A Fast Obstacle Detection Method based on Optical Flow *

Nicola Ancona

Tecnopolis CSATA Novus Ortus 70010 Valenzano - Bari - Italy.
e-mail: ancona@minsky.csata.it

Abstract. This paper presents a methodology, based on the estimation of the optical flow, to detect static obstacles during the motion of a mobile robot. The algorithm is based on a correlation scheme. At any time, we estimate the position of the focus of expansion and stabilize it by using the Kalman filter. We use the knowledge of the focus position of the flow field computed in the previous time to reduce the search space of corresponding patches and to predict the flow field in the successive one. Because of its intrinsic recursive aspect, the method can be seen as an on-off reflex which detects obstacles lying on the ground during the path of a mobile platform. No calibration procedure is required. The key aspect of the method is that we compute the optical flow only on one row of the image, that is relative to the ground plane.

1 Introduction

The exploitation of robust techniques for visual processing is certainly a key aspect in robotic vision application. In this work, we investigate an approach for the detection of static obstacles on the ground, by evaluation of optical flow fields. A simple way to define an obstacle is by a plane lying on the ground, orthogonal to it and high enough to be perceived. In this framework, we are interested in the changes happening on the ground plane, rather than in the environmental aspect of the scene. Several constraints help analysis and in tackling the problem. Among them:

1. the camera attention is on the ground plane, sensibly reducing the amount of required computational time and data;
2. the motion of a robot on a plane exhibits only three degrees of freedom; further, the height of the camera from this plane remains constant in time.

The last constraint is a powerful one on the system geometry, because, in pure translational motion, only differences of the vehicle's velocity and depth variations can cause changes in the optical flow field. Then the optical flow can be analysed looking for the *anomalies* with respect to a predicted velocity field [SA1].

A number of computational issues have to be taken into account: a) the *on-line* aspect, that is the possibility to compute the optical flow using at most two frames; b)the capability of detecting obstacles on the vehicle's path in a *reliable* and *fast* way; c)the possibility of *updating* the system status when a new frame is available. The above considerations led us to use a recursive token matching scheme as suitable for the problem at hand. The developed algorithm is based on a correlation scheme [LI1] for the estimation

* **Acknowledgements**: this paper describes research done at the Robotic and Automation Laboratory of the Tecnopolis CSATA. Partial support is provided by the Italian PRO-ART section of PROMETHEUS.

of the optical flow fields. It uses two frames at a time to compute the optical flow and so it is a suitable technique for on-line control strategies. We show how the estimation of the optical flow on only one row of reference on the image plane is robust enough to predict the presence of an obstacle. We estimate the flow field in the next frame using a predictive Kalman filter, in order to have an *adaptive* search space of corresponding patches, according to the environmental conditions. The possibility of changing the search space is one of the key aspects of the algorithm's performance. We have to enlarge the search space only when it is needed, that is only when an obstacle enters the camera's field of view. Moreover, it is important to point out that no calibration procedure is required. The developed methodology is different from [SA1] and [EN1] because we are interested in the temporal evolution of the predicted velocity field.

2 Obstacle model

Let us suppose that a camera moves with pure translational motion on a floor plane, β, (fig. 1a), and that the distance h between the optical center $\mathbf{C}$ and the reference plane stays constant in time. Let $\mathbf{V}(V_x, V_y, V_z)$ be the velocity vector of $\mathbf{C}$ and let us suppose that it is *constant* and parallel to the plane β. Let us consider a plane γ (obstacle) orthogonal to β, having its normal parallel to the motion direction. Moreover, let us consider a point $\mathbf{P}(P_x, P_y, P_z)$ lying on β and let $\mathbf{p}(p_u, p_v)$ be its perspective projection on the image plane: $\mathcal{T}(\mathbf{P}) = \mathbf{p}$. When the camera moves, γ intersects the ray projected by $\mathbf{p}$ in a point $\mathbf{Q}(Q_x, Q_y, Q_z)$, with $Q_z < P_z$. In other words, $\mathbf{Q}$ lies on the straight line through $\mathbf{C}$ and $\mathbf{P}$. It is a worth to point out that the points $\mathbf{P}$ and $\mathbf{Q}$ are acquired from the element $\mathbf{p}$ at different temporal instants, because we have assumed the hypothesis of opacity of the objects' surfaces in the scene.

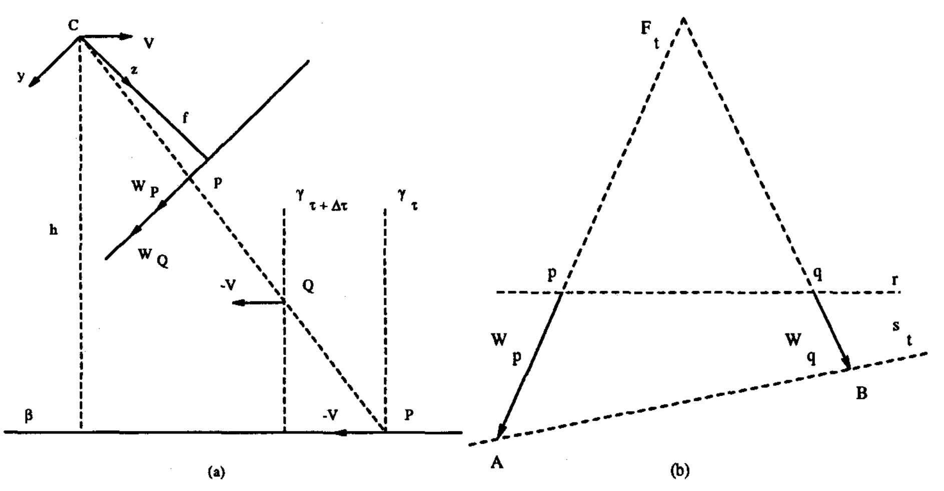

Fig. 1. (a) The geometrical aspects of the adopted model. The optical axis of the camera is directed toward the floor plane β, forming a fixed angle with it. (b) The field vectors relatives to the points $\mathbf{p}$ and $\mathbf{q}$ on r in a general situation.

Let us consider the field vectors, $\mathbf{W}_P$ and $\mathbf{W}_Q$, projected on the image plane by the camera motion, relative to $\mathbf{P}$ and $\mathbf{Q}$. At this point, let us make some useful considerations:

1. $\mathbf{W}_P$ *and* $\mathbf{W}_Q$ *have the same direction.* This statement holds because, in pure translational motion, the vectors of the $2D$ motion field converge to the focus of expansion $\mathbf{F}$, independently from the objects' surface into the scene. Then, $\mathbf{W}_P$ and $\mathbf{W}_Q$ are parallel and so the following proposition holds: $\exists\lambda > 0 \ni' \mathbf{W}_Q = \lambda\mathbf{W}_P$.
2. As $Q_z < P_z$, the following relation holds: $\|\mathbf{W}_P\| < \|\mathbf{W}_Q\|$

So we can claim that, under the constraint of a constant velocity vector, a point $\mathbf{Q}$ (obstacle), rising from the floor plane, does not change the direction of the vector flow, with respect to the corresponding point on the floor plane $\mathbf{P}$, but it only increases its length. So, the *variation of the modulus* of the optical flow at one point, generated by the presence of a point rising from the floor plane, is a *useful indicator* to detect obstacles along the robot path.

We know that estimation of the optical flow is very sensitive to noise. To this aim let us take a row on the image plane rather than single points, where to extract the flow field. The considerations as above still hold. Let us suppose the X axis of the camera coordinate system to be parallel to the plane β. We consider a straight line $r : v = k$ on the image plane and let w be the corresponding line on β obtained by back projecting r: $w = \mathcal{T}^{-1}(r) \cap \beta$. Under these constraints, all points on w have the same depth value. Let us consider two elements of r: $\mathbf{p}(p_u, k)$ and $\mathbf{q}(q_u, k)$, and let $\mathbf{P}(P_x, P_y, z)$ and $\mathbf{Q}(Q_x, Q_y, z)$ be two points on w such that: $\mathcal{T}(\mathbf{P}) = \mathbf{p}, \mathcal{T}(\mathbf{Q}) = \mathbf{q}$ We can state that the end points of the $2D$ motion field, $\mathbf{W}_p$ and $\mathbf{W}_q$, lie on the straight line s (fig. 1b). In particular, when the camera views the floor plane without obstacles, the straight lines r and s stay parallel and maintain the same displacement during the time. When an obstacle enters into the camera's field of view, the line parameters change and they can be used to detect the presence of obstacles.

3 Search space reduction

A first analysis of the optical flow estimation process shows that the performances of the algorithm are related to the magnitude of δ, the expected velocity of a point on the image plane. This quantity is proportional to the *search space* (SS) of corresponding brightness patches, in two successive frames. In other words, SS is the set of possible displacements of a point during the time unit. We focus our attention on the search space reduction, one of the key aspects of many correspondence based algorithms, to make the performance of our approach close to real time and to obtain more reliable results. The idea is to *adapt* the size of the search space according to the presence or absence of an obstacle in the scene.

Let us consider (fig. 1b) a row r on the image plane and let $\mathbf{p}$ and $\mathbf{q}$ be two points on it. Let $\mathbf{W}_p$ and $\mathbf{W}_q$ be the relative field vectors. Moreover, let s_t be the straight line where all of end points of the field vectors lie, at the time t. The search space SS_t, at the time t, is defined by the following rectangular region:

$$SS_t = \left[\min_u\{\mathbf{W}_p, \mathbf{W}_q\}, \max_u\{\mathbf{W}_p, \mathbf{W}_q\}\right] \times \left[\min_v\{\mathbf{W}_p, \mathbf{W}_q\}, \max_v\{\mathbf{W}_p, \mathbf{W}_q\}\right] \quad (1)$$

We want to stress out that SS_t is constrained by $\mathbf{F}_t$ and by the straight line s_t. For the sake of these considerations, wishing to predict $SS_{t+\Delta t}$, at the time $t + \Delta t$, it is enough to predict $s_{t+\Delta t}$, knowing $\mathbf{F}_t$ and s_t at the previous time. To realize this step, that we call *optical flow prediction*, we assume the *temporal continuity constraint* to be true, in other words the possibility of using an high sample rate of input data hold.

Suppose we know an estimate of the FOE $\hat{F}_t$ at time t [AN1]. The end points of $\mathbf{W}_p$ and $\mathbf{W}_q$ determine a straight line s_t, whose equation is $y = m_t x + n_t$. As we are only considering pure translational motion, the straight lines determined by the vectors $\mathbf{W}_p$ and $\mathbf{W}_q$ converge to $\hat{F}_t$. So, let us consider: $l_p = \left[\hat{F}_t, \mathbf{p}\right], l_q = \left[\hat{F}_t, \mathbf{q}\right]$ determined by $\hat{F}_t$ and $\mathbf{p}$ and $\mathbf{q}$ respectively. We denote by A and B the intersections of l_p and l_q with s_t. As these points lie on s_t, to predict $s_{t+\Delta t}$, the position of s_t at the instant $t + \Delta t$, it is sufficient to predict the position of A and B at the instant $t + \Delta t$. In the following, we consider only the point A, because the same considerations hold for the point B. As the position of $\hat{F}$ and $\mathbf{p}$ are constant, we can affirm that A moves on the line l_p. To describe the kinematic equation relative to this point, let us represent l_p in a parametric way. So the motion of A on l_p can be described using the temporal variations of the real parameter λ:

$$\lambda(t) = \lambda(\tau) + \dot{\lambda}(t-\tau) + \frac{1}{2}\ddot{\lambda}(t-\tau)^2 \quad where\ \tau < t \tag{2}$$

This equation, describing the temporal evolution of A, can be written in a recursive way. Setting $\tau = (k-1)T$ and $t = kT$, where T is the unit time, we get:

$$\begin{cases} x_1(k) = x_1(k-1) + x_2(k-1)T + \frac{1}{2}a(k-1)T^2 \\ x_2(k) = x_2(k-1) + a(k-1)T \end{cases} \tag{3}$$

where $x_1(k)$ denotes the value of the parameter λ, $x_2(k)$ its velocity and $a(k)$ its acceleration. In this model, $a(k)$ was regarded as a white noise. Using a vectorial representation, the following equation holds: $\mathbf{x}(k) = \mathbf{A}\mathbf{x}(k-1) + \mathbf{w}(k-1)$ describing the dynamical model of the signal. At each instant of time it is possible to know only the value of λ, so the observation model of the signal is given by the following equation: $y(k) = \mathbf{C}\mathbf{x}(k) + v(k)$ and $\mathbf{C} = (1, 0)$, where $E[v(k)] = 0$ and $E[v^2(k)] = \sigma_v^2$. The last two equations, describing the system and observation model, can be solved by using the predictive Kalman filtering equations. At each step, we get the best prediction of the parameter λ for A and B and so we are able to predict the estimate of the optical flow field, for all of points of r.

4 Experimental results

The sequence, fig. 2, was acquired from a camera mounted on a mobile platform moving at a speed of 100 mm/sec. The camera optical axis was pointing towards the ground. In this experiment, we used human legs as obstacle. The size of each image is made of 64×256 pixels. The estimation of the optical flow was performed only on the central row (32^{nd}) of each image. The fig. (3) shows the parameters m and n of s during the sequence. It is possible to note that at the beginning of the sequence, the variations of the parameters m and n are not very strong. Only when the obstacle is close to the camera, the perception module can detect the presence of the obstacle. This phenomena is due to the experimental set-up: camera's focal length and angle between optical axis and ground plane. The algorithm perceives the presence of an obstacle when one of the above parameters increase or decrease in a monotonous way. Our implementation run on a Risk 6000 IBM at the rate of 0.25 sec.

Acknowledgements: we would like to thank Piero Cosoli for helpful comments on the paper. Antonella Semerano checked the English.

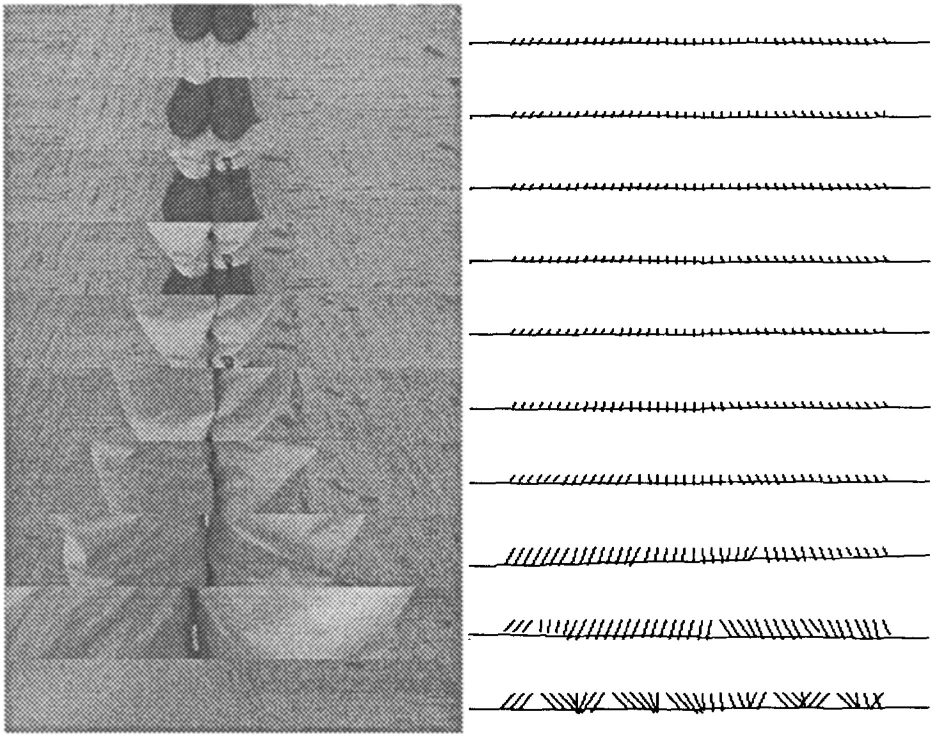

Fig. 2. Ten images of the sequence and the relative flow fields computed on the reference row.

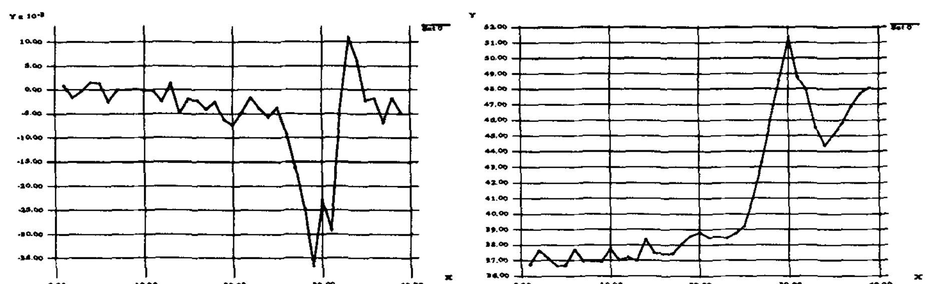

Fig. 3. The values of the parameters m and n of s computed during the sequence.

References

[LI1] Little J., Bulthoff H. and Poggio T.: Parallel Optical Flow Using Local Voting. IEEE 2nd International Conference in Computer Vision, 1988

[SA1] Sandini G. and Tistarelli M.: Robust Obstacle Detection Using Optical Flow. IEEE Workshop on Robust Computer Vision, 1-3 October 1990, Seattle - USA

[AN1] Ancona N.: A First Step Toward a Temporal Integration of Motion Parameters. IECON'91, October 28 1991, Kobe - Japan

[EN1] Enkelmann W.: Obstacle Detection by Evaluation of Optical Flow Fields from Image Sequences. First European Conference on Computer Vision, April 1990, Antibes - France