.2 shows the active vision robot, which has two kinetic degrees of freedom; one is pan-motion with the rotational vertical axis and the other is the translation with the horizontally translational axis. Fig.3 illustrates the experimental setup. In the initial posture of the camera, the optical axis, defined as *Z*-axis of the camera coordinate system, is vertically set to the translational axis of the active vision robot, which is defined as *X*-axis. In the experiments, the *Z*-axis is slanted with a fixed angle θ=35 [deg] from the initial posture. The robot is controlled only to translate the camera motion with constant speed.

Fig.4 shows the structure of the control system. The robot is controlled with PC (Pentium III 1MHz). The generated motion command in the PC is sent to the servo driver via D/A board. The driver provides DC current proportional to the command into each motor, and then the active camera is actuated. The rotating angles of the motors are measured with an encoder attached in each motor. The controller adopts PD control law with disturbance observer so as to realize the robust control system (Ohnishi, 1996). Though the camera motion is only in translation, very precise motion has been achieved, owing to the robust controller. The precise motion contributes to generate accurate optical flows.

Fig.2. Active vision robot

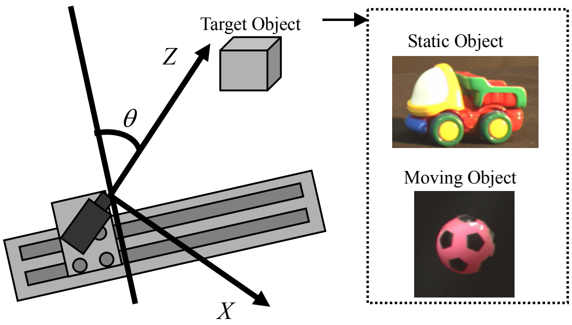


Fig.3. Top view of the experimental setup

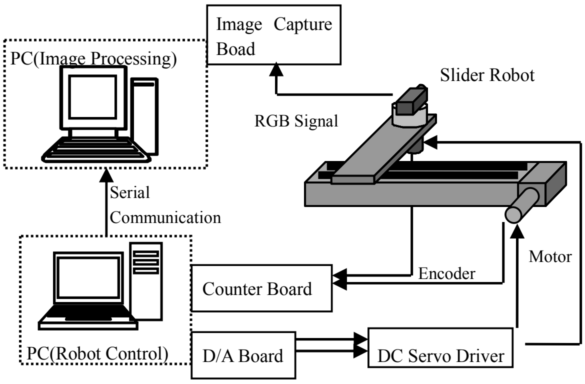


Fig. 4. Structure of control system

On the other hand, the images captured with CCD camera are sent to PC for image processing per 33 [msec]. For generating optical flows in our method, the information of camera motion is required, so that PC for controlling robot sends the information to PC for image processing via serial communication.

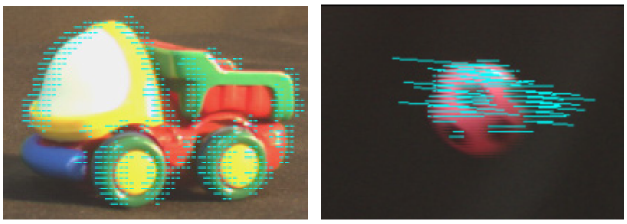
4.2 Experiments and results

In the physical experiments, the camera moves only in translation with constant velocity without pan-motion, where the pan angle is fixed in an offset angle θ=35 [deg], that is, the camera motion is restricted along a segmental path. When the constant speed along the path is specified as $\dot{\boldsymbol{d}}$, the translational velocity of the active camera in camera coordinate system is represented as follows, due to the offset angle θ.

$$\dot{\boldsymbol{t}} = \left[\dot{d}\cos\theta \quad 0 \quad \dot{d}\sin\theta\right]^T \tag{16}$$

And then, since the camera gets closer to the target object due to the offset angle, the captured object image becomes bigger with the camera motion.

Fig. 5 shows the experimental results; (a) is for a static object and (b) for moving. In the figures, the generated optical



(a) static object (b) moving object

Fig.5. Results of generated optical flow

flows are superimposed on the object, respectively.

When the theoretical angle of the optical flow is estimated in (14), the angle ψ between the theoretical one and the generated one is calculated in (15). If the angle is in the range $\psi = 0 \pm 0.1$ [deg], the flows are regarded as 'match', and, if not, 'mismatch', vice versa.

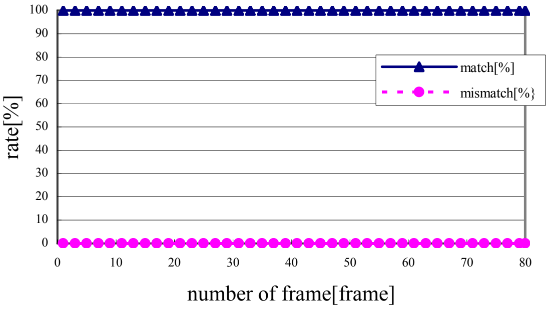
Fig.6 shows the examined results on static object and moving object. The respective results are examined 80 times in each frame of the captured image. In the results on static object, the angles between the theoretical and the generated optical flows are completely matched in whole the frames. On the other hand, many mismatched data is obtained in the results on moving object.

4.3 *Moving object detection*

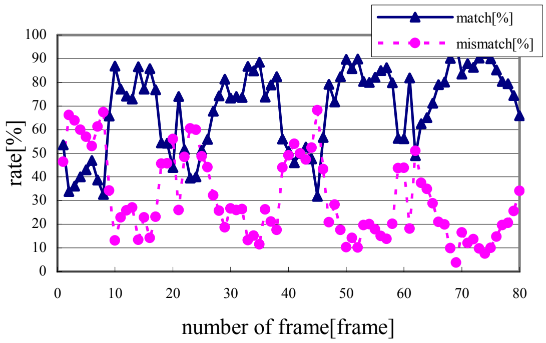
In the former section, it is found that the optical flow on a static object completely matches with the theoretical one, while some flows on moving object mismatches. By finding the region containing mismatched flows, the moving object is detected in the captured images.

In order to detect the moving object detection, the captured image is subdivided into small squares. Each size of the sub-region is specified to 20×20[pixel]. In case that the sub-region contains enough mismatched flows, it is regarded that the region indicates a part of a moving object.

The motion of active camera motion is controlled to keep a constant speed 2[mm/sec] in translation. The view angles of the camera lens are 25.22 [deg] along vertical axis and 33.24 [deg] along horizontal, respectively. The static objects are settled at about 0.7[m] from the camera.



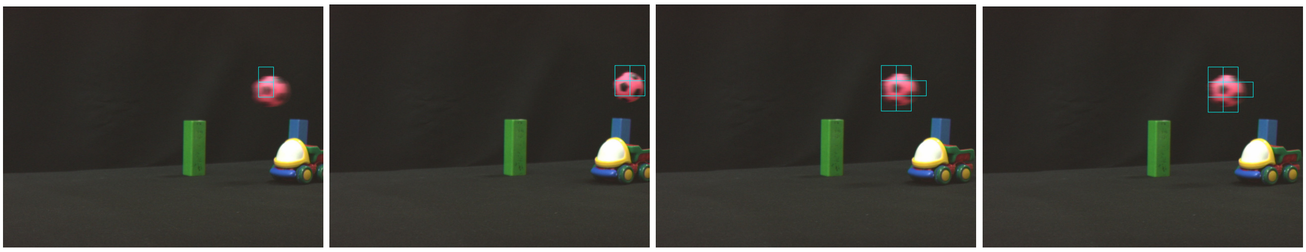
(a) static object



(b) moving object

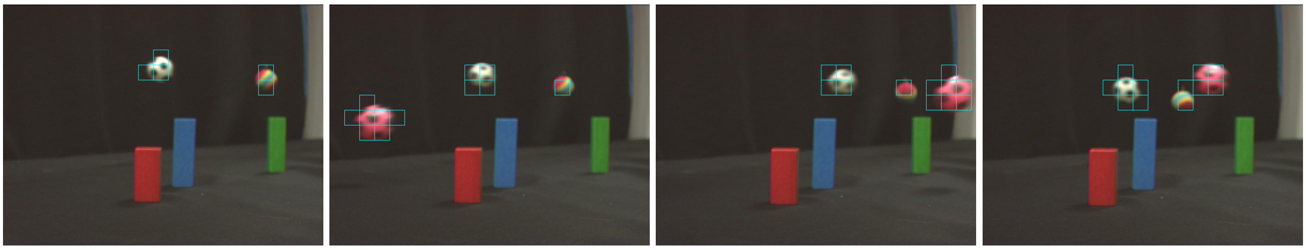
Fig.6. Matching rate between theoretical flow and estimated flow

Fig. 7 shows the experimental results, where a hung pink ball is moving in the air among static colored bricks. It is found that several squared region indicate the moving object, while there is no region on the static objects.



(a) 1st frame (b) 20th frame (c) 40th frame (d) 60th frame

Fig.7. Experimental results on a moving object detection



(a) 1st frame (b) 20th frame (c) 40th frame (d) 60th frame

Fig.8. Experimental results on plural moving objects detection

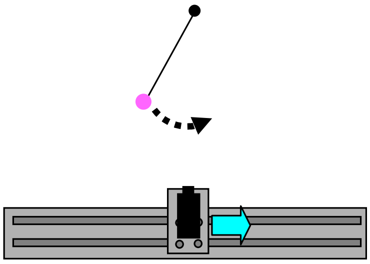


Fig. 9. Experimental setup for examination of relative speed

Fig. 8 shows the results in the case that there exist plural moving objects in the workspace. Even if the moving object does not exist initially, nor if existing objects would get out of the image, each moving object is surely detected. And also, if the moving objects make occlusion each others, or if the static object hides the moving one, the detection has been achieved, so that it means that our proposed method is valid in any case.

4.4 Performance depending on target speed

In our proposed method, the moving object is detected if it has enough speed. In this section, it is examined how much speed is required to detect the moving object.

Fig.9 shows the experimental setup for examination of relative speed between the camera and the target object. The object is made a part of a pendulum. Though the length of the pendulum is made constant to be 0.76[m], the swing angle is adjusted to achieve specified target speed. The optical axis of the camera is vertically set to the translation direction.

The camera speed is specified to 1.00[mm/sec], while the various speeds of the target are specified with adjusting the swing angle. Though the camera speed seems very slow, but the aim of this experiment is to exam the affects of the relative speed only, so that the camera speed is not so important even if the camera does not move.

Fig.10 shows the experimental results of relative speed examination. The faster the relative speed gets, the higher the mismatch rate, that is, the rate of the moving object detection grows up.

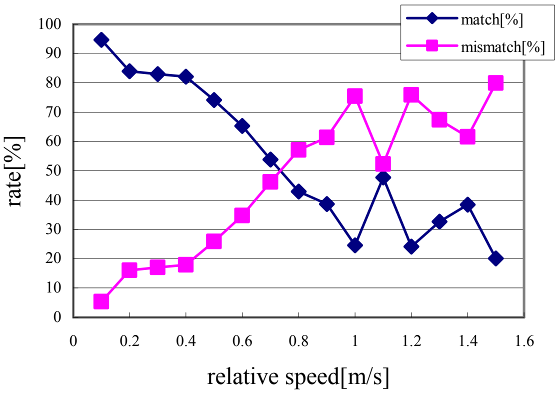


Fig.10. Matching rate for relative speed

The mismatch rate is referred for the decision whether the region containing mismatch flows should be on the moving object. But, if even a few mismatch flows are contained, it must be regarded that the region should be on the moving object because no mismatch flow is contained in the region on the static object. Therefore, it is found that the moving object whose speed is more than 0.1[m/s] can be detected. At the speed 0.1[m/s], the swing angle of the pendulum has been set to 2.3[deg].

5. CONCLUSIONS

A novel method for moving object detection for an active camera system is proposed based on optical flow distortion. Such the detection is not so easy in conventional approaches due to the camera motion. In our proposed approach, the optical flow is employed and the object motion is detected with comparing between the estimated flows and the theoretical ones. Though the comparison requires precise estimation of the flows, precise flows are generated in our method since the flow is derived from the image information and the camera motion.

In the experiments, it is achieved that plural moving objects have been detected during the statically fixed objects. And then, the performance of our method depending on the target speed is examined. The experimental results have proved the validity of our proposed method.

REFERENCES

G. Adiv (1985), Determining 3-D motion and structure from optical flow generalized by several moving objects, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **Vol.7**, pp.384-401

A. Bruhn, J. Weickert and C. Schnorr (2005), Lucas/Kanade Meets Horn/Schunck: Combining Local and Global Optic Flow Methods, *International Journal of Computer Vision*, **Vol. 61**, **No.3**, pp.211-231

H-J Chen, Y. Shirai and M. Asada (1993), Detecting Multiple Rigid Image Motions from an Optical Flow Obtained with Multi-Scale, Multi-Orientation Filters, *IEICE Transactions on Information and Systems*, **Vol.E76-D**, **No.10**, pp.1253-1262

K. Ohnishi, M. Shibata, T. Murakami (1996), Motion control for advanced mechatronics, *IEEE/ASME Transactions on Mechatronics,* **Vol.1**, **No.1**, pp.56-67

A. Shio and J. Sklansky (1991), Segmentation of people in motion, *Proc. IEEE Workshop on Visual Motion,* pp.325-332

M. Yachida, M. Asada and S. Tsuji (1981), Automatic analysis of moving image, *IEEE Transactions on Pattern Analysis and Machine Intelligence,*.**PAMI-3**, **No.1**, pp.12-20