COMPUTER VISION, GRAPHICS, AND IMAGE PROCESSING 39, 328–354 (1987)

Camera Rotation Invariance of Image Characteristics

KEN-ICHI KANATANI*

Center for Automation Research, University of Maryland, College Park, Maryland 20742

Received May 9, 1986; accepted August 5, 1986

The image transformation due to camera rotation relative to a stationary scene is analyzed, and the associated transformation rules of "features" given by weighted averaging of the image are derived by considering infinitesimal generators on the basis of group representation theory. Three-dimensional vectors and tensors are reduced to two-dimensional invariants on the image plane from the viewpoint of projective geometry. Three-dimensional invariants and camera rotation reconstruction are also discussed. The result is applied to the shape recognition problem when camera rotation is involved. © 1987 Academic Press, Inc.

1. INTRODUCTION

The problem we consider in this paper is as follows. Suppose the camera is rotated by a certain angle around its lens center relative to a stationary scene. Then, a different projected image is seen on the image plane. However, since a point on the image plane corresponds to a *ray* in the 3D scene, occlusion is not affected by camera rotation. If the amount of camera rotation is known, the original image can be recovered. (Here, we do not consider the effect of the image boundary. We assume that the image plane is sufficiently large and that the object or scene of interest is always included in the field of view.) This means that the *information content* of the image is not affected by the 2D image transformation induced by the camera rotation.

Suppose the viewed image is characterized by a finite number of parameters or *features*. If the camera is rotated, the image is also changed so that the features change their values. If the set of features is *invariant* in the sense that these new values are completely determined by the original values and the amount of the camera rotation, we can predict the values of the features which would be obtained if the camera were rotated by a given amount. Conversely, if we are given two views of the same object obtained from different camera orientations, we can reconstruct the amount of camera rotation R which would transform the values of the features to prescribed values. An important fact is that in this process *we need not know the point-to-point correspondence*. All computations are based on the observed features, which are global quantities.

These considerations are very important in many problems of computer vision and pattern recognition when the camera orientation is controlled by a computer. Even if the camera is fixed, various types of analysis of the image become easy if we apply to the image the transformation equivalent to camera rotation. This technique is used for the shape-from-texture problem by Kanatani and Chou [7] and for the interpretation of lengths and angles by Kanatani [6]. A similar analysis is done when the object is moving and we are observing the optical flow (Kanatani [5]). In this paper, we will discuss, as a typical example, the center of gravity and principal axes

*Current address: Department of Computer Science, Gunma University, Kiryu, Gunma 376, Japan.

0734-189X/87 $3.00

of a given region to see how the invariant properties can be utilized to recognize the shape and to reconstruct the (actual or hypothetical) camera rotation.

2. CAMERA ROTATION AND INVARIANT FEATURES

Let f be the *focal length* of the camera. The camera image is thought of as the projection onto an image plane located at distance f from the viewpoint O; a point P in the scene is projected onto the intersection of the image plane with the *ray*, connecting point P and the viewpoint O. Let us choose an XYZ-coordinate system such that the viewpoint O is at the origin and the Z axis coincides with the camera optical axis. Choose an xy-coordinate system in such a way that the x and y axes are parallel to the X and Y axes with $(0, 0, f)$ as the origin. This xy plane plays the role of the image plane (Fig. 1). A point (X, Y, Z) in the scene is projected onto (x, y) on the image plane, where

$$x = fX/Z, \qquad y = fY/Z. \tag{2.1}$$

Consider a camera rotation around O (lens center) and the induced transformation of the image. Suppose the camera is rotated by rotation matrix R, which is an orthogonal matrix, i.e., $RR^{\mathrm{T}} = I$. Then, the point in the scene which was seen at (x, y) now moves to another point (x', y') given by

THEOREM 1. *The image transformation induced by camera rotation $R = (r_{ij})$ is given by*

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

Proof. A rotation of the camera by R is equivalent to the rotation of the scene in the opposite sense. If the scene is rotated by $R^{-1}(= R^{\mathrm{T}})$, where $^{\mathrm{T}}$ denotes transpose, point (X, Y, Z) moves to point (X', Y', Z'), where

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

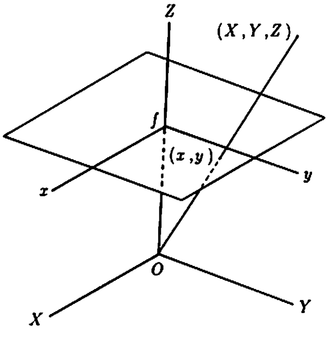

FIG. 1. The XYZ-coordinate system is fixed to the camera, the origin O being the camera lens center. The image plane is taken to be $Z = f$, where f is the camera focal length. A point (X, Y, Z) in the scene is projected onto point (x, y) on the image plane.

This point is projected to (x', y') on the image plane, where $x' = fX'/Z'$ and $y' = fY'/Z'$. Combining this with Eqs. (2.1), we obtain Eq. (2.2).

It should be emphasized that the image transformation due to camera rotation does not require any knowledge about the scene and that the transformation has an inverse which is obtained by interchanging R and R^T. This means that transformations of the form of Eq. (2.2), which form a subgroup of the 2D *projective transformation group*, do not alter the information content of the image as long as the image boundary is ignored. (In this paper, we always regard the portion of the image near the boundary as *unimportant*.) In the following, some basic results from projective geometry are summarized in a way that is convenient in our consideration of the image plane transformation.

Suppose the image is characterized by a finite number of parameters J_i, $i = 1, 2, \ldots, N$, which we call *features* of the image (Amari [1, 2]). (They are called *properties* in Rosenfeld and Kak [9].) If the image is transformed by Eqs. (2.2) as a result of camera rotation R, these features take different values J_i', $i = 1, \ldots, N$. We say that a set of features J_i, $i = 1, \ldots, N$, is *invariant* if the values of J_i', $i = 1, \ldots, N$, are determined by the values of J_i, $i = 1, \ldots, N$, and the amount of camera rotation R alone (cf. Weyl [15]). This definition suggests that an invariant set of features describes some aspects of the image that are "inherent to the scene itself" and are independent of the camera orientation.

Let J_i, $i = 1, \ldots, N$, be an invariant set of features. We say the set is *reducible* if it splits, after an appropriate rearrangement, into two or more sets of features, each of which is itself invariant separately. If no further reduction is possible, we say that the set of features is *irreducible* (cf. Weyl [15]). This definition suggests that an irreducible invariant set of features describes a "single" characteristic inherent to the scene while a reducible set describes two or more different characteristics at the same time.

If a quantity c does not change its value under transformation (2.2), i.e.,

$$c' = c, \tag{2.4}$$

under camera rotation R, we call it a *scalar*. Obviously, a scalar is itself an invariant and is irreducible. Hence, it describes a characteristic inherent to the scene.

If a pair, a, b of numbers is transformed as x, y of transformation (2.2), i.e.,

$$a' = f\frac{r_{11}a + r_{21}b + r_{31}f}{r_{13}a + r_{23}b + r_{33}f}, \qquad b' = f\frac{r_{12}a + r_{22}b + r_{32}f}{r_{13}a + r_{23}b + r_{33}f}, \tag{2.5}$$

we call it a *point*. Note that any pair of numbers can be interpreted as a position *on the image plane*. However, it is interpreted as indicating a position *in the scene* if and only if it is transformed as a point. It is easily proved that a point is also an invariant set of features and is irreducible.

A line on the image plane is expressed in the form

$$Ax + By + C = 0. \tag{2.6}$$

Here, the ratio $A : B : C$ alone has a geometrical meaning; A, B, C and cA, cB, cC for a non-zero scalar c define one and the same line. In order to emphasize this fact, let us write $A : B : C$ to express a line. If transformation (2.2) is applied, line (2.6) is mapped into

$$A'x' + B'y' + C' = 0, \tag{2.7}$$

as in

THEOREM 2. *A line $A : B : C$ on the image plane is transformed by camera rotation R into the line*

$$\begin{aligned} A' : B' : C' = {} & r_{11}A + r_{21}B + r_{31}C/f : r_{12}A \\ & + r_{22}B + r_{32}C/f : f(r_{13}A + r_{23}B) + r_{33}C. \end{aligned} \tag{2.8}$$

Proof. In view of Eqs. (2.1), Eq. (2.6) is written as $A(fX/Z) + B(fY/Z) + C = 0$, or

$$[A\ B\ C/f]\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = 0. \tag{2.9}$$

From Eq. (2.3), we find that $A, B, C/f$ are transformed as a vector, i.e.,

$$\begin{bmatrix} A' \\ B' \\ C'/f \end{bmatrix} = R^{\mathrm{T}} \begin{bmatrix} A \\ B \\ C/f \end{bmatrix}, \tag{2.10}$$

from which Eq. (2.8) is obtained.

If the ratio of three given quantities A, B, C is transformed by Eq. (2.8) under camera rotation, we call it a *line* and write it as $A : B : C$. It is easily proved that it is an invariant set of features and is irreducible. As in the case of a point, any triplet of numbers can be interpreted as a line on the image plane, but it is interpreted as a line in the scene if and only if it is transformed as a line.

All the invariant properties considered in this paper are invariant with respect to the "projective transformations" of the form of Eqs. (2.2). In traditional "projective geometry," all equations are written in terms of "homogeneous coordinates" defined in a "projective space" (cf. Naeve and Eklundh [8]). If we regard the xy-image plane (with the "line at infinity" added) as a two-dimensional projective space and introduce homogeneous coordinates, Eqs. (2.2) are rewritten as a linear transformation. The "point" and "line" defined here are mutually "dual" and expressed exactly dually in homogeneous coordinates.

However, the purpose of this paper is to deal with applications of the ideas of projective geometry, and in dealing with real images the xy-Cartesian coordinate system is most convenient. Therefore, in the following, we express all the invariant properties in terms of the xy-"inhomogeneous" coordinates of the image plane. The aim of this paper is to translate the results known in projective geometry into "managable" forms and to demonstrate the practical use of this type of knowledge.

3. IRREDUCIBLE REDUCTION OF 3D VECTORS AND TENSORS

Consider three quantities a, b, c which are transformed as a 3D *vector*, i.e.,

$$\begin{bmatrix} a' \\ b' \\ c' \end{bmatrix} = R^{\mathrm{T}} \begin{bmatrix} a \\ b \\ c \end{bmatrix}, \tag{3.1}$$

for camera rotation R. (Note that the rotation matrix R is transposed because we adopted the convention that R is the amount of "camera rotation".) This is an invariant set of features but is not irreducible because

LEMMA 1. *If a, b, c are transformed as a 3D vector, then the length $\sqrt{a^2 + b^2 + c^2}$ is a scalar.*

There are two ways, mutually dual, to interpret a 3D vector a, b, c as irreducible sets of features. One way is to regard $fa/c, fb/c$ as a point and the length $\sqrt{a^2 + b^2 + c^2}$ as its *intensity*, which is a scalar. We can easily check from Theorem 1 that

LEMMA 2. *If a, b, c are transformed as a 3D vector, then $fa/c, fb/c$ are transformed as a point.*

Hence a pair $fa/c, fb/c$ has an interpretation as a point invariant on the image plane in the sense described above. Here, we allow the case $c = 0$, regarding it as a point located at infinity. We also make the convention that the intensity is negative if $c < 0$. If we imagine that the 3D vector (a, b, c) is emanating from the origin O (or the camera lens center) of the XYZ-coordinate system, the point $(fa/c, fb/c)$ is the intersection of the image plane with the ray defined by the 3D vector (a, b, c).

Another way to represent a 3D vector on the image plane is to regard $a : b : fc$ as a line and the length $\sqrt{a^2 + b^2 + c^2}$ as its intensity. We can easily check from Theorem 2 that

LEMMA 3. *If a, b, c are transformed as a 3D vector, then $a : b : fc$ is transformed as a line.*

Hence, equation $az + by + fc = 0$ has an interpretation as a line invariant on the image plane in the sense described above. If we imagine that the 3D vector (a, b, c) is emanating from the origin O (or the camera lens center) of the XYZ-coordinate system, the line $ax + by + fc = 0$ is the intersection of the image plane with the plane passing through the origin O and perpendicular to (a, b, c). As before, we allow the case of $a = b = 0$, regarding the line as located at infinity, and make the convention that the intensity is negative if $c < 0$.

The above results are summarized as follows:

THEOREM 3. *A 3D vector is an invariant feature set. It can be irreducibly reduced into a point and a scalar or into a line and a scalar on the image plane.*

Next, consider nine elements A_{ij}, $i, j = 1, 2, 3$, which are transformed by camera rotation R as a 3D *tensor*, i.e.,

$$\begin{bmatrix} A'_{11} & A'_{12} & A'_{13} \\ A'_{21} & A'_{22} & A'_{23} \\ A'_{31} & A'_{32} & A'_{33} \end{bmatrix} = R^{\mathrm{T}} \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix} R. \tag{3.2}$$

By definition, this is an invariant set of features. However, it is reducible. First, it can be decomposed into a *symmetric part* and an *antisymmetric part* (or *skew part*)

$$\begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix}$$
$$= \begin{bmatrix} A_{11} & (A_{12}+A_{21})/2 & (A_{31}+A_{13})/2 \\ (A_{12}+A_{21})/2 & A_{22} & (A_{23}+A_{32})/2 \\ (A_{31}+A_{13})/2 & (A_{23}+A_{32})/2 & A_{33} \end{bmatrix}$$
$$+ \begin{bmatrix} 0 & (A_{12}-A_{21})/2 & -(A_{31}-A_{13})/2 \\ -(A_{12}-A_{21})/2 & 0 & (A_{23}-A_{32})/2 \\ (A_{31}-A_{13})/2 & -(A_{23}-A_{32})/2 & 0 \end{bmatrix}, \quad (3.3)$$

and each part is transformed as a 3D tensor by Eq. (3.2) separately. Moreover, it can be verified that the three independent elements $(A_{23} - A_{32})/2, (A_{31} - A_{13})/2, (A_{12} - A_{21})/2$ of the antisymmetric part are transformed as a 3D vector. Hence, they are, from Theorem 3, irreducibly reduced into a point and a scalar or into a line and a scalar.

Suppose $A = (A_{ij})$ is already a symmetric 3D tensor. As is well known, such a tensor is represented by three mutually perpendicular unit vectors e_1, e_2, e_3 indicating the principal axes and the corresponding principal values $\sigma_1, \sigma_2, \sigma_3$ in the form

$$A = \sigma_1 e_1 e_1^{\mathrm{T}} + \sigma_2 e_2 e_2^{\mathrm{T}} + \sigma_3 e_3 e_3^{\mathrm{T}}. \quad (3.4)$$

Here, this representation does not change if e_1 (or e_2 or e_3) is replaced by $-e_1$ (or $-e_2$ or $-e_3$). (If two of $\sigma_1, \sigma_2, \sigma_3$ are identical, the corresponding principal axes are not unique and can be arbitrarily rotated rigidly around the remaining one. If all of $\sigma_1, \sigma_2, \sigma_3$ are identical, the orientations of e_1, e_2, e_3 are completely arbitrary as long as they are mutually orthogonal.)

The three principal values are scalars, each of which is an invariant irreducible feature. On the other hand, if we determine the orientations of two of the three principal axis orientations, say e_1 and e_2, the orientation of the remaining one is uniquely determined. (e_3 and $-e_3$ indicate the same orientation.) As is shown in Theorem 3, the orientations of e_1 and e_2 are represented by two points on the image plane. (If we replace e_1 (or e_2) by $-e_1$ (or $-e_2$), the corresponding points are unchanged as desired.) However, since e_1 and e_2 are perpendicular, one of the two points and the line connecting the two points are sufficient; if one point on the image plane and a line through it are given, the three orientations are determined (Appendix A). Thus, we obtain

THEOREM 4. *A* 3D *tensor is invariantly reduced to its symmetric part and its antisymmetric part. The antisymmetric part is irreducibly reduced into a point and a scalar or a line and a scalar. The symmetric part is irreducibly reduced to three scalars, a point and a line passing through it.*

4. INFINITESIMAL GENERATORS OF THE IMAGE TRANSFORMATION

Let $F(x, y)$ represent an observed image. This may be the intensity of the gray-level or a vector-valued function corresponding to R, B, and G. Here, the value of $F(x, y)$ is assumed to be inherent to the scene and independent of the viewing orientation. Chromaticity, for example, has this property. Furthermore, $F(x, y)$ is assumed to be of *finite support*, i.e., $F(x, y)$ is zero at a sufficiently large distance from the origin of the image plane.

Let us write the transformation of Eq. (2.2), which is determined by the rotation matrix R, symbolically as

$$(x', y') = M[R](x, y). \tag{4.1}$$

Then, we can see the (transposed) *homomorphism* in the sense that

$$M[R_2] \circ M[R_1] = M[R_1 R_2]. \tag{4.2}$$

Now, define the *rotation operator* T_R acting on image $F(x, y)$ by

$$T_R F(x, y) \equiv F(M[R^{\mathrm{T}}](x, y)). \tag{4.3}$$

In view of our assumptions of image value constancy and finite support, the function $T_R F(x, y)$ describes the image we observe if the image plane undergoes the transformation (2.2). Operator T_R induces a (transposed) *representation* of the 3D rotation group $SO(3)$ in the sense that

$$T_{R_2 R_1} = T_{R_1} \circ T_{R_2}. \tag{4.4}$$

As is well known, this representation is completely determined once its behavior for infinitesimal rotations (i.e., its *Lie algebra*) is known, since $SO(3)$ is a compact Lie group.

A 3D rotation is specified by the *rotation axis* (n_1, n_2, n_3), which is taken to be a unit vector, and the *rotation angle* Ω (rad) screwwise around it. As is well known, the corresponding rotation matrix is given by

$$R = \begin{bmatrix} \cos\Omega + (1-\cos\Omega)(n_1)^2 & (1-\cos\Omega)n_1 n_2 - \sin\Omega n_3 & (1-\cos\Omega)n_1 n_3 + \sin\Omega n_2 \\ (1-\cos\Omega)n_2 n_1 + \sin\Omega n_3 & \cos\Omega + (1-\cos\Omega)(n_2)^2 & (1-\cos\Omega)n_2 n_3 - \sin\Omega n_1 \\ (1-\cos\Omega)n_3 n_1 - \sin\Omega n_2 & (1-\cos\Omega)n_3 n_2 + \sin\Omega n_1 & \cos\Omega + (1-\cos\Omega)(n_3)^2 \end{bmatrix}. \tag{4.5}$$

If the rotation is *infinitesimally small*, i.e., Ω is infinitesimally small, the rotation matrix takes the form $R = I + \delta R + o(\Omega)$, where I is the unit matrix, δR is the matrix given by

$$\delta R = \begin{bmatrix} 0 & -\Omega_3 & \Omega_2 \\ \Omega_3 & 0 & -\Omega_1 \\ -\Omega_2 & \Omega_1 & 0 \end{bmatrix}, \tag{4.6}$$

and $o(\Omega)$ denotes higher order terms in Ω. (We let the context indicate whether these terms are scalars, vector, or tensors.) Here, we put $\Omega_1 = \Omega n_1$, $\Omega_2 = \Omega n_2$, and $\Omega_3 = \Omega n_3$.

If the rotation is infinitesimal, the transformation of Eqs. (2.2) becomes $x' = x + \delta x + o(\Omega)$ and $y' = y + \delta y + o(\Omega)$, where

$$\begin{aligned} \delta x &= -f\Omega_2 + \Omega_3 y + \frac{1}{f}(-\Omega_2 x + \Omega_1 y)x, \\ \delta y &= f\Omega_1 - \Omega_3 x + \frac{1}{f}(-\Omega_2 x + \Omega_1 y)y. \end{aligned} \tag{4.7}$$

Then, the image $F(x, y)$ also undergoes an infinitesimal change and becomes

$$F(x - \delta x, y - \delta y) = F(x, y) + \delta F(x, y) + o(\Omega), \tag{4.8}$$

and $\delta F(x, y)$ is given by

$$\begin{aligned} \delta F(x, y) &= -\frac{\partial F}{\partial x}\delta x - \frac{\partial F}{\partial y}\delta y \\ &= -\left[-f\Omega_2 + \Omega_3 y + \frac{1}{f}(-\Omega_2 x + \Omega_1 y)x\right]\frac{\partial F}{\partial x} \\ &\quad -\left[f\Omega_1 - \Omega_3 x + \frac{1}{f}(-\Omega_2 x + \Omega_1 y)y\right]\frac{\partial F}{\partial y} \\ &= -(\Omega_1 D_1 + \Omega_2 D_2 + \Omega_3 D_3)F(x, y), \end{aligned} \tag{4.9}$$

where the *infinitesimal generators* are defined by

$$\begin{aligned} D_1 &= \frac{xy}{f}\frac{\partial}{\partial x} + \left(f + \frac{y^2}{f}\right)\frac{\partial}{\partial y}, \qquad D_2 = -\left(f + \frac{x^2}{f}\right)\frac{\partial}{\partial x} - \frac{xy}{f}\frac{\partial}{\partial y}, \\ D_3 &= y\frac{\partial}{\partial x} - x\frac{\partial}{\partial y}. \end{aligned} \tag{4.10}$$

Hence, operator T_R becomes, for infinitesimal rotations,

$$T_R = I - (\Omega_1 D_1 + \Omega_2 D_2 + \Omega_3 D_3) + o(\Omega), \tag{4.11}$$

where I is the identity operator.

It can be checked easily that these infinitesimal generators satisfy the *commutation relations*

$$[D_1, D_2] = D_3, \qquad [D_2, D_3] = D_1, \qquad [D_3, D_1] = D_2, \tag{4.12}$$

where the *commutator* is defined by $[A, B] \equiv AB - BA$. Hence, a set of functions can be found which induces a representation of the 3D rotation group $SO(3)$ [3, 4].

As is well known, a set of functions which induces an *irreducible representation* is obtained as eigenfunctions of the *Casimir operator*

$$H \equiv -\left(D_1^2 + D_2^2 + D_3^2\right). \tag{4.13}$$

The eigenvalue is $l(l + 1)$ and the eigenspace is $(2l + 1)$ dimensional, where l is an integer or half-integer called the *weight* of the irreducible representation (cf. Gel'fand, Minlos, and Shapiro [3], Hammermesh [4]). In other words, the differential equation

$$HF = l(l+1)F, \qquad \text{or} \qquad \left(D_1^2 + D_2^2 + D_3^2\right)F + l(l+1)F = 0, \quad (4.14)$$

has $(2l+1)$ independent solutions, which become the basis of the irreducible representation D_l of weight l (Appendix B).

5. ADJOINT ROTATION AND FEATURE TRANSFORMATION

Let J be a feature of the image. To be precise, a feature is a *functional* mapping the image function $F(x, y)$ into a real number $J[F(x, y)]$. Consider a *linear* feature obtained by weighted averaging or *filtering*

$$J[F(x,y)] = \int m(x,y)F(x,y)\,dx\,dy. \quad (5.1)$$

Here, $m(x, y)$ is the filter weight function and integration is performed over the entire image plane. (Recall our assumption of finite support of $F(x, y)$.) If the camera is rotated by R, the image becomes $T_R F(x, y)$ by Eq. (4.2) and hence the corresponding feature becomes

$$J[T_R F(x,y)] = \int m(x,y)T_R F(x,y)\,dx\,dy. \quad (5.2)$$

We define the *adjoint rotation operator* T_R^* by

$$J[T_R F(x,y)] = \int T_R^* m(x,y)F(x,y)\,dx\,dy. \quad (5.3)$$

From this definition, we can set that operator T_R^* induces a representation of the 3D rotation group in the sense that

$$T_{R_2 R_1}^* = T_{R_2}^* \circ T_{R_1}^*. \quad (5.4)$$

Once we know how this adjoint rotation operator T_R^* acts, the transformation of such features is immediately computed for any given image. This is done by just considering infinitesimal transformations.

If the image is infinitesimally changed as in Eq. (4.8), feature J also undergoes an infinitesimally small change $J \to J + \delta J + o(\Omega)$. Substitution of Eq. (4.9) and integration by parts yield

$$\delta J = \int (\Omega_1 D_1^* + \Omega_2 D_2^* + \Omega_3 D_3^*)m(x,y)F(x,y)\,dx\,dy, \quad (5.5)$$

where D_1^*, D_2^*, and D_3^* are the *adjoint infinitesimal generators* defined by

$$D_1^* = \frac{3y}{f} + \frac{xy}{f}\frac{\partial}{\partial x} + \left(f + \frac{y^2}{f}\right)\frac{\partial}{\partial y},$$
$$D_2^* = -\frac{3x}{f} - \left(f + \frac{x^2}{f}\right)\frac{\partial}{\partial x} - \frac{xy}{f}\frac{\partial}{\partial y}, \tag{5.6}$$
$$D_3^* = y\frac{\partial}{\partial x} - x\frac{\partial}{\partial y}.$$

In Eq. (5.5), no boundary terms appear due to our assumption of finite support for $F(x, y)$. Hence, operator T_R^* becomes, for infinitesimal rotations,

$$T_R^* = I + \Omega_1 D_1^* + \Omega_2 D_2^* + \Omega_3 D_3^* + o(\Omega). \tag{5.7}$$

It can be checked easily that these adjoint infinitesimal generators satisfy the commutation relations

$$[D_1^*, D_2^*] = D_3^*, \qquad [D_2^*, D_3^*] = D_1^*, \qquad [D_3^*, D_1^*] = D_2^*. \tag{5.8}$$

Hence, we can find a set of functions which induces a representation of the 3D rotation group $SO(3)$. Then, operator T_R^* acts as a linear transformation on them (cf. Gel'fand *et al.* [3]). As before, a basis of the irreducible representation D_l of weight l is obtained as $(2l + 1)$ eigenfunctions of the (adjoint) Casimir operator

$$H^* \equiv -\left(D_1^{*2} + D_2^{*2} + D_3^{*2}\right), \tag{5.9}$$

i.e., as $(2l + 1)$ independent solutions of the differential equation

$$H^* m = l(l+1)m, \qquad \text{or} \qquad \left(D_1^{*2} + D_2^{*2} + D_3^{*2}\right)m + l(l+1)m = 0. \tag{5.10}$$

From Appendix C, we find that

$$J = \int \frac{F(x, y)\,dx\,dy}{\sqrt{(x^2 + y^2 + f^2)^3}} \tag{5.11}$$

is an invariant (i.e., it is transformed as a scalar). This implies that

$$\rho(x, y) = \frac{1}{\sqrt{(x^2 + y^2 + f^2)^3}} \tag{5.12}$$

is an *invariant measure* (Appendix D).

We also see from Appendix C that

$$J_1 = \int \frac{xF(x, y)\,dx\,dy}{(x^2 + y^2 + f^2)^2}, \qquad J_2 = \int \frac{yF(x, y)\,dx\,dy}{(x^2 + y^2 + f^2)^2},$$
$$J_3 = \int \frac{fF(x, y)\,dx\,dy}{(x^2 + y^2 + f^2)^2} \tag{5.13}$$

are transformed as a 3D vector. Hence, they are irreducibly reduced to a scalar $\sqrt{(J_1)^2 + (J_2)^2 + (J_3)^2}$ and a point $fJ_1/J_3, fJ_2/J_3$ (or a line $J_1 : J_2 : fJ_3$) on the image plane. This scalar and point (or line) are invariant in the sense that they describe characteristics inherent to the scene.

Also from Appendix C, we find that

$$J_{11} = \int \frac{x^2 F(x, y)\, dx\, dy}{\sqrt{(x^2 + y^2 + f^2)^5}}, \quad J_{12} = \int \frac{xyF(x, y)\, dx\, dy}{\sqrt{(x^2 + y^2 + f^2)^5}}, \quad J_{13} = \int \frac{fxF(x, y)\, dx\, dy}{\sqrt{(x^2 + y^2 + f^2)^5}},$$
$$J_{21} = \int \frac{xyF(x, y)\, dx\, dy}{\sqrt{(x^2 + y^2 + f^2)^5}}, \quad J_{22} = \int \frac{y^2 F(x, y)\, dx\, dy}{\sqrt{(x^2 + y^2 + f^2)^5}}, \quad J_{23} = \int \frac{fyF(x, y)\, dx\, dy}{\sqrt{(x^2 + y^2 + f^2)^5}},$$
$$J_{31} = \int \frac{fxF(x, y)\, dx\, dy}{\sqrt{(x^2 + y^2 + f^2)^5}}, \quad J_{33} = \int \frac{fyF(x, y)\, dx\, dy}{\sqrt{(x^2 + y^2 + f^2)^5}}, \quad J_{33} = \int \frac{f^2 F(x, y)\, dx\, dy}{\sqrt{(x^2 + y^2 + f^2)^5}} \tag{5.14}$$

are transformed as a 3D (symmetric) tensor. Hence, they are irreducibly reduced to three scalars, a point, and a line passing through it on the image plane. They are invariant and describe characteristics inherent to the scene.

6. INVARIANT CHARACTERISTICS OF A SHAPE

As an application of the results in the previous sections, let us consider the characterization of a shape on the image plane. Consider a region S on the image plane. Its characteristic function

$$F(x, y) = \begin{cases} 1 & \text{if } (x, y) \in S \\ 0 & \text{otherwise} \end{cases} \tag{6.1}$$

is taken as the image function $F(x, y)$.

The simplest characteristics of the region S may be its area

$$\bar{S} = \int_S dx\, dy \left(= \int F(x, y)\, dx\, dy \right). \tag{6.2}$$

However, this area is not invariant with respect to camera rotation. Suppose the region S is located far away from the image origin. If we move it so that it comes to the center of the image plane by appropriately controlling the camera orientation, the area of Eq. (6.2) changes. Consequently, Eq. (6.2) is not considered to be a characteristic inherent to the scene itself. In short, Eq. (6.2) is not a scalar.

On the other hand, if Eq. (6.2) is replaced by

$$C = f^3 \int_S \frac{dx\, dy}{\sqrt{(x^2 + y^2 + f^2)^3}}, \tag{6.3}$$

this is a scalar as was shown in the previous section. If S is a small region located around the image origin, i.e., $x \approx 0$, $y \approx 0$ in S, then C is approximately equal to this area. We call C the *invariant area* of region S. It is interpreted as the area the region would have if the region were removed to the center of the image plane by changing the camera orientation. Geometrically speaking, this quantity is nothing but an expression of the solid angle the object makes with respect to the viewer.

Another simple but important characteristic is the center of gravity of the region S

$$\bar{x} = \int_S x\,dx\,dy \bigg/ \int_S dx\,dy, \qquad \bar{y} = \int_S y\,dx\,dy \bigg/ \int_S dx\,dy. \tag{6.4}$$

Again these quantities do not have invariant meanings. Namely, if region S is moved to another region by camera rotation and $(\bar{x}', \bar{y}')$ is its center of gravity, $(\bar{x}, \bar{y})$ is not mapped into $(\bar{x}', \bar{y}')$ by the same camera rotation. In short, $\bar{x}, \bar{y}$ is not a point.

On the other hand, we know from Section 5 that

$$a_1 = \int_S \frac{x\,dx\,dy}{(x^2+y^2+f^2)^2}, \qquad a_2 = \int_S \frac{y\,dx\,dy}{(x^2+y^2+f^2)^2},$$
$$a_3 = f\int_S \frac{dx\,dy}{(x^2+y^2+f^2)^2}, \tag{6.5}$$

are transformed as a 3D vector. Hence, fa_1/a_3, fa_2/a_3 are transformed as a point. If the region S is a small region located around the image origin and $x \approx 0$, $y \approx 0$ in S, then $(fa_1/a_3, fa_2/a_3)$ is approximately the center of gravity of the region. We call, $(fa_1/a_3, fa_2/a_3)$ the *invariant center of gravity* of region S. It is interpreted as the point which would be mapped into the center of gravity if the region were moved to the center of the image plane by changing the camera orientation. Geometrically, this point corresponds to the center of the solid angle the object makes with respect to the viewer.

Another useful characteristics is the moment tensor (M_{ij}), $i, j = 1,2$, defined by

$$M_{11} = \int_S (x-\bar{x})^2\,dx\,dy, \qquad M_{12} = M_{21} = \int_S (x-\bar{x})(y-\bar{y})\,dx\,dy,$$
$$M_{22} = \int_S (y-\bar{y})^2\,dx\,dy. \tag{6.6}$$

Its principal values indicate the amount of elongation of the region S along the corresponding principal axes. However, as described above, this tensor does not have invariant properties. Namely, the principal values of (M_{ij}) are not scalars, and its principal axes are not lines on the image plane.

On the other hand, we know from the previous section that

$$B_{11} = \int_S \frac{x^2\,dx\,dy}{\sqrt{(x^2+y^2+f^2)^5}}, \qquad B_{12} = \int_S \frac{xy\,dx\,dy}{\sqrt{(x^2+y^2+f^2)^5}}, \qquad B_{13} = f\int_S \frac{x\,dx\,dy}{\sqrt{(x^2+y^2+f^2)^5}},$$
$$B_{21} = \int_S \frac{xy\,dx\,dy}{\sqrt{(x^2+y^2+f^2)^5}}, \qquad B_{22} = \int_S \frac{y^2\,dx\,dy}{\sqrt{(x^2+y^2+f^2)^5}}, \qquad B_{23} = f\int_S \frac{y\,dx\,dy}{\sqrt{(x^2+y^2+f^2)^5}},$$
$$B_{31} = f\int_S \frac{x\,dx\,dy}{\sqrt{(x^2+y^2+f^2)^5}}, \qquad B_{32} = f\int_S \frac{y\,dx\,dy}{\sqrt{(x^2+y^2+f^2)^5}}, \qquad B_{33} = f^2\int_S \frac{dx\,dy}{\sqrt{(x^2+y^2+f^2)^5}}, \tag{6.7}$$

are transformed as a 3D (symmetrical) tensor. Since this tensor is positive definite as long as region S is not empty, it has three positive principal values $\sigma_1, \sigma_2, \sigma_3$. Let σ_3 be the maximum principal value. Let e_1, e_2, e_3 be the corresponding unit eigenvectors (determined except for sign). Let (g_1, g_2) be the point corresponding to vector e_3. Let l_1 be the line passing through (g_1, g_2) and the point corresponding to vector e_1 (or the line representing vector e_2). Similarly, let l_2 be the line passing through (g_1, g_2) and the point corresponding to vector e_2 (or the line representing vector e_1). By our method of construction, scalars σ_1, σ_2, point (g_1, g_2) and lines l_1, l_2 are all invariant quantities. It can be checked that lines l_1, l_2 are approximately the principal axes, and σ_1, σ_2 are approximately the corresponding principal values if S is a sufficiently small region around the origin. Hence, scalars σ_1 and σ_2 are the principal values the region would have if it were moved to the center of the image plane by camera rotation, and l_1, l_2 are lines which would be mapped onto the principal axes. We call point (g_1, g_2) the *invariant center of inertia*, lines l_1, l_2 the *invariant principal axes*, and σ_1, σ_2 the corresponding *invariant principal values*.

7. INVARIANTS AND CAMERA ROTATION RECONSTRUCTION

In the previous section, scalar C defined by Eq. (6.3), 3D vector $a = (a_i)$ defined by Eqs. (6.5) and 3D tensor $B = (B_{ij})$ defined by Eqs. (6.7) are interpreted as a set of two dimensional invariant quantities on the image plane. Here, let us consider their three dimensional aspects.

First, since C, a, and B are transformed as a scalar, a vector, and a tensor, respectively, by camera rotation, we can extract *invariants* that do not change their values when the camera is rotated. Obviously, scalar C itself is an invariant.

Second, since a is a 3D vector, it has, as was discussed in Section 3, only one invariant, namely its *length* $\|a\|$, or equivalently $a^{\mathrm{T}}a$.

On the other hand, B is a 3D symmetric tensor, and hence it has, as was described in Section 3, three invariants, namely the three principal values $\sigma_1, \sigma_2, \sigma_3$, or equivalently any three independent algebraic expressions formed from them such as the *fundamental symmetric forms* $\sigma_1 + \sigma_2 + \sigma_3$, $\sigma_1\sigma_2 + \sigma_2\sigma_3 + \sigma_3\sigma_2$, $\sigma_1\sigma_2\sigma_3$. In terms of the components of the original tensor B, they are, respectively,

$$B_{11} + B_{22} + B_{33}(= \mathrm{Tr}(B)), \qquad \begin{vmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{vmatrix} + \begin{vmatrix} B_{22} & B_{23} \\ B_{32} & B_{33} \end{vmatrix} + \begin{vmatrix} B_{33} & B_{31} \\ B_{13} & B_{11} \end{vmatrix},$$

$$\begin{vmatrix} B_{11} & B_{12} & B_{13} \\ B_{21} & B_{22} & B_{23} \\ B_{31} & B_{32} & B_{33} \end{vmatrix} (= \det B). \tag{7.1}$$

Alternatively, we can use $\sigma_1 + \sigma_2 + \sigma_3$, $\sigma_1^2 + \sigma_2^2 + \sigma_3^2$, $\sigma_1^3 + \sigma_2^3 + \sigma_3^3$. This set is equal to

$$\mathrm{Tr}(B), \qquad \mathrm{Tr}(B^2), \qquad \mathrm{Tr}(B^3). \tag{7.2}$$

Finally, there are invariants describing the relationship between 3D vector a and 3D tensor B. As was discussed in Section 3, a 3D vector is geometrically thought of as a directed axis to which its length is attached and a 3D symmetric tensor as three mutually perpendicular (undirected) axes to which their respective principal values are attached. Now that the length and the principal values have been counted, the remaining invariants are those specifying the orientation of the vector relative to the three mutually perpendicular axes. Hence, two invariants exist. We can choose, say, $a^{\mathrm{T}}Ba$ and $a^{\mathrm{T}}B^2a$ (Smith [10], Spencer [11], Wang [12–14].) Of course, the choice is not unique as stated above, and other choices are also possible.

We say that two regions S and S' on the image plane are *equivalent* if one region can be transformed into the other by a camera rotation, i.e., by changing the camera orientation. If the two regions are equivalent, the above invariants must have identical values. If they have different values, the two regions cannot be equivalent. On the other hand, if the two regions are known to be equivalent, the camera rotation which would take one region into the other can be reconstructed by observing the invariant center of gravity and the invariant moment tensor alone. This is done as follows.

Suppose we observe a and B for region S and a' and B' for region S'. Assume that B (hence B' as well) has three distinct eigenvalues and $a \neq 0$. Let e_1, e_2, and e_3 be the associated eigenvectors of B. Since the eigenvectors are determined except for sign and magnitude, choose one set such that e_1, e_2, e_3 are mutually perpendicular unit vectors forming a right-hand system in that order. Construct a matrix R_1 having e_1, e_2, e_3 as its columns in that order. Let e_1', e_2', e_3' be the corresponding unit eigenvectors of B' forming a right-hand system. Since the signs of the eigenvectors are arbitrary, there are four possibilities to make a right-hand system. For each case, construct the corresponding matrix R_2. Then, the rotation matrix which transforms B to B' is given by

$$R = R_1 R_2^{\mathrm{T}}. \tag{7.3}$$

Finally, choose one out of those eight possible R's that transforms a to a'.

If B (hence B' as well) has only two distinct eigenvalues (a single root and a pair of multiple roots), let e_1 be the eigenvector associated with the single root. Suppose a is neither parallel nor perpendicular to e_1. Since the sign of e_1 is arbitrary, choose it so that a and e_1 make an acute angle. Then, we can construct three mutually orthogonal vectors forming a right-hand system e_1, $e_2 = e_1 \times a/\|e_1 \times a\|$, $e_3 = e_1 \times e_2$. We can form R_1 and R_2 as described above, and the desired rotation is given by Eq. (7.3). If a is perpendicular to e_1, there exist two solutions. If a is parallel to e_1, or if B (hence B' as well) has one eigenvalue (i.e., B $(= B')$ is a multiple of I), R is any rotation that maps a to a' and we can add any rotation around a'. The case where $a = a' = 0$ is treated similarly. These observations can be summarized as

THEOREM 5.

$$C,\ a^{\mathrm{T}}a,\ \mathrm{Tr}(B),\ \mathrm{Tr}(B^2),\ \mathrm{Tr}(B^3),\ a^{\mathrm{T}}Ba,\ a^{\mathrm{T}}B^2a \tag{7.4}$$

exhaust all the invariants constructed from C, a, and B. If two regions are equivalent, the amount of camera rotation which take one region into the other can be reconstructed from a and B alone.

An important fact is that both the equivalence test and the camera rotation reconstruction *do not require knowledge of point-to-point correspondence*, since the computation is solely based on features (6.3), (6.5), (6.6), which are obtained by integration over the regions under consideration.

Theoretically, the camera rotation is exactly reconstructed as described above. In practice, however, the invariant center of gravity $(fa_1/a_3, fa_2/a_3)$ and the invariant center of inertia (g_1, g_2) are usually located very near, and vector a and vector e_3 are very close to each other. Therefore, the last step of choosing one out of four possible R's by checking Ra may become difficult if much noise is involved. In this case, the final choice is done by applying the transformation (2.2) to region S in four ways and choosing the one which make region S sufficiently overlapping S'. (Since we are focusing on the principal axes, the four possibilities correspond to the four possible (skewed) "mirror image" (including identity) with respect to the principal axes.)

EXAMPLE. Consider the three regions S_0, S_1, S_2 on the image plane (Fig. 2a). We use a scaling such that the focal length f is unity. Computing the integrations of Eqs. (6.5) and (6.7), we find their invariant centers of gravity (Fig. 2b) and principal axes (Fig. 2c) as

S_0	S_1	S_2
$(-0.081, -0.202)$	$(0.464, 0.076)$	$(-0.470, 0.346)$
$y = -2.814x - 0.431$	$y = 1.667x - 0.697$	$y = -0.079x + 0.310$
$y = 0.382x - 0.171$	$y = -0.476x + 0.297$	$y = -16.522x - 7.424$

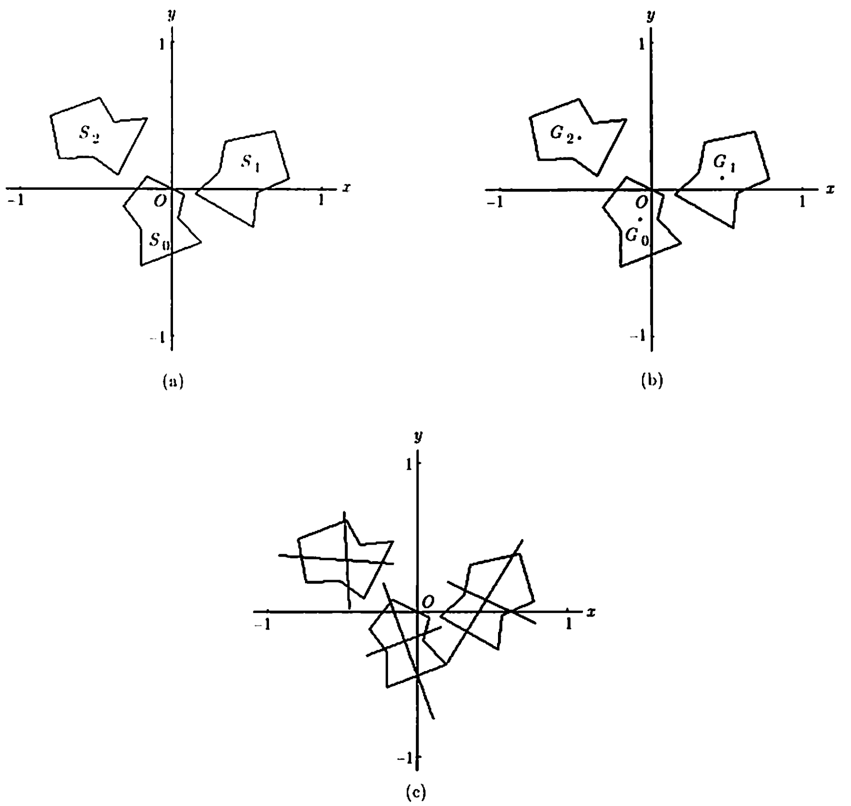

FIG. 2. (a) Three regions S_0, S_1, S_2 to be tested for equivalence. (b) Computed invariant centers of gravity G_0, G_1, G_2 of regions S_0, S_1, S_2. (c) Computed invariant principal axes of regions S_0, S_1, S_2.

The invariants of (7.4) become

	S_0	S_1	S_2
C	0.1440	0.1440	0.1121
$a^{T}a$	0.0202	0.0202	0.0123
$\mathrm{Tr}(B)$	0.1440	0.1440	0.1121
$\mathrm{Tr}(B^2)$	0.0197	0.0197	0.0121
$\mathrm{Tr}(B^3)$	0.0028	0.0028	0.0013
$a^{T}Ba$	0.0028	0.0028	0.0014
$a^{T}B^2a$	0.0004	0.0004	0.0001

From this result, we can conclude that regions S_0 and S_1 can be equivalent but region S_2 is not equivalent to either. (Here, the data are exact up to rounding. If the data are affected by a large amount of error, a statistical method such as hypothesis testing becomes necessary.) By the procedure described in the previous section, the camera rotation which could map region S_0 onto region S_1 is constructed to be

$$R = \begin{bmatrix} 0.573 & -0.761 & -0.296 \\ 0.567 & 0.631 & -0.530 \\ 0.591 & 0.136 & 0.795 \end{bmatrix}.$$

This is the rotation around the axis of orientation $(0.384, -0.512, 0.768)$ by angle $60°$ screwwise.

8. CONCLUDING REMARKS

In this paper, we have presented invariant properties of an image with respect to camera rotation, introducing the notions of "invariance" and "irreducibility" and translating results from projective geometry in terms of the (inhomogeneous) image coordinate system. We also gave an example, computing the invariant center of gravity and the invariant principal axes and reconstructing the camera rotation. The procedure does not required the knowledge of point-to-point correspondence on the image plane. Many other applications are also possible.

Consider the problem of shape recognition. Suppose we have a reference image obtained from a certain camera orientation. If a test image is obtained from a different camera orientation, the two images cannot be compared directly due to projective distortion. However, Theorem 5 provides an easy test for their equivalence. Namely, as is also shown in the previous example, if the invariants of (7.4) have different values, the two regions cannot be equivalent and the test shape is rejected.

If C, a, and B alone are sufficient to characterize the set of test shapes in question completely, the equivalence is already determined at this stage. Otherwise, we can move the test shape into the position of the reference shape in such a way that both have the same a and B. Then, the rest of the shape characteristics are compared to test for the equivalence. The necessary camera rotation is reconstructed as described in Section 7, and the corresponding image transformation is performed either by actually moving the camera or by numerically computing the image transformation (2.2).

We say that a region on the image plane is in the *standard position*, if the invariant center of inertia center (g_1, g_2) coincides with the origin of the image plane and the invariant principal axes coincide with the x and y axes. Any region on the image plane can be moved into the standard position by camera rotation R such that

(i) B is diagonalized in the form

$$R^{T}BR = \begin{bmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \\ 0 & 0 & \sigma_3 \end{bmatrix},$$

where σ_3 is the largest principal value and

(ii) if

$$R^{T}a = \begin{bmatrix} a_1' \\ a_2' \\ a_3' \end{bmatrix}$$

then $a_3' > 0$.

Evidently, shape recognition becomes easier if the test shapes are always moved into the standard position (either by actually rotating the camera or by computation). However, this technique is not restricted to shape recognition. If a camera is tracking a moving object while the camera position is fixed, or if a camera attached to a robot or an autonomous vehicle is aiming at a fixed object in the stationary scene, the technique described above can be used so that the object in question is always seen in the standard position.

On the other hand, testing the equivalence is also viewed as detecting *active motion*. When an object image moves on the image plane, we call the motion *passive* if that motion is induced by camera rotation alone and *active* otherwise. When the camera orientation is changed, object images move on the image plane, but those objects may also have moved in the scene independently of the camera. According to the procedure described above, we can detect active motion even if the angle and orientation of camera rotation is not known. If the corresponding two object images are not equivalent, the object must have moved actively. If they are equivalent, the object has not moved in the scene, although motion is observed on the image plane. In the previous example, if three regions S_0, S_1, S_2 are images of the same object, we can conclude that an active motion took place between S_0 (or S_1) and S_2 while no such motion took place between S_0 and S_1.

Another possible application is camera orientation registration. Even if the camera is rotated by an unknown angle around an unknown axis, the camera orientation can be determined as long as one particular region corresponding to a stationary object is identified on the image plane before and after the camera rotation. Thus, the principle we have described has a wide range of applications to many problems.

APPENDIX A: DUALITY AND CONJUGACY

Consider a line l on the image plane which does not pass through the origin. Let

$$x\cos\theta + y\sin\theta = d \tag{A.1}$$

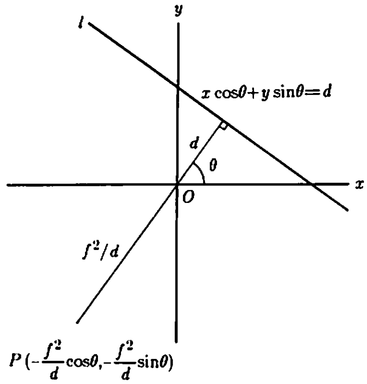

FIG. A1. Line l: $\cos\theta + y\sin\theta = d$ and point $P(-(f^2/d)\cos\theta, -(f^2/d)\sin\theta)$ are mutually dual with respect to the origin O.

($d > 0$) be its equation. We say that point

$$P\left(-\frac{f^2}{d}\cos\theta, -\frac{f^2}{d}\sin\theta\right) \tag{A.2}$$

is *dual* to line l with respect to the origin. Conversely, line l is said to be dual to point P with respect to the origin. In other words, if we draw a line passing through the origin and perpendicular to line l, and if d is the distance between the origin and line l, the dual point P is located on the other side of the perpendicular line and at distance f^2/d from the origin (Fig. A1). If $d = 0$, point P is interpreted as located at infinity (at $(\cos\theta, \sin\theta, 0)$ in homogeneous coordinates), and similarly the line at infinity is regarded as the dual line of the origin O.

Consider a line l and a point P on it on the image plane. Let H be the foot of the perpendicular line drawn from the origin to line l, and let d be the distance between point P and point H. Consider a point Q on the other side of line l at distance f^2/d from point H (Fig. A2). We say that point A is *conjugate* to point P on line l and conversely point P is conjugate to point Q on line l. If $d = 0$, Q is regarded as located at infinity.

As stated in Theorem 3, a 3D vector is represented as a point or as a line on the image plane. By definition, the point and the line are easily shown to be mutually dual. Hence, if one is known, the other is obtained immediately.

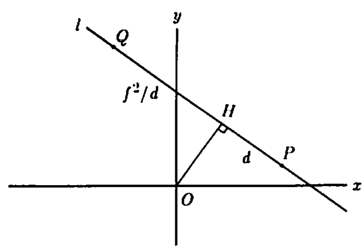

FIG. A2. Points P, Q on line l are mutually conjugate with respect to the foot H of the perpendicular line drawn from the origin O to line l.

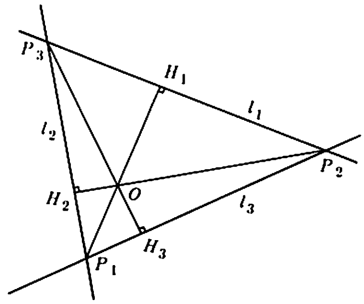

FIG. A3. Point P_1 and line l_1, point P_2 and line l_2, point P_3 and line l_3 are mutually dual, and points P_2 and P_3 on line l_1, points P_3 and P_1 on line l_2, points P_1 and P_2 on line l_3 are mutually conjugate.

As stated in Theorem 4, a 3D symmetric tensor is represented by three scalars, a point, and a line passing through it. Let e_1, e_2, e_3 be the unit vectors of the principal axes (determined up to sign). Let P_1, P_2, P_3 be the points corresponding to them, and let l_1 be the line connecting points P_2 and P_3, l_2 the line connecting points P_3 and P_1, and l_3 the line connecting points P_1 and P_2 (Fig. A3). Then, it is easy to see that point P_1 and line l_1 are mutually dual, and so are point P_2 and line l_2, and point P_3 and line l_3. It is also seen that points P_2 and P_3, points P_3, P_1, and points P_1 and P_2 are conjugate on lines l_1, l_2, and l_3, respectively. Hence, if point P_3 and line l_1 are given, line l_3 and point P_1 are obtained as their duals. Point P_2 is given as the intersection between lines l_1 and l_3, and line l_2 is given as the line connecting points P_3 and P_1. Thus, a point and a line passing through it are sufficient to represent on the image plane the orientations of the three principal axes of a 3D symmetric tensor.

APPENDIX B: FUNCTION BASIS OF IRREDUCIBLE REPRESENTATIONS

From Eqs. (4.10), the Casimir operator becomes

$$H = -\left(f + \frac{x^2+y^2}{f}\right)\left[\left(f + \frac{x^2}{f}\right)\frac{\partial^2}{\partial x^2} + \frac{2xy}{f}\frac{\partial^2}{\partial x\,\partial y} + \left(f + \frac{y^2}{f}\right)\frac{\partial^2}{\partial y^2}\right] - \left(f + x + \frac{2x(x^2+y^2)}{f^2}\right)\frac{\partial}{\partial x} - \left(f + y + \frac{2y(x^2+y^2)}{f^2}\right)\frac{\partial}{\partial y} \tag{B.1}$$

so that Eq. (4.14) becomes

$$\left(f + \frac{x^2+y^2}{f}\right)\left[\left(f + \frac{x^2}{f}\right)F_{xx} + \frac{2xy}{f}F_{xy} + \left(f + \frac{y^2}{f}\right)F_{yy}\right] + \left(f + x + \frac{2x(x^2+y^2)}{f}\right)F_x + \left(f + x + \frac{2y(x^2+y^2)}{f}\right)F_y + l(l+1)F = 0. \tag{B.2}$$

Since representations of half-integer weights are not interesting because the same image must be obtained after a rotation of 2π (the sign is reversed after a rotation of 2π if the weight is a half-integer), we consider only irreducible representations of integer weights.

For $l = 0$ ($l(l+1) = 0$), one solution ($2l + 1 = 1$) is easily found:

$$F_0^1(x, y) = 1. \tag{B.3}$$

Obviously, this is invariant with respect to rotation

$$D_1 F_0^1 = 0, \qquad D_2 F_0^1 = 0, \qquad D_3 F_0^1 = 0, \tag{B.4}$$

and hence

$$T_R F_0^1 = F_0^1. \tag{B.5}$$

For $l = 1$ ($l(l+1) = 2$), the following three solutions ($2l + 1 = 3$) are found:

$$F_1^1(x, y) = \frac{x}{\sqrt{x^2 + y^2 + f^2}}, \qquad F_1^2(x, y) = \frac{y}{\sqrt{x^2 + y^2 + f^2}},$$
$$F_1^3(x, y) = \frac{f}{\sqrt{x^2 + y^2 + f^2}}. \tag{B.6}$$

Application of the infinitesimal generators D_1, D_2, D_3 yields

$$D_1 \begin{bmatrix} F_1^1 \\ F_1^2 \\ F_1^3 \end{bmatrix} = -A_1 \begin{bmatrix} F_1^1 \\ F_1^2 \\ F_1^3 \end{bmatrix}, \qquad D_2 \begin{bmatrix} F_1^1 \\ F_1^2 \\ F_1^3 \end{bmatrix} = -A_2 \begin{bmatrix} F_1^1 \\ F_1^2 \\ F_1^3 \end{bmatrix},$$
$$D_3 \begin{bmatrix} F_1^1 \\ F_1^2 \\ F_1^3 \end{bmatrix} = -A_3 \begin{bmatrix} F_1^1 \\ F_1^2 \\ F_1^3 \end{bmatrix}, \tag{B.7}$$

where

$$A_1 = \begin{bmatrix} & & \\ & & -1 \\ & 1 & \end{bmatrix}, \quad A_2 = \begin{bmatrix} & & 1 \\ & & \\ -1 & & \end{bmatrix}, \quad A_3 = \begin{bmatrix} & -1 & \\ 1 & & \\ & & \end{bmatrix}, \tag{B.8}$$

and the commutation relations are satisfied:

$$[A_1, A_2] = A_3, \qquad [A_2, A_3] = A_1, \qquad [A_3, A_1] = A_2. \tag{B.9}$$

Consequently, for infinitesimal rotations, we have

$$T_R \begin{bmatrix} F_1^1 \\ F_1^2 \\ F_1^3 \end{bmatrix} = I + (\Omega_1 A_1 + \Omega_2 A_2 + \Omega_3 A_3) \begin{bmatrix} F_1^1 \\ F_1^2 \\ F_1^3 \end{bmatrix} + o(\Omega). \tag{B.10}$$

This implies that F_1^1, F_1^2, F_1^3 are transformed as

$$T_R \begin{bmatrix} F_1^1 \\ F_1^2 \\ F_1^3 \end{bmatrix} = R \begin{bmatrix} F_1^1 \\ F_1^2 \\ F_1^3 \end{bmatrix}. \tag{B.11}$$

For $l = 2$ ($l(l+1) = 6$), the following five solutions ($2l + 1 = 5$) are found:

$$F_2^1(x, y) = \frac{2x^2 - y^2 - f^2}{3(x^2 + y^2 + f^2)}, \qquad F_2^2(x, y) = \frac{-x^2 + 2y^2 - f^2}{3(x^2 + y^2 + f^2)},$$
$$F_2^3(x, y) = \frac{xy}{x^2 + y^2 + f^2}, \qquad F_2^4(x, y) = \frac{fx}{x^2 + y^2 + f^2},$$
$$F_2^5(x, y) = \frac{fy}{x^2 + y^2 + f^2}. \tag{B.12}$$

Application of the infinitesimal generators D_1, D_2, D_3 yields

$$D_1 \begin{bmatrix} F_2^1 \\ F_2^2 \\ F_2^3 \\ F_2^4 \\ F_2^5 \end{bmatrix} = -A_1 \begin{bmatrix} F_2^1 \\ F_2^2 \\ F_2^3 \\ F_2^4 \\ F_2^5 \end{bmatrix}, \quad D_2 \begin{bmatrix} F_2^1 \\ F_2^2 \\ F_2^3 \\ F_2^4 \\ F_2^5 \end{bmatrix} = -A_2 \begin{bmatrix} F_2^1 \\ F_2^2 \\ F_2^3 \\ F_2^4 \\ F_2^5 \end{bmatrix}, \quad D_3 \begin{bmatrix} F_2^1 \\ F_2^2 \\ F_2^3 \\ F_2^4 \\ F_2^5 \end{bmatrix} = -A_3 \begin{bmatrix} F_2^1 \\ F_2^2 \\ F_2^3 \\ F_2^4 \\ F_2^5 \end{bmatrix}, \tag{B.13}$$

where

$$A_1 = \begin{bmatrix} & & & & \\ & & & & -2 \\ & & & -1 & \\ & & 1 & & \\ 1 & 2 & & & \end{bmatrix}, \qquad A_2 = \begin{bmatrix} & & & 2 & \\ & & & & \\ & & & & 1 \\ -2 & -1 & & & \\ & & -1 & & \end{bmatrix},$$
$$A_3 = \begin{bmatrix} & & -2 & & \\ & & 2 & & \\ 1 & -1 & & & \\ & & & & -1 \\ & & & 1 & \end{bmatrix}, \tag{B.14}$$

and the commutation relations are satisfied,

$$[A_1, A_2] = A_3, \qquad [A_2, A_3] = A_1, \qquad [A_3, A_1] = A_2. \tag{B.15}$$

Consequently, for infinitesimal rotations, we have

$$T_R \begin{bmatrix} F_2^1 \\ F_2^2 \\ F_2^3 \\ F_2^4 \\ F_2^5 \end{bmatrix} = I + (\Omega_1 A_1 + \Omega_2 A_2 + \Omega_3 A_3) \begin{bmatrix} F_2^1 \\ F_2^2 \\ F_2^3 \\ F_2^4 \\ F_2^5 \end{bmatrix} + o(\Omega). \tag{B.16}$$

This implies that if we put

$$\begin{aligned} F_{11} &= F_2^1, & F_{22} &= F_2^2, & F_{33} &= -F_2^1 - F_2^2, \\ F_{12} = F_{21} &= F_2^3, & F_{31} = F_{13} &= F_2^4, & F_{32} = F_{23} &= F_2^5, \end{aligned} \tag{B.17}$$

functions $F_2^1, F_2^2, F_2^3, F_2^4, F_2^5$ are transformed as

$$T_R \begin{bmatrix} F_{11} & F_{12} & F_{13} \\ F_{21} & F_{22} & F_{23} \\ F_{31} & F_{32} & F_{33} \end{bmatrix} = R \begin{bmatrix} F_{11} & F_{12} & F_{13} \\ F_{21} & F_{22} & F_{23} \\ F_{31} & F_{32} & F_{33} \end{bmatrix} = R^{\mathrm{T}}. \tag{B.18}$$

Solutions for $l = 3, 4, \ldots,$ are constructed similarly. In fact, function $F_l^1, F_l^2, \ldots, F_l^{2l+1}$ are just the lth spherical harmonics projected onto the image xy plane.

APPENDIX C: FEATURE BASIS OF IRREDUCIBLE REPRESENTATIONS

From Eqs. (5.6), the Casimir operator becomes

$$\begin{aligned} H^* = &-\left(f + \frac{x^2 + y^2}{f}\right) \\ &\times\left[\left(f + \frac{x^2}{f}\right)\frac{\partial^2}{\partial x^2} + \frac{2xy}{f}\frac{\partial^2}{\partial x\,\partial y} + \left(f + \frac{y^2}{f}\right)\frac{\partial^2}{\partial y^2} + 8x\frac{\partial}{\partial x} + 8y\frac{\partial}{\partial y}\right] \\ &-6 - \frac{12(x^2 + y^2)}{f^2}, \end{aligned} \tag{C.1}$$

so that Eq. (5.10) becomes

$$\begin{aligned} &\left(f + \frac{x^2 + y^2}{f}\right)\left[\left(f + \frac{x^2}{f}\right)m_{xx} + \frac{2xy}{f}m_{xy} + \left(f + \frac{y^2}{f}\right)m_{yy} + 8xm_x + 8ym_y\right] \\ &\quad + \left[l(l+1) + 6 + \frac{12(x^2 + y^2)}{f^2}\right]m = 0. \end{aligned} \tag{C.2}$$

For $l = 0$ ($l(l+1) = 0$), the solution ($2l + 1 = 1$) is found

$$m_0^1(x, y) = \frac{1}{\sqrt{(x^2 + y^2 + f^2)^3}}. \tag{C.3}$$

Application of the infinitesimal generators D_1^*, D_2^*, D_3^* yields

$$D_1^* m_0^1 = 0, \qquad D_2^* m_0^1 = 0, \qquad D_3^* m_0^1 = 0, \tag{C.4}$$

and consequently m_0^1 is invariant for T_R^*,

$$T_R^* m_0^1 = m_0^1. \tag{C.5}$$

Feature J of Eq. (5.11) is obtained by

$$J = \int m_0^1(x, y) F(x, y)\, dx\, dy. \tag{C.6}$$

From Eq. (C.5), this is a scalar.

For $l = 1$ ($l(l+1) = 2$), the three solutions ($2l + 1 = 3$) are found:

$$m_1^1(x, y) = \frac{x}{(x^2 + y^2 + f^2)^2}, \qquad m_1^2(x, y) = \frac{y}{(x^2 + y^2 + f^2)^2},$$
$$m_1^3(x, y) = \frac{f}{(x^2 + y^2 + f^2)^2}. \tag{C.7}$$

Application of the infinitesimal generators D_1^*, D_2^*, D_3^* yields

$$D_1^* \begin{bmatrix} m_1^1 \\ m_1^2 \\ m_1^3 \end{bmatrix} = -A_1^* \begin{bmatrix} m_1^1 \\ m_1^2 \\ m_1^3 \end{bmatrix}, \qquad D_2^* \begin{bmatrix} m_1^1 \\ m_1^2 \\ m_1^3 \end{bmatrix} = -A_2^* \begin{bmatrix} m_1^1 \\ m_1^2 \\ m_1^3 \end{bmatrix},$$
$$D_3^* \begin{bmatrix} m_1^1 \\ m_1^2 \\ m_1^3 \end{bmatrix} = -A_3^* \begin{bmatrix} m_1^1 \\ m_1^2 \\ m_1^3 \end{bmatrix}, \tag{C.8}$$

where

$$A_1^* = \begin{bmatrix} & & \\ & & -1 \\ & 1 & \end{bmatrix}, \qquad A_2^* = \begin{bmatrix} & & 1 \\ & & \\ -1 & & \end{bmatrix}, \qquad A_3^* = \begin{bmatrix} & -1 & \\ 1 & & \\ & & \end{bmatrix}, \tag{C.9}$$

and the commutation relations are satisfied

$$[A_1^*, A_2^*] = A_3^*, \qquad [A_2^*, A_3^*] = A_1^*, \qquad [A_3^*, A_1^*] = A_2^*. \tag{C.10}$$

Consequently, for infinitesimal rotations, we have

$$T_R^* = \begin{bmatrix} m_1^1 \\ m_1^2 \\ m_1^3 \end{bmatrix} = I - (\Omega_1 A_1^* + \Omega_2 A_2^* + \Omega_3 A_3^*) \begin{bmatrix} m_1^1 \\ m_1^2 \\ m_1^3 \end{bmatrix} + o(\Omega). \tag{C.11}$$

This implies that m_1^2, m_1^2, m_1^3 are transformed as

$$T_R^* \begin{bmatrix} m_1^1 \\ m_1^2 \\ m_1^3 \end{bmatrix} = R^{\mathrm{T}} \begin{bmatrix} m_1^1 \\ m_1^2 \\ m_1^3 \end{bmatrix}. \tag{C.12}$$

Features J_i, $i = 1, 2, 3$, of Eqs. (5.13) are obtained by

$$J_i = \int m_1^i(x, y) F(x, y)\, dx\, dy \tag{C.13}$$

From Eq. (C.12), they are transformed as a 3D vector.

For $l = 2$ ($l(l+1) = 6$), we can find the five solutions ($2l + 1 = 5$),

$$\begin{aligned} m_2^1(x, y) &= \frac{2x^2 - y^2 - f^2}{3\sqrt{(x^2+y^2+f^2)^5}}, & m_2^2(x, y) &= \frac{-x^2 + 2y^2 - f^2}{3\sqrt{(x^2+y^2+f^2)^5}}, \\ m_2^3(x, y) &= \frac{xy}{\sqrt{(x^2+y^2+f^2)^5}}, & m_2^4(x, y) &= \frac{fx}{\sqrt{(x^2+y^2+f^2)^5}}, \\ m_2^5(x, y) &= \frac{fy}{\sqrt{(x^2+y^2+f^2)^5}}. \end{aligned} \tag{C.14}$$

Application of the infinitesimal generators D_1, D_2, D_3 yields

$$D_1^* \begin{bmatrix} m_2^1 \\ m_2^2 \\ m_2^3 \\ m_2^4 \\ m_2^5 \end{bmatrix} = -A_1^* \begin{bmatrix} m_2^1 \\ m_2^2 \\ m_2^3 \\ m_2^4 \\ m_2^5 \end{bmatrix}, \quad D_2^* \begin{bmatrix} m_2^1 \\ m_2^2 \\ m_2^3 \\ m_2^4 \\ m_2^5 \end{bmatrix} = -A_2^* \begin{bmatrix} m_2^1 \\ m_2^2 \\ m_2^3 \\ m_2^4 \\ m_2^5 \end{bmatrix},$$

$$D_3^* \begin{bmatrix} m_2^1 \\ m_2^2 \\ m_2^3 \\ m_2^4 \\ m_2^5 \end{bmatrix} = -A_3^* \begin{bmatrix} m_2^1 \\ m_2^2 \\ m_2^3 \\ m_2^4 \\ m_2^5 \end{bmatrix}, \tag{C.15}$$

where

$$A_1^* = \begin{bmatrix} & & & & \\ & & & & -2 \\ & & & -1 & \\ & & 1 & & \\ 1 & 2 & & & \end{bmatrix}, \quad A_2^* = \begin{bmatrix} & & & 2 & \\ & & & & 1 \\ -2 & -1 & & & \\ & & -1 & & \\ & & & & \end{bmatrix},$$

$$A_3^* = \begin{bmatrix} & & -2 & & \\ & & 2 & & \\ 1 & -1 & & & \\ & & & & -1 \\ & & & 1 & \end{bmatrix}, \tag{C.16}$$

and the commutation relations are satisfied

$$[A_1^*, A_2^*] = A_3^*, \qquad [A_2^*, A_3^*] = A_1^*, \qquad [A_3^*, A_1^*] = A_2^*. \tag{C.17}$$

Consequently, for infinitesimal rotations, we have

$$T_R^* \begin{bmatrix} m_2^1 \\ m_2^2 \\ m_2^3 \\ m_2^4 \\ m_2^5 \end{bmatrix} = I - (\Omega_1 A_1^* + \Omega_2 A_2^* + \Omega_3 A_3^*) \begin{bmatrix} m_2^1 \\ m_2^2 \\ m_2^3 \\ m_2^4 \\ m_2^5 \end{bmatrix} + o(\Omega). \tag{C.18}$$

This implies that if we put

$$\begin{aligned} m_{11} &= m_2^1, & m_{22} &= m_2^2, & m_{33} &= -m_2^1 - m_2^2, \\ m_{12} &= m_{21} = m_2^3, & m_{31} &= m_{13} = m_2^4, & m_{32} &= m_{23} = m_2^5, \end{aligned} \tag{C.19}$$

functions $m_2^1, m_2^2, m_2^3, m_2^4, m_2^5$ are transformed as

$$T_R^* \begin{bmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{bmatrix} = R^{\mathrm{T}} \begin{bmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{bmatrix} R. \tag{C.20}$$

Features J_{ij}, $i, j = 1, 2, 3$, of Eqs. (6.7) are obtained by

$$J_{ij} = \int \left(m_{ij}(x, y) + \tfrac{1}{3} m_0^1(x, y) \delta_{ij} \right) F(x, y)\, dx\, dy, \tag{C.21}$$

where δ_{ij} is the Kronecker delta. From Eqs. (C.5) and (C.20), they are transformed as a 3D tensor.

Solutions for $l = 3, 4, \ldots$ are constructed similarly. In fact, we can check, by substitution, that the solution is given by

$$m_l^i(x, y) = \frac{F_l^i(x, y)}{\sqrt{(x^2 + y^2 + f^2)^3}}, \qquad i = 1, 2, \ldots, 2l + 1. \tag{C.22}$$

APPENDIX D: INVARIANT MEASURE

We say that $\rho(x, y)\, dx\, dy$ is an *invariant measure* if for any image $F(x, y)$

$$\int T_R F(x, y) \rho(x, y)\, dx\, dy = \int F(x, y) \rho(x, y)\, dx\, dy. \tag{D.1}$$

In view of Eqs. (5.1) and (5.3), this is equivalent to

$$T_R^* \rho(x, y) = \rho(x, y). \tag{D.2}$$

Hence, ρ is given by the solution of Eq. (5.10) with $l = 0$. From Eq. (C.3) of Appendix C, we obtain

$$\rho(x, y) = \frac{1}{\sqrt{(x^2 + y^2 + f^2)^3}}. \tag{D.3}$$

This result can be interpreted intuitively in terms of fluid dynamics. Suppose the camera is rotating with rotation velocity $(\omega_1, \omega_2, \omega_3)$, namely rotating around an axis of orientation $(\omega_1, \omega_2, \omega_3)$ with angular velocity $\sqrt{(\omega_1)^2 + (\omega_2)^2 + (\omega_3)^2}$ (rad/sec) screwwise. (Here, $\omega_1, \omega_2, \omega_3$ are also interpreted as instantaneous angular velocities around the x, the y, the z axis, respectively.) The *optical flow* induced on the image plane is obtained by dividing both sides of Eqs. (4.7) by δt

$$u = -f\omega_2 + \omega_3 y + \frac{1}{f}(-\omega_2 x + \omega_1 y)x, \qquad v = f\omega_1 - \omega_3 x + \frac{1}{f}(-\omega_2 x + \omega_1 y)y. \tag{D.4}$$

If this flow is regarded as a fluid flow with density $\rho(x, y)$, the necessary and sufficient condition that the fluid is neither created or annihilated in the course of flowing is, as is well known, given by the *equation of continuity*

$$\frac{\partial(\rho u)}{\partial x} + \frac{\partial(\rho v)}{\partial y} = 0. \tag{D.5}$$

If Eqs. (D.4) are substituted, Eq. (D.5) becomes

$$(\omega_1 D_1^* + \omega_2 D_2^* + \omega_3 D_3^*)\rho = 0. \tag{D.6}$$

This equation must be satisfied for arbitrary $\omega_1, \omega_2, \omega_3$. Hence, the invariant measure $\rho(x, y)$ is given as a solution of the differential equations

$$D_1^*\rho = 0, \qquad D_2^*\rho = 0, \qquad D_3^*\rho = 0. \tag{D.7}$$

ACKNOWLEDGMENTS

The author thanks Professor Azriel Rosenfeld and Professor Larry S. Davis of the University of Maryland and Professor Shun-ichi Amari of the University of Tokyo for helpful discussions and comments. He also thanks Mr. T.-C. Chou of the University of Maryland for doing numerical experiments. 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 DAAK70-83-K-0018.

REFERENCES

1. S. Amari, Invariant structures of signal and feature spaces in pattern recognition problems, *RAAG Mem.* **4**, 1968, 553–566.
2. S. Amari, Feature spaces which admit and detect invariant signal transformations, in *Proceedings, 4th Int. Joint Conf. Pattern Recog., Tokyo, 1978*, pp. 452–456.
3. I. M. Gel'fand, R. A. Minlos, and Z. Ya. Shapiro, *Representation of the Rotation and Lorentz Groups and Their Applications*, Pergamon, Oxford, 1963.

4. M. Hammermesh, *Group Theory and Its Application to Physical Problems*, Addison–Wesley, Reading, MA, 1964.
5. K. Kanatani, Transformation of optical flow by camera rotation, in *Proceedings, IEEE Workshop on Motion: Representation and Analysis, Charleston, SC, 1986*, pp. 113–118.
6. K. Kanatani, *Constraints on Length and Angle*, Technical Report No. CAR-TR-200, Center for Automation Research, University of Maryland, 1986.
7. K. Kanatani and T.-C. Chou, Shape from texture: General principle, *Proceedings, IEEE Conf. Comput. Vision Pattern Recog., Miami Beach, FL, 1986*, pp. 578–583.
8. A. Naeve, and J.-O. Eklundh, *On Projective Geometry and the Recovery of 3-D Structure*, Technical Report No. CAR-TR-154, Center for Automation Research, University of Maryland, 1985.
9. A. Rosenfeld and A. Kak, *Digital Picture Processing*, Vol. 2, Academic Press, Orlando, FL, 1982.
10. G. F. Smith, On isotropic functions of symmetric tensors, skew-symmetric tensors and vectors, *Internat. J. Engrg. Sci.* **9**, 1971, 899–916.
11. A. J. M. Spencer, Theory of invariants, in *Continuum Physics*, Vol. 1 (A. C. Eringen, Ed.), pp. 239–353, Academic Press, New York, 1971.
12. C.-C. Wang, A new representation theorem for isotropic functions, Part 1, *Arch. Rational Mech. Anal.* **36**, 1970, 166–197.
13. C.-C. Wang, A new representation theorem for isotropic functions, Part 2, *Arch. Rational Mech. Anal.* **36**, 1970, 198–223.
14. C.-C. Wang, Corrigendum to my papers on representation for isotropic functions, *Arch. Rational Mech. Anal.* **43**, 1971, 392–395.
15. H. Weyl, *The Classical Groups, Their Invariants and Representations*, 2nd ed., Princeton Univ. Press, Princeton, NJ, 1946.