Geometric Approach for Egomotion Estimation using Normal Flow

César Silva[⋆] **José Santos-Victor**

Instituto de Sistemas e Robótica - Instituto Superior Técnico,
Av. Rovisco Pais, 1096 Lisboa Codex, PORTUGAL
e-mail: {cesar,jasv}@isr.ist.utl.pt

Abstract. In this paper, we address the problem of egomotion estimation for a monocular moving observer, under arbitrary translation and rotation. The method we propose is uniquely based on the spatio-temporal image derivatives, and has $O(n^2)$ complexity. The approach consists in searching the image for particular geometric properties of the normal flow. The algorithm is performed in two successive steps: the **Ψ-line search algorithm** determines the direction of the Focus of Expansion (FOE) and computes some information on the rotation parameters; the **Φ-circle search algorithm** determines the FOE, and the individual values of the rotation. Various experiments with synthetic and real images are presented and discussed.

1 Introduction

We address the problem of egomotion estimation for a monocular moving observer. This is a key step for the navigation of any autonomous system.

The problem of egomotion estimation consists in determining the 3D motion parameters, by observing an image sequence over time. Due to the well known depth-speed ambiguity, we can only aim at retrieving the angular velocity, $\boldsymbol{\omega}$, and the direction of linear motion, $\mathbf{t}$, in a total of 5 unknowns ($\mathbf{t}$, can only be fully recovered provided that depth is known).

The first step to estimate egomotion or structure from motion is the computation of displacement vector fields of an image sequence acquired by a moving camera. Usual approaches are based either on correspondences between special image features (edge points, corners, etc) along the sequence [1]; or on the estimation of the dense motion field, identified by the optical flow [4]; or finally in the so-called direct methods [2, 7, 9, 14]. The correspondence problem or the optical flow estimation are, in general, ill-posed [3, 5], and it is usually necessary to make very restrictive assumptions about the observed scenes [15, 16]. The solutions obtained are often unstable and require large amounts of computation. The direct methods are less demanding than the previous ones, and use the image brightness information directly to recover the motion parameters. These methods are usually based on normal flow [10, 12] or the spatio-temporal image derivatives [11].

For the first stage, several direct methods [2, 7, 9, 14] have been proposed, based on the normal-flow field. The computation of the normal flow does not require any special assumption about the scene structure. It represents the only image motion component that can be estimated from local measurements alone, due to the "aperture problem" [10].

In the second stage, the motion parameters are recovered from the image flow data. Several methods estimate only the heading direction, for bounded rotation. Aloimonos and Duric [2] proposed a qualitative solution to the problem, by using a set of geometric constraints from the normal flow measurements. Similarly, Sinclair et al. [14] presented an algorithm to recover the Focus of Expansion from the normal flow, that is tolerant to rotational motion. These methods demonstrated some robustness, but their efficiency is strongly dependent on the angular velocity values. On the other hand, they do not estimate the rotation motion parameters.

Fermuller and Aloimonos in [9, 7, 8] introduce a method which is based on a selection of image points that form global patterns in the image plane, using the orientation and the sign of the global normal-flow vectors. This approach considers the complete egomotion estimation problem as a pattern recognition problem, introducing a set of geometric constraints that reveal a global structure hidden in normal-flow fields. Considering only the qualitative part of the algorithm presented in [7], the method does not compute the third component of rotational movement (the rotation around the Z axis) and has an appreciable complexity hidden on the pattern recognition algorithm.

⋆ César Silva is supported by a grant from PRAXIS-XXI research program.This work was partially funded by projects PRAXIS/3/3.1/TPR/23/94 and JNICT-PBIC/TPR/2550/95.

The method proposed here for egomotion estimation is uniquely based on spatio-temporal image derivatives. We assume that the observer can move freely in any direction with general rotation, in an arbitrary static environment with textured objects. The egomotion parameter estimation is made in a passive way, although it might benefit from considering an active observer.

The approach we follow consists in searching the image for particular geometric properties of the normal flow, tightly connected to the egomotion parameters. Similarly to the work of Fermuller, rather than considering the whole set of image flow data, we use only image sites, that have special geometric properties, which convey relevant information about the observer motion. This algorithmic approach has some similarities with previous work [13].

Throughout the paper, we consider a camera-centered coordinate system, as depicted in Figure 1. The image

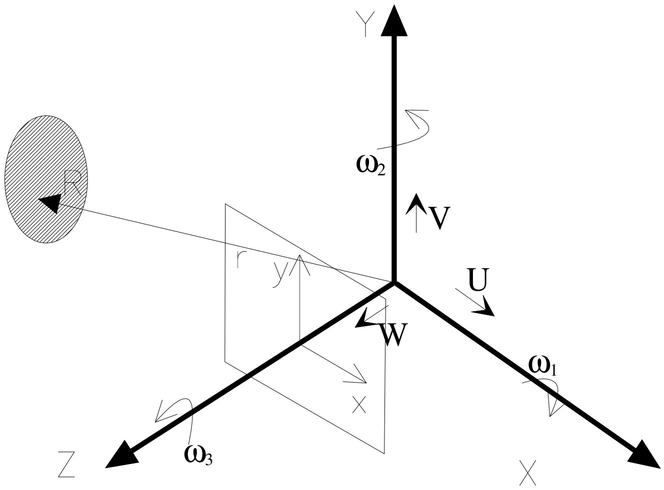

Fig. 1.: Imaging geometry and motion representation.

formation is modeled as a perspective projection.

Given the camera linear velocity $\mathbf{t} = [U\ V\ W]^T$, angular velocity $\boldsymbol{\omega} = [\omega_1\ \omega_2\ \omega_3]^T$ and the scene depth, Z, of each corresponding scene point, the motion field induced in the image plane, can be calculated for every image point (x, y) by the following well-known equations [6] :

$$\begin{cases} u = \frac{-Uf+xW}{Z} + \omega_1 \frac{xy}{f} - \omega_2(\frac{x^2}{f} + f) + \omega_3 y \\ v = \frac{-Vf+yW}{Z} + \omega_1(\frac{y^2}{f} + f) - \omega_2 \frac{xy}{f} - \omega_3 x \end{cases}, \tag{1}$$

where u and v are the optic flow components along the X and Y axis.

The Focus of Expansion (FOE) is defined as $(\sigma,\ \eta) = (fU/W, fV/W)$, which corresponds to the projection of the observer linear velocity, $\mathbf{t}$, in the image plane. In order to simplify the notation we define the function $\rho(x, y)$ as the inverse of the time-to-crash, given by $\rho(x, y) = W/Z(x, y)$.

Our method works in two steps :

1. The **Ψ-line search algorithm** determines the direction of the FOE vector defined by $(\sigma,\ \eta)$, estimates ω_3, and establishes a linear constraint in $(\omega_1,\ \omega_2)$. This algorithm uses the normal flow vectors that are perpendicular to the lines passing through the image center. These vectors are called the *circular normal flow*.
2. The **Φ-circle search algorithm** finds a circle (the Φ-circle) containing the FOE and uses the normal flow vectors that are collinear with the lines passing through the origin. We call these vectors the *radial normal flow*. In this step, we use the results provided by the Ψ-line search algorithm. The remaining 3-D motion parameters are herein retrieved.

The egomotion parameters are estimated via an algorithm of $O(n^2)$ complexity (where the image size is $n \times n$ pixels). The estimation method will be explained in the following sections.

2 Estimating the Motion Parameters

2.1 Determining the FOE direction : the Ψ-line search algorithm

Let us assume that the Focus of Expansion (FOE) is located along a direction of Ψ radians. We define the *Ψ-line* as the line joining the image center and the FOE, hence making an angle Ψ with the horizontal axis.

We compute the *circular* component of the optical flow, along lines passing through the image center with a general direction ψ :

$$\forall(x,y) : (x,y) = r(\cos\psi, \sin\psi), \qquad U_\psi(r) = (u,v)\cdot(-\sin\psi, \cos\psi), \tag{2}$$

where r denotes the distance from point (x,y) to the image center. Therefore, U_ψ gives the component of the flow that is perpendicular to the line passing through the image center with direction ψ. Figure 2 illustrates a sample of a normal flow field, and the circular normal flow field (the normal flow vectors perpendicular to the set of lines passing through the image center).

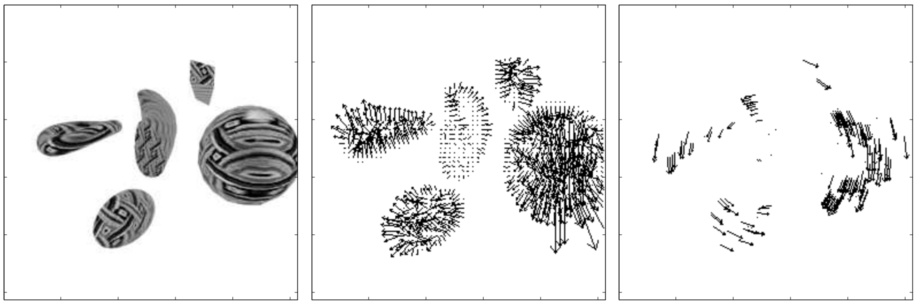

Fig. 2.: Sample image from the sequence; corresponding normal-flow field; and the circular normal flow.

Computing U_ψ for $\psi \in [-\pi/2, \pi/2[$, we have :

$$\begin{aligned} &\forall(x,y) \ : \ (x,y) = r(\cos\psi, \sin\psi) \\ &U_\psi(r) = -\rho(x,y)(\sigma,\eta)\cdot(-\sin\psi,\cos\psi) + f(\omega_1\cos\psi + \omega_2\sin\psi) - \omega_3 r \end{aligned} \tag{3}$$

U_ψ is composed by an affine term in r, and a term that, in general, is non-linear in r. However, this non-linear component vanishes both when ψ corresponds to the FOE direction (i.e. $\psi = \Psi$), or when the FOE is located at the image origin. Then U_Ψ becomes affine in r :

$$\forall(x,y) : (x,y) = r(\cos\Psi, \sin\Psi), \qquad U_\Psi(r) = f(\omega_1\cos\Psi + \omega_2\sin\Psi) - \omega_3 r. \tag{4}$$

The rationale of the algorithm is to search a straight line passing through the image origin, such that $U_\psi(r)$ is affine in r. To estimate the Ψ-line parameters, we need at least three points, but it is convenient to use more points, to increase robustness. The direction corresponding to the FOE minimizes the mean square error, when approximating $U_\psi(r)$ to an affine function of r. The *Ψ-line algorithm* "visits" only once each pixel of the image. Therefore the algorithm has a $O(n^2)$ complexity in a $n \times n$ image[2].

A key aspect is that for the affine model fitting, we only use the image points where the normal-flow vectors correspond to the *circular* component of optical flow. The problem that remains to be solved is that of estimating the individual values of ω_1, ω_2, and the FOE, as only FOE direction, ω_3 and a constraint on ω_1 and ω_2 are known. This problem will be addressed in the second phase of our algorithm.

2.2 Solving the motion parameters : the Φ-circle search algorithm

In previous section, we computed the *circular* component of optical-flow to apply the Ψ-line search algorithm. In order to solve the remaining parameters, we can compute the *radial* component of optical flow as follows:

$$\forall(x,y) : (x,y) = r(\cos\psi, \sin\psi), \qquad V_\psi(r) = (u,v)\cdot(\cos\psi, \sin\psi), \tag{5}$$

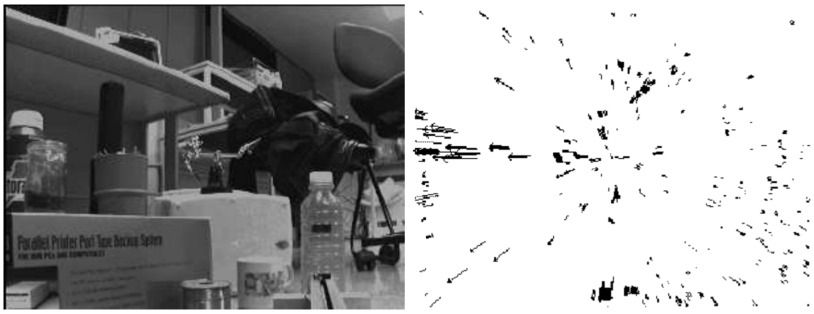

Fig. 3.: Sample image from the sequence and the radial normal flow.

where r denotes the distance from point (x, y) to the image center. V_ψ gives the radial component of the flow in the direction ψ. Figure 3 exemplifies the corresponding radial flow field.

Computing V_ψ for $\psi \in [0, 2\pi[$, we have :

$$\begin{aligned} \forall(x,y)\ &:\ (x,y) = r(\cos\psi, \sin\psi) \\ V_\psi(r) &= \underbrace{\rho(x,y)(r - (\sigma,\eta)\cdot(\cos\psi,\sin\psi))}_{translational\ component} + \underbrace{\left(\frac{r^2}{f}+f\right)(\omega_1\sin\psi - \omega_2\cos\psi)}_{rotational\ component} \end{aligned} \tag{6}$$

We can observe that (see Figure 4) :

1. By definition, the *radial* component of optical flow lies on one line passing through the image center.
2. The component of the flow due to translation radiates from the FOE, hence lying on a line that contains the FOE.

If we consider a circle (the Φ-circle) with the FOE and the image center as diametrical opposite points, then the lines referred in 1. and 2. are perpendicular, as Figure 4 shows. Therefore the component of V_ψ due to translation vanishes for every point on the Φ-circle.

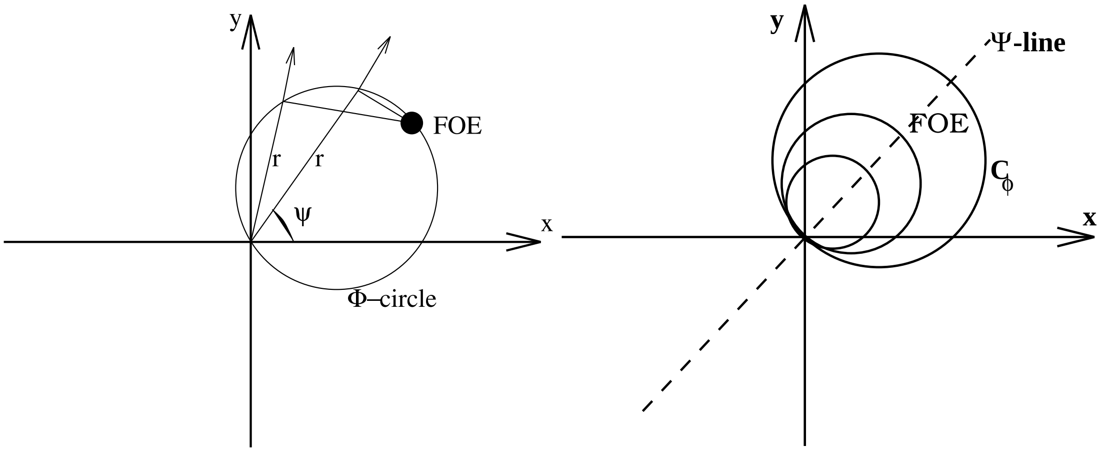

Fig. 4.: a) Radial component of optical flow on Φ-circle. b) Sequence of circles C_ϕ.

Given the known FOE-direction Ψ, we can define the set of circles C_ϕ containing the image center and having diameter ϕ, along the Ψ-line, as shown in Figure 4.

[2] In this paper, we choose the least mean squares as general estimation method, maintaining the complexity order.

Suppose that we compute the equation (6) for the circle C_ϕ and divide both terms by $\frac{r^2}{f}+f$. Then we have:

$$\forall(x,y):(x,y)\in C_\phi,\ T_\phi(\psi)=\frac{V_\psi(r)}{\frac{r^2}{f}+f}=\rho(x,y)\frac{r-(\sigma,\eta)\cdot(\cos\psi,\sin\psi)}{\frac{r^2}{f}+f}+\omega_1\sin\psi-\omega_2\cos\psi, \tag{7}$$

where r is dependent on ϕ and ψ to ensure that $T_\phi(\psi)$ is computed on circle C_ϕ.

We can conclude from (7) that if C_ϕ corresponds to the Φ-circle (i.e. $\phi=\Phi$), then $T_\Phi(\psi)$ will become sinusoidal in ψ. Based on this property, the algorithm consists of searching a circle C_Φ, such that $T_\Phi(\psi)$ is sinusoidal. To estimate the T_Φ parameters, we need three points, but more points are needed to increase the algorithmic robustness. When approximating $T_\Phi(\psi)$ to a sinusoidal function, the Φ-circle minimizes the mean square error. Known the Ψ-line, the *Φ-circle algorithm* "visits" only once each point of the image, such as the *Ψ-line algorithm*. Therefore the algorithm has also a $O(n^2)$ complexity.

In this algorithm we only use the radial normal flow, solving the remaining egomotion parameters, namely the FOE and the individual values of ω_1 and ω_2.

3 Results

The algorithm we proposed in this paper, was tested in a series of experiments of increasing difficulty, using both synthetic and real image data. The synthetic images size is 256×256 pixels and the focal length, f, is 302 pixels. The real experiment images size is 322×244 pixels and the focal length is 229 pixels. All the images are initially convolved with a low-pass Gaussian filter (5×5 mask, 1.0 standard deviation). Both the temporal and spatial derivatives are obtained using a 5-point derivation mask : $[1\ \ -8\ \ 0\ \ 8\ \ -1]/12$. We will now present results obtained for various kinds of scenes and different camera motions.

3.1 Linear motion along Z and rotation around Z

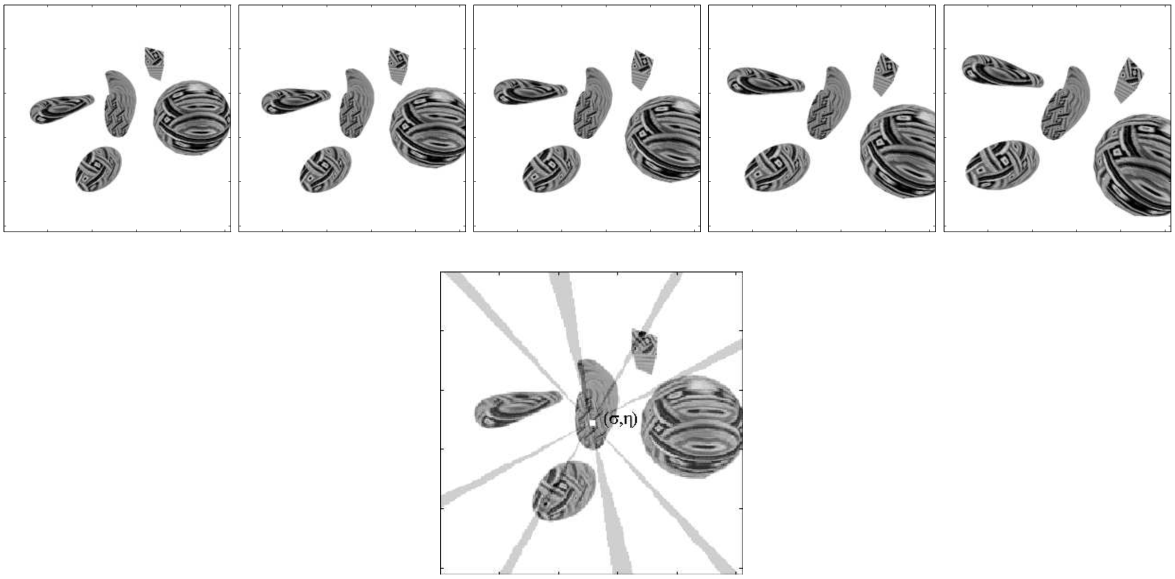

Fig. 5.: a) Image sequence, with rotation around Z axis. b) Estimation results.

In the first experiment, using synthetic images, the camera moves along the Z axis and rotates around the optical axis at 1^o/frame. The image sequence is shown in Figure 5.

In this example, the FOE is located at the image origin, represented in Figure 5 by a square mark. Therefore, all the lines passing through the image origin verify the constraints in the Ψ-line search algorithm. Hence, multiple Ψ-lines are detected, therefore the FOE is located at the origin. The Ψ-lines detected by the search algorithm are shown in Figure 5.

In this particular situation, the first part of the algorithm calculates all the motion parameters. If we have two Ψ-lines with directions Ψ_1 and Ψ_2, then we can estimate all the rotation parameters, from the following equations :

$$U_{\Psi_1}(r) = f(\omega_1 \cos\Psi_1 + \omega_2 \sin\Psi_1) - \omega_3 r$$
$$U_{\Psi_2}(r) = f(\omega_1 \cos\Psi_2 + \omega_2 \sin\psi_2) - \omega_3 r,$$

For each Ψ-line selected, $U_\Psi(r)$ was estimated as : $U_\Psi(r) = -0.0178r$. The constant term that corresponds to $\omega_1 \cos\Psi_1 + \omega_2 \sin\Psi_1$, is for all the lines selected, in the order of magnitude of 10^{-4} rad/frame. Thus, we conclude that ω_1 and ω_2 are negligible and $\omega_3 = 1.02^o/frame$. Therefore the rotation was recovered with about a 2% error.

3.2 Linear and angular motion

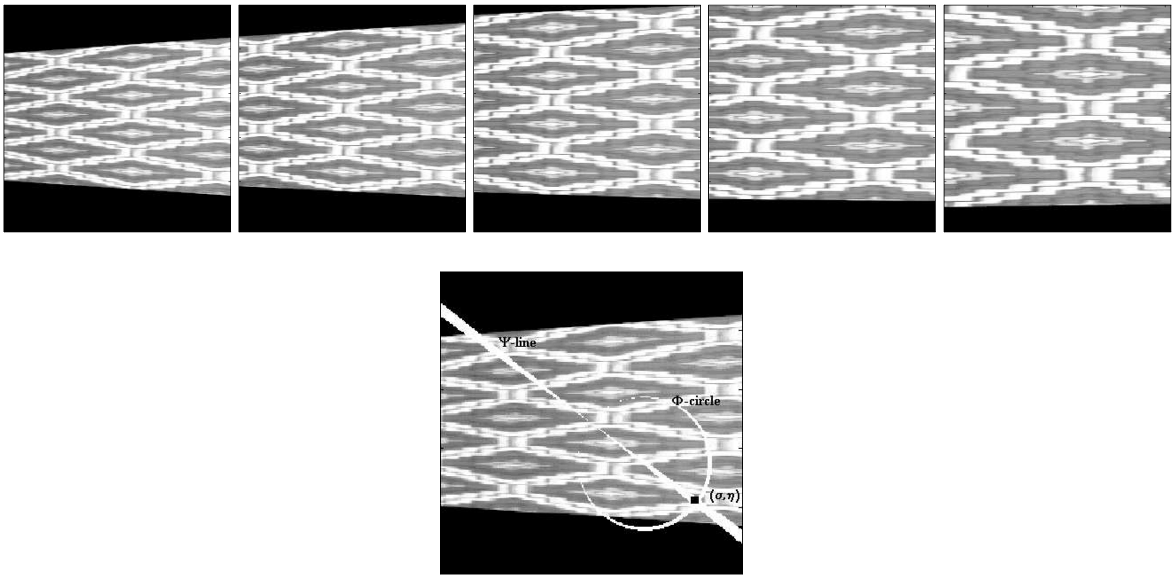

Fig. 6.: a) Image sequence, with translational and rotational motion. b) Estimation results.

In the second experiment, the motion is more complex, as the camera rotates around the Y axis at $0.45^o/frame$, while translating. The corresponding image sequence is shown in Figure 6.

As usual, the true FOE is shown in Figure 6, together with Ψ-line found by our algorithm. The affine model found for $U_\Psi(r)$ is given by : $U_\Psi(r) = -0.0044f - 0.0010r$. By comparison with equation (4), we can already conclude that $\omega_1 \cos\Psi + \omega_2 \sin\Psi = -0.0044$ and $\omega_3 = 0.0010$ rad/frame. By applying the second part of the algorithm, we found the Φ-circle shown in Figure 6, which intersects the Ψ-line close to the true location of the FOE. The angular motion estimates were $\boldsymbol{\omega} = [-0.034 \;\; 0.44 \;\; 0.06]^T$ deg/frame, which is close to the ground truth values.

3.3 Real world experiment

In the final experiment, with real images, the camera undergoes a translational motion in a real cluttered scene, yielding the image sequence shown in Figure 7.

As before, we represented the true FOE in Figure 7, and $\boldsymbol{\omega} \approx \mathbf{0}$. The Ψ-line found by our algorithm is shown in Figure 7.

The estimated $U_\Psi(r)$ associated to Ψ line is : $U(r) = 0.000191f - 0.00051r$. This implies that $\omega_3 = 0.00051$ rad/frame, which is approximately zero. Finally, by applying the Φ-circle search algorithm, the line found intersects the Ψ-line near to true FOE location, as shown in the same figure. The resulting rotation parameter estimates are $(\omega_1, \omega_2) = (0.00042, 0.00010)$ rad/frame, which are very small values, as expected.

Fig. 7.: a) Image sequence in a real set-up. b) Estimation results.

4 Conclusions

In this paper, we addressed the problem of egomotion estimation for a monocular observer, moving under arbitrary translation and rotation. As opposed to other methods that rely on the complete optical flow field, our approach is uniquely based on spatio-temporal image derivatives. There are no assumptions made on the structure of the scene and there are no constraints relating the observer position/pose and the environment, such as fixation.

The approach adopted, rather than considering the whole set of image flow data, relies on the image sites with special geometric properties, that convey relevant information about the observer motion. As a result, the algorithm has complexity $O(n^2)$, for $n \times n$ images. Our method works in two successive steps. First, the **Ψ-line search algorithm** is used to determine the direction of FOE. As an outcome, we also estimate the rotation component around the Z axis, and a linear constraint in the remaining two rotational components. Finally, in the second phase, we use the **Φ-circle search algorithm**, to determine the FOE, and the individual values of the angular velocity. The Ψ-line search algorithm uses the *circular normal flow* and the Φ-circle search algorithm uses the *radial normal flow*. Therefore we observe an interesting duality between the algorithms.

In the paper, we presented a series of experiments using both synthetic and real images, for various kinds of motion, that illustrate the robustness of the approach.

Future research will address more sophisticated estimation methods that may increase the system robustness and performance. Additionally, by considering an active observer, one can use further constraints in the egomotion estimation process, and the algorithms proposed can be used in a closed loop navigation system.

References

1. J. Aloimonos and A. Basu. Determining the translation of a rigidly moving surface, without correspondence. In *Proc. IEEE Conf. CVPR*, June 1986.
2. Y. Aloimonos and Z. Duric. Estimating the heading direction using normal flow. *Int. Journal of Computer Vision*, 13(1):33–56, September 1994.
3. Y. Aloimonos, I. Weiss, and A. Banddophaday. Active vision. *Int. Journal of Computer Vision*, 1(4):333–356, January 1988.
4. D. H. Ballard and O. A. Kimball. Rigid body motion from depth and optical flow. *Comput. Vision Graphics Image Processing*, 22(1), April 1983.
5. M. Bertero, T. Poggio, and V. Torre. Ill-posed problems in early vision. *Proceedings of the IEEE*, 76(8):869–889, 1988.
6. A. Bruss and B.K.P. Horn. Passive navigation. *Computer Vision, Graphics and Image Processing*, 21(1):3–20, January 1983.
7. C. Fermuller. Passive navigation as a pattern recognition problem. *Int. Journal of Computer Vision*, 14(2):147–158, March 1995.

8. C. Fermuller. Qualitative egomotion. *Int. Journal Comp. Vision*, 15(1/2):7–29, 1995.
9. C. Fermuller and Y. Aloimonos. The role of fixation in visual motion analysis. *Int. Journal of Computer Vision*, 11(2):165–186, October 1993.
10. B.K.P. Horn and B. Shunck. Determining optical flow. *Art. Intelligence*, 17:185–203, 1981.
11. S. Negahdaripour and S. Lee. Motion recovery from images sequences using only first order optical flow information. *Int. Journal of computer Vision*, 9(3):163–184, 1992.
12. J. Santos-Victor, G. Sandini, F. Curotto, and S. Garibaldi. Divergent stereo in autonomous navigation : From bees to robots. *Int. Journal of Comp. Vision*, 14(2):159–177, 1995.
13. C. Silva and J. Santos-Victor. Direct egomotion estimation. In *Proc. of the Int. Association for Pattern Recognition*, Vienna, Austria, August 1996.
14. D. Sinclair, A. Blake, and D. Murray. Robust estimation of egomotion from normal flow. *Int. Journal of Computer Vision*, 13(1):57–69, September 1994.
15. D. Terzopoulos. Regularization of inverse visual problems involving discontinuities. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8(4):413–424, July 1986.
16. A. Tikhonov and V. Arsenin. *Solution of ill-posed problems.* Wash. DC: Winston, 1977.