CAR-TR-160
CS-TR-1576

DAAK70-83-K-0018
October 1984
Revised June 1985

ANALYSIS OF STRUCTURE AND MOTION FROM OPTICAL FLOW
PART I: ORTHOGRAPHIC PROJECTION

Ken-ichi Kanatani*
Department of Computer Science
Gunma University
Kiryu, Gunma 376
Japan

COMPUTER VISION LABORATORY

CENTER FOR AUTOMATION RESEARCH

UNIVERSITY OF MARYLAND
COLLEGE PARK, MARYLAND
20742

CAR-TR-160
CS-TR-1576

DAAK70-83-K-0018
October 1984
Revised June 1985

ANALYSIS OF STRUCTURE AND MOTION FROM OPTICAL FLOW PART I: ORTHOGRAPHIC PROJECTION

Ken-ichi Kanatani*

Department of Computer Science
Gunma University
Kiryu, Gunma 376
Japan

ABSTRACT

The 3D structure and motion of an object is determined from its optical flow under orthographic projection. First, the image domain is divided into planar or almost planar regions by checking the flow. For each region, parameters of the flow are determined. Transformation rules under coordinate changes and hydrodynamic analogies are also discussed. The 3D structure and motion are determined in explicit forms in terms of irreducible parameters deduced from group representation theory. The solution is not unique, containing an indeterminate scale factor and comprising true and spurious solutions. Their geometrical interpretations are also studied. The spurious solution disappears if two or more regions of the object are observed.

The support of the Defense Advanced Research Projects Agency and the U.S. Army Night Vision and Electro-Optics Laboratory under Contract DAAK70-83-K-0018 (DARPA Order 3206) is gratefully acknowledged.

*Visiting the Center for Automation Research, University of Maryland, during 1985-6.

1. INTRODUCTION

Determination of the 3D structure and motion of an object from its projected 2D images is one of the most important tasks of computer vision. This can be done by introducing various heuristic or a priori "constraints" such as planarity or smoothness of the object and rigidity of the motion coupled with various other sources of information like texture, shading, etc. (e.g., [1]). Basically there are two approaches. One is first to seek "correspondence" of points, i.e., knowledge of which point moves to which one, between two sequential images, resulting in a so-called "optical flow," and various techniques have been tried to detect the optical flow (e.g., see [2 - 6]). The other approach does not use the point correspondence or the optical flow but directly measures some sorts of "features" of the image. The 3D motion is detected from these features and their time changes alone if a particular model is assumed for the object. For example, the 3D motion of a planar surface can be detected from statistical properties of its texture [7] or its contour [8, 9].

In this paper, we take the first approach and assume that an optical flow is already obtained. There have been many studies of schemes computing from a given optical flow the 3D motion of the object or the observer seeing a stationary environment ("egomotion" from the "motion parallax" [10]). Studies of this type have often been associated with the "computational approach" to human perception [2, 11, 12]. Roughly speaking, procedures are classified into two groups; "correspondence-based" approachs and "flow-based" ones. The correspondence-based approach picks up several correspondence pairs out of the optical flow, and subsequent computation is based on their coordinates, assuming no specific model about the object except the rigidity of the motion. Then, the 3D structure and motion are recovered numerically, e.g., see [13 - 15] for orthographic projection and [16 - 23] for central projection. On the other hand, the flow-based one assumes a certain model about the object or the scene and examines the flow pattern, extracting some features like the "focus of expansion" or "vanishing point," taking spatial derivatives, estimating parameters by global model fitting and sometimes employing hydrodynamic analogies, e.g., see [24 - 34] for central projection.

In this paper, we present a flow-based approach under orthographic projection, which is the case when the size of the object is small compared with the distance from the viewer. So far, no attempts have been made of the flow-based approach under orthographic projection. Probably, it has been believed that not enough clues are obtained as to the 3D structure and motion under orthographic projection, since no looming up or shrinking away is caused by motions. It is true that much information is lost by orthographic projection. For instance, we cannot tell the absolute distance from the viewer and whether it is approaching or receding. However, it is possible to extract information about the relative depth, its translation and 3D rotation. The solution is not unique, as was pointed out by Sugihara and Sugie [15]. Yet, we can describe geometrical relationships among those indeterminate solutions. Here, we try to extract here as much knowledge as is possible in "analytical terms" under such "imperfect information." This is very important in practice, because we often fail to observe the effects of central projection because, say, the object is too small or is located far away or the focal length of the camera is not small enough. In these cases, even partial knowledge is useful, for it can be supplemented by other sources of information.

We first divide the image domain into regions which can be regarded as being planar or almost planar. A criterion for it is discussed. Then, we pick up each region and compute its surface orientation and motion. We take a Cartesian xy-coordinate system on the image plane and the z-axis perpendicular to it. However, the choice of the xy-coordinate system is completely arbitrary in principle. Hence, interpretations based on different xy-coordinate systems must coincide. In other words, interpretations of 3D structure and motion must be "invariant." In order to make our scheme invariant, we first study the transformation rules when one coordinate system is replaced by another. This consideration has a practical significance, for it is sometimes convenient to take different coordinate systems for different regions of the same object, taking the coordinate origin in each region, for instance. Then, quantities associated with different regions must be compared after appropriate transformations.

We invoke group representation theory [35 - 38] to extract "irreducible parameters" with respect to coordinate changes. We also make use of "hydrodynamic analogies," viewing the optical flow as if it were a flow of real fluid. Since hydrodynamics is usually described in invariant forms, it gives a clear understanding of the invariant nature of our interpretation. Then, we express the surface orientation and the motion in an "explicit" form, which is made possible by the use of the irreducible parameters. It turns out that there are two types of solutions for each planar region. One is the true one and the other is a "spurious" one, each containing one indeterminate scale parameter. We describe the geometrical relationship between the true and the spurious solutions in an invariant manner. We also show that the spurious solution disappears if two or more different regions of the same object are observed.

Our flow-based approach to reconstruct the 3D structure and motion is a generalization of the correspondence-based approach of Sugihara and Sugie [15], which can be viewed as a special case of ours. In order to apply their correspondence-based approach, the velocity measurement must be accurate. In contrast, the flow-based approach extracts global quantities, which are in general less sensitive to noise and possible misdetection of correspondence. Thus, our flow-based approach bears a practical significance. Moreover, many important observations such as the invariance, transformation rules, hydrodynamic analogies, the "spurious" solution and its geometrical interpretation cannot be easily realized from a correspondence-based approach like that of Sugihara and Sugie [15]. Their reasoning is also insufficient to give the degree of indeterminacy. This is because we solve the problem in analytical terms, while the scheme of Sugihara and Sugie [15] gives solutions only numerically. Another big advantage of the flow-based approach is that the procedure can be applied to the case where no optical flow is available or no correspondence of points is detected. This is because the formulation rests on the "flow parameters" extracted from the optical flow. Sometimes, these parameters can be computed directly from the image sequence itself without detecting point correspondence (Kanatani [7 - 9]).

2. IDENTIFICATION OF OPTICAL FLOW

Suppose a plane is moving in the scene and we are looking at its image on the xy-plane orthographically projected along the z-axis. Let $z = px + qy + r$ be the equation of the plane. The orientation of the plane is specified by the two parameters p and q, which are often referred to as the "gradients" because $p = \partial z/\partial x$ and $q = \partial z/\partial y$. We take (0, 0, r), the intersection between the plane and the z-axis, as a reference point and assume that the reference point is translating with translation velocity (a, b, c) and is rotating with rotation velocity (w1, w2, w3), i.e., rotating by $\mathrm{sqr}((w1)^2 + (w2)^2 + (w3)^2)$ rad/sec screwwise around an axis along (w1, w2, w3) at the reference point (Fig. 1). (Here, sqr(.) stands for the square root.) Since the absolute depth r and the velocity c in the z-direction are indiscernible, our goal is to determine the gradients p and q, the translation velocities a and b and the rotation velocities w1, w2 and w3 by observing the projected image.

If a plane with gradients p, q is moving with translation velocities a, b and with rotation velocities w1, w2, w3 at (0, 0, r), an elementary calculation shows that the x- and y-components of the velocity of a point (x, y, z) on the plane are given by

$$u(x, y) = a + (pw2)x + (qw2 - w3)y, \quad (2.1)$$

$$v(x, y) = b - (pw1 - w3)x - (qw1)y, \quad (2.2)$$

respectively. This is called the "optical flow." As was stated before, we assume that the optical flow is already available at particular feature points. We first try to fit the following form to the observed flow:

$$u(x, y) = a + Ax + By, \quad (2.3)$$

$$v(x, y) = b + Cx + Dy. \quad (2.4)$$

Here, we call parameters a, b, A, B, C, D the "flow parameters." The simplest way may be to use the least square method to minimize

$$M = \iint[(a + Ax + By - u(x, u))^2 + (b + Cx + Dy - v(x, y))^2]dxdy \;. \quad (2.5)$$

From $\partial M/\partial a = 0$, $\partial M/\partial b = 0$, $\partial M/\partial A = 0$, $\partial M/\partial B = 0$, $\partial M/\partial C = 0$, $\partial M/\partial D = 0$, we obtain the flowing set of equations called the "normal equations:"

$$(\iint dxdy)a + (\iint xdxdy)A + (\iint ydxdy)B = \iint u(x, y)dxdy,$$
$$(\iint dxdy)b + (\iint xdxdy)C + (\iint ydxdy)D = \iint v(x, y)dxdy,$$
$$(\iint xdxdy)a + (\iint x^2dxdy)A + (\iint xydxdy)B = \iint xu(x, y)dxdy,$$
$$(\iint ydxdy)a + (\iint xydxdy)A + (\iint y^2dxdy)B = \iint yu(x, y)dxdy, \quad (2.6)$$
$$(\iint xdxdy)b + (\iint x^2dxdy)C + (\iint xydxdy)D = \iint xv(x, y)dxdy,$$
$$(\iint ydxdy)b + (\iint xydxdy)C + (\iint y^2dxdy)D = \iint yv(x, y)dxdy.$$

From these, we can generally determine estimates of the flow parameters a, b, A, B, C, D. In particular, if the domain of observation is symmetric with respect to both the x- and the y-axes and has area S, the estimates are explicitly given by

$$a = \iint u(x, y)dxdy/S, \qquad b = \iint v(x, y)dxdy/S$$
$$A = \iint xu(x, y)dxdy/\iint x^2dxdy, \qquad B = \iint yu(x, y)dxdy/\iint y^2dxdy, \quad (2.7)$$
$$C = \iint xv(x, y)dxdy/\iint x^2dxdy, \qquad D = \iint yv(x, y)dxdy/\iint y^2dxdy.$$

Of course, we must replace the integrals by appropriate summations, since the velocity is observed only at a finite number of points.

For the estimates a, b, A, B, C, D, the "residual" M of (2.5) becomes

$$M = \iint u^2dxdy - ((\iint udxdy)a + (\iint xudxdy)A + (\iint yudxdy)B)$$
$$+ \iint v^2dxdy - ((\iint vdxdy)b + (\iint xvdxdy)C + (\iint yvdxdy)D). \quad (2.8)$$

An optical flow can be that of a plane motion if and only if $M = 0$. From this, we obtain a "criterion of planarity," taking account of possible errors. Namely, we can view a computed optical flow as that of plane motion if M given by (2.8) is less than a certain threshold. (Note that the optical flow is uniquely determined if velocities are given at at least three feature points, cf. Section 8.) This suggests the following procedure. Namely, starting from three or more feature points where the residual is very small, we can add to them other points from around one by one, each time recomputing the flow parameters and checking the residual M, as long as the residual M does not exceed a prescribed threshold. Then, we end up with a region which is almost planar. The procedure is repeated for the rest of the regions, and the image is roughly decomposed into almost planar small regions. (Exact boundaries of these small regions are not necessary. They are reconstructed by the procedures of Section 7.)

In the following, we assume that a given optical flow can be regarded as that of plane motion and the flow parameters a, b, A, B, C, D are already estimated. Thus, the flow parameters a, b, A, B, C, D are the only available data. Hence, any information must be expressed in terms of a, b, A, B, C, D. We also use the matrix form

$$u = a + Ar, \tag{2.9}$$

where $a = (a, b)$, $r = (x, y)$ and

$$A = \begin{bmatrix} A & B \\ C & D \end{bmatrix}. \tag{2.10}$$

We do not make particular distinction between a column vector and a row vector because we can easily tell which is which.

Note that we need not necessarily know the optical flow or detect point correspondence, because all we need is the the flow parameters, not the flow itself. For example, if we use the methods of Kanatani [7 - 9], the flow parameters are determined directly without knowing correspondence.

On the other hand, consider an extreme case where each almost-planar patch consists only of three points. This amounts to polyhedral approximation of the object. Then, our flow-based approach is equivalent to the correspondence-based approach (cf. Example 3 in Section 8). In other words, our flow-based approach is a generalization of the correspondence-based approach, including it as an extreme case.

3. COORDINATE CHANGE AND INVARIANCE

As was stated in the previous section, all information we get is the flow parameters a, b, A, B, C, D alone, from which we want to compute the gradients p, q and the motion parameters a, b, w1, w2, w3. However, all of these parameters are defined with respect to a given Cartesian xy-coordinate system on the image plane, and hence, if we use another xy-coordinate system, we obtain different values. Since we can take an arbitrary Cartesian coordinate system on the image plane, the interpretation of the structure and motion must be "invariant" with respect to coordinate systems. To be precise, structure and motion parameters computed from different values of the flow parameters due to coordinate change must coincide with those obtained by transforming the original interpretation accordingly. In short, interpretation must "commute" with coordinate change (Fig. 2). In the following, we consider a group-theoretical way to exploit this invariance. Although it may seem too pompous to invoke a sophisticated mathematics for a very simple problem like the present one, this consideration provides us with a transparent viewpoint, an elegant closed formulation and also a strong guidance to cope with more complicated problems like images under central projection, etc. Besides, it is sometimes convenient to use different coordinate systems for different regions. Then, we need the transformation rules for coordinate changes in order to combine the results.

Suppose we take an x'y'-coordinate system by rotating the original xy-coordinate system by angle t counterclockwise and by translating it by (h1, h2) (Fig. 3). The origin of the x'y'-coordinate system O' is at (h1, h2) of the xy-coordinate system. If (x, y) are the coordinates with respect to the former coordinate system and (x',y') with respect to the latter, their relationship is given by

$$r' = R(r - h), \tag{3.1}$$

where r = (x, y), r' = (x', y'), h = (h1, h2) and

$$R = \begin{bmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{bmatrix}. \tag{3.2}$$

The coordinate transformations of this type, which we denote by (R, h), form a group known as the 2D "Euclidean transformation group." Consider another transformation $r'' = R'(r' - h')$. If we operate (3.1) followed by this, we obtain $r'' = R'(R(r - h) - h') = R'R(r - (R^T h' + h))$, which is transformation $(R'R, R^T h' + h)$, where T designates the transpose. (Note that R is an orthogonal matrix so that $R^{-1} = R^T$.) In other words, the group operation is given by

$$(R', h')(R, h) = (R'R, R^T h' + h). \tag{3.3}$$

If u, v are the velocities with respect to the original coordinate system and u', v' with respect to the new one, it is easy to see

$$u' = Ru, \tag{3.4}$$

where u = (u, v) and u' = (u', v'). In other words, u and v are transformed "as a vector." Next, consider the gradients p, q. Since $p = \partial z/\partial x$ and $q = \partial z/\partial y$, they must be transformed as a vector, i.e.,

$$p' = Rp, \tag{3.5}$$

where p = (p, q) and p' = (p', q'). Consider the rotation velocities w1, w2, w3. Obviously, w3 is an invariant under coordinate changes or a "scalar," and w1, w2 are transformed as a vector, i.e.,

$$w' = Rw, \qquad (3.6)$$

where $w = (w1, w2)$ and $w' = (w1', w2')$, because they are projections of a 3D vector. The fact that (p, q) and (w1, w2) are transformed as vectors implies that they have invariant meanings. In fact, (p, q) indicates the direction of "maximum gradient" or the "steepest ascent." The magnitude of (p, q) is the incline along that direction. On the other hand, (w1, w2) indicates the "axis of rotation" in the xy-plane. In other words, if rotation w3 around the z-axis is not considered, the rotation is realized by rotating the plane around that axis by $sqr(w1^2 + w2^2)$ (rad/sec) screwwise.

On the other hand, a, b are "not" transformed as a vector. In fact, the rotation velocities w1, w2, w3 are defined at the reference point (0, 0, r), and this rotation induces translation velocity (w1, w2, w3) × (x, y, z - r) at point (x, y, z). Since the new reference point goes to (h1, h2, r + ph1 + qh2), the translational velocity induced there is

$$\begin{bmatrix} w1 \\ w2 \\ w3 \end{bmatrix} \times \begin{bmatrix} h1 \\ h2 \\ ph1 + qh2 \end{bmatrix} = \begin{bmatrix} pw2 & qw2 - w3 \\ - pw1 + w3 & - qw1 \\ w2 & - w1 \end{bmatrix} \begin{bmatrix} h1 \\ h2 \end{bmatrix}. \qquad (3.7)$$

Comparing this with eqns (2.1) and (2.2), we obtain the transformation rule

$$a' = R(a + Ah), \qquad (3.8)$$

where $a' = (a', b')$.

Finally, note that if parameters s1, s2 are transformed as a vector, the complex number s1 + is2 is transformed with "weight" - 1, i.e., by multiplication of e(- t). (Here, i is the imaginary unit, and we use abbreviation e(.) for exp(i.).) Hence, if we put

$$P = p + iq, \qquad W = w1 - w2, \qquad (3.9)$$

we have

$$P' = e(-t)P, \qquad W' = e(-t)W. \tag{3.10}$$

The significance of this consideration will become clear in subsequent sections. In the following, we regard the complex numbers P and W as also 2D vectors.

4. IRREDUCIBLE PARAMETERS AND THEIR INVARIANCE

As was stated earlier, the optical flow is first expressed by eqn (2.9). After the coordinate change (3.1), it becomes $u' = Ru = R(a + Ar) = R(a + A(R^T r' + h)) = (Ra + RAh) + RAR^T r'$. Comparing this with eqn (2.9), we find that the new flow parameters a', b', A', B', C', D' are given by

$$a' = Ra + RAh, \qquad A' = RAR^T, \tag{4.1}$$

where a' = (a', b') and A' is the matrix of A', B', C', D' like eqn (2.10). Here, we again obtain eqn (3.8), the transformation of a, b. We can also see that A is transformed "as a tensor."

Apparently, the transformation of the flow parameters from a, b, A, B, C, D to a', b', A', B', C', D' is a linear transformation, which we denote by rep[(R, h)]. This gives a six dimensional "representation" of the 2D Euclidean group. Indeed, if we transform a', b', A', B', C', D' into a", b", A", B", C", D" by transformation (R', h') we get $a'' = R'a' + R'A'h' = R'(Ra + RAh) + R'(RAR^T)h' = (R'R)a + (R'R)A(R^T h' + h)$ and $A'' = R'A'R'^T = R'RAR^T R'^T = (R'R)A(R'R)^T$. In view of eqns (4.1), this composite transformation is $\mathrm{rep}[(R'R, R^T h' + h)]$, which is equal to rep[(R', h')(R, h)] by eqn (3.3). Hence,

$$\mathrm{rep}[(R', h')]\mathrm{rep}[(R, h)] = \mathrm{rep}[(R', h')(R, h)], \tag{4.2}$$

and rep[(R, h)] is really a representation.

From eqns (4.1), we find that this representation is "reducible" into a and A because parameters A are transformed among themselves. (However, this representation is not "completely reducible," since the remaining parameters a are not tranformed among themselves.) In other words, A induces a four dimensional representation of the 2D rotation group. This representation is completely reducible because the rotation group is topologically a "compact" group. It is also seen that it is decomposed into "irreducible representations," all of which are one-dimensional, i.e., multiplication of e(nt) with integer n ("weight"). This is a consequence of Schur's lemma and the fact that the 2D rotation group is a "commutative" group (cf. [35, 36]).

As is well known, the decomposed irreducible representations are given by calculating the "character" of the representation. Let us compute the "trace" of the transformation $A' = RAR^T$. If A = 1, B = C = D = 0, then $A' = \cos^2 t$, which is the first diagonal element. Likewise, we can see that all the diagonal elements are $\cos^2 t$. Hence, the character is $4\cos^2 t = 2 + 2\cos 2t = 2 + e(2t) + e(-2t)$. Thus, there are two "absolute invariants" (weight 0) and two "relative invariants" of weight 2 and - 2. The last two are mutual complex conjugate because the representation space is real.

The process of decomposition is given by Weyl's principle of exploiting the symmetry of tensor A (cf. [37, 38]). It is decomposed into symmetric and antisymmetric parts as

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} A & (B+C)/2 \\ (B+C)/2 & D \end{bmatrix} + \begin{bmatrix} 0 & -(C-B)/2 \\ (C-B)/2 & 0 \end{bmatrix}. \quad (4.3)$$

This decomposition is "invariant," i.e., independent of the coordinate change. Hence, A, B + C, D are transformed among themselves, and R = C - B is an absolute invariant. The symmetric part is further decomposed into the "scalar" and "deviator" parts as

$$\begin{bmatrix} A & (B+C)/2 \\ (B+C)/2 & D \end{bmatrix} = \frac{A+D}{2}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} (A-D)/2 & (B+C)/2 \\ (B+C)/2 & -(A-D)/2 \end{bmatrix}, \quad (4.4)$$

which is also invariant, and $T = A + D$ is an absolute invariant. The set of $A - D$, $B + C$ is transformed by

$$\begin{bmatrix} A' - D' \\ B' + C' \end{bmatrix} = \begin{bmatrix} \cos 2t & \sin 2t \\ -\sin 2t & \cos 2t \end{bmatrix} \begin{bmatrix} A - D \\ B + C \end{bmatrix}. \quad (4.5)$$

The decomposition so far corresponds to the expression $2 + 2\cos 2t$ of the character. If we consider complex parameter $S = (A - D) + i(B + C)$, it is transformed with weight -2, i.e., $S' = e(-2t)S$. We call the three invariant parameters

$$T = A + D, \qquad R = C - B, \qquad S = (A - D) + i(B + C), \quad (4.6)$$

the "irreducible parameters" of the optical flow.

Lastly, let us consider "hydrodynamic analogies." If the optical flow (4.1) is regarded as the flow of real fluid, each of the above invariants has a physical meaning. For example, $T = \partial u/\partial x + \partial v/\partial y$ is the "divergence," and the first term of the right-hand side of eqn (4.4) describes a flow like Fig. 4. Similarly, $R = \partial v/\partial x - \partial u/\partial y$ is the "rotation" or "vorticity" of the flow, and the second term of the right-hand side of eqn (4.3) describes the flow of Fig. 5 (cf. [7, 25, 26]). The second term of the right-hand side of eqn (4.4) describes a "pure shear flow" like Fig. 6. Consider the polar representation abs(S)e(arg(S)) of S. Since S rotates by 2t clockwise around the origin on the complex plane when the xy-coordinate system is rotated by t counterclockwise, $Q1 = e(\arg(S)/2)$ and $Q2 = iQ1$ both rotate by t clockwise. This means that they are transformed "as vectors" with weight -1, and hence their orientations have invariant meanings. As a matter of fact, Q1 and Q2 represent the directions of "maximum extension" and "maximum compression," respectively, in hydrodynamics. This becomes clear if we rotate the coordinate system by arg(S)/2 counterclockwise so that Q1 and Q2 coincide with the x- and the y-axis, respectively. Then, the last term on the right-hand side of eqn (4.4), which describes the "pure shear flow," is diagonalized as

$$\mathrm{abs}(S)\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}/2, \tag{4.7}$$

which is the "canonical form" of the pure shear flow (Fig. 7). The orientations of Q1 and Q2 are called the "principal axes" of the flow. The magnitude $\mathrm{abs}(S) = \mathrm{sqr}((A - D)^2 + (B + C)^2)$ is an absolute invariant and is called the "shear strength" in hydrodynamics.

NOTE. An easiest way to see that quantities of eqns (4.6) are really invariants is to consider the "infinitesimal transformations." Let d denote differentiation with respect to t at t = 0. From $A' = RAR^T$, we immediately obtain dA = (dR)A - A(dR), where

$$dR = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}. \tag{4.8}$$

Hence, we obtain

$$dA = B + C, \quad dB = -A + D, \quad dC = -A + D, \quad dD = -B - C, \tag{4.9}$$

from which results

$$dT = 0, \qquad dR = 0, \qquad dS = -2idS. \tag{4.10}$$

Thus, we can confirm that T and R are really absolute invariants while S is a relative invariant of weight - 2.

5. DETERMINATION OF THE ROTATION AROUND THE z-AXIS

Now that we have prepared necessary mathematical preliminaries, we proceed to determining the surface and motion parameters. Since a, b are directly obtained, we only have to compute p, q, w1, w2, w3 from A, B, C,

D. Comparing eqns (2.1) and (2.2) with eqn (2.9), we have a set of equations to solve as follows:

$$A = pw2, \quad B = qw2 - w3, \quad C = - pw1 + w3, \quad D = - qw1. \tag{5.1}$$

In terms of the irreducible parameters, eqns (5.1) become

$$T = pw2 - qw1, \qquad R = 2w3 - pw1 - qw2, \tag{5.2}$$

$$S = pw2 + qw1 + i(qw2 - pw1). \tag{5.3}$$

Eqns (5.2) are combined together if we consider a complex number R + iT, which is also an absolute invariant. We get

$$R + iT = 2w3 - pw1 - qw2 + i(pw2 - qw1). \tag{5.4}$$

Hence, the given equations (5.1) are equivalent to eqns (5.4) and (5.3). If we use the complex forms of eqns (3.9) for p, q, w1, w2, the right hand sides of eqns (5.4) and (5.3) become 2w3 - PW* and - iPW, respectively, where * designates the complex conjugate. Hence, solving eqns (5.1) is equivalent to solving

$$PW^* = 2w3 - (R + iT), \tag{5.5}$$

$$PW = iS, \tag{5.6}$$

with P, W and w3 as unknowns. Note that P is of weight - 1 and W* is of weight 1, so that PW* is of weight 0 or an absolute invariant. Hence, both sides of eqn (5.5) are an absolute invariant (weight 0). Likewise, both sides of eqn (5.6) are of weight 2.

First, consider w3. It is determined from the fact that the left hand sides of eqns (5.5) and (5.6) have the same magnitude. Hence, we get (2w3 - (R + iT))(2w3 - (R - iT)) = SS*, so that w3 is given as a root of the quadratic equation

$$X^2 - RX + (T^2 + R^2 - SS^*)/4 = 0. \tag{5.7}$$

Since this is of an absolutely invariant form (note SS* is an absolute invariant), the solution w3 is a scalar as is expected. Eqn (5.7) has two roots

$$X = (R \pm \mathrm{sqr}(SS^* - T^2))/2. \tag{5.8}$$

In order that the solutions be real, the discriminant must be non-negative. Namely,

$$\mathrm{abs}(T) \leq \mathrm{abs}(S). \tag{5.9}$$

In terms of hydrodynamic analogies:

<u>LEMMA</u> 1. The magnitude of divergence should not be greater than the shear strength.

Here, we have obtained a "criterion of rigidity." Namely, if the observed values of A, B, C, D do not satisfy inequality (5.9), the flow cannot be interpreted as that of rigid plane motion.

If inequality (5.9) is satisfied, eqn (5.8) gives two real roots. The two solutions are

$$X = w3 \quad \text{and} \quad w3 - (pw1 + qw2). \tag{5.10}$$

This can be checked by substituting eqns (5.1) in eqn (5.7). We get $X^2 - (2w3 - pw1 - qw2)X + w3(w3 - (pw1 + qw2)) = 0$, or $(X - w3)(X - w3 + pw1 + qw2) = 0$. Thus, one of the two roots gives the true solution while the other gives a "spurious solution," and we cannot tell one from the other for a given optical flow. Then, as we show in the next section, p, q, w1, w2 are determined for each of these solutions, resulting in two types of solutions, the "true" and the "spurious" one. However, the spurious solution disappears if two plane faces of the same object are observed. This will be discussed later. Note that pw1 + qw2 is a scalar because it is the inner product of two "vectors" P = p + iq and W = w1 + iw2. Also note that we do not have the spurious solution only when the equality of (5.10) holds, in which case pw1 + qw2 = 0, or P = p + iq and W = w1 + iw2 are mutually "orthogonal."

6. DETERMINATION OF SURFACE ORIENTATION AND ROTATION

All the equations to be solved for $P = p + iq$ and $W = w1 + iw2$ are eqns (5.5) and (5.6). It is immediately observed that the magnitude of W cannot be determined uniquely, since W multiplied by a scale factor together with P divided by that factor also satisfies the original equation. Hence, we can take the magnitude $k = abs(W)$ as an indeterminate scale factor. Of course, we could take w1 as an indeterminate scale factor and express the rest in terms of w1, or we could take w2, or $(w1 + w2)/2$, etc. However, these quantities are not invariants and hence the interpretations in terms of them do not have invariant meanings, while k is an invariant and hence leads to invariant geometrical interpretations.

Now, take the ratio of eqn (5.6) to eqn (5.5). We get

$$W/W^* = iS/(2w3 - (R + iT)). \qquad (6.1)$$

Similarly, if we take the ratio of eqn (5.6) to the complex conjugate of eqn (5.5), we get

$$P/P^* = iS/(2w3 - (R - iT)). \qquad (6.2)$$

The left hand sides of eqns (6.1) and (6.2) are $e(2arg(W))$ and $e(2arg(P))$, respectively. From eqn (6.1), we conclude that $2arg(W) = \pi/2 + arg(S) - arg(2w3 - (R + iT))$. There exist two values for $arg(W)$ mutually opposite with respect to the origin. However, we can pick up one of them arbitrarily, say

$$arg(W) = \pi/4 + arg(S)/2 - arg(2w3 - (R + iT))/2, \qquad (6.3)$$

if we allow the scale factor k to be negative. This does not lose the uniqueness of the expression $W = ke(arg(W))$. Thus, we have completely determined w1 and w2, since the scale factor k is an essential indeterminate. Namely,

$$W = ke(\pi/4 + arg(S)/2 - arg(2w3 - (R + iT))/2). \qquad (6.4)$$

Now that we have obtained W, the remaining P is determined from eqn

(5.6) by

$$P = iS/W$$

$$= Se(\pi/4 - arg(S)/2 + arg(2w3 - (R + iT))/2)/k. \qquad (6.5)$$

Note that S is of weight - 2, while W is of weight - 1. Hence, P is of weight - 1, i.e., transformed as a vector, as expected. Thus, we obtain

THEOREM 1.

$$w3 = (R \pm sqr(SS^* - T^2))/2, \qquad (5.8)$$

$$W = ke(\pi/4 + arg(S)/2 - arg(2w3 - (R + iT))/2), \qquad (6.4)$$

$$P = Se(\pi/4 - arg(S)/2 + arg(2w3 - (R + iT))/2)/k, \qquad (6.5)$$

where k is an arbitrary real number.

On the other hand, we find, from eqn (6.5), that

$$arg(P) = arg(S) - arg(W) \pm \pi/2, \qquad (6.6)$$

where the double sign corresponds to the sign of the scale factor k. Therefore, we see that

$$(arg(P) + arg(W))/2 = arg(S)/2 \pm \pi/4. \qquad (6.7)$$

This implies a simple interpretation in hydrodynamic analogies. Recall that arg(S)/2 is the direction of maximum extension. Hence, the orientations designated by eqn (6.7) bisect the angle made by the directions of maximum extension and maximum compression. These orientations are known as the directions of "maximum shearing," because the viscosity becomes maximum along these directions. Thus, we conclude:

COROLLARY 1. The orientations of $P = p + iq$ and $W = w1 + iw2$ are symmetric with respect to the direction of maximum shearing.

This statement, of course, has an invariant meaning irrespective of the choice of the coordinate system.

So far, we have assumed that w3 is the true rotation velocity. Suppose it is the spurious one w3 - (pw1 + qw2). If we replace w3 in eqn (6.1) by w3 - (pw1 + qw2), eqn (6.1) becomes

$$W^*/W = - iS/(2w3 - (R - iT)). \qquad (6.8)$$

Comparing this with eqn (6.2), we see that spurious 2arg(W) is opposite to true 2arg(P). In other words, the orientation of spurious w is orthogonal to that of true p and we cannot say any more about its orientation because the magnitude k of w is indeterminate including the signature. If we obtain spurious W by eqn (6.4), spurious P is again determined by eqn (6.5). It can be immediately seen that the orientation of spurious p is orthogonal to that of true w, and the above observation is still valid for spurious p and w. At the same time, we obeserve the following with respect to the true and spurious solutions.

COROLLARY 2. The orientations of true and spurious W are symmetric with respect to the principal axes of the flow, and so are the orientations of true and spurious P.

This statement also has an invariant meaning.

EXAMPLE 1. Consider the flow of Fig. 8. The flow parameters are

a = 0.1, b = 0.1, A = 0.0873, B = - 0.2269, C = 0.0873, D = 0.0524,

and hence

T = 0.1397, R = 0.3142, S = 0.0349 - 0.1396i.

Since abs(S) = 0.1439, we see abs(T) < abs(S) and hence the flow can be regarded as that of rigid motion. First, from eqn (5.5), we obtain w3 = 10, 8 (deg/sec). The remaining components of rotation and the gradients are

$W1 = (0.7061 + 0.7081i)k$ (rad/sec), $P1 = (0.1233 - 0.1484i)/k$,

$W2 = (0.5157 + 0.8568i)k$ (rad/sec), $P2 = (0.1019 - 0.1016i)/k$,

where k is the indeterminate scale factor. One set of these is the true solution and the other is the spurious one. Fig. 9 illustrates these when $k = 0.5$. There, the principal axes and the directions of maximum shearing are also indicated.

7. ADJACENT TWO OPTICAL FLOWS

Now, we consider two regions of an image which have different optical flows, i.e., different flow parameters a, b, A, B, C, D. Obviously, this arises if the object is a polyhedron. However, the object can have a smoothly varying surface, in which case we divide the surface image into small regions each of which can be regarded as almost planar, say according to the criterion of planarity discussed in Section 2.

Let a, b, A, B, C, D be the parameters of one region and a', b', A', B', C', D' those of the other, and assume that these two sets are not identical. If the two regions are planar and adjacent to each other, their intersection must be a straight line, on which u, v must be continuous, i.e.,

$$[a] + [A]x + [B]y = 0, \qquad [b] + [C]x + [D]y = 0, \qquad (7.1)$$

where [ ] designates the difference, e.g., $[a] = a' - a$. The necessary and sufficient condition that eqns (7.1) define one and the same line is

$$[a] : [b] = [A] : [C] = [B] : [D]. \qquad (7.2)$$

Eqn (7.2) gives a "criterion of adjacency." In other words, if eqn (7.2) is not satisfied (within a certain error), the two regions cannot be regarded as being adjacent to each other. If eqn (7.2) is satisfied, eqns (7.1) define the "intersection line." If the two regions are two adjacent faces of a polyhedron, the intersection line may be directly observed. However, even when intersection lines are missing due to noise or some technical difficulties, we can still recover them once we can successfully estimate the flow parameters on each region, say by eqns (2.6). Moreover, even when the two regions are neighboring "almost planar" patches of a smoothly varying surface, the "hypothetical intersection line" is still defined.

Next, we must check if the two adjacent planar regions are "rigidly" connected. Obviously, a "criterion of rigid adjacency" is given by testing if we can determine common motion parameters a, b, w1, w2, w3. (If the computation is done with respect to different coordinate systems for the two regions, we must compare them after appropriately transforming them as is discussed in Section 3.)

Suppose two regions are images of two planes $z = px + qy + r$ and $z = p'x + q'y + r'$. Since $z = px + qy + r = p'x + q'y + r'$ on the intersection line, its equation on the image plane becomes

$$[p]x + [q]y + [r] = 0. \qquad (7.3)$$

Hence, we see that

<u>LEMMA</u> 2. Vector [P] is perpendicular to the intersection line.

According to Section 5, we can first compute w3 for the two regions separately, ending up with two solutions for each region, the true and the spurious one. If the two regions are rigidly connected, the true one must be common to them, and we can pick up the common one as the true w3. If both the true and spurious solutions are common, we have $pw1 + qw2 = p'w1 + q'w2$ according to eqn (5.10). This means $[p]w1 + [q]w2 = 0$ and hence [P] is perpendicular to $W = w1 + iw2$. Since the intersection line is always perpendicular to [P], W must be parallel to the intersection line. As was pointed in the previous section, we can only determine W's orientation as an undirected axis, and hence we can say that W is determined. Then, we can pick up the correct value of w3 so that eqn (6.1) or (6.2) is satisfied. Once we have determined w3 uniquely, we can compute W for each of the regions. If the two orientations of W do not coincide (within a certain error), the two regions cannot be regarded as being rigidly connected. If they coincide, the scale factor k can be taken to be common to both.

Now, if $y = mx + n$ is the computed intersection line, we must have, in comparison with eqn (7.3), that $[p] : [q] : [r] = m : -1 : n$. Hence,

LEMMA 3. If the intersection line is $y = mx + n$, the equations of the planes are

$$z = px + qy + r, \qquad z = p'x + q'y + (r - [q]n). \tag{7.4}$$

This can be also extend to other regions. Hence, we have established the following fact.

THEOREM 2. The structure and motion of an object are determined from its optical flow under orthographic projection only up to an unknown absolute depth r and an indeterminate scale factor k aside from the translation in the z-direction.

We have also given explicit formulae of computation.

EXAMPLE 2. Consider the flow of Fig. 10. According to the discussion in Section 2, we cannot conclude that this is a flow induced by a motion of a single plane but that this consists of two different flows with flow parameters

$$a = -0.1,\ b = 0.2, \qquad a' = -0.1489,\ b' = 0.2244,$$

$$A = 0.2094,\ B = -0.1047,\ C = 0.0698,\ D = -0.0349$$

$$A' = -0.1396,\ B' = -0.3490,\ C' = 0.2443,\ D' = 0.0873.$$

Eqn (7.2) is satisfied within rounding error, and the intersection of the plane surfaces must be $y = -1.4286x - 0.2$ which is indicated in the figure. From the former flow (upper right) we obtain w3 = 10, 0 (deg/sec) and from the latter (lower left) w3 = 24, 10 (deg/sec). Hence, we conclude that w3 = 10 (deg/sec), and the remaining rotation components become

$$W = (0.4472 + 0.8944i)k \text{ (rad/sec)},$$

where k is the indeterminate scale factor. The gradients are given by

$$P = (0.2341 + 0.0780i)/k, \qquad P' = (-0.1561 - 0.1951i)/k,$$

respectively. They are indicated in Fig. 11 when k = 0.5. Note that [P] is always perpendicular to the intersection line. The equations of the two planes are

$$z = (0.2341x + 0.0780y)/k + r,$$

$$z = (-0.1561x - 0.1951y)/k + (r - 0.0546/k).$$

8. CONCLUDING REMARKS

In this paper, we have exhausted all that can be known given an optical flow under orthographic projection of a rigid object moving in 3D space. First, the image is divided into small regions which are either planar faces or almost-planar patches of a smoothly varying surface. A criterion for it was also discussed. Analyzing each of these regions, we have reached a conclusion that the motion can be recovered up to a common absolute depth and a common scale factor. We also presented explicit formulae of computation. This conclusion was partly pointed out by Sugihara and Sugie [15], who took a correspondence-based approach and considered a finite number of rigidly moving points. They proved that one scale factor must be involved in addition to the indeterminate scale factor, but they failed to show that the solution is unique once the scale factor and the absolute depth are given. Moreover, their algorithm is not perfect and it may produce physically impossible solutions.

On the other hand the correspondence-based approach can be incorporated in our flow-based approach. Consider the case where each almost-planar patch consists of only three points, which amounts to polyhedral approximation of the object. Then, observing the velocities of three points is equivalent to observing the optical flow of the plane spanned by these three points. For example, suppose we measured velocities (u, v), (u', v') and (u", v") at three points (x, y), (x', y') and (x", y"), respectively. Then, the flow parameters a, b, A, B, C, D are determined by solving simultaneous equations

$$\begin{bmatrix} 1 & x & y \\ 1 & x' & y' \\ 1 & x'' & y'' \end{bmatrix} \begin{bmatrix} a \\ A \\ B \end{bmatrix} = \begin{bmatrix} u \\ u' \\ u'' \end{bmatrix}, \tag{8.1}$$

$$\begin{bmatrix} 1 & x & y \\ 1 & x' & y' \\ 1 & x'' & y'' \end{bmatrix} \begin{bmatrix} b \\ C \\ D \end{bmatrix} = \begin{bmatrix} v \\ v' \\ v'' \end{bmatrix}. \tag{8.2}$$

These give a unique solution for a, b, A, B, C, D unless the determinant

$$\begin{vmatrix} 1 & x & y \\ 1 & x' & y' \\ 1 & x'' & y'' \end{vmatrix} \tag{8.3}$$

vanishes, i.e., which is a condition for collinearity of the three points. Hence, if velocities at three non-collinear points are observed, the flow parameters are determined.

EXAMPLE 3. Suppose we measured velocities at

$$r = (0.6, 0.2), \quad r' = (-0.2, -0.4), \quad r'' = (-0.4, 0.8),$$

and observed

$$u = (-0.0416, 0.1052), \; u' = (-0.0975, 0.1767), \; u'' = (0.077, 0.1593),$$

respectively (Fig. 12). From eqns (8.1) and (8.3),

$$a = -0.0486, \qquad b = 0.1523,$$

$$A = -0.0349, \; B = 0.1396, \; C = -0.0698, \; D = -0.0262,$$

$$T = -0.0611, \qquad R = -0.2094, \qquad S = -0.0087 + 0.0698i.$$

The corresponding flow is shown in Fig. 13. By the procedure we showed before, we see that the two solutions are

$$w3 = -5 \text{ (deg/sec)}, \qquad W = (0.4477 + 0.8942i)k \text{ (rad/sec)},$$
$$P = (-0.0390 + 0.0585i)/k,$$
$$w3 = -5 \text{ (deg/sec)}, \qquad W = (0.8319 + 0.5549i)k \text{ (rad/sec)},$$
$$P = (-0.0629 + 0.0315i)/k.$$

Hence, the equation of the plane is

$$z = -0.0390x/k + 0.0585y/k + r \quad \text{or} \quad z = -0.0629x/k + 0.0315y/k + r.$$

Therefore, the z-coordinates of the three points are

$$z = -0.0117/k + r, \qquad z' = -0.0156/k + r, \qquad z'' = 0.0624/k + r$$

or $$z = -0.0314/k + r, \qquad z' = r, \qquad z'' = 0.0504/k + r,$$

where k is the common scale factor and r is the common absolute depth.

Thus, our flow-based approach is a generalization of the correspondence-based approach, including it as a special case. In practice, however, observed velocities of a small number of points are unreliable due to noise and misdetection of point correspondence, as was pointed out earlier. Hence, it seems a wise policy to base the whole computation on the flow parameters a, b, A, B, C, D obtained by the process of taking sums or averages of a number of data, which is less sensitive to local errors in general. Therefore, our present formulation seems preferable for actual processing. Moreover, since our flow-based approach starts with the flow parameters, we do not necessarily have the optical flow or detect point correspondence. For example, if we use the methods of Kanatani [7 - 9], the flow parameters are determined directly without knowing correspondence. As we have seen, indeterminacy is involved in one optical flow. However, the indeterminacy is reduced if a sequence of successive optical flows of the same object, because p, q, r, a, b, c (the velocity along the z-axis if not zero), w1, w2 and w3 cannot evolve arbitrarily. Namely, we have the following "compatibility conditions"

$$dp/dt = pqw1 - (p^2 + 1)w2 - qw3, \tag{8.4}$$

$$dq/dt = (q^2 + 1)w1 - pqw2 + pw3. \tag{8.5}$$

$$dr/ft = c - pa - qb. \tag{8.6}$$

Taking a flow-based approach rather than the correspondence-based approach of Sugihara and Sugie [15] has also led to various other useful concepts and interpretations. Our flow-based analysis enabled us to study the transformation properties under coordinate changes and to express the quantities, formulations and interpretations in frame indifferent manners. The concept of invariance is important not only for consistent and elegant descriptions but also for practical applications, because it allows us to choose a specific coordinate system suitable to each different region. Furthermore, the concept of invariance has naturally lead to hydrodynamic analogies which make clear the intuitive meanings of our interpretations. Taking full advantage of our invariant approach, we expressed the solution in explicit forms, while Sugihara and Sugie [15] gave a scheme only to compute numerically. In the course of our analysis, we showed the existence of the spurious solution and gave its geometrical interpretaion. We also showed that it disappears if flows of two different regions of the same object are observed.

REFERENCES

1. D. H. Ballard and C. M. Brown, Computer Vision, Prentice-Hall, Englewood Cliffs, NJ, 1982.
2. S. Ullman, The Interpretation of Visual Motion, MIT Press, Cambridge, Mass., 1979.
3. R. Jain, Dynamic scene analysis using pixel-based processes, Computer, 14-8 (1981), 12 - 18.
4. W. B. Thompson, Lower-level estimation and interpretation of visual motion, Computer, 14-8 (1981), 20 - 28.
5. B. K. P. Horn and B. G. Schunk, Determining optical flow, Artificial Intelligence, 17 (1981), 185 - 203.
6. J. M. Prager and M. A. Arbib, Computing the optic flow: the MATCH algorithm and prediction, Computer Vision, Graphics, and Image Processing, 24 (1983), 271 - 304.
7. K. Kanatani, Detection of surface orientation and motion from texture by a stereological technique, Artificial Intelligence, 23 (1984), 213 - 237.
8. K. Kanatani, Tracing planar surface motion from projection without knowing the correspondence, Computer Vision, Graphics, and Image Processing, 29 (1985), 1 - 12.
9. K. Kanatani, Detecting the motion of a planar surface by line and surface integrals, Computer Vision, Graphics, and Image Processing, 29 (1985), 13 - 22.

10. D. Regan, K. Beverley and M. Cynader, The visual perception of motion in depth, Scientific American, 241-1 (1979), 122 - 133.

11. S. Ullman, Analysis of visual motion by biological and computer systems, Computer, 14-8 (1981), 57 - 69.

12. D. Marr, Vision: A Computational Investigation into the Human Representation and Processing of Visual Information, Freeman, San Francisco, 1982.

13. S. Ullman, The interpretation of structure from motion, Proc. R. Soc. Lond., B-203 (1979), 405 - 426.

14. D. D. Hoffman and B. E. Flinchbaugh, The interpretation of biological motion, Biol. Cybern., 42 (1982), 195 - 204.

15. K. Sugihara and N. Sugie, Recovery of rigid structure from orthographically projected optical flow, Computer Vision, Graphics, and Image Processing, 27 (1984), 309 - 320.

16. J. W. Roach and J. K. Aggarwal, Detemining the movement of objects from a sequence of images, IEEE Trans. Pattern Anal. Machine Intelligence, PAMI-2-6 (1980), 554 - 562.

17. K. Prazdny, Egomotion and relative depth map from optical flow, Biol. Cybernetics, 36 (1980), 87 - 102.

18. H.-H. Nagel, Representation of moving rigid objects based on visual observations, Computer, 14-8 (1981), 29 - 39.

19. H. C. Longuet-Higgins, A computer algorithm for reconstructing a scene from two projections, Nature, 239-10 (1981), 133 - 135.

20. T. S. Huang and R. Y. Tsai, Image sequence analysis: Motion estimation, Image Sequence Analysis (ed. T. S. Huang), Springer, Berlin/Heidelberg/New York, 1981, pp. 1 - 16.

21. B. L. Yen and T. S. Huang, Determining 3-D motion and surface of a rigid body using the spherical projection, Computer Vision, Graphics,

and Image Processing, 21 (1983), 21 - 32.

22. J.-Q. Fang and T. S. Huang, Solving three-dimensional small-rotation motion equations: Uniqueness, algorithms, and numerical results, Computer Vision, Graphics, and Image Processing, 26 (1984), 183 - 206.

23. R. Y. Tsai and T. S. Huang, Uniqueness and estimation of three-dimensional motion parameters of rigid objects with curved surfaces, IEEE Trans. Pattern Anal. Machine Intelligence, PAMI-6-1 (1984), 13 - 27.

24. D. A. Gordon, Static and dynamic visual fields in human space perception, J. Opt. Soc. Am., 55-10 (1965), 1296 - 1303.

25. 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, 22-9 (1975), 773 - 791.

26. J. J. Koenderink and A. J. van Doorn, Local structure of movement parallax of the plane, J. Opt. Soc. Am., 66-7 (1976) 717 - 723.

27. J. J. Koenderink and A. J. van Doorn, Visual perception of rigidity of solid shape, J. Math. Biol., 3 (1976), 79 - 85.

28. W. F. Clocksin, Perception of surface slant and edge labèls from optical flow: A computational approach, Perception, 9 (1980), 253 - 269.

29. H. C. Longuet-Higgins and K. Prazdny, The interpretation of a moving retinal image, Proc. R. Soc. Lond., B-208 (1980), 385 - 397.

30. J. J. Koenderink and A. J. van Doorn, Exterospecific component of the motion parallax field, J. Opt. Soc. Am., 71-8 (1981), 953 - 957.

31. K. Prazdny, Determining the instantaneous direction of motion from optical flow generated by a curvilinearly moving observer, Computer Graphics and Image Processing, 17 (1981), 238 - 248.

32. J. H. Rieger, Information in optical flows induced by curved paths of observation, J. Opt. Soc. Am., 73-3 (1983), 339 - 344.

33. A. R. Bruss and B. K. P. Horn, Passive navigation, Computer Graphics and Image Processing, 21 (1983), 3 - 20.

34. H. C. Longuet-Higgins, The visual ambiguity of a moving plane, Proc. R. Soc. Lond., B-223 (1984), 165 - 175.

35. I. M. Gel'fand, R. A. Minlos and Z. Ya. Shapiro, Representation of the Rotation and Lorentz Groups and their Applications, Pergamon, Oxford, 1963.

36. M. Hammermesh, Group Theory and its Application to Physical Problems, Addison-Wesley, Reading, Mass., 1964.

37. H. Weyl, Classical Groups, their Invariants and Representations, 2nd ed., Princeton Univ. Press, Princeton, NJ, 1946.

38. H. Weyl, The Theory of Groups and Quantum Mechanics, Dover, New York, 1950.

FIGURE CAPTIONS

Fig. 1 A plane of equation $z = px + qy + r$ is moving with translation velocity (a, b, 0) at (0, 0, r) and rotation velocity (w1, w2, w3) around it. An optical flow is induced on the xy-plane by orhtographic projection along the z-axis with (0, 0, - f) as the viewpoint.

Fig. 2 Interpretation must be "invariant" with respect to coordinate changes, i.e., it must "commute" with coordinate changes.

Fig. 3 A new x'y'-coordinate system is taken by rotating the xy-coordinate system by t cunterclockwise and translating it by (h1, h2). The new origin O' is at (h1, h2) and the new x'-axis makes angle t with the old x-axis.

Fig. 4 Divergent flow.

Fig. 5 Rotational flow.

Fig. 6 Pure shear flow with two principal axes Q1 (maximum extension) and Q2 (maximum compression).

Fig. 7 The canonical form of pure shear flow. The principal axes coincide with the coordinate axes.

Fig. 8 An example of optical flow. The flow parameters are a = 0.1, b = 0.1, A = 0.0873, B = - 0.2269, C = 0.0873, D = 0.0524.

Fig. 9 The result of our analysis of the flow of Fig. 8. Two solutions are possible, the true one and the spurious one. The principal axes and the direction of maximum shearing are also indicated.

Fig. 10 Another example of optical flow. This flow cannot be regarded as that of a single plane. It consists of two planar regions. The dashed line is the intersection line computed from the flow.

Fig. 11 The result of our analysis of the flow of Fig. 10. The spurious solution does not appear.

Fig. 12 Observation of three moving points on the image plane. The optical flow of the plane spanned by these three points are uniquely determined unless the three points are colinear.

Fig. 13 The computed hypothetical optical flow associated with the motion of Fig. 12.

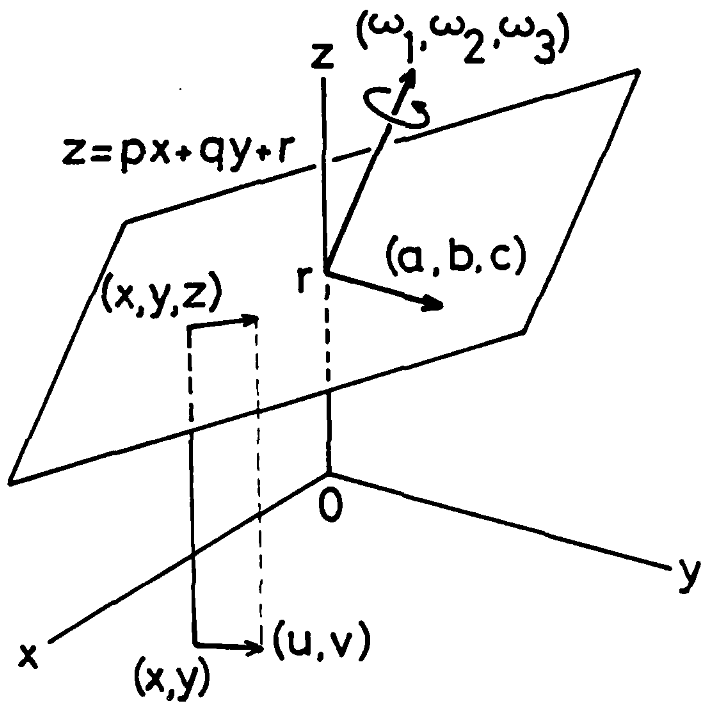

Fig. 1

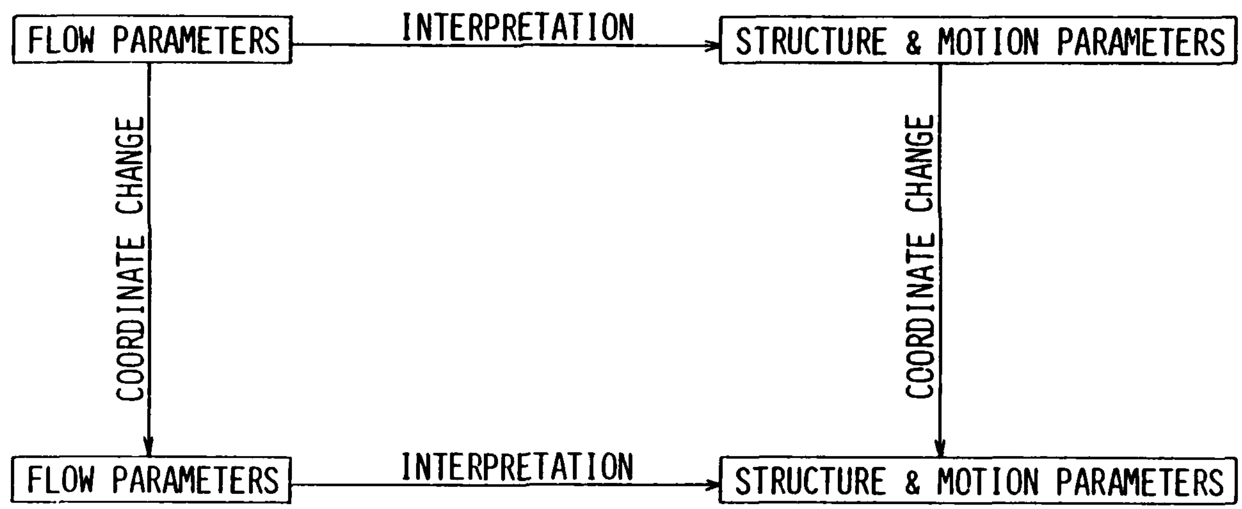
FLOW PARAMETERS
INTERPRETATION
STRUCTURE & MOTION PARAMETERS
COORDINATE CHANGE
COORDINATE CHANGE
FLOW PARAMETERS
INTERPRETATION
STRUCTURE & MOTION PARAMETERS

Fig. 2

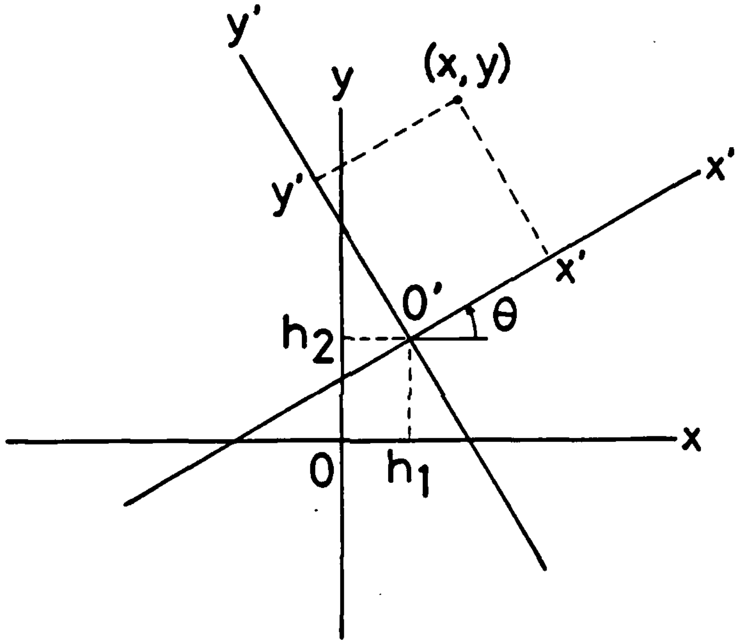

Fig. 3

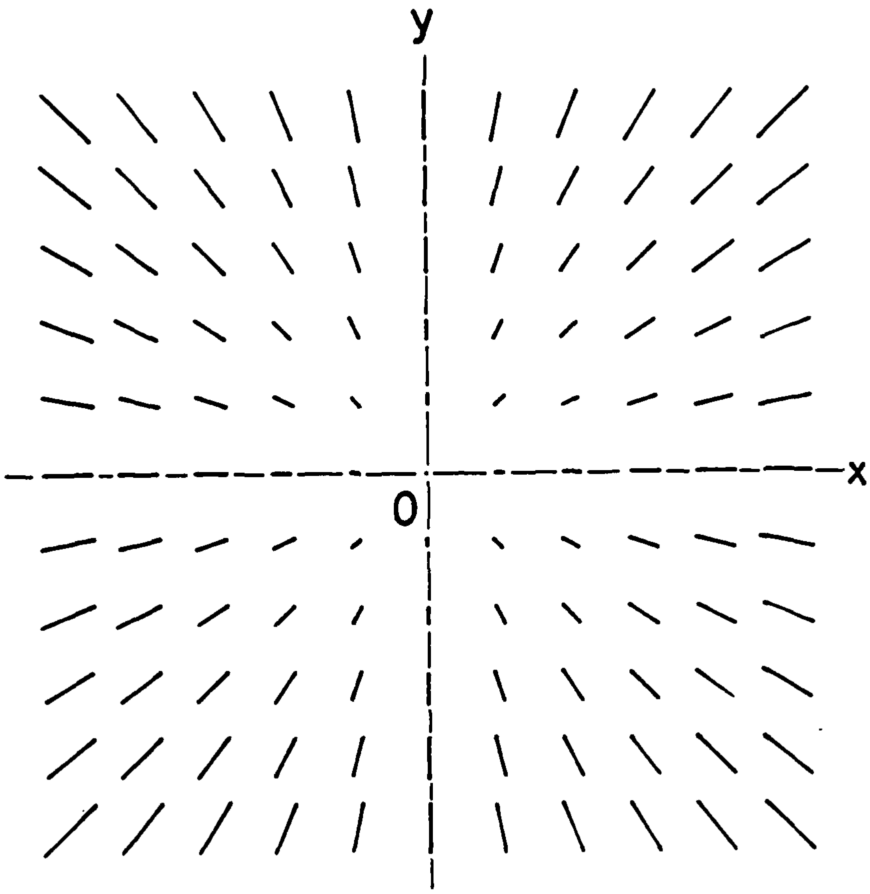

Fig. 4

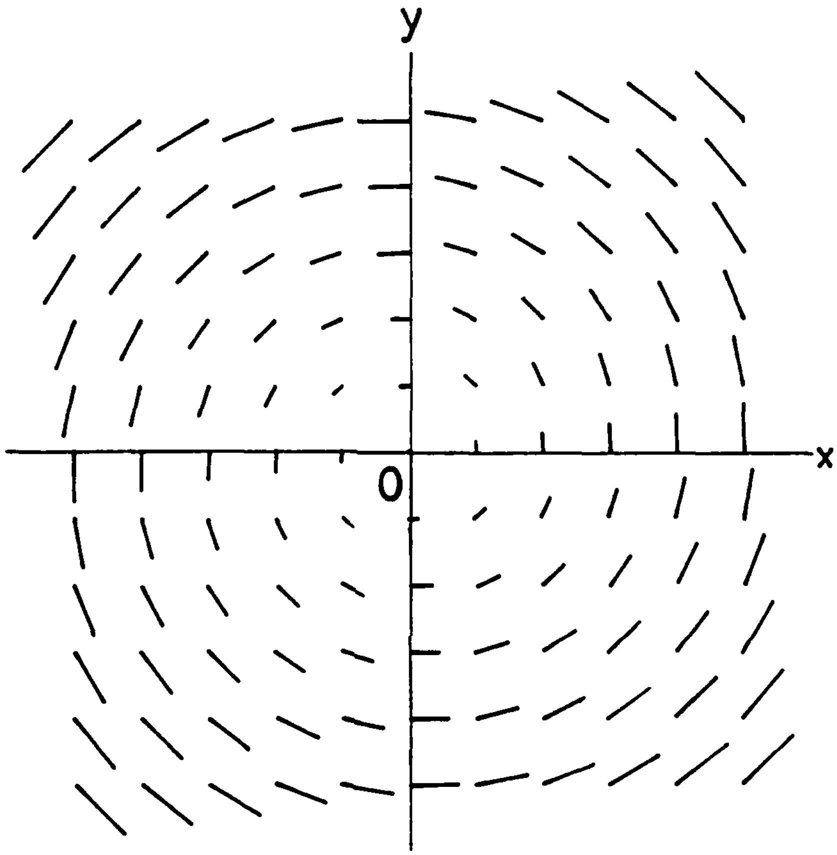

Fig. 5

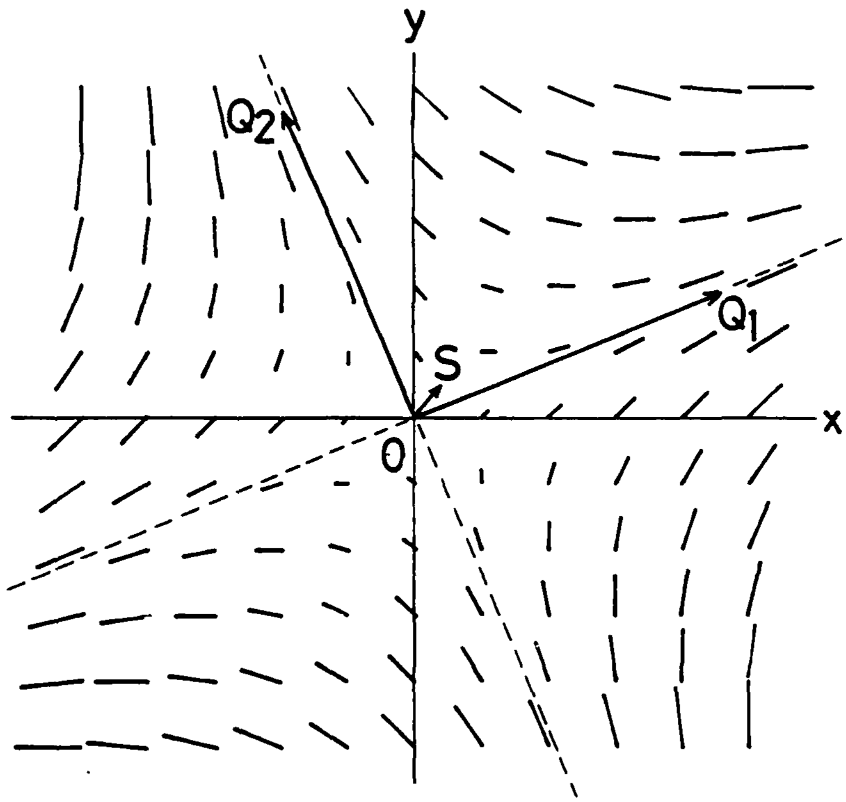

Fig. 6

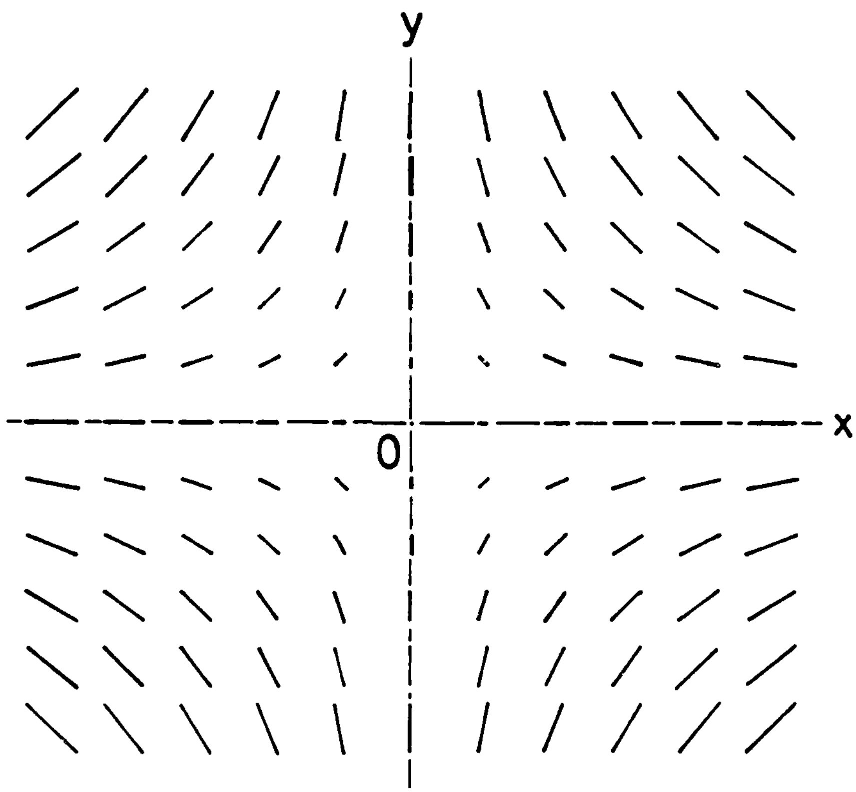

Fig. 7

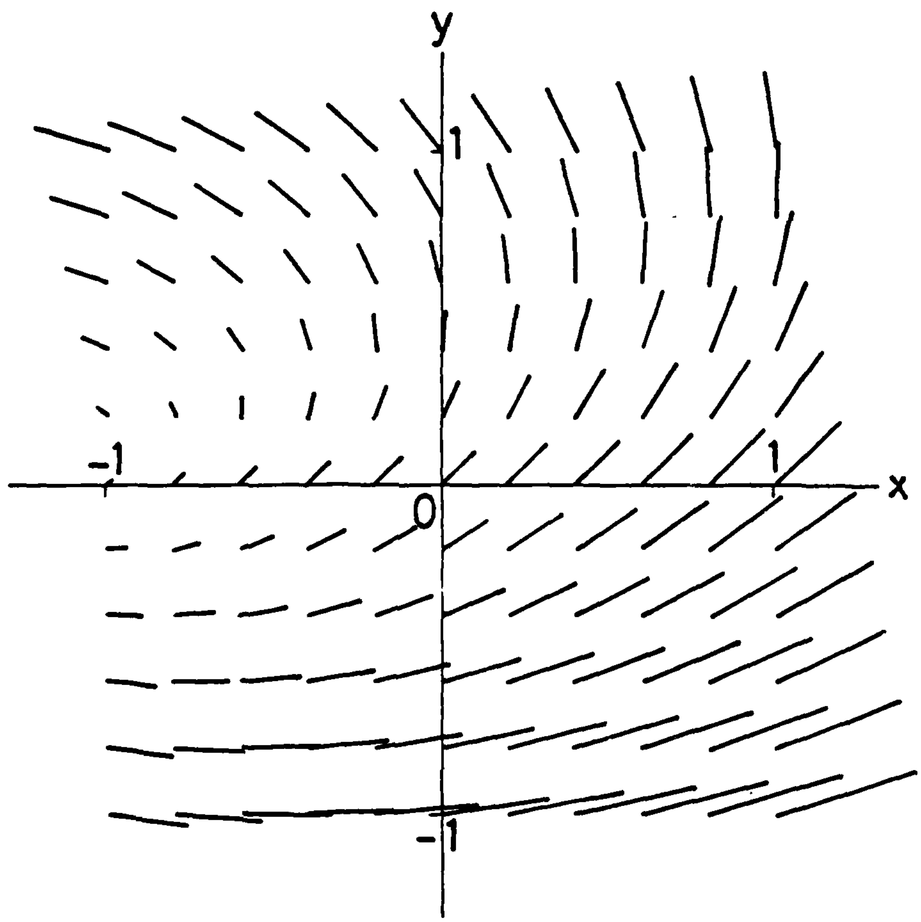

Fig. 8

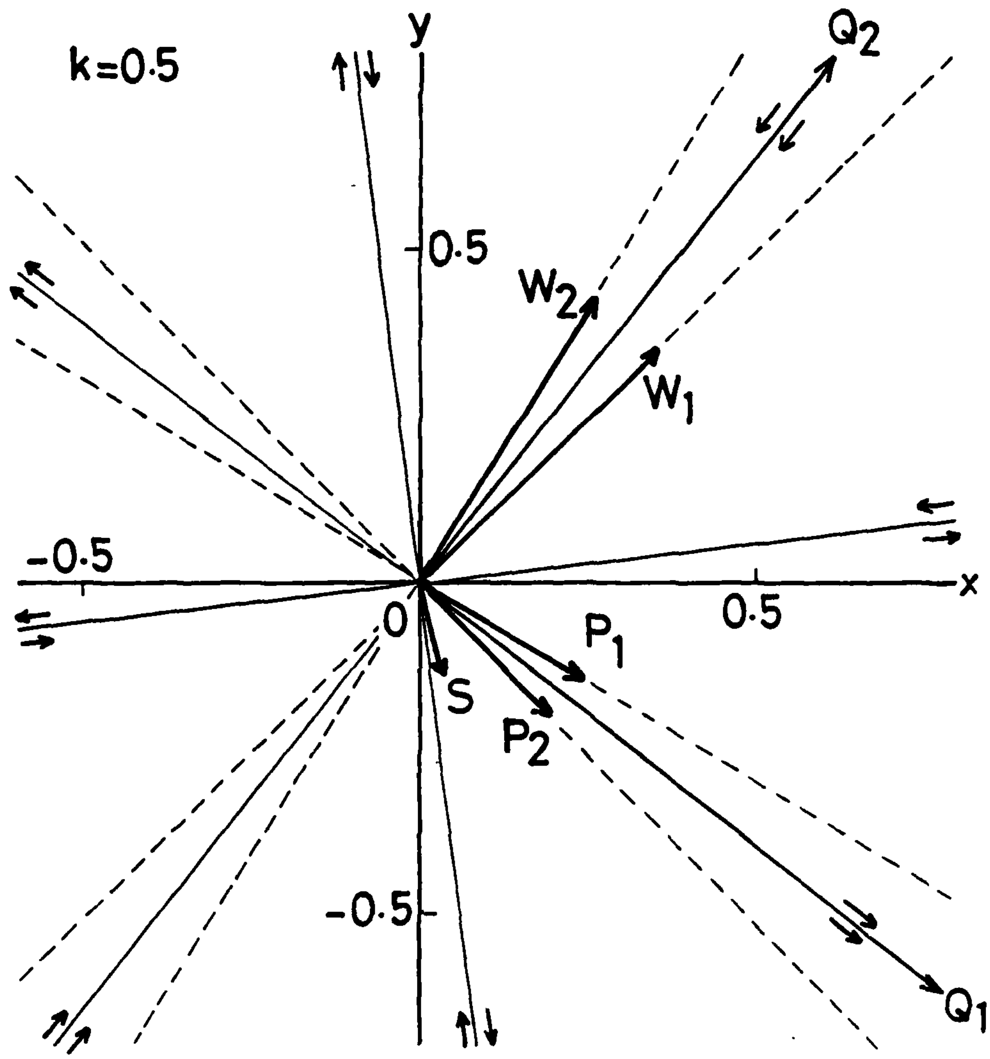

Fig. 9

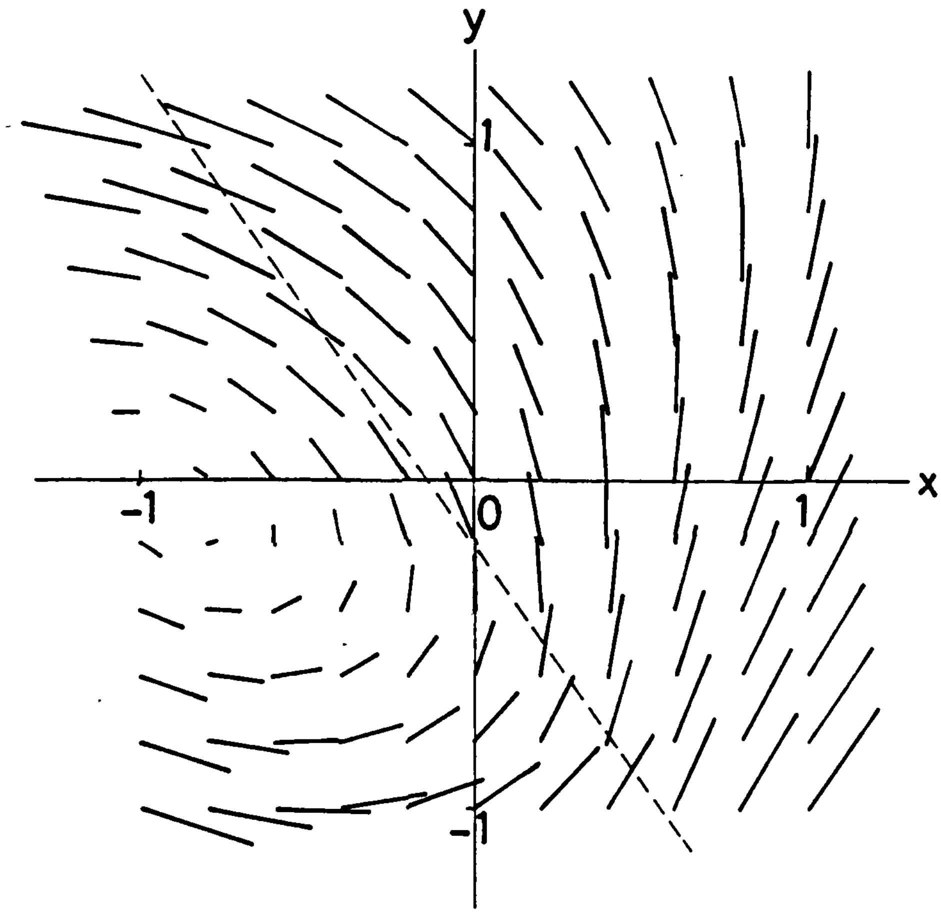

Fig. 10

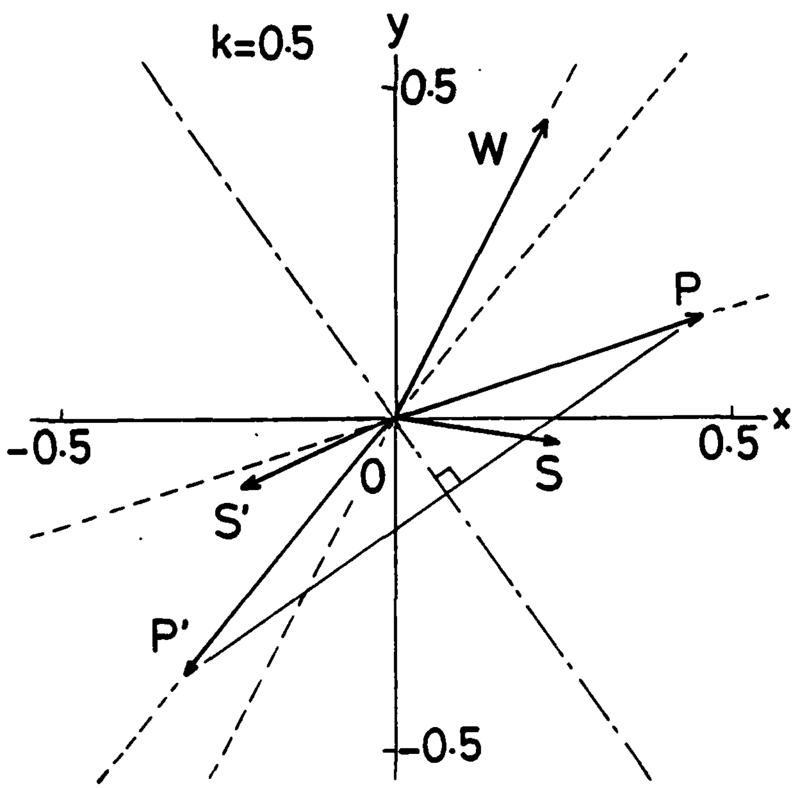

Fig. 11

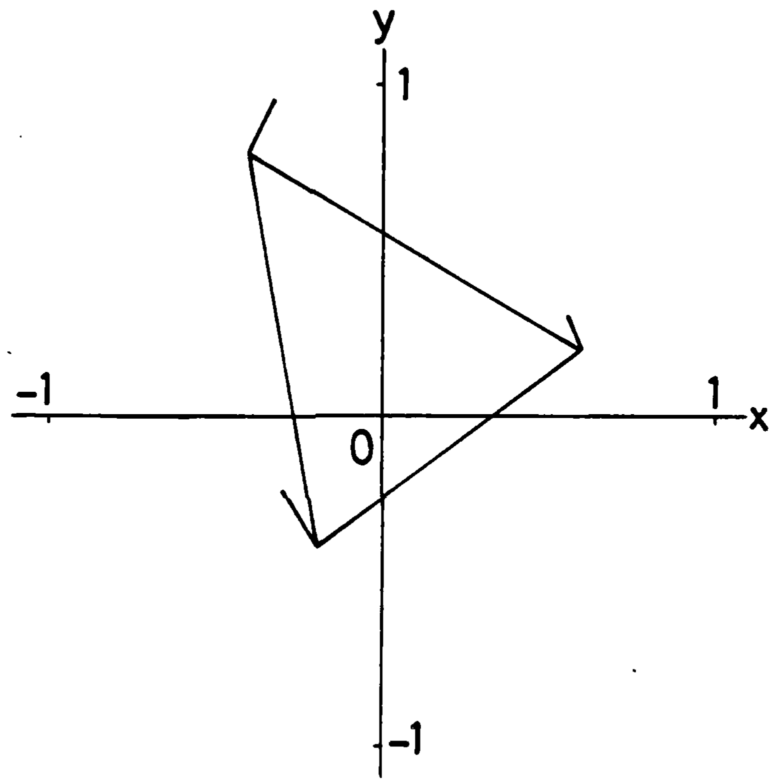

Fig. 12

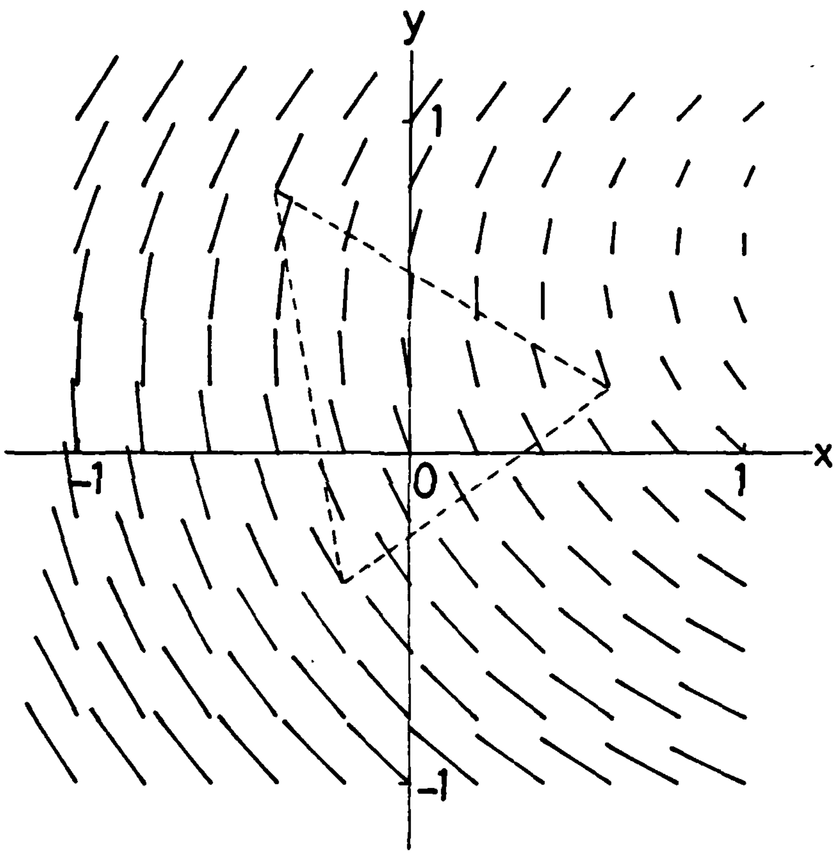

Fig. 13

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE

AD-A171 612

REPORT DOCUMENTATION PAGE

1a REPORT SECURITY CLASSIFICATION: UNCLASSIFIED

1b RESTRICTIVE MARKINGS: N/A

2a SECURITY CLASSIFICATION AUTHORITY: N/A

2b DECLASSIFICATION/DOWNGRADING SCHEDULE: N/A

3. DISTRIBUTION/AVAILABILITY OF REPORT: Approved for public release; distribution unlimited

4 PERFORMING ORGANIZATION REPORT NUMBER(S): CAR-TR-160, CS-TR-1576

5. MONITORING ORGANIZATION REPORT NUMBER(S): N/A

6a NAME OF PERFORMING ORGANIZATION: University of Maryland

6b. OFFICE SYMBOL (If applicable): N/A

7a. NAME OF MONITORING ORGANIZATION: Army Night Vision and Electro-Optics Laboratory

6c ADDRESS (City, State and ZIP Code): Center for Automation Research, College Park, MD 20742

7b. ADDRESS (City, State and ZIP Code): Fort Belvoir, VA 22060

8a NAME OF FUNDING/SPONSORING ORGANIZATION: Defense Advanced Research Projects Agency

8b OFFICE SYMBOL (If applicable): IPTO

9 PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER: DAAK70-83-K-0018

8c ADDRESS (City, State and ZIP Code): 1400 Wilson Blvd., Arlington, VA 22209

10 SOURCE OF FUNDING NOS.

PROGRAM ELEMENT NO	PROJECT NO	TASK NO	WORK UNIT NO

11 TITLE (Include Security Classification): Analysis of Structure and Motion from Optical Flow. Part I: Orthographic Projection

12. PERSONAL AUTHOR(S): Ken-ichi Kanatani

13a TYPE OF REPORT: Technical

13b. TIME COVERED: FROM ______ TO N/A

14. DATE OF REPORT (Yr., Mo., Day): June 1985

15. PAGE COUNT: 46

16 SUPPLEMENTARY NOTATION:

17 COSATI CODES

FIELD	GROUP	SUB. GR.

18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)

19 ABSTRACT (Continue on reverse if necessary and identify by block number)

The 3D structure and motion of an object is determined from its optical flow under othrographic projection. First, the image domain is divided into planar or almost planar regions by checking the flow. For each region, parameters of the flow are determined. Transformation rules under coordinate changes and hydrodynamic analogies are also discussed. The 3D structure and motion are determined in explicit forms in terms of irreducible parameters deduced from group representation theory. The solution is not unique, containing an indeterminate scale factor and comprising true and spurious solutions. Their geometrical interpretations are also studied. The spurious solution disappears if two or more regions of the object are observed.

20 DISTRIBUTION/AVAILABILITY OF ABSTRACT: UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT. ☐ DTIC USERS ☐

21. ABSTRACT SECURITY CLASSIFICATION: UNCLASSIFIED

22a NAME OF RESPONSIBLE INDIVIDUAL:

22b TELEPHONE NUMBER (Include Area Code):

22c OFFICE SYMBOL:

DD FORM 1473, 83 APR EDITION OF 1 JAN 73 IS OBSOLETE

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE