MASSACHUSETTS INSTITUTE OF TECHNOLOGY
ARTIFICIAL INTELLIGENCE LABORATORY
and
CENTER FOR BIOLOGICAL INFORMATION PROCESSING
WHITAKER COLLEGE

A.I. Memo 917 December 1986

C.B.I.P. Memo No. 22

MOTION FIELD AND OPTICAL FLOW: QUALITATIVE PROPERTIES

Alessandro Verri and Tomaso Poggio

ABSTRACT: In this paper we show that the *optical flow*, a 2-D field that can be associated with the variation of the image brightness pattern, and the 2-D *motion field*, the projection on the image plane of the 3-D velocity field of a moving scene, are in general different, unless very special conditions are satisfied. The optical flow, therefore, is ill-suited for computing structure from motion and for reconstructing the 3-D velocity field, problems that require an accurate estimate of the 2-D motion field. We then suggest a different use of the optical flow. We argue that stable qualitative properties of the 2-D motion field give useful information about the 3-D velocity field and the 3-D structure of the scene, and that they can be usually obtained from the optical flow. To support this approach we show how the (smoothed) optical flow and 2-D motion field, interpreted as vector fields tangent to flows of planar dynamical systems, may have the same qualitative properties from the point of view of the theory of structural stability of dynamical systems.

This report describes research done within the Artificial Intelligence Laboratory. Support for the A.I. Laboratory's artificial intelligence research is provided in part by the Advanced Research Projects Agency of the Department of Defense under Office of Naval Research contract N00014–85–K–0124. Support for this research is also provided by a grant from the Office of Naval Research, Engineering Psychology Division and by gift of the Artificial Intelligence Center of Hughes Aircraft Corporation to T. Poggio.

1. Introduction

A key task for many vision systems is to extract information from a sequence of images. This information can be useful to solve important problems such as recovering the 3-D velocity field, or segmenting the image into parts corresponding to different moving objects, or reconstructing the 3-D structure of surfaces. The recovery of the 2-D *motion field*, (that we define as the projection on the image plane of the 3-D velocity field) is thought to be an essential step in the solution of these problems. The data available, however, are temporal variations in the brightness pattern. These variations are usually associated with a perceived motion field, called *optical flow* (Gibson, 1950; Fennema and Thompson, 1979; Horn and Schunck, 1981). In order to recover the 2-D motion field, the assumption that the 2-D motion field and the optical flow coincide has often been made. It must be noted though, that this assumption is clearly satisfied only in the case in which variations in the brightness pattern correspond to features of the visible, 3-D surfaces. In fact several authors have developed algorithms to reconstruct the 2-D motion field from optical flow data defined only at locations of features in the image (Hildreth, 1984a,b; Waxman, 1986). Examples in which this assumption does not hold are known (Horn, 1986), but they have been regarded as pathological cases. As a matter of fact, algorithms that deal with the recovery of the 2-D motion field from dense optical flow data have been proposed, with the more or less explicit assumption that the two fields are the same (Horn and Schunck, 1981; Nagel, 1984; Kanatani, 1985).

In this paper we show that the optical flow and the motion field are in general different, unless very special conditions are satisfied. We explicitly compute the difference between their normal components (the component along the direction of the gradient) under broad assumptions. We show

that they are arbitrarily close where the image gradient is sufficiently strong. Hence, feature-based matching algorithms that rely on edges of various types (including texture edges) are more appropriate than point-to-point ones to solve problems that rely on accurate recovery of the 2-D motion field, such as structure from motion. One may then ask, what is the optical flow for? In the second part of the paper we suggest that meaningful information about the 3-D velocity field and the 3-D structure can be obtained from qualitative properties of the 2-D motion field. We then argue that this information can be retrieved directly from the optical flow or its normal components. We describe a specific approach that exploits results from the theory of stability of dynamical systems. A more detailed analysis of this approach will be presented in a forthcoming paper by V. Torre and coworkers.

The paper is divided in two parts. In the first, we define the problem and we state explicitly the assumptions that we have used. In particular, we consider in detail how image irradiance can be related to scene radiance in the case of a scene consisting of non-lambertian surfaces. We describe, then, a method that allows us to show that the optical flow and the motion field are almost always different. We analytically compute the difference between the normal components of the two fields assuming, first, the lambertian model of reflectance and then a more realistic one for arbitrary rigid motion of a generic surface. We also calculate how this difference depends on the image gradient and the 3-D velocity of moving objects. In the second part we show how both the optical flow and the motion field can be processed to become vector fields tangent to flows of dynamical systems. The optical flow then, can be considered as a perturbed motion field under the conditions determined in the first part. Results from the theory of stability of dynamical systems suggest that qualitative, stable properties of the motion field hold for the optical flow. We sketch some example of these properties and how

they can be used in a description of the 3-D velocity field. We finally discuss briefly some connections with biological systems.

2. Preliminaries

In this chapter we review the definitions of motion field and optical flow, and we state the assumptions that we used throughout the paper. In particular, we consider in detail how image irradiance can be related to scene radiance in the case of a scene consisting of non-lambertian surfaces.

2.1. Definitions

Let us define notations and summarize definitions that will be useful in what follows. For more details on the geometry of perspective projection see Appendix *A*1. Throughout the following we will assume, if it is not otherwise stated, that any expression can be differentiated as many times as nedeed. Let

$$\mathbf{x}_p = \frac{f}{f + \mathbf{x} \cdot \mathbf{n}} \big(\mathbf{x} - (\mathbf{x} \cdot \mathbf{n})\mathbf{n}\big) \tag{2.1.1}$$

be the equation defining the projection of a generic point on the image plane, where $\mathbf{x}_p = (x_p, y_p, 0)$ is the position vector of the projected point, $\mathbf{x} = (x, y, z)$ is the position vector of the point, $\mathbf{n}$ is the unit vector normal to the image plane (projection plane) and f is the focal length (see Figure 1). Notice that the origin O is on the image plane, the focus of projection F is located at $(0, 0, -f)$, and $f\mathbf{n} + \mathbf{x}$ is the vector pointing from F to the point.

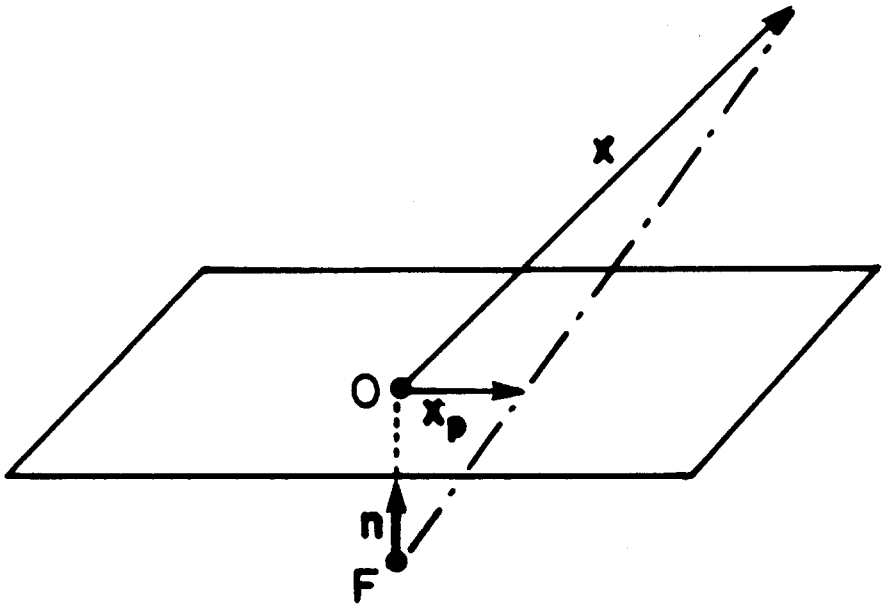

Figure 1. The geometry of perspective projection.

The motion field $\mathbf{v}_p$ can be obtained differentiating (2.1.1) with respect to the time. If $\mathbf{v} = d\mathbf{x}/dt$ we have [1]

$$\mathbf{v}_p = \frac{f}{f + \mathbf{x} \cdot \mathbf{n}}(\mathbf{v} - (\mathbf{v} \cdot \mathbf{n})\mathbf{n} - \frac{\mathbf{v} \cdot \mathbf{n}}{f + \mathbf{x} \cdot \mathbf{n}}(\mathbf{x} - (\mathbf{x} \cdot \mathbf{n})\mathbf{n})) \qquad (2.1.2)$$

Notice that in (2.1.2) $\mathbf{v}_p$ is given in terms of $\mathbf{x}$ and $\mathbf{v}$, position and velocity of the moving points in the scene, which are not known.

Let $E = E(x_p, y_p, t)$ be the image irradiance, that is the intensity of light at the point (x_p, y_p) of the image plane at the time t. If ∇_p is the gradient with respect to the image coordinates, then

$$\frac{dE}{dt} = \frac{\partial E}{\partial t} + \nabla_p E \cdot \mathbf{v}_p \qquad (2.1.3)$$

Now if

$$\frac{dE}{dt} = 0 \qquad (2.1.4)$$

[1] It can be easily shown that the perspective projection of the 3-D velocity vector is equal to the velocity of the projected point on the image plane, since both the vector are defined in terms of infinitesimal. This is not true for a generic, *finite* vector

then, if $\|\nabla_p E\| = 0$

$$-\frac{\partial E/\partial t}{\|\nabla_p E\|} = \frac{\nabla_p E \cdot \mathbf{v}_p}{\|\nabla_p E\|} \tag{2.1.5}$$

Therefore, if (2.1.4) holds, the projection of the motion field along the direction of the gradient can be given in terms of derivatives of the image irradiance (which can be computed). In what follows, this component will be called $\mathbf{v}_\perp$, or the *normal component*; thus

$$\mathbf{v}_\perp = \frac{\mathbf{v}_p \cdot \nabla_p E}{\|\nabla_p E\|} \frac{\nabla_p E}{\|\nabla_p E\|} \tag{2.1.6}$$

Equation (2.1.5) can be interpretated as an instance of the well-known *aperture* problem (Marr and Ullman, 1981; Horn and Schunck, 1981): that is the information available at each point of a sequence of frames is only the component of the motion field along the direction of the image gradient. To recover a full and unique motion field, some other constraint is needed: Horn and Schunck (1981), for example, showed that there is only one 2-D field whose normal component coincides with (2.1.6) and which is the smoothest of all possible ones. Examples for which (2.1.5) is not true are well known (Horn and Schunck, 1981). Consider, for instance, a rotating sphere with no texture on it (*i.e.* with uniform albedo) under arbitrary, fixed illumination. Since the image irradiance at each image location does not change with time, the left-hand side of (2.1.5) is identically equal to zero, while the right-hand side is different from zero almost everywhere. Notice that keeping the sphere fixed and moving the light source (2.1.5) is again wrong. In this case, however, the left-hand-side is different from zero while $v_\perp$ is zero everywhere. In both cases the perceived motion in the image is different from the motion field. It is worthwhile, then, to introduce a new field, called the *minimal optical flow*, related to the perceived motion in the image, and not necessarily equal to the normal component of the motion field. Notice that the perceived motion in the examples above agrees qualitatively

with the left-hand-side of (2.1.5). Indeed, the optical flow in the first case is identically equal to zero, while in the second is different from zero almost everywhere. Therefore let us *define* the normal component $\mathbf{O}_F$ of the optical flow as:

$$\mathbf{O}_F = -\frac{\partial E/\partial t}{\|\nabla_p E\|}\frac{\nabla_p E}{\|\nabla_p E\|} \tag{2.1.7}$$

Hence, with respect to this definition, the minimal optical flow and the normal component of the motion field are always directed along the gradient and they coincide if and only if (2.1.4) holds.

Remark: in the literature, it is usually assumed that (2.1.4) holds. As a consequence, the normal components of the motion field and of the optical flow are the same and the latter can be used as a constraint to recover the 2-D motion field.

2.2. Scene Radiance and Image Irradiance

Let us review briefly some definitions of photometry and make explicit the constraints under which the image irradiance is related to the scene radiance. The image irradiance E is the power per unit area of light at each point (x_p, y_p) of the image plane: thus $E = E(x_p, y_p)$. The scene radiance L is the power per unit area of light that can be thought emitted by each point of a surface S in the scene in a particular direction. This surface can be fictitious, or it may be the actual radiating surface of a light source, or the illuminated surface of a solid. The scene radiance can be thought as a function of the point of the surface and of the direction in space. If (a, b) are intrinsic coordinates of the surface and (α, β) polar coordinates determining a direction in space with respect to the normal to the surface, we can write $L = L(a, b, \alpha, \beta)$. Given the scene radiance it is possible, in principle, to compute the expected image irradiance. For example in the

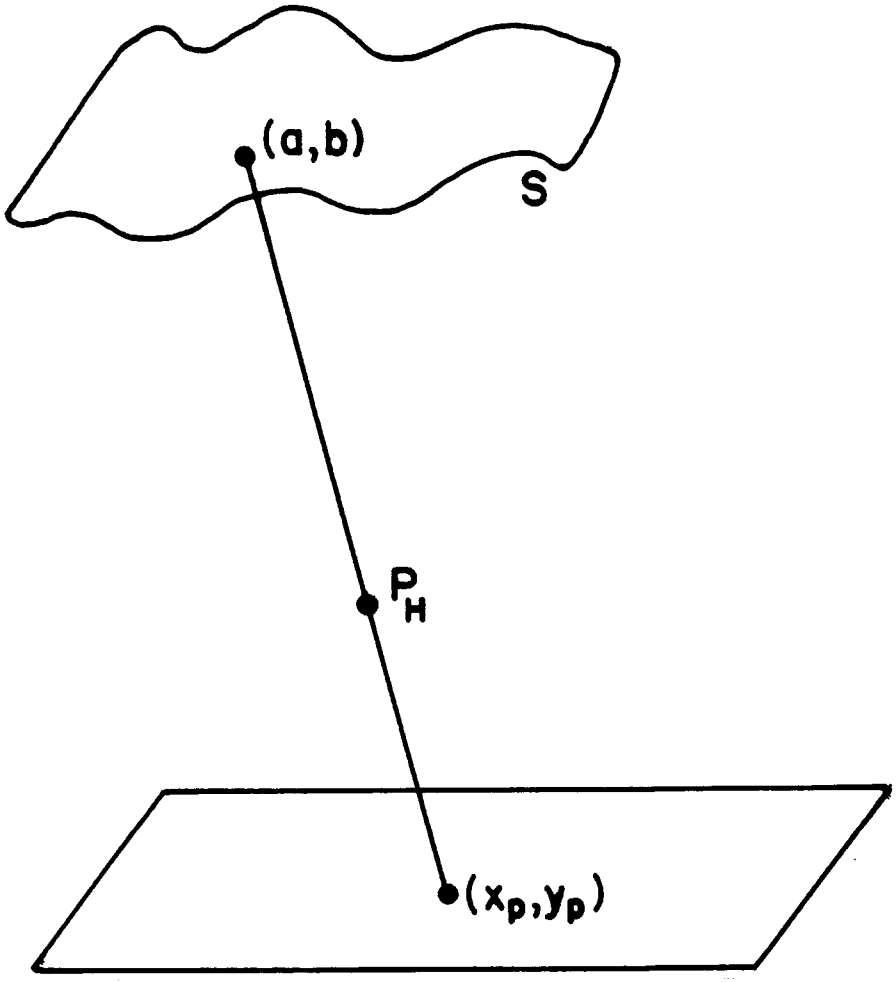

Figure 2. Scene radiance and image irradiance in the pinhole approximation: the image irradiance at the point (x_p, y_p) is given by the scene radiance at the point (a, b) on the surface in the direction of the line connecting the two points and passing through the pinhole P_H.

case of pinhole camera approximation, that is assuming that the camera has an infinitesimally small aperture, the image irradiance at a point (x_p, y_p) is proportional to the scene radiance at the point (a, b) on the surface in the direction of the pinhole, say (α^0, β^0), where (x_p, y_p), (a, b) and the pinhole lie on the same line (see Figure 2). Therefore we have

$$E(x_p(a,b), y_p(a,b)) = L(a, b, \alpha^0, \beta^0) \tag{2.2.1}$$

where $(x_p(a,b), y_p(a,b))$ is the image point that lies on the line connecting (a,b) to the pinhole. In practice, however, the aperture of any real optical device is finite and not very small (ultimately to avoid diffraction effects): thus (2.2.1) does not generally hold. Assuming that the surface is lambertian, *i.e.* $L(a,b,\alpha,\beta) = L(a,b)$, that there are not losses within the system and that the angular aperture (on the image side) is small it can be proved (Born and Wolf, 1959) that

$$E(x_p(a,b), y_p(a,b)) = L(a,b)\Omega\cos^4\varphi \tag{2.2.2}$$

where Ω is the solid angle corresponding to the angular aperture and φ is the angle between the principal ray (that is the ray passing through the center of the aperture) and the optical axis. With the further assumption that the aperture is much smaller than the distance of the viewed surface, the lambertian hypothesis can be relaxed to give (Horn and Sjoberg, 1979)

$$E(x_p(a,b), y_p(a,b)) = L(a,b,\alpha^0,\beta^0)\Omega\cos^4\varphi \tag{2.2.3}$$

where α^0 and β^0 are the polar coordinates of the direction of the principal ray. It must be pointed out that (2.2.3) holds if L is continuous with respect to α and β. In what follows we will assume that this is the case. Furthermore, we will assume that the optical system has been calibrated so that (2.2.3) can be rewritten as (2.2.1). Finally, notice that

$$\nabla_p E \cdot \mathbf{v}_p = \nabla_S L \cdot (\frac{da}{dt}, \frac{db}{dt}), \tag{2.2.4}$$

where ∇_S is the gradient with respect to the surface coordinates, since differentiating (2.2.1) we have

$$\nabla_p E \cdot (dx_p, dy_p) = \nabla_S L \cdot (da, db). \tag{2.2.5}$$

3. Minimal optical flow and Motion Field

We describe a general method that allows us to show that the minimal optical flow and the normal component of the motion field are almost always different, or equivalently that (2.1.4) does not hold. We compute the difference between the normal components of the two fields, assuming first the Lambertian model of reflectance and then a more realistic one for pure translation, pure rotation and general rigid motion of a generic surface. It turns out that the two fields are equal only under very special conditions, which can be explicitly stated. We also show that the difference is smaller where the image gradient is stronger, justifying the use of feature-based algorithms. Of course, this argument does not imply that feature-based algorithms *should* be used: it says, however, that locations of edges (meant here as sharp changes in intensity) contain most of the correct information.

3.1. Computing the Minimal Optical Flow

Consider a rigid surface S moving in space from (2.2.1). The image irradiance E at the time t at the point (x_p, y_p) is equal to the scene radiance L at the point (a, b) on S, *i.e.* $E(x_p, y_p, t) = L(a, b)$. The image irradiance at the time $t + \Delta t$ is given by the scene radiance of the surface at the time $t + \Delta t$. As shown in Figure 3, the point[1] on S that radiates toward (x_p, y_p) at the time $t + \Delta t$ is the point $(a - \Delta a, b - \Delta b)$.1 The normal $\mathbf{N}$ to S at the time $t + \Delta t$ at the point $(a - \Delta a, b - \Delta b)$, $\mathbf{N}_{t+\Delta t}(a - \Delta a, b - \Delta b)$, will be

$$\mathbf{N}_{t+\Delta t}(a - \Delta a, b - \Delta b) = \mathbf{N}_t(a - \Delta a, b - \Delta b) + \Delta \mathbf{N} \qquad (3.1.1)$$

1We assume that the surface corresponds to a moving convex body to avoid self-occlusions due to the motion. In fact, the computation that follows holds for any convex surface patch.

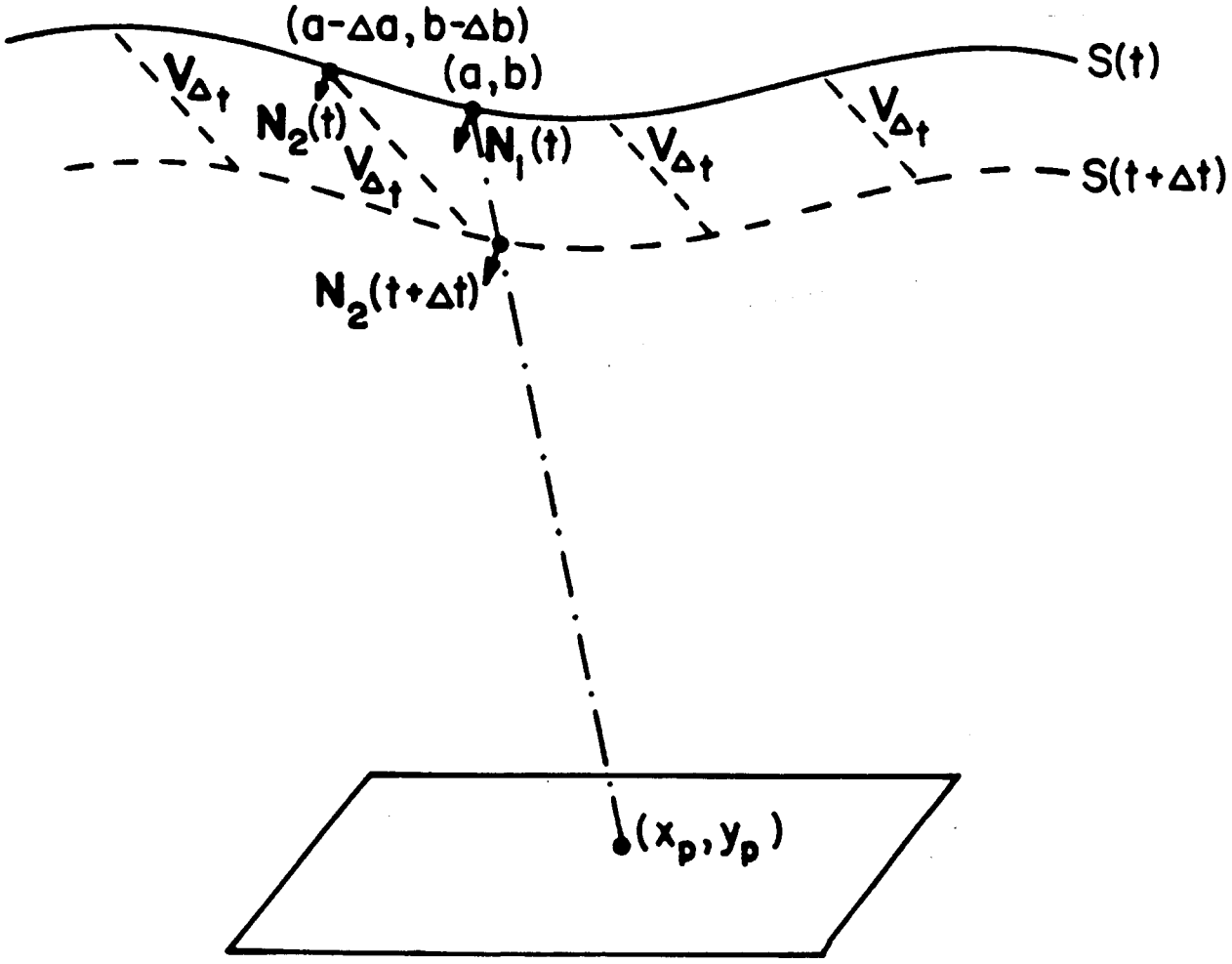

Figure 3 Computing the minimal optical flow: the point (a, b) on S radiates toward (x_p, y_p) at time t. The point $(a - \Delta a, b - \Delta b)$ radiates toward the same point at time $t + \Delta t$. The normal $\mathbf{N}_1$ is the normal to the S at the point (a, b) and $\mathbf{N}_2$ at $(a - \Delta a, b - \Delta b)$.

where $\Delta\mathbf{N}$ is the first order variation of $\mathbf{N}$ due to the motion of S during the time interval Δt. Now in the case of translation

$$\Delta\mathbf{N} = 0 \tag{3.1.2}$$

while in case of rotation with angular velocity ω

$$\Delta\mathbf{N} = \omega \times \mathbf{N}\Delta t \tag{3.1.3}$$

Notice that (3.1.3) can be considered as the expression of $\Delta\mathbf{N}$ for any

kind of motion. Similarly, for each argument A of the scene radiance, we can write

$$A_{t+\Delta t}(a,b) = A_t(a,b) + \Delta A. \tag{3.1.4}$$

To compute ΔA, let us distinguish between arguments of L that are intrinsic function of the surface coordinates (a,b), such as texture and albedo, and those that are in fact function of the space coordinates (x,y,z), (such as the illumination and the point of view) and that are expressed in terms of (a,b) only for convenience. If A is an intrinsic function of the surface coordinates, it follows immediately that

$$\Delta A = 0, \tag{3.1.5}$$

while if A is a function of the space coordinates, from the Taylor expansion we have

$$\Delta A = \nabla A \cdot \mathbf{v}\Delta t, \tag{3.1.6}$$

where ∇ is the gradient operator with respect to the space coordinates. Let us assume that L can be written as a function of m arguments A^i, $i = 1,...,m$ and of $\mathbf{N}$. Then, taking into account (3.1.3) and (3.1.4), (2.2.1) becomes

$$E(x_p,y_p,t+\Delta t) = L\big(A_t^i(a-\Delta a, b-\Delta b) + \Delta A^i, \mathbf{N}_t(a-\Delta a, b-\Delta b) + \Delta\mathbf{N}\big) \tag{3.1.7}$$

at time $t + \Delta t$ and

$$E(x_p,y_p,t) = L\big(A_t^i(a,b), N_t(a,b)\big) \tag{3.1.8}$$

at time t. Therefore, using (3.1.6) and (3.1.7),

$$\frac{\partial E}{\partial t} =$$

$$\lim_{\Delta t\to 0}\frac{1}{\Delta t}\Big(L\big(A_t^i(a-\Delta a, b-\Delta b) + \Delta A^i, \mathbf{N}_t(a-\Delta a, b-\Delta b) + \Delta\mathbf{N}\big) - L\big(A_t^i(a,b), N_t(a,b)\big)\Big), \tag{3.1.9}$$

where the ΔA^i are computed using (3.1.5) or (3.1.6) according to the kind of argument. From (3.1.9), the minimal optical flow can be derived easily. To simplify notation, let us suppress the subscript t from Equation (3.1.9).

From (3.1.9) we easily get

$$\frac{\partial E}{\partial t} = -\nabla_s L \cdot \left(\frac{da}{dt}, \frac{db}{dt}\right) + \sum_{i=1}^{p} \frac{\partial L}{\partial A_i} \nabla A^i \cdot V + \frac{\partial L}{\partial N} \omega \times N$$

if p of the $A^i (i = 1, \ldots, m)$ require the use of (3.1.6) to compute ∇A and $\frac{\partial L}{\partial N} = \left(\frac{\partial L}{\partial N_x}, \frac{\partial L}{\partial N_y}, \frac{\partial L}{\partial N_z}\right)$ for $N = (N_x, N_y, N_z)$.

Therefore, using (2.1.6), (2.1.7), and (2.2.4), we can write

$$v_\perp - 0_F = \sum_{i=1}^{p} \frac{\partial L}{\partial A_i} \nabla A^i \cdot v + \frac{\partial L}{\partial N} \cdot \omega \times N \tag{3.1.10}$$

Thus, the normal components of the two fields are different if the surface undergoes a motion with a rotational component, or the reflectance function contains arguments depending on space coordinates.

Let us consider now some interesting examples in detail.

3.2. Translation of a Lambertian Surface

Consider a lambertian surface S. The scene radiance due to S will be

$$L = \rho \mathbf{I} \cdot \mathbf{N} \tag{3.2.1}$$

where ρ is the albedo of S, $\mathbf{I}$ the unit vector in the direction of the illumination and $\mathbf{N}$ is the unit normal to the surface. Let us compute the difference (3.1.10) between the normal components of the optical flow and of the motion field corresponding to a translation of S in space with velocity $\mathbf{v}$ under uniform fixed illumination. Substituting (3.2.1) in (3.1.10) and changing the sign, we have

$$v_\perp = O_F, \tag{3.2.2}$$

since $\omega = 0$ and none of the arguments of L in (3.2.1) depends on space constraints (I is constant). Therefore, the minimal optical flow of a translating lambertian surface uniformly illuminated is exactly equal to the motion field.

Remark: in the case of non-uniform illumination the right hand side of (3.2.2) contains an extra term due to $\Delta\mathbf{I}$. Using (3.1.6) to compute the components of $\Delta\mathbf{I}$, (3.1.10) yields

$$v_\perp - O_F = \frac{1}{\|\nabla_p E\|}\rho\Big(\frac{\partial \mathbf{I}}{\partial x}\frac{dx}{dt} + \frac{\partial \mathbf{I}}{\partial y}\frac{dy}{dt} + \frac{\partial \mathbf{I}}{\partial z}\frac{dz}{dt}\Big)\cdot\mathbf{N},$$

which can be rewritten

$$v_\perp - O_F = \frac{1}{\|\nabla_p E\|}\rho\frac{d\mathbf{I}}{dt}\cdot\mathbf{N}, \tag{3.2.3}$$

since $\partial\mathbf{I}/\partial t = 0$ (the illumination is supposed to be fixed). Let us consider now the case of a rotating lambertian surface.

3.3. Rotation of a Lambertian Surface

Let S be a lambertian surface rotating in space with angular velocity $\boldsymbol{\omega}$. Let $\mathbf{I}$ be again uniform. Applying the same argument of the previous section but taking into account the constraint (3.1.3) for ∇N, we get

$$v_\perp - O_F = \frac{\rho\mathbf{N}\cdot\mathbf{I}\times\boldsymbol{\omega}}{\|\nabla_p E\|}. \tag{3.3.1}$$

In the case of rotation, therefore, even under uniform illumination, the minimal optical flow and the normal component of the motion field are different. They are equal for any surface only if $\boldsymbol{\omega}$ and $\mathbf{I}$ are parallel. This corresponds to the case of a surface rotating around an axis parallel to the direction of uniform illumination. In the case of non-uniform illumination, an extra term like the one in (3.2.3) must be added to (3.3.1). *Remark:* it is worth considering analytically the example of the rotating sphere of the previous section. Due to rotational symmetry we have

$$\mathbf{N}(a-\Delta a, b-\Delta b) + \boldsymbol{\omega}\times\mathbf{N}(a-\Delta a, b-\Delta b)\Delta t = \mathbf{N}(a,b) \tag{3.3.2}$$

$\forall a, \forall b$ on the sphere. Furthermore,

$$\mathbf{I}_{t+\Delta t}(a-\Delta a, b-\Delta b) = \mathbf{I}_t(a-\Delta a, b-\Delta b) + \Delta\mathbf{I} = \mathbf{I}_t(a,b), \tag{3.3.3}$$

since in this case the displacement in space, $\mathbf{v}\Delta t$, is equal to the displacement on the surface, $(\Delta a, \Delta b)$. Therefore, if ρ is uniform,

$$-\frac{\partial E}{\partial t} = 0. \tag{3.3.4}$$

The minimal optical flow is then, as expected, equal to zero under any illumination.

3.4. Translation of a Specular Surface

Let us consider now a model of reflectance more realistic than the lambertian one. Following Phong (1975; see also Horn and Sjoberg, 1979) we define the scene radiance as a linear combination of a lambertian and a specular term, *i.e.*

$$L = L_{\text{lamb}} + L_{\text{spec}}. \tag{3.4.1}$$

The lambertian term is equal to the one used before, while the specular term is

$$E = \frac{s\mathbf{D} \cdot \mathbf{R}}{D}, \tag{3.4.2}$$

where s is the fraction of light reflected by the surface, $\mathbf{D} = f\mathbf{n} + \mathbf{x}$ is the vector pointing from the focus to the radiating point and

$$\mathbf{R} = \mathbf{I} - 2(\mathbf{I} \cdot \mathbf{N})\mathbf{N} \tag{3.4.3}$$

is the unit vector in the direction of the perfect specular reflection. Let us assume that s is not a function of the direction of the incident light and that it is constant on the surface. The specular term is thus proportional to the cosine of the angle between the direction of specular reflection and the line of sight.

Since we are computing derivatives and L is a linear combination of L_{lamb} and L_{spec} we can compute separately the contributions to the minimal optical flow due to the lambertian and the specular term, adding the

results afterward. Therefore, we only need to compute now the specular one. Let us consider, first, the case of pure translation of a surface S radiating accordingly to (3.4.2) and let us call S a *specular* surface. If S is translating with velocity $\mathbf{v}$ and $\mathbf{I}$ is uniform, substituting (3.4.2) into (3.1.10) and taking into account the constraint (3.1.2), we have

$$v_1 - 0_F = \frac{s}{D^3} \frac{(D^2\mathbf{v}\cdot\mathbf{R} - (\mathbf{D}\cdot\mathbf{v})(\mathbf{D}\cdot\mathbf{R}))}{\|\nabla_p E\|}, \tag{3.4.4}$$

since from (3.1.6)

$$\lim_{\Delta t\to 0} \frac{\Delta\mathbf{D}}{\Delta t} = \frac{\partial\mathbf{D}}{\partial x}\frac{dx}{dt} + \frac{\partial\mathbf{D}}{\partial y}\frac{dy}{dt} + \frac{\partial\mathbf{D}}{\partial z}\frac{dz}{dt} = \frac{d\mathbf{D}}{dt} = \frac{d\mathbf{x}}{dt} = \mathbf{v}. \tag{3.4.5}$$

Using again the two fields we get a well known vector identity:

$$v_\perp - O_F = \frac{s}{D^3} \frac{(\mathbf{v}\times\mathbf{D})\cdot(\mathbf{R}\times\mathbf{D})}{\|\nabla_p E\|}. \tag{3.4.6}$$

Thus, in the case of translation of a specular surface, the minimal optical flow and the normal component of the motion field are always different.

Remark: let us consider the case of orthographic projection. When $f \to \infty$, (3.4.6) becomes

$$v_\perp - O_F = 0,$$

since when $f \to \infty$, $D \to \infty$. Therefore, in the othographic limit, the minimal optical flow of a translating specular surface is equal to the normal component of the 2-D motion field.

3.5. Rotation of a Specular Surface

Consider now the same *specular* surface S rotating in space with angular velocity $\boldsymbol{\omega}$. Then, substituting (3.4.2) into (3.1.10) and taking into account the constraint (3.1.3), we have

$$\mathbf{v}_\perp - 0_F = \frac{-s}{\|\nabla_p \mathbf{E}\|\mathbf{D}^3}\Big(2D^2\big((\mathbf{I}\cdot\mathbf{N})(\mathbf{D}\cdot\boldsymbol{\omega}\times\mathbf{N}) + (\mathbf{D}\cdot\mathbf{N})(\mathbf{I}\cdot\boldsymbol{\omega}\times\mathbf{N})\big) - \\ -\big((\boldsymbol{\omega}\times\mathbf{x})\times\mathbf{D}\big)\cdot(\mathbf{R}\times\mathbf{D}) + \\ \big((\boldsymbol{\omega}\times\mathbf{x}_0)\times\mathbf{D}\big)\cdot(\mathbf{R}\times\mathbf{D})\Big), \qquad (3.5.1)$$

since $\mathbf{v} = \boldsymbol{\omega}\times\mathbf{x}$ and $\mathbf{x}_0$ gives the location of the axis of rotation. Now (3.5.1) gives

$$\mathbf{v}_\perp - 0_F = \frac{-s}{\|\nabla_p \mathbf{E}\|\mathbf{D}^3}\Big(2D^2(\mathbf{I}\cdot\mathbf{N})(\mathbf{D}\cdot\boldsymbol{\omega}\times\mathbf{N}) + 2D^2(\mathbf{D}\cdot\mathbf{N})(\mathbf{I}\cdot\boldsymbol{\omega}\times\mathbf{N}) + \\ + D^2\big((\boldsymbol{\omega}\times\mathbf{x})\cdot\mathbf{R}\big) + (\mathbf{D}\cdot\mathbf{R})(\mathbf{D}\cdot\boldsymbol{\omega}\times\mathbf{x}) + 2D^2\big((\boldsymbol{\omega}\times\mathbf{x}_0)\times\mathbf{D}\big)\cdot(\mathbf{R}\times\mathbf{D})\Big)$$

This expression can be simplified in the following way: since $\mathbf{D} = f\mathbf{n}+\mathbf{x}$,

$$\mathbf{v}_\perp - 0_F = \frac{-s}{\|\nabla_p \mathbf{E}\|\mathbf{D}^3}\Big(2fD^2(\mathbf{I}\cdot\mathbf{N})(\mathbf{n}\cdot\boldsymbol{\omega}\times\mathbf{N}) + 2D^2(\mathbf{I}\cdot\mathbf{N})(\mathbf{x}\cdot\boldsymbol{\omega}\times\mathbf{N}) + \\ + 2D^2(\mathbf{D}\cdot\mathbf{N})(\mathbf{I}\cdot\boldsymbol{\omega}) + 2D^2\big((\boldsymbol{\omega}\times\mathbf{x}_0)\times\mathbf{D}\big)\cdot \\ (\mathbf{R}\times\mathbf{D})\Big),$$

that can be rearranged to give

$$\mathbf{v}_\perp - 0_F = \frac{-s}{\|\nabla_p \mathbf{E}\|\mathbf{D}^3}\Big(f(\mathbf{n}\times\boldsymbol{\omega})\cdot\big(2D^2(\mathbf{I}\cdot\mathbf{N})\mathbf{N} + (\mathbf{D}\cdot\mathbf{R})\mathbf{x}\big) + \\ + D^2(\mathbf{I}\times\boldsymbol{\omega})\cdot\big(2(\mathbf{D}\cdot\mathbf{N})\mathbf{N} - \mathbf{x}\big) + 2D^2\big((\boldsymbol{\omega}\times\mathbf{x}_0)\times\mathbf{D}\big)\cdot(\mathbf{R}\times\mathbf{D})\Big),$$

but $\mathbf{x} = f\mathbf{n} - \mathbf{D}$; therefore,

$$\mathbf{v}_\perp - 0_F = \frac{-s}{\|\nabla_p \mathbf{E}\|\mathbf{D}^3}\Big((\mathbf{n}\times\boldsymbol{\omega})\cdot\big(2fD^2(\mathbf{I}\cdot\mathbf{N})\mathbf{N} + f(\mathbf{D}\cdot\mathbf{R})\mathbf{D}\big) + \\ + (\mathbf{I}\times\boldsymbol{\omega})\cdot\big(2D^2(\mathbf{D}\cdot\mathbf{N})\mathbf{N} - D^2\mathbf{D} + fD^2\mathbf{n}\big) + 2D^2\big((\boldsymbol{\omega}\times\mathbf{x}_0)\times\mathbf{D}\big)\cdot(\mathbf{R}\times\mathbf{D})\Big).$$

That is,

$$\mathbf{v}_\perp - 0_F = \frac{-s}{\|\nabla_p \mathbf{E}\|\mathbf{D}^3}\Big((\mathbf{n}\times\omega)\cdot\big(2fD^2(\mathbf{I}\cdot\mathbf{N})\mathbf{N} + f(\mathbf{D}\cdot\mathbf{R})\mathbf{D}\big) +$$

$$+(\mathbf{I}\times\omega)\cdot\big(-D^2(\mathbf{D}-2(\mathbf{D}\cdot\mathbf{N})\mathbf{N}) + fD^2\mathbf{n}\big) + 2D^2\big((\omega\times\mathbf{x}_0)\times\mathbf{D}\big)\cdot(\mathbf{R}\times\mathbf{D})\Big).$$

Since

$$(\mathbf{I}\times\omega)\cdot\mathbf{N} = -(\mathbf{N}\times\omega)\cdot\mathbf{I},$$

we have

$$\mathbf{v}_\perp - 0_F = \frac{-s}{\|\nabla_p \mathbf{E}\|\mathbf{D}^3}\Big((\mathbf{n}\times\omega)\cdot\big(2fD^2(\mathbf{I}\cdot\mathbf{N})\mathbf{N} + f(\mathbf{D}\cdot\mathbf{R})\mathbf{D} - fD^2\mathbf{I}\big) +$$
$$+(\mathbf{I}\times\omega)\cdot\big(-D^2(\mathbf{D}-2(\mathbf{D}\cdot\mathbf{N})\mathbf{N})\big) +$$
$$+2D^2\big((\omega\times\mathbf{x}_0)\times\mathbf{D}\big)\cdot(\mathbf{R}\times\mathbf{D})\Big),$$

but $\mathbf{I} - 2(\mathbf{I}\cdot\mathbf{N})\mathbf{N} = \mathbf{R}$, and so

$$v_\perp - O_F = \frac{s}{D^3\|\nabla_p E\|}\Big(D^2(\mathbf{I}\times\omega)\cdot(\mathbf{D}-2(\mathbf{D}\cdot\mathbf{N})\mathbf{N}) -$$
$$-f(\mathbf{n}\times\omega)\cdot\big(\mathbf{D}\times(\mathbf{D}\times\mathbf{R})\big) + 2D^2\big((\omega\times\mathbf{x}_0)\times\mathbf{D}\big)\cdot(\mathbf{R}\times\mathbf{D})\Big). \qquad (3.5.2)$$

The minimal optical flow, therefore, is equal to the motion field for any specular surface only when $\mathbf{I}$, ω and $\mathbf{n}$ are parallel.

Remark: let us consider, again, the orthographic limit. Taking into account that as $f \to \infty$, $D \to \infty$ and $\mathbf{D}/D \to \mathbf{n}$, (3.5.2) becomes

$$v_\perp - O_F = \frac{2s}{\|\nabla_p E\|}\big((\mathbf{n}\cdot\mathbf{N})(\omega\times\mathbf{I}\cdot\mathbf{N}) + (\mathbf{I}\cdot\mathbf{N})(\omega\times\mathbf{n}\cdot\mathbf{N})\big). \qquad (3.5.3)$$

Therefore, even under othographic limit, the two fields are different.

3.6. General Case

Let us consider, now, the general case. We will assume (3.4.1) as scene radiance of a surface S undergoing a given rigid motion (composition of a

rotation and a translation) in space. Adding together (3.3.1), (3.4.6) and (3.5.2), we obtain the difference between the motion field and the minimal optical flow for a surface in the general case under uniform illumination, *i.e.*:

$$v_\perp - O_F = \frac{\rho \mathbf{N}\cdot\mathbf{I}\times\boldsymbol{\omega}}{\|\nabla_p E\|} + \frac{s}{D^3}\frac{(\mathbf{v}\times\mathbf{D})\cdot(\mathbf{R}\times\mathbf{D})}{\|\nabla_p E\|} + \\ + \frac{s}{D^3\|\nabla_p E\|}\Big(D^2(\mathbf{I}\times\boldsymbol{\omega})\cdot(\mathbf{D}-2(\mathbf{D}\cdot\mathbf{N})\mathbf{N}) - \\ -f(\mathbf{n}\times\boldsymbol{\omega})\cdot(\mathbf{D}\times(\mathbf{D}\times\mathbf{R})) + 2D^2((\boldsymbol{\omega}\times\mathbf{x}_0)\times\mathbf{D})\cdot(\mathbf{R}\times\mathbf{D})\Big). \quad (3.6.1)$$

The right-hand side of (3.6.1) is generally different from zero. In fact, there are no general conditions under which it is identically equal to zero. Notice, however, that if ω and v are bounded

$$\lim_{|\nabla_p E|\to\infty} |v_\perp - O_F| = 0, \quad (3.6.2)$$

Equation 3.6.2 shows that the points in the image where the gradient is stronger are the points where the minimal optical flow is closer to the motion field. These points are characterized by sharp changes in intensity - edges -, that usually correspond to important physical events on surfaces, such as boundaries, orientation discontinuities and especially surface markings. Thus, to solve problems such as structure from motion, or the recovery of the 3-D velocity field, which require an accurate estimate of the 2-D motion field, edge-based algorithms seem more suitable than algorithms based on spatial and temporal derivatives of the image brightness. As a consequence, in order to obtain a precise reconstruction of the 2-D motion field, algorithms based on the solution of the correspondence problem among edges may be used. Notice that *matching* can be best performed between frames that are closely spaced in time whereas the *structure from motion* computation is best performed between widely spaced frames. The whole argument agrees

with the fact that, as intuitively expected, the minimal optical flow and the motion field at image *features* corresponding to precise locations on the 3-D surfaces coincide. It must be pointed out that in this analysis we have not considered shadows and self-shadow effects. They usually give rise to edges in the image that do not correspond to features in the scene. Furthermore, the Phong model of reflectance does not include sharp intensity changes due to specularities.

4. Qualitative Properties of the Minimal optical flow

Traditionally, the optical flow has been considered as the first step for recovering 3-D structure and 3-D motion. In this chapter we suggest a different use of the minimal optical flow. We argue that qualitative properties of the 2-D motion field give useful information about the 3-D velocity and the 3-D structure of surfaces and that these qualitative properties can be usefully inferred from the obtainable minimal optical flow. As an example of this approach, we introduce the qualitative properties associated with 2-D dynamical systems and show how to process minimal optical flow and motion field for making them equivalent to flows of dynamical systems on the plane. We then suggest, from properties of structural stability of dynamical systems, that the minimal optical flow may be equivalent to the motion field in terms of qualitative properties.

4.1. What is the minimal optical flow for?

In the previous section we have shown that the minimal optical flow and the motion field are different almost everywhere. As a consequence, the

minimal optical flow cannot be used to solve problems such as structure from motion and recovery of the 3-D velocity field, whose solutions rely on precise reconstruction of the 2-D motion field. We have also proved that the two fields are very similar at locations where the image gradient is strong. This led to the suggestion that feature-based algorithms may provide more reliable solutions to those problems.

Here we argue that the minimal optical flow, as a field defined almost everywhere, can be used to retrieve meaningful information about the 3-D velocity field and the 3-D structure of the scene. In particular, we consider qualitative properties of the 2-D motion fields which can be connected to significative events in the scene. Such properties are likely to be found in the corresponding minimal optical flows as well. As an example, consider an object moving toward the image plane. This kind of motion generates a focus of expansion in the 2-D motion field. The presence of a focus of expansion on the image plane, therefore, may be related to an object moving toward the plane itself. As we have seen, however, the information available is not the motion field, nor its normal component, but the minimal optical flow (or its nomal component). If the difference between the two fields is sufficiently small, we expect to find a focus of expansion also in the minimal optical flow. In the next sections we will show how the 2-D motion field and the optical flow can be considered vector fields tangent to flows of some dynamical system: it becomes then possible to establish a suggestive analogy between the theory of structural stability of dynamical systems and the qualitative description of the two fields. A focus of expansion of a dynamical system, for example, is a stable property for small perturbations of the system: this means that given a vector field with a focus of expansion, every field obtained from it by means of a sufficiently small perturbation will also show a focus of expansion.

4.2. Smoothing the Optical Flow and the Motion Field

In order to establish a connection with the theory of stability of dynamical systems, we must insure that the optical flow and the motion field have an appropriate degree of smoothness. This is not always the case, because of discontinuities arising at object boundaries or to noise affecting the optical flow data. We suggest to use a filtering step to smooth the field. It is worthwhile noticing that a filtering step on the normal component of a dense motion field is a (regularization) method to recover the whole 2-D motion field. [2]

4.3. Qualitative Descriptions of Dynamical Systems

For a rigorous and thorough review on dynamical systems see Hirsch and Smale (1974). Here, for the sake of completeness, we summarize the main definitions and results.

A *dynamical system* is a C^1 map ϕ: $R \times A \rightarrow A$, where A is an open set of an Euclidean space and writing $\phi(t,x) = \phi_t(x)$, the map ϕ_t: $A \rightarrow A$ satisfies:

(a) ϕ_0: $A \rightarrow A$ is the identity;

(b) the composition $\phi_t(\phi_s(x)) = \phi_{t+s}$ for each $t, s \in R$.

A dynamical system ϕ_t on A gives rise to a differential equation on A, that is a vector field y: $A \rightarrow E$ defined as follows:

$$y(x) = \frac{d}{dt}\phi_t(x)\Big|_{t=0}. \qquad (4.3.1)$$

[2] This "smoothed" 2-D motion field may not be the same recovered using standard algorithms, but its qualitative properties are likely to be preserved. The analogy we are about to present, indeed, will support this argument (and the equivalence between qualitative properties of the 2-D motion field and the optical flow as well).

Thus, for every x, $y(x)$ is the tangent vector to the curve $t \to \phi_t(x)$ at $t = 0$. Equation (4.3.1) can be rewritten in a more conventional way as

$$\frac{dx}{dt} = y(x). \tag{4.3.2}$$

Under suitable conditions on $y(x)$, there exists a dynamical system associated to (4.3.2) as a differential equation. Namely, a sufficient condition on $y(x)$ is that it is a C^1 function defined on an open subset of R^2. Intuitively a dynamical system can be thought as a one-parameter family of transformation ϕ_t: $A \to A$ describing the motion of the points in A as the time passes. The trajectories of the points are given by the solution curves to equation (4.3.2). Since equation (4.3.2) is *autonomous* (that is, the right-hand side does not depend explicitly on time), if $y(x^0) = 0$, then $x = x^0$ is a solution to it. Without loss of generality, we can assume that x^0 coincides with the origin. For obvious reasons, we will restrict our attention to planar systems, (*i.e.* in what follows, A will be an open set in R^2). Solutions like x^0 are called *equilibrium points* or *equilibria.* In the case of linear systems, useful qualitative information about the behaviour of the solution to (4.2.2) can be obtained from the eigenvalues of the matrix M of the coefficients of the differential equation. The restriction to planar systems reduces the classification to four fundamental cases:

I : M has real eigenvalues of opposite signs. In this case the origin is called a *saddle*: the equilibrium is unstable (an equilibrium is *stable* if any nearby solutions to it stays nearby for all the future time. It is *unstable* otherwise).

II : The eigenvalues have negative real parts. The origin is called a *sink* and it is stable equilibrium. The main property of a sink is that

$$\lim_{t \to \infty} x(t) = 0$$

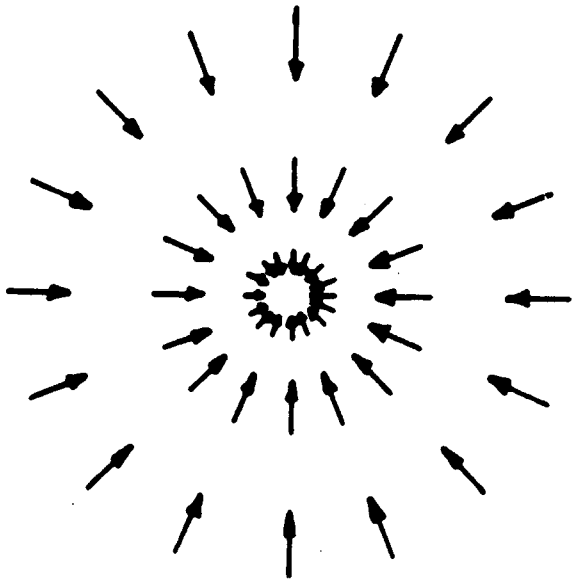

Figure 4. Vector field tangent to a planar *sink*: all the solutions curves are pointing toward the origin.

Qualitatively, the *phase portrait* of the solutions, that is, the family of the solutions curves as a subset of R^2, looks like Figure 4, where only some tangent vectors of some solutions curves have been drawn. Sinks can be classified depending on further characteristics of the eigenvalues. A *focus* (Figure 4), for example, represents the case of coincident eigenvalues (M is supposed to be diagonalizable); a *node*, the case of different real eigenvalue; a *spiral*, the case of complex conjugates eigenvalues. A sink-increasing rotational component corresponds to each different case.

III : The eigenvalues have positive real parts. The origin is called a *source*. The main property of a source is that

$$\lim_{t\to\infty} |x(t)| = \infty$$

and

$$\lim_{t\to-\infty} |x(t)| = 0.$$

A source can be considered as the dual case of a sink: the phase portrait of a source and of the correponding sink are the same except that for the

direction of the motion which must be reversed. Reversing the arrows in Figure 4, for example, obtain the phase portrait of a system with coincident real positive eigenvalues. A source is obviously an unstable equilibrium.

IV : The eigenvalues are pure imaginary. The origin is called a *center.* All the solutions are *periodic* with the same period. A center is a stable equilibrium. For a reason that will be made clear soon, this last case is of little practical interest, since even a small perturbation of the field will make the orbits spiral inward to (or outward from) the origin, changing the qualitative properties of the solution's curves. In other words, a center is not a structurally stable property.

The crucial point is that this classification is *exhaustive.* Every solution to Equation (4.3.2) (in the linear case) looks like a saddle, a sink, a source, or a center. The same classification holds for the non-linear case with respect to the eigenvalues of the derivative of the right-hand side of (4.3.2), considered as a linear operator. This is equivalent to consider a linear approximation of the system in the neighbors of the origin. However non-linear systems are interesting in theirself, since they can show also a different qualitative behavior. A non-linear system, indeed, can have in addition *limit cycles.* Intuitively, a limit cycle is a closed orbit towards which other solutions' curves spiral with the same asymptotic period. Defining a *ω-limit set*, $L_\omega(x)$, as $L_\omega(x) = \{a \in A$ such that $\exists t_n \to \infty$ with $x(t_n) \to a\}$ and similarly an *α-limit set* $L_\alpha(x)$ as $L_\alpha(x) = \{b \in A$ such that $\exists t_n \to -\infty$ with $x(t_n) \to b\}$, a limit cycle is a closed orbit γ such that $\gamma \subset L_\omega(x)$ or $\gamma \subset L_\alpha(x)$ for some $x = \gamma$. Under somewhat more restrictive conditions, a limit cycle can be a *periodic attractor* (for a rigorous definition of it see Hirsch and Smale 1974). Intuitively, a periodic attractor is a limit cycle such that nearby trajectories

not only have the same asymptotic period but also are in phase. Saddles, sinks, sources and periodic attractors are very important for a qualitative description of planar systems. Indeed it can be shown that such properties are *structurally stable*, that is they persist after a perturbation of the right-hand side of (4.3.2). As far as planar systems are concerned they also *fully* characterize limit sets. By means of the Poincaré-Bendixon theorem it can be shown that compact limit sets other than limit cycles are saddles, or sinks, or sources or trajectories joining them.

4.4. Equilibria and their Interpretations

In the definition of dynamical system the right-hand side of Equation (4.3.2) can be interpretated as a vector field tangent to the family of curves in the plane, solutions to (4.3.2) itself. It is straightforward to see that both the smoothed optical flow and motion field (i.e. after the filtering operation) can be considered as istances of such a vector field [3]. Indeed, it is sufficient to insure that both the fields are continuous with continuous first derivatives. The classification of the solutions can now be interpreted in terms of characteristic points of the 2-D motion field. A source, for example, corresponds to a focus of expansion of the field. The structural stability of the source, in turn, says that a focus of expansion persists even if the field is perturbed. From this perspective a focus of expansion is expected to be detectable in a 2-D motion field reconstructed with different algorithms and in the optical flow as well, when they can be considered as perturbed examples of the

[3] We stress the fact that the analogy with the dynamical system is between *phase portraits* of dynamical systems and motion flows. The parameter t in the definition of dynamical system is *not* the physical time. We considered motion flows, such as the 2-D motion field or the optical flow at a fixed time, comparing them with the vector field tangent to the phase portrait of some system: we are not interested in the physical meaning of the underlying dynamical system.

"true" 2-D motion field.

4.5. Discussion

If our point of view is correct, the only critical property of the optical flow is that it have the same qualitative properties of the 2-D velocity field. Notice that this requirement also satisfies two important uses of the optical flow: *to detect discontinuities* and to *help long-range matching* of the stereo type, needed for the computation of structure-from-motion. Quantitative equivalence, which is impossible in general, is in any case irrelevant for this use of the optical flow. As a consequence, *many different "optical flows" may be defined.* Equation (2.1.6) does not have any priviliged role: other definitions could be preferred on the basis of criteria such as computability (from image data) or ease of implementation (for given hardware constraints).

This point of view has clear implications for biological visual systems: movement detecting cells (say, in the retina) do not have to compute the specific minimal optical flow defined by equation (2.1.6): other, possibly simpler, estimates of the velocity field that preserve its qualitative properties are equally good candidates (such as correlation-like algorithms). This argument may explain why the models proposed to explain motion dependent behaviour in insects (Hassenstein and Reichardt, 1956), motion perception in humans (Van Santen and Sperling, 1984) and physiology of cells (Barlow and Levick, 1965; Torre and Poggio, 1978) are all implementing computations quite different from the minimal optical flow as it is usually defined (see equation 2.1.6). In addition all these models do not typically measure velocity – not even in the case of uniform translation in a frontoparallel plane. Even for simple motions of the latter type the output of models such as the correlation models depends on both the velocity and the spatial structure of

the moving pattern. One is tempted to consider this as a weakness of these models compared to the definition of minimal optical flow, Equation 2.1.7. Our results, however, show that this is not the case: first, the minimal optical flow is correct only in a very special situation; second, all these models may have the same qualitative properties of the motion field, which, from our point of view, is the only critical requirement for a "good" measurement of motion. The next question is of course whether these biological models are in fact "close" enough to the motion field to share the same qualitative properties. We do not know the answer yet. We conjecture, however, that they are indeed usually similar enough to preserve the main qualitative properties of the motion field. The conjecture is based on results (Poggio and Reichardt, 1973 and Poggio, 1985) showing that most of the biological models proposed so far can be considered as special instances or approximations of a general class of nonlinear models (characterized as Volterra systems of the second order); and that the minimal optical flow, as defined in equation, is also approximately a Volterra functional of the second order (Poggio, 1985).

It is important to stress that the approach outlined in the second part of this paper for classifying the qualitative properties of the optical flow is only one of the possible methods. While we plan to develop further that particular approach, others should be explored as well: in particular flows that do not correspond to dynamical systems on the plane may be better suited for capturing important and stable properties of the velocity field such as motion discontinuities. In this case, the classification of qualitative properties should take place without a preliminary smoothing operation.

In addition to the *classification* of *stable* qualitative properties of the velocity field, much work needs to be done at the level of their *interpretation* in terms of 3-D structure and 3-D velocity. Some of the qualitative properties

of the (smoothed) velocity field have an easy interpretation in those terms: an obvious example is again a focus of expansion that is usually related to "crashing" motion. It is likely that many, more subtle relations exist between the qualitative properties of the flow and the underlying 3-D motion and structure. For example, preliminary results by Torre et al. (personal communication) suggest that the number of *focuses* in the (smoothed) field may be characteristic for the rigidity of motion in the visible scene.

Finally, we should mention an obvious extension of the approach described in the second part of the paper. We have only considered so far the velocity field "frozen" at a given instant of time. The succession of image frames provides in fact a time-dependent field: the evolution in time of the qualitative properties we have described – how they are created, disappear and transform – should be characterized in qualitative terms, for instance using the language of catastrophe and bifurcation theory. The use of time-dependent fields should be practically much more robust, because of the redundant information available in a sequence of *very closely* spaced frames (in time). Our analysis should be extended to qualitative properties that are structurally stable not only at a given time but also in the time dependent field.

Acknowledgments We would like to thank Vincent Torre, Davi Geiger, B. Caprile, Berthold K. P. Horn, and especially Bror Saxberg for useful discussions. This work has been supported in part by a grant from the ONR (Psychology Division) to Tomaso Poggio and a grant from the Sloan Foundation to Tomaso Poggio and Shimon Ullman.

4.6. Appendix A1: Perspective and Orthographic Projections

In this section we explain in more detail the geometry of perspective projection used in the paper. Let $\mathbf{n}$ be the unit normal to the projection plane and f the focal length. In order to obtain the orthographic projection as the limit of the perspective one for $f \to \infty$, the focus cannot be located at the origin of the system of coordinates. To simplify the geometry without losing in generality, let the origin lie on the projection plane. The vector pointing from the focus to a point $\mathbf{x} = (x, y, z)$ is now $f\mathbf{n} + \mathbf{x}$. To obtain the expression of the projected point $\mathbf{x}_p$ notice that from Figure 1 is easy to see that

$$\frac{f\mathbf{n} + \mathbf{x}}{(f\mathbf{n} + \mathbf{x}) \cdot \mathbf{n}} = \frac{f\mathbf{n} + \mathbf{x}_p}{f}.$$

From that, we have

$$\mathbf{x}_p = f\frac{f\mathbf{n} + \mathbf{x}}{f + \mathbf{x} \cdot \mathbf{n}} - f\mathbf{n},$$

and finally

$$\mathbf{x}_p = \frac{f}{f + \mathbf{x} \cdot \mathbf{n}}\big(\mathbf{x} - (\mathbf{x} \cdot \mathbf{n})\mathbf{n}\big)$$

or

$$\mathbf{x}_p = \frac{f}{f + \mathbf{x} \cdot \mathbf{n}}\big(\mathbf{n} \times (\mathbf{x} \times \mathbf{n})\big).$$

The orthographic projection equation can be easily obtained for $f \to \infty$, *i.e.*

$$\mathbf{x}_{ort} = \lim_{f \to \infty} \mathbf{x}_p = \frac{f}{f + \mathbf{x} \cdot \mathbf{n}}\big(\mathbf{n} \times (\mathbf{x} \times \mathbf{n})\big) = \big(\mathbf{n} \times (\mathbf{x} \times \mathbf{n})\big).$$

Combining the last two equations, we obtain the general relationship between perspective and orthographic projection, that is

$$\mathbf{x}_p = \frac{f}{f + \mathbf{x} \cdot \mathbf{n}}\mathbf{x}_{ort}.$$

4.7. References

Barlow, H.B. and R.W. Levick, (1965) The mechanism of directional selectivity in the rabbit's retina, J. Physiol. **173:** 477–504.

Born, M. and Wolf, E. 1959 *Principles of Optics.* New York: Pergamon Press.

Gibson, J. J. 1950. *The Perception of the Visual World.* Boston: Houghton Mifflin.

Fennema, C. L., Thompson, W. B. 1979. Velocity determination in scenes containing several moving objects. *Comput. Graph. Image Proc.* 9:301–315.

Hassenstein, B. and Reichardt, W. (1956) Systemtheoretische Analyse der Zeit-, Reihenfolgen- und Vorzeichenauswertung bei der Bewgungs-perzeption der Russelkafers, Chlorophanus. Z. Naturforsch. IIb: 513–524.

Hildreth, E. C. 1984a. *The Measurement of Visual Motion.* Cambridge: MIT Press.

Hildreth, E. C. 1984b. The Computation of the velocity field. *Proc. R. Soc. London B* 221:189–220.

Hirsch, M.W. and Smale, S. 1974 *Differential Equations, Dynamical Systems, and Linear Algebra.* New York: Academic Press.

Horn, B.K.P. *Robot Vision*, Cambridge and New York: Massachusetts Institute of Technology Press and McGraw-Hill, 1986.

Horn, B. K. P., Schunck, B. G. 1981. Determining optical flow. *Artif. Intell.* 17:185–203.

Horn, B.K.P. and Sjoberg, R.W. 1979. Calculating the reflectance map. *Applied Optics* 18: 1770-1779.

Kanatani, K. 1985. Structure from motion without correspondence: general principle. *Proc. Image Understanding Workshop*, Miami, FL, pp. 107–116.

Marr, D., Ullman, S. 1981. Directional selectivity and its use in early visual processing. *Proc. R. Soc. London Ser. B* 211:151–180.

Nagel, H.–H. 1984. Recent advances in image sequence analysis. *Proc. Premier Colloque Image — Traitement, Synthese, Technologie et Applications*, Biarritz, France, May, pp. 545–558.

Phong, B.T. 1975 Illumination for computed generated pictures. *Communications of the ACM* 18:311-317.

Poggio, T. 1985. Cold Spring harbor lecture, unpublished.

Poggio, T., Reichardt, W. 1973. Considerations on models of movement detection. *Kybernetik 13*, 223-227.

Schunck, B. G., Horn, B. K. P. 1981. Constraints on optical flow computation. *Proc. IEEE Conf. Patt. Recog. Image Proc.*, August, pp. 205–210.

Torre, V. and T. Poggio, (1978) A Synaptic mechanism possibly underlying directional selectivity to motion. Proc. R. Soc. Lond. B **202:** 409–416.

Van Santen, J. P. H., Sperling, G. 1984. A temporal covariance model of motion perception. *J. Opt. Soc. Am. A*, 1:451–473.

Waxman, A. M. 1986. Image flow theory: A framework for 3–D inference from time–varying imagery. In *Advances in Computer Vision*, ed. C. Brown, New Jersey: Erlbaum. *In press.*

CS-TR Scanning Project
Document Control Form

Date : 10/18/95

Report # AIM-917

Each of the following should be identified by a checkmark:
Originating Department:

☒ Artificial Intellegence Laboratory (AI)
☐ Laboratory for Computer Science (LCS)

Document Type:

☐ Technical Report (TR) ☒ Technical Memo (TM)
☐ Other:_____

Document Information

Number of pages: 32 (38-IMAGES)
Not to include DOD forms, printer instructions, etc... original pages only.

Originals are:
☒ Single-sided or
☐ Double-sided

Intended to be printed as :
☐ Single-sided or
☒ Double-sided

Print type:
☐ Typewriter ☐ Offset Press ☒ Laser Print
☐ InkJet Printer ☐ Unknown ☐ Other:_____

Check each if included with document:

☒ DOD Form (2) ☐ Funding Agent Form ☐ Cover Page
☐ Spine ☐ Printers Notes ☐ Photo negatives
☐ Other: _____

Page Data:

Blank Pages (by page number): _____

Photographs/Tonal Material (by page number): _____

Other (note description/page number):

Description :	Page Number:
(A) IMAGE MAPS: (1-32) UNNUM'ED TITLE PAGE, 2-32 (33-38) SCAN CONTROL, DOD(2), TRGT'S(3)	
(B) CUT+PASTE F.G.'S ON PAGES, 5, 8, 11, 14	

Scanning Agent Signoff:
Date Received: 10/18/95 Date Scanned: 10/24/95 Date Returned: 10/26/95

Scanning Agent Signature: Michael W. Cook

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: 917

2. GOVT ACCESSION NO.:

3. RECIPIENT'S CATALOG NUMBER: AD-A183727

4. TITLE (and Subtitle): Motion Field and Optical Flow: Qualitative Properties

5. TYPE OF REPORT & PERIOD COVERED: Memo

6. PERFORMING ORG. REPORT NUMBER:

7. AUTHOR(s): Alessandro Verri & Tomaso Poggio

8. CONTRACT OR GRANT NUMBER(s): N00014-85-K-0124

9. PERFORMING ORGANIZATION NAME AND ADDRESS:
Artificial Inteligence Laboratory
545 Technology Square
Cambridge, MA 02139

10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS:

11. CONTROLLING OFFICE NAME AND ADDRESS:
Advanced Research Projects Agency
1400 Wilson Blvd.
Arlington, VA 22209

12. REPORT DATE: December 1986

13. NUMBER OF PAGES: 32

14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office):
Office of Naval Research
Information Systems
Arlington, VA 22217

15. SECURITY CLASS. (of this report): Uncalssified

15a. DECLASSIFICATION/DOWNGRADING SCHEDULE:

16. DISTRIBUTION STATEMENT (of this Report):

Distribution is unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report):

18. SUPPLEMENTARY NOTES:

None

19. KEY WORDS (Continue on reverse side if necessary and identify by block number):

Optical Flow
Motion Field

20. ABSTRACT (Continue on reverse side if necessary and identify by block number):

In this paper we show that the optical flow, a 2-D field that can be associated with the variation of the image brightness pattern, and the 2-D motion field, the projection on the image plane of the 3-D velocity field of a moving scene, are in general different, unless very special conditions are satisfied. The optical flow, therefore, is ill-suited for computing structure from motion and for reconstructing the 3-D velocity field, problems that require an accurate estimate of the 2-D motion field. We then suggest a different use of

---->

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-014-6601

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

the optical flow. We argue that stable qualitative properties of the 2-D motion field give useful information about the 3-D velocity field and the 3-D structure of the scene, and that they can usually be obtained from the optical flow. To support this approach we show how the (smoothed) optical flow and 2-D motion field, interpreted as vector fields tangent to flows of planar dynamical systems, may have the same qualitative properties from the point of view of the theory of structural stability of dynamical systems.

Scanning Agent Identification Target

Scanning of this document was supported in part by the **Corporation for National Research Initiatives,** using funds from the **Advanced Research Projects Agency** of the **United states Government** under Grant: **MDA972-92-J1029.**

The scanning agent for this project was the **Document Services** department of the **M.I.T Libraries.** Technical support for this project was also provided by the **M.I.T. Laboratory for Computer Sciences.**