Estimation of Visual Motion Parameters Used for Bio-inspired Navigation

Mohamad T. Alkowatly, Victor M. Becerra, and William Holderbaum
University of Reading/School of Systems Engineering, Reading, UK
Email: m.t.alkowatly@pgr.reading.ac.uk; {v.m.becerra,w.holderbaum}@reading.ac.uk

***Abstract*—Visual motion cues play an important role in animal and humans locomotion without the need to extract actual ego-motion information. This paper demonstrates a method for estimating the visual motion parameters, namely the Time-To-Contact (TTC), Focus of Expansion (FOE), and image angular velocities, from a sparse optical flow estimation registered from a downward looking camera. The presented method is capable of estimating the visual motion parameters in a complicated 6 degrees of freedom motion and in real time with suitable accuracy for mobile robots visual navigation.**

***Index Terms*—time-to-contact, focus of expansion, image angular velocity, optic flow models**

I. Introduction

A. Motivation

Since Gibson's work in [1] visual information and optic flow have been gaining increasing research interest in order to understand and solve visual locomotion problems. Gibson has highlighted the importance of some of the visual motion parameters like the location of the Focus of Expansion (FOE) and the dilation of optic flow to avoid or achieve a contact with an object. Image dilation plays an important role in estimating the Time-To-Contact (TTC) with observed objects [2]. Image angular velocity of ventral flow has been found to be employed by bees to regulate their speed and and distance to surrounding walls and their height while landing [3].

Such studies highlighted the importance of visual motion parameters and their role in locomotion without the need to estimate actual motion parameters such as the velocity or the position or even the knowledge of the structure of the scene. Many robotics systems implementa- tions exploited only visual motion information to achieve autonomous navigation. Image angular velocity has been used in [4] [5] to achieve horizontal autonomous landing. Image dilation has been used by [6] to achieve vertical landing. FOE and TTC have been used in [7] to implement a collision warning system and by [8] to implement a visual collision avoidance algorithms for mobile robots.

Manuscript received February 9, 2013; revised April 16, 2013.

B. Related Work

A good deal of work has been done on the estimation of the ego-motion parameters from an image sequence. A method proposed by the authors in [9] involves the calculation of the direction of motion, effectively the FOE, from the optic flow difference at points representing edges of depth variations, then use this point to calculate the motion angular velocity and the depth map, but not image angular velocities. This requires a dense cluttered scene and a reliable method of calculating the optic flow on such points. Heeger and Jepson [10] proposed a method to retrieve the translational and rotational motion as well as a depth map of the scene by solving for the direction of motion from a sub-sampled solution space then solving for rotation and depth. Their proposed method adds an unwanted complexity to retrieve the actual ego-motion parameters and depth map and their off-line computed coefficient scheme requires the optic flow estimates at constant points, something that might introduce bad optic flow measurements if such points have low contrast. A recursive method proposed by Barron and Eagleson [11] to solve for the translational and rotational ego-motion velocities, angular acceleration as well as depth map. The method is only tested in a restricted motion profiles where rotation takes place on one axis only.

Camus in [12] presented a method of solving for the FOE then the time-to-contact. He calculates the FOE by averaging the translational optic flow signs along the horizontal and vertical image directions separately. Then he calculates the time-to-contact from rate of expansion of optic flow from the FOE assuming that the motion is straight forward with no lateral motion. Similar approach taken by Sund are swaran et al. [7] to calculate the FOE and the TTC using the more reliable normal optic flow measurements. Both methods prefer a uniform distribution of optic flow measurements for an unbiased FOE location. They also assume a forward only motion for the FOE, and hence the TTC, to be calculated correctly.

A mathematical framework for computing the TTC and the camera angular velocities is presented by Micheli et al. [13]. The method finds the motion parameters from the eigenvalues of the Jacobian of the optic flow evaluated at the FOE. Unfortunately the most general motion this method can handle is an axial motion with

doi: 10.12720/joig.1.3.120-124

rotation axis constant and coincides with the translation axis besides it does not demonstrate a method for calculating the image angular velocities.

A method for calculating the FOE has been proposed by Teshima et al. [14] that does not rely on optic flow. Instead, the method iteratively estimates the expected new video frame from the previous frame and a motion model that depends on the location of the FOE by minimising the sum of absolute differences (SAD). Direct gradient based methods exploiting brightness constancy have been used to calculate the TTC without feature tracking in [15], however only translational motion is considered. Recently, a new method of estimating the TTC from the SIFT features scales is presented in [16]. In all these methods considerable computation is performed to estimate only one visual motion parameter. Having a separate method for calculating different visual motion parameters consecutively both exhausts available on-board processing power and propagates the errors from one stage as a ground truth for the next stage leading to undesirable and unrecoverable error accumulation. The aim of this paper is to present a simple method for simultaneously estimating all visual motion parameters, namely the FOE, TTC, and image angular velocities or ventral flow in an unrestricted 6 degrees of freedom (6-DoF) motion of a flying robot carrying a downward looking camera. The method should be able to operate in real-time and be flexible with the distribution of the textures in the image. An overview of the ego-motion and visual motion parameters will be presented in the next section, the proposed method is presented in section 3 and evaluation of the proposed method is presented in section 4.

II. Visual Motion Parameters

In a Cartesian coordinates frame, optic flow or the image velocities (u, v) at image point $p = (x, y) \in R^2$ is a function of the viewed object's real-world coordinates $P = (X, Y, Z) \in R^3$, the relative translational and angular velocities of the observer and viewed object $V = (V_x, V_y, V_z)$ and $W = (W_x, W_y, W_z)$ respectively and the projection plane distance from the projection point, or the focal length $f > 0$. This function has been called the ego-motion image velocities function.

The image velocities function due to ego-motion has been first described in [9]. Horizontal and vertical image velocities (u, v) can be described as a function of ego motion and projection focal length as follows:

$$u = -f(\frac{V_x}{Z} + W_y) + x\frac{V_z}{Z} + yW_z - x^2\frac{W_y}{f} + xy\frac{W_x}{f} \quad (1)$$

$$v = -f(\frac{V_y}{Z} - W_x) + y\frac{V_z}{Z} - xW_z + y^2\frac{W_x}{f} - xy\frac{W_y}{f} \quad (2)$$

where image point $x = X/Z, y = Y/Z$ is the projection of world point $P = (X, Y, Z)$ on the image plane.

If the optic flow has no rotational component ($W = (0, 0, 0)$) then the translational optic flow can be written as:

$$u_T = -f\frac{V_x}{Z} + x\frac{V_z}{Z} \quad (3)$$

$$v_T = -f\frac{V_y}{Z} + y\frac{V_z}{Z} \quad (4)$$

In this case the FOE is the only point where the optic flow vectors all coincides; hence it will be the only vanishing point of the optical flow vectors. We can find the image point coordinated x_{foe}, y_{foe} of the FOE from (3) and (4) by setting the optical flow vectors and angular velocities to zero.

$$x_{foe} = f\frac{V_x}{V_z} \qquad y_{foe} = f\frac{V_y}{V_z} \quad (5)$$

which is only defined when V_z is non-zero.
From (3, 4, 5) the optical flow translational velocities can be written as

$$u_T = (x - x_{foe})\frac{V_z}{Z} \quad (6)$$

$$v_T = (y - y_{foe})\frac{V_z}{Z} \quad (7)$$

which shows that the translational optical flow vectors exhibit pure dilation about the FOE. Equations (6) and (7) show the direction relationship between optical flow estimation and the direction of motion presented by the FOE as well as the TTC presented by image dilation d given by:

$$d = \frac{V_z}{Z} \quad (8)$$

which is clearly the inverse of the TTC.

Image angular velocity ω [rad s^{-1}] or the ventral flow has been defined by Srinivasan et al. [3] as:

$$\omega = \frac{V}{D} \quad (9)$$

where V [m s^{-1}] is the translational velocity along which the image angular velocity is measured and D [m] is the distance from the eye to the surface generating the visual features on the retina. Image angular velocities in both image directions ω_x, ω_y can be defined by using V_x, V_y respectively in equation 9. It is clear that in the case of a downward looking camera the image angular velocities are the scaled lateral velocities presented as the first terms in (1) and (2).

III. Optic Flow Modelling

By careful inspection of the ego-motion equations (1) And (2), the following parameters can be defined:

$$\begin{aligned} a_1 &= -f\left(\frac{V_x}{Z} + W_y\right) & a_2 &= \frac{V_z}{Z} \\ a_3 &= W_z & a_4 &= \frac{W_x}{f} \\ a_6 &= -f\left(\frac{V_y}{Z} - W_x\right) & a_5 &= -\frac{W_y}{f} \end{aligned} \tag{10}$$

which allows modelling the estimated optic flow using the ego-motion equations and the above model parameters as:

$$u = a_1 + a_2 x + a_3 y + a_4 xy + a_5 x^2 \tag{11}$$

$$v = a_6 - a_3 x + a_2 y + a_5 xy + a_4 y^2 \tag{12}$$

If the camera focal length **f** is known then the angular velocities of the observer can be deduced as follows:

$$\begin{aligned} W_x &= f a_4 \\ W_y &= -f a_5 \\ W_z &= a_3 \end{aligned} \tag{13}$$

The equations for TTC (T_c), lateral ventral flows (ω_x, ω_y), and the FOE location (x_{foe}, y_{foe}) can be found as follows:

$$\begin{aligned} T_c &= \frac{1}{a_2} \\ \omega_x &= -\frac{a_1}{f} - W_y \\ \omega_y &= -\frac{a_6}{f} + W_x \\ x_{foe} &= f\frac{\omega_x}{a_2} \\ y_{foe} &= f\frac{\omega_y}{a_2} \end{aligned} \tag{14}$$

Ego-motion model equations (11), (12) have six unknowns thus optic flow estimation at a minimum of three points are required to solve for the model parameters. This assumes that there is no depth variation in the scene, however with large number of optic flow measurements a minimal variation in the depth could be tolerated if the solution is found in a least square sense. The following system of equations could be written for n points:

$$\begin{bmatrix} u_1 \\ v_1 \\ : \\ u_n \\ v_n \end{bmatrix} = \begin{bmatrix} 1 & x_1 & y_1 & x_1 y_1 & x_1^2 & 0 \\ 0 & y_1 & -x_1 & y_1^2 & x_1 y_1 & 1 \\ : & : & : & : & : & : \\ 1 & x_n & y_n & x_n y_n & x_n^2 & 0 \\ 0 & y_n & -x_n & y_n^2 & x_n y_n & 1 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \\ a_6 \end{bmatrix} \tag{15}$$

And the values of model parameters a_1 to a_6 can be found using the least squares solution of the over-determined system defined in (15).

In addition to facilitate visual motion parameters estimation, the achieved optic flow model in a previous frame helps providing an initial guess to the optic flow estimation algorithm in the next frame by projecting the points in question using (1) and (2). By doing this both the reliability and the performance of the optic flow estimation is enhanced especially when the magnitude of optic flow becomes large due to large displacement or when the depth becomes small.

IV. Results

A dynamic virtual environment is built using VRML and integrated in a simulation environment using Simulink and Simulink 3D animation toolbox. The images gener- ated from the toolbox are (720×576) pixels and simulates a camera with (71°×49°)FoV. Optic flow measurements from generated image stream are calculated using Pyrami- dal implementation of Lukas and Kanade sparse optic flow at a maximum of 200 features chosen for best tracking using Harris corner [17] finder. Visual motion parameters are estimated and compared with ground truth values from the Simulink environment. A side-view of the simulated environment is shown in Fig. 1 where the simulated camera is carried below the aerial robot (top of Fig (1) and facing the ground. The tests are performed on an Intel Core i5 2.5 Ghz and was able to achieve a frame rate of 33 fps inclusive of the optic flow estimation.

The accuracy of the estimation is tested under two different motion profiles. The first is achieved by directly moving the camera through an axial motion with V =(0.75, 0.5, 1)[m s^{-1}] and W= (0, 0, 0.1745)[rad s^{-1}]. The theoretical vs. the visually registered values for visual motion parameters TTC ω_x and x_{foe} are shown in Fig 2, 3, 4 respectively

The second motion profile is achieved by controlling a helicopter platform carrying the camera to move in a sinusoidal translational and rotational motion resulting in

Motion profile shown in Fig. 5. The theoretical vs. the visually registered values for visual motion parameters TTC, ω_x and x_{foe} are shown in Fig. 6, 7, 8 respectively.

It is clear that the method produces very good estimates in the first motion profile while the quality of the estimated degrades slightly in the second motion profile. This reduced accuracy is due to invalidating the assumption of a uniform depth values for all image points when the camera tilts. However if the camera orientation is not expected to vary significantly the achieved accuracy is suitable for the Purpose of robot navigation.

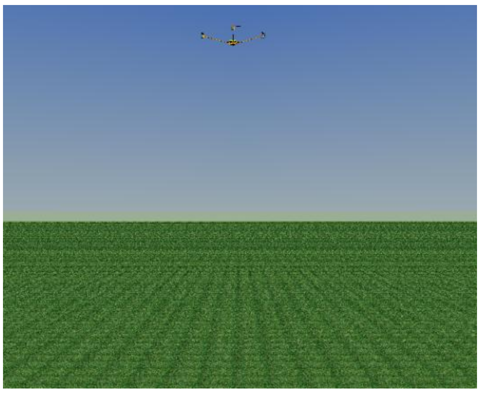

Figure 1. Side-view of the simulated environment

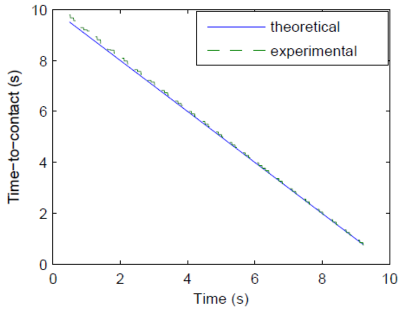

Figure 2. Theoretical vs experimental TTC values in motion profile 1

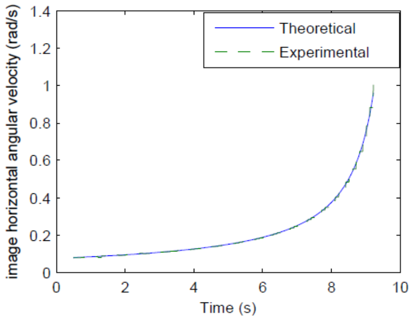

Figure 3. Theoretical vs experimental ω_x values in motion profile 1

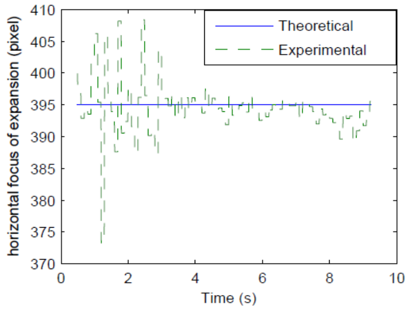

Figure 4. Theoretical vs experimental x_{foe} values in motion profile 1

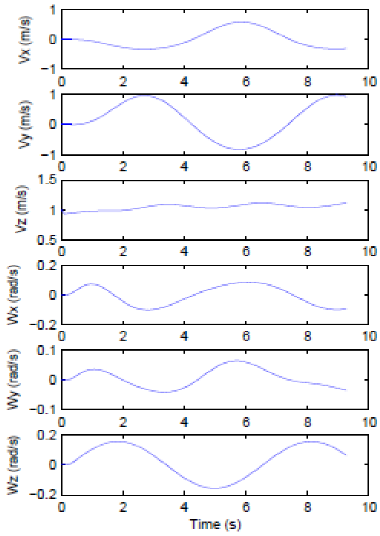

Figure 5. Translational and angular velocities of test motion profile 2

In order to quantitatively measure the estimation accuracy, the root mean square error (RMSE) of the five estimated visual motion parameters against their theoretic cal values is calculated and shown in Table I for the two motion profiles defined above. The horizontal and vertical directionhx, hy in degrees are included for convenience.

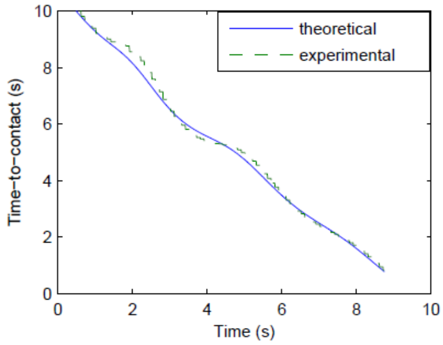

Figure 6. Theoretical vs experimental TTC values

V. Conclusions and Future Work

In this paper a simple closed form method for estimating visual motion parameters, namely the time-to-contact, focus of expansion, and image angular velocities from a general 6 degrees of freedom camera motion. The proposed method uses sparse optic flow estimates at arbitrary image location allowing exploiting image textures in each frame. All parameters are estimated simultaneously rather than in stages to prevent error accumulation. The method managed to accurately estimate the required parameters in real-time.

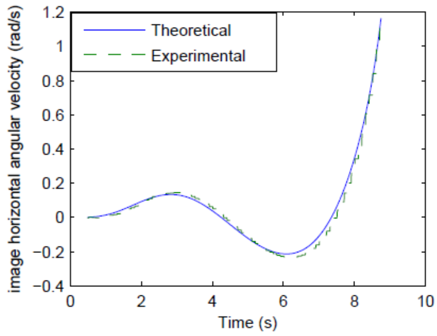

Figure 7. Theoretical vs experimental ω_x values

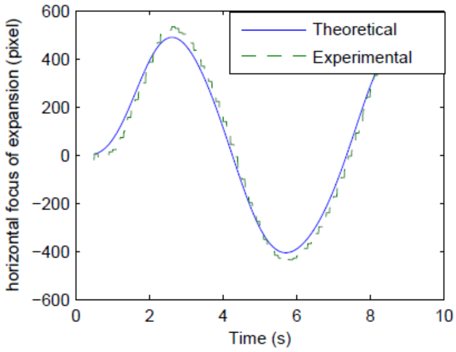

Figure 8. Theoretical vs experimental x_{foe} values

TABLE I. ROOT MEAN SQUARE OF THE ESTIMATED PARAMETERS

Parameter	Motion 1	Motion 2	Unit
T_c	0.097	0.1981	seconds
ω_x	0.0075	0.0284	rad s^{-1}
ω_y	0.0052	0.0127	rad s^{-1}
x_{foe}	4.3924	36.2498	pixels
y_{foe}	7.6814	16.9793	pixels
h_x	0.4331	3.5746	degree
h_y	0.7841	1.7333	degree

Future work should find ways to address degradation in estimation accuracy due to variation in depth due to multiple planar objects or slant surfaces possibly due to camera orientation without adding the complexity of resolving depth itself.

REFERENCES

[1] J. J. Gibson, "Visually controlled locomotion and visual orienttation in animals," *British Journal of Psychology*, vol. 49, no. 3, pp. 182–194, 1958.

[2] D. N. Lee, "A theory of visual control of braking based on information about time-to-collision," *Perception*, vol. 5, no. 4, pp. 37–459, 1976.

[3] M. Srinivasan, S. Zhang, M. Lehrer, and T. Collett, "Honeybee navigation en route to the goal: visual flight control and odometry," *Journal of Experimental Biology*, vol. 199, no. 1, pp. 237–244, 1996.

[4] M. V. Srinivasan and M. Ibbotson, "Biologically inspired strategies, algorithms and hardware for visual guidance of autonomous helicopters," *Tech. Re*p. 2011.

[5] F. Valette, F. Ruffier, S. Viollet, and T. Seidl, "Biomimetic optic flow sensing applied to a lunar landing scenario," *IEEE, 2010*, pp. 2253–2260.

[6] X. B. Herisse, T. Hamel, R. Mahony, and F. X. Russotto, "Landing a VTOL unmanned aerial vehicle on a moving platform using optical flow," *IEEE Transactions on Robotics*, vol. 28, no. 1, pp. 77–89, 2012.

[7] V. Sundareswaran, S. A. Beardsley, and L. M. Vaina, "12. Fast Pro-cessing of image motion patterns arising from 3-D translational motion," in *Optic Flow and Beyond*, 2004, vol. 324, pp. 273.

[8] K. Souhila and A. Karim, "Optical flow based robot obstacle avoidance," *International Journal of Advanced Robotic Systems*, vol. 4, no. 1, pp. 13–16, 2007.

[9] H. C. Longuet-Higgins and K. Prazdny, "The interpretation of a moving retinal image," in *Proc. Royal Society of London. Series B. Biological Sciences*, vol. 208, no. 1173, pp. 385–397, 1980.

[10] D. J. Heeger and A. D. Jepson, "Subspace methods for recovering rigid motion I: Algorithm and implementation," *International Journal of Computer Vision*, vol. 7, no. 2, pp. 95–117, 1992.

[11] J. Barron and R. Eagleson, "Recursive estimation of time-varying motion and structure parameters," *Pattern Recognition*, vol. 29, no. 5, pp. 797–818, May 1996.

[12] T. Camus, "Calculating time-to-contact using real-time quantized optical flow," *Tech. Rep.*, 1995.

[13] E. Micheli, V. Torre, and S. Uras, "The accuracy of the computation of optical flow and of the recovery of motion parameters," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 15, no. 5, pp. 434–447, May 1993.

[14] T. Teshima, H. Saito, S. Ozawa, K. Yamamoto, and T. Ihara, "Estimation of foe without optical flow for vehicle lateral position detection," in *Proc. IAPR Conference on Machine Vision Applications*, 2005, pp. 406–409.

[15] B. K. Horn, Y. Fang, and I. Masaki, "Time to contact relative to a planar surface," in *IEEE Intelligent Vehicles Symposium*, Jun. 2007, pp. 68–74.

[16] G. de Croon, D. Izzo, and G. Schiavone, "Time-to-Contact estimation in landing scenarios using feature scales," *Acta Futura*, vol. 5, pp. 73–82, 2012.

[17] C. Harris and M. Stephens, "A combined corner and edge detector," in *Alvey Vision Conference*, vol. 15. Manchester, UK, 1988, pp. 50.

Mohamad T. Alkowatly is a current PhD student in Cybernetics researching visual bio-inspired aerial guidance and navigation strategies. He has an MSc in Aeronautics Engineering (2009, UK) and a BSc in Informatics Engineering (2006, Syria).

Victor Becerra is a Professor in Automatic Control. His research interests include computational optimal control, nonlinear control, system identification, robotics, and artificial intelligence.

William Holderbaum is a Senior Lecturer in Mathematical Engineering. His research interests include boolean input systems, geometric control, rehabilitation engineering, and smart grid systems. The are all with the School of Systems Engineering at University of Reading, UK.