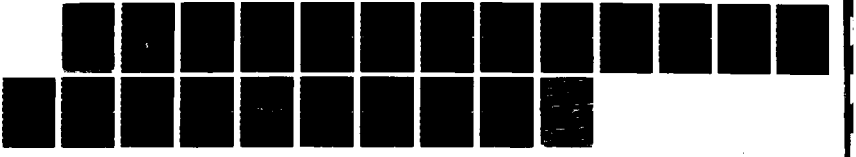
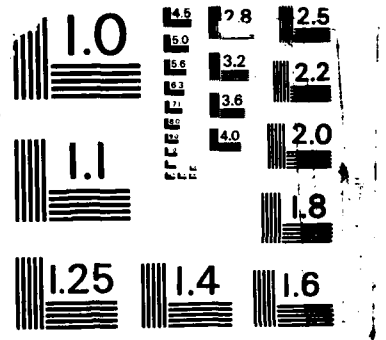
AD-A181 200 AMBIGUITIES OF A MOTION FIELD(U) MASSACHUSETTS INST OF 1/1
TECH CAMBRIDGE ARTIFICIAL INTELLIGENCE LAB
S NEGAHDARIPOUR JAN 87 AI-M-940 N00014-75-C-0643
UNCLASSIFIED F/G 12/9 NL





MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS - 1963 - A

DTIC FILE COPY

(12)

REPORT DOCUMENTATION PAGE

1. REPORT NUMBER	2. GOVT ACCESSION N
AI-MEMO 940	

AD-A181 200

4. TITLE (and Subtitle)

Ambiguities of a Motion Field

AI Memo

6. PERFORMING ORG. REPORT NUMBER

7. AUTHOR(s)

Shahriar Negahdaripour

8. CONTRACT OR GRANT NUMBER(s)

N00014-75-C-0643

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

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

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

12. REPORT DATE
Jan 87

13. NUMBER OF PAGES
20

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

15. SECURITY CLASS. (of this report)
UNCLASSIFIED

15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Distribution is unlimited.

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

DTIC
ELECTE
JUN 11 1987
S D
D

18. SUPPLEMENTARY NOTES

None

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

Motion Field, Optical Flow, Quadratic Surfaces Asymptotic Lines

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

Abstract: We study the conditions under which a perspective motion field can have multiple interpretations and present analytical expressions for the relationship among these interpretations. Furthermore, ~~we show that~~, in most cases, the ambiguity in the interpretation of a motion field can be resolved by imposing the physical constraint that depth is positive over the image region onto which the surface projects.

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
ARTIFICIAL INTELLIGENCE LABORATORY

A.I. Memo No. 940 January, 1987

Ambiguities of a Motion Field

Shahriar Negahdaripour

Abstract: In this paper we study the conditions under which a perspective motion field can have multiple interpretations, and present analytical expressions for the relationship among these interpretations. It is shown that, in most cases, the ambiguity in the interpretation of a motion field can be resolved by imposing the physical constraint that depth is positive over the image region onto which the surface projects.

Key Words: Motion Field, Optical Flow, Quadratic Surfaces, Asymptotic Lines

© Massachusetts Institute of Technology, 1987

This report describes research done at the Artificial Intelligence Laboratory of the Massachusetts Institute of Technology. Support for the laboratory's artificial intelligence research is provided in part by the Defense Advanced Research Projects Agency under Office of Naval Research contract No. N00014-75-C-0643.

DTIC
COPY
INSPECTED

A-1

1 Introduction

One of the primary tasks of a computer vision system is to reconstruct from two-dimensional images certain three-dimensional properties of a scene such as motion, shape, and the spatial arrangment of objects in the scene. In monocular vision, the goal is to recover, from time-varying images, the relative motion between a viewer and the environment as well as the structure of the environment. The structure of the scene is usually defined in terms of the relative depth of points on the visible part of the surface of the scene.

An important issue in motion vision is whether the solution can be determined uniquely. Alternatively, one may ask if there exist situations that give rise to an ambiguity in the interpretation of three-dimensional motion and shape.

Apparently, Hay [1966] was the first to report the observation that two planar surfaces each undergoing different rigid motions can produce the same instantaneous motion in the image plane. The same observation and the proof of the existence of at most two solutions in the case of planar surfaces has since been reported by Tsai et al. [1982], Waxman & Ullman [1985], Longuet-Higgins [1984], Maybank [1984], and Negahdaripour & Horn [1985].

In the case of curved surfaces, two types of approaches, based on local and global representation of three-dimensional surfaces, have been pursued. In the local approach, the surface is represented by its Taylor series expansion in some neighborhood of the fixation point, based on the assumption that the surface has continuous derivatives up to some order n in that region. In the global approach, no special model is assumed and the depth values are allowed to vary arbitrarily from one point to the next.

Using a local second-order analysis, Longuet-Higgins & Prazdny [1980] show that three interpretations are possible for the motion parameters and the local structure of the surface of the scene (tangent plane orientation and surface curvature). Waxman et al. [1986] derive the special cases which give rise to the three-fold ambiguity observed by Longuet-Higgins & Prazdny. In addition, they show that other situations can give rise to a two-fold ambiguity (similar results were derived by Negahdaripour & Yuille [1986]). Negahdaripour [1986] shows that only certain hyperboloids of one sheet and circular cylinders can give rise to an image motion field with multiple interpretations. The ambiguity of hyperboloids of one sheet can be either two-fold or three-fold. The ambiguity associated with circular cylinders is two-fold and can be viewed as a degenerate case of the three-fold ambiguity of hyperboloids of one sheet. Negahdaripour [1986] also shows that most of the ambiguities observed by Waxman et al. are the shortcomings of a local second-order analysis of the motion field.

The same problem has been addressed using a global analysis. Fang & Huang [1984] use the correspondence of nine image points in two frames to show that the motion parameters can be determined uniquely unless the points lie on a second-order surface passing through the viewing point. Tsai & Huang [19[illegible] show that the correspondence of

seven image points is sufficient to recover motion uniquely unless the points lie on either two planes with one passing through the viewing point or a cone that passes through the viewing point. In an elegant derivation, Horn [1986b] shows that the class of curved surfaces that may give rise to the same motion field is restricted to certain hyperboloids of one sheet that are viewed from a point on their surface. He also shows that the special cases observed by Tsai & Huang are degenerate cases of his results.

In this paper, we present some results related to the ambiguities in the interpretation of three-dimensional motion of curved surfaces. These results can be summarized as follows:

(1) Only certain hyperboloids of one sheet or cylindrical surfaces that are viewed by an observer moving parallel to the image plane can give rise to an ambiguity in the interpretation of a perspective motion field. These ambiguities occur under two rare conditions and can be either two-fold or three-fold. In either case, the resulting motion field is quadratic.

(2) When multiple solutions exist, the relationship among them can be derived in closed-form.

(3) With a large field of view, it is generally possible to identify the correct solution by imposing the constraint that depth is positive over the image region onto which the surface projects.

2 Motion Field

We assume a viewer-centered coordinate system. The optical axis is chosen along the z-axis, and the image is formed on the plane $z = 1$; that is, without loss of generality, the focal length is chosen to be unity (Figure 1). Let $\mathbf{R} = [X, Y, Z]^T$ be a point in the scene that projects onto the point $\mathbf{r} = [x, y, 1]^T$ in the image. Assuming a perspective projection, we have

$$\mathbf{r} = \frac{1}{\mathbf{R} \cdot \hat{\mathbf{k}}} \mathbf{R},$$

where $Z = \mathbf{R} \cdot \hat{\mathbf{k}}$ is the distance of the point $\mathbf{R}$ from the viewer along the optical axis. Suppose the viewer moves with relative translational and rotational velocities $\mathbf{t}$ and $\boldsymbol{\omega}$ with respect to the scene. Then a point in the scene moves with respect to the viewer with velocity

$$\mathbf{R}_t = -\mathbf{R} \times \boldsymbol{\omega} - \mathbf{t}.$$

The corresponding image point moves with velocity (Negahdaripour & Horn [1987])

$$\mathbf{r}_t = -\hat{\mathbf{k}} \times \left(\mathbf{r} \times \left(\mathbf{r} \times \boldsymbol{\omega} - \frac{1}{\mathbf{R} \cdot \hat{\mathbf{k}}} \right) \right.$$

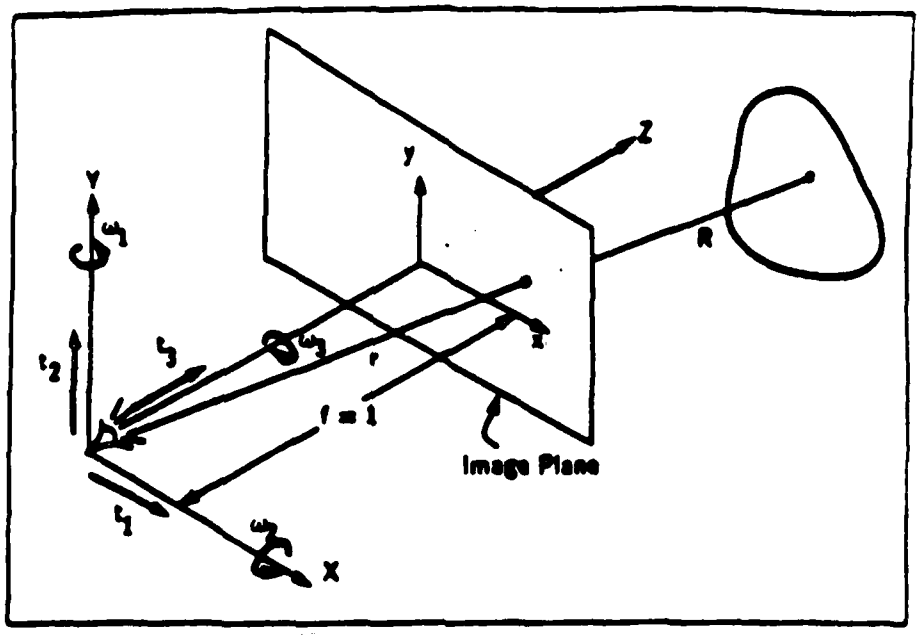


Figure 1. Viewer-centered coordinate system.

Note that $\mathbf{r}_t \cdot \hat{\mathbf{z}} = 0$ and, therefore, the third component of the 3D vector $\mathbf{r}_t$ is always zero, as expected. The velocities of all image points taken collectively defines a two-dimensional vector field that we call the image motion field (Negahdaripour & Horn [1987]). This vector field has been referred to elsewhere as the optical flow field (see Horn [1986a] or Negahdaripour [1986] for the distinction between the motion field and optical flow).

The motion field remains unchanged if the translational vector t and the depth values Z are multiplied by the same constant factor. Therefore, we can recover depth and the translational motion from image motion field only up to a scale factor; this has been referred to as the scale-factor ambiguity in motion vision.

3 The Uniqueness Issue

The motion field is a purely geometric concept. It is uniquely defined in terms of the observer motion and the scene structure. More precisely, once we specify the motion of the viewer as well as the structure of the scene, the motion field is unique as given by the earlier equation; that is,

$$\text{3D motion and structure} \xrightarrow{\text{unique}} \text{2D image motion}$$

Ideally, the motion field over some region of the image can be used to recover the relative motion, $\boldsymbol{\omega}$ and t, as well as the structure of the scene (defined in terms of the depth values of the points on the surface of the scene, Z) up to a scale factor. The issue

we address is the uniqueness of the solution:

$$\text{2D image motion} \xrightarrow{\text{unique?}} \text{3D motion and structure}$$

More precisely, the question we answer is: Given the motion field over some region of the image, when can we recover the motion parameters and the scene structure uniquely (up to the scale-factor ambiguity)? Note that the solution may be non-unique because the relationship between the two-dimensional image motion field and the underlying three-dimensional structure and motion is non-linear and, therefore, there need not be a one-to-one correspondence. Alternatively, we are interested to know what circumstances may give rise to an ambiguity in the interpretation of the motion field. In general, the problem of robust recovery of 3D structure and motion can be more complicated (than deriving the uniqueness results) since we cannot compute the motion field accurately enough from time-varying images. The inaccuracy in the estimate of the motion field obtained from noisy data may then introduce additional ambiguities in the interpretation of three-dimensional motion and shape (Ullman [1983], Jerian & Jain [1983], Adiv [1985]). In this paper, we restrict attention to the types of ambiguities that arise even when the exact motion field is used.

4 Surface Representation

One of the issues in deriving uniqueness results for curved surfaces is the choice of an appropriate representation of general three-dimensional surfaces. For example, the surfaces of a scene may be represented by the relationship $Z = Z(X,Y)$, where $\mathbf{R} = [X,Y,Z]^T$ is a point on the surface of the scene. This will be referred to as a global representation since we can define the whole surface in this form. There is no constraint on the relationship among the depth values of neighboring points and the surface need not have a particular structure in local regions; for example, it need not be smooth. The surfaces of most physical objects, however, possess some degree of "regularity" at least in local regions; the regularity or smoothness of a surface can be measured in terms of the continuity in the surface function and its derivatives. In other words, most surfaces have continuous derivatives up to some order n in local regions. Then we can represent a surface by its Taylor series expansion up to order n in a local region, assumed, for simplicity, to be around the fixation point:

$$Z = Z_0 + Z_X X + Z_Y Y + \frac{1}{2} Z_{XX} X^2 + Z_{XY} XY + \frac{1}{2} Z_{YY} Y^2 + \ldots + \frac{1}{n!} Z_{X\ldots X} X^n + \ldots + \frac{1}{n!} Z_{Y\ldots Y} Y^n + O(\epsilon).$$

Alternatively, the surface may be represented by a Taylor series expansion in terms of the image coordinates. This is a more convenient representation when we deal with images. Finally, we may write the expansion of $d = 1/Z$, which is a measure related to

the disparity function in stereopsis), instead of the expansion of Z, since the motion field equation depends on $1/Z$ (as a result of perspective projection). Therefore, the surface can be represented by the equation

$$d = d_0 + d_x x + d_y y + \frac{1}{2} d_{xx} x^2 + d_{xy} xy + \frac{1}{2} d_{yy} y^2 + \ldots + \frac{1}{n!} d_{x \ldots x} x^n + \ldots + \frac{1}{n!} d_{y \ldots y} y^n + O(e).$$

Ignoring the error term, this can be written more compactly using tensor notation,

$$d = d_0 + d_i r_i + \frac{1}{2} d_{ij} r_i r_j + \ldots + \frac{1}{n!} d_{ij \ldots} r_i r_j \ldots ,$$

where $r_i = [x, y]^T$. Since depth can be recovered only up to a scale factor, we can set $d_0 = 1$. Therefore, we can write

$$d = 1 + d_i r_i + \frac{1}{2} d_{ij} r_i r_j + \ldots + \frac{1}{n!} d_{ij \ldots} r_i r_j \ldots .$$

This will be referred to as a local representation since it may be the true representation of the surface only in a local region near the fixation point. One justification for using a local representation is that we only need a finite number of parameters, namely, the coefficients of the Taylor series (d_i, d_{ij}, ...) to represent the surface, where the number of parameters is related to the order of the Taylor series. We may not need to impose any restriction on n in addressing the uniqueness issue since we deal solely with a purely theoretical problem. The problem of robust recovery of motion and shape, however, is a separate issue. In practice, we cannot robustly determine the coefficients of the surface function beyond the linear terms due to noise in the data (Adiv [1985], Le Guilloux [1986]). Furthermore, the resulting non-linear problem is usually ill-conditioned.

Using a local representation of curved surfaces, the motion vision problem reduces to estimating a finite number of unknowns, namely, the motion and surface parameters, from the image data. Similarly, the uniqueness issue translates into the following question: How many sets of motion/surface parameters can give rise to the same motion field over the image region of interest?

5 Ambiguities in the Interpretation of the Motion Field

If we substitute the equation for $d = 1/Z$ into the motion field equation, we arrive at the Taylor series representation of the motion field, $\mathbf{r}_t$ [illegible] $u, v, 0]^T$, in the image region under

consideration; that is,

$$u = u_0 + u_x x + u_y + \frac{1}{2}u_{xx}x^2 + u_{xy}xy + \frac{1}{2}u_{yy}y^2 + \frac{1}{6}u_{xxx}x^3 + \frac{1}{2}u_{xxy}x^2y + \frac{1}{2}u_{xyy}xy^2 + \frac{1}{6}u_{yyy}y^3 + \ldots$$

$$v = v_0 + v_x x + v_y + \frac{1}{2}v_{xx}x^2 + v_{xy}xy + \frac{1}{2}v_{yy}y^2 + \frac{1}{6}v_{xxx}x^3 + \frac{1}{2}v_{xxy}x^2y + \frac{1}{2}v_{xyy}xy^2 + \frac{1}{6}v_{yyy}y^3 + \ldots$$

where the coefficients of the Taylor series expansion of the motion field are given by

$$\begin{aligned}
u_0 &= -w_y - t_x & v_0 &= w_x - t_y \\
u_x &= t_z - d_x t_x & v_x &= -w_x - t_y d_x \\
u_y &= w_z - t_x d_y & v_y &= t_z - t_y d_y \\
u_{xx} &= -w_y + t_z d_x - 2t_x d_{xx} & v_{xx} &= -2t_y d_{xx} \\
u_{xy} &= w_x + t_z d_y - t_x d_{xy} & v_{xy} &= -w_y + t_z d_x - t_y d_{xy} \\
u_{yy} &= -2t_x d_{yy} & v_{yy} &= w_x + t_z d_y - 2t_y d_{yy} \\
u_{xxx} &= 3t_z d_{xx} - t_x d_{xxx} & v_{xxx} &= -t_y d_{xxx} \\
u_{xxy} &= 2t_z d_{xy} - t_x d_{xxy} & v_{xxy} &= t_z d_{xx} - t_y d_{xxy} \\
u_{xyy} &= t_z d_{yy} - t_x d_{xyy} & v_{xyy} &= 2t_z d_{xy} - t_y f d_{xyy} \\
u_{yyy} &= -t_x d_{yyy} & v_{yyy} &= 3t_z d_{yy} - t_y d_{yyy} \\
&\vdots & &\vdots
\end{aligned}$$

The question regarding the uniqueness of the solution can be rephrased as follows: Under what circumstances can we obtain the same set of motion field coefficients up to some order n for different sets of motion/surface parameters?

Considering the motion field coefficients up to the second-order terms, Longuet-Higgins & Prazdny [1980] showed that three interpretations for the motion and local structure of the surface (tangent plane orientation and surface curvature) are generally possible. They arrived at this conclusion by reducing the problem of determining the motion and surface parameters to solving a cubic equation; however, they did not show when the cubic equation can possess three real solutions. Waxman et al. [1986] derived the conditions that give rise to the ambiguity observed by Longuet-Higgins & Prazdny (similar results were derived independently by Negahdaripour & Yuille [1986]). These conditions can be categorized as follows:

(1) Three solutions are obtained when $d_x = d_y = 0$, the surface has a negative Gaussian curvature ($d_{xx}d_{yy} - d_{xy}^2 < 0$), and the mean-scaled curvature is unity ($\frac{1}{2}(d_{xx} + d_{yy}) = 1$).

(2) Two solutions are obtained if the surface has a non-positive Gaussian curvature ($d_{xx}d_{yy} - d_{xy}^2 \leq 0$), the mean-scaled curvature is unity ($\frac{1}{2}(d_{xx} + d_{yy}) = 1$), and the translation along the line of sight vanishes ($t_z = 0$). When $d_x/d_y = t_x/t_y$, the two solutions become equivalent.

(3) Two solutions are obtained regardless of the sign of the Gaussian curvature when $d_x/d_y = t_x/t_y$, (this condition was referred to as the structure-motion coincidence). The two solutions degenerate to a unique solution when the translation is along the surface normal or when $t_z = 0$.

(4) In all other cases, the solution is unique.

Longuet-Higgins & Prazdny claim that these ambiguities are usually only instantaneous, and can be resolved at the next time instant. Negahdaripour [1986] shows that an ambiguity in the interpretation of the motion field arises only in the case of some quadratic surfaces with a non-positive Gaussian curvature at the fixation point. More importantly, these surfaces have to be viewed by an observer translating perpendicular to the viewing direction. Interestingly, the restriction on the motion of the observer is peculiar to the local representation of curved surfaces. Furthermore, the resulting motion field is second-order; but this obviously does not imply that any second-order motion field is ambiguous. The ambiguity in the case of quadratic surfaces that are viewed by an observer moving parallel to the image plane (t_z=0) is either

(1) two/three-fold when the surface gradient vanishes ($d_x = d_y = 0$), the Gaussian curvature is zero/negative ($d_{xx}d_{yy} - d_{xy}^2 \leq 0$), and the mean-scaled curvature is unity ($\frac{1}{2}(d_{xx} + d_{yy}) = 1$), or

(2) two-fold when the Gaussian curvature is negative ($d_{xx}d_{yy} - d_{xy}^2 < 0$), the mean-scaled curvature is unity ($\frac{1}{2}(d_{xx} + d_{yy}) = 1$), and the surface normal, the optical axis, and one of the asymptotic lines of the quadratic surface are in the same plane; that is,

$$\text{either} \quad \frac{d_x}{d_y} = \frac{-d_{xy} + \sqrt{d_{xy}^2 - d_{xx}d_{yy}}}{d_{xx}} \quad \text{or} \quad \frac{d_x}{d_y} = \frac{-d_{xy} - \sqrt{d_{xy}^2 - d_{xx}d_{yy}}}{d_{xx}}.$$

When $d_x/d_y = t_x/t_y$, the ambiguity is resolved because the two solutions become identical.

6 Surfaces That Give Rise to an Ambiguity

We have given the conditions under which there may be an ambiguity in the interpretation of a given motion field resulting from the relative motion between an observer and a curved surface. Since the ambiguity is restricted to quadratic surfaces, we can ignore the higher order terms in the surface function given earlier. Therefore, an "ambiguous surface" is given by

$$d = 1 + d_i r_i + \frac{1}{2} d_{ij} r_i r_j,$$

which can be written

$$d = \frac{1}{2}\mathbf{r}^T\mathbf{D}\mathbf{r},$$

where

$$\mathbf{D} = \begin{pmatrix} d_{xx} & d_{xy} & d_x \\ d_{xy} & d_{yy} & d_y \\ d_x & d_y & 2 \end{pmatrix}.$$

If we multiply both sides of the equation for $d = 1/Z$ by Z^2, we obtain

$$Z = \frac{1}{2}\mathbf{R}^T\mathbf{D}\mathbf{R}$$

or

$$-\hat{\mathbf{z}}^T\mathbf{R} + \frac{1}{2}\mathbf{R}^T\mathbf{D}\mathbf{R} = 0,$$

which is the equation of a quadratic surface passing through the origin (viewing point). This, however, is an artifact resulting from the multiplication of both sides of the equation for d by Z^2; that is, by multiplying both sides of the equation for d by Z^2, we have artificially made the origin ($\mathbf{R} = \mathbf{0}$) a point on the surface. To be precise, we should define the surface by

$$-\hat{\mathbf{z}}^T\mathbf{R} + 2\mathbf{R}^T\mathbf{D}\mathbf{R} = 0, \qquad \mathbf{R} \neq \mathbf{0}.$$

This brings up an interesting issue since some previous results derived using a global representation of curved surfaces suggest that an "ambiguous surface" should pass through the viewing point (Fang & Huang [1984], Tsai & Huang [1984], Horn [1986b]). The inconsistency in the interpretation of the results can be attributed to the difference between the local and global representations of the surface and can be explained through the following example: Consider an observer on the inside surface of a cylinder viewing a point on the inside surface directly across from him/her (as before, a viewer-centered coordinate system is assumed). In the global representation, we consider the whole surface, whereas in the local representation, we only consider some neighborhood of the fixation point. The viewing point is a point on the surface in the global representation. It will not be so in the local representation. The difficulty, however, arises when we consider the behavior of the motion field at the origin of the image plane. The question is: Which point projects onto the origin of the image plane, the viewing point or the point directly across? Consequently, is it the motion of the viewing point or the point across that we observe (or measure) in the image? Clearly, it is the latter. Whether we do or do not include the origin as a point on the surface seems to depend on the choice of the representation we use.

The signs of the eigenvalues of the matrix $\mathbf{D}$ determine the type of the quadratic surface. When an ambiguity exists, the eigenvalues of $\mathbf{D}$, in the ascending order, are given by (see the Appendix)

$$\lambda_- = 1 - \sqrt{1 + p + q}, \quad \lambda_0 = 2, \quad \text{and} \quad \lambda_+ = 1 + \sqrt{1 + p + q},$$

where

$$p = d_x^2 + d_y^2 \quad \text{and} \quad q = d_{xy}^2 - d_{xx}d_{yy}.$$

In the case of a three-fold ambiguity, we have $p = 0$ and $q > 0$. The eigenvalues of $\mathbf{D}$ become

$$\lambda_- = 1 - \sqrt{1+q} < 0, \quad \lambda_0 = 2, \quad \text{and} \quad \lambda_+ = 1 + \sqrt{1+q} > 2.$$

Since the ambiguous surface has one negative and two positive eigenvalues, it has to be a hyperboloid of one sheet. When $q = 0$, that is, the Gaussian curvature vanishes, the ambiguity degenerates to a two-fold one. In this case, we have

$$\lambda_- = 0, \quad \lambda_0 = 2, \quad \text{and} \quad \lambda_+ = 2.$$

and the hyperboloid of one sheet degenerates to a circular cylinder. The other ambiguous situation arises when the optical axis, the surface normal, and an asymptotic line are in the same plane. In this case, $p > 0$ and $q > 0$, and we obtain

$$\lambda_- = 1 - \sqrt{1+p+q} < 0, \quad \lambda_0 = 2, \quad \text{and} \quad \lambda_+ = 1 + \sqrt{1+p+q} > 2.$$

Again, the surface is a hyperboloid of one sheet. Note that, in every case, we have $\lambda_- + \lambda_+ = \lambda_0$ and, therefore, there is a constraint among the eigenvalues. We conclude that not every hyperboloid of one sheet or circular cylinder can give rise to an ambiguity.

7 Relationship Among Multiple Solutions

The conditions given above allow us to determine when we may expect an ambiguity in the interpretation of a motion field and the number of possible interpretations. Generally, we do not know the correct interpretation in advance. Therefore, we need to obtain all of the possible solutions before we can identify the correct one; it is often difficult, if not impossible, to obtain a robust closed-form solution in the case of curved surfaces. We may need to rely on some type of iterative method to recover motion from a set of non-linear equations. An iterative method, however, can converge to only one of up to three possible solutions (when multiple solutions exist) depending on the initial condition. Several runs of the iterative algorithm may be necessary before we obtain every possible solution; this is expensive computationally. Furthermore, in some cases, a solution may be hard to obtain if it has a small radius of convergence (this is important if it is the true solution). Finally, the true solution may not be an optimum one (in the least-squares sense) with noisy data. Therefore, it is not only important to know what circumstances can give rise to an ambiguity and the number of possible solutions, it is equally important to know the relationship between the true and spurious solutions.

Suppose we have determined a motion/surface pair, $\{\mathbf{t}, \omega\}$ and $d = 1/Z$, that is consistent with the data, where the surface is given by

$$d = 1 + d_i r_i + \frac{1}{2} d_{ij} r_i r_j.$$

Furthermore, suppose we know that the solution is not unique (that is, the solution we have obtained either fully or, due to noise, approximately satisfies the conditions for an ambiguous case). Then we can find another (two other) motion/surface pair(s), $\{\tilde{\mathbf{t}}, \tilde{\omega}\}$ and $\tilde{d}$, that is (are) consistent with the data, where the surface $\tilde{d}$ is given by

$$\tilde{d} = 1 + \tilde{d}_i r_i + \frac{1}{2}\tilde{d}_{ij} r_i r_j.$$

Negahdaripour [1986] has derived the relationship among the multiple solutions in the two potentially ambiguous cases described earlier.

7.1 Case One: Three-Fold Ambiguity of Hyperboloids of One Sheet

There may be three interpretations if the surface of the scene is a hyperboloid of one sheet. The "surface of ambiguity" is characterized by the following conditions:

(1) the surface gradient is zero ($d_x = d_y = 0$),

(2) the Gaussian curvature is negative ($d_{xx}d_{yy} - d_{xy}^2 < 0$), and

(3) the mean-scaled curvature is unity ($\frac{1}{2}(d_{xx} + d_{yy}) = 1$).

The other solutions, in terms of the first solution, are given by

$$\begin{aligned}
\tilde{t}_x &= \alpha \tilde{t}_y & \tilde{\omega}_x &= w_x - k d_{yy}/(2\alpha\tilde{t}_y) \\
\tilde{t}_y &= (t_x d_{yy} + \alpha t_y d_{xx})/(2\alpha) & \tilde{\omega}_y &= w_y - k d_{xx}/(2\tilde{t}_y) \\
\tilde{t}_z &= 0 & \tilde{\omega}_z &= w_z \\
\tilde{d}_0 &= 1 & \tilde{d}_{xx} &= t_y d_{xx}/\tilde{t}_y \\
\tilde{d}_x &= 0 & \tilde{d}_{xy} &= -(\alpha t_x d_{xx} + t_y d_{yy})/(2\alpha\tilde{t}_y) \\
\tilde{d}_y &= 0 & \tilde{d}_{yy} &= t_x d_{yy}/(\alpha\tilde{t}_y)
\end{aligned}$$

where

$$\alpha = \frac{-d_{xy} \pm \sqrt{d_{xy}^2 - d_{xx}d_{yy}}}{d_{xx}} \quad \text{and} \quad k = (\alpha t_y - t_x)\tilde{t}_y.$$

For each solution of α, we obtain one "spurious" solution from these equations. We can alternatively determine both spurious solutions using only one of the solutions for α. This requires that we derive another set of relationships for the dual solutions. These are given by

$$\begin{aligned}
\tilde{t}_x &= (\alpha t_x + t_y) d_{yy}/(2\alpha) & \tilde{\omega}_x &= w_x - k\alpha/(\alpha t_x + t_y) \\
\tilde{t}_y &= (\alpha t_x + t_y) d_{xx}/2 & \tilde{\omega}_y &= w_y - k/(\alpha t_x + t_y) \\
\tilde{t}_z &= 0 & \tilde{\omega}_z &= w_z \\
\tilde{d}_0 &= 1 & \tilde{d}_{xx} &= 2t_y/(\alpha t_x + t_y) \\
\tilde{d}_x &= 0 & \tilde{d}_{xy} &= -(t_x + \alpha t_y)/(\alpha t_x + t_y) \\
\tilde{d}_y &= 0 & \tilde{d}_{yy} &= 2\alpha t_x/(\alpha t_x + t_y)
\end{aligned}$$

where

$$k = -(\alpha t_x + t_y)(\alpha t_x d_{xx} - t_y d_{yy})/(2\alpha).$$

To summarize, we can determine the two "spurious" solutions either (1) by substituting the two values obtained for α into one of the two sets of relationships given above or (2) by substituting one of the two solutions of α into both sets of equations. When the Gaussian curvature is zero (with other conditions being the same), the surface degenerates to a circular cylinder and the ambiguity reduces to a two-fold one since the two solutions for α become identical.

7.2 Case Two: Two-Fold Ambiguity of Hyperboloids of One Sheet

Certain hyperboloids of one sheet can give rise to a two-fold ambiguity. The surface of ambiguity is characterized by the following conditions:

(1) the Gaussian curvature is negative ($d_{xx}d_{yy} - d_{xy}^2 < 0$),

(2) the mean-scaled curvature is unity ($\frac{1}{2}(d_{xx} + d_{yy}) = 1$), and

(3) the surface normal, the optical axis, and one of the asymptotic lines of the quadratic surface are in the same plane; that is,

$$\text{either} \quad \frac{d_x}{d_y} = \frac{-d_{xy} + \sqrt{d_{xy}^2 - d_{xx}d_{yy}}}{d_{xx}} \quad \text{or} \quad \frac{d_x}{d_y} = \frac{-d_{xy} - \sqrt{d_{xy}^2 - d_{xx}d_{yy}}}{d_{xx}}.$$

The dual solution is given by

$$\begin{aligned}
\widetilde{t}_x &= \alpha \widetilde{t}_y & \widetilde{\omega}_x &= w_x - k d_{yy}/(2\alpha \widetilde{t}_y) \\
\widetilde{t}_y &= (t_x d_{yy} + \alpha t_y d_{xx})/(2\alpha) & \widetilde{\omega}_y &= w_y - k d_{xx}/(2\widetilde{t}_y) \\
\widetilde{t}_z &= 0 & \widetilde{\omega}_z &= w_z + k d_y/\widetilde{t}_y \\
\widetilde{d}_0 &= 1 & \widetilde{d}_{xx} &= t_y d_{xx}/\widetilde{t}_y \\
\widetilde{d}_x &= t_x d_y/\widetilde{t}_y & \widetilde{d}_{xy} &= -(\alpha t_x d_{xx} + t_y d_{yy})/(2\alpha \widetilde{t}_y) \\
\widetilde{d}_y &= t_y d_y/\widetilde{t}_y & \widetilde{d}_{yy} &= t_x d_{yy}/(\alpha \widetilde{t}_y)
\end{aligned}$$

where $\alpha = d_x/d_y$ and $k = (\alpha t_y - t_x)\widetilde{t}_y$. When $d_x/d_y = t_x/t_y$, the ambiguity is resolved because the two solutions become identical.

8 Imposing Depth-Positiveness Constraint to Resolve Ambiguity

In many cases, an incorrect interpretation, when more than one exists, may be ruled out due to the fact that it violates physical constraints. One such constraint is that the depth value for a point that is projected onto the image should be positive; that is, a point can be seen only if it is in front of the camera (see also Horn [1986b]).

Consider, once again, the surface

$$d = 1 + d_i r_i + \frac{1}{2} d_{ij} r_i r_j.$$

Since $d = 1/Z$, then $d = 0$ implies that $Z \to \infty$. Therefore, the conic section

$$1 + d_i r_i + \frac{1}{2} d_{ij} r_i r_j = 0$$

is the image of points at infinity. It is also the boundary between regions with positive and negative depth values. The region(s) with positive depth values can be the image of the surface under consideration whereas the regions with negative depth values cannot. A negative depth value corresponds to a point on a surface that either is behind the camera (or translates in the opposite direction). The point cannot belong to the same physical surface that is projected onto the region of the image with positive depth values. The regions with negative depth values are the image(s) of the background and/or other objects in the scene. Therefore, the conic section is the boundary between the object under consideration and the background or other objects in the scene. We can use these boundaries to identify the correct solution, provided that the field of view is large enough so that the image includes some portion of these boundaries.

It is more instructive to study an ambiguous situation. We consider the case that gives rise to a three-fold ambiguity since it is easier to derive the equations for the minimum size of the field of view. In this case, the surface is given by

$$d = 1 + \frac{1}{2} d_{xx} x^2 + d_{xy} xy + \frac{1}{2} d_{yy} y^2$$

with the constraints

$$d_{xx} d_{yy} - d_{xy}^2 < 0 \quad \text{and} \quad \frac{1}{2}(d_{xx} + d_{yy}) = 1.$$

Suppose the image plane is circular with radius r (half-angle field of view is $\tan^{-1}(r)$ degrees). In order to have some portion of the boundaries of the true surface within the field of view, we should have $r > r_c$, where r_c is the shortest distance from the origin of the image plane to a point on the conic section

$$1 + \frac{1}{2} d_{xx} x^2 + d_{xy} xy + \frac{1}{2} d_{yy} y^2 = 0.$$

Ideally, we want a larger field of view so that we can have as much of the boundary in the image as possible.

When $d_{xy} \neq 0$, it can be shown that $r_c = \sqrt{x_c^2 + y_c^2}$, where

$$x_c = \pm \sqrt{\frac{-2}{d_{yy} m^2 + 2 d_{xy} m + d_{xx}}} \quad , \quad y_c = m x_c,$$

and

$$m = \frac{d_{yy} - d_{xx}}{2 d_{xy}} \pm \sqrt{\left(\frac{d_{yy} - d_{xx}}{2 d_{xy}}\right)^2 + 1}.$$

Only one of the two signs gives a value of m that makes the solution for x_c and y_c real-valued (the proof follows easily from the geometry of the problem). When $d_{xy} = 0$, then either $d_{xx} < 0$ or $d_{yy} < 0$ (because $d_{xx}d_{yy} - d_{xy}^2 = d_{yy}d_{xx} < 0$). In this case, we instead have either

$$\left(x_c = 0\ ,\ y_c = \sqrt{\frac{-2}{d_{yy}}}\right) \quad \text{if} \quad d_{yy} < 0$$

or

$$\left(y_c = 0\ ,\ x_c = \sqrt{\frac{-2}{d_{xx}}}\right) \quad \text{if} \quad d_{xx} < 0.$$

We can derive similar expressions for $\tilde{r}_c$, the shortest distance from the origin of the image plane to a point on the conic section of either spurious solution. Now if $\tilde{r}_c > r > r_c$, the image includes some portion of the boundaries of the true surface but no part of the boundaries of the spurious solution or if $r > r_c$ and $r > \tilde{r}_c$, then the image should include some portion of the boundaries of both the true and spurious solutions. It is only when $r < \min(r_c, \tilde{r}_c)$ that we cannot identify the true solution since then the depth values are positive everywhere in the image for every possible solution.

Example: Consider a viewer moving with translational velocity $\mathbf{t} = [1, 2, 0]^T$ with respect to surface d defined by

$$d = 1 + 0.5x^2 + 10xy + 0.5y^2.$$

Using the equations given earlier, the spurious solutions are:

(1) An observer moving with translational and rotational velocities $\tilde{\mathbf{t}}_1 = [.445, -8.98, 0]^T$ and $\tilde{\omega}_1 = [-11.0, .550, 0]^T$ with respect to surface $\tilde{d}_1$ given by

$$\tilde{d}_1 = 1 + -.112x^2 - 2.17xy + 1.11y^2.$$

(2) An observer moving with translational and rotational velocities $\tilde{\mathbf{t}}_2 = [-19.5, .975, 0]^T$ and $\tilde{\omega}_2 = [-1.03, 20.5, 0]^T$ with respect to surface $\tilde{d}_2$ given by

$$\tilde{d}_2 = 1 + +1.03x^2 - .461xy - .026y^2.$$

The boundaries of the three surfaces are shown in Figure 2. For each solution, the regions of negative depth values are shown by hatched lines. The resulting second-order motion field is given by

$$\mathbf{r}_t = \begin{pmatrix} -1 - 0.5x^2 - 10xy - 0.5y^2 \\ -2 - x^2 - 20xy - y^2 \\ 0 \end{pmatrix},$$

which is shown in Figure 3 (the image plane is a unit square; that is, the field of view is $2\tan^{-1}(0.5) \approx 54$ degrees). Note that the velocity vectors are all parallel emanating from

the focus of expansion at infinity in the region where the depth values are positive (the region including the bottom-left, the center, and the top-right of the image). They point in the opposite direction in the bottom-right and top-left of the image. This implies that, for points in these regions, either the motion is in the opposite direction (to the true one) or the depth values are negative. In either case, these regions cannot be the images of parts of the same object that is imaged into the region with positive depth values.

The apices of the conic section of the true surface are located at $\pm(1/3,\ 1/3)$. If the field of view is larger than $2\tan^{-1}(\sqrt{2}/3) \approx 51$ degrees, we can identify the correct solution by matching the boundaries of depth discontinuity in the image with the conic section of the true surface.

The two spurious solutions involve a viewer rotating about an axis parallel to one of the asymptotic directions and translating parallel to the other asymptotic direction of the true surface. This is quite counter-intuitive since the motion field suggests that the underlying 3D motion is purely translational.

9 Summary

In this paper I have presented some results concerning the ambiguity in the interpretation of the motion of curved surfaces. These results suggest that only certain hyperboloids of one sheet or circular cylinders viewed by an observer moving parallel to the image can give rise to an ambiguity in the interpretation of the underlying motion. In the case of hyperboloids of one sheet, the ambiguity can be either two-fold or three-fold, whereas there can be at most two solutions in the case of circular cylinders. In either case, the resulting motion field is second-order. I have also given analytical expressions for the relationship among multiple solutions. In most cases, an ambiguity can be resolved by imposing the positive-depth constraint.

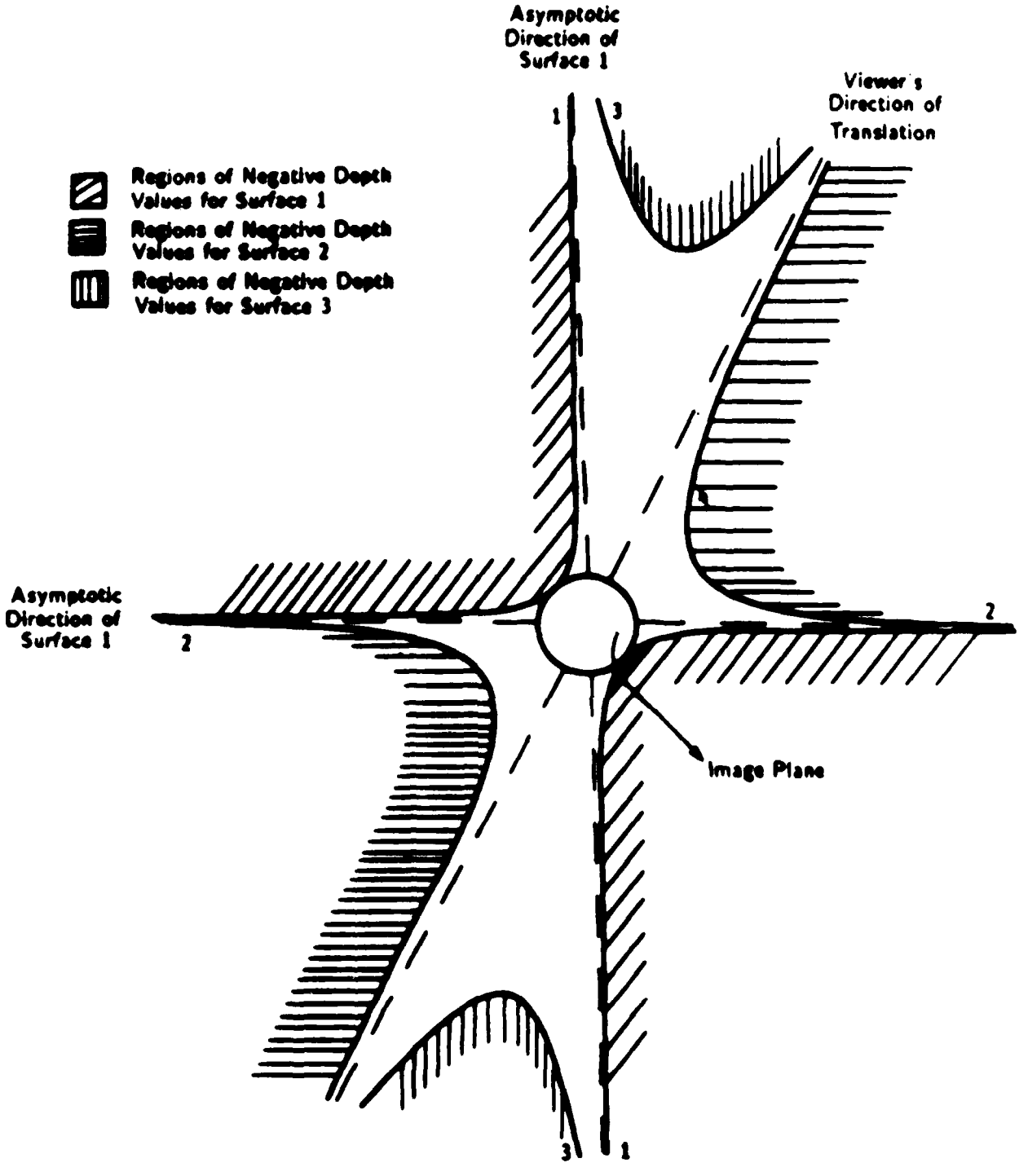


Figure 2. The images of the boundaries of three hyperboloids of one sheet responsible for a three-fold ambiguity. The ambiguity in a motion field is usually restricted to small region of the image. For a sufficiently large field of view, the ambiguity can be resolved by matching the depth discontinuity boundaries with the conic section of the true surface.

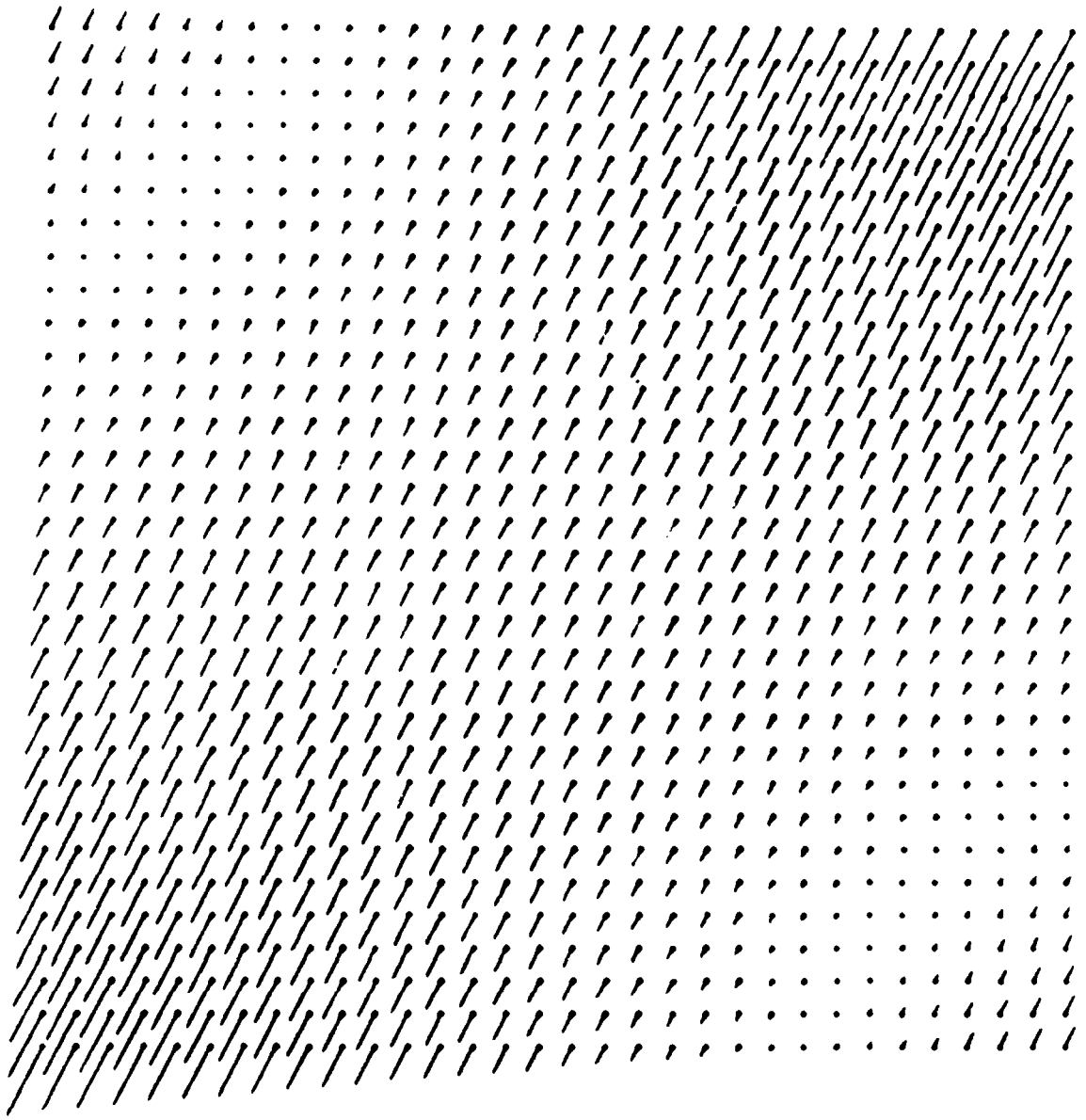


Figure 3. A motion field with three rigid body motion interpretations.

Appendix

The eigenvalues of

$$\mathbf{D} = \begin{pmatrix} d_{zz} & d_{zy} & d_z \\ d_{zy} & d_{yy} & d_y \\ d_z & d_y & 2 \end{pmatrix}$$

are the solutions of the characteristic equation

$$\lambda^3 + a_2\lambda^2 + a_1\lambda + a_0 = 0,$$

where

$$\begin{aligned} a_2 &= -(2 + d_{zz} + d_{yy}), \\ a_1 &= 2(d_{zz} + d_{yy}) + (d_{zz}d_{yy} - d_{zy}^2) - (d_z^2 + d_y^2), \\ a_0 &= -2(d_{zz}d_{yy} - d_{zy}^2) + (d_{zz}d_y^2 - 2d_{zy}d_zd_y + d_{yy}d_z^2). \end{aligned}$$

One of the conditions in the case of an ambiguity is that the mean-scaled curvature is unity; that is, $\frac{1}{2}(d_{zz} + d_{yy}) = 1$. Therefore, we have

$$\begin{aligned} a_2 &= -4, \\ a_1 &= 4 + (d_{zz}d_{yy} - d_{zy}^2) - (d_z^2 + d_y^2), \\ a_0 &= -2(d_{zz}d_{yy} - d_{zy}^2) + (d_{zz}d_y^2 - 2d_{zy}d_zd_y + d_{yy}d_z^2). \end{aligned}$$

Another condition is that either the surface gradient should vanish ($d_z = d_y = 0$) or the optical axis, the surface normal and one of the asymptotic lines of the surface are in the same plane. In either case, we obtain

$$d_{zz}d_y^2 - 2d_{zy}d_zd_y + d_{yy}d_z^2 = 0.$$

Using this in the earlier equations and simplifying the results, we can show that the characteristic equation simplifies to

$$(\lambda - 2)(\lambda^2 - 2\lambda - (p + q)) = 0,$$

where

$$p = d_z^2 + d_y^2 \quad \text{and} \quad q = d_{zy}^2 - d_{zz}d_{yy}.$$

Therefore, the eigenvalues of **D**, in the ascending order, are

$$\lambda_- = 1 - \sqrt{1 + p + q}, \quad \lambda_o = 2, \quad \text{and} \quad \lambda_+ = 1 + \sqrt{1 + p + q}.$$

References

Adiv, G. (1985) "Inherent Ambiguities in Recovering 3-D Motion and Structure From a Noisy Flow Field," *Proceedings of the IEEE Computer Vision and Pattern Recognition Conference*, San Francisco, June.

Fang, J.Q., & T.S. Huang (1984) "Solving Three-Dimensional Small-Rotation Motion Equations: Uniqueness, Algorithms, and Numerical Results," *Computer Vision, Graphics, and Image Processing*, Vol. 26, No 2, May.

Hay, J.C. (1966) "Optical Motion and Space Perception: An Extension of Gibson's Analysis," *Psychology Review*, Vol. 73, No 6.

Horn, B.K.P. (1986a) *Robot Vision*, MIT Press, Cambridge and McGraw-Hill, New York.

Horn, B.K.P. (1986b) "Rareness of Ambiguities in Motion Vision," submitted to *International Journal of Computer Vision.*

Jerian, C., & R. Jain (1984) "Determining Motion Parameters for Scenes with Translation and Rotation," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 6, No 4, July.

Le Guilloux, Y. (1986) "Structure from Motion, Acceleration and Taylor Series," *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, Miami, FL, June.

Longuet-Higgins, H.C., & K. Prazdny (1980) "The Interpretation of a Moving Retinal Image," *Proceedings of the Royal Society of London*, Series B, Vol. 208.

Longuet-Higgins, H.C. (1984) "The Visual Ambiguity of a Moving Plane," *Proceedings of the Royal Society of London*, Series B, Vol. 223.

Maybank, S.J. (1984) "The Angular Velocity Associated with the Optical Flow Field due to a Single Moving Rigid Plane," *Proceedings of the Sixth European Conference in Artificial Intelligence*, Pisa, Italy, September.

Negandaripour, S., & B.K.P. Horn (1985) "Direct Passive Navigation," AI Memo 821, MIT Artificial Intelligence Laboratory, Cambridge, MA, February.

Negandaripour, S., & B.K.P. Horn (1987) "Direct Passive Navigation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 9, No 1, January.

Negahdaripour, S., & A. Yuille (1986) "Direct Passive Navigation: Analytical Solution for Quadratic Surfaces," unpublished work.

Negahdaripour, S., (1986) "Direct Passive Navigation," Ph.D. Thesis, Department of Mechanical Engineering, MIT, Cambridge, MA, November.

Tsai, R.Y., T.S. Huang, & W.L. Zhu (1982) " Estimating Three-Dimensional Motion Parameters of a Rigid Planar Patch, II: Singular Value Decomposition," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, Vol. ASSP-30, No 4, August

Tsai, R.Y., & T.S. Huang (1984) "Uniqueness and Estimation of Three-Dimensional Motion Parameters of Rigid Objects with Curved Surfaces," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 6, No 1, January.

Ullman, S. (1983) "Computational Studies in the Interpretation of Structure from Motion: Summary and Extension," AI Memo 706, MIT AI Lab, Cambridge, MA, March.

Waxman, A.M., & S. Ullman (1985) "Surface Structure and Three-Dimensional Motion from Image Flow Kinematics," *The International Journal of Robotics Research*, Vol. 4, No 3, Fall.

Waxman, A.M., B. Kamgar-Parsi, & M. Subbarao (1986) "Closed-Form Solutions to Image Flow Equations for 3-D Structure and Motion," CAR-TR-190, Computer Vision Laboratory, Center for Automation Research, University of Maryland, College Park, MD, February.

END

7-87

DTIC