2009 International Symposium on Nonlinear Theory and its Applications
NOLTA'09, Sapporo, Japan, October 18-21, 2009

Depth Detection Using Optical Flow Obtained by Rotational Camera

Xuebing Wang† and Kazuo Ishii†

†Graduate School of Life Science and Systems Engineering, Kyushu Institute of Technology
2-4 Hibikino, Wakamatsu-ku, Kitakyushu 808-0196, Japan
Email: wang-xuebing@edu.brain.kyutech.ac.jp, ishii@brain.kyutech.ac.jp

Abstract– This paper provides a new strategy for estimation of absolute depth perception to objects from a sequence of successive frames by a single rotational camera. Motion parallax which is a big cue for depth perception in monocular vision is used to estimate depth of object point present in the scene. In the simulation, the camera model is designed to accelerate along a trajectory of a circle by a constant radius. The transformation between global coordinate and camera coordinate is performed to obtain the coordinate of objects on the successive images. So the objects' acceleration on the image can be estimated. In the algorithm we only use the object acceleration on the image and the camera's acceleration to estimate the absolute depth from objects.

1. Introduction

There are several ways for a vision system to obtain depth information from a complex visual scene, primarily could be categorized into monocular vision and binocular vision [1]. At close range, most work on 3-d distance detection has been focused on binocular vision which is an accurate method to tell the distance for the objects near to the observer. Binocular cues are based on the disparity between the different views of the world from the two eyes, which can be obtained using a stereoscope [2] and a detailed overview of stereoscopes and their design principles is given by howard & Rogers [3]. In fact, for distant objects (beyond about 20 feet) stereopsis contributes little to depth perception; monocular cues play a more important role, especially for micro robot or small unmanned vehicles which eyes are close together therefore do not have good stereoscopic vision. On the other hand, stereopsis provides only relative depth information, while absolute estimates of distance require monocular cues [4].

The monocular cues include interposition, relative size, perspective and motion parallax [5]. The former three cues are static depth cues, whereas the relative apparent motion of objects within your field of view as you move (motion parallax) can provide vivid monocular depth information from observer to objects in motion field [6]. A system that measures distance for a stationary object by moving the observer has been demonstrated for a 1D sensor array [7]. In our former work, we did a Monte-Carlo simulation for depth from motion parallax, also an experiment is performed by fusing a 3-D inertial sensor and a silicon retina which acceleration is created by system's gravity along a slide to percept depth of a high contrast object in the range of 3 meters [8]. Also the verification experiment is performed under outdoor situation [9]. However these applications are evaluated in the translational field, i.e. only provides a depth perception algorithm for one dimension, and it is difficult to use this system on robots in respect that the space of acceleration is big. In 2008, a team of researchers led by Greg DeAngelis have discovered that the depth estimation processing of the image from a single eye and the motion of our bodies happened in small part of the brain [10]. The neurons are combining monocular vision cues and non-visual cues to come up with a unique way to create an approximation of the three-dimensional world in our minds. The non-visual cues should be included in the 2-D field such as motion information of rotation. Many researchers are concentrated on monocular vision for depth using rotation observer such as in [11]. However as the algorithm is close to the binocular disparity and the distance from previous view point to next one is short, result in big errors in estimation and need many continuous samples to make a modification means that the algorithm is not fit for the real time vision task. Our research is emphasized on the motion parallax as an independent cue for depth perception by using only three samples.

This paper is concentrated on the new trajectory for depth perception by only using observer's rotation acceleration and monocular vision in order to map the entire surrounding environment view by distance information. In the simulation, the camera model is designed to accelerate along a trajectory of a circle by a constant radius. The transformation between global coordinate and camera coordinate is performed to obtain the coordinate of objects on the successive images. So the objects' acceleration on the image can be estimated. Because if the perceived acceleration of objects on the focal image and the observer's acceleration are needed for depth detection. The algorithm is simple so that it is easy to be embedded in a robot system. And if the rotation radius is small enough also the system can be used in the small mobile robot or unmanned vehicles.

In a word, the aim of this work is to make a real-time, low cost, small depth detection system using motion parallax.

2. Depth from motion parallax

For a monocular system would correspond to motion parallax or dynamic occlusion etc. In the real world as you move in the environment around this object will demonstrate a phenomenon called motion parallax. Objects closer to you have fixated will appear to move more slowly but in the same direction as yourself. The relative apparent motion of objects within your field of view (as you move) provides a strong cue to the relative distance of objects from the observer. If the velocity of the observer is known the distance to static objects can be determined. If also the acceleration of the observer is known the distance can be estimated to objects moving with constant velocity. The depth estimation equation as follows detailed discussed in [8].

$$D = \frac{a_{obs} f}{\alpha_p} + f \qquad (1)$$

Where D is the distance from observer (camera) to objects, a_{obs} is observer acceleration, a_p is perceived acceleration of objects on the focal image and f is focal length of camera. From the depth equation, in translational field the depth estimation only depends on the acceleration of observer and perceived acceleration of objects. Only the translational field depends on the distance of moving points in the image, while the rotational field does not [6]. So we propose to improve the algorithm from the translational field to rotational field by using an approximate measurement analysis through equation 1. Fig.1 shows the system model and camera model. The camera model is to turn α degree clockwise around global coordinate origin O with a same acceleration. The relation between global coordinate and camera coordinate is also illustrated. Then the projected point (X_f, Y_f) of object P(X,Y,Z) on the focal image give a transformation as follows:

$$X_f = \frac{f}{d_x} \frac{X\cos\alpha - Z\sin\alpha}{X\sin\alpha + Z\cos\alpha - r} + C_x$$
$$Y_f = \frac{f}{d_y} \frac{Y}{X\sin\alpha + Z\cos\alpha - r} + C_y \qquad (2)$$

Where d_x, d_y are one pixel size of focal image in X and Y axis respectively, f is the focal length of camera, r is the rotation radius, (C_x, C_y) is the center of the focal image.

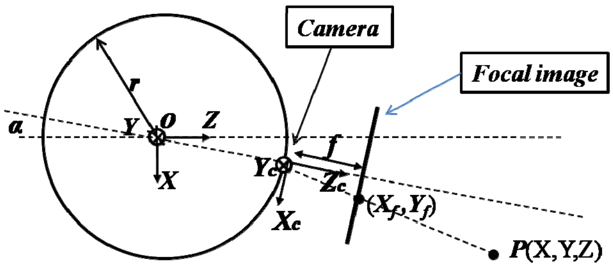

Fig.1 System model include a camera accelerated according to same value by radius r

3. Simulation

3.1. Simulation set-up

As mentioned in above, camera model is accelerated along the global coordinate origin O with a same acceleration. Camera model parameter set-up follows the settings of a real camera. The system settings and camera model settings are shown as below. System model is built in MATLAB.

- Focal length is 5.5mm
- Resolution is 640*480 [pixels]
- d_x= 0.0075mm, d_y= 0.00728mm
- Center coordinate of focal image (C_x C_y)=(320,240)
- Frame rate is 30fps
- Angle acceleration is set to $\pi/3$ rad/s^2
- Rotation radius is set to r= 0.5m

3.2. Perceived acceleration estimation

The transformation between global coordinate and camera coordinate is performed to obtain the coordinate of objects on the successive images as (2). In simulation through this equation the perceived velocity vector of objects on the focal image (optical flow) can be obtained. So the objects' acceleration on the image also can be estimated through successive three images. In translational field case under same observer acceleration, constant acceleration can be obtained. However as a result of rotational component added in the trajectory, only while objects pass through the center area of focal image we can tell the value of perceived acceleration is close to the one in case of translation. In other area a compensatory model is needed. In experimental stage we would like to use a 2-D optical flow sensor [12] to embed this system into mobile robot, the adjustment from the level of velocity vector of objects on the focal image is needed. First, assume the camera model rotate with a constant velocity around global coordinate origin. The object P is placed at coordinates (0, 0, 1.5), which means the distance between object and observer is 1m while the object is projected in the center of the focal image. The result of velocity and acceleration shows in Fig.2. From the Fig.2 although the camera velocity is constant, the perceived velocity is changed from big to small then to big as a curve of second degree. When the object passes through the center area of the focal image, the velocity is smallest. This indicates that the velocity changing is based on the location of objects on the image. Fig.3 shows the relation between velocity and location of objects. A approximate quadratic equation can be obtained as follows:

$$y = 2.7646 \times 10^{-5} x^2 - 0.0183x + 22.217 \qquad (3)$$

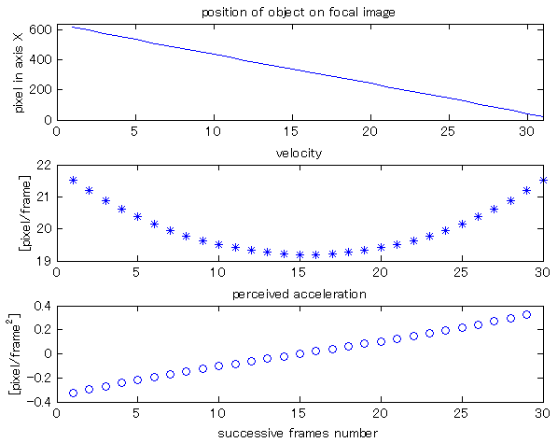

Fig.2 position of object on focal image, perceived velocity and perceived acceleration

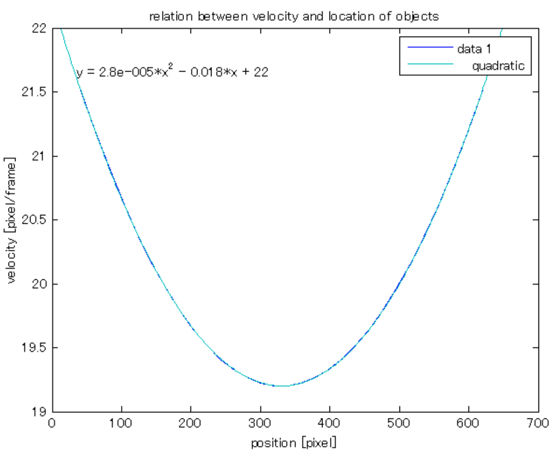

Fig.3 relation between velocity and location of objects

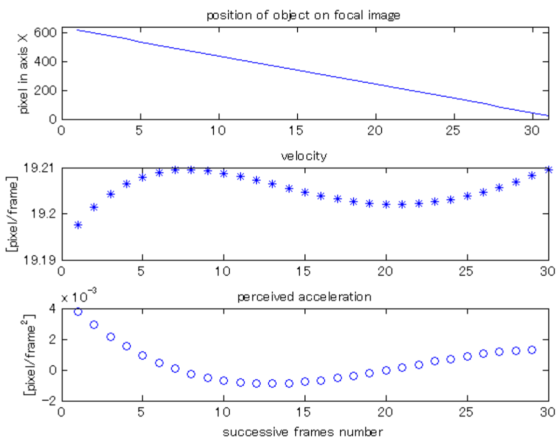

Fig.4 position of object on focal image, perceived velocity and perceived acceleration after adjustment.

Fig.4 shows the position of object on focal image, perceived velocity and perceived acceleration after adjustment. The perceived acceleration is close to zero. Then the position of object in axis Z is changed from 0.1m to 2m at intervals by 0.1m. Fig.5 shows the result of relation between perceived acceleration and true distance. So this relation is also can be used by looking for the corresponding depth from estimated acceleration and just need to approximate the estimated acceleration data to an approximated curve.

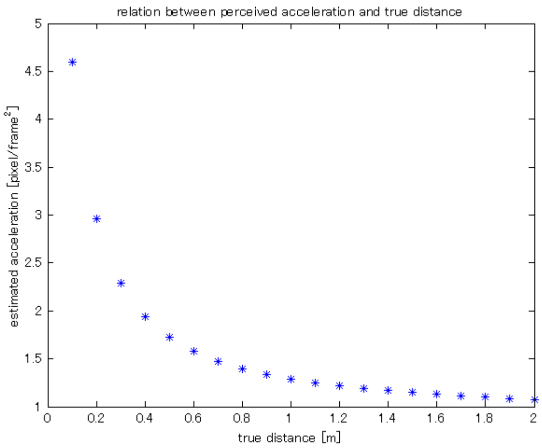

Fig.5 Relation between estimated acceleration and true distance

3.3. Depth estimation

Depth estimation is performed by using estimated acceleration and camera acceleration as (1). Fig.6 shows the result. Depth estimation data spread out from the reference line and intersect about 0.4m. The equation (1) belongs to inverse proportion function as y=1/x, so when the error is added into denominator, measurement values present the property of diverging. The relative errors from 0.1 to 1m are shown in Fig.7. The errors from 0.3 to 1.0m are within 40% and error at 0.4m close to 0. However in other area the relative errors are so big that cannot use in experimental stage, the improvement of algorithm is needed to adjust the estimated value.

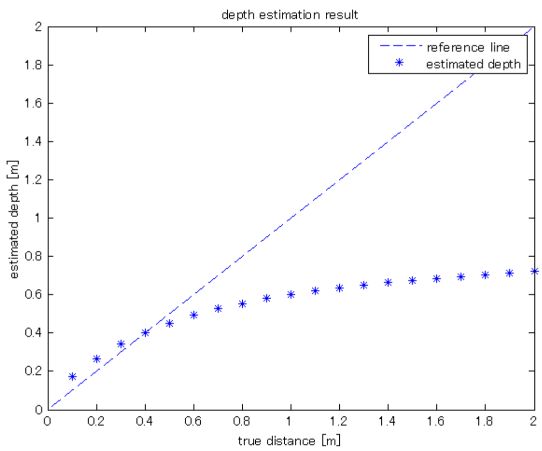

Fig.6 depth estimation result

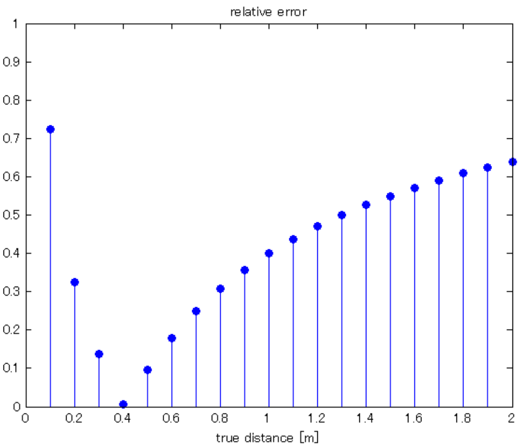

Fig.7 Relative error of estimated depth

4. Conclusions and discussion

Motion parallax is used to generate a feasible sensing strategy for determining the position between a moving observer and stationary objects or moving non-accelerating objects in translational direction. Depth estimation is performed just with perceived acceleration and observer acceleration. The simulation of depth estimation shows the feasibility of this algorithm. The big relative error area can also be adjusted using particle filter by combining other algorithm like stereopsis.

Acknowledgments

This research is supported by Intellectual Cluster Project (2nd) of MEXT, Japan.

References

[1] Martin J. Tovee, “An introduction to the visual system,” Cambridge University Press, 1996.
[2] C.Wheatstone, “Some remarkable phenomena of vision I,” Philosophical transactions of the eroyal society, pp.371-395, 1838.
[3] I. P. Howard, B. J. Rogers, “Binocular vision and stereopsis,” Martin J. Tovee, “An introduction to the visual system,” Series No 29, Oxford University Press, 1995.
[4] S. H. Ferris, “Motion parallax and absolute distance,” Journal of experimental psychology, 1972.
[5] R.Sekuler, R. Blake, “Perception(3rd Ed),” McGraw-Hill Inc,1994.
[6] H. A. Mallot, “Computational vision information processing in perception and visual behavior,” The MIT Press, 2000.
[7] S.S.Karri, A.H.Titus, “Toward an analog very large integration system for perceiving depth through motion parallax,” Optical Engineering 44(5), 2005.
[8] V. Becanovic, Xuebing Wang, “A vision system for depth perception that uses inertial sensing and motion parallax,” Mechatronics and Vision in Practice, pp.65-76, Springer, 2007.
[9] Xuebing Wang, Kazuo Ishii, “Depth perception using a monocular vision system,” ICONIP 2008, Part I, LNCS 5506, pp.778-785, Springer, 2009
[10] Nadler JW, Angelaki DE, DeAngelis GC, “A neural representation of depth from motion parallax in macaque visual cortex,” *Nature, 452*: 642-645, 2008
[11] A.Iguchi, H. Kudo, N. Ohnishi, “depth estimation using the knowledge of camera rotation”, Proceedings of the IEICE General Conference, pp,284 20030303,2003
[12] A.A.Stocker, “Analog integrated 2D optical flow sensor,” Analog integrated circuits and signal processing, Springer; vol.46, no.2, pp.121-138, 2006.