International Journal of Computer Vision, 13:1, 33–56 (1994)

Estimating the Heading Direction Using Normal Flow

YIANNIS ALOIMONOS AND ZORAN DURIC*
Computer Vision Laboratory, Center for Automation Research, Department of Computer Science and Institute for Advanced Computer Studies, University of Maryland College Park, MD 20742-3411

Received April 13, 1992. Revised September 16, 1993.

Abstract

If an observer is moving rigidly with bounded rotation then normal flow measurements (i.e., the spatiotemporal derivatives of the image intensity function) give rise to a constraint on the oberver's translation. This novel constraint gives rise to a robust, qualitative solution to the problem of recovering the observer's heading direction, by providing an area where the Focus of Expansion lies. If the rotation of the observer is large then the solution area is large too, while small rotation causes the solution area to be small, thus giving rise to a robust solution. In the paper the relationship between the solution area and the rotation and translation vectors is studied and experimental results using synthetic and real calibrated image sequences are presented. This work demonstrates that the algorithm developed in (Horn and Weldon 1987) for the case of pure translation, if appropriately modified, results in a robust algorithm that works in the case of general rigid motion with bounded rotation. Subsequently, it has the potential to replace expensive accelerometers, inertial systems and inaccurate odometers in practical navigational systems for the problem of kinetic stabilization, which is a prerequisite for any other navigational ability.

1 Introduction

The problem of passive navigation has attracted a lot of attention in the past ten years (Bruss and Horn 1983; Longuet-Higgins 1981; Longuet-Higgins and Prazdny 1980; Spetsakis and Aloimonos 1988; Tsai and Huang 1984; Ullman 1979) because of the generality of a potential solution. The problem has been formulated as follows: Given a sequence of images taken by a monocular observer undergoing unrestricted rigid motion in a stationary environment, to recover the 3-D motion of the observer. In particular, if (U, V, W) and $(\omega_x, \omega_y, \omega_z)$ are the translation and rotation, respectively, comprising the general rigid motion of the observer, the problem is to recover the following five numbers: the direction of translation $(\frac{U}{W}, \frac{V}{W})$ and the rotation $(\omega_x, \omega_y, \omega_z)$. (See Figure 1 for a pictorial description of the geometric model of the observer; O is the nodal point of the eye).

The problem has thus been formulated as the general 3-D motion estimation problem (kinetic depth or structure from motion) and its solution would solve a series of problems (for example target pursuit, visual rendezvous, etc.) as simple applications. In this paper we study the problem of passive navigation in the framework of purposive vision (Aloimonos 1990a; Aloimonos 1992). Our basic thesis is that we must seek a robust solution for the problem under consideration only. If our proposed solution for the passive navigation problem also solves the problem of determining the 3-D motion of an object moving in the field of view of a static observer, then we have solved a more general problem than the one we initially considered. In addition, the technique has qualitative characteristics. For an example of qualitative approaches to visual motion problems, see Burger and Bhanu (1990),

*This work was funded in part by ARPA, ONR, NSF (under a Presidential Young Investigator Award, Grant IRI-90-57934), Alliant Techsystems, Inc., Texas Instruments, Inc., and Sony Corporation. Thanks to Sara Larson for her expert help in preparing this paper, to Zoran Duric for performing the experiments, and to the anonymous reviewers whose comments significantly improved the manuscript.

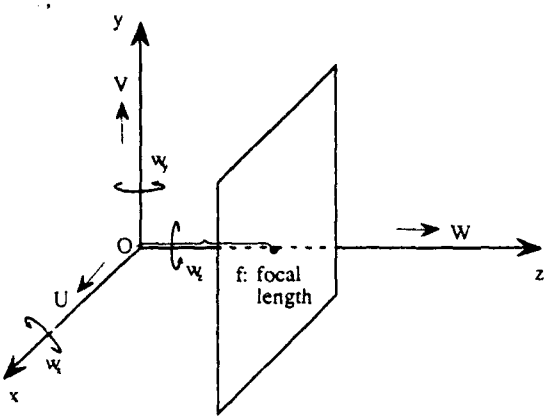


Fig. 1. Image plane perpendicular to the optical axis OZ.

Francois and Bouthemy (1990), Thompson and Kearney (1986), Thompson and Painter (1992), Tistarelli and Sandini (1992), Weinshall (1990), Weinshall (1991), Zisserman and Cipolla (1990), and Fermüller (1993a).

2 Previous Work

Previous research can be classified into two broad categories: methods based on optic flow or correspondence and direct methods.[1]

In the first category, whose uniqueness properties are well understood (Faugeras and Maybank 1990), under the assumption that optic flow or correspondence is known with some uncertainty, finding the best solution results in a non-linear optimization problem. One develops an error measure (usually a function of the input error) that is minimized in some way. Treating the problem as one of statistical estimation has given rise lately to very sophisticated approaches. Although such research on general recovery is making tremendous progress, the existing general recovery results cannot yet survive in the real world, because small amounts of error in the input can produce very large errors in the output (Spetsakis and Aloimonos 1988; Horn 1990; Young and Chellappa 1988; Weng, Huang, and Ahuja 1987). Although it is true that if a human operator corresponded features in the successive image frames,[2] most of these algorithms would give practical results, it is highly questionable that these algorithms could be used in a real time navigational system, when an average of 1% input noise is enough to create an error of 100% in the output,[3] and especially when the problem of computing optic flow or displacements (correspondence) is ill-posed and any algorithm for computing them must rely on assumptions about the world that might not always be valid. There is no doubt that research on the topic will continue and will shed more light on the difficulties associated with the general problem of 3-D motion computation.

In the second category, direct methods attempt to recover 3-D motion using as input the spatiotemporal derivatives of the image intensity function, thus getting rid of the correspondence problem. These techniques were pioneered in (Aloimonos and Brown 1984) for the case of pure rotation and developed much further by Horn and his associates (Horn and Weldon 1987; Negahdaripour 1986; White and Weldon 1987) for the case of translation only. Recently Fermüller addressed the general case (unrestricted rigid motion) (Fermüller 1993b) by discovering geometric constraints on the normal flow signs that take the form of global patterns in the image plane. Here, we treat the general problem but for the case where the rotation is bounded. When this paper was under review it came to our attention that the same result was developed independently in (Blake, Murray, and Sinclair 1992).

3 Kinetic Stabilization

Consider a monocular observer as in Figure 2. We assume that the observer moves only forward (see Figure 3).[4] It is assumed that the observer is equipped with inertial sensors which provide the rotation$(\omega_x, \omega_y, \omega_z)$ of the observer at any time. As the observer moves in its environment, normal flow fields are computed in real time. Since optic flow due to rotation does not depend on depth but on image position (x, y), we know (and can compute in real time) its value (u^R, v^R) at every image point along with the normal flow.[5] That means that we know the normal flow due to translation (see Figure 3a). In other words, since we can derotate, we assume that the normal flow is due to translation only. In later sections we analyze the case where rotation is present. When the observer moves forward[6] in a static scene, it is approaching anything visible in the scene and the flow is expanding. From Figure 3b, it is clear that the focus of expansion (FOE) = $(\frac{U}{W}, \frac{V}{W})$ (when the gradient space of directions is superimposed on the

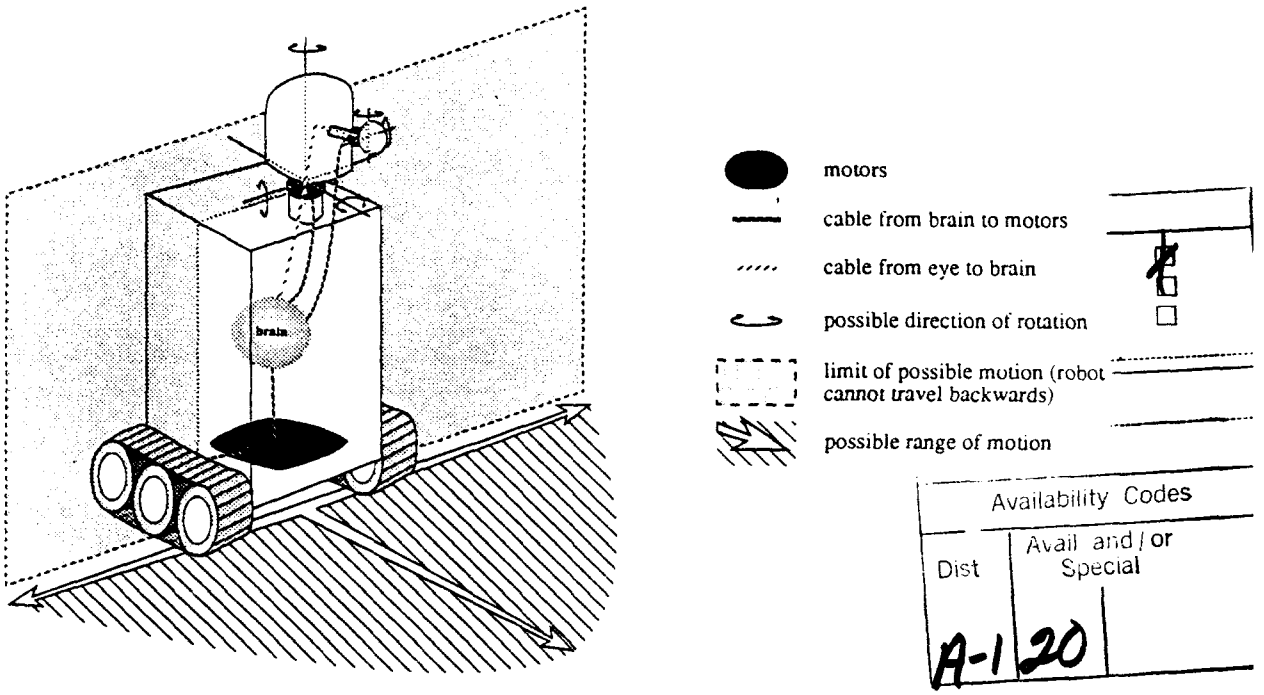


Fig. 2.

image space) lies in the half plane defined by line (ϵ); thus every point in that half space receives one vote for being the FOE. Clearly, at every point we obtain a constraint-line which constrains the FOE to lie in a half plane. If the FOE lies on the image plane (i.e. the direction of translation is anywhere in the solid sector $OABCD$ (Figure 4)) then the FOE is constrained to lie in an area on the image plane and thus it can be localized (see Figure 5). When the FOE does not lie inside the image, a closed area cannot be found, but the votes collected by the half planes indicate its general direction. By making a "saccade", i.e. a rotation of the camera, the observer can then bring the FOE inside the image and localize it (Figure 6 explains the process).

An algebraic way to derive the same constraint (Horn and Weldon 1987) is as follows: If $f(x, y, t)$ is the image intensity function, then we have $f_x u + f_y v + f_t = 0$, where u, v is the flow. If we only have translation (or we know the rotation), then we get $f_x\left(\frac{-U+xW}{Z}\right) + fy\left(\frac{-V+yW}{Z}\right) + f_t = 0$ or $f_x\frac{W}{Z}\left(x - \frac{U}{W}\right) + f_y\frac{W}{Z}\left(y - \frac{V}{W}\right) + f_t = 0$ and if $\frac{W}{Z} > 0$, $\left(f_x\left(x - \frac{U}{W}\right) + f_y\left(y - \frac{V}{W}\right)\right)/f_t < 0$.

This linear inequality in $\frac{U}{W}$, $\frac{V}{W}$, (i.e. the FOE) constrains the FOE to lie on one side of the line normal to (f_x, f_y). The contribution of this paper is to show that this simple constraint intersection technique, when appropriately modified, works even in the presence of rotation.

4 The Algorithm

We assume that the computation of the normal flow, the voting and the localization of the area containing the highest number of votes can be done in real time. In this paper we don't get involved with real time implementation issues as we wish to analyze the theoretical aspects of the technique. However it is quite clear that computation of normal flow can be done in real time (there already exist chips performing edge detection). According to the literature on connectionist networks (Ballard 1984), voting can also be done in real time. Let S denote the area with the highest number of votes. Let $L(S)$ be a Boolean function that is true when the intersection of S with the image boundary is the null set, and false otherwise. Then, the following algorithm finds the area S, i.e., solves

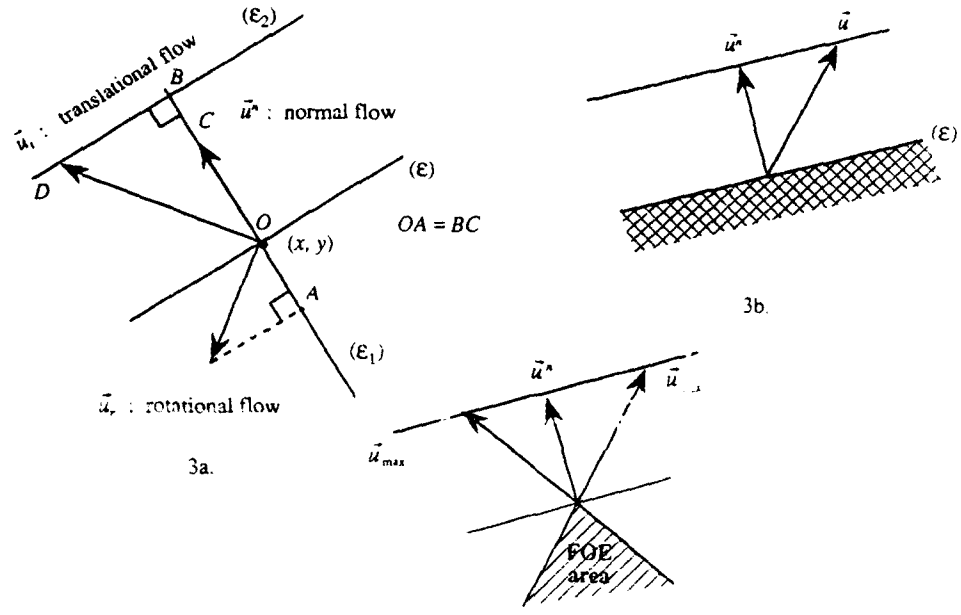


Fig. 3. Given the normal flow $\vec{u}^n$ and the rotational flow $\vec{u}_R$ at a point $O(x, y)$, and given that the projection of the sum $\vec{u}_R + \vec{u}_t$ on (ϵ_1) should equal $\vec{u}^n$, we conclude that the transitional flow is OD, where D is anywhere on (ϵ_2). Clearly, in such a case, the focus of expansion lies on the half plane defined by (ϵ) that does not contain $\vec{u}_t$. This statement is equivalent to the following algebraic inequality (Horn and Weldon 1987). If $f(x, y, t)$ is the image intensity function, then we have $f_x u + f_y v + f_t = 0$, where u, v is the flow. If we only have translation (or we have rotation), then we get $f_x\left(\frac{-U+xW}{Z}\right) + fy\left(\frac{-V+yW}{Z}\right) + f_t = 0$ or $f_x\frac{W}{Z}\left(x - \frac{U}{W}\right) + f_y\frac{W}{Z}\left(y - \frac{V}{W}\right) + f_t =)$ and if $\frac{W}{Z} > 0$, $\left(f_x\left(x - \frac{U}{W}\right) + f_y\left(y - \frac{V}{W}\right)\right)/f_t < 0$. However, thinking in terms of normal flow due to translation is as in Figure 3b, the FOE must lie in the half plane (dotted line) of (ϵ). But this assumes that the flow $\vec{u}$ can be arbitrarily large, which is absurd. If there is a bound on the flow, then the FOE is constrained further (Figure 3c).

the passive navigation problem. We assume that the inertial sensors provide the rotation and thus we know the normal flow due to translation.

```
1. begin {
2.         find area S
3.         repeat until L(S)
4.         { rotate camera around x, y axes so that
             the optical axis passes through the
             center of S (saccade)
5.           find area S
           }
6.         output S
       }
```

If the camera has a wide angle lens, then image points can represent many orientations, and only one saccade (or none) may be necessary. But if we have a small angle lens, then we may have to make more than one saccade.

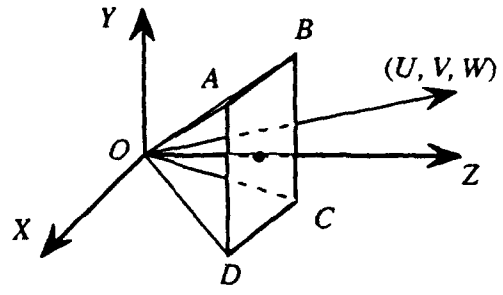


Fig. 4. Consider the camera coordinate system. If the translation vector (U, V, W) is anywhere inside the solid $OABCD$ defined by the nodal point of the eye and the boundaries of the image, then the FOE is somewhere on the image.

5 Analysis of the Method

We have assumed that the inertial sensors will provide the observer with accurate information about rotation.

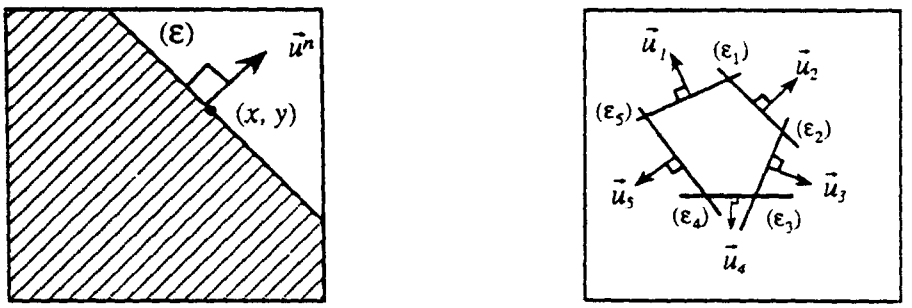


Fig. 5. (a) From a measurement of $\vec{u}^n$ of the normal flow due to translation at a point (x, y) of the image, every point of the image belonging to the half plane defined by (ϵ) that does not contain $\vec{u}$ is a candidate for the position of the focus of expansion, and collects one vote. The voting is done in parallel for every image measurement. (b) If the FOE lies within the image boundaries, then the area containing the highest number of votes is the area containing the FOE. Using only a few measurements can result in a large area. Using many measurements (all possible) results in a small area (in our experiments a small area means a few pixels, usually at most three or four).

Although expensive accelerometers can achieve very high accuracy, the same is not true for inexpensive inertial sensors and so we are bound to have some error. Thus we must assume that some unknown rotational part still exists and contributes to the value of the normal flow. As a result, the method for finding the FOE (previous section) which is based on translational normal flow information (since we have "derotated") might be affected by the presence of some rotational flow. In this section, we study the effect of rotation (the error of the inertial sensor) on the technique for finding the FOE.

In order to avoid artificial problems introduced by perspective distortions in the case of a planar retina and to simplify the formulas without loss of generality, we employ a spherical retina. Let a sphere with radius f and center O (Figure 7) represent the spherical retina (with O the nodal point of the eye) and a coordinate system $OXYZ$ attached to it.

Let

$$\vec{r}_\omega = (X, Y, Z) \text{ be a world point}$$

and

$$\vec{r} = (x, y, z) \text{ be its image on the image plane.}$$

Then

$$\frac{\vec{r}}{f} = \frac{\vec{r}_\omega}{R}, \quad R = \|\vec{r}_\omega\| = \sqrt{\vec{r}_\omega \cdot \vec{r}_\omega}$$

It can easily be shown (Koenderink and van Doorn 1975; Maybank 1985) that

$$\dot{\vec{r}} = -\frac{\vec{t}f}{R} - \vec{\omega} \times \vec{r} + \frac{\vec{r}}{Rf}(\vec{t} \cdot \vec{r}) \qquad (1)$$

$$= \frac{1}{R}\left[-\vec{t}f + \frac{\vec{r}}{f}(\vec{t} \cdot \vec{r})\right] - \vec{\omega} \times \vec{r}$$

Thus, the translational flow is

$$\vec{u}_t = \frac{1}{R}\left[-\vec{t}f + \frac{\vec{r}(\vec{t} \cdot \vec{r})}{f}\right]$$

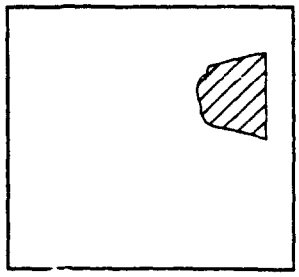
while the rotational flow is given by

$$\vec{u}_R = -\vec{\omega} \times \vec{r}.$$

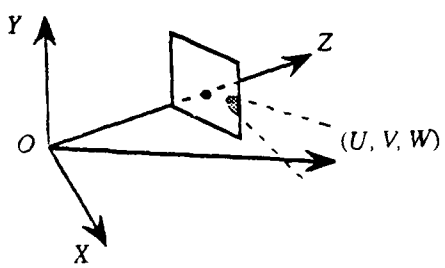
Without loss of generality we can set $f = 1$.

At this point we define two quantities that will be of use later. They are $\tau = \frac{R}{\|\vec{t}\|}$, which is related to the time to collision, and $k = \frac{\|\vec{\omega}\|}{\|\vec{t}\|}R = \|\vec{\omega}\|\tau$, which represents the effective ratio of rotation and translation.

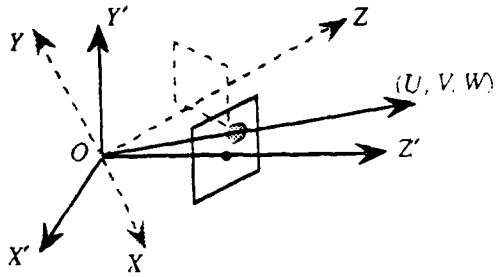
The geometry of the spherical projection is then given in Figure 8. It has been shown (Nelson and Aloimonos 1988) that a full (360°) visual field simplifies motion analysis. However, what we usually have is just a piece of the surface of the sphere (due to a limited field of view). Consider then that the image (the part that we see) is projected on the surface patch S. Obviously, voting for the estimation of the FOE can be performed for all points on S.



(a)



(b)



(c)

Fig. 6. (a) If the area containing the highest number of votes has a piece of the image boundary as part of its boundary, then the FOE is outside the image plane (see 6b). (b) The position of the area containing the highest number of votes indicates the general direction in which the translation vector lies. (c) The camera ("eye") rotates so that the area containing the highest number of votes becomes centered. With a rotation around the x and y axes only, the optical axis can be positioned anywhere in space. The process stops when the highest vote area is entirely inside the image.

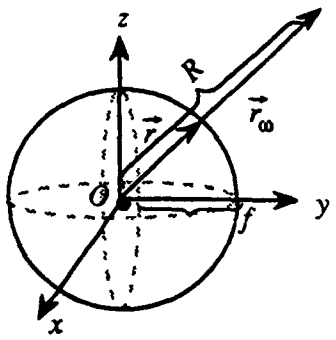


Fig. 7.

5.1 Principles of Voting

Consider

$\vec{r}_i = (x, y, z)$, a point in S,

$\vec{n}_i = (n_x, n_y, n_z)$, the image gradient direction at point $\vec{r}_i$,

$\dot{\vec{r}}_i = \vec{u}_i = (u_x, u_y, u_z)$, the flow at point $\vec{r}_i$, and

$\vec{u}_i^n = (\vec{n}_i \cdot \vec{u}_i) \cdot \vec{n}_i$, the normal flow at $\vec{r}_i$.

Then (see Figure 9) if $\vec{r} = (x, y, z)$ is a point in S, a feature point $\vec{r}_i$ will vote for $\vec{r}$ being the FOE (direction of translation) iff $\vec{u}_i^n(\vec{r} - \vec{r}_i) < 0$ (see Figure 9).

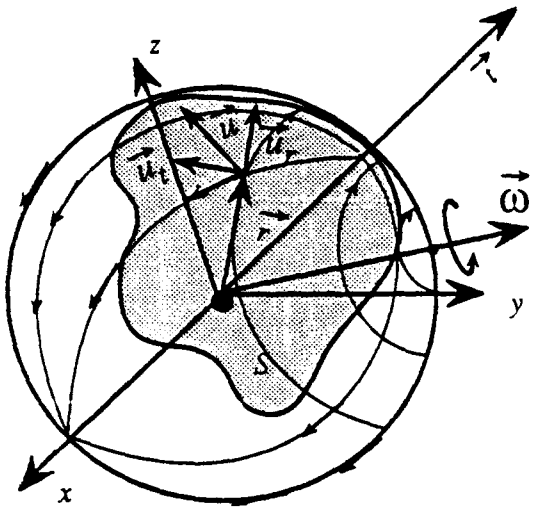


Fig. 8.

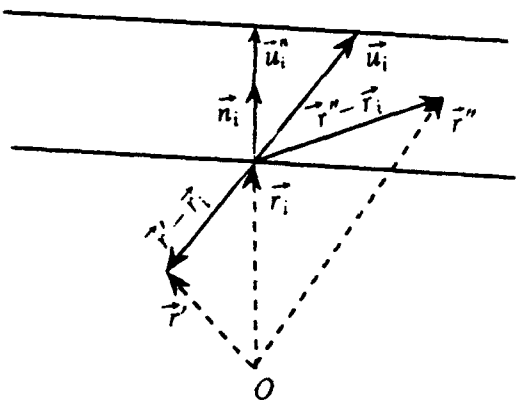


Fig. 9.

If $V[\vec{r}]$ represents the number of votes collected at point $\vec{r}$, then it is easy to see that

$$V[\vec{r}] = \sum_{\vec{r}_i \in S} U\left[\vec{u}_i^n(\vec{r}_i - \vec{r})\right]$$

where

$$U(x) = \begin{cases} 1, & x > 0 \\ 0, & x \leq 0 \end{cases}$$

Let $S' = \left\{\vec{r} | \forall \vec{r}' \in S,\ V[\vec{r}] \geq V[\vec{r}']\right\}$ be the set of points that have acquired the maximum number of votes. There are two cases:

Case 1: S' does not intersect the border of S, in which case the FOE is in S'.

Case 2: S' touches the border of S, in which case the FOE could be outside of S.

It should be clear that if there is no rotation, then S' will always contain the FOE or give the direction of the FOE—i.e. the direction towards which we need to rotate. The size of S' depends on the distribution of features.

In the sequel we investigate the performance of the voting scheme in the presence of rotation. In particular we ask how large area S is when rotation is present. It will be shown that this depends on the angle θ_ω between the direction of translation and the axis of rotation as well as on the rotation-to-translation ratio k. In particular, θ_ω distorts area S' and k enlarges it as it grows. The rest of the paper quantifies this interaction.

Before we proceed with the analysis we introduce a natural coordinate system that greatly simplifies the calculations.

5.2 A Natural Coordinate System

Since spherical projection is symmetrical we can choose a coordinate system that facilitates our analysis. We define a new orthonormal coordinate system with unit vectors $\vec{\sigma}_x = \vec{i}, \vec{\sigma}_y = \vec{j}, \vec{\sigma}_z = \vec{k}$ (defined by Figure 10):

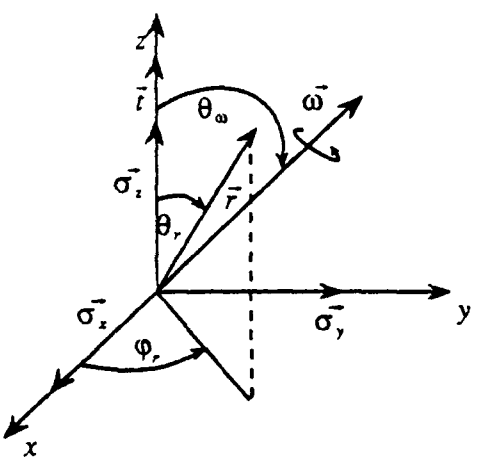


Fig. 10.

$$\vec{\sigma}_z = \frac{\vec{t}}{\|\vec{t}\|};$$

$$\vec{\sigma}_x = \begin{cases} \frac{\vec{\omega}\times\vec{t}}{\|\vec{\omega}\times\vec{t}\|} \text{ if } \vec{\omega}\times\vec{t} \neq 0 \\ \text{any unit vector such that} \\ \qquad \sigma_x \cdot \sigma_z = 0 \text{ otherwise.} \end{cases}$$

$$\vec{\sigma}_y = \vec{\sigma}_z \times \vec{\sigma}_x.$$

In spherical coordinates we get

$$\vec{t} = \|\vec{t}\|(0, 0, 1)$$
$$\vec{\omega} = \|\vec{\omega}\|(0, \sin\theta_\omega, \cos\theta_\omega)$$
$$\vec{r} = (\cos\varphi_r \sin\theta_r, \sin\varphi_r \sin\theta_r, \cos\theta_r)$$

So,

$$\vec{t} = (0, 0, W) := W\vec{k} \text{ for some } W$$
$$\vec{\omega} = (0, B, C) := B\vec{j} + C\vec{k} \text{ for some } B \text{ and } C.$$

Similarly, we define a coordinate system (ξ, η) which lies on the plane tangent to the sphere at point $\vec{r} = (x, y, z)$. This tangent plane is spanned by the vectors

$$\vec{n}_\xi = \frac{\vec{u}_t}{\|\vec{u}_t\|} \text{ and } \vec{n}_\eta = \frac{\vec{k}\times\vec{r}}{\|\vec{k}\times\vec{r}\|} \text{ with } \vec{n}_\xi \cdot \vec{n}_\eta = 0.$$

Any flow vector lies on the tangent plane; therefore it will be a linear combination of vectors $\vec{n}_\xi$ and $\vec{n}_\eta$.[7]

Now we are ready to express normal flow in the new coordinate system. Consider a feature with gradient direction

$$\vec{n} = \vec{n}_\xi \cos\alpha + \vec{n}_\eta \sin\alpha$$

The translational normal flow is

$$\vec{n} \cdot \vec{u}_t = \left(\frac{\vec{u}_t}{\|\vec{u}_t\|}\cos\alpha + \vec{n}_\eta \sin\alpha\right)\vec{u}_t =$$
$$= \frac{\|\vec{u}_t\|^2}{\|\vec{u}_t\|}\cos\alpha + \vec{n}_\eta \cdot \vec{u}_t \sin\alpha = \|\vec{u}_t\|\cos\alpha.$$

Also,

$$\|\vec{u}_t\| = \left[\frac{\|t\|^2}{R^2} + \frac{(\vec{t}\cdot\vec{r})^2}{R^2}(\vec{r}\cdot\vec{r}) - \frac{2(\vec{r}\cdot\vec{t})^2}{R^2}\right]^{\frac{1}{2}}$$
$$= \frac{\|t\|}{R}\sin\theta_r$$

and

$$\vec{n} \cdot \vec{u}_t = \frac{\|t\|}{R}\sin\theta_r \cos\alpha$$

For rotational normal flow we get

$$\vec{n} \cdot \vec{u}_R = -(\vec{\omega}\times\vec{r})\vec{n}_\xi \cos\alpha - (\vec{\omega}\times\vec{r})\vec{n}_\eta \sin\alpha$$

Also,

$$\begin{aligned} -(\vec{\omega}\times\vec{r})\vec{n}_\xi &= -\frac{(\vec{\omega}\times\vec{r})\vec{u}_t}{\|\vec{u}_t\|} \\ &= -\frac{(\vec{\omega}\times\vec{r})}{\|\vec{u}_t\|}\left[-\frac{\vec{t}}{R} + \frac{\vec{r}(\vec{t}\cdot\vec{r})}{R}\right] \\ &= \frac{(\vec{\omega}\times\vec{r})\vec{t}}{\|\vec{u}_t\|}\cdot\frac{1}{R} = -\|\vec{\omega}\|\sin\theta_\omega\cos\varphi_r \end{aligned}$$

and

$$-(\vec{\omega}\times\vec{r})\vec{n}_\eta$$
$$= -\frac{(\vec{\omega}\times\vec{r})(\vec{k}\times\vec{r})}{\|\vec{k}\times\vec{r}\|}$$
$$= \frac{-\sin\theta_r\|\vec{\omega}\|[\cos\theta_\omega\sin\theta_r - \sin\theta_\omega\cos\theta_r\sin\varphi_r]}{\sin\theta_r}$$

so that

$$\vec{n} \cdot \vec{u}_R = \|\vec{\omega}\|[-\cos\theta_\omega\sin\theta_r\sin\alpha$$
$$+ \sin\theta_\omega\cos\theta_r\sin\varphi_r\sin\alpha - \sin\theta_\omega\cos\varphi_r\cos\alpha]$$

5.3 Correctness of Voting in the Presence of Rotation

The normal flow (as well as actual flow) is very small in the region close to the FOE, and in the directions close to orthogonal to the directions of the flow. Consequently, even when only translation is present, in order to avoid inaccuracies that might arise in the estimated direction of the normal flow—numerical manipulation of very small quantities is unstable—we are going to discard any normal flow whose magnitude is less than some threshold T_t. Later, it will turn out that choosing this threshold greatly facilitates the geometrical analysis of the technique. Considering an actual flow $\vec{u}$ at a point A (see Figure 11) we can compute the locus of gradient directions $\vec{n}$ along which the normal flow (i.e. the projection of $\vec{u}$ on $\vec{n}$) is bigger than the threshold T_t. In Figure 11 they are all directions inside angle BAC defined by

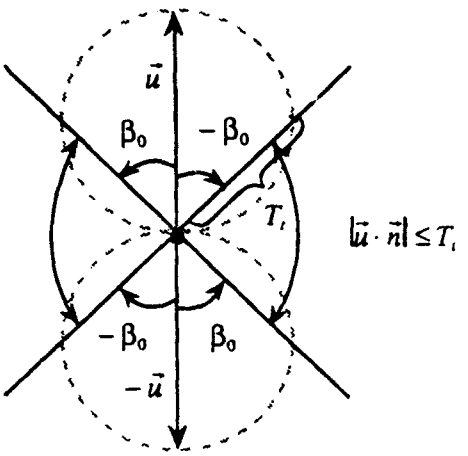


Fig. 11.

$\beta_0 = \arccos \frac{T_t}{\|\vec{u}\|}$ for $\frac{T_t}{\|\vec{u}\|} \le 1$, or there are no such directions for $\frac{T_t}{\|\vec{u}\|} > 1$.

We now develop a condition that needs to be satisfied in order for voting at a point to be correct in the presence of rotation.

Voting will clearly be correct only if the direction of the translational normal flow is the same as the direction of the actual normal flow, that is when

$$(\vec{n}\cdot\vec{u}_t)(\vec{n}\cdot\vec{u}) > 0 \tag{2}$$

In addition, since we consider only normal flows greater than threshold, we need

$$|\vec{n}\cdot\vec{u}| > T_t \tag{3}$$

Inequality (2) becomes

$$\begin{aligned}(\vec{n}\cdot\vec{u}_t)(\vec{n}\cdot\vec{u}) &= (\vec{n}\cdot\vec{u}_t)(\vec{n}\cdot\vec{u}_t + \vec{n}\cdot\vec{u}_R) = \\ &= (\vec{n}\cdot\vec{u}_t)^2 + (\vec{n}\cdot\vec{u}_t)(\vec{n}\cdot\vec{u}_R) > 0\end{aligned} \tag{4}$$

So, if we set $|\vec{n}\cdot\vec{u}_R| = T_t$, then there are two possibilities: either $|\vec{n}\cdot\vec{u}|$ is below the threshold, in which case it is of no interest to voting, or the sign of $\vec{n}\cdot\vec{u}$ is the same as the sign of $\vec{n}\cdot\vec{u}_t$. In other words, if we can set the threshold equal to the maximum value of the normal rotational flow, *then our voting will always be correct*. But at point $\vec{r}$ of the sphere the rotational flow is

$$\begin{aligned}|\vec{n}\cdot\vec{u}_R| &\le \|\vec{n}\|\cdot\|\vec{u}_R\| = \|\vec{u}_R\| = \|\vec{\omega}\times\vec{r}\| = \\ &= \|\vec{\omega}\|\cdot\|\vec{r}\|\cdot|\sin(\angle\vec{\omega},\vec{r})| \le \|\vec{\omega}\|\end{aligned}$$

Thus if we choose $T_t = \|\vec{\omega}\|$, then the sign of $\vec{n}\cdot\vec{u}$ (actual normal flow) is equal to the sign of $\vec{u}_t\cdot\vec{n}$ (translational normal flow) for any normal flow of magnitude greater than T_t.

5.4 The Geometry of the Solution Area

The introduction of the threshold T_t into our analysis has a beneficial side-effect, since this constrains the possible gradient directions at every point where we can vote. As a consequence we can estimate the size of the smallest possible area that we might find as a solution. We need to caution the reader that in the case of pure translation the solution area (the area containing the FOE) contains the uncertainty area (the area where the values of the normal flow do not allow voting to be performed). However, when rotation is present then the solution area, in general, does not contain the uncertainty area. The reason for this is that points far away from the position of the FOE might constrain the solution area more than the uncertainty area does. The following two sections quantify this analysis.

The Case of Translation Depending on the threshold T_t, there is a point closest to the FOE for which a feature (normal flow) can be registered. It is obtained when

$$\frac{\|\vec{r}\|}{R}|\sin\theta_r\cos\alpha| = T_t \tag{5}$$

It is obvious that for θ_r smaller than some threshold θ_{r_0} (5) never holds; for $\theta_r = \theta_{r_0}$, (5) holds only for $\alpha = 0$, when θ_r grows this cone of directions for $\vec{n}$ (gradient) or range of α's grows and for $\theta_r = \frac{\pi}{2}$ it reaches $|\alpha| \in [0, \frac{\pi}{2} - \theta_{r_0}]$ if the R's are the same at both points.

This increasing range of α's for increasing θ_r can be viewed as an increasing density of features (features are registered only for α's in the cone). But it is easy to show that feature points $\vec{r}_i$ with increasing θ will vote for some points with $\theta > \theta_{r_i}$ and will vote for some points with $\theta < \theta_{r_i}$. This happens when for $\alpha > 0$ the FOE constraint line gets slanted. The effect is shown in Figure 12.

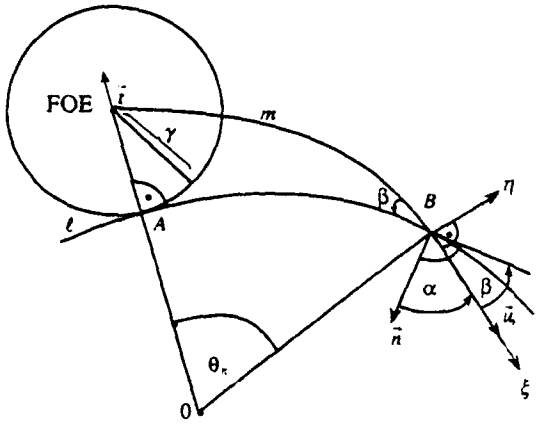


Fig. 12.

Using spherical trigonometry (the law of sines) (Korn and Korn 1968) for triangle AB (FOE), we get

$$\sin\gamma = \sin\beta \sin\theta_{r_i} \tag{6}$$

Since $\vec{n}$ is normal to great circle ℓ, we have $\alpha + \beta = \frac{\pi}{2}$ and (6) becomes

$$\sin\gamma = \cos\alpha \sin\theta_{r_i} \tag{7}$$

The normal translational flow is given by

$$\vec{u}_t \cdot \vec{n} = \frac{\|\vec{t}\|}{R} \sin\theta_r \cos\alpha$$

And so, setting $|\vec{u}_t \cdot \vec{n}| = T_t$ and using (7), we find the most restrictive voting, i.e. the smallest possible γ:

$$\sin\gamma = \frac{T_t \cdot R}{\|\vec{t}\|} = \sin\theta_{r_0} \tag{8}$$

We prove here that the voting function $V[\vec{r}]$ (introduced in Section 5.1) is non-decreasing on any great circle, as we move from the south to the north pole where the FOE is assumed to be.[8] To remind the reader, voting at a feature point $\vec{r}_i$ increases by one the votes of every point in the northern hemisphere[9] (see Figure 13), defined by the great circle normal to the gradient at the feature point $\vec{r}_i$. Consider a great circle SA(FOE). All points on the arc SA receive zero votes, while each point on A (FOE) receives one vote. Consequently, since each voting process can only increase the votes, the number of votes is non-decreasing as we move closer to the FOE. This simply means that the solution area (i.e. S') on the sphere will always be closed and contain the FOE. Its size, however, could be large if the distribution of features is not favorable. In addition, if voting is done only on a surface patch S (a limited visual field), S' could be open.

The Case of Nonzero Rotation In the presence of rotation, voting will still yield a closed area on the sphere which can in general be larger. Here we study properties of the shape and size of the solution area. Due to the use of threshold $T_t = \|\vec{\omega}\|$, points in an area around the FOE will not be used for voting. The size and shape of the solution area will depend on the angle θ_ω between $\vec{t}$ and $\vec{\omega}$ and the threshold $T_t = \|\vec{\omega}\|$. We first consider the case where $\vec{t}$ and $\vec{\omega}$ are parallel. Then the normal flow is given by (see Section 5.2)

$$\vec{n} \cdot \vec{u} = \sin\theta_r \left[\frac{\|\vec{t}\|}{R} \cos\alpha - \|\vec{\omega}\| \sin\alpha \right]$$

Now we find the angular distance θ_{r_0} between the FOE and the closest point to it on the sphere that can vote, i.e. the closest point with normal flow equal to $T_t = \|\vec{\omega}\|$. It is clear that the point closest to the FOE that can vote is the one at which the maximum possible normal flow is equal to T_t. Maximum normal flow is obtained when the direction of the gradient is the same as the direction of the actual flow. That

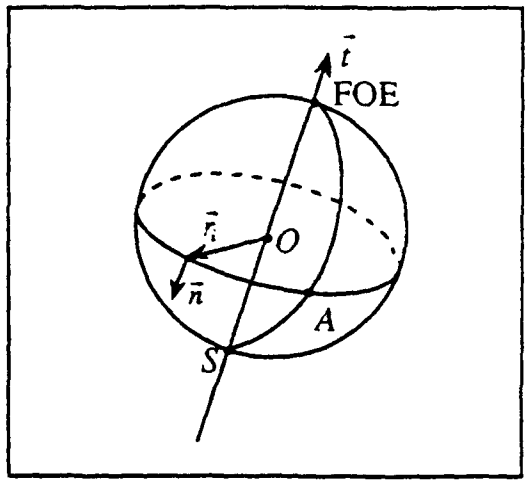


Fig. 13.

happens for the angle α for which

$$\vec{n} \cdot \vec{u} = \sin\theta_r \left[\frac{\|\vec{t}\|}{R} \cos\alpha - \|\vec{\omega}\| \sin\alpha \right]$$

is maximized, i.e. for

$$\alpha_0 = -\arctan\left(\frac{\|\vec{\omega}\|}{\|\vec{t}\|} R \right)$$

Thus, the angular distance θ_{r_0} obeys

$$\vec{n} \cdot \vec{u} = \sin\theta_{r_0} \left[\frac{\|\vec{t}\|}{R} \cos\alpha_0 - \|\vec{\omega}\| \sin\alpha_0 \right] = \|\vec{\omega}\|$$

or

$$\tan\theta_{r_0} = \frac{\|\vec{\omega}\|}{\|\vec{t}\|} R.$$

As was said before, if the uncertainty area is not contained in the image, then the solution area will not be closed; otherwise things depend on the distribution of features, and any feature point might further constrain the solution area more, as was shown in Figure 12.

The rest of this section describes various properties of the solution area as the relevant parameters vary.

If $\theta_r = \theta_{r_0}$ then $\alpha = \theta_{r_0}$ (Figure 12) and the constraint line (circle) will come as close as γ to the FOE where

$$\begin{aligned} \sin\gamma &= \sin\theta_{r_0} \sin\beta = \\ &= \sin\theta_{r_0} \sin(\tfrac{\pi}{2} - \alpha) = \tfrac{1}{2} \sin 2\theta_{r_0} \end{aligned}$$

If, however, $\alpha \neq \theta_{r_0}$ (we might even have $\alpha = \frac{\pi}{2}$), as the cone of normal flows around the flow grows (Figure 11), this constrains the area around the north pole (FOE) even more. Unlike the case of translation only, voting further restricts the FOE when feature points are moving away from it and the area for which voting is maximum becomes smaller than the uncertainty area. On the other hand, if $\vec{\omega} \times t \neq \vec{0}$ but $\vec{\omega} = (0, B, C)$, things become unsymmetrical around the FOE. Using the already-defined coordinate system $Oxyz$, $\vec{\omega}$ is defined by $\|\vec{\omega}\|$, $\varphi_\omega = \frac{\pi}{2}$ (the analog of φ_r—see Figure 10), and θ_ω.

When the angle θ_ω becomes greater than zero and acquires a small value, a subtle change in the uncertainty area occurs. The flow values at the points on the border of that area increase or decrease depending on their positions. Since flow is continuous in $\vec{\omega}$, the point or points for which $\|\vec{u}\| = \|\vec{\omega}\|$ will stay close to the border. The effect is that the borderline $\|\vec{u}\| = \|\vec{\omega}\|$ will change its shape in the same way. If the flow increases, the border of the uncertainty area will shrink closer to the north pole (FOE) and if flow decreases, it will stretch the border away from the north pole

It happens that with growth of θ_ω and $\|\vec{\omega}\|$ this area stretches away in direction $\vec{\omega} \times \vec{t}$ and it shrinks in the opposite direction. The exact shape of the area for a given $\vec{\omega}, \vec{t}$ and $R = R(\varphi, \theta)$ can be computed numerically (the border is defined by $\|\vec{u}\| = \|\vec{u}_t + \vec{u}_R\| = \|\vec{\omega}\|$). The effect of the change in shape of the area is as if the FOE moved in the $x - z$ plane ($\varphi_r = 0$). Again, if the area were not completely closed (was intersected by the image patch S), the solution area (with maximum voting) would not be closed. If the area were in the image, feature points outside it would further constrain the area which contains the FOE because of the slant of the FOE constraint lines in each feature point. Figure 14 demonstrates graphically the evolution of the uncertainty area for different values of θ_ω and k. Each figure is produced by projecting every point of the northern hemisphere on the plane tangent to the north pole (FOE) with the south pole as the center of projection (stereographic projection). Figure 14(a) shows the evolution of the uncertainty area for $k = 0.25$ and $\theta_\omega = 0, \frac{\pi}{6}, \frac{\pi}{3}, \frac{\pi}{2}$, with the uncertainty area centered for $\theta_\omega = 0$, and completely offset for $\theta_\omega = \frac{\pi}{2}$ (with all other values in between). Similarly, Figures 14(b) and 14(c) show the same results for $k = 0.5$ and 0.75, respectively.

The Case of Dominant Rotation Although the technique described in this paper was derived to solve the problem of kinetic stabilization it turns out that it has general applicability. It can be modified to handle the case of dominant rotation with translation.

For the case of pure rotation and a spherical retina the optical flow will correspond to vectors tangent to the circles around the axis of rotation $\vec{\omega}$. The point at which the axis of rotation passes through the image will be called the AOR. If there is circular optical flow in the image (due to pure rotation) the center of all the circles is the AOR. If we take an arbitrary optical flow vector $\vec{u}_R$ at the point $\vec{r}_i$ then we can say that a point $\vec{r}$ is a candidate for the AOR if

$$(\vec{r}_i \times \vec{u}_R)\vec{r} < 0.$$

This inequality expresses the fact that the feature point and the flow vector at the point span the plane p

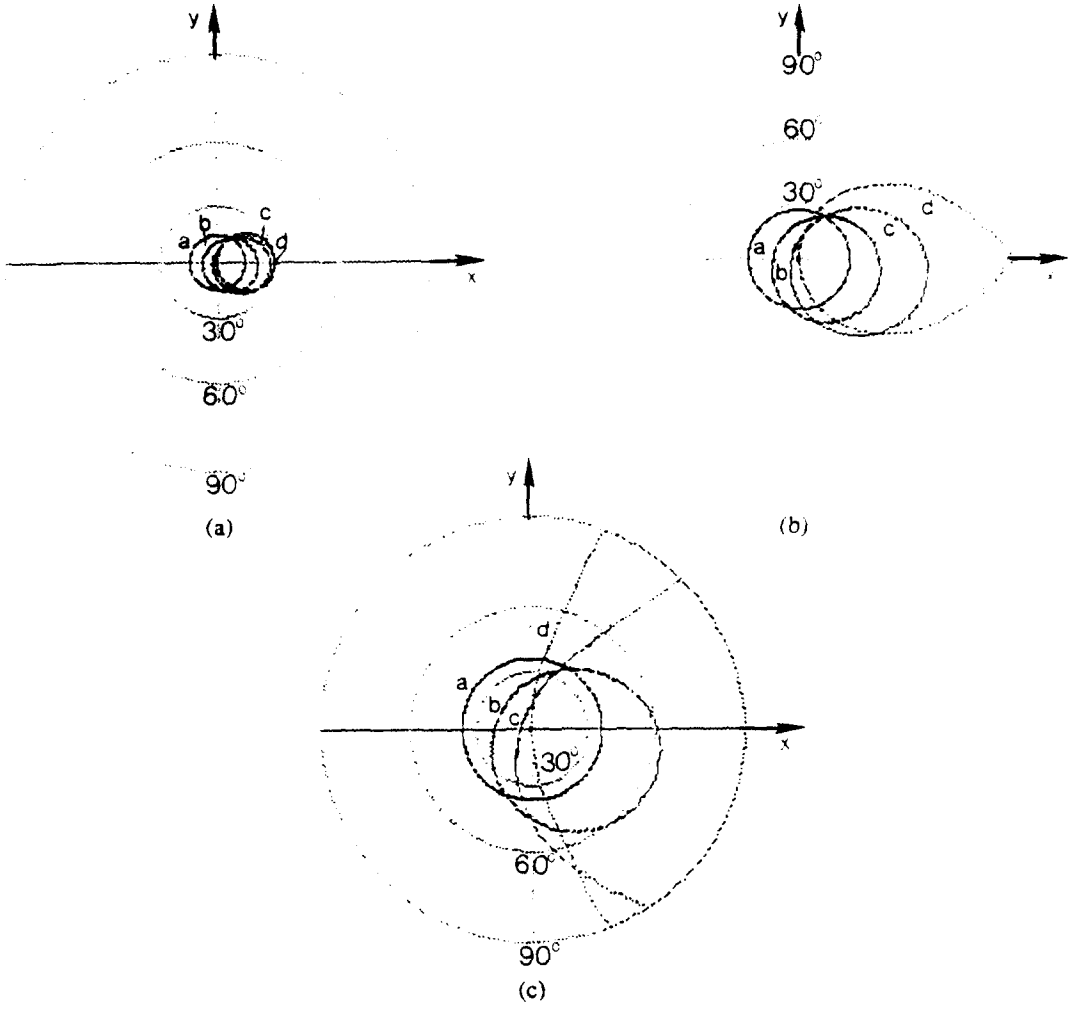


Fig. 14. (a) The evolution of the uncertainty area for $k = 0.25$ and $\theta_\omega = 0, \frac{\pi}{6}, \frac{\pi}{3}, \frac{\pi}{2}$, with the uncertainty area centered for $\theta_\omega = 0$, and completely offset for $\theta_\omega = \frac{\pi}{2}$ (with all other values in between). (b) The same results for $k = 0.5$. (c) The same results for $k = 0.75$.

which cuts the sphere in two hemispheres where one contains all possible candidate points for the AOR (and all of them satisfy the previous inequality). Furthermore, all possible positions of the AOR lie on the great circle which is normal (on the sphere) to the great circle which is the intersection of the plane p and the image sphere. In other words if we replace $\vec{u}_R$ with the normal flow $\vec{u}_R^n$ the inequality will still hold.

Very similar reasoning applies in the case of a flat retina (perspective projection). Given an optical flow (u, v) at the feature point (x_i, y_i) all possible candidate points for the AOR are on the right of the line passing through (x_i, y_i) and parallel to (u, v). Furthermore, they all lie on the line normal to (u, v) and originating at (x_i, y_i). In other words candidate points (x, y) for the AOR satisfy the inequality

$$((u, v, 0) \times (x - x_i, y - y_i, 0))(0, 0, 1) < 0.$$

This inequality indicates that the z component of the vector product of the optical flow vector and the difference of the candidate AOR point and the feature point must be negative. As in the case of a spherical retina this holds even when the optical flow (u, v) is replaced by the normal flow (u^n, v^n). As was done

in the case of translation, voting can be performed. Points with maximum votes are candidates for the AOR. If a minimum is sought then the opposite direction will be found. If the area is closed then the AOR is localized as before; otherwise its general direction will be indicated by the area with maximum votes.

An analysis (on a spherical retina) similar to the one performed for the case of dominant translation can be performed again. This time, however, the threshold should be set to $T_t = \tau = \frac{R}{\|\vec{t}\|}$. If the magnitude of the normal flow is greater then T_t then it must have the same sign (and direction) as rotational normal flow.

When $\vec{\omega}$ and $\vec{t}$ are parallel the angular radius of the uncertainty region is equal to θ_{r_0} where $\cot\theta_{r_0} = \frac{\|\vec{\omega}\|}{\|\vec{t}\|}R$. The difference in the angular radii of the uncertainty areas around the FOE and the AOR is that the tangent is replaced by the cotangent. When $\theta_\omega > 0$ the uncertainty area around the AOR changes shape in a similar manner as the uncertainty area around the FOE. It extends in the direction $\vec{\omega} \times \vec{t}$ with the growth of θ_ω and gets closer to the AOR in the opposite direction.

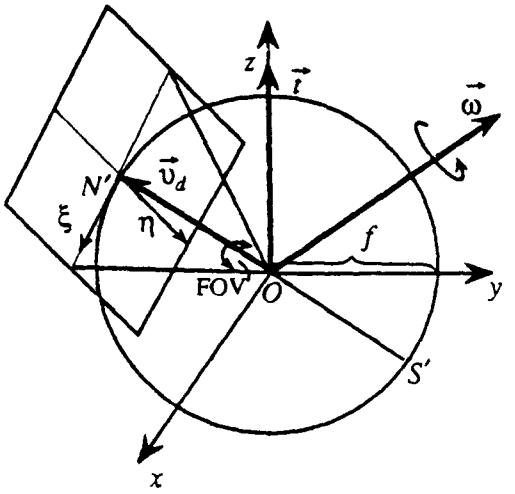


Fig. 15. Sphere $OXYZ$ represents a spherical retina (frame $OXYZ$ is the frame of the observer). The translation vector $\vec{t}$ is along the z axis and the rotation axis lies on the plane OZY. Although a spherical retina is used here, information is used only from a patch of the sphere defined by the solid angle FOV containing the viewing direction $\vec{v}_d$ (defined by the two angles θ and φ—see Figure 10). The spherical image patch is projected stereographically with center S' on the plane P tangent to the sphere at N', and having a natural coordinate system (ξ, η). All results (solution areas, voting functions, actual and normal flow fields) are projected and shown on the tangential plane.

6 Experimental Results

We have performed several experiments with both synthetic and real image sequences in order to demonstrate the stability of the method. From experiments on real images it was found that in the case of pure translation or pure rotation the method computes the Focus of Expansion or the Axis of Rotation very accurately. In the case of general motion it was found from experiments on synthetic data that the behavior of the method is as predicted by our theoretical analysis.

6.1 Synthetic Data

We considered a set of features at random depths (uniformly distributed in a range $R_{\min}$ to $R_{\max}$). The scene was imaged using a spherical retina as in Figure 15. Optic flow and normal optic flow were computed on the sphere and then projected onto the tangent plane (see Figure 15). Normal flow was computed by considering features whose orientations were produced using a uniform distribution.

Figures 16 to 20 show one set of experiments. Figure 16 shows the optic flow field for $\theta_\omega = 0°$, viewing angles $(\theta, \varphi) = (0°, 0°)$, $R_{\min} = 10$ and $R_{\max} = 20$ in units of focal length, $\|\vec{t}\| = 1$, $k = \frac{\|\vec{\omega}\|}{\|\vec{t}\|} \cdot \frac{R_{\min}+R_{\max}}{2} = 0.1$ and $FOV = 56°$. Figure 17 shows the corresponding normal flow. Similarly Figures 18, 19 show optical and normal flow fields for the same conditions as before with the exception that $k = 0.75$ which is obtained by growing $\|\vec{\omega}\|$. Under the above viewing conditions, the FOE as well as AOR is in the center of the image. Figure 20 shows results of voting for determining the FOE. In the first column thresholding precedes voting, with $T_t = \|\vec{\omega}\| f$, and in the second column there is no thresholding. In the first row, only the area with the maximum number of votes is shown, while in the second row the whole voting function is displayed (black is maximum). Clearly, the solution is a closed area (except for the biggest k) whose size grows with k.

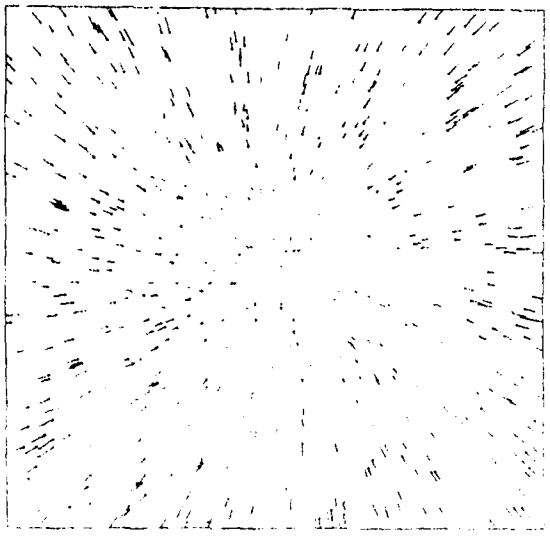


Fig. 16.

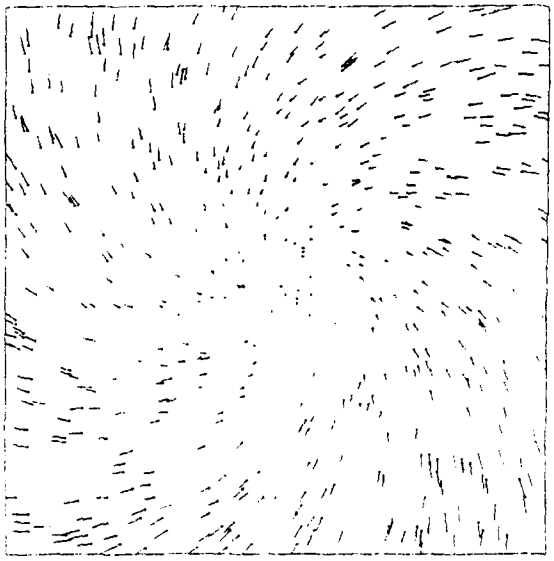


Fig. 18.

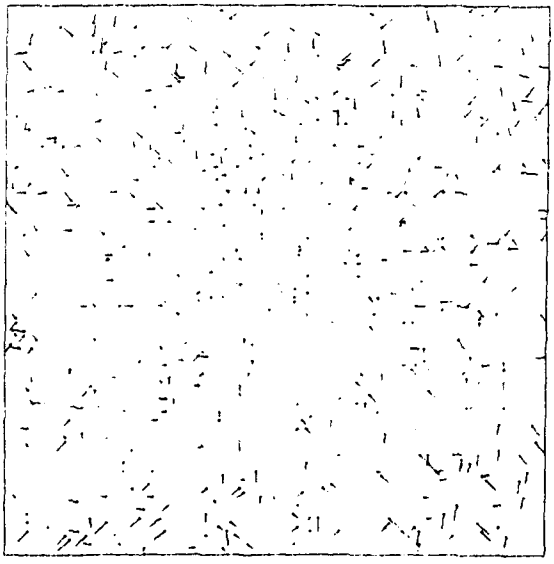


Fig. 17.

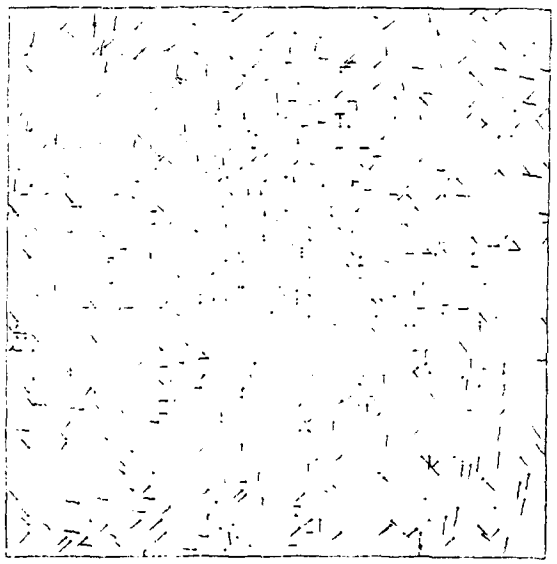


Fig. 19.

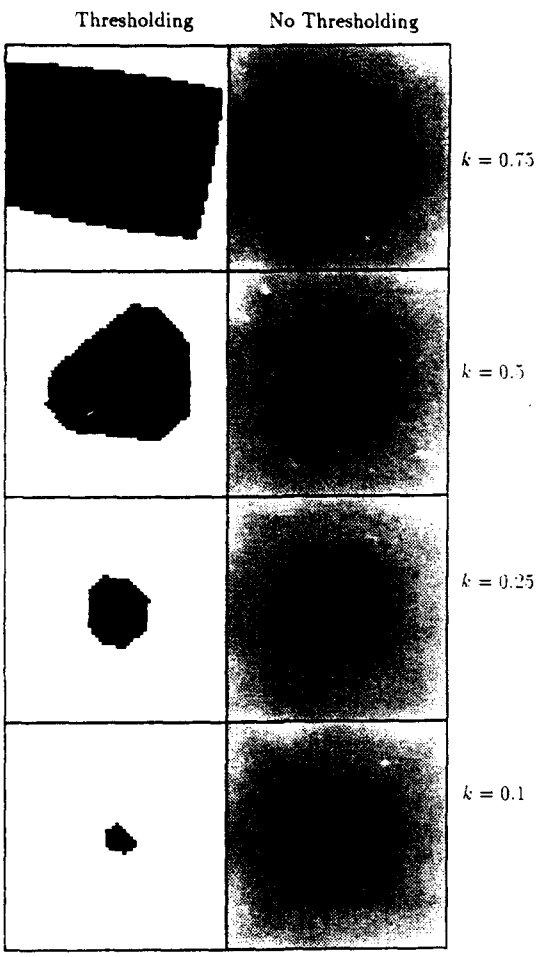


Fig. 20.

Figures 21 to 25 show the second set of experiments. The only change from the first set is that $FOV = 106°$. The FOE is in the center of the image. The uncertainty is quite small due to bigger field of view. This was predicted in our analysis.

Finally, Figures 26 to 30 show the third set of experiments. The change from the first set is that $\theta_\omega = 45°$. As was predicted the solution area gets distorted and for bigger $\|\vec{\omega}\|$ it becomes open. In case when there is no thresholding before voting (second row) this appears as shift in the estimated position of the FOE.

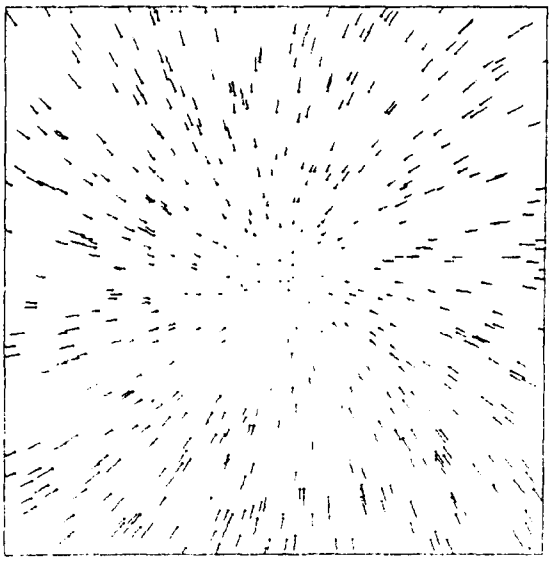


Fig. 21.

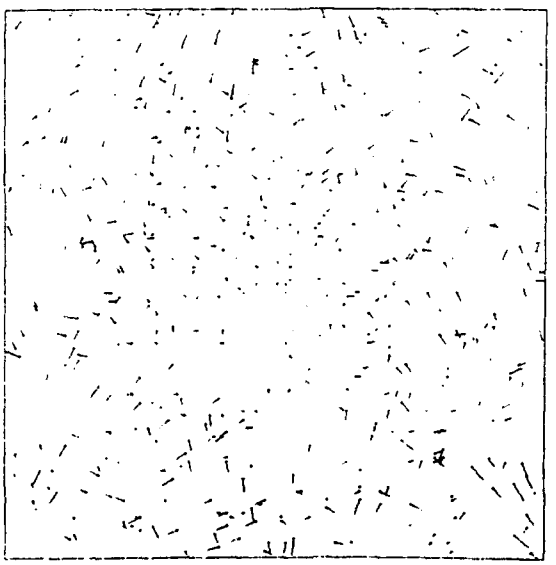


Fig. 22.

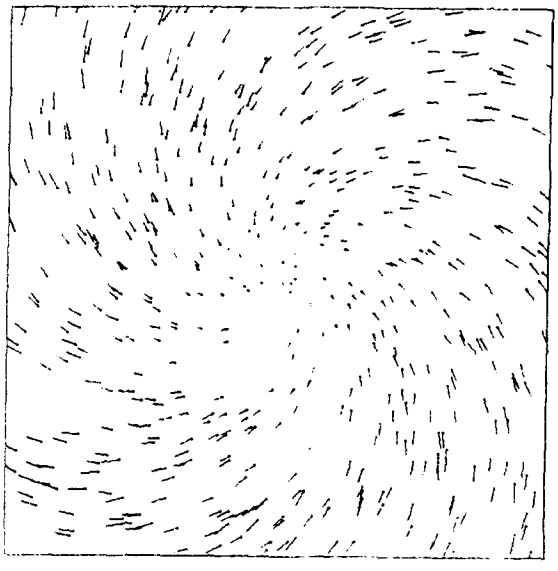


Fig. 23.

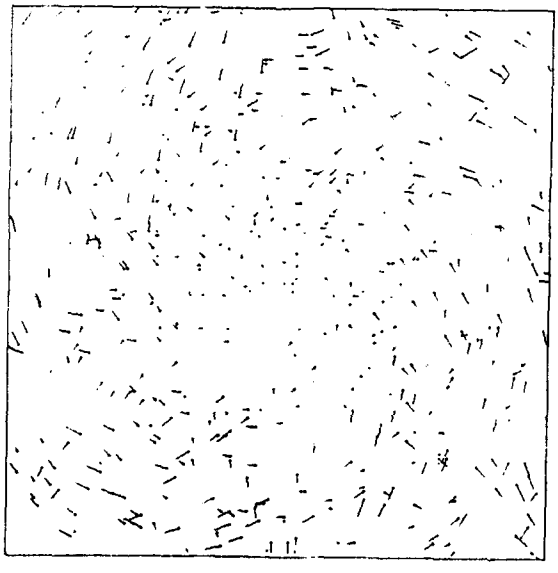


Fig. 24.

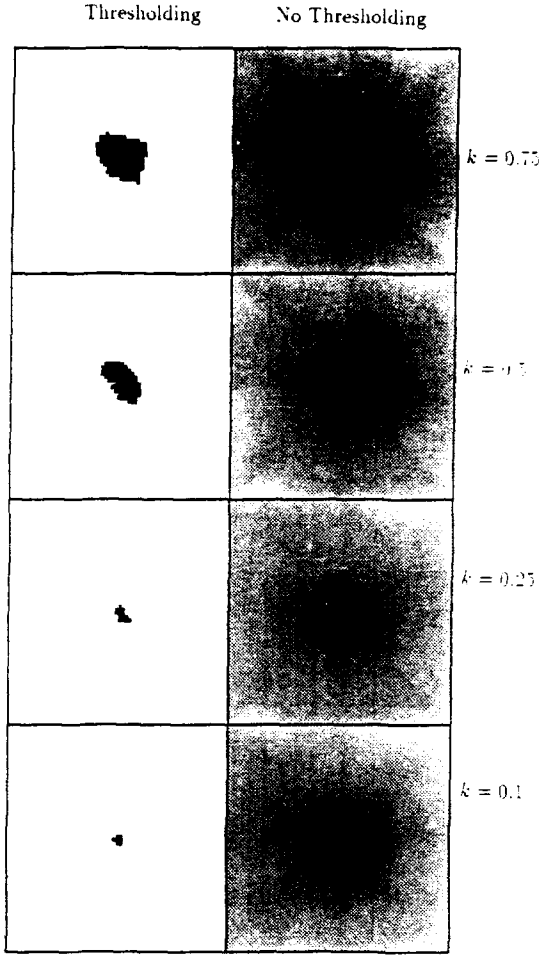


Fig. 25.

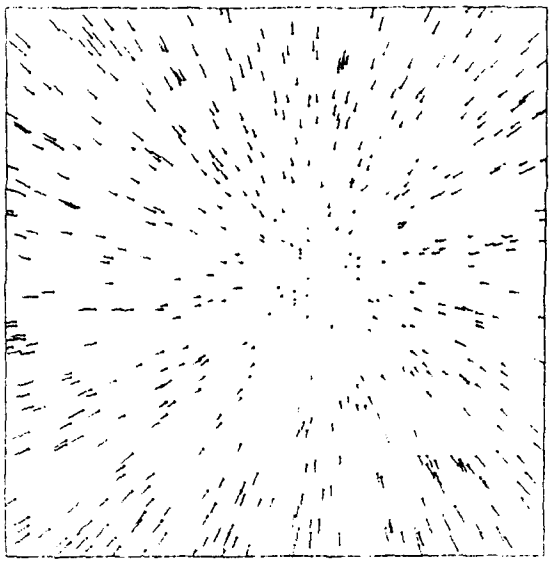


Fig. 26.

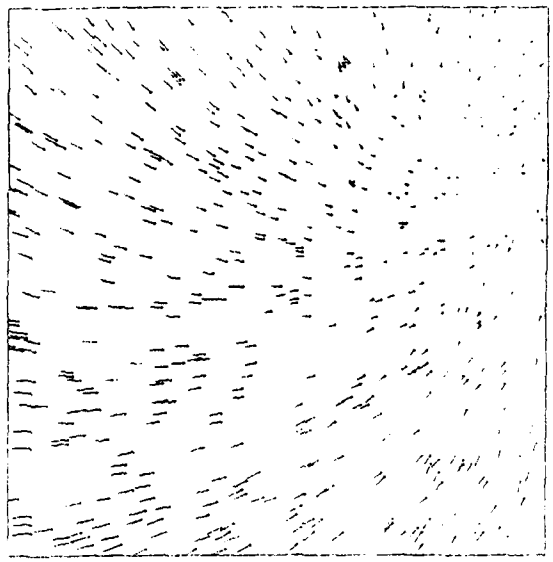


Fig. 28.

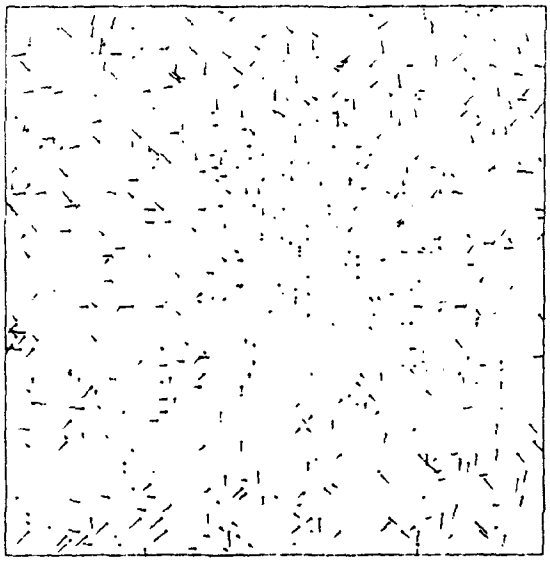


Fig. 27.

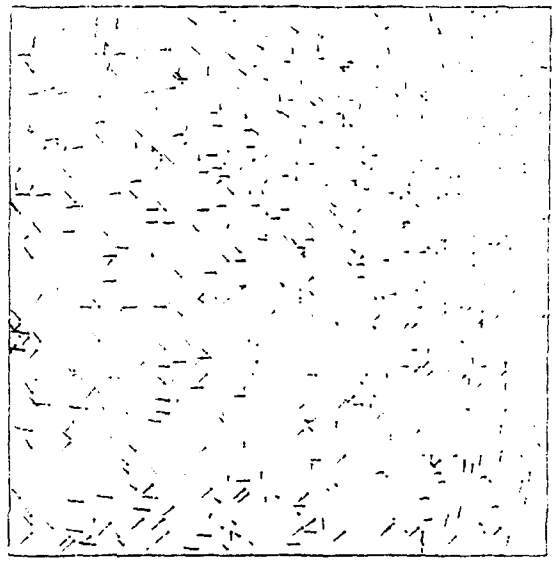


Fig. 29.

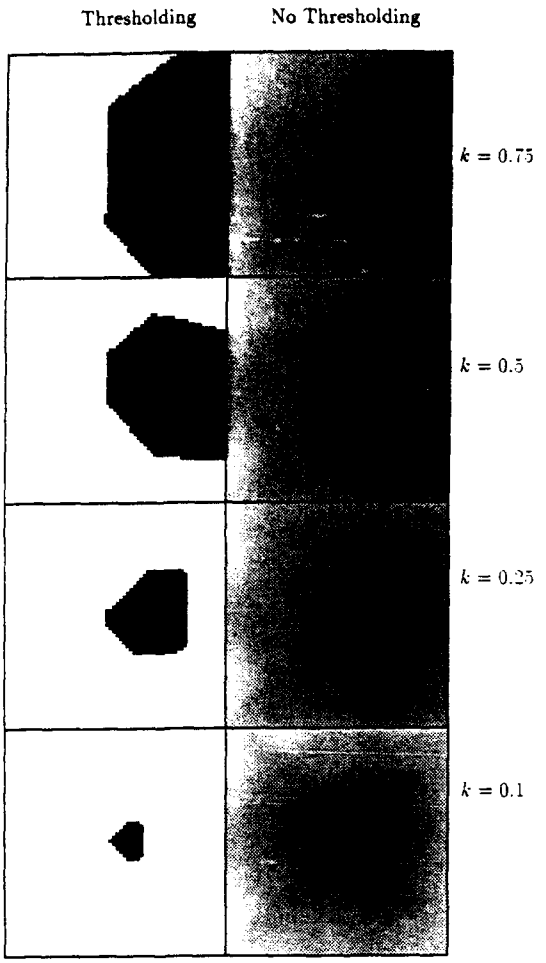


Fig. 30.

Fig. 31.

6.2 Real Data

Figure 31 shows one of the images from a dense sequence collected in our laboratory using an American Merlin Robot Arm carrying a miniature CCD Sony T.V. camera and translating along the camera's optical axis (Figure 32). Figure 33 shows the last frame of the sequence and Figure 34 shows the normal flow field obtained. Finally, Figure 35 shows the first frame with the solution area (where the FOE lies), which agrees with the ground truth.

Figure 36 shows the first from a series of images acquired and made public for the IEEE 1991 Workshop on Motion by NASA Ames Research Center. The camera is moving forward (FOE = (232,240), which is in our images (4 times reduced) in the middle of the white area of the Coca-Cola can). Figure 37 shows a normal flow field acquired from this sequence and Figure 38 shows the solution area. Figure 39 shows the solution area superimposed with the first frame, which contains the actual solution.

Figure 40 shows the first of a series of images collected by the University of Massachusetts at Amherst and made public for the IEEE 1991 Workshop on Motion. The camera was mounted on a robot arm. The upper arm of the robot (shoulder to elbows) is approximately along the viewing direction. The lower arm (elbow to gripper) is normal to the upper arm (90 deg.). The camera is traveling along the circle centered at the elbow and the axis of the camera is parallel to the upper arm. Since the scene is 5–10m away the effect is one of the rotation about the axis parallel to the viewing direction and small translation normal to it (FOE at infinity, dominant rotation, k approximately equal to 0.1. Figure 41 shows the last frame of the sequence. Figure 42 represents the normal flow estimated using frames 3, 4 and 5. Figure 43 shows the results of voting for the position of the AOR and Figure 44 shows the position of the AOR superimposed on the first frame.

The above experiments produced very good results (actual solution always inside the solution area) because there was either dominant translation or dominant rotation. The experiment below demonstrates noisy results for the FOE and the AOR because translation and rotation have about the same proportion on the image.[10] Figure 45 shows the first from a series of images of a box rotating around a vertical axis passing through the middle of the up-face, and collected by

Fig. 32.

Fig. 33.

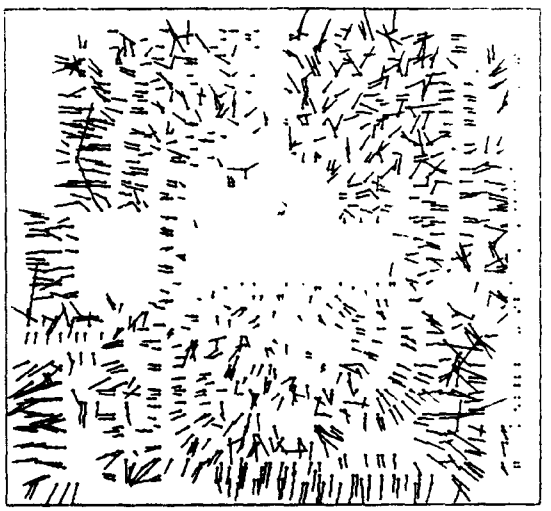


Fig. 34.

Fig. 35.

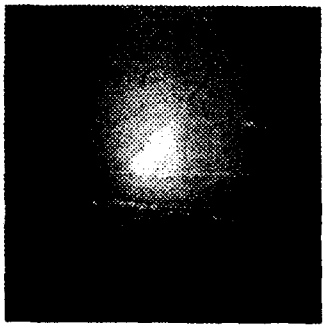


Fig. 38.

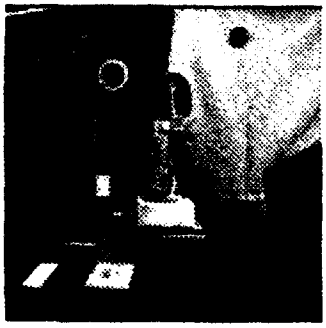


Fig. 36.

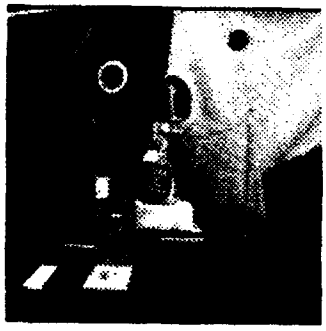


Fig. 39.

Fig. 37.

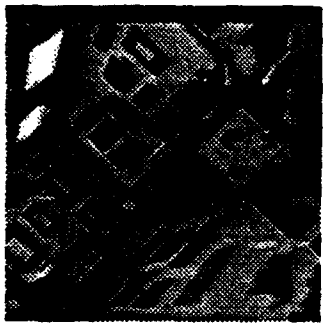


Fig. 40.

Fig. 41.

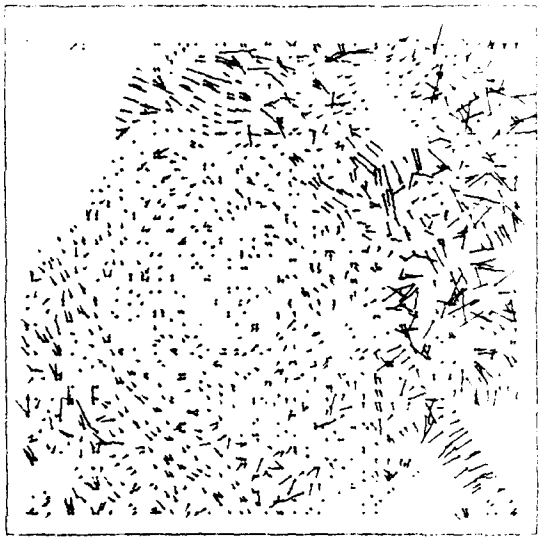


Fig. 42.

the University of Massachusetts at Amherst and made public for the 1991 Workshop on Motion. The box is rotating around the shaft. To compare our algorithm's results with ground truth we need to understand the object's motion in a camera-centered coordinate system. Since there is a distance between the box and the camera of 600 mm, the retinal motion produced by the box's rotation is equivalent to the motion produced by a general motion of the observer, consisting of both rotation and translation. The axis of rotation is parallel to the shaft pointing downward and the translation is along a circle centered at the shaft with radius 600 mm. The circle lies on a plane normal to the shaft, and the direction of translation is to the left side (with the FOE at infinity). In this case rotation and translation are of about the same proportion ($k \simeq 1$) and thus the results for both the FOE and the AOR are noisy. Figure 46 shows the normal flow field obtained from the first three images of the sequence, and figures 47 and 48 display the solutions for the FOE and the AOR respectively, superimposed on the original image.

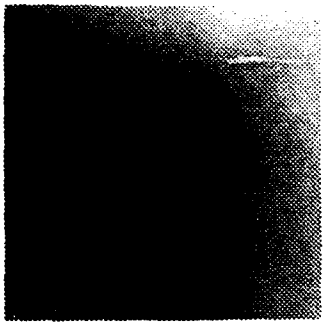


Fig. 43.

7 Conclusions

A technique was presented for computing the direction of motion of a moving observer using as input the normal flow field. In particular, for the actual computation only the direction of the normal flow is used. We showed theoretically that the method works robustly even when some amount of rotation is present, and we quantified the relationship between time-to-collision and magnitude of rotation that allows the method to work correctly. It has been shown that the position of the estimated

Fig. 44.

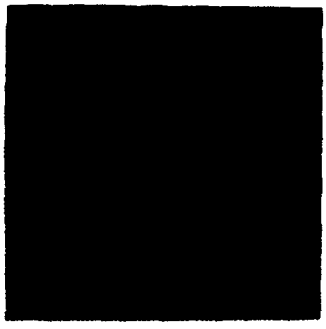


Fig. 45.

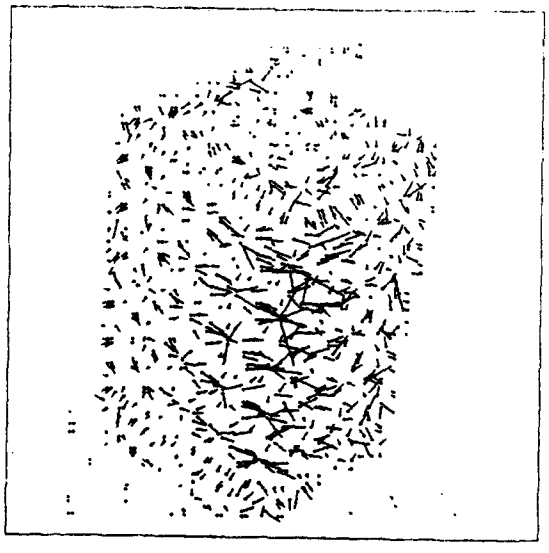


Fig. 46.

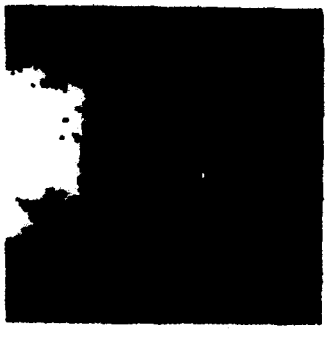


Fig. 47.

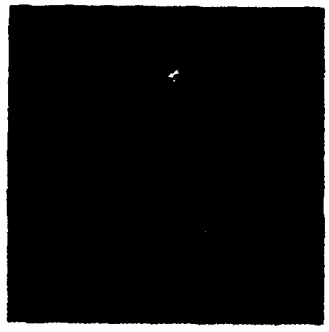


Fig. 48.

FOE is displaced in the presence of rotation and this displacement has been explained. The practical significance of this research is that if we have at our disposal an inertial sensor whose error bounds are known, we can use the method described in this paper to obtain a machine vision system that can robustly compute the heading direction. However, if rotation is not large,[11] then the method can still reliably compute the direction of motion, without using inertial sensor information. The technique cannot be used for determining the translation of a rigidly moving object, simply because the area on the image where voting could be performed is relatively small. See, for example, Figure 49, where an object is translating parallel to the optical axis (a), but the solution area is open (b) (in this case the FOE = (0,0)). Finally, the same analysis described here has been carried out for a different coordinate system (Duric et al., 1993).

Notes

1. One can also differentiate a category of methods that use correspondence of macrofeatures (contours, lines, sets of points, etc.) (Aloimonos and Shulman 1989; Spetsakis and Aloimonos 1990), but we don't discuss them here, due to the lack of literature on the stability of such techniques.
2. As in photogrammetry, for example, for solving the problem of relative orientation (Horn 1990).
3. Since measurements are in focal length units, 1% error in displacements amounts to about 3–8 pixels for most commercially available cameras.
4. In the case of backward movement the situation is symmetric (maximum – minimum) and handled similarly.

(a) (b)

Fig. 49.

5. If computation of normal flow at some points is unreliable, we just don't compute normal flow there (see Section 6).
6. In the sense of Figure 2; we assume that the observer usually looks towards where it is moving.
7. These two vectors are not defined at the FOE, i.e. at $\vec{r} = \vec{k}$, $\vec{u}_t = 0$, $\vec{r} \times \vec{k} = \vec{0}$, but this is the only singular point.
8. This is true regardless of rotation, and it is a general geometric result.
9. The hemisphere that contains the FOE.
10. It is worth noting that the algorithm in ([?]) will produce accurate results in this case.
11. How large is large has been quantified in Section 5.

References

J. Aloimonos, "Purposive and qualitative active vision", *Proc. ARPA Image Understanding Workshop*, 1990a, 816–828.

J. Aloimonos (Ed.), *CVGIP: Image Understanding* **56**, 1992.

J. Aloimonos and C.M. Brown, "Direct processing of curvilinear sensor motion from a sequence of perspective images", *Proc. Workshop on Computer Vision: Representation and Control*, Annapolis, MD, 72–77, 1984.

J. Aloimonos and D. Shulman, *Integration of Visual Modules: An Extension of the Marr Paradigm*, Academic Press, Boston, 1989.

J. Aloimonos, I. Weiss, and A. Bandopadhay, "Active vision", *Int'l. J. Comp. Vision* **2**, 1988, 333–356.

D.H. Ballard, "Parameter networks", *Artificial Intelligence* **22**, 1984, 235–267.

A. Blake, S. Murray and A. Sinclair, "Egomotion from normal flow", personal communication, 1992.

A. Bruss and B.K.P. Horn, "Passive navigation", *Computer Vision, Graphics Image Processing* **21**, 1983, 3–20.

W. Burger and B. Bhanu, "Estimating 3D egomotion from perspective image sequences", *IEEE Trans. PAMI* **12**, 1990, 1040–1058.

Z. Duric, A. Rosenfeld and L.S. Davis, "Passive navigation using the Frenet-Secret model," Proc. 4th ICCV, Berlin, Germany, 1993, pp. 363–369.

O. Faugeras and S. Maybank, 'Motion from point matches: Multiplicity of solutions", *Int'l. J. Computer Vision* 4, 1990, 225–246.

C. Fermüller, "Basic Visual Capabilities," Ph.D. Thesis, Institute for Automation, Technical University of Vienna, 1993a (also available as Technical Report 668, Center for Automation Research, University of Maryland).

C. Fermüller, "Navigational preliminaries", in *Active Perception*, Y. Aloimonos (Ed.), Lawrence Erlbaum Associates, Inc., Publishers, 1993b.

E. Francois and P. Bouthemy, "Derivation of qualitative information in motion analysis", *Image and Vision Computing* **8**, 1990, 279–288.

B.K.P. Horn, "Relative Orientation", *Int'l J. Comp. Vision* **4**, 1990, 59–78.

B.K.P. Horn and E.J. Weldon, "Computationally efficient methods of recovering translational motion", *Proc. International Conference on Computer Vision*, 1987, 2–11.

J.J. Koenderink 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**, 1975, 773–791.

G.A. Korn and T.M. Korn, *Mathematical Handbook for Engineers and Scientists*, McGraw Hill, 1968.

H.C. Longuet-Higgins, "A computer algorithm for reconstructing a scene from two projections", *Nature* **293**, 1981, 133–135.

H.C. Longuet-Higgins and K. Prazdny, "The interpretation of a moving retinal image", *Proc. Royal Soc. London B* **208**, 1980, 385–397.

D. Marr, *Vision*, W.H. Freeman, San Francisco, 1982.

S.J. Maybank, "The angular velocity associated with the optical flow field arising from motion through a rigid environment", *Proc. Royal Soc. London A* **401**, 1985, 317–326.

S. Negahdaripour, Ph.D. Thesis, MIT Artificial Intelligence Laboratory, 1986.

R.C. Nelson and J. Aloimonos, "Finding motion parameters from spherical flow fields (Or the advantages of having eyes in the back of your head)", *Biological Cybernetics* **58**, 1988, 261–273.

R. Nelson and J. Aloimonos, "Using flow field divergence for obstacle avoidance in visual navigation", *IEEE Trans. PAMI* **11**, 1989, 1102–1106.

M.E. Spetsakis and J. Aloimonos, "Optimal computing of structure from motion using point correspondences in two frames", *Proc. International Conference on Computer Vision*, 1988.

M.E. Spetsakis and J. Aloimonos, "Unification theory of structure from motion", Technical Report CAR-TR-482, Computer Vision Laboratory, Center for Automation Research, University of Maryland, College Park, 1989.

M.E. Spetsakis and J. Aloimonos, "Structure from motion using line correspondences", *Int'l. J. Computer Vision* **4**, 1990, 171–183.

W.B. Thompson and J.K. Kearney, "Inexact vision", *Proc. Workshop on Motion* 1986, 15–22.

W.B. Thompson and J.S. Painter, "Qualitative constraints for structure from motion", in Y. Aloimonos (Ed.), *CVGIP: Image Understanding* **56**, 1992, 69–77.

M. Tistarelli and G. Sandini, "Dynamic aspects in active vision", in Y. Aloimonos (Ed.), *CVGIP: Image Understanding*, **56**, 1992, 87–104.

R.Y. Tsai and T.S. Huang, "Uniqueness and estimation of three dimensional motion parameters of rigid objects with curved surfaces", *IEEE Trans. PAMI* **6**, 1984, 13–27.

S. Ullman, *The Interpretation of Visual Motion*, MIT Press, Cambridge, MA, 1979.

D. Weinshall, "Qualitative depth from stereo with applications", *CVGIP* **49**, 1990, 222–241.

D. Weinshall, "Direct computation of qualitative 3D shape and motion invaraiants", *IEEE Trans. PAMI* **13**, 1991, 1236–1240.

J. Weng, T.S. Huang and N. Ahuja, "A two step approach to optimal motion and structure estimation", *Proc. IEEE Computer Society Workshop on Computer Vision*, 1987.

G. White and E. Weldon. Utilizing Gradient Vector Distributions to Recover Motion Parameters. In *Proc. IEEE Workshop on Computer Vision*, pages 132–137, 1987.

G.S. Young and R. Chellappa, "3-D motion estimation using a sequence of noisy stereo images", *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, 1988.

A. Zisserman and R. Cipolla, "Qualitative surface shape from deformations of image curves", *Proc. Workshop on Qualitative Vision, AAAI '90*, 1990, 41–45.