To Appear:
Proceedings of the Eighth Israeli Conference on Artificial Intelligence and Computer Vision, Dec. 1991

Computing Motion Parameters using the Curl of the Flow Field

by

V. Sundareswaran
Courant Institute of Mathematical Sciences

Robert Hummel
Courant Institute of Mathematical Sciences
and The Center for Neural Science

New York University
Courant Institute
251 Mercer Street
New York, NY 10012
sundare@cs.nyu.edu hummel@cs.nyu.edu

This research was supported, in part, by a grant from Grumman Corporation of Bethpage, New York. We thank Brian Wetton, P. Anandan and Allan Dobbins for comments and suggestions.

Computing Motion Parameters using the Curl of the Flow Field

V. Sundareswaran
Courant Institute of Mathematical Sciences
New York University
Robert Hummel
Courant Institute of Mathematical Sciences
and The Center for Neural Science
New York University

Abstract

We consider the problem of determining the parameters of motion of a sensor in a fixed scene, given the optical flow field induced on the perspectively projected image. Specifically, we consider the problem of finding the rotational parameters of motion. Although this problem has been much studied, most methods make use of local flow field data, rather than integrating information from the entire image plane. When information is incorporated from many locations in the image domain, typical methods require the solution of large systems of equations. We suggest a new method, based on measurements of the curl of the flow field. We show that under a variety of assumptions, the flow due to translational motion is cancelled, and the rotational parameters of motion can be obtained by fitting a linear function to the available data throughout the entire image plane. The regression process gives a completely trivial method for the estimation of the rotational parameters. We further show that noise in the measurement of the curl data is inconsequential for two reasons: (1) we may instead use curl values that are averaged over large regions, by measuring circulation values about cycles, and that these measurements serve equally well as point values of the curl; and (2) the regression process attenuates noise in the measurement of the curl values, so that in practical situations, accurate estimates are obtained. We demonstrate the viability of the approach with a number of calculations for typical situations, and a number of experiments using synthetically-generated motion data.

1. Introduction

As a camera moves around in a fixed scene, a motion is induced on the image plane. The instantaneous flow field depends upon the parameters of the instantaneous motion of the sensor, as well as the distances from the sensor to the surface points of objects observed at each pixel in the image. For many applications, it is useful to determine, from the observed motion of objects in the image, the parameters of motion of the sensor. In particular, the motion of the sensor can be described by six parameters: three parameters determine the translational velocity of the focal point of the sensor, and three additional parameters give the angular velocity of the coordinate system attached to the sensor. The difficulty in determining these parameters, when they are unknown, is that typically the induced motion on the image plane must also be determined, and further, since the depths to the objects are unknown, the determination of the parameters from the

This research was supported, in part, by a grant from Grumman Corporation of Bethpage, New York. We thank Brian Wetton, P. Anandan and Allan Dobbins for comments and suggestions.

image motion is nontrivial.

A related problem arises when a sensor observes rigid objects that are moving about in a scene. Potentially, the sensor is also moving. In this case, there are motion parameters that describe the relative motion of each object with respect to the camera. Once again, it is often desired to determine the parameters of this motion for each object.

Determining motion parameters can be useful for tracking objects, for segmenting objects, and for coarse determination of depths. Often, knowledge about the relative motion of the objects is the primary objective of the image analysis process.

There are two phases to the motion parameter determination. In the first phase, the image sequence is analyzed, and the image flow field is determined. That is, the image sequence is converted to a sequence of image vector fields, where the vector field at any instant indicates the velocity of motion of each point on the image plane at that time. In the second phase, the vector fields are used to determine the motion parameters, through a process of solving an inverse problem, where the depths to the imaged points are typically unknown parameters that must also be determined. In this paper, we will concentrate on the second phase.

Both phases are well-studied problems. For the first phase, numerous methods have been suggested for determining the vector field flow from an image sequence, including the method of Horn and Schunk [1], Barnard and Thompson [2], Hildreth [3], and the more recent methods of Anandan [4] and Heeger [5].

In the second phase, the flow parameters must be determined from the optical field. In special cases, this has been extensively studied, for example, in the work of Lawton [6], Longuet-Higgins and Prazdny [7], and Prazdny [8, 9]. Solution methods based on the use of a sufficient number of correspondences are given by Waxman *et al.* [10], Waxman and Wohn [11], and Faugeras [12]. More recently, Heeger and Jepson [13] have suggested a simple linear approach based on a search for the focus of expansion. We give more details of their approach in Section 3.

We now present an overview of the proposed method. A complete formulation of the problem is given in the subsequent two sections, and a derivation of the method is given in Section 4. Section 5 presents the method in a succinct and complete fashion, and the final two sections present an analysis of the potential errors, first analytically, then empirically.

Suppose the the optical flow field in the image is $V = (u(x,y), v(x,y))$, and that feature detectors in the image measure the curl of V at a number of locations or the average of the curl of V over a number of patches in the image. The latter values, the average of the curl over a patch, will be seen to be equal to the *circulation* of V about the patch. Providing the variation in the depths are small compared to the overall depth of the objects in the scene, or providing any of a number of other conditions hold true for each sample location or patch in the image from which the curls are obtained, the sample values should lie on a linear surface over the image, i.e, $g(x,y) = ax+by+c$. The coefficients of this linear surface, which we may obtain by fitting a surface to the data, are directly proportional to the parameters of rotational velocity of the image sensor. The remaining unknowns, the translational parameters and approximate depths to points in the objects, can then be obtained from observation of the flow field V, by subtracting off the flow field due to the rotational motion that has already been determined.

The curl of the flow field has long been recognized as an important property of the motion field. However, the observation that the curl of the field is approximated by a linear function, and the use of that approximation to determine the rotational parameters, appears to be new. Koenderink and Van Doorn [14] note that the gradient of the velocity field can be decomposed into fixed fields with elementary coefficients, namely from the divergence, the curl, and the deformation of the field. We begin with their computation of the curl (converted to image coordinates), but make use of the function to solve for the rotational parameters. They instead studied the properties of these elementary fields in the case of an observer moving with respect to a plane

[15, 16]. The existence of receptive fields sensitive to the curl was hypothesised by Koenderink and Van Doorn [14] and Longuet-Higgins and Prazdny [7]. Regan and Beverly [17] studied this possibility by conducting experiments and concluded that the existence of vorticity receptors is very likely. The approximation based on constant depth surfaces, however, does not appear in these works, since the interest typically was in deriving surface structure directly.

The approach is complementary to the approach of Rieger and Lawton [18] which is based on finding depth discontinuities, where the jump in the vector flow field cancels the component of the field due to the rotational parameters, and leaves only a component dependent on the translational parameters. If confronted with a scene without regions of sufficiently constant depth, then there are likely to be sudden discontinuities, and an approach based on vector flow jumps can be used. As we will see, the assumption of constant depth, or any of a number of other sufficient assumptions, for the purposes of computing rotational parameters, is not unreasonable in many circumstances.

2. The optical flow equations

For the case of perspective projection of a scene on a planar imaging sensor, the optical flow equations are well-known and understood. Derivations for a moving sensor in a stationary environment may be found in [7, 13, 19] and yield the image optical flow (u,v) at the image point (x,y)

$$u(x,y) = \frac{1}{Z(x,y)}[-ft_1 + x \cdot t_3] + \omega_1 \left(\frac{xy}{f} \right) - \omega_2 \left(f + \frac{x^2}{f} \right) + \omega_3 y,$$

$$v(x,y) = \frac{1}{Z(x,y)}[-ft_2 + y \cdot t_3] + \omega_1 \left(f + \frac{y^2}{f} \right) - \omega_2 \left(\frac{xy}{f} \right) - \omega_3 x,$$

where (t_1, t_2, t_3) are the components of the three-dimensional translational motion of the sensor, and $(\omega_1, \omega_2, \omega_3)$ is the rotational velocity of the sensor coordinate system. Here, f is the focal length of the camera system, so that the transformation from spatial coordinates to image coordinates is governed by the equations

$$x = fX/Z, \quad y = fY/Z,$$

where $(X,Y,Z) = (X(x,y), Y(x,y), Z(x,y))$ is the position of the surface in 3-space that is imaged at the image coordinate position (x,y).

Similar equations hold for the case of nonstationary objects; then, a flow equation must be written for each object, since there are different relative translations and rotations for each such object.

As is common, we note that the flow equations may be grouped into the sum of two terms: the first term gives a flow field due to the translational components, and is modulated by the inverse depths, and the second term is a flow field due to the rotational components, and is independent of the depths:

$$V(x,y) = \begin{pmatrix} u(x,y) \\ v(x,y) \end{pmatrix} = V_{\mathbf{t}}(x,y) + V_{\boldsymbol{\omega}}(x,y),$$

with

$$V_{\mathbf{t}}(x,y) = \frac{t_3}{Z(x,y)} \begin{bmatrix} x-\tau \\ y-\eta \end{bmatrix}, \quad \tau = \frac{ft_1}{t_3}, \quad \eta = \frac{ft_2}{t_3},$$

and

$$V_{\omega}(x,y) = \begin{bmatrix} \frac{xy}{f} & -f-\frac{x^2}{f} & y \\ f+\frac{y^2}{f} & \frac{-xy}{f} & -x \end{bmatrix} \begin{bmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \end{bmatrix}.$$

Note also that the flow due to the translational components has a radial structure, expanding or contracting about a "focus of expansion" at location (τ,η), and with a magnitude modulated by the distance from the focus of expansion, the component of translation in the viewing direction (t_3), and the inverse depth to each pixel, $\rho(x,y) = 1/Z(x,y)$.

3. Motion parameter determination

The motion parameter determination problem is the following. We are given a collection of image locations $\{(x_i,y_i)\}_{i=1}^{N}$ and approximate flow field measurements at those points, $\{(u_i,v_i)\}_{i=1}^{N}$. That is, we have noisy knowledge of $(u(x_i,y_i),v(x_i,y_i)) = (u_i,v_i)$ at the points (x_i,y_i). We wish to determine the six parameters of motion of the sensor: $\mathbf{t} = (t_1,t_2,t_3)$, $\boldsymbol{\omega} = (\omega_1,\omega_2,\omega_3)$, such that there exists a collection of inverse depths ρ_i, $i=1, \cdots ,N$ satisfying, as best as possible, the flow equations:

$$\begin{bmatrix} u_i \\ v_i \end{bmatrix} = t_3 \cdot \rho_i \begin{bmatrix} x_i-\tau \\ y_i-\eta \end{bmatrix} + \begin{bmatrix} \frac{x_i y_i}{f} & -f-\frac{x_i^2}{f} & y_i \\ f+\frac{y_i^2}{f} & \frac{-x_i y_i}{f} & -x_i \end{bmatrix} \begin{bmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \end{bmatrix},$$

where (τ,η) is as defined above. Although the primary unknowns are the parameters $\mathbf{t}$, $\boldsymbol{\omega}$, the resulting inverse depths $\{\rho_i\}$ give some useful information about the structure of the object by means of the equations $Z(x_i,y_i) = 1/\rho_i$, and thus are often used in subsequent processing.

We observe that the problem has $N+6$ unknowns in $2N$ equations. Thus at least 6 points (12 equations) are needed. We further observe that since t_3 and the ρ_i are unknown, we can at best solve for the collection $t_3 \cdot \rho_i$. That is, only proportional depths can be obtained. Moreover, the unknowns for the translational parameters are expressed in terms of t_3, τ, and η; it is impossible to determine t_3 unless at least one actual depth value is known. In general, many more than 6 points are used in order to achieve stability and invariance to noise in the determination of τ, η, $\boldsymbol{\omega}$, and proportional inverse depths.

As mentioned before, many solution methods have been studied. We briefly outline the method suggested by Heeger and Jepson [13]. They observe that for fixed (τ,η), the equations are linear in the remaining unknowns $\boldsymbol{\omega}$, $\{t_3 \cdot \rho_i\}_{i=1}^{N}$. We thus have $2N$ linear equations in $N+3$ unknowns, in the case when (τ,η) is fixed. As long as $N \geq 4$, these are easily solved to give the least mean square error, and that error is also easily determined. The residual error in the least mean square solution can be used as a measures of the quality of the estimate (τ,η) for the focus of expansion — in noise-free circumstances, there should be zero residual error if τ and η are the correct values. Equivalently, we see that the data $\{(u_i,v_i)\}_{i=1}^{N}$, regarded as a vector in $2N$-space, should lie on an $N+3$-dimensional hyperplane defined by the fixed (τ,η); the extent that the vector lies off this hyperplane measure the noise in the data and the correctness of the (τ,η) estimate. Heeger and Jepson then simply check a large array of possible (τ,η) values, computing the error at each such position. Where the error is minimized, they declare the correct (τ,η) to be found, and the least mean square solution to the other variables to give the solution. For methods based on a similar approach, but set in the continuous domain thus providing more insight into the structure of the problem, see R. Hummel and V. Sundareswaran [20].

4. Using the curl of the flow field

4.1. Derivation

The *curl* (or *rot*) of a two-dimensional vector field $V(x,y) = (u(x,y), v(x,y))$ is a scalar function defined by

$$\text{curl}(V) = \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y}.$$

The curl of V is also denoted $\nabla \times V$. If we apply the curl operator to both sides of the optical flow equation, we obtain:

$$\nabla \times V(x,y) = t_3 \cdot \left[\frac{\partial \rho}{\partial x} \cdot (y - \eta) - \frac{\partial \rho}{\partial y} \cdot (x - \tau) \right] - \frac{1}{f} \left[x\omega_1 + y\omega_2 + 2f\omega_3 \right]. \tag{4.1}$$

4.2. Regression

Let us assume, for the moment, that $Z(x,y)$ is locally constant. Then $\rho(x,y)$ is also locally constant, and $\partial \rho / \partial x = \partial \rho / \partial y = 0$, and the first term on the right hand side of (4.1) vanishes. We then obtain:

$$\nabla \times V = \frac{-1}{f} \left[x\omega_1 + y\omega_2 + 2f\omega_3 \right]. \qquad \text{(Constant depth)} \quad (4.2)$$

We will later weaken the assumption of constant depth.

Now, suppose that instead of being given a collection of flow velocities $\{(u_i, v_i)\}$, we are given a collection of measurements of the curl of the flow velocity:

$$\gamma(x,y) = \nabla \times V(x,y),$$

so that we have values $\{\gamma_i\}_{i=1}^{N}$ at locations $\{(x_i, y_i)\}_{i=1}^{N}$. We then have (for the constant depth case) the linear equations:

$$\gamma_i = \frac{-1}{f} \left[x_i, y_i, 2f \right] \begin{bmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \end{bmatrix} \quad i=1, \cdots, N. \tag{4.3}$$

This system of equations is easily solved (as long as the points (x_i, y_i) are not collinear). It is more revealing, however, to observe that the right hand side of (4.2) is a linear function in x and y. Thus: *the rotational parameters of the motion can be obtained from the coefficients of a linear approximation to the curl of the flow field in a region of constant depth.*

Let us denote the linear function in x and y by g:

$$g(x,y) = \frac{-1}{f} \left[\omega_1 x + \omega_2 y + 2\omega_3 f \right].$$

We have seen that the curl, $\gamma(x,y)$, of the vector field should give an approximation to $g(x,y)$, at least in regions of constant depth. Note also that ω_1, ω_2, and ω_3 are the coefficients of the functions $\phi_1(x,y) = x/f$, $\phi_2(x,y) = y/f$, and $\phi_3(x,y) = 2$, which are mutually orthogonal functions when defined on a symmetric domain D (such as the unit disk $\{x^2+y^2 \le 1\}$, or a rectangular image $\{-a \le x \le a,\ -b \le y \le b\}$). Thus, there exist constants c_1, c_2, and c_3 (depending on the domain D) such that the coefficients satisfy

$$\omega_1 = c_1 \iint_D x g(x,y) dxdy, \quad \omega_2 = c_2 \iint_D y g(x,y) dxdy, \quad \omega_3 = c_3 \iint_D g(x,y) dxdy.$$

In the case where we have a good estimate (by virtue of constant depth) of $g(x,y)$ from $\gamma(x,y)$ over an entire domain D (which is symmetric about the origin of the image plane), then we obtain the striking result that the parameters of ω may be obtained from that estimate by integration against three specific basis functions.

4.3. Conditions for linearity

We now weaken the restrictions. We first observe that the Equation (4.3) requires that $\rho(x,y)$ be constant in a neighborhood of (x_i,y_i), but is independent of the value of $\rho(x_i,y_i)$. Thus if the collection of points $\{(x_i,y_i)\}$ are at different depths, but that each is obtained from a region where $\rho(x,y)$ is locally constant, then the same set of equations result. Further, if there are points (x_i,y_i) where $\rho(x,y)$ is not locally constant, then the data $\gamma(x_i,y_i)$ may lie off of the linear surface $g(x,y) = -(\omega_1 x + \omega_2 y + 2f\omega_3)/f$, but that this defect can be observed as an outlier providing there are sufficient number of points lying on the linear surface so as to deduce the coefficients.

Next, observe that constant depth is a sufficient condition, but that there are other situations where the first term of the right hand side of Equation (4.1) will vanish. Specifically:

- If the translational velocity **t**, is zero, then the term vanishes;
- Near the focus of expansion (τ,η), the term is small, and at the focus of expansion, the term vanishes; and
- If the vector

$$\left[\frac{\partial\rho}{\partial x}, \frac{\partial\rho}{\partial y}\right] \tag{4.4}$$

  is proportional to the vector $(x-\tau, y-\eta)$, then the term vanishes at that point. The vector (4.4) is the same as

$$\frac{-1}{Z^2(x,y)}\left[\frac{\partial Z}{\partial x}, \frac{\partial Z}{\partial y}\right],$$

  which is very nearly proportional to the "tilt" of the surface in space (X,Y,Z) that is imaged on the pixel (x,y). By tilt, we mean the projection of the surface normal onto the $Z=0$ plane. Thus the condition is that the tilt of the surface imaged at the point (x,y) should be the same direction as a vector from (x,y) to the focus of expansion.
- If the distance $Z(x,y)$ to the object is large compared to the variation in the Z with respect to x and y, then even if the distance is not constant, the first term will be negligible. Note that the variation in Z is divided by $Z^2(x,y)$.

In each of these situations, the equation (4.2) is valid, and so the vector field curl values $\{\gamma_i\}$ will lie on the planar surface defined by $g(x,y)$.

4.4. A robust method for obtaining curl values

Finally, we consider the problem of obtaining the values of the curl of the vector field. Since the flow vector field is already a velocity field involving first derivatives, the curl of the vector field involves a difference of second derivatives, and is therefore likely to be quite noisy. Suppose that instead of being given $\gamma(x_i,y_i) = \nabla\times V(x_i,y_i)$ at a fixed point (x_i,y_i), we are given the average of γ over a region D_i (which need not be symmetric):

$$\tilde{\gamma}_i = \frac{1}{|D_i|}\iint_{D_i} \gamma(x,y)\,dx\,dy,$$

(see Fig. 1). Assuming that ρ is constant over D_i, or any of the other above conditions hold over D_i, we then have that

$$\tilde{\gamma}_i = \frac{1}{|D_i|}\iint_{D_i} \frac{-1}{f}[x\omega_1+y\omega_2+2f\omega_3]dxdy = \frac{-1}{f}[\tilde{x}_i\omega_1+\tilde{y}_i\omega_2+2f\omega_3],$$

where $(\tilde{x}_i,\tilde{y}_i)$ is the centroid of the region D_i in the image coordinates. Thus, because the equation is linear in x and y, the curl data may be averaged over regions, and the same equations hold, where the point location is evaluated at the centroid of the region.

Obtaining the average of the curl over a region in the image is plausible due to Stoke's theorem:

$$\iint_D \mathrm{curl}(V)dxdy = \oint_{\partial D} V(x,y)ds.$$

That is, the average of the curl over a domain can be replaced by a normalized (i.e., scalar multiple of the) contour integral around the boundary of the domain. Thus the average values of the curl can be obtained from contour integrals of the vector flow field $(u(x,y),v(x,y))$ around the boundary of the regions. To be explicit, we recall that the contour integral of a vector field can be defined in terms of an arc-length parameterization of a (smooth) closed curve, say $(x(s),y(s))$, for $s\in[0,L]$. Then the contour integral and thus the i^{th} sample is given by

$$\tilde{\gamma}_i = \int_0^L [u(x(s),y(s))\frac{dx}{ds}(s) + v(x(s),y(s))\frac{dy}{ds}(s)]ds.$$

In fact, this value is independent of the parameterization of the closed curve. We will call this value the *circulation* of the flow field about the cycle ∂D_i. For actually obtaining the circulation values, there are three possibilities: (1) computing the contour integral from the flow field, as indicated above; (2) computing or sensing the curl of the flow field at all points in D_i and then averaging the values to obtain the circulation; or (3) sensing the circulation directly, using a collection of "hard-wired" circulation detectors. We do not here advocate one method in preference to the others.

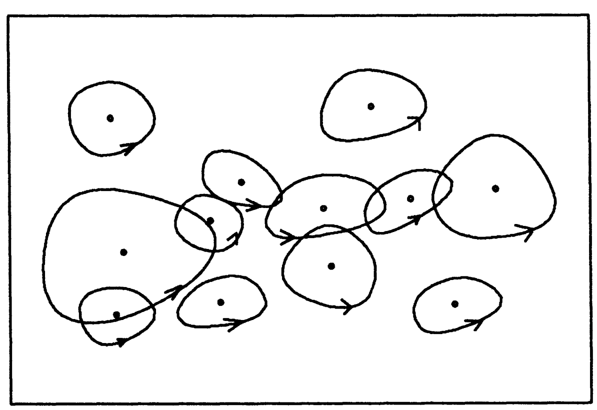

Figure 1. Contours of regions D_i and their centroids.

5. The proposed algorithm

We summarize and discuss the proposed algorithm for computing motion parameters.

We assume we are given a (perhaps coarsely sampled) collection of estimates of the image flow field $\{(u_i,v_i)\}_{i=1}^{N}$, at associated locations $\{(x_i,y_i)\}_{i=1}^{N}$, and *additionally* a collection of measures of the circulation of the flow field about various cycles:

$$\tilde{\gamma}_i = \oint_{\partial D_i} (u,v) ds,$$

for cycles ∂D_i, $i = 1, \cdots, M$ in the image domain. We further assume that either the translation velocity $\mathbf{t}$ is zero, or that each image region D_i is predominantly characterized by points where any of the following conditions hold:

(1) The surface imaged at that point has locally constant depth;

(2) The depth of the surface imaged at that point is large compared to the variation in the depths in the region about that point; or

(3) The tilt of the surface imaged at that point is directed away or toward the focus of expansion — equivalently, the normal to the surface at the point imaged at the point intersects in 3D space the line from the sensor extending in the direction of translational velocity of the sensor.

It is not required that all points in D_i satisfy the same one of these conditions; rather, we only need that each point $(x,y) \in D_i$ satisfies at least one of the conditions, or else that the measure of the set of points in D_i where none of the conditions is satisfied is small. Further, violations will not create any problems providing there is sufficient variation in the tilts of the surfaces, so that errors cancel out.

We note that the existence of perceptual filters capable of detecting local circulation of the flow field is not inconsistent with the results of [17].

For each region D_i, we determine the centroid of the region $(\tilde{x}_i,\tilde{y}_i)$, for $i = 1, \cdots, M$. We use the collection of data

$$(\tilde{x}_i,\tilde{y}_i,\tilde{\gamma}_i), \quad i = 1, \cdots M,$$

to fit a linear surface $g(x,y) = -(\omega_1 x + \omega_2 y + \omega_3 2f)/f$. The coefficients ω_1, ω_2, and ω_3 can be obtained in a variety of ways, but will in any case be an adaptive weighted sum of the $\tilde{\gamma}_i$ data. In particular, assuming that the collection of centroids is reasonably symmetric about the origin of the image plane, ω_3 will be a weighted average of the $\tilde{\gamma}_i$ data, ω_1 will be a weighted average of the same data with weights proportional to the x-coordinate position of the centroid, and ω_2 will likewise be a weighted average with weights proportional to the y-coordinate of the centroid. Of course, outliers can be detected and discarded before the averaging.

Of course, once the rotational parameters are known, it is not difficult to solve for other motion parameters and inverse depths.

6. An analytic analysis of the sensitivity

We calculate the range of errors that might be expected for the proposed algorithm. In particular, we are interested in computing the degree to which violation of the assumptions affect the linearity of the curl of the flow field.

We assume that units of measurement are in terms of the focal length of the system, so that $f = 1$. In these units, the range of x and y for a reasonably wide angle camera (60 degrees) satisfies $-1/2 \le x,y \le 1/2$. For the purposes of the analysis, let us consider the imaging and flow of a flat surface patch at location (X_0,Y_0,Z_0) with surface normal (n_1,n_2,n_3):

$$(X-X_0)\cdot n_1 + (Y-Y_0)\cdot n_2 + (Z-Z_0)\cdot n_3 = 0.$$

Since our analysis will be entirely local, our conclusions on the relative sizes of the parameters for which errors are small will apply to more complicated surfaces, where (X_0,Y_0,Z_0) is the location of the patch, and (n_1,n_2,n_3) refers to the local normal to the surface (in 3D space). We then have

$$\frac{\partial \rho}{\partial x}(x,y) = \frac{n_1}{n_1X_0+n_2Y_0+n_3Z_0}, \quad \frac{\partial \rho}{\partial y}(x,y) = \frac{n_2}{n_1X_0+n_2Y_0+n_3Z_0}.$$

Typical values for the magnitudes of the rotational parameters will range from zero to 0.5 radians per second. The t_3 translational parameter varies greatly according to the application. Values of t_3 in the range of 0 to 100 focal length units per second are plausible.

Recall again Equation (4.1), the formula for the curl of the flow field:

$$\nabla\times V(x,y) = t_3\cdot\left[\frac{\partial \rho}{\partial x}\cdot(y-\eta) - \frac{\partial \rho}{\partial y}\cdot(x-\tau)\right] - \left[x\omega_1 + y\omega_2 + 2\omega_3\right],$$

where we have set $f = 1$, and t_3 is to be measured in focal length units per second. Setting $R = n_1X_0+n_2Y_0+n_3Z_0$, we then have

$$\nabla\times V(x,y) = t_3\cdot\left[\frac{n_1\cdot(y-\eta) - n_2\cdot(x-\tau)}{R}\right] - \left[x\omega_1 + y\omega_2 + 2\omega_3\right].$$

Regrouping, we have

$$\nabla\times V(x,y) = x\left[-\omega_1-\frac{n_2\cdot t_3}{R}\right] + y\left[-\omega_2 + \frac{n_1\cdot t_3}{R}\right] + \left[-2\omega_3 - \frac{(n_1\cdot t_2-n_2\cdot t_1)}{R}\right]$$

Clearly, our ability to estimate the parameters ω_1, ω_2, and ω_3 will depend on the size of the remaining terms in the corresponding coefficients. That is, we need that the terms

$$\frac{n_2\cdot t_3}{R}, \quad \frac{n_1\cdot t_3}{R}, \quad \frac{(n_1\cdot t_2-n_2\cdot t_1)}{2R}$$

are small in magnitude, respectively, relative to the expected sizes of ω_1, ω_2, and ω_3. Note that we have substituted t_1 and t_2 for $t_3\tau$ and $t_3\eta$ respectively. More realistically, it is the *average* values of n_1 and n_2 throughout the image that influences the accuracy of the linear regression that is used to estimate the components of $\boldsymbol{\omega}$. Although typical values of the average surface normal tilts must be determined empirically, many scenes are composed of a variety of tilt directions. Fig. 2 shows plots of n_1/R and n_2/R for a typical scene of a road, computed using depth data. The values of n_1/R are well-distributed about zero, and average to 0.0040. The values of n_2/R, on the other hand, are predominantly positive, and average to 0.0103. The result is that using this scene, the value of the horizontal rotational component ω_1 is likely to be estimated with greater error when using the flow circulation algorithm.

We also note that if the components of $\mathbf{t}$ are bounded, then these terms will be small as long as R is sufficiently large, since $-1\le n_1,n_2\le 1$ always. For example, if $|t_3|\le 1$ (a fairly small forward velocity), then as long as R is greater than 50 or so (focal units), then ω_1 and ω_2 components on the order of 0.1 or 0.2 radians per second should be reasonably accurately deducible. However, note that R is the distance from the origin to the planar surface obtained by extending the surface patch in all directions, measured in focal distance units at the point of closest approach (see Fig. 3). Thus a nearby patch will always lead to a relatively small R, whereas a distant patch could cause a problem if it is tilted so that the plane passes near the origin. However, such patches are precisely the surface elements that will be greatly foreshortened, and thus

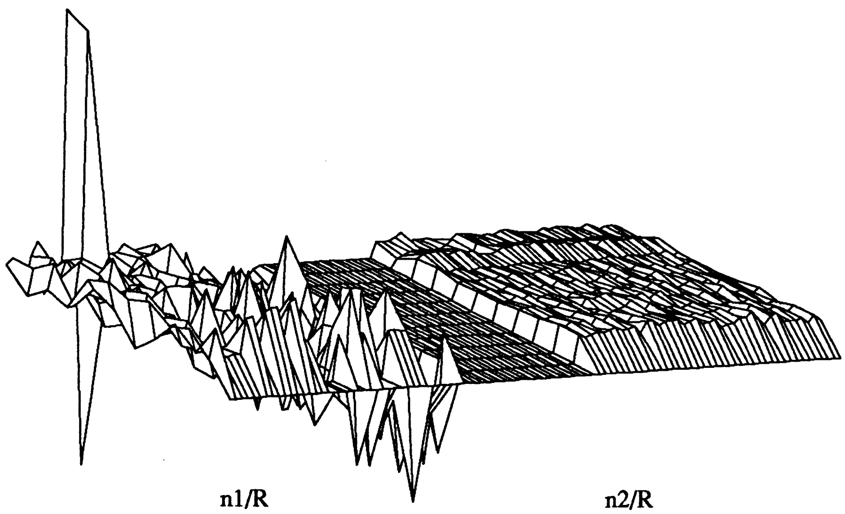

Figure 2. The plots for n1/R and n2/R for an outdoor scene.

will occupy only a small area on the image plane! Accordingly, as long as we confine ourselves to distant surface patches, we can expect that much of the image area will be dominated by surface patches with large values of R.

7. Experimental results

We have applied the algorithm to synthetically-generated vector flow fields. First, consider the image of an ellipsoid, defined by

$$\left[\frac{X-X_0}{a}\right]^2 + \left[\frac{Y-Y_0}{b}\right]^2 + \left[\frac{Z-Z_0}{c}\right]^2 = 1.$$

We will assume that the ellipsoid center is 150 focal units away (along the Z-axis), and the semiradii are 50, 30, and 70 focal units respectively. We assume the ellipsoid is in front of a flat planar background. With the observer moving with a velocity of (0.3,0,2) focal units per second, and rotating with an angular velocity of (0.2,0.1,0.5) radians per second, the curl of the vector flow field will have values as graphed in Figure 4. The true linear surface is seen in the background region, the distortion caused by the translational velocity in conjunction with the surface tilts. When a linear regression is used to fit a surface to the data, the resulting coefficients indicate an estimated rotational velocity of (0.2, 0.1008, 0.5) radians per second. If there is error in the determination of the curl of the vector field, then the resulting surface shown in Figure 4 will be similarly perturbed. However, the surface fit by regression will not be sensitive to random independent noise. That is, estimates of the linear surface will be based on averaged values over regions, and not on local gradients of the displayed surface.

Next, we consider a corridor (see Figure 5). The sensor moves with fixed velocity along the central line of the corridor, with a slight sideways drift. The view down the corridor is as shown in Figure 5. The translational velocity satisfies $t_1 = 1$, $t_2 = 0$, and $t_3 = 5$ focal units per second,

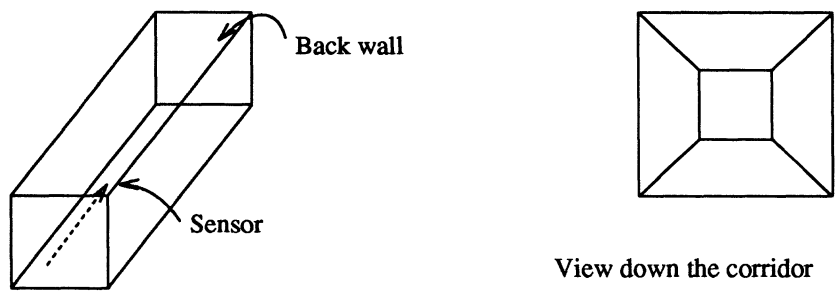

Figure 5. The corridor, and the view from the sensor.

and $\omega = (0.2, 0.1, 0.5)$ radians per second. The actual curl of the vector field is shown in Figure 6; on the back wall of the corridor, the curl is precisely linear and is entirely due to the rotational parameters; on the sides, there is a deviation due to the translational parameters and the tilt of the surfaces. In some sense, a corridor is a worst-case scenario for the algorithm, due to the abundance of surface area with small R. However, clearly, the use of large-scale averages, or the best fit linear surface, is likely to lead to the correct parametric estimates. In this case, the regression fit of a surface to the observed data leads to the estimate $\omega = (0.2, 0.1379, 0.5)$ radians per second. Once again, noise in sensing the curl values will perturb the surface shown in Figure 6. However, if circulation values are calculated about circuits, then the surface will be locally averaged, accordingly. The accuracy in determining the rotational parameters will depend on the accuracy with which the parameters of the plane defined by the surface in the central region (where the back face of the corridor is imaged).

References

[1] Horn, B. K. P. and B. Schunk, "Determining optical flow," *Artificial Intelligence* **17**, pp. 185-203 (1981).

[2] Barnard, S. and W. Thompson, "Disparity analysis of images," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, p. 333 (1980).

[3] Hildreth, E.C., "Computations underlying the measurement of visual motion," *Artificial Intelligence* **23**, pp. 309-354 (1984).

[4] Anandan, P., "A computational framework and an algorithm for the measurement of visual motion," *International Journal of Computer Vision* **2**, pp. 283-310 (1989).

[5] Heeger, David J., "Optical flow using spatiotemporal filters," *International Journal of Computer Vision* **1**, pp. 279-302 (1988).

[6] Lawton, D., "Processing translational image sequences," *Computer Vision, Graphics and Image Processing* **22**, pp. 116-144 (1983).

[7] Longuet-Higgins, H. C. and K. Prazdny, "The interpretation of a moving retinal image," *Proc. Royal Society of London, B* **208**, pp. 385-397 (1980).

[8] Prazdny, K., "Egomotion and relative depth map from optical flow," *Biological Cybernetics* **36**, pp. 87-102 (1980).

[9] Prazdny, K., "Determining the instantaneous direction of motion from optical flow generated by a curvilinearly moving observer," *Computer Graphics and Image Processing* **17**, pp. 238-248

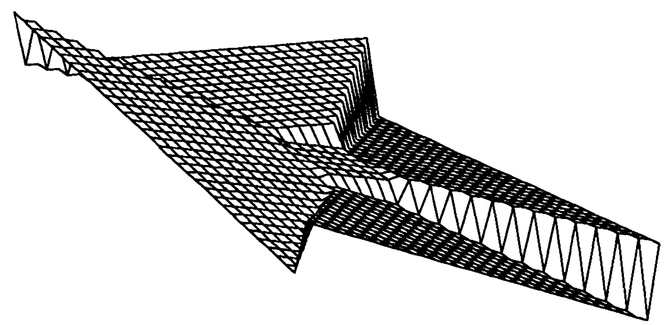

Figure 6. The curl of the flow field for the image of a corridor.

(1981).

[10] Waxman, A., B. Kamgar-Parsi, and M. Subbarao, "Closed form solutions to image flow equations," *Proceedings of the First International Conference on Computer Vision*, pp. 12-24 (1987).

[11] Waxman, A. and K. Wohn, "Image flow theory: A framework for 3-D inference from time varying imagery," pp. 165-224 in *Advances in Computer Vision*, Christopher Brown (Eds.), Lawrence Erlbaum Associates (1988).

[12] Faugeras, O., F. Lustman, and G. Toscani, "Motion and structure from motion from point and line matches," *Proceedings of the First International Conference on Computer Vision*, pp. 25-34 (1987).

[13] Heeger, D. and A Jepson, "Subspace methods for recovering rigid motion I: Algorithm and implementation," University of Toronto, Research in Biological and Computational Vision Tech Rep, RBCV-TR-90-35, (). Submitted to *International Journal of Computer Vision.*

[14] Koenderink, J.J. and A.J. Van Doorn, "Invariant properties of the motion parallax field due to the movement of rigid bodies relative to an observer," *Optica Acta* **22** (9), pp. 773-791 (1975).

[15] Koenderink, J.J. and A.J. Van Doorn, "Local structure of movement parallax of the plane," *Journal of the Optical Society of America* **66** (7),(1976).

[16] Koenderink, J.J. and A.J. Van Doorn, "Exterospecific component of the motion parallax field," *J. Opt. Soc. Am.* **71** (8), pp. 953-957 (1981).

[17] Regan, D. and K.I. Beverley, "Visual responses to vorticity and the neural analysis of optic flow," *J. Opt. Soc. Am. A* **2** (2), pp. 280-283 (Feb 1985).

[18] Rieger, J.H. and D.T. Lawton, "Determining the instantaneous axis of translation from optical flow generated by arbitrary sensor motion," *Proc. ACM interdisciplinary workshop on motion, Toronto, Canada,* pp. 33-41 (1983).

[19] Horn, B. K. P., *Robot Vision,* MIT Press, Cambridge (1986).

[20] Hummel, R. and V. Sundareswaran, "Motion parameter estimation from global flow field data," NYU Robotics Lab Technical Report, (December, 1990). To appear, *IEEE Transactions on Pattern Analysis and Machine Intelligence.*