NISTIR 5442

A Visual-Motion Fixation Invariant

Daniel Raviv
Florida Atlantic University
The Robotics Center and
The Electrical Engineering Department
Boca Raton, Florida
and
Sensory Intelligent Group
Intelligent Systems Division

and

Nissim Ozery
Florida Atlantic University
The Robotics Center and
The Electrical Engineering Department
Boca Raton, Florida

U.S. DEPARTMENT OF COMMERCE
Technology Administration
National Institute of Standards
and Technology
Bldg. 220 Rm. B124
Gaithersburg, MD 20899

NIST

NISTIR 5442

A Visual-Motion Fixation Invariant

Daniel Raviv
Florida Atlantic University
The Robotics Center and
The Electrical Engineering Department
Boca Raton, Florida
and
Sensory Intelligent Group
Intelligent Systems Division

and

Nissim Ozery
Florida Atlantic University
The Robotics Center and
The Electrical Engineering Department
Boca Raton, Florida

U.S. DEPARTMENT OF COMMERCE
Technology Administration
National Institute of Standards
and Technology
Bldg. 220 Rm. B124
Gaithersburg, MD 20899

June 1994

U.S. DEPARTMENT OF COMMERCE
Ronald H. Brown, Secretary

TECHNOLOGY ADMINISTRATION
Mary L. Good, Under Secretary for Technology

NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY
Arati Prabhakar, Director

A VISUAL-MOTION FIXATION INVARIANT

Daniel Raviv[1,2] and Nissim Ozery[1]

[1]Robotics Center and Department of Electrical Engineering,
Florida Atlantic University (FAU), Boca Raton, FL 33431
e-mail: ravivd@acc.fau.edu
[2]Intelligent Systems Division, National Institute of Standards and Technology (NIST)
Bldg. 220, Room B124, Gaithersburg, MD 20899

ABSTRACT

This paper deals with a visual-motion fixation invariant. We show that during fixation there is *a measurable nonlinear function* of optical flow that produces the *same* value for all points of a stationary environment *regardless* of the 3-D shape of the environment. During fixated camera motion relative to a rigid object, e.g., a stationary environment, the projection of the fixated point remains (by definition) at the same location in the image, and all other points located on the 3-D rigid object can only rotate relative to the 3-D fixation point. This rotation rate of the points is *invariant* for all points that lie on the particular environment, and it is measurable from a sequence of images. This new invariant is obtained from a set of monocular images, and is expressed explicitly as a *closed form solution*. In this paper we show how to *extract* this invariant analytically from a sequence of images using optical flow information, and we present results obtained from real data experiments.

This work was supported in part by a grant from the National Science Foundation, Division of Information, Robotics and Intelligent Systems, Grant # IRI-9115939

4. It exists during fixation, and the fixation point can be chosen *arbitrarily.*
5. It is *a measurable,* nonlinear function of optical flow, i.e., it is measured from visual data in camera coordinates.
6. It produces the *same* value for all points of a 3-D rigid object *regardless* of its 3-D shape, i.e., it is valid for any structure in the stationary environment.
7. There is no need for 3-D reconstruction.
8. The invariant is obtained from a closed form expression, and it is measured in *time* units rather than *distance* units.
9. Using a logarithmic retina, this invariant can be obtained directly, i.e., *without* many additional computations.

2: A fixation invariant

The following assumptions are made in the derivation of the invariant: a) The camera axis of rotation is perpendicular to the line that connects the fixation point with the pinhole point of the camera. b) There is no relative translation between the fixation point and the camera. c) The angular velocity of the camera is constant. d) The angles extracted from the image are approximately equal to the angles extracted from an orthographic projection plane. Practically, in order to achieve similar projection, a narrow field of view camera should be used, and the observed object has to be placed far away from the camera.

Figure 1 shows a moving camera fixated on a stationary object. Without loss of generality, we chose to derive the invariant for a stationary camera fixated on a point located on a rotating 3-D object (Figure 2). This choice simplifies the derivation of the invariant, as well as the experimental set up.

In Figure 2 points A, B and F are points on the 3-D object, where point F marks the 3-D fixation point. Note that points A, B and F need not be on the same horizontal plane. Points F', A' and B' are the projections of points F, A and B on the projection plane, respectively, and points F*, A* and B* are the projections of points F', A' and B' on the image plane, respectively. The axis of rotation of the object is perpendicular to the page . Note that for the chosen projection the following derivation is independent of the distance **l** between F*-F'.

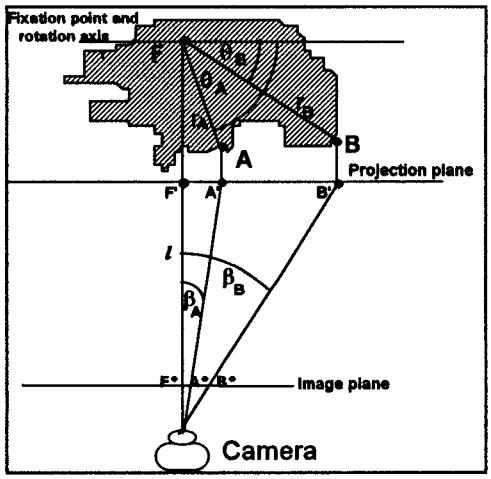

Figure 2: Angles extracted from projection plane.

From Figure 2, for any point on the rigid body, (assuming $\beta \neq 0$, and using $\beta = \beta(t)$ and $\theta = \theta(t)$) the following relationship holds:

$$\tan\beta = \frac{r\cos\theta}{l} \tag{1}$$

Taking the derivative of both sides of Equation (1) with respect to time, we get:

$$\frac{\dot{\beta}}{\cos^2\beta} = \frac{-r}{l}\omega\sin\theta \tag{2}$$

where the dot above the variable represents derivative with respect to time, and $\omega = \frac{d\theta}{dt}$ is the unknown angular velocity of the rigid object about the fixation point, which is unknown. Dividing Equation (2) by Equation (1) and after manipulating the results we obtain:

$$\frac{\dot{\beta}}{\cos^2\beta \cdot \tan\beta} = -\omega \cdot \tan\theta \tag{3}$$

Therefore,

$$\theta = \tan^{-1}\left(\frac{-2\dot{\beta}}{\omega\sin(2\beta)}\right) \tag{4}$$

Specifically, for two visible points A and B at time instant t, we can write:

$$\theta_A(t) = \tan^{-1}\left(\frac{-2\dot{\beta}_A(t)}{\omega\sin[2\beta_A(t)]}\right), \tag{5}$$

and

$$\theta_B(t) = \tan^{-1}\left(\frac{-2\dot{\beta}_B(t)}{\omega\sin[2\beta_B(t)]}\right), \tag{6}$$

where subscripts A and B correspond respectively to points A and B of the rigid object respectively at time instant t. Subtracting Equation (6) from Equation (5) at time instants t_1 and t_2 yield Equations (7) and (8) respectively:

$$\theta_A(t_1) - \theta_B(t_1) = \tan^{-1}\frac{2\dot{\beta}_A(t_1)}{\omega\sin(2\beta_A(t_1))} - \tan^{-1}\frac{2\dot{\beta}_B(t_1)}{\omega\sin(2\beta_B(t_1))} \tag{7}$$

$$\theta_A(t_2) - \theta_B(t_2) = \tan^{-1}\frac{2\dot{\beta}_A(t_2)}{\omega\sin(2\beta_A(t_2))} - \tan^{-1}\frac{2\dot{\beta}_B(t_2)}{\omega\sin(2\beta_B(t_2))} \tag{8}$$

Since $(\theta_A(t) - \theta_B(t))$ is *constant at all time instants*, one can equate Equations (7) and (8) to get:

$$\tan^{-1}\frac{2\dot{\beta}_A(t_1)}{\omega\sin[2\beta_A(t_1)]} - \tan^{-1}\frac{2\dot{\beta}_B(t_1)}{\omega\sin[2\beta_B(t_1)]} = \tan^{-1}\frac{2\dot{\beta}_A(t_2)}{\omega\sin[2\beta_A(t_2)]} - \tan^{-1}\frac{2\dot{\beta}_B(t_2)}{\omega\sin[2\beta_B(t_2)]}. \tag{9}$$

Using the identity

$$\tan(\alpha - \beta) = \frac{\tan\alpha - \tan\beta}{1 + \tan\alpha\tan\beta}, \tag{10}$$

Equation (9) can be written as follows:

$$\frac{\frac{a_1}{\omega}-\frac{b_1}{\omega}}{1+\frac{a_1 b_1}{\omega^2}}=\frac{\frac{a_2}{\omega}-\frac{b_2}{\omega}}{1+\frac{a_2 b_2}{\omega^2}} \tag{11}$$

where:

$$a_1=\frac{2\dot{\beta}_A(t_1)}{\sin[2\beta_A(t_1)]} \tag{12}$$

$$b_1=\frac{2\dot{\beta}_B(t_1)}{\sin[2\beta_B(t_1)]} \tag{13}$$

$$a_2=\frac{2\dot{\beta}_A(t_2)}{\sin[2\beta_A(t_2)]} \tag{14}$$

$$b_2=\frac{2\dot{\beta}_B(t_2)}{\sin[2\beta_B(t_2)]} \tag{15}$$

Manipulations of Equation (11) yield the following:

$$\boxed{\omega^2=\frac{a_1 b_1(a_2-b_2)-a_2 b_2(a_1-b_1)}{(a_1-b_1)-(a_2-b_2)}} \tag{16}$$

Note that a_1, b_1, a_2 and b_2 have to be finite numbers, and $(a_1-b_1)\neq(a_2-b_2)$.

The meaning of Equation (16) is that during fixation there is a *measurable* nonlinear function of optical flow that produces the *same* value (invariant) for all points of a stationary environment *regardless* of the 3-D structure of the environment. *This invariant is* ω. Note that the solution does not distinguish between clockwise and counter-clockwise rotations.

An interesting fact arises from the derivation of the invariant. The expressions in Equations (12)-(15) can be obtained linearly using a *logarithmic retina,* since the variables a_1, a_2, b_1 and b_2 are derivatives of a logarithmic expressions. Specifically, we have

$$\frac{2\dot{\beta}}{\sin(2\beta)}=\frac{d}{dt}\{\log[\tan(\beta)]\} \tag{17}$$

i.e., these coefficients can be obtained directly from a linear change of visual data as measured by a logarithmic retina [6].

3: The experiment

In this section we describe: 1) Simulations used to verify the result of the derivation and to test the theoretical limitations of the method. 2) Real data measurements using a Theodolite to test some basic practical aspects of the approach. 3) Real data measurements using a CCD video camera.

3.1: Simulation

We used computer simulations to verify the result in Equation (16), and to test the theoretical limitations of the method. The location of two points on an arbitrary rotating object were simulated, and the angular velocity of the object was obtained using Equations (12)-(16). The points on the object (A and B) were selected arbitrarily with random radii (r_A and r_B), and with random initial phase (θ_A and θ_B). During the simulated motion the program selects four consecutive time instants (t_1, t_2, t_3 and t_4), where $\Delta t=(t_2-t_1)=(t_4-t_3)\ll(t_3-t_2)$. At each one of these four time instants the program obtained

the spatial location of the points A and B, and using simple vector analysis it calculated four angles: $\beta_A(t_1)$, $\beta_B(t_1)$, $\beta_A(t_3)$ and $\beta_B(t_3)$, and their approximated derivatives:

$$\dot{\beta}_A(t_1) \cong \frac{\beta_A(t_2)-\beta_A(t_1)}{\Delta t} \tag{18}$$

$$\dot{\beta}_B(t_1) \cong \frac{\beta_B(t_2)-\beta_B(t_1)}{\Delta t} \tag{19}$$

$$\dot{\beta}_A(t_3) \cong \frac{\beta_A(t_4)-\beta_A(t_3)}{\Delta t} \tag{20}$$

$$\dot{\beta}_B(t_3) \cong \frac{\beta_B(t_4)-\beta_B(t_3)}{\Delta t} \tag{21}$$

These angles and their derivatives were substituted in Equations (12)-(15) to obtain the final result which is the *angular velocity* of the rotating object.

We varied some simulation factors such as: (t_3-t_2), $\omega\Delta t$, l, the ratio (r_A/r_B), the ratio $(\Delta t/\omega)$ and the ratio (θ_1/θ_2). Table 1 shows the effect of different values of Δt on the results for angular velocity of 10 rad/s. Values of Δt vary from 10^{-5} to 10^{-1} seconds, and as expected the error increases as Δt increases with larger Δt. However, even for the case of $\Delta t = 0.1$ seconds, i.e., $\omega \cdot \Delta t = 1$ radian, the error in the result is only about 4%.

These various tests verified the theoretical results, and showed some limitations of the method. There are three situations that may result in large errors in the analysis of the angular velocity:

(a) $(t_3 - t_2)\cdot\omega = n\pi$, where $n = \pm1, \pm2, \pm3, \cdots$.
In this case, points A, B appear at opposite locations relative to the rotation axis, and therefore no new information about the motion of the object is gained.

(b) $\theta_A + \theta_B = \pi + 2n\pi$, where $n = \pm1, \pm2, \pm3, \cdots$.
This causes a singularity point in Equation (16).

(c) Δt is not "small enough". This results in an error in estimating the optical flow $d\beta/dt$.

3.2: Theodolite measurements

In this part of the experiment a theodolite is used to measure the β angles. Figure 3 shows an object held by a robot. Points A, B and F were marked on the object by the tip of pins that were inserted into the object like in Figure 4, and the theodolite was located about 1.2 meters away from the axis of rotation.

During the experiment, the robot rotated the object to four different orientations. The β angles, viewed by the theodolite, were recorded at each orientation to yield the data that is necessary for the analysis of the angular velocity, i.e., angles $\beta_A(t_1)$, $\beta_B(t_1)$, $\beta_A(t_2)$, $\beta_B(t_2)$, $\beta_A(t_3)$, $\beta_B(t_3)$, $\beta_A(t_4)$, and $\beta_B(t_4)$.

Note that there is no real time dimension in this experiment since we stop the robot for every measurement. If one assumes that the angular velocity is ω, then the time dimension can be obtained by calculating Δt using $\Delta t = \Delta\theta/\omega$. $\dot{\beta}$ can be approximated using the angles extracted from the theodolite divide by Δt as in Equations (18)-(21). Using these equations the ω obtained from the visual data and the calculated Δt, can be compared with the assumed angular velocity.

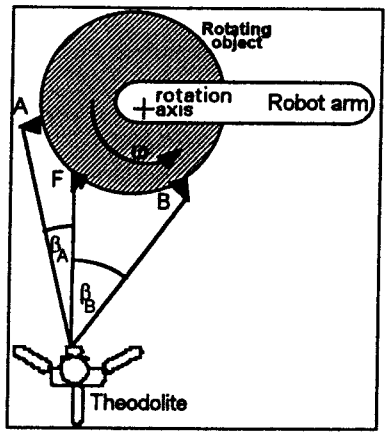

Figure 3: The theodolite experimental setup

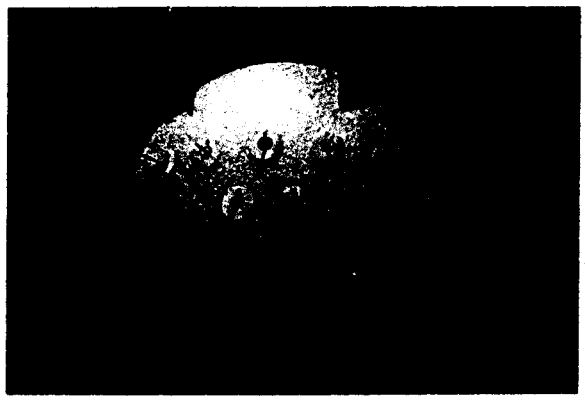

Figure 4: The object used in the theodolite experiment.

The real data results show that the method is quite robust. For 24 different experiments the average error in ω is 7.4% as shown in Table 2.

3.3: Camera measurements

In this part of the experiment we used the suggested method to extract the angular velocity of a rotated object using a CCD video camera. The experimental set-up is shown in Figure 5.

The following are some details of the experimental set-up: The distance between camera and the fixation point is 150 cm, and the camera's field of view is about 9 degrees. The object has a minimum radius of 22 cm. In order to simplify the image processing, the object has well defined dark vertical strips located on a white background, for a high contrast, as shown in Figure 6. The object is rotated at an angular velocity of 0.021 rad/s.

During the experimental process we found some practical limitations: (a) The value of $\omega\Delta t$ should be large enough so that the point being tracked passes through a "sufficient number of pixels" to decrease the effect of the discretization error, yet theoretical limitations impose an upper bound on Δt.

(b) Points that produce large value of optical flow cause larger errors in the computed ω. This is due to the fact that these points produce little change in optical flow.

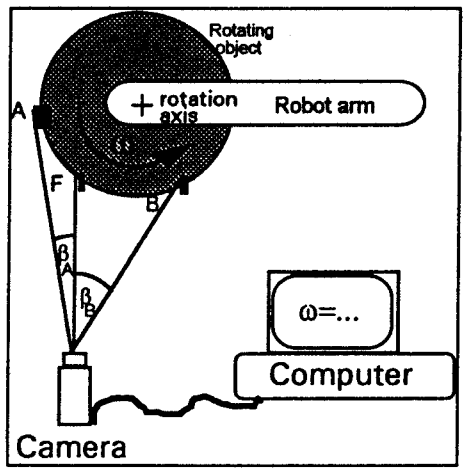

Figure 5: Camera measurements experimental setup

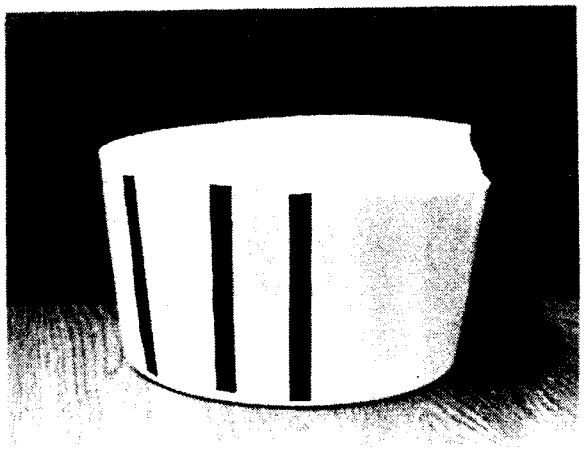

Figure 6: The object used in the camera experiment.

Under these experimental conditions the average error in ω over 24 experiments is about 3.5% as shown in Table 3. In every experiment the processing was done over 20 different sets of points simultaneously, and the results were averaged. The real angular velocity of the robot was 0.021 rad/s.

ω =	10.0	10.0	10.0	10.0	10.0	10.0	10.0	10.0
Δt =	0.00001	0.00001	0.001	0.001	0.01	0.01	0.1	0.1
Result =	10.0005	10.0005	10.0431	10.0387	9.9753	9.9493	10.2746	10.4344

Table 1: The effect of Δt on the error in ω. True ω is 10 rad/s.

Exp. #	1	2	3	4	5	6	7	8
Result	1.006	1.010	0.956	1.092	1.016	1.140	0.882	1.225

Exp. #	9	10	11	12	13	14	15	16
Result	1.103	0.958	1.002	0.845	1.059	0.989	0.945	0.953

Exp. #	17	18	19	20	21	22	23	24
Result	1.005	0.997	1.007	1.446	0.979	0.966	1.071	0.978

Table 2: Results of ω using data obtained from theodolite. True ω is 1 rad/s.

Exp. #	1	2	3	4	5	6	7	8
Result	0.0203	0.0210	0.0224	0.0213	0.0228	0.0195	0.0226	0.0225

Exp. #	9	10	11	12	13	14	15	16
Result	0.0219	0.0249	0.0221	0.0220	0.0221	0.0226	0.0231	0.0247

Exp. #	17	18	19	20	21	22	23	24
Result	0.0215	0.0218	0.0201	0.0199	0.0210	0.0216	0.0198	0.0201

Table 3: Results of ω using data obtained from camera. True ω is 0.021 rad/s.

4: Conclusions and Future work

In this paper, it has been shown that during the process of camera fixation there is a scene independent visual motion invariant. The result is stated in a closed form, and can be obtained using the optical flow of only two points at two different time instants. The results of the described experiments, obtained from both simulated and real data, are highly encouraging. Currently we are extending the method to a general motion and textured environments using a feature tracking method [10].

5: Acknowledgment

The authors would like to thank Marty Herman, Jim Albus, Ernie Kent and Don Orser of NIST for fruitful discussions regarding this invariant as well as other invariants [5] and Srivatsan Krishnan and Sridhar Kundur of FAU for important comments on this paper.

6: References

[1] Albus, J. S., "Outline for a Theory of intelligence", IEEE Transactions On Systems, Man and Cybernetics, vol. 21, No 3, 1991.

[2] Cutting, J., "Perception With an Eye for Motion", MIT Press, 1986.

[3] Gibson, J.J., "New Reasons for Realism", Synthesis 17, pp. 162-167,1967.

[4] Raviv, D., "A Quantitative Approach To Camera Fixation", In Proceedings of Computer Vision and Pattern Recognition (CVPR) Conference, Hawaii, June 1991.

[5] Raviv, D., "Invariants In Visual Motion", National Institute of Standards and Technology, Technical Report NISTIR 4722, 1991.

[6] Raviv, D., Albus, J.S., Orser, D., "On Logarithmic Retinae", National Institute of Standards and Technology, Technical Report NISTIR 92-4807.

[7] Robert Cipolla, Andrew Black, "Surface Orientation and Time to Contact from Image Divergence and Deformation", Department of Engineering science., University of Oxford, OX1, England, Proceedings of ECCV, 1992.

[8] Weinshall D. , "Direct Computation of Qualitative 3-D shape and Motion invariants", IEEE Transactions On Pattern Analysis and Machine Intelligence, Vol. 13, No 12, December 1991.

[9] Weinshall D. , "Model-Based Invariant for 3-D Vision ", IEEE Computer society Conference On Computer Vision and Pattern Recognition, 1993.

[10] Krishnan S. and Raviv D., "Adaptive 2-D Feature Tracking Algorithm for Finding The Focus of Expansion", To be presented at the Conference on Recent Advances in Robotics, 1994.