The Analysis of time varying image sequences

E. De Micheli[(*)], S. Uras & V. Torre
[(*)]Istituto di Cibernetica e Biofisica - C.N.R., Genova, Italy
Dipartimento di Fisica, Genova, Italy

Introduction

The analysis of time-varying image sequences is a classical problem of machine vision [1] [5], which is likely to be rather useful in several fields, such as robotics and passive navigation.

In this paper a general approach to the problem is presented, which is based on the computation of optical flow obtained by a similar procedure to the one proposed by Girosi et al. [2]. By exploiting mathematical properties of the 2D motion field [6] several 3D motion parameters, such as *time-to-collision* and angular velocity, can be recovered with a precision which is dependent on the image texture.

The Computation of Optical Flow

When an opaque object is moving in front of an artificial or a biological eye, it defines (in an appropriate system of reference, possibly solid to the image plane of the eye) a 3D velocity field $\vec{V}(\vec{R}) = (V_x, V_y, V_z)$ where $\vec{R} = (R_x, R_y, R_z)$. Because of the imaging device the 3D velocity field $\vec{V}$ is transformed into a 2D motion field $\vec{v} = (v_x, v_y)$ on the image plane [4] [6]. The available information, however, is not the 2D motion field $\vec{v}$ but the scalar field $E(x, y, t)$ of the image brightness at location (x, y) on the image plane at time t. By optical flow we mean any 2D vector field derived from $E(x, y, t)$ which is close to the 2D motion field $\vec{v} = (v_x, v_y)$. It is therefore evident that many different optical flows exist each of which has different properties and behaviour, according to the computing algorithm and the closeness criteria. It has recently been shown [2] that by assuming $\frac{d}{dt}$ grad $E = 0$ it is possible to obtain an optical flow $\vec{u} = (u_x, u_y)$ computed as:

$$\vec{u} = -H^{-1} \frac{\partial}{\partial t} \text{grad } E \tag{1}$$

where $H = (\frac{\partial^2 E}{\partial x_i \partial x_j})$ is the Hessian matrix of $E(x, y, t)$ and $\vec{u}$ is related to the 2D velocity field $\vec{v}$, by the equation:

$$\vec{u} = \vec{v} + H^{-1} \left(J_{\vec{v}}^T \cdot \text{grad } E - \text{grad } \frac{dE}{dt} \right) \tag{2}$$

where $J_{\vec{v}}^T$ is the transpose of the Jacobian matrix $(\frac{\partial v_i}{\partial x_j})$ of $\vec{v}$ and $\frac{dE}{dt}$ is the total derivative of the image brightness $E(x, y, t)$. Equation (2), which is just an identity, shows that the optical flow computed from equation (1) usually differs from the true 2D motion field, but also indicates that, when $\frac{dE}{dt}$ and $J_{\vec{v}}^T$ are bounded, the optical flow $\vec{u}$ will approach the true 2D motion field $\vec{v}$ whenever the entries of the matrix H^{-1} are small. Since $\frac{dE}{dt}$ and $J_{\vec{v}}^T$ are likely to be usually bounded, with the exception of those points near object boundaries or motion discontinuities, we can determine whether the computed optical flow $\vec{u}$ is close to $\vec{v}$ by simply evaluating H^{-1}. It is evident that the entries of H^{-1} will be small whenever the two real eigenvalues λ_1 and λ_2 of the symmetric matrix H are large.

An inspection of equation (1), from which the optical flow is derived, shows that numerical stability of the computation of optical flow is guaranteed whenever the inversion of the matrix H is numerically stable. This condition is fulfilled when det H is large and the conditioning number c_H of H is close to 1 [3]. Since the matrix H is symmetric, we have that $c_H = |\lambda_1/\lambda_2|$ where λ_1 and λ_2 are the two real eigenvalues of H with largest and smallest absolute value respectively. Consequently, it is evident that det H large and $c_H \sim 1$ imply both numerical stability in the computation of $\vec{u}$ and similarity between optical flow $\vec{u}$ and 2D motion field $\vec{v}$. As a result when det H is large and $c_H \sim 1$ an optical flow is obtained, which is numerically stable and almost correct (i.e. close to the true 2D motion field).

The Recovery of Motion Parameters

The obvious test of any procedure for motion analysis is the comparison between 3D motion parameters directly measured with those recovered from the analysis of the image sequence. Here we discuss the recovery of 3D motion parameters in two special, but practically relevant cases: pure translation (Fig. 1) and pure rotation (Fig. 2).

It has been shown [6] that in the case of pure translation the 2D motion field has at most one singular point, which is a focus and does not change its location on the image plane with time. Moreover if P_T is the singular point the *time-to-collision* between the image plane and the point projected into P_T is simply $1/\lambda$, where λ is the value of the two coincident eigenvalues of $J_{\vec{v}}$ computed at P_T. In the case of a pure axial rotation, that is when the rotation axis is orthogonal to the 3D surface, the angular velocity ω can be obtained from

$$\omega^2 = \det J_{\vec{v}}|_{P_R} \tag{3}$$

where P_R is the immobile point of the pure rotation [6] which is a center.

By using the sparse optical flow obtained from eq. 1 it is possible to locate the singular point $P = (\bar{x}, \bar{y})$ and to analyse the nature of the singular point by estimating the 6 parameters $\bar{x}, \bar{y}, a, b, c$ and d, such that

$$\begin{aligned} u_x &= a(x - \bar{x}) + b(y - \bar{y}) \\ u_y &= c(x - \bar{x}) + d(y - \bar{y}) \end{aligned}$$

represents the least square approximation of the flow in a suitable neighbourhood of P.

In the case of a pure translation we expect b and c to be negligible and the values of a and d to be very similar, whereas for axial motion a and d close to zero and b and c opposite in sign.

An extensive experimentation on sequences of images of different objects has shown that an accuracy of about 95 % can be obtained in the special case of a highly textured plane parallel to the image plane. This configuration, which is optimal from a theoretical point of view (the 2D motion field becomes linear), also proved experimentally to be the most favourable. For scenes with little texture and strongly departing from a planar structure, the agreement between computed and directly measured 3D motion parameters deteriorates and may become poorer.

We conclude that the proposed technique for the analysis of image sequence is suitable for the vision system of a mobile robot, and for many industrial applications.

References

1 - J.K. Aggarwal, & N. Nandhakumar. On the computation of motion from sequences of images - A review. *Proceedings of the IEEE*, 76, No. 8, 917-935, 1988.

2 - F. Girosi, A. Verri, & V. Torre. Constraints in the Computation of Optical Flow. *Proceedings of the IEEE Workshop on Visual Motion*, Irvine CA, 1989.

3 - C. Lanczos, : Linear Differential Operators. D. Van Nostrand Company, London 1961.

4 - H.C. Longuet-Higgins, & K. Prazdny. The interpretation of a moving retinal image. In *Proc. R. Soc. Lond. B* **208** 385-397, 1980.

5 - S. Ullmann. The Interpretation of Visual Motion. MIT Press, Cambridge, 1979

6 - A. Verri, F. Girosi, and V. Torre. Mathematical properties of the two-dimensional motion field: from Singular Points to Motion Parameters. JOSA A, Vol. 6, No. 5, pp.698-712, 1989.

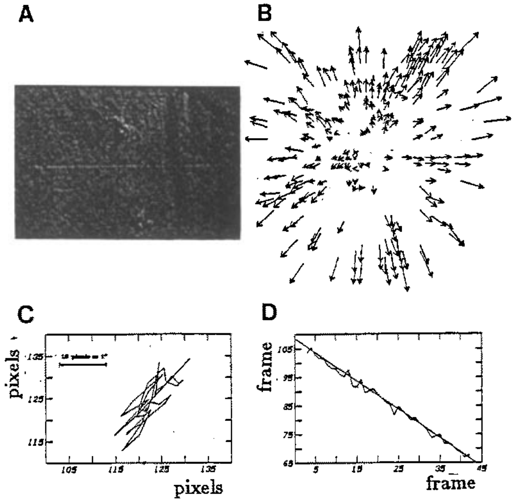

Fig. 1: Pure translation. A: An image of a photograph of two climbers. The image sequence was composed of 44 frames. The camera was slid on the rail of an optical bench by 1 cm between each image acquisition. B: The sparse optical flow relative to the image 15 of the sequence by solving eq. 1. C: The localization of the focus of expansion. The focal length of the camera was 8 mm and the width of a pixel was 0.014 mm so an angular displacement of 1 degree corresponds to a displacement on the image plane of about 10 pixels. D: Comparison between the true *time-to-collision* (straight line) and the computed *time-to-collision* (polygonal line).

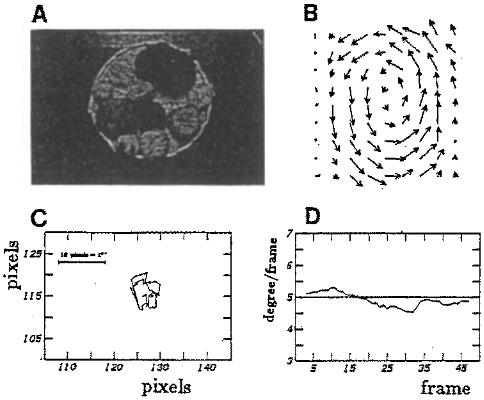

Fig. 2: Pure rotation. The image sequence was composed of 50 frames. The camera was viewing from above and pointing towards a rotating platform, on which different objects were mounted. The platform was rotated by 5 degrees between each frame. B: The smoothed optical flow. C: The localization of the immobile point (i.e. the singular point). D: Comparison between the true (straight line) and the computed (polygonal line) angular velocity.