

Transformation of Optical Flow by Camera Rotation

KEN-ICHI KANATANI

Abstract—This paper deals with a mathematical problem about the effect of camera rotation on the description of optical flow generated by a planar surface in motion. The transformation law of the parameters is explicitly given by analyzing infinitesimal generators and irreducibly reducing the induced representation of the 3-D rotation group $SO(3)$. The parameter space is decomposed into invariant subspaces, and the optical flow resulting from planar surface motion is accordingly decomposed into two parts, from which an invariant basis is deduced. A procedure is presented to test the equivalence of two optical flows and to reconstruct the camera rotation. The relationship with the analytical expressions for 3-D recovery is also discussed.

Index Terms—Camera rotation, Casimir operator, commutator, computer vision, infinitesimal generators, invariant basis, irreducible representation, Lie algebra, Lie group, optical flow, rotation group, 3-D recovery

I. INTRODUCTION

THIS paper deals with a mathematical problem extracted from a topic of computer vision. Although the result of our analysis is in itself very interesting and has some direct applications in computer vision problems, the mathematical techniques introduced here and their implications have large potential usefulness in a wide range of problems. We will introduce results from abstract mathematics such as theories of Lie group, Lie algebra, group representation, and invariance, and show how such abstract concepts can be used in *actual* computational problems of computer vision.

Traditionally, results of abstract mathematics such as Lie group theory often have been confined in books on mathematics and physics and read almost exclusively by mathematicians and physicists; they are seldom read by a wide spectrum of researchers in computer science. One reason is that materials in these books are not arranged in a way accessible to researchers in engineering science. This paper is intended to bridge such a gap and bring useful mathematical knowledge and computational techniques into researches of computer vision and related studies.

Specifically, the problem we consider in this paper is as follows. Suppose the camera is rotated by a certain angle around the center of the lens relative to a stationary

scene (Fig. 1). As a result, a different image is seen on the image plane. However, a point on the image plane corresponds to a "ray" of light coming from the scene, and the rays coming from a particular object are identical no matter how the camera is rotated.¹ Hence, the "information content" of the image is not affected, and the original image can be recovered if the angle of camera rotation is known. Occlusion, for instance, is not affected by camera rotation.

Suppose the image we observe can be characterized by a finite number of parameters. This happens when the object we are viewing has a well defined shape and are specified by a finite number of parameters, which is often the case in computer vision applications. If the camera is rotated, the image changes so that these parameters also change their values. However, as implied by the above considerations, the new values must be completely determined by the original values and the amount of camera rotation. Hence, a rule of parameter transformation associated with the camera rotation should be obtained. Using this transformation rule of finite number of parameters, we can predict what the image would look like if the camera were to be rotated; we need neither actually rotate the camera nor generate a new image by computation. This technique plays an important role in many real problems as well as in theoretical purposes.

Since the parameters undergo a transformation rule associated with the camera rotation, there may exist *invariants*, i.e., expressions in the parameters which do not change their values. If so, we can test for the *equivalence* of two images; we can check whether or not two given images represent one and the same object viewed from different camera angles. This is done by measuring the invariants; if they take on different values, these two images cannot be equivalent. We may choose a *complete* set of invariance, or *invariance basis*, in such a way that two images are necessarily equivalent if they take on the same values.

The above consideration applies not only to stationary scenes but to moving objects as well. In this paper, we consider optical flows resulting from a planar surface in 3-D motion. First, we derive the transformation law which predicts what optical flow would be obtained if the camera had a different orientation. We will show, by using the theory of Lie group and Lie algebra, that we can derive

Manuscript received March 12, 1986; revised August 7, 1987. This work was supported in part by the Defense Advanced Research Projects Agency and the U.S. Army Night Vision and Electro-Optics Laboratory under Contract DAAB07-86-K-F073.

The author is with the Department of Computer Science, Gunma University, Kiryu, Gunma 376, Japan. Part of this work was performed while he was visiting the Center for Automation Research, University of Maryland, College Park.

IEEE Log Number 8718603.

¹In the following, we ignore the image boundary, assuming that the image plane is sufficiently large and that the object we are interested in is always included in the field of view.

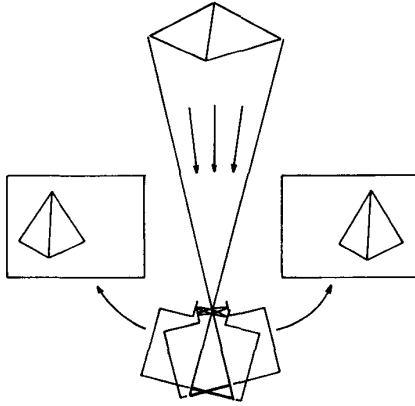


Fig. 1. Incoming rays of light are identical however the camera is rotated around the center of the lens. Hence, the object image retains the same information as it had before the camera rotation (as long as we ignore the film boundary.)

such transformation rules for finite rotations simply by studying infinitesimal transformations.

One of the consequences of our analysis is that the optical flow is decomposed into two parts, each of which has an "invariant" meaning in the sense that each part is transformed independently of the other; one part is characterized by a vector and the other by a tensor. An invariant basis is obtained by constructing invariants from this vector and tensor. These invariants can be used to test for equivalence when two optical flows are given. Furthermore, if the two optical flows are shown to be equivalent, representing one and the same moving surface viewed differently, we can compute the relative orientation of the two cameras. We will also give a brief discussion on the relationship between our analysis and the analytical 3-D recovery problem.

II. CAMERA ROTATION TRANSFORMATION

The camera imaging geometry can be modeled as perspective projection. Take an XYZ -coordinate system, and let the origin O be the center of projection and $Z = f$ be the image plane. A point P in the scene is projected onto the intersection of the image plane with the "ray" connecting the point P and the origin O (Fig. 2). The constant f is determined by the camera: it corresponds to the distance between the center of the lens and the surface of the film. If we choose an xy -coordinate system on the image plane in such a way that the x - and y -axes are parallel to the X - and Y -axes with origin $(0, 0, f)$, a point (X, Y, Z) in the scene is projected onto (x, y) on the image plane, where

$$x = fX/Z, \quad y = fY/Z. \quad (2.1)$$

Suppose the camera is rotated around the center of the lens by an orthogonal matrix R . Then, the point (x, y) moves to another point (x', y') as follows.²

²In this paper, we regard the image plane as a continuous field over which continuous values are defined. Namely, we ignore the sampling of images into discrete pixels, or quantization of gray levels.

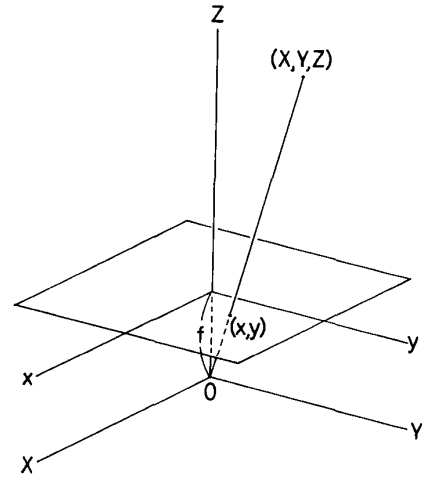


Fig. 2. A point (X, Y, Z) is projected onto point (x, y) on the image plane $Z = f$ by perspective projection from viewpoint O .

Proposition 1: The image transformation induced by camera rotation $R = (r_{ij})$ is given by

$$x' = f \frac{r_{11}x + r_{21}y + r_{31}f}{r_{13}x + r_{23}y + r_{33}f}, \quad y' = f \frac{r_{12}x + r_{22}y + r_{32}f}{r_{13}x + r_{23}y + r_{33}f}. \quad (2.2)$$

Proof: A rotation of the XYZ -coordinate system (i.e., the camera) by R is equivalent to the rotation of the scene in the opposite sense, i.e., by $R^{-1} = R^T$. ("T" denotes matrix transpose.) Then, point (X, Y, Z) in the scene moves to

$$\begin{bmatrix} X' \\ Y' \\ Z' \end{bmatrix} = \begin{bmatrix} r_{11} & r_{21} & r_{31} \\ r_{12} & r_{22} & r_{32} \\ r_{13} & r_{23} & r_{33} \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}. \quad (2.3)$$

(See Fig. 3.) This point is projected onto (x', y') on the image plane by $x' = fX'/Z'$, $y' = fY'/Z'$. Combination of this with (2.1) results in (2.2).

We call the transformation (2.2) the *camera rotation transformation*.³ It should be emphasized that this image transformation does not require any "knowledge" about the scene. This transformation has its inverse, which is obtained by interchanging R and R^T . This means that the "information content" is not altered, since we can freely move between two images by computation.

If the camera rotation is infinitesimal, the camera rotation transformation becomes $x' = x + \delta x$, $y' = y + \delta y$, where

$$\begin{aligned} \delta x &= -f\Omega_2 + \Omega_3y + \frac{1}{f}(-\Omega_2x + \Omega_1y)x + O(\Omega^2), \\ \delta y &= f\Omega_1 - \Omega_3x + \frac{1}{f}(-\Omega_2x + \Omega_1y)y + O(\Omega^2), \end{aligned} \quad (2.4)$$

³The same transformation is mentioned as the "gaze transformation" or "perspectivity" in [19], [22]. The transformations of the form of (2.2) form a subgroup of the 2-D projective transformation group.

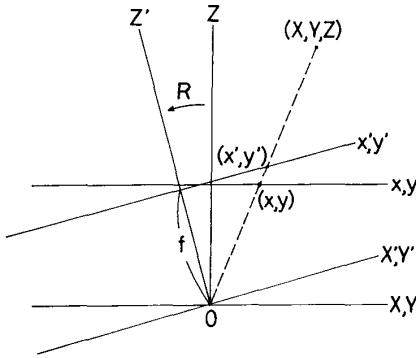


Fig. 3. Geometrical relationship of the camera rotation. A point on the image plane $Z = f$ the intersection of the corresponding ray with the new image plane $Z' = f$.

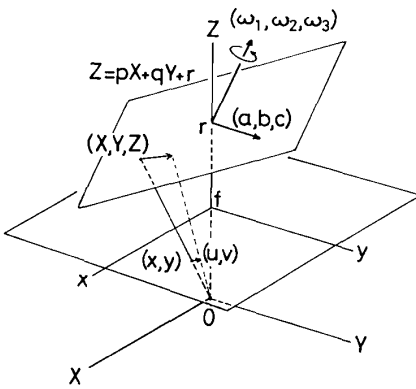


Fig. 4. A planar surface whose equation is $Z = pX + qY + r$ is moving with translation velocity (a, b, c) at $(0, 0, r)$ and rotation velocity $(\omega_1, \omega_2, \omega_3)$ around it. An optical flow is induced on the image xy -plane by perspective projection from the origin O .

where $\Omega_1, \Omega_2, \Omega_3$ are, respectively, the rotation angles of the XYZ -coordinate system (i.e., the camera) around the X -, Y -, and Z -axes screw-wise. The last terms on the right-hand sides designate terms of orders equal to or higher than 2 in $\Omega_1, \Omega_2, \Omega_3$. Equations (2.4) are obtained by substituting $r_{11} = r_{22} = r_{33} = 1 + O(\Omega^2)$, $r_{32} = -r_{23} = \Omega_1 + O(\Omega^2)$, $r_{13} = -r_{31} = \Omega_2 + O(\Omega^2)$, $r_{21} = -r_{12} = \Omega_3 + O(\Omega^2)$ into (2.2).

III. OPTICAL FLOW OF A MOVING PLANAR SURFACE

Suppose a planar surface whose equation is $Z = pX + qY + r$ is moving rigidly in the scene. The instantaneous rigid motion is specified by the velocity (a, b, c) at a reference point—which is called the *translation velocity*—and the *rotation velocity* $(\omega_1, \omega_2, \omega_3)$ screw-wise around that reference point. We take the reference point to be $(0, 0, r)$, the intersection of the surface with the Z -axis.⁴ (Fig.

⁴The choice of the reference point is mathematically arbitrary, and many authors take it at the origin, i.e., the center of the lens. However, the choice above is desirable from the viewpoint of computational robustness. See the discussion in [7].

4.) Then, the *optical flow* induced on the image plane is given as follows.⁵

Proposition 2: The optical flow resulting from a moving planar surface is given by

$$\begin{aligned} u(x, y) &= u_0 + Ax + By + (Ex + Fy)x, \\ v(x, y) &= v_0 + Cx + Dy + (Ex + Fy)y, \end{aligned} \quad (3.1)$$

where

$$\begin{aligned} u_0 &= \frac{fa}{r}, & v_0 &= \frac{fb}{r}, \\ A &= p\omega_2 - \frac{pa + c}{r}, & B &= q\omega_2 - \omega_3 - \frac{qa}{r}, \\ C &= -p\omega_1 + \omega_3 - \frac{pb}{r}, & D &= -q\omega_1 - \frac{qb + c}{r}, \\ E &= \frac{1}{f} \left(\omega_2 + \frac{pc}{r} \right), & F &= \frac{1}{f} \left(-\omega_1 + \frac{qc}{r} \right). \end{aligned} \quad (3.2)$$

The proof is given in Appendix A. (See also [4], [7], [12], [15], [19]–[23].)

Thus, we are viewing a very restricted form of image motion specified by eight parameters $u_0, v_0, A, B, C, D, E, F$; if these parameters are the same, the motion seems identical to the viewer. (This fact has already been pointed out by many researchers, e.g., Longuet-Higgins [12], Waxman and Ullman [23].) These parameters may be computed from a given optical flow by fitting the form of (3.1) by the least square method. If the optical flow itself is not available, we can also compute them from the time changes of some global characteristics such as the surface boundary contour [9], [21], [22]. Let us call these parameters *flow parameters*.

Now, what flow would we observe if the camera were oriented differently? One thing is certain: *we would still observe the motion of a planar surface and hence observe an optical flow which has the form*

$$\begin{aligned} u' &= u'_0 + A'x' + B'y' + (E'x' + F'y')x', \\ v' &= v'_0 + C'x' + D'y' + (E'x' + F'y')y'. \end{aligned} \quad (3.3)$$

This is because a planar surface is always a planar surface no matter where it is viewed from. Hence, the “form” of the optical flow is always the same. Suppose the new camera orientation is given by rotating the camera by R from the original orientation around the viewpoint. The following questions naturally arise.

1) How are the new flow parameters u'_0, v'_0, A', B', C' ,

⁵The term “optical flow” is used with many different meanings depending on the situation. Image processing techniques to detect “optical flow” are supposed to detect small (but finite) displacements of “corresponding points” between two image frames. However, the correspondence is often established between “points of the same light intensity” which do not necessarily correspond to the same point of the object. Here, we use this term to mean the instantaneous velocity field produced on the image plane. In [15], [19]–[23], the term “image flow” is used with the same meaning.

D' , E' , F' expressed in terms of the original flow parameters u_0 , v_0 , A , B , C , D , E , F and the camera rotation R ?

2) What expressions of the flow parameters are *invariant* with respect to the camera rotation?

3) Given two optical flows resulting from planar motion, how can we test their *equivalence*?

4) If equivalence is confirmed, how can we find a rotation R which would transform one camera orientation into the other?

We can answer these questions simply by analyzing *infinitesimal transformations*. In other words, consideration of "infinitesimal" camera rotation automatically leads to transformation laws for "finite" rotations. To show this is one of the main purposes of this paper.

IV. TRANSFORMATION OF FLOW PARAMETERS

If the camera rotation is infinitesimal, the image is also altered infinitesimally according to (2.4). The transfor-

$$\begin{aligned} b_{11} &\equiv \frac{1}{3}(2A - D), & b_{22} &\equiv \frac{1}{3}(2D - A), & b_{33} &\equiv -\frac{1}{3}(A + D), \\ b_{12} = b_{21} &\equiv \frac{1}{2}(B + C), & b_{23} = b_{32} &\equiv \frac{1}{2}(v_0/f - fF), & b_{31} = b_{13} &\equiv \frac{1}{2}(u_0/f - fE), \end{aligned} \quad (4.6)$$

mation of the flow parameters is also infinitesimal and given as follows.

Lemma 1: The infinitesimal transformations $u'_0 = u_0 + \delta u_0$, \dots , $F' = F + \delta F$ of the flow parameters u_0 , v_0 , A , B , C , D , E , F are given by

$$\delta \begin{bmatrix} u_0 \\ v_0 \\ A \\ B \\ C \\ D \\ E \\ F \end{bmatrix} = \begin{bmatrix} & & \Omega_3 & f\Omega_2 & -f\Omega_1 & & & & \\ & -\Omega_3 & & & & f\Omega_2 & -f\Omega_1 & & \\ & -2\Omega_2/f & \Omega_1/f & & \Omega_3 & \Omega_3 & & 2f\Omega_2 & -f\Omega_1 \\ & \Omega_1 & & -\Omega_3 & & \Omega_3 & & f\Omega_2 & \\ & & -\Omega_2/f & -\Omega_3 & & \Omega_3 & -f\Omega_1 & & f\Omega_2 \\ & -\Omega_2/f & 2\Omega_1/f & & -\Omega_3 & -\Omega_3 & & f\Omega_2 & -2f\Omega_1 \\ & & & -\Omega_2/f & & \Omega_1/f & & & \Omega_3 \\ & & & & -\Omega_2/f & & \Omega_1/f & -\Omega_3 & \end{bmatrix} \begin{bmatrix} u_0 \\ v_0 \\ A \\ B \\ C \\ D \\ E \\ F \end{bmatrix} + O(\Omega^2). \quad (4.1)$$

The proof is given in Appendix B.

Lemma 2: The flow parameters u_0 , v_0 , A , B , C , D , E , F are transformed linearly by the (finite) camera rotation transformation, namely

$$\begin{bmatrix} u'_0 \\ \vdots \\ F' \end{bmatrix} = T(R) \begin{bmatrix} u_0 \\ \vdots \\ F \end{bmatrix}, \quad (4.2)$$

and hence define a representation of $SO(3)$. In other words, if R_1 , R_2 are two 3-D rotations, then we have a homomorphism from $SO(3)$ into the group of eight-

dimensional matrices under matrix multiplication:

$$T(R_1) T(R_2) = T(R_2 R_1). \quad (4.3)$$

The proof is given in Appendix C.

Lemma 3: The representation $T(R)$ is equivalent to the direct sum of the irreducible representations \mathfrak{D}_1 and \mathfrak{D}_2 of weights 1 and 2, respectively:

$$T(R) \equiv \mathfrak{D}_1 \oplus \mathfrak{D}_2. \quad (4.4)$$

The proof is given in Appendix D.

Lemma 4: If we define a_1 , a_2 , a_3 by

$$a_1 \equiv -\frac{1}{2}(v_0/f + fF),$$

$$a_2 \equiv \frac{1}{2}(u_0/f + fE),$$

$$a_3 \equiv \frac{1}{2}(C - B), \quad (4.5)$$

they are transformed as a vector, and if we define

they are transformed as a deviator tensor. Namely,

$$\begin{bmatrix} a'_1 \\ a'_2 \\ a'_3 \end{bmatrix} = R^T \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}, \quad (4.7)$$

$$\begin{bmatrix} b'_{11} & b'_{12} & b'_{13} \\ b'_{21} & b'_{22} & b'_{23} \\ b'_{31} & b'_{32} & b'_{33} \end{bmatrix} = R^T \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{bmatrix} R. \quad (4.8)$$

The proof is given in Appendix E.

Now, we obtain the final result about the transformation of the flow parameters as follows.

Proposition 3: Given the flow parameters u_0 , v_0 , A , B , C , D , E , F , define the 3×3 matrix $C = (c_{ij})$ by

$$\begin{aligned}
c_{11} &= \frac{1}{3}(2A - D), & c_{12} &= B, & c_{13} &= \frac{1}{f}u_0, \\
c_{21} &= C, & c_{22} &= \frac{1}{3}(2D - A), & c_{23} &= -fF, \\
c_{31} &= -fE, & c_{32} &= \frac{1}{f}v_0, & c_{33} &= -\frac{1}{3}(A + D),
\end{aligned} \tag{4.9}$$

Transform this matrix as a tensor:

$$\begin{bmatrix} c'_{11} & c'_{12} & c'_{13} \\ c'_{21} & c'_{22} & c'_{23} \\ c'_{31} & c'_{32} & c'_{33} \end{bmatrix} = \mathbf{R}^T \begin{bmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{bmatrix} \mathbf{R}. \tag{4.10}$$

Then, the transformed flow parameters are given as follows:

$$\begin{aligned}
u'_0 &= fc'_{13}, & v'_0 &= fc'_{32}, \\
A' &= c'_{11} - c'_{33}, & B' &= c'_{12}, \\
C' &= c'_{21}, & D' &= c'_{22} - c'_{33}, \\
E' &= -\frac{1}{f}c'_{31}, & F' &= -\frac{1}{f}c'_{23}.
\end{aligned} \tag{4.11}$$

The proof is given in Appendix F.

V. INVARIANT BASIS OF PLANAR OPTICAL FLOW

From Lemma 4, we see that an optical flow resulting from planar motion is characterized by a (real) vector $\mathbf{a} = (a_i)$, and a tensor $\mathbf{B} = (b_{ij})$ which is a (real) traceless symmetric matrix. The vector \mathbf{a} has only one invariant, namely its "length" $\|\mathbf{a}\|$, or equivalently

$$\mathbf{a}^T \mathbf{a}, \tag{5.1}$$

while the real symmetric tensor \mathbf{B} has three invariants, namely its three *real* principal values $\sigma_1, \sigma_2, \sigma_3$, or equivalently any three independent algebraic expressions formed from these such as the "fundamental symmetric polynomials":

$$\sigma_1 + \sigma_2 + \sigma_3, \quad \sigma_2\sigma_3 + \sigma_3\sigma_1 + \sigma_1\sigma_2, \quad \sigma_1\sigma_2\sigma_3 \tag{5.2}$$

(cf. [13], [14], [16]–[18], [24]). Alternatively, we can use the following sums of powers:

$$\sigma_1 + \sigma_2 + \sigma_3, \quad \sigma_1^2 + \sigma_2^2 + \sigma_3^2, \quad \sigma_1^3 + \sigma_2^3 + \sigma_3^3. \tag{5.3}$$

This set is computationally convenient, since they are expressed in terms of the components of matrix \mathbf{B} as

$$\text{Tr } \mathbf{B} (= 0), \quad \text{Tr } \mathbf{B}^2, \quad \text{Tr } \mathbf{B}^3, \tag{5.4}$$

where Tr denotes the trace of a matrix. If the tensor is symmetric and traceless, there remain only two invariants, since the first invariant (trace) vanishes.

Thus, the vector \mathbf{a} has one invariant, and the symmetric tensor \mathbf{B} has three invariants. However, if we consider a *set* of the vector \mathbf{a} and the tensor \mathbf{B} , we must have, in addition to the invariants of each of them, invariants de-

scribing the *relative relationship* between the vector \mathbf{a} and the tensor \mathbf{B} .

Geometrically speaking, a vector is thought of as a *directed* axis to which its "magnitude" (i.e., the length) is attached, while a symmetric tensor is thought of as three mutually perpendicular *undirected* axes (i.e., the *principal axes*) to which their respective "magnitudes" (i.e., the *principal values*) are attached. It is intuitively clear that for a vector its length is the only invariant, since the orientation of its axis can be changed arbitrarily by a 3-D rotation. It is also clear that for a symmetric tensor the three principal values are the only invariants, since the orientations of its principal axes as a whole are changed by a 3-D rotation.

However, if we have a vector and a tensor, there appears another invariant aspect which is not changed by a 3-D rotation of *both* of them at the same time: namely, the orientation of the vector *relative* to the principal axes of the tensor. Since orientation can be specified by two parameters (say, the longitude and latitude on a unit sphere), we have two additional invariants. The choice is not unique, but a computationally convenient one is given by the following two (cf. [13], [14], [16]–[18]).

$$\mathbf{a}^T \mathbf{B} \mathbf{a}, \quad \mathbf{a}^T \mathbf{B}^2 \mathbf{a} \tag{5.5}$$

Now, let us consider the *equivalence* of two images; we say that two images are *equivalent* if they are obtained by projection of the same object for different camera orientations. Since we have already obtained an invariant basis, we can conclude as follows.

Proposition 4: Two optical flows resulting from planar motion are equivalent if and only if invariants

$$\mathbf{a}^T \mathbf{a}, \quad \text{Tr } \mathbf{B}^2, \quad \text{Tr } \mathbf{B}^3, \quad \mathbf{a}^T \mathbf{B} \mathbf{a}, \quad \mathbf{a}^T \mathbf{B}^2 \mathbf{a}, \tag{5.6}$$

take the same values, where $\mathbf{a} = (a_i)$ and $\mathbf{B} = (b_{ij})$ are given by (4.5) and (4.6).

VI. RECONSTRUCTION OF CAMERA ROTATION

If two optical flows are equivalent,⁶ we can find a rotation matrix \mathbf{R} which takes one camera orientation into the other. This is done as follows.

Suppose we observe vector \mathbf{a} and tensor \mathbf{B} from one flow, and similarly \mathbf{a}' and \mathbf{B}' from the other. For simplicity, we assume that none of them is zero.

⁶Of course, in real situations, we must take noise and error into consideration, and allow a certain tolerance for testing the equivalence by the values of the invariants.

First, assume that \mathbf{B} (hence \mathbf{B}' as well) has three distinct principal values $\sigma_1, \sigma_2, \sigma_3$. Let $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ be the mutually orthogonal unit vectors indicating the corresponding principal axes of \mathbf{B} . Since they are determined except for sign, fix one set such that $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ form a right-hand system in that order. Construct a matrix \mathbf{R}_1 having $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ as its three columns in this order.

Similarly, let $\mathbf{e}'_1, \mathbf{e}'_2, \mathbf{e}'_3$ be the mutually orthogonal unit vectors indicating the corresponding principal axes of \mathbf{B}' such that they form a right-hand system in this order. Since the signs are arbitrary, there exist four possibilities. (The remaining four are left-hand systems.) For each case, construct the corresponding matrix $\mathbf{R}_2^{(i)}$, $i = 1, 2, 3, 4$. Then, the camera rotation which transforms \mathbf{B}' to \mathbf{B} is given by

$$\mathbf{R}^{(i)} = \mathbf{R}_2^{(i)} \mathbf{R}_1^T, \quad i = 1, 2, 3, 4. \quad (6.1)$$

(The proof is given in Appendix G.) Finally, choose among these four, the one which transforms \mathbf{a}' to \mathbf{a} .

If \mathbf{B} (hence \mathbf{B}' as well) has only two distinct eigenvalues, (i.e., a single root and a pair of multiple roots), let \mathbf{e}_1 be the unit eigenvector associated with the single root. Suppose \mathbf{a} is neither parallel nor perpendicular to \mathbf{e}_1 . Since the sign of \mathbf{e}_1 is arbitrary, choose it so that \mathbf{a} and \mathbf{e}_1 make an acute angle. Then, we can obtain three mutually orthogonal vectors forming a right-hand system by constructing

$$\mathbf{e}_1, \quad \mathbf{e}_2 = \frac{\mathbf{e}_1 \times \mathbf{a}}{\|\mathbf{e}_1 \times \mathbf{a}\|}, \quad \mathbf{e}_3 = \mathbf{e}_1 \times \mathbf{e}_2. \quad (6.2)$$

We can construct matrices $\mathbf{R}_1, \mathbf{R}_2$ as described above, and the camera rotation is given by $\mathbf{R} = \mathbf{R}_2 \mathbf{R}_1^T$.

The remaining cases yield multiple solutions. If \mathbf{a} is perpendicular to \mathbf{e}_1 , there exist two solutions. If \mathbf{a} is parallel to \mathbf{e}_1 , or if \mathbf{B} (hence \mathbf{B}' as well) has one eigenvalue, i.e., \mathbf{B} ($= \mathbf{B}'$) is a multiple of the unit matrix \mathbf{I} , the solution for \mathbf{R} is any rotation that maps \mathbf{a}' to \mathbf{a} , and we can add any rotation around \mathbf{a} .

VII. INVARIANT DECOMPOSITION OF OPTICAL FLOW

According to Proposition 3, if we define c_{ij} , $i, j = 1, 2, 3$, by (4.9), the optical flow equations (3.1) are written as

$$\begin{aligned} u &= fc_{13} + (c_{11} - c_{33})x + c_{12}y - \frac{1}{f}(c_{31}x + c_{32}y)x, \\ v &= fc_{23} + c_{21}x + (c_{22} - c_{33})y - \frac{1}{f}(c_{31}x + c_{32}y)y, \end{aligned} \quad (7.1)$$

and c_{ij} , $i, j = 1, 2, 3$, are transformed as a tensor [i.e., by (4.10)] by the camera rotation transformation.

On the other hand, Lemma 4 asserts that if we define a_i , $i = 1, 2, 3$, by (4.5) and b_{ij} , $i, j = 1, 2, 3$, by (4.6), the optical flow equations (3.1) are decomposed into two

parts $u = u_a + u_b$, $v = v_a + v_b$, where

$$\begin{aligned} u_a &= fa_2 - a_3y + \frac{1}{f}(a_2x - a_1y)x, \\ v_a &= -fa_1 + a_3x + \frac{1}{f}(a_2x - a_1y)y, \end{aligned} \quad (7.2)$$

$$\begin{aligned} u_b &= fb_{31} + (b_{11} - b_{33})x + b_{12}y \\ &\quad + \frac{1}{f}(-b_{31}x - b_{23}y)x, \\ v_b &= fb_{32} + b_{12}x + (b_{22} - b_{33})y \\ &\quad + \frac{1}{f}(-b_{31}x - b_{23}y)y. \end{aligned} \quad (7.3)$$

This decomposition is *unique* because a_i , $i = 1, 2, 3$, and b_{ij} , $i, j = 1, 2, 3$, are computed from the original flow parameters $u_0, v_0, A, B, C, D, E, F$ by (4.5) and (4.6). Moreover, this decomposition is *invariant* in the sense that each flow is transformed by the camera rotation transformation independently of the other (4.7), (4.8). At the same time, this decomposition is *irreducible* in the sense that no further decomposition is possible.⁷ Hence, *they should have separate geometrical meanings independent of the other*, cf. [8]. Let us call the flow (u_a, v_a) the *vector part* and the flow (u_b, v_b) the *tensor part* of the original optical flow.

Suppose the entire scene, including all the objects in it, is stationary, and the camera is rotating around the center of the lens with rotation velocity $(\omega_1, \omega_2, \omega_3)$. Consider what "optical flow" we should observe on the image plane. Since the rotation velocity components $\omega_1, \omega_2, \omega_3$ are obtained by dividing infinitesimal camera rotation angles $\Omega_1, \Omega_2, \Omega_3$ by the time δt it took, we divide both sides of the infinitesimal camera rotation transformation (2.4) by δt , take the limit of $\delta t \rightarrow 0$, and obtain

$$\begin{aligned} u &= -f\omega_2 + \omega_3y + \frac{1}{f}(-\omega_2x + \omega_1y)x, \\ v &= f\omega_1 - \omega_3x + \frac{1}{f}(-\omega_2x + \omega_1y)y. \end{aligned} \quad (7.4)$$

Comparing these equations to (7.2), we see that the vector part (u_a, v_a) is the flow we will observe if the camera is rotated with rotation velocity $(-a_1, -a_2, -a_3)$ relative to the stationary scene, or equivalently if the planar surface (or any other object or scene) is *orbiting* around the origin (i.e., the center of the lens), the configuration relative to the camera kept fixed, with rotation velocity (a_1, a_2, a_3) .

This means that this vector part essentially does not contain any information about the 3-D structure and motion of the object. We may alternatively call this flow the "uninformative part." The information about the 3-D structure and motion is contained in the tensor part (u_b, v_b) .

⁷For the transformation rules of (4.7) and (4.8) correspond to irreducible representations \mathfrak{D}_1 and \mathfrak{D}_2 of $SO(3)$, cf. Appendix E.

v_b). The decomposition of the flow into two parts as (7.2) and (7.3) is simply the separation of the flow component due to the "orbiting" motion from the remaining flow components. This observation implies that 3-D recovery from optical flow can be performed from the tensor part alone.

The recovery of 3-D shape and motion from optical flow has been studied by many researchers. If the object is a planar surface, the solution is obtained in an analytically closed form [6], [7], [12], [15], [20]. If the result is rewritten in terms of our a_i , $i = 1, 2, 3$, and b_{ij} , $i, j = 1, 2, 3$, it can be shown that the surface gradient components p , q are completely determined by tensor (b_{ij}) alone. In other words, suppose we determine the translation velocity components (scaled by the depth r) a/r , b/r , c/r (the solution is unique) and the rotation velocity components ω_1 , ω_2 , ω_3 (two sets of solution exist),⁸ assuming that the vector part is zero: $a_i = 0$, $i = 1, 2, 3$. It can be proved that addition of a nonzero vector part only modifies the solution in the form

$$a/r \rightarrow a/r + a_2, \quad b/r \rightarrow b/r - a_1, \quad (7.5)$$

$$\omega_1 \rightarrow \omega_1 + a_1, \quad \omega_2 \rightarrow \omega_2 + a_2, \quad \omega_3 \rightarrow \omega_3 + a_3. \quad (7.6)$$

Namely, the vector part simply adds the effect of the "orbiting" of the surface around the center of the lens. (The proof is given in Appendix H.)

Finally, note that if the camera is instantaneously rotating with some angular velocity, the flow (7.4) is superimposed onto the original optical flow. However, if we decompose the flow into its vector and tensor parts, the tensor part is not altered at all, since the superimposition only adds to the vector part.

VIII. CONCLUDING REMARKS

In this paper, we analyzed the optical flow resulting from planar motion, and derived the parameter transformation rule when the camera is rotated by a *finite* amount. However, invoking the Lie group theory, we can deduce finite transformation properties solely from analysis of *infinitesimal* transformations.

Our results are interesting in its own right from a theoretical point of view. More importantly, however, the mathematical tools developed here will provide a powerful means of analysis in a wide range of problems in computer vision. For example, the same technique can be applied to static scenes, characterizing object images in terms of "invariant features," i.e., quantities invariant to camera rotation. This is done in [9].

Another example is the 3-D recovery problem. Since perspective distortion of image increases as we go away

⁸As pointed out earlier, a rigid motion is specified by the translation velocity and the rotation velocity with respect to an arbitrary chosen reference point. Our reference point is the intersection of the surface with the Z-axis, and in this case the translation velocity (scaled by the depth) is uniquely determined, while the rotation velocity has two solutions. Some authors take it at the coordinate origin, i.e., the viewpoint. Then, their "translation velocity" is a linear combination of our translation velocity and rotation velocity, and hence the solution is not unique. See also the discussion in [7].

from the image origin, it is convenient first to move the image into the center of the image by the camera rotation transformation. Then, we analyze the 3-D shape, position, motion, etc. there, and transform the obtained 3-D interpretation back into the original configuration. This technique is used for the shape-from-texture problem [11] and for the interpretation of length and angle projected onto the image plane [10]. A similar approach is suggested for the 3-D recovery of curved surfaces from optical flow [19], [22].

In this paper, we did not discuss the 3-D recovery from optical flow—its significance and analytical techniques. There exist many papers on this issue. A general philosophy is given by [19], [22], for example. Closed form analytical solution is given for planar surface motion [6], [7], [12], [15]. The general case of curved surfaces is also analyzed, and closed form solution is given [20]–[23]. However, accurate detection of optical flow from a sequence of real images is in general a very difficult problem. Although many techniques have been proposed and tested, we do not refer to them, since the technical issue is not the scope of this paper and many papers are currently being reported one after another. Analytical techniques for detecting optical flow or the object motion directly from global characteristics, such as the object contour, without first detecting the point-to-point correspondence, have also been proposed and tested [1], [4], [5], [21], [22].

APPENDIX A PROOF OF PROPOSITION 2

If we combine the perspective projection relationship (2.1) with the equation $Z = pX + qY + r$ of the planar surface, we obtain the following one-to-one correspondence between the surface and the image plane:

$$\begin{aligned} X &= \frac{rx}{f - px - qy}, & Y &= \frac{ry}{f - px - qy}, \\ Z &= \frac{rf}{f - px - qy}. \end{aligned} \quad (A.1)$$

According to the definition of the translation and rotation velocities, the velocity of point (X, Y, Z) in the scene is given by

$$\begin{bmatrix} \dot{X} \\ \dot{Y} \\ \dot{Z} \end{bmatrix} = \begin{bmatrix} a \\ b \\ c \end{bmatrix} + \begin{bmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \end{bmatrix} \times \begin{bmatrix} X \\ Y \\ Z - r \end{bmatrix}. \quad (A.2)$$

Substituting $Z = pX + qY + r$ into this, we obtain

$$\begin{aligned} \dot{X} &= a + p\omega_2 X + (q\omega_2 - \omega_3)Y, \\ \dot{Y} &= b + (\omega_3 - p\omega_1)X - q\omega_1 Y, \\ \dot{Z} &= c - \omega_2 X + \omega_1 Y. \end{aligned} \quad (A.3)$$

Differentiating both sides of (2.1), we obtain the velocity

of the image point as follows:

$$\begin{aligned}\dot{x} &= \frac{f\dot{X}}{Z} - \frac{fX\dot{Z}}{Z^2} = \frac{f\dot{X}}{Z} - \frac{x\dot{Z}}{Z}, \\ \dot{y} &= \frac{f\dot{Y}}{Z} - \frac{fY\dot{Z}}{Z^2} = \frac{f\dot{Y}}{Z} - \frac{y\dot{Z}}{Z}.\end{aligned}\quad (\text{A.4})$$

From (2.1) and (A.3), we obtain

$$\begin{aligned}\frac{f\dot{X}}{Z} &= \frac{fa}{Z} + p\omega_2x + (q\omega_2 - \omega_3)y, \\ \frac{f\dot{Y}}{Z} &= \frac{fb}{Z} + (\omega_3 - p\omega_1)x - q\omega_1y, \\ \frac{f\dot{Z}}{Z} &= \frac{fc}{Z} - \omega_2x + \omega_1y.\end{aligned}\quad (\text{A.5})$$

Substituting these into (A.4), and eliminating Z by using the last equation of (A.1), we obtain the optical flow in the form of (3.1) and (3.2).

APPENDIX B PROOF OF LEMMA 1

Velocity components u, v are transformed under transformation $x \rightarrow x', y \rightarrow y'$ by

$$\begin{bmatrix} u' \\ v' \end{bmatrix} = \begin{bmatrix} \partial x'/\partial x & \partial x'/\partial y \\ \partial y'/\partial x & \partial y'/\partial y \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix}.\quad (\text{B.1})$$

In other words, the velocity vector is multiplied by the "Jacobi matrix." If we substitute the infinitesimal camera rotation transformation $x' = x + \delta x, y' = y + \delta y$ given by (2.4), and apply the optical flow equations (3.1), we obtain $u' = u + \delta u, v' = v + \delta v$, where

$$\begin{aligned}\delta u &= \Omega_3 v_0 + \left(-\frac{1}{f}(2\Omega_2 u_0 - \Omega_1 v_0) + \Omega_3 C\right)x \\ &+ \left(\frac{1}{f}\Omega_1 u_0 + \Omega_3 D\right)y - \frac{1}{f}(2\Omega_2 A - \Omega_1 C)x^2 \\ &- \left(\frac{1}{f}(\Omega_1 A - 2\Omega_2 B + \Omega_1 D) + \Omega_3 E\right)xy \\ &+ \left(\frac{1}{f}\Omega_1 B + \Omega_3 F\right)y^2 - \frac{2}{f}\Omega_2 E x^3 \\ &+ \frac{2}{f}(\Omega_1 E - \Omega_2 F)x^2 y + \frac{2}{f}\Omega_1 F x y^2 + O(\Omega^2),\end{aligned}$$

$$\begin{aligned}\delta v &= -\Omega_3 u_0 - \left(\frac{1}{f}\Omega_2 v_0 + \Omega_3 A\right)x \\ &+ \left(-\frac{1}{f}(\Omega_2 u_0 - 2\Omega_1 v_0) - \Omega_3 B\right)y \\ &- \left(\frac{1}{f}\Omega_2 C + \Omega_3 E\right)x^2 \\ &+ \left(-\frac{1}{f}(\Omega_2 A - 2\Omega_1 C + \Omega_2 D) - \Omega_3 E\right)xy \\ &- \frac{1}{f}(\Omega_2 B - 2\Omega_1 D)y^2 - \frac{2}{f}\Omega_2 E x^2 y \\ &+ \frac{2}{f}(\Omega_1 E - \Omega_2 F)xy^2 + \frac{2}{f}\Omega_1 F y^3 + O(\Omega^2).\end{aligned}\quad (\text{B.2})$$

Now, we know that the optical flow equations must have the form of (3.3) after the infinitesimal camera rotation transformation. If we substitute $u' = u + \delta u, v' = v + \delta v$ in the left-hand sides of (3.3), and $x' = x + \delta x, y' = y + \delta y, u'_0 = u_0 + \delta u_0, \dots, F' = F + \delta F$ into the right-hand sides, we obtain

$$\begin{aligned}\delta u &= \delta u_0 + \delta A x + A \delta x + \delta B y + B \delta y + \delta E x^2 + 2 E x \delta x \\ &+ \delta F x y + F y \delta x + F x \delta y + O(\Omega^2), \\ \delta v &= \delta v_0 + \delta C x + C \delta x + \delta D y + D \delta y + \delta E x y + E y \delta x \\ &+ E x \delta y + \delta F y^2 + 2 F y \delta y + O(\Omega^2).\end{aligned}\quad (\text{B.3})$$

Substituting the infinitesimal camera rotation transformation (2.4), we obtain

$$\begin{aligned}\delta u &= \delta u_0 - f(\Omega_2 A - \Omega_1 B) \\ &+ (\delta A - \Omega_3 B - f(2\Omega_2 E - \Omega_1 F))x \\ &+ (\delta B + \Omega_3 A - f\Omega_2 F)y \\ &+ \left(\delta E - \frac{1}{f}\Omega_2 A - \Omega_3 F\right)x^2 \\ &+ \left(\delta F + \frac{1}{f}(\Omega_1 A - \Omega_2 B) + 2\Omega_3 E\right)xy \\ &+ \left(\frac{1}{f}\Omega_1 B + \Omega_3 F\right)y^2 - \frac{2}{f}\Omega_2 E x^3 \\ &+ \frac{2}{f}(\Omega_1 E - \Omega_2 F)x^2 y + \frac{2}{f}\Omega_1 F x y^2 + O(\Omega^2), \\ \delta v &= \delta v_0 - f(\Omega_2 C - \Omega_1 D) + (\delta C - \Omega_3 D + \Omega_1 E)x \\ &+ (\delta D + \Omega_3 C - f(\Omega_2 E - 2\Omega_1 F))y \\ &- \left(\frac{1}{f}\Omega_2 C + \Omega_3 E\right)x^2 \\ &- \left(\delta E + \frac{1}{f}(\Omega_1 C - \Omega_2 D) - 2\Omega_3 F\right)xy \\ &+ \left(\delta F + \frac{1}{f}\Omega_1 D + \Omega_3 E\right)y^2 \\ &- \frac{2}{f}\Omega_2 E x^2 y + \frac{2}{f}(\Omega_1 E - \Omega_2 F)xy^2 \\ &+ \frac{2}{f}\Omega_1 F y^3 + O(\Omega^2).\end{aligned}\quad (\text{B.4})$$

Comparing (B.3) and (B.4), we obtain

$$\begin{aligned}\delta u_0 &= \Omega_3 v_0 + f\Omega_2 A - f\Omega_1 B + O(\Omega^2), \\ \delta v_0 &= -\Omega_3 u_0 + f\Omega_2 C - f\Omega_1 D + O(\Omega^2), \\ \delta A &= -\frac{2}{f}\Omega_2 u_0 + \frac{1}{f}\Omega_1 v_0 + \Omega_3 B + \Omega_3 C \\ &+ 2f\Omega_2 E - f\Omega_1 F + O(\Omega^2), \\ \delta B &= \frac{1}{f}\Omega_1 u_0 - \Omega_3 A + \Omega_3 D + f\Omega_2 F + O(\Omega^2), \\ \delta C &= -\frac{1}{f}\Omega_2 v_0 - \Omega_3 A + \Omega_3 D - f\Omega_1 E + O(\Omega^2),\end{aligned}$$

APPENDIX E
PROOF OF LEMMA 4

A "vector" is a set of three components a_1, a_2, a_3 , which are transformed by the camera rotation \mathbf{R} as

$$\begin{bmatrix} a'_1 \\ a'_2 \\ a'_3 \end{bmatrix} = \mathbf{R}^T \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}. \quad (\text{E.1})$$

This is a linear mapping, and hence it defines a representation of $SO(3)$, which is called the *vector representation*. If the camera rotation is infinitesimal, the values of a_1, a_2, a_3 are also changed infinitesimally into $a_1 + \delta a_1, a_2 + \delta a_2, a_3 + \delta a_3$, as follows:

$$\begin{aligned} \delta \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} &= \begin{bmatrix} 0 & \Omega_3 & -\Omega_2 \\ -\Omega_3 & 0 & \Omega_1 \\ \Omega_2 & -\Omega_1 & 0 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} + O(\Omega^2) \\ &= (\Omega_1 A_1 + \Omega_2 A_2 + \Omega_3 A_3) \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} + O(\Omega^2). \end{aligned} \quad (\text{E.2})$$

The infinitesimal generators A_1, A_2, A_3 are given by

$$\begin{aligned} A_1 &= \begin{bmatrix} & & -1 \\ & 1 & \\ -1 & & \end{bmatrix}, \quad A_2 = \begin{bmatrix} & -1 & \\ & & \\ 1 & & \end{bmatrix}, \\ A_3 &= \begin{bmatrix} & 1 & \\ -1 & & \\ & & \end{bmatrix}. \end{aligned} \quad (\text{E.3})$$

It can immediately be confirmed that they satisfy the commutation relations $[A_2, A_3] = -A_1, [A_3, A_1] = -A_2, [A_1, A_2] = -A_3$. The Casimir operator $H = -(A_1^2 + A_2^2 + A_3^2)$ becomes

$$H = \begin{bmatrix} 2 & & \\ & 2 & \\ & & 2 \end{bmatrix}. \quad (\text{E.4})$$

Since this is $1(1+1)$ times the $(2 \times 1 + 1)$ -dimensional unit matrix, the vector representation is the irreducible representation \mathfrak{D}_1 of weight 1 of $SO(3)$.

Consider a traceless symmetric tensor $\mathbf{B} = (b_{ij}), i, j = 1, 2, 3$. The components are transformed by the camera rotation \mathbf{R} as

$$\begin{bmatrix} b'_{11} & b'_{12} & b'_{13} \\ b'_{21} & b'_{22} & b'_{23} \\ b'_{31} & b'_{32} & b'_{33} \end{bmatrix} = \mathbf{R}^T \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{bmatrix} \mathbf{R}. \quad (\text{E.5})$$

This is a linear mapping from $b_{ij}, i, j = 1, 2, 3$, onto $b'_{ij}, i, j = 1, 2, 3$, and hence it defines a representation of $SO(3)$. Since tensor \mathbf{B} is traceless and symmetric, there are only five independent elements. Let us take $b_{11}, b_{22}, b_{12}, b_{31}, b_{32}$. If the rotation is infinitesimal, we obtain

$$\begin{aligned} \delta \begin{bmatrix} b_{11} \\ b_{22} \\ b_{12} \\ b_{31} \\ b_{32} \end{bmatrix} &= \begin{bmatrix} & & 2\Omega_3 & -2\Omega_2 & \\ & & -2\Omega_3 & & 2\Omega_1 \\ -\Omega_3 & \Omega_3 & & \Omega_1 & -\Omega_2 \\ 2\Omega_2 & \Omega_2 & -\Omega_1 & & \Omega_3 \\ -\Omega_1 & -2\Omega_1 & \Omega_2 & -\Omega_3 & \end{bmatrix} \begin{bmatrix} b_{11} \\ b_{22} \\ b_{12} \\ b_{31} \\ b_{32} \end{bmatrix} + O(\Omega^2) \\ &= (\Omega_1 A_1 + \Omega_2 A_2 + \Omega_3 A_3) \begin{bmatrix} b_{11} \\ b_{22} \\ b_{12} \\ b_{31} \\ b_{32} \end{bmatrix} + O(\Omega^2). \end{aligned} \quad (\text{E.6})$$

The infinitesimal generators A_1, A_2, A_3 are given by

$$\begin{aligned} A_1 &= \begin{bmatrix} & & & 2 \\ & & 1 & \\ & & & -1 \\ -1 & -2 & & \end{bmatrix}, \\ A_2 &= \begin{bmatrix} & & -2 & \\ & & & -1 \\ 2 & 1 & & \\ & & & 1 \end{bmatrix}, \\ A_3 &= \begin{bmatrix} & & 2 & \\ & & -2 & \\ -1 & 1 & & \\ & & & 1 \\ & & & & -1 \end{bmatrix}, \end{aligned} \quad (\text{E.7})$$

and it is easy to confirm that these infinitesimal generators satisfy the commutation relations $[A_2, A_3] = -A_1$, $[A_3, A_1] = -A_2$, $[A_1, A_2] = -A_3$. The Casimir operator $H = -(A_1^2 + A_2^2 + A_3^2)$ becomes

$$H = \begin{bmatrix} 6 & & & & \\ & 6 & & & \\ & & 6 & & \\ & & & 6 & \\ & & & & 6 \end{bmatrix}. \quad (\text{E.8})$$

Since this is $2(2 + 1)$ times the $(2 \times 2 + 1)$ -dimensional unit matrix, the representation defined by the traceless symmetric tensor is the irreducible representation \mathcal{D}_2 of weight 2 of $SO(3)$.

Now, let us go back to (4.1). Take the following set of vectors as a basis of the eight-dimensional representation space:

$$\begin{bmatrix} -f \\ \\ \\ -1/f \end{bmatrix} \begin{bmatrix} f \\ \\ \\ 1/f \end{bmatrix} \begin{bmatrix} \\ -1 \\ 1 \\ \\ \end{bmatrix} \begin{bmatrix} \\ \\ \\ 2 \\ 1 \end{bmatrix} \begin{bmatrix} \\ \\ \\ 1 \\ 2 \end{bmatrix} \begin{bmatrix} \\ \\ \\ 1 \\ 1 \end{bmatrix} \begin{bmatrix} f \\ \\ \\ -1/f \end{bmatrix} \begin{bmatrix} \\ f \\ \\ -1/f \end{bmatrix} \quad (\text{E.9})$$

With respect to this basis, the Casimir operator of (D.2) is diagonalized as follows:

$$H' = \begin{bmatrix} 2 & & & & & & & \\ & 2 & & & & & & \\ & & 2 & & & & & \\ & & & 6 & & & & \\ & & & & 6 & & & \\ & & & & & 6 & & \\ & & & & & & 6 & \\ & & & & & & & 6 \end{bmatrix}. \quad (\text{E.10})$$

The vector consisting of the flow parameters u_0, \dots, F is expressed as the following linear combination of the new basis vectors:

$$\begin{bmatrix} u_0 \\ v_0 \\ A \\ B \\ C \\ D \\ E \\ F \end{bmatrix} = -\frac{v_0/f + fF}{2} \begin{bmatrix} -f \\ \\ \\ -1/f \end{bmatrix} + \frac{u_0/f + fE}{2} \begin{bmatrix} f \\ \\ \\ 1/f \end{bmatrix} + \frac{C - B}{2} \begin{bmatrix} \\ -1 \\ 1 \\ \\ \end{bmatrix} + \frac{2A - D}{3} \begin{bmatrix} \\ \\ \\ 2 \\ 1 \end{bmatrix} + \frac{2D - A}{3} \begin{bmatrix} \\ \\ \\ 1 \\ 2 \end{bmatrix} \\ + \frac{B + C}{2} \begin{bmatrix} f \\ \\ \\ -1/f \end{bmatrix} + \frac{u_0/f - fE}{2} \begin{bmatrix} f \\ \\ \\ -1/f \end{bmatrix} + \frac{v_0/f - fF}{2} \begin{bmatrix} \\ f \\ \\ -1/f \end{bmatrix} \quad (\text{E.11})$$

Hence, the new components with respect this new basis are given by a_1, a_2, a_3 of (4.5), and $b_{11}, b_{22}, b_{12}, b_{31}, b_{32}$ of (4.6). In terms of these new components, (4.1) is rewritten as

$$\delta \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ b_{11} \\ b_{22} \\ b_{13} \\ b_{31} \\ b_{32} \end{bmatrix} = \begin{bmatrix} \Omega_3 & -\Omega_2 & & & & & & \\ -\Omega_3 & & \Omega_1 & & & & & \\ \Omega_2 & -\Omega_1 & & & & & & \\ \hline & & & 2\Omega_3 & -2\Omega_2 & & & \\ & & & -2\Omega_3 & & 2\Omega_1 & & \\ & & & & & & \Omega_1 & -\Omega_2 \\ -\Omega_3 & \Omega_3 & & & & & & \\ 2\Omega_2 & \Omega_2 & -\Omega_1 & & & & \Omega_3 & \\ -\Omega_1 & -2\Omega_1 & \Omega_2 & \Omega_3 & & & & \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ b_{11} \\ b_{22} \\ b_{13} \\ b_{31} \\ b_{32} \end{bmatrix} + O(\Omega^2). \quad (\text{E.12})$$

The coefficient matrix is written as $\Omega_1 A'_1 + \Omega_2 A'_2 + \Omega_3 A'_3$, where the infinitesimal generators A'_1, A'_2, A'_3 are given by

$$A'_1 = \begin{bmatrix} & 1 & & & & & & \\ -1 & & & & & & & \\ \hline & & & & & & & \\ & & & & & & 2 & \\ & & & & & & & 1 \\ & & & & & & & -1 \\ & & & & & & & -1 \\ -1 & -2 & & & & & & \end{bmatrix},$$

$$A'_2 = \begin{bmatrix} & -1 & & & & & & \\ 1 & & & & & & & \\ \hline & & & & & & & \\ & & & & & & -2 & \\ & & & & & & & -1 \\ & & & & & & & 2 \\ & & & & & & 1 & \\ & & & & & & & 1 \end{bmatrix},$$

$$A'_3 = \begin{bmatrix} & 1 & & & & & & \\ -1 & & & & & & & \\ \hline & & & & & & & \\ & & & & & & 2 & \\ & & & & & & -2 & \\ & & & & & & & -1 \\ & & & & & & -1 & 1 \\ & & & & & & & -1 \\ & & & & & & & 1 \end{bmatrix}. \quad (\text{E.13})$$

Since these infinitesimal generators are the direct sums of

those for a vector (E.3) and those for a tensor (E.7), a_1, a_2, a_3 are transformed as a vector and $b_{ij}, i, j = 1, 2, 3$, are transformed as a traceless symmetric tensor.

APPENDIX F PROOF OF PROPOSITION 3

Equation (E.1) is also written as the following tensor transformation rule for an antisymmetric tensor.

$$\begin{bmatrix} 0 & -a'_3 & a'_2 \\ a'_3 & 0 & -a'_1 \\ -a'_2 & a'_1 & 0 \end{bmatrix} = \mathbf{R}^T \begin{bmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{bmatrix} = \mathbf{R}. \quad (\text{F.1})$$

Hence, we can define a new tensor C by

$$\begin{bmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{bmatrix} = \begin{bmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{bmatrix}. \quad (\text{F.2})$$

Thus, we obtain (4.9)-(4.11).

APPENDIX G PROOF OF (6.1)

Consider the mutually orthogonal unit vectors $i = (1, 0, 0)^T, j = (0, 1, 0)^T, k = (0, 0, 1)^T$ along the X-, Y-, and Z-axes. By the definition of \mathbf{R}_1 , if the scene is rotated by \mathbf{R}_1 , the vectors i, j, k move onto vectors e_1, e_2, e_3 . Similarly, if the scene is rotated by $\mathbf{R}_1^{(i)T}$, the vectors i, j, k move onto vectors e'_1, e'_2, e'_3 . Hence, if the scene is rotated by $\mathbf{R}_2^{(i)} \mathbf{R}_1^T$, the vectors e_1, e_2, e_3 move onto the vectors e'_1, e'_2, e'_3 . Conversely, if the camera is rotated by $\mathbf{R}_2^{(i)} \mathbf{R}_1^T$, the vectors e'_1, e'_2, e'_3 move onto the vectors e_1, e_2, e_3 .

APPENDIX H PROOF OF (7.5) AND (7.6)

We follow the formulation of Kanatani [7]. The 3-D recovery solution is given in terms of $a_i, i = 1, 2, 3$, and

b_{ij} , $i, j = 1, 2, 3$ as follows. First, define T, S, L by

$$T = b_{33}, \quad S = (b_{11} - b_{22}) + 2ib_{12}, \quad L = b_{31} + ib_{32}, \quad (\text{H.1})$$

where i is the imaginary unit.¹¹ Then, the translation velocity components scaled by the depth r are given by

$$a/r = \text{Re}[L] + a_2, \quad b/r = \text{Im}[L] - a_1, \quad c/r = \alpha, \quad (\text{H.2})$$

where α is the middle of the three real roots of the following cubic equation,¹² and Re and Im denote the real and the imaginary parts.

$$X^3 - 3TX^2 + \frac{1}{4}(9T^2 - |S|^2 - 4|L|^2)X + \frac{1}{2}(\text{Re}[L^2S^*] + 3T|L|^2) = 0. \quad (\text{H.3})$$

Here, $*$ and $|\cdot|$ denote the complex conjugate and the modulus of the complex number. Next, using the solution α , compute the following two complex quantities.

$$P = \frac{1}{\alpha}(L \pm \sqrt{L^2 - \alpha S}), \quad W = i(L \mp \sqrt{L^2 - \alpha S}). \quad (\text{H.4})$$

Then, the surface gradient components p, q and the rotation velocity components $\omega_1, \omega_2, \omega_3$ are given as follows.

$$\begin{aligned} p &= \text{Re}[P], & q &= \text{Im}[P], \\ \omega_1 &= \text{Re}[W + iL] + a_1, & \omega_2 &= \text{Im}[W + iL] + a_2, \\ \omega_3 &= \frac{1}{2}\text{Re}[PW^*] + a_3. \end{aligned} \quad (\text{H.6})$$

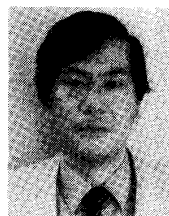
From these results follows the observation in Section VII.

ACKNOWLEDGMENT

Part of this work was done while the author was visiting the Center for Automation Research, University of Maryland. He thanks A. Rosenfeld and L. Davis for the support and helpful discussions. He also thanks A. Waxman and M. Subbarao for discussions on optical flow analysis. He also received helpful comments from S. Amari and K. Sugihara of the University of Tokyo.

REFERENCES

- [1] T.-C. Chou and K. Kanatani, "Recovering 3D rigid motions without correspondence," in *Proc. 1st Int. Conf. Computer Vision*, London, June 1987, pp. 534-538.
- [2] I. M. Gel'fand, R. A. Minlos, and Z. Y. Shapiro, *Representation of the Rotation and Lorentz Groups and Their Applications*. Oxford, England: Pergamon, 1963.
- [3] M. Hamermesh, *Group Theory and Its Application to Physical Problems*. Reading, MA: Addison-Wesley, 1964.
- [4] K. Kanatani, "Detecting the motion of a planar surface by line and surface integrals," *Comput. Vision Graphics Image Processing*, vol. 29, pp. 13-22, 1985.
- [5] —, "Structure from motion without correspondence: General principle," in *Proc. DARPA Image Understanding Workshop*, Miami Beach, FL, Dec. 1985, pp. 107-116.
- [6] —, "Structure and motion from optical flow under orthographic projection," *Comput. Vision Graphics Image Processing*, vol. 35, pp. 181-199, 1986.
- [7] —, "Structure and motion from optical flow under perspective projection," *Comput. Vision Graphics Image Processing*, vol. 38, pp. 122-146, 1987.
- [8] —, "Coordinate rotation invariance of image characteristics for 3D shape and motion recovery," in *Proc. 1st Int. Conf. Computer Vision*, London, June 1987, pp. 55-64.
- [9] —, "Camera rotation invariance of image characteristics," *Comput. Vision Graphics Image Processing*, vol. 39, pp. 328-354, 1987.
- [10] —, "Constraints on length and angle," *Comput. Vision Graphics Image Processing*, to be published.
- [11] K. Kanatani and T.-C. Chou, "Shape from texture: General principle," in *Proc. IEEE Conf. Comput. Vision Pattern Recognition*, Miami Beach, FL, June 1986, pp. 578-583.
- [12] H. C. Longuet-Higgins, "The visual ambiguity of a moving plane," *Proc. Roy. Soc. London*, vol. B-223, pp. 165-175, 1984.
- [13] G. F. Smith, "On isotropic functions of symmetric tensors, skew-symmetric tensors and vectors," *Int. J. Eng. Sci.*, vol. 9, pp. 899-916, 1971.
- [14] A. J. M. Spencer, "Theory of invariants," in *Continuum Physics*, vol. 1, A. C. Eringen, Ed. New York: Academic, 1971, pp. 239-353.
- [15] M. Subbarao and A. M. Waxman, "Closed form solution to image flow equations for planar surfaces in motion," *Comput. Vision Graphics Image Processing*, vol. 35, pp. 181-199, 1986.
- [16] C.-C. Wang, "A new representation theorem for isotropic functions, Part 1," *Arch. Ration. Mech. Anal.*, vol. 36, pp. 166-197, 1970.
- [17] C.-C. Wang, "A new representation theorem for isotropic functions, Part 2," *Arch. Ration. Mech. Anal.*, vol. 36, pp. 198-223, 1970.
- [18] —, "Corrigendum to my papers on representation for isotropic functions," *Arch. Ration. Mech. Anal.*, vol. 43, pp. 392-395, 1971.
- [19] A. M. Waxman, "An image flow paradigm," in *Proc. 2nd IEEE Workshop Computer Vision: Representation and Control*, Annapolis, MD, Apr. 1984, pp. 49-57.
- [20] A. M. Waxman, B. Kamgar-Parsi, and M. Subbarao, "Closed-form solutions to image flow equations," in *Proc. 1st Int. Conf. Computer Vision*, London, June 1987, pp. 12-24.
- [21] A. M. Waxman and K. Wahn, "Contour evolution, neighborhood deformation, and global image flow: Planar surfaces in motion," *Int. J. Robotics Res.*, vol. 4, no. 3, pp. 95-108, 1985.
- [22] —, "Image flow theory: A framework for 3-D inference from time-varying imagery," in *Advances in Computer Vision*, C. M. Brown, Ed. Hillsdale, NJ: Erlbaum, 1987.
- [23] A. M. Waxman and S. Ullman, "Surface structure and three-dimensional motion from image flow kinematics," *Int. J. Robotics Res.*, vol. 4, no. 3, pp. 72-94, 1985.
- [24] H. Weyl, *The Classical Groups, Their Invariants and Representations*, 2nd ed. Princeton, NJ: Princeton University Press, 1946.



Ken-ichi Kanatani was born in Japan in 1947. From 1967 to 1979, he studied applied mathematics at the Department of Mathematical Engineering and Instrumentation Physics, Faculty of Engineering, University of Tokyo, Tokyo, Japan, and received the B.Eng., M.Eng., and Ph.D. degrees in 1972, 1974, and 1979, respectively. From 1969 to 1970 of that period, he studied physics at Case Western Reserve University, Cleveland, OH.

In 1979 he became Assistant Professor at the Department of Computer Science, Gunma University, Kiryu, Gunma, Japan, and in 1983 Associate Professor there. He teaches applied mathematics, numerical analysis, image processing, and computer vision. From 1985 to 1986, he visited the Center for Automation Research, University of Maryland, College Park. His interests are applications of advanced mathematics to various engineering problems. His research areas have included the mechanics of microstructured materials such as polycrystalline metals and granular materials, and their application to soil mechanics and powder technology. His current interest is image understanding and 3-D reconstruction problems of computer vision.