2011 International Conference on Software and Computer Applications
IPCSIT vol.9 (2011) © (2011) IACSIT Press, Singapore

Detection of Moving Objects with a Mobile Omni-directional Camera

Chi-Cheng Cheng [1 +], Chia-Chih Chou [1] and Po-Wen Hsueh [2]

[1] Department of Mechanical and Electro-Mechanical Engineering, National Sun Yat-sen University Kaohsiung, Taiwan 80424, Republic of China

[2] Metal Industries Research & Development Centre, Kaohsiung, Taiwan 81160, Republic of China

Abstract. The omni-directional camera has a full range in all directions, which gains the complete field of view. In the past, a moving object can be detected, only when the camera is static or moving with a known speed. In this paper, we assume an omni-directional camera is mounted on a moving platform, which travels with a planar motion. The region of floor in the omni-directional image and the brightness constraint equation are applied to estimate the ego-motion. Using the estimated ego-motion, the optical flow caused by the floor motion can be computed. Therefore, comparing its direction with the direction of the optical flow on the image leads to detection of a moving object. Even if the camera is in the condition that combining translational and rotational motions, a moving object can still be accurately identified.

Keywords: ego-motion, motion detection, omni-directional cameras, optical flow.

1. Introduction

The compound eye of insects owns different structure from that of humans, which consists of thousands of ommatidia and even more. One special feature of the compound eye is the extremely broad view angle to cover almost all direction around the insect body. This uniqueness plays a significant role in keeping away from hostile surroundings and locating possible preys. Research results imply that the compound eye determines three dimensional motion of moving objects in terms of optical flows [1]. Therefore, the compound eye is more sensitive to moving objects than static ones.

There have been three basic approaches to expand viewing range of a camera [2]. They are (i) single camera revolving 360 degrees, (ii) multiple cameras, and (iii) an omni-directional camera. The omni-directional camera is able to acquire an omni-directional image at one time without worrying about correspondence problems for images taken in different viewing directions or from different cameras.

Detection of moving objects for a static camera has been extensively studied for many years. But for the case with a moving camera and moving objects, in order to distinguishing moving objects, these two independent motions need to be decomposed. A computational method to calculate the camera's rotation speed for static background without prior knowledge of position relationship between an object and the camera was presented in 1981 [3]. Then a similar approach was applied to determine parameters of translational movement of a moving vehicle [4]. Gandhi and Trivedi installed an omni-directional camera on a mobile platform and moving objects could be differentiated if their corresponding image information did not match the floor movement model [5].

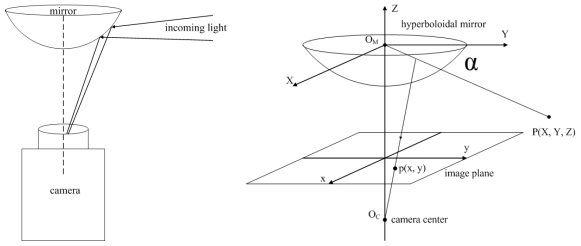
By comparing estimation performance of ego-motion, it was concluded that the camera should be pointed downwards because translation and rotation components could be easily decomposed [6]. In order to overcome similar optical flow velocities caused by translation and rotation using conventional cameras, a

+ Corresponding author. Tel.: + 886-7-525-4236; fax: +886-7-525-4299.
E-mail address: chengcc@mail.nsysu.edu.tw.

Jacobian transformation matrix was used to have the optical flow projected onto a sphere or a curved surface [7][8]. Nevertheless, these mapping techniques were restricted in static environment.

The omni-directional visual system discussed in this paper consists of a CCD camera and a reflective hyperboloid mirror, which allows image of the surrounding environment to be projected onto the camera as shown in Fig. 1(a). Fig. 1(b) illustrates the imaging of the reflective omni-directional visual system. A reference coordinate system $(O_M; X, Y, Z)$ is established with its origin at the focal point of the hyperboloid surface and the *Z* axis coincides with the central axis. As a result,

$$Z = \sqrt{X^2 + Y^2} \tan\alpha \ . \tag{1}$$



(a) (b)

Fig. 1: (a) The reflective omni-directional visual system and (b) its corresponding imaging illustration.

Although equations for ideal projection geometry can be derived straightforwardly, they can only be applied if the collinearity of the optical axis of the camera and the central axis of the hyperboloid mirror can be assured. Therefore, through a calibration technique presented by Scaramuzza, *et al*. [9], the component in Z direction of the vector along PO_M has the following form:

$$[x, y, f(r)]^T \ , \tag{2}$$

where $f(r) = a_0 + a_1 r + \ldots + a_N r^N$, $r = \sqrt{x^2 + y^2}$, and (x, y) is the coordinate on the image plane.

This paper aims to apply an omni-directional camera to get a panorama vision like insects do. Then under the circumstance of a moving camera, solve the ego-motion of the camera itself and further detect moving objects in the scene.

2. Estimation of Ego-Motion

Assume that the omni-directional camera is installed on a mobile vehicle traveling with a linear translational velocity $\mathbf{T}_C$ and a rotational velocity $\boldsymbol{\omega}_C$. When the vehicle moves, there will be relative motion between the vehicle and the static environment. The ego-motion of the vehicle can therefore be determined by the resulting optical flow on the image. The depth problem will be solved by the height between the vehicle and the ground, which should be easily known in advance.

For the omni-directional image, the full surrounding scene will be projected onto a circular region on the image plane. Therefore, a polar coordinate system instead of a regular Cartesian reference system will be applied for deriving the equation of brightness constancy raised by Horn and Schunck [10], i.e.,

$$\frac{\partial E}{\partial r}\dot{r} + \frac{\partial E}{\partial \theta}\dot{\theta} + \frac{\partial E}{\partial t} = 0 \ , \tag{3}$$

where $E(r,\theta,t)$ denotes the brightness at a point with the polar coordinate (r,θ) on the image plane at time t.

Assume a point **P**(*X*, *Y*, *Z*) is on the ground, $\mathbf{T}_C = [T_X, T_Y, T_Z]^T$, and $\boldsymbol{\omega}_C = [\omega_X, \omega_Y, \omega_Z]^T$. Since only planar motion exists, $T_Z = \omega_X = \omega_Y = 0$,

$$\dot{X} = Y\omega_Z - T_X \ , \quad \dot{Y} = -X\omega_Z - T_Y \ , \text{and } \dot{Z} = 0 \ . \tag{4}$$

Because $\theta = \tan^{-1}(Y/X)$,

$$\dot{\theta} = \frac{X\dot{Y} - \dot{X}Y}{X^2 + Y^2} \ . \tag{5}$$

Furthermore, according to (1) and (2), the following expression can be obtained:

$$\alpha = \tan^{-1}\frac{Z}{\sqrt{X^2 + Y^2}} = \tan^{-1}\frac{f(r)}{r} \ . \tag{6}$$

Taking time derivative of (6) and incorporating both (4) and (5) lead to

$$\dot{\theta} = \frac{-Y \cdot T_X + X \cdot T_Y}{X^2 + Y^2} + \omega_Z \text{ and } \dot{r} = G(r) \cdot F_\alpha(X,Y,Z) \cdot (X \cdot T_X + Y \cdot T_Y) \ , \tag{7}$$

where

$$F_\alpha(X,Y,Z) = \frac{-Z}{(X^2 + Y^2 + Z^2)\sqrt{X^2 + Y^2}} \text{ and } G(r) = \frac{r^2 + [f(r)]^2}{-a_0 + a_2 r^2 + \cdots + (N-2)a_{N-1}r^{N-1} + (N-1)a_N r^N} \ . \tag{8}$$

As a result, the brightness constancy equation (3) can be rewritten as:

$$[E_r \cdot G(r) \cdot F_\alpha(X,Y,Z) \cdot X - \frac{E_\theta \cdot Y}{X^2 + Y^2}]T_X + [E_r \cdot G(r) \cdot F_\alpha(X,Y,Z) \cdot Y + \frac{E_\theta \cdot X}{X^2 + Y^2}]T_Y + E_\theta \omega_Z + E_t = 0 \ . \tag{9}$$

The equation above will be employed to estimate the motion components T_X, T_Y, and ω_Z. Nevertheless, the position of the point *P* in three-dimensional space is still unknown.

Let *H* represent the height of the focal point of the hyperboloid surface, which is usually known in advance. Therefore,

$$X = -H\cot\alpha\cos\theta \ , \quad Y = -H\cot\alpha\sin\theta \ , \text{and } Z = -H \ . \tag{10}$$

Since $\theta = \tan^{-1}(y/x)$ and $\alpha = \tan^{-1}[f(r)/r]$, its corresponding location (X,Y,Z) on the ground can be solved by (10). Therefore, all parameters in (9) are determined. But just one equation is not enough to resolve three unknowns. Then the motion components T_X, T_Y, and ω_Z will be estimated by applying the least squared approach to more than two pixels' information.

3. Motion Estimation of a Moving Object

Any point (*x*, *y*) on the image plane can be represented in a polar form of $(r\cos\theta, r\sin\theta)$. Differentiating the polar expression with respect to time yields

$$\dot{x} = \dot{r}\cos\theta - r\dot{\theta}\sin\theta \ \text{ and } \ \dot{y} = \dot{r}\sin\theta + r\dot{\theta}\cos\theta \ . \tag{11}$$

Assume that the corresponding point **P** in three-dimensional scene is denoted by

$$\mathbf{P} = [X \quad Y \quad Z]^T = R[X_P \quad Y_P \quad Z_P]^T \ , \tag{12}$$

where R is measure of the depth and $X_P^2 + Y_P^2 + Z_P^2 = 1$. Incorporating (12) into (7) leads to

$$\dot{\theta} = \frac{1}{R}\left(\frac{-Y_P \cdot T_X + X_P \cdot T_Y}{{X_P}^2 + {Y_P}^2} + R\omega_Z\right) \text{ and } \dot{r} = \frac{1}{R}[G(r) \cdot F_\alpha(X_P, Y_P, Z_P) \cdot (X_P \cdot T_X + Y_P \cdot T_Y)] \ . \tag{13}$$

After putting (13) into (11), the direction of optical flow can be determined by solving the ratio of $\dot{y}/\dot{x}$. Detection of a moving object will be accomplished by comparing the directions of the computed optical flow and the optical flow on the floor.

4. Simulation Studies

An approach proposed in [11] was adopted to generate test images for the hyperboloid-type omni-camera used in this work. Fig. 2(a) shows the hyperboloid-type omni-directional camera employed in this paper. A real image taken by the omni-directional camera and the simulated floor image are illustrated in Fig. 2(b) and 2(c), respectively.

(a) (b) (c)

Fig. 2: (a) The hyperboloid-type omni-directional camera used for simulation studies, (b) a real image taken by the omni-directional camera and (b) the simulated floor image.

There is a blind spot for r < 60 pixels in the real image. Therefore, information in this region will not be used. The region near the camera seldom contains moving objects and provides better image quality because of shorter depth. An annular region with r = 60~100 pixels will be applied for estimation of ego-motion. Besides, another annular region with r = 100~190 pixels will be chosen for detection of moving objects.

Three different simulation tests, translation only, rotation only, and combination of translation and rotation, were investigated for estimation of ego-motion. Let $\mathbf{T}_a$ and ω_a denote actual translation velocity and rotational speed, and $\mathbf{T}_e$ and ω_e stand for estimated translation velocity and rotational speed. The errors of the translation velocity and the rotation speed are respectively defined as:

$$e_t = \cos^{-1}\frac{\mathbf{T}_a \cdot \mathbf{T}_e}{|\mathbf{T}_a||\mathbf{T}_e|} \text{ and } e_\omega = |\omega_e - \omega_a| \ . \tag{14}$$

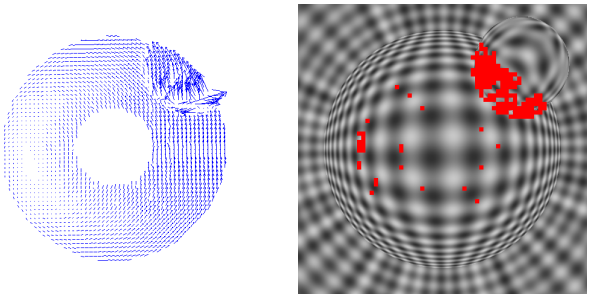
Simulation results are listed in Table 1. It appears that excellent estimation performance for ego-motion can be reached for pure translation and pure rotation. However, estimation performance of ego-motion for motion combing translation and rotation is still satisfactory for general applications.

Table 1: Simulation results of estimation of ego-motion.

	$\mathbf{T}_a$	$\mathbf{T}_e$	e_t	ω_a	ω_e	e_ω
Translation only	(0, -30)	(0.0111, -30.0384)	0.0212	0	-0.0028	0.0028
Rotation only	(0, 0)	-	-	-10	-9.9519	0.0481
Translation and rotation	(0, -30)	(0.1783, -30.2220)	0.3380	-10	-10.0496	0.0496

A moving object was simulated as a sphere with a radius of 60 at (100, 100, -40). Assume the sphere travels with a planar velocity of (20, 20, 0) and the camera moves with three different conditions including translation with (0, -30), rotation with -10 deg/sec, and combination of the translation and the rotation. It

appears that the outliers based on the computed optical flow over a given threshold value is able to indicate moving objects for all three cases. The results for the condition of hybrid motion are illustrated in Fig. 3.



(a) (b)

Fig. 3: (a) The computed optical flow, and (b) distribution of outliers for motion with translation and rotation.

5. Conclusions

This paper utilizes a hyperboloid-type omni-directional camera to expand fields of view by mimicking compound eye of insects. Ego-motion of the camera itself is estimated using the floor image. If the direction of the optical flow for an object in the scene is significantly different from that of the static floor, the object can be justified to be moving. Furthermore, when the camera owns a motion either pure translation or pure rotation, better estimation result for the ego-motion will be obtained than that for the motion of the camera combining translation and rotation. Large deviation of the optical flows between the static scene and a moving object brings about better accuracy on detection of moving objects. However, when a moving object produces a similar optical flow as that of the static scene, then the object cannot be differentiated as a moving object from the static background.

6. Acknowledgements

This work was supported in part by the Metal Industries Research & Development Centre, Taiwan, Republic of China.

7. References

[1] T. Collett. Insect vision: Controlling actions through optical flow. *Current Biology* 2002, **12** (18): R615-R617.

[2] S. K. Nayar. Catadioptric omnidirectional camera. In: *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*. 1997, pp. 482-488.

[3] K. Prazdny. Determining the instantaneous direction of motion from optical flow generated by a curvilinearly moving observer. Computer Graphics and Image Processing 1981, **17**: 238-248.

[4] W. Burger, and B. Bhanu. Estimating 3-D egomotion from perspective image sequences. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 1990, **12** (11): 1040-1058.

[5] T. Gandhi, and M. M. Trivedi. Motion based vehicle surround analysis using an omni-directional camera. In: *Proceedings of IEEE Intelligent Vehicles Symposium* 2004, pp. 560-565.

[6] Q. Ke, and T. Kanade. Transforming camera geometry to a virtual downward-looking camera: Robust ego-motion estimation and ground-layer detection. In: *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition* 2003, pp. 390- 397

[7] J. Gluckman, and S. K. Nayar. Ego-motion and omnidirectional cameras. In: *Proceedings of the International Conference on Computer Vision* 1998, pp. 999-1005.

[8] O. Shakernia, R. Vidal, S. Sastry. Omnidirectional egomotion estimation from back-projection flow. In: *Proceedings of IEEE Workshop on Omnidirectional Vision* 2003, pp. 82.

[9] D. Scaramuzza, A. Martinelli, R. Siegwart. A flexible technique for accurate omnidirectional camera calibration and structure from motion. In: *Proceedings of IEEE International Conference on Computer Vision Systems* 2006, pp. 45.

[10] B. Horn and, and B. Schunck. Determining optical flow. *Artificial Intelligence* 1981, **17**: 185-203.

[11] S. Hirunyaphisutthikul, and A. Tunwannarux. Synthesizing test image sequences for omnicamera: Translating and rotating spheres. In: *Proceedings of IEEE International Symposium on Communications and Information Technology* 2004, pp. 421-426.