CVGIP: IMAGE UNDERSTANDING
Vol. 54, No. 3, November, pp. 333–348, 1991

Computational Projective Geometry

KENICHI KANATANI

Department of Computer Science, Gunma University, Kiryu, Gunma 376, Japan

Received April 27, 1989; accepted April 12, 1991

A computational formalism is given to computer vision problems involving *collinearity* and *concurrency* of points and lines on a 2-D plane from the viewpoint of projective geometry. The image plane is regarded as a *2-D projective space*, and points and lines are represented by unit vectors consiting of homogeneous coordinates, called *N-vectors*. Fundamental notions of projective geometry such as *collineations*, *correlations*, *polarities*, *poles*, *polars*, and *conis* are reformulated as "computational" processes in terms of N-vectors. They are also given 3-D interpretations by regarding 2-D images as perspective projection of 3-D scenes. This N-vector formalism is further extended to infer 3-D translational motions from 2-D motion images. Stereo is also viewed as a special type of translational motion. Three computer vision applications are briefly discussed—interpretation of a rectangle, interpretation of a road, and interpretation of planar surface motion. © 1991 Academic Press, Inc.

1. INTRODUCTION

One of the basic building blocks of *projective geometry* is *perspective projection:* roughly speaking, a 2-D (or generally n-D) *projective space* is a perspectively projected image of a 3-D (or $(n + 1) - D$) space. Naturally, projective geometry is expected to play a central role in the study of computer vision, which aims to obtain 3-D interpretations by analyzing 2-D camera images of 3-D scenes. In fact, there have been many attempts to analyze perspective projection images by invoking projective geometry [3, 5, 11, 12, 22, 32]. In doing so, however, we immediately encounter many *computational* problems. For example, if we want to compute the intersection of nearly parallel lines in the image, the computation may overflow, or if not, introduce a considerable amount of error in subsequent computation because computation by a computer consists of fixed-length numerical operations.

To mathematicians, projective geometry is a purely abstract mathematics, unifying Euclidean geometry dating back to ancient Greece and non-Euclidean geometries developed by such mathematicians as Gauss, Lobachevskii, Bolyai, and Riemann. Given rigorous axiomatization by Klein, Hilbert and others, it is one of the most elegant constructs of mathematics. As a result, the computational aspect is not central. Indeed, the very motivation of projective geometry is to *pursue logical consistencies by ignoring computational anomalies*. It follows that in order to fully utilize projective geometry for computer vision problems, we must reformulate the entire structure of projective geometry from a *computational* viewpoint.

A means to overcome computational breakdown is already provided by projective geometry itself, namely the use of *homogeneous coordinates*, by which points and lines are represented by three real numbers. Since homogeneous coordinates can be multiplied by any nonzero number, it is computationally most convenient to scale them so that they constitute a unit vector. We call such a vector an *N-vector*. The aim of this paper is to

(i) reformulate basic concepts of projective geometry as *computational processes* expressed in terms of N-vectors, and

(ii) relate them to 3-D interpretations of the scene.

The most fundamental concepts of projective geometry are *collinearity* of points (i.e., points lying on a common line) and *concurrency* of lines (i.e., lines passing through a common point), from which such notions as *collineations*, *correlations*, *polarities*, *poles*, *polars*, *conjugacy*, and *conics* are defined. Two facts are essential abut these notions:

• All these concepts do not involve *metrics* (length, angle, area, etc.). In fact, projective geometry is regarded as the most general geometry because, as Klein pointed out, all known classical geometries—both Euclidean and non-Euclidean—are obtained by introducing particular metrics.

• There exists a *duality* such that for every statement we have a corresponding statement where the roles of points and lines are interchanged. This means that the entire structure of projective geometry is unchanged if the roles of points and lines are interchanged.

In this paper, we redefine such concepts as mentioned above as *computational processes* in terms of N-vectors in such a way that the inherent duality becomes explicit.

1049-9660/91 $3.00

We then show that the use of N-vectors not only resolves the computational problems but also provides straightforward 3-D interpretations to such notions as *vanishing points* of lines, *vanishing lines* of planar surfaces, and *focuses of expansion* of translational motions. We also show that 3-D interpretations involving parallelism and orthogonality in the scene are succinctly expressed in terms of N-vectors.

Translation of the camera or objects occurs naturally in many industrial environments. Alternatively, we can control the camera *actively* for the purpose of understanding 3-D environments, which is known as the *active vision* paradigm [1]. *Stereo* can also be viewed as a special kind of translational motion. We present an N-vector formalism for analyzing translational motion and stereo.

In appendices, we will briefly discuss three applications of our formalism:

1. The constrains on a projection image of a rectangle in a scene are expressed in terms of N-vectors. This result is useful for not only 3-D object recognition but also camera calibration [2, 4, 14, 18, 19], which is essential in implementing computer vision techniques.

2. The constraints on a projection image of an ideal road in a scene are expressed in terms of N-vectors. This result is useful for autonomously navigating land vehicles, which are currently major national projects across the world [6, 17, 23, 27, 30, 31].

3. A mathematical analysis, in terms of N-vectors, is presented for computing the 3-D rigid motion of a planar surface from two projection images. This problem is one of the central theoretical issues of computer vision [3, 7, 10, 21, 25, 28, 29].

2. PERSPECTIVE PROJECTION AND *N*-VECTORS

The image plane is regarded as a 2-D *projective space*. This means that a *point* is designated by a triplet (m_1, m_2, m_3) of real numbers, not all of them being 0. These three numbers are called *homogeneous coordinates*. If $m_3 \neq 0$, point (m_1, m_2, m_3) is identified with the point (fm_1/m_3, fm_2/m_3) on the image plane; $x = fm_1/m_3$ and $y = fm_2/m_3$ are called *inhomogeneous coordinates* (the meaning of the constant f will be explained shortly). If $m_3 = 0$, point (m_1, m_2, 0) is regarded as located at *infinity* and called an *ideal point*. The set of all points at infinity is called the *line at infinity* or the *ideal line;* its "equation" is $m_3 = 0$.

A *line* in the 2-D projective space is also defined by a triplet (m_1, n_1, n_3) of real numbers, not all of them being 0. These three numbers are also called *homogeneous coordinates*. If n_1 or n_2 is not zero, the line appears on the image plane as $n_1x + n_2y + n_3f = 0$. If $n_1 = n_2 = 0$, the line is interpreted to be the ideal line at infinity.

The homogeneous coordinates can be multiplied by an arbitrary nonzero number, and the point or line that they represent is still the same. This means that their magnitudes can be kept within a reasonable range by multiplying them by an appropriate constant whenever they become two large or too small. The most convenient way is to keep the homogeneous coordinates (m_1, m_2, m_3) to be a "unit vector": We call such a vector an *N-vector.*

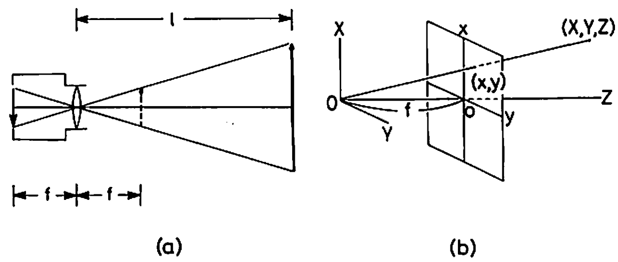

FIG. 1. (a) The camera imaging model, (b) Perspective projection of the scene.

If we use N-vectors, representation of points and lines is unique except for sign; we hereafter ignore the sign of N-vectors unless specified, meaning that if $\mathbf{m}$ (or $\mathbf{n}$) is an N-vector $-\mathbf{m}$ (or $-\mathbf{n}$) is also an N-vector representing the same point (or line). According to the above definition. the N-vector of point (a, b) is[1]

$$\mathbf{m} = \pm N\left[\begin{pmatrix} a \\ b \\ f \end{pmatrix}\right], \tag{2.1}$$

and the N-vector of line $Ax + By + C = 0$ is

$$\mathbf{n} = \pm N\left[\begin{pmatrix} A \\ B \\ C/f \end{pmatrix}\right]. \tag{2.2}$$

The reason the definition of the N-vector involves constant f comes from the following 3-D interpretation of *perspective projection*. Take an XYZ coordinate system fixed to the camera so that the origin O corresponds to the center of the lens, which we call the *viewpoint*, and the Z-axis corresponds to the optical axis of the camera. Then the plane $Z = f$ is identified with the image plane, where f is the distance between the center of the lens and the surface of the film (Fig. 1 (a)). We assume that f is a known constant, and call it the *focal length*, although this may not exactly be the focal length of the lens (it coincides with the focal length of the lens if the camera is focused at infinity).

A point (X, Y, Z) in the scene is projected onto the

[1] In this paper, $N[\mathbf{u}] = \mathbf{u}/\|\mathbf{u}\|$ denotes the normalization of vector $\mathbf{u}$, and $\|\mathbf{u}\|$ denotes its norm.

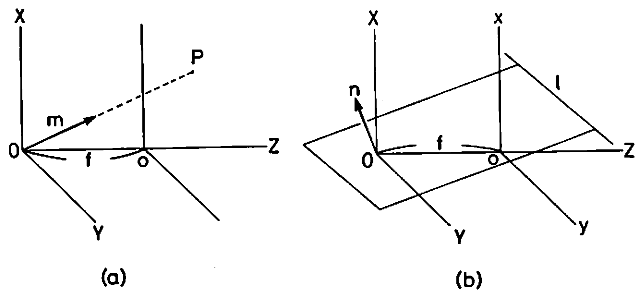

FIG. 2. (a) The N-vector **m** of a point P on the image plane. (b) The N-vector **n** of a line l on the image plane.

intersection of the image plane $Z = f$ with the ray starting from the viewpoint O and passing through that point. Take an image xy coordinate system on the image plane such that the x- and the y-axes are respectively parallel to the X- and the Y-axes, $(0, 0, f)$ being the image coordinate origin (Fig. 1(b)). Then the image coordinates (x, y) of point (X, Y, Z) are given by

$$x = f\frac{X}{Z}, \quad y = f\frac{Y}{Z}, \tag{2.3}$$

and its homogeneous coordinates are simply (X, Y, Z). That is, *scene coordinates can be identified with homogeneous coordinates on the image plane.*

Consider a plane $AX + BY + CZ = 0$ passing through the viewpoint O. Vector (A, B, C) designates the surface normal to this plane. This plane intersects the image plane $Z = f$ along the line $Ax + By + Cf = 0$, whose homogeneous coordinates are (A, B, C).

In summary:

- the N-vector **m** of a point P can be interpreted as the unit vector starting from the viewpoint O and pointing toward P (Fig. 2(a)), and
- the N-vector **n** of a line l can be interpreted as the unit vector normal to the plane passing through the viewpoint O and intersecting the image plane along l (Fig. 2(b)).

3. VANISHING POINTS AND VANISHING LINES

Let us begin with the well known fact that projections of parallel lines in the scene meet at a common "vanishing point." Formally, the *vanishing point* of a line in the scene is the limit of the projection of a point that moves along the line indefinitely in one direction (both directions define the same vanishing point). From Fig. 3(a), it is easy to confirm the following theorem (the formal proof is an easy exercise):

THEOREM 1. *A line in the scene extending along unit vector* **m** *has, when projected, a vanishing point of N-vector* **m**.

Since the vanishing point is determined by the 3-D orientation of the line alone, irrespective of its location in the scene, we see that

COROLLARY 1. *Projections of parallel lines in the scene intersect at a common vanishing point.*

As is well known, projections of planar surfaces mutually parallel in the scene define a common "vanishing line." Formally, the *vanishing line* of a planar surface in the scene is the set of all the vanishing points of lines lying in the surface. From Fig. 3(b), we can easily obtain the following theorem (again the formal proof is easy):

THEOREM 1'. *A planar surface in the scene whose unit surface normal is* **n** *has, when projected, a vanishing line of N-vector* **n**.

Since the vanishing line is determined by the 3-D orientation of the planar surface alone, irrespective of its location in the scene, we see that

COROLLARY 1'. *Projections of planar surfaces mutually parallel in the scene define a common vanishing line.*

In summary, if a vanishing point is detected on the image plane, its N-vector indicates the 3-D orientation of the line, and if a vanishing line is detected on the image plane, its N-vector indicates the surface normal to the planar surface.

4. FUNDAMENTAL DUALITY OF *N*-VECTORS

The following two theorems play an essential role in the subsequent discussions.

THEOREM 2. *The N-vector* **m** *of the intersection P of two lines l and l' whose N-vectors are* **n** *and* **n'**, *respectively, is given by*

$$\mathbf{m} = \pm N[\mathbf{n} \times \mathbf{n}']. \tag{4.1}$$

Proof. Vector **n** is normal to the plane passing through the viewpoint O and intersecting the image plane along line l. Similarly, vector **n'** is normal to the plane passing through the viewpoint O and intersecting the image plane along line l'. The N-vector **m** of their intersec-

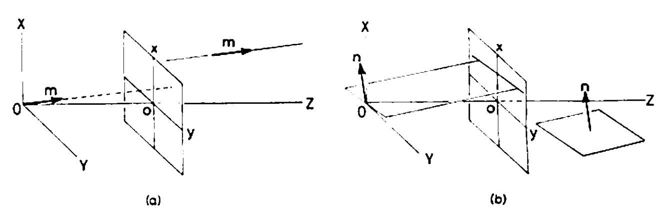

FIG. 3. (a) The vanishing point of a line in the scene. (b) The vanishing line of a planar surface in the scene.

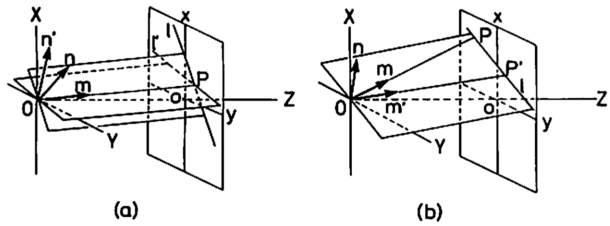

FIG. 4. (a) The N-vector **m** of the intersection P of two lines l and l'. (b) the N-vector **n** of the line l passing through two points P and P'.

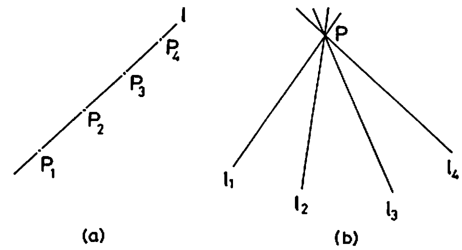

FIG. 5. (a) Collinearity of points. (b) Concurrency of lines.

tion P is perpendicular to both **n** and **n**′ (Fig. 4(a)). It follows that the N-vector **m** is obtained by normalizing $\pm\mathbf{n} \times \mathbf{n}'$. ■

THEOREM 2'. *The N-vector* **n** *of the line l passing through two points P and P′ whose N-vectors are* **m** *and* **m**′, *respectively, is given by*

$$\mathbf{n} = \pm N[\mathbf{m} \times \mathbf{m}'] \qquad (4.2)$$

Proof. Vector **m** indicates the 3-D orientation of the line starting from the viewpoint O and passing through point P. Similarly, vector **m**′ indicates the 3-D orientation of the line starting from the viewpoint O and passing through point P'. The N-vector **n** of the line l passing through them is perpendicular to both **m**′ and **m**′ (Fig. 4(b)). It follows that the N-vector **n** is obtained by normalizing $\pm\mathbf{m} \times \mathbf{m}'$. ■

Since Eqs. (4.1) and (4.2) do not involve divisions except for the final normalization, the computation is always kept within a finite domain. Even if lines l and l' are parallel on the image plane, the N-vector of their intersection is correctly computed as an ideal point as lone as l and l' are distinct. Similarly, wherever two points P and P' are located on the image plane (even at infinity), the N-vector of the line passing through them is correctly computed as long as P and P' are distinct.

Comparing Theorems 1 and 1' or Theorems 2 and 2', we immediately notice a striking similarity: *The roles of points and lines are interchangeable*. This *duality* is one of the most fundamental characteristics of projective geometry.

5. COLLINEARITY OF POINTS AND CONCURRENCY OF LINES

Formally, a *2-D projective space* is the set of all points and lines designated by *all N-vectors*, including *points at infinity (ideal points)* and the *line at infinity* (the *ideal line*): a point is regarded as an ideal point if the third component of its N-vector is 0, while a line is regarded as the ideal line if the first and the second components of its N-vector are both 0.

A point P of N-vector **m** and a line l of N-vector **n** are said to be *incident* to each other if[2]

$$(\mathbf{m}, \mathbf{n}) = 0 \qquad (5.1)$$

When point P and line l are incident to each other, we also say that point P is *on* line l, and line l passes through point P. Points are said to be *collinear* if there exists a line passing through all of them (Fig. 5(a)). Lines are said to be *concurrent* if there exists a point that is on all of them (Fig. 5(b)).

As pointed out earlier, *collinearity* and *concurrency* are the most fundamental concepts of projective geometry. However, their definitions are given above do not suit our *computational* approach because they are based on *existence* of something. There are many possibilities to redefine them as *constructive* computational processes. For example,

PROPOSITION 1. *Points P_α $\alpha = 1, \ldots, N$, are collinear if and only if the rank of their N-vectors $\mathbf{m}_\alpha$, $\alpha = 1, \ldots, N$, is less than* 3.

Proof. Note that the rank is defined as the maximum number of linearly independent vectors. In three dimensions, three vectors are linearly dependent if and only if they are coplanar. By definition, points P_α of N-vectors $\mathbf{m}_\alpha$ are collinear if and only if there exists a unit vector **n** such that $(\mathbf{m}_\alpha, \mathbf{n}) = 0$, $\alpha = 1, \ldots, N$. This implies that vectors $\mathbf{m}_\alpha$ are all perpendicular to the vector **n** and hence all lie on the plane perpendicular to **n**, which means that any three of them are linearly dependent. Hence, the rank is less than three. Conversely, if the rank is less than three, any three of the vectors $\mathbf{m}_\alpha$ are coplanar, which means that they all lie in a common plane. If we let **n** be the unit surface normal to it, we have $(\mathbf{n}_\alpha, \mathbf{n}) = 0$, $\alpha = 1, \ldots, N$. ■

THEOREM 3. *Points P_α, $\alpha = 1, \ldots, N$, are collinear if and only if the smallest eigenvalue of the moment matrix*,[3]

[2] In this paper, $(\mathbf{a}, \mathbf{b})$ designates the inner product of vectors **a** and **b**.
[3] In this paper, $\mathbf{a}^T$ (or $\mathbf{A}^T$) designates the transpose of vector **a** (or matrix **A**).

$$M = \sum_{\alpha=i}^{N} w_\alpha \mathbf{m}_\alpha \mathbf{m}_\alpha^{\mathrm{T}}, \tag{5.2}$$

is 0. The associated unit eigenvector **n** *is the N-vector of the line passing through the points* P_α, *where* w_α *are positive constants.*

Proof. The condition that $(\mathbf{m}_\alpha, \mathbf{n}) = 0$, $\alpha = 1, \ldots, N$, is equivalently rewritten as $\sum_{\alpha=1}^{N} w_\alpha(\mathbf{m}_\alpha, \mathbf{n})^2 = 0$, where w_α are positive constants. In terms of the matrix defined by Eq. (5.2), this condition is rewritten as

$$(\mathbf{n}, \mathbf{M}\mathbf{n}) = 0. \tag{5.3}$$

Since **M** is a symmetric and positive semi-definite matrix, this condition holds if and only if the smallest eigenvalue of **M** is 0 and **n** is the associated eigenvector. ■

COROLLARY 3. *Points* P_α, $\alpha = 1, \ldots, N$, *are collinear if and only if the moment matrix* (5.2) *is singular.*

In view of computational considerations, Theorem 3 is the most convenient among many alternative forms. For one thing, it not only provides a means to judge collinearity of points as a *computational* process but also computationally defines the N-vector of the line passing through all the points. Another reason is the *robustness* to noise. Suppose, for example, the data are not exactly accurate due to noise in the image. Then three linearly dependent vectors generally become linearly independent even if the noise is infinitesimally small. As a result, Proposition 1 cannot be used as a robust criterion to judge collinearity, since the rank of three or more vectors is almost always 3.

Consider Theorem 3, on the other hand. If the noise is small, the eigenvalues of the moment matrix (5.2), which are all nonnegative, are expected to change by a small amount. Hence, we can judge collinearity of points by checking how small the smallest eigenvalue of the moment matrix (5.2) is. If we check the proof of Theorem 3 carefully, it is easy to see that the associated unit eigenvector is the N-vector of the line *fitted* to point P_α by the *least-squares criterion*.

$$\sum_{\alpha=1}^{N} w_\alpha(\mathbf{m}_\alpha, \mathbf{n})^2 \rightarrow \min, \tag{5.4}$$

where w_α is the weight for the αth data. It is also easy to confirm that the smallest eigenvalue of the moment matrix (5.2) equals the *residual* $\sum_{\alpha=1}^{N} w_\alpha(\mathbf{m}_\alpha, \mathbf{n})^2$ for the resulting best fit.

Now that we have formulated collinearity of points, we automatically have a formulation of concurrency of lines thanks to the inherent duality of projective geometry: all we need to do is *interchange the roles of points and lines.* The formal proof runs exactly the same as for points.

PROPOSITION 1'. *Lines* l_α, $\alpha = 1, \ldots, N$, *are concurrent if and only if the rank of their N-vectors* $\mathbf{n}_\alpha$, $\alpha = 1, \ldots, N$, *is less than* 3.

THEOREM 3'. *Lines* l_α, $\alpha = 1, \ldots, N$, *are concurrent if and only if the smallest eigenvalue of the moment matrix*

$$N = \sum_{\alpha=1}^{N} w_\alpha \mathbf{n}_\alpha n_\alpha^{\mathrm{T}} \tag{5.5}$$

is 0. *The associated unit eigenvector* **m** *is the N-vector of the common intersection of lines* l_α, *where* w_α *are positive constants.*

COROLLARY 3'. *Lines* l_α, $\alpha = 1, \ldots, N$, *are concurrent if and only if the moment matrix* (5.5) *is singular.*

Again, Theorem 3' is the most convenient: It not only provides a means to judge concurrency of lines as a *computational* process but also computationally defines the N-vector of the common intersection of all the lines. It is also robust to noise: Theorem 3' can also be viewed as estimation of the common intersection of lines l_α by the *least-squares criterion*

$$\sum_{\alpha=1}^{N} w_\alpha(\mathbf{n}_\alpha, \mathbf{m})^2 \rightarrow \min, \tag{5.6}$$

where w_α is the weight for the αth data.

6. COLLINEATIONS AND CORRELATIONS

A one-to-one mapping that maps the set of all points to the set of all points *and* the set of all lines to the set of all lines is said to be a *collineation* if

(i) collinear points are mapped to colinear points,

(ii) concurrent lines are mapped to concurrent lines, and

(iii) the incidence is preserved (i.e., if a point is on a line, the mapped point is on the mapped line).

It can be proved that a collineation is written as a linear mapping of N-vectors in the form

$$\mathbf{m}' = \pm N[\mathbf{a}^{\mathrm{T}}\mathbf{m}], \quad \mathbf{n}' = \pm N[\mathbf{A}^{-1}\mathbf{n}], \tag{6.1}$$

where **A** is a three-dimensional nonsingular matrix, and **m** and **n** the N-vectors of the original point and line, while **m'** and **n'** are the N-vectors of the point and line after the mapping (we omit the proof).[4] Conversely, any three-

[4] The proof involves the invariance of the *cross ratio* (or *anharmonic ratio*) and *projective coordinates* defined in terms of the cross ratio under collineations. This is in fact one of the most sophisticated theorems of projective geometry.

dimensional nonsingular matrix **A** defines a collineation in the form of Eqs. (6.1), which is easily checked. The matrix **A** representing a collineation is unique up to a single scale factor.

In terms of inhomogeneous coordinates (i.e., image coordinates), the first of Eqs. (6.1) for $\mathbf{A} = (A_{ij})$, $i, j = 1, 2, 3$, is rewritten as

$$x' = f\frac{A_{11}x + A_{21}y + A_{31}f}{A_{13}x + A_{23}y + A_{33}f},$$

$$y' = f\frac{A_{12}x + A_{22}y + A_{32}f}{A_{13}x + A_{23}y + A_{33}f}, \qquad (6.2)$$

The set of all collineations of a 2-D projective space forms a group of transformations—called the *group of 2-D projective transformations*.

Since the matrix **A** has a scale indeterminacy, it has eight degrees of freedom. Hence, it can be determined if we find a correspondence of four points in general position or four lines in general position over two images.[5] Namely, *there exists a unique collineation that maps arbitrary four points in general position to arbitrary four points in general position, or arbitrary four lines in general position to arbitrary four lines in general position.* This is one of the most fundamental theorems in projective geometry.

A one-to-one mapping that maps the set of all points to the set of all "lines" *and* the set of all lines to the set of all "points" is said to be a *correlation* if

(i) collinear points are mapped to concurrent lines,

(ii) concurrent lines are mapped to collinear points, and

(iii) the incidence is preserved (i.e., if a point is on a line, the mapped line passes through the mapped point).

It can also be proved that a correlation is written as a linear mapping of N-vectors in the form

$$\mathbf{n}' = \pm N[\mathbf{Q}^{\mathrm{T}}\mathbf{m}], \quad \mathbf{m}' = \pm N[\mathbf{Q}^{-1}\mathbf{n}], \qquad (6.3)$$

where **Q** is a three-dimensional nonsingular matrix, and **m** and **n** the N-vectors of the original point and line, while **n'** and **m'** are the N-vectors of the line and point after the mapping (we omit the proof). Conversely, any three-dimensional nonsingular matrix **Q** defines a correlation in the form of Eqs. (6.3), which is easily checked. The matrix **Q** representing the correlation is unique up to a single scale factor.

[5] By *general position*, we mean that no three points are collinear or no three lines are concurrent.

7. POLARITIES, CONJUGACY, AND CONICS

A correlation is a *polarity* if whenever point P is mapped to line l, line l is also mapped to point P. In matrix representation, it is easy to see that *a correlation is a polarity if and only if the matrix* **Q** *is symmetric:*

$$\mathbf{Q}^{\mathrm{T}} = \mathbf{Q}. \qquad (7.1)$$

If point P is mapped to line l and line l is mapped to point P by a polarity, the point P is said to be the *pole* of the line l, and the line l the *polar* of the point P with respect to the polarity. If three lines l, l', and l'' are the polars of points P, P', and P'', respectively, the triangle defined by the three lines l, l', and l'' is called the *polar triangle* of triangle $\Delta PP'P''$. A triangle that is a polar triangle of itself is called a *self-polar triangle*.

Two points P and P' are mutually *conjugate* if the polar of P passes through P', and the polar of P' passes through P. Two lines l and l' are mutually *conjugate* if the pole of l is on l', and the pole of l' is on l. Hence, the three vertices (or edges) of a self-polar triangle are conjugate to each other. A point that is conjugate to itself is said to be *self-conjugate*. For a given polarity, the set of all self-conjugate points is called the *conic* of the polarity. It is easy to see that the conic of a polarity represented by matrix **Q** is the set of points whose N-vectors satisfy the homogeneous quadratic equation

$$(\mathbf{m}, \mathbf{Qm}) = 0. \qquad (7.2)$$

In terms of the inhomogeneous coordinates (i.e., image coordinates), the conic (7.2) for $\mathbf{Q} = (Q_{ij S2A})$, $i, j = 1, 2, 3$, is rewritten as

$$Q_{11}x^2 + 2Q_{12}xy + Q_{22}y^2 + 2f(Q_{31}x + Q_{32}y) + Q_{33}f^2 = 0. \qquad (7.3)$$

Conversely, a conic expressed as a homogeneous quadratic equation in N-vector **m** in the form of Eq. (7.2), or equivalently an inhomogeneous quadratic equation in image coordinates (x, y) in the form of Eq. (7.3), uniquely defines a polarity represented by the coefficient matrix $\mathbf{Q} = (Q_{ij})$, $i, j = 1, 2, 3$.

The set of all collineations that map conic (7.2) (or (7.3)) to itself forms a subgroup of the group of 2D projective transformations—called the subgroup *associated* with this conic (or the corresponding polarity).

8. POLES, POLARS, AND DUALITY THEOREM

Consider the following mapping (Fig. 6(a)):

A point is mapped to a line that has the same N-vector, and a line is mapped to a point that has the same N-vector.

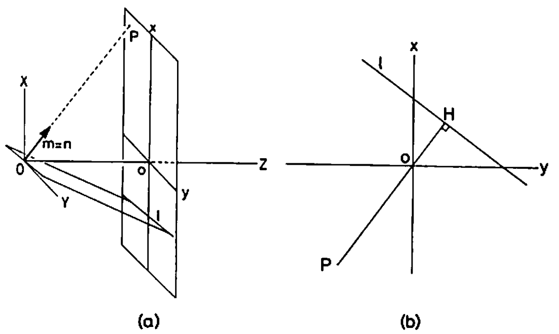

FIG. 6. (a) The standard polarity: Point P is the pole of line l, and line l is the polar of point P. (b) 2-D interpretation of the standard polarity.

According to the definition in the preceding section, this mapping is a correlation. Evidently, this correlation is also a polarity. Let us call this polarity the *standard polarity*. For the standard polarity, the *pole* of a line whose N-vector is **n** has N-vector **n**, and the *polar* of a point whose N-vector is **m** has N-vector **m**. In terms of image coordinates,

PROPOSITION 2. *The pole of line* $Ax + By + C = 0$ *is* $(f^2A/C, f^2B/C)$.[6]

PROPOSITION 2'. *The polar of point* (a, b) *is* $ax + by + f^2 = 0$.[7]

The standard polarity is also interpreted as the following 2-D geometric relationship on the image plane (the proof is easy; see Fig. 6(b)):

PROPOSITION 3. *Given a point P and a line l on the image plane, draw a line passing through the image origin o and the point P. Let H be the intersection of this line with l. Point P is the polar of line l, and line l is the polar of point P if and only if*

(*i*) *line PH is orthogonal to line l,*

(*ii*) *point P and point H are on the opposite sides of the image origin o, and*

(*iii*)
$$\overline{oP} \cdot \overline{oH} = f^2. \tag{8.1}$$

From the definition of polarity, we immediately obtain the following two theorems, on which the line detection algorithm known as the *Hough transform* is based (see Fig. 5).

THEOREM 4. *The poles of concurrent lines are collinear, and the common line passing through them is the polar of the common intersection.*

[6] If $C = 0$, the pole is understood to be an ideal point.

[7] If $a = b = 0$, the polar is also understood to be the ideal line.

THEOREM 4'. *The polars of collinear points are concurrent, and the common intersection is the pole of the common line.*

Since the matrix representing the standard polarity is the unit matrix **I**, the corresponding conic is $(\mathbf{m}, \mathbf{m}) = 0$, or in terms of image coordinates $x^2 + y^2 + f^2 = 0$. This conic describes an *imaginary circle* centered at the image origin o with *imaginary radius if*. This conic is known as the *absolute conic* [3]. It can be proved that the subgroup of the group of 2-D projective transformations associated with the absolute conic consists of the image transformations induced by *rotating the camera around the center of the lines relative to a stationary scene*. Such transformations are called *camera rotation transformations*, and their invariance properties play an important role in 3-D interpretation of objects and scenes [8–10, 12].

9. ORTHOGONALITY CRITERION IN TERMS OF CONJUGACY

For the standard polarity, points having N-vectors **m** and **m**' are *conjugate* to each other if and only if

$$(\mathbf{m}, \mathbf{m}') = 0, \tag{9.1}$$

and lines having N-vectors **n** and **n**' are *conjugate* to each other if an only if

$$(\mathbf{n}, \mathbf{n}') = 0. \tag{9.2}$$

(Fig. 7). In terms of image coordinates,

PROPOSITION 4. *Points* (a, b) *and* (a', b') *are conjugate to each other if and only if*

$$aa' + bb' + f^2 = 0. \tag{9.3}$$

PROPOSITION 4'. *Lines* $Ax + By + C = 0$ *and* $A'x + B'y + C' = 0$ *are conjugate to each other if and only if*

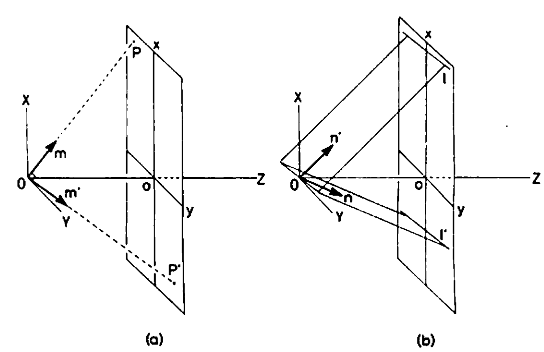

FIG. 7. (a) Points P and P' are conjugate to each other. (b) Lines l and l' are conjugate to each other.

$$AA' + BB' + \frac{CC'}{f^2} = 0. \quad (9.4)$$

It is not difficult to obtain the following 2-D interpretation of the conjugacy relationship on the image plane (we omit the proof; see Fig. 8):

PROPOSITION 5. *Let P and P' be two points on the image plane, and l the line passing through them. Let H be the foot of the perpendicular line drawn from the image origin o to l. Points P and P' are conjugate to each other if and only if they are on the opposites sides of H and*

$$\overline{PH} \cdot \overline{P'H} = \overline{oH}^2 + f^2. \quad (9.5)$$

PROPOSITION 5'. *Let l and l' be two lines on the image plane. Let P be their intersection, and o the image origin. Lines l and l' are conjugate to each other if and only if the angles they make with line oP from the opposite sides satisfy*

$$\cot \angle oPl \cot \angle oPl' = 1 + \left(\frac{\overline{oP}}{f}\right)^2. \quad (9.6)$$

Combining these results with the 3-D interpretation of the vanishing point and vanishing line (Theorems 1 and 1'), we obtain the following theorems concerning 3-D interpretation of the scene (we omit the proof; see Fig. 9(a)):

THEOREM 5. *Two lines are orthogonal to each other in the scene if and only if their vanishing points are conjugate to each other on the image plane.*

COROLLARY 5. *Three lines in the scene are mutually orthogonal if and only if their vanishing points define a self-polar triangle on the image plane.*

THEOREM 5'. *Two planar surfaces are orthogonal to each other in the scene if and only if their vanishing lines are conjugate to each other on the image plane.*

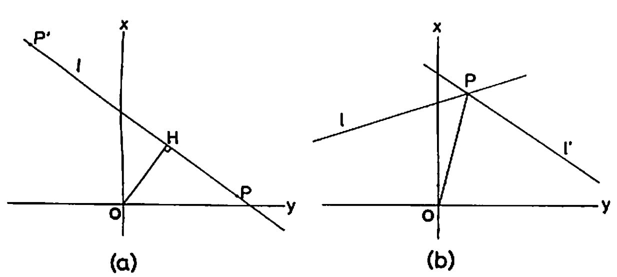

FIG. 8. (a) 2-D interpretation of two conjugate points. (b) 2-D interpretation of two conjugate lines.

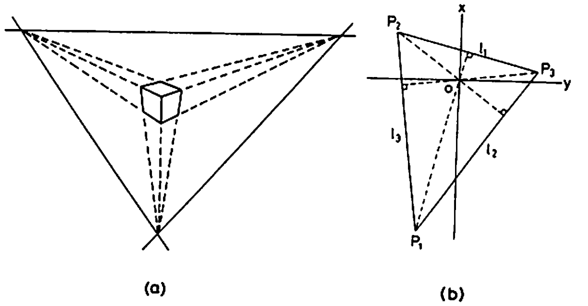

FIG. 9. (a) Interpretation of orthogonality and parallelism in the scene. (b) The orthocenter of a self-polar triangle.

COROLLARY 5'. *Three planar surfaces in the scene are mutually orthogonal if and only if their vanishing lines define a self-polar triangle on the image plane.*

In summary, once vanishing points and vanishing lines are detected on the image plane, we can immediately check the orthogonality of the corresponding lines and planar surfaces in the scene.

Theoretically, we can also determine the origin o of the image plane by applying the following proposition (Fig. 9(b)), which is a direct consequence of Proposition 3, but in practice the computation is very sensitive to image noise:

PROPOSITION 6. *The orthocenter of any self-polar triangle is at the image origin o.*

10. IMAGE SEQUENCE AND *N*-VECTORS

When observing a sequence of images of points moving rigidly in the scene, can we compute the 3-D structure of the points and their 3-D rigid motion from these images? This problem is known as *shape from motion,* and many computational theories have been presented. In the following, we show that the computation of structure and motion becomes very easy if the motion is limited to translations only.

Translational motions occur in many application domains, a typical situation being when objects are conveyed on a conveyer belt in an industrial environment or a mobile robot is proceeding along a straight path. Instead of "passively" exploiting such translations, the camera can also be "actively" controlled for the purpose of 3-D recognition—this paradigm is known as *active vision* [1]. *Stereo* can also be viewed as a motion because the use of two cameras whose optical axes are parallel is equivalent to observing a translation of the scene relative to a single camera.

Consider a point moving in the scene, and let $(x(t), y(t))$ be its perspective projection. The time derivative of the

N-vector of this point is given by

$$\dot{\mathbf{m}}(t) = \frac{1}{\sqrt{x^2 + y^2 + f^2}} \begin{pmatrix} \dot{x}(t) \\ \dot{y}(t) \\ 0 \end{pmatrix} - \frac{x(t)\dot{x}(t) + y(t)\dot{y}(t)}{\sqrt{x^2 + y^2 + f^2}} \begin{pmatrix} x(t) \\ y(t) \\ f \end{pmatrix} \qquad (10.1)$$

This is immediately obtained by differentiating $\mathbf{m}(t) = N[(x(t), y(t), f)^{\mathrm{T}}] = (x(t), y(t), f)^{\mathrm{T}}/\sqrt{x^2 + y^2 + f^2}$. Let us call $\dot{\mathbf{m}}(t)$ the *N-velocity* of point $(x(t), y(t))$. Since the N-vector $\mathbf{m}(t)$ is a unit vector, differentiation of $(\|\mathbf{m}(t)\|^2 =)\ (\mathbf{m}(t), \mathbf{m}(t)) = 1$ yields

PROPOSITION 7. *The N-vector and the N-velocity of a moving point are orthogonal to each other:*

$$(\mathbf{m}(t), \dot{\mathbf{m}}(t)) = 0. \qquad (10.2)$$

If a point is translating in the scene, its projection defines a straight trajectory in the course of its motion on the image plane.

PROPOSITION 8. *If* $\mathbf{m}(t)$ *and* $\dot{\mathbf{m}}(t)$ *are the N-vector and the N-velocity, respectively, of a projection of a translating point in the scene, the N-vector of its trajectory on the image plane is given by*

$$\mathbf{n} = \pm N[\mathbf{m}(t) \times \dot{\mathbf{m}}(t)]. \qquad (10.3)$$

Proof. Consider the plane passing through the viewpoint O and intersecting the image plane along the trajectory. Since $\mathbf{m}(t)$ and $\dot{\mathbf{m}}(t)$ are both contained in this plane (Fig. 10), the unit surface normal to this plane (the N-vector of the trajectory) is given by Eq. (10.3). ■

11. FOCUS OF EXPANSION

If we observe a sequence of images of points translating in the scene, all the image points seem to be moving, on the image plane, away from or toward a fixed point, which is known as the *focus of expansion* (Fig. 11). This

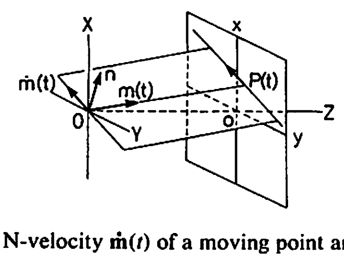

FIG. 10. The N-velocity $\dot{\mathbf{m}}(t)$ of a moving point and the N-vector $\mathbf{n}$ of its trajectory.

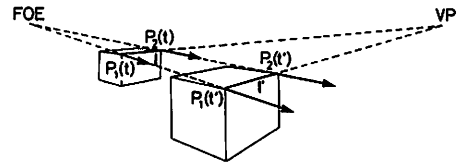

FIG. 11. The focus of expansion (*FOE*) of a translational motion and the vanishing point (*VP*) of a translating line segment.

is obvious from Theorem 1 and Corollary 1, since the focus of expansion is simply the vanishing point of mutually parallel trajectories in the scene. Thus,

THEOREM 6. *A point translating in the scene in the direction of unit vector* $\mathbf{u}$ *has, when projected onto the image plane, a focus of expansion whose N-vector is* $\mathbf{u}$.

COROLLARY 6. *Projections of points rigidly translating in the scene have a common focus of expansion on the image plane.*

The focus of expansion is easily computed as follows (including the case when the focus of expansion is an ideal point at infinity, which occurs when the lines connecting corresponding points are all parallel on the image plane):

PROPOSITION 9. *If two points* $P_1(t)$ *and* $P_2(t)$ *whose respective N-vectors are* $\mathbf{m}_1(t)$ *and* $\mathbf{m}_2(t)$ *in the first image correspond to points* $P_1(t')$ *and* $P_2(t')$ *whose respective N-vectors are* $\mathbf{m}_1(t')$ *and* $\mathbf{m}_2(t')$ *in the second image, the N-vector of the focus of expansion is given by*

$$\mathbf{u} = \pm N[N[\mathbf{m}_1(t) \times \mathbf{m}_1(t')] \times N[\mathbf{m}_2(t) \times \mathbf{m}_2(t')]] \qquad (11.1)$$

provided no two of $P_1(t)$, $P_2(t)$, $P_1(t')$, *and* $P_2(t')$ *coincide.*

Proof. The N-vector of the line passing through points $P_1(t)$ and $P_1(t')$ is $\pm N[\mathbf{m}_1(t) \times \mathbf{m}_1(t')]$, and the N-vector of the line passing through points $P_2(t)$ and $P_2(t')$ is $\pm N[\mathbf{m}_2(t) \times \mathbf{m}_2(t')]$ (Theorem 2'). The N-vector of their intersection is given by Eq. (11.1) (Theorem 2). ■

PROPOSITION 9'. *If images of two distinct points* $P_1(t)$ *and* $P_2(t)$ *whose N-vectors are* $\mathbf{m}_1(t)$ *and* $\mathbf{m}_2(t)$, *respectively, are moving on the image plane with N-velocities* $\dot{\mathbf{m}}_1(t)$ *and* $\dot{\mathbf{m}}_2(t)$, *respectively, the N-vector of the focus of expansion is given by*

$$\mathbf{u} = N[N[\mathbf{m}_1(t) \times \dot{\mathbf{m}}_1(t)] \times N[\mathbf{m}_2(t) \times \dot{\mathbf{m}}_2(t)]]. \qquad (11.2)$$

Proof. The focus of expansion is the intersection of the trajectories of the points on the image plane. The N-vectors of the trajectories of points $P_1(t)$ and $P_2(t)$ are respectively $\pm N[\mathbf{m}_1(t) \times \dot{\mathbf{m}}_1(t)]$ and $\pm N[\mathbf{m}_2(t) \times \dot{\mathbf{m}}_2(t)]$ (Proposition 8). The N-vector of the intersection of their trajectories is given by Eq. (11.2) (Theorem 2). ■

12. STRUCTURES FROM TRANSLATIONAL MOTION

The most important consequence of restricting the motion to translations is that the 3-D structure of objects can be computed by a very simple procedure. The following two propositions are most fundamental.

PROPOSITION 10. *Let $P_1(t)P_2(t)$ be the projection, at time t, of a line segment translating in the scene, and let $P_1(t')P_2(t')$ be the projection of the same line segment at time t'. Let $\mathbf{m}_1(t)$, $\mathbf{m}_2(t)$, $\mathbf{m}_1(t')$, and $\mathbf{m}_2(t')$ be the N-vectors of points $P_1(t)$, $P_2(t)$, $P_1(t')$, and $P_2(t')$, respectively. If no two of these four points coincide, the 3-D orientation of the line segment in the scene is given by unit vector*

$$\mathbf{m}_{12} = \pm N[N[\mathbf{m}_1(t) \times \mathbf{m}_2(t)] \times N[\mathbf{m}_1(t') \times \mathbf{m}_2(t')]]. \tag{12.1}$$

Proof. See Fig. 11. The N-vector of the line l passing through points $P_1(t)$ and $P_2(t)$ is $\mathbf{n} = \pm N[\mathbf{m}_1(t) \times \mathbf{m}_2(t)]$, and the N-vector of the line l' passing through points $P_1(t')$ and $P_2(t')$ is $\mathbf{n}' = \pm N[\mathbf{m}_1(t') \times \mathbf{m}_2(t')]$ (Theorem 2'). Since l and l' are projections of parallel lines in the scene, their intersection on the image plane is their vanishing point. The N-vector of the vanishing point is $m_{12} = \pm N[\mathbf{n} \times \mathbf{n}']$ (Theorem 2). Since the N-vector of the vanishing point indicates the 3-D orientation of these lines (Theorem 1), we obtain the assertion. ∎

PROPOSITION 10'. *If $\mathbf{m}_1(t)$ and $\mathbf{m}_2(t)$ are the N-vectors of points $P_1(t)$ and $P_2(t)$, respectively, and if $\dot{\mathbf{m}}_1(t)$ and $\dot{\mathbf{m}}_2(t)$ are their respective N-velocities, the 3-D orientation of the line segment in the scene is*[8]

$$\mathbf{m}_{12} = \pm N[|\mathbf{m}_1(t), \mathbf{m}_2(t), \dot{\mathbf{m}}_2(t)|\mathbf{m}_1(t) - |\mathbf{m}_1(t), \mathbf{m}_2(t), \dot{\mathbf{m}}_1(t)|\mathbf{m}_2(t)]. \tag{12.2}$$

Proof. If we put $\Delta t = t' - t$, and substitute

$$\begin{aligned} \mathbf{m}_1(t') &= \mathbf{m}_1(t) + \dot{\mathbf{m}}_1(t)\Delta t + O(\Delta t^2), \\ \mathbf{m}_2(t') &= \mathbf{m}_2(t) + \dot{\mathbf{m}}_2(t)\Delta t + O(\Delta t^2) \end{aligned} \tag{12.3}$$

into Eq. (12.1), we obtain

$$\begin{aligned} \mathbf{m}_{12} &= \pm N[N[\mathbf{m}_1(t) \times \mathbf{m}_2(t)] \times N[(\mathbf{m}_1(t) + \dot{\mathbf{m}}_1(t)\Delta t + O(\Delta t^2)) \times (\mathbf{m}_2(t) + \dot{\mathbf{m}}_2(t)\Delta t + O(\Delta t^2))]] \\ &= \pm N[(\mathbf{m}_1(t) \times \mathbf{m}_2(t)) \times [\mathbf{m}_1(t) \times \mathbf{m}_2(t) + (\dot{\mathbf{m}}_1(t) \times \mathbf{m}_2(t) + \mathbf{m}_1(t) \times \dot{\mathbf{m}}_2(t))\Delta t + O(\Delta t^2)]] \\ &= \pm N[[(\mathbf{m}_1(t) \times \mathbf{m}_2(t)) \times (\dot{\mathbf{m}}_1(t) \times \mathbf{m}_2(t)) + (\mathbf{m}_1(t) \times \mathbf{m}_2(t)) \times (\mathbf{m}_1(t) \times \dot{\mathbf{m}}_2(t))]\Delta t + O(\Delta t^2)], \end{aligned} \tag{12.4}$$

[8] In this paper, $|\mathbf{a}, \mathbf{b}, \mathbf{c}| = (\mathbf{a} \times \mathbf{b}, \mathbf{c}) = (\mathbf{b} \times \mathbf{c}, \mathbf{a}) = (\mathbf{c} \times \mathbf{a}, \mathbf{b})$ denotes the *scalar triple product* of three vectors **a**, **b**, and **c**.

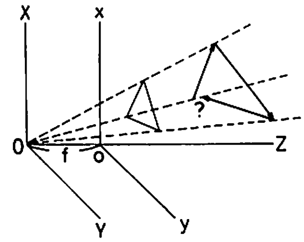

FIG. 12. Inconsistency can arise in 3-D reconstruction if the image data are not exact.

where we have removed the inner normalization operations (the outermost one assures the same result). Since the operand of the normalization operator N can be multiplied by any scalar, we can obtain Eq. (12.2) by dividing the operand of the normalization by Δt, taking the limit of $\Delta t \to 0$, and using the identity $(\mathbf{a} \times \mathbf{b}) \times (\mathbf{a} \times \mathbf{c}) = |\mathbf{a}, \mathbf{b}, \mathbf{c}|\mathbf{a}$. ∎

In summary, the 3-D configuration of points rigidly translating in the scene is, in principle, uniquely determined *up to a single scale factor* if point-to-point correspondence is known over two images, or if image velocities are observed in one image. This is because the 3-D orientation of the line segment connecting any two points is computed by Proposition 10 or 10', and its 3-D configuration is reconstructed by placing it in the scene according to the computed 3-D orientation in such a way that its projection coincides with the observed image. Thus, the entire 3-D configuration is successively reconstructed once the depth of the initial segment is fixed.

In practice, however, images are often not accurate, and inconsistencies may arise. For example, the starting point and the ending point of line segments forming a closed loop may not coincide (Fig. 12). For a unique and robust 3-D reconstruction, we need an *optimization* technique (see [11, 12, 26] for details).

13. DISPARITY MAPS AND DEPTH MAPS FOR STEREO

Since translation of the scene is equivalent to translation of the camera relative to the scene, *stereo* can be treated in the same way as translation of the scene. Here, we consider two cameras whose optical axes are parallel (this condition is not really necessary, as we will discuss later). Choose one camera as a reference, and define an XYZ coordinate system based on it with origin O at the center of its lens and the Z-axis along its optical axis. Let **b** be the vector indicating the center of the lens of the other camera with respect to this coordinate system (Fig. 13(a)). Let us call it the *base-line vector*.

If the image planes of these two cameras are identified, the two images can be analyzed in the same way as for a translational motion. For instance, the lines connecting corresponding points, called *epipolars*, meet at a single

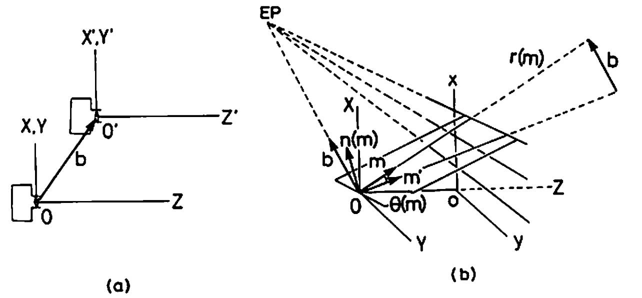

FIG. 13. (a) Stereo configuration and the base-line vector **b**. (b) The epipole (*EP*) and epiolars of a stereo system.

point on the image plane, which is called the *epipole* (Fig. 13(b)). Theorem 6 and Corollary 6 are respectively rephrased as follows:

THEOREM 7. *For a stereo system of base-line vector* **b**, *a pair of corresponding points on the image plane define an epipole whose N-vector is* $\mathbf{u} = \pm N[\mathbf{b}]$.

COROLLARY 7. *All epipolars meet at a common epipole.*

Thus, once point-to-point correspondence are detected, we can immediately compute the epipole and hence the orientation of the base-line vector **b**. Since the base-line vector **b** is usually known from the camera setting geometry, the above fact can also be used as a means of detecting the point-to-point correspondences over the two images. Viewed in this way, the above fact is called the *epipolar constraint*.

For a given point, the epipolar passing through it is easily computed. The following proposition is obvious (see Fig. 13(b)).

PROPOSITION 11. *For a stereo system of base-line vector* **b**, *the N-vector* $\mathbf{n}(\mathbf{m})$ *of the epipolar passing through a point of N-vector* **m** *is*

$$\mathbf{n}(\mathbf{m}) = \pm N[\mathbf{b} \times \mathbf{m}]. \tag{13.1}$$

Since $\mathbf{n}(\mathbf{m})$ can also be viewed as a function to give a point its epipolar, it is called the *epipolar map*. If a point of N-vector **m** in the reference image corresponds to a point of N-vector $\mathbf{m}'$ in the other image, the angle $\theta(\mathbf{m}) = \cos^{-1}(\mathbf{m}, \mathbf{m}')$ is called the *disparity*. Since $\theta(\mathbf{m})$ can be viewed as a function to give a point its disparity, it is also called the *disparity map*.

In the following, we assume that N-vectors of points are assumed to be signed so that their Z-components are nonnegative. Let $r(\mathbf{m})$ be the distance, from the viewpoint O of the reference camera, of the point whose N-vector is **m**. Since $r(\mathbf{m})$ can be viewed as a function to give a point its depth, it is called the *depth map*. The depth map $r(\mathbf{m})$ is uniquely computed from the disparity map $\theta(\mathbf{m})$ as follows (we omit the proof):

THEOREM 8. *For a stereo system of base-line vector* **b**, *the depth map* $r(\mathbf{m})$ *is given by*

$$r(\mathbf{m}) = (\mathbf{b}, \mathbf{m}) + \|\mathbf{b} \times \mathbf{m}\| \cot \theta(\mathbf{m}), \tag{13.2}$$

where $\theta(\mathbf{m})$ *is the disparity map.*

Combining this with Theorem 7, we see that the depth map $r(\mathbf{m})$ *is computed up to a scale factor* from the disparity map $\theta(\mathbf{m})$ *without the knowledge of the base-line vector* **b**. Namely,

COROLLARY 8. *If the N-vector* **u** *of the epipole is signed so that it extends in the direction of the base-line vector* **b**, *the depth map* $r(\mathbf{m})$ *is computed from the disparity map* $\theta(\mathbf{m})$ *by*

$$r(\mathbf{m}) = k[(\mathbf{u}, \mathbf{m}) + \|\mathbf{u} \times \mathbf{m}\| \cot \theta(\mathbf{m})], \tag{13.3}$$

where k is a positive scale factor.

So far, we have assumed that the optical axes of the two cameras are parallel, but this condition is not necessary. If the second camera is rotated by **R** with respect to the reference camera, a vector $\mathbf{m}'$ with respect to the second camera equals $\mathbf{R}\mathbf{m}'$ with respect to the reference camera. Hence, all we need to do is replace N-vector $\mathbf{m}'$ for the second camera by vector $\mathbf{R}\mathbf{m}'$ (this is exactly the *camera rotation transformation* [8–10, 12] mentioned earlier). Thus, as long as the relative rotation **R** between the two cameras is known, the orientations of the optical axes of the two cameras are irrelevant.

14. CONCLUDING REMARKS

We have reformulated projective geometry so that it can be used as a tool for 3-D analysis of images, emphasizing the *computational* aspects. In projective geometry, all concepts such as collineations, correlations, polarities, poles, polars, conics, and conjugacy are defined in abstract terms, while here they are interpreted as *computational processes* in terms of *N-vectors*. We also gave these mathematical concepts their *3-D interpretations* by regarding 2-D images as perspective projection of 3-D scenes, and presented a mathematical formalism for analyzing *translational motion* and *stereo* in terms of N-vectors.

As typical applications of our formalism, we give a brief description of the interpretation of an image of a rectangle in Appendix A, the 3-D road shape reconstruction in Appendix B, and the 3-D motion analysis of a planar surface in Appendix C.

In this paper, we have assumed that all image data are accurate. If noise is involved, a means of estimating and testing true configurations becomes necessary. This problem is studied in detail in [13] in the framework presented here.

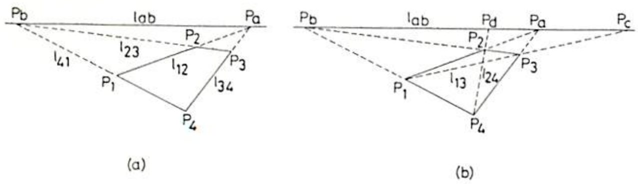

FIG. A. (a) A projection image of a rectangle. (b) A projection image of a square.

APPENDIX A. INTERPRETATION OF A RECTANGLE

Consider a projection of a rectangle in a scene. Let the four vertices on the image plane be labeled clockwise as P_1, P_2, P_3, and P_4 (Fig. A(a)). Let $\mathbf{m}_1$, $\mathbf{m}_2$, $\mathbf{m}_3$, and $\mathbf{m}_4$ be their respective N-vectors. If the four edges are labeled as l_{12}, l_{23}, l_{34}, and l_{41} as shown in the figure, their respective N-vectors are given by Theorem 2' as

$$\mathbf{n}_{12} = \pm N[\mathbf{m}_1 \times \mathbf{m}_2], \quad \mathbf{n}_{23} = \pm N[\mathbf{m}_2 \times \mathbf{m}_3],$$
$$\mathbf{n}_{34} = \pm N[\mathbf{m}_3 \times \mathbf{m}_4], \quad \mathbf{n}_{41} = \pm N[\mathbf{m}_4 \times \mathbf{m}_1]. \tag{A.1}$$

Since lines l_{12} and l_{34} are projections of parallel lines in the scene, their intersection P_a on the image plane is their vanishing point. Similarly, the intersection P_b of lines l_{23} and l_{41} is their vanishing point. Their N-vectors are given by Theorem 2 as

$$\mathbf{m}_a = \pm N[\mathbf{n}_{12} \times \mathbf{n}_{34}], \quad \mathbf{m}_b = \pm N[\mathbf{n}_{23} \times \mathbf{n}_{41}], \tag{A.2}$$

respectively. It follows that the N-vector of the vanishing line l_{ab} of this rectangle is given by Theorem 2' as

$$\mathbf{n}_{ab} = \pm N[\mathbf{m}_a \times \mathbf{m}_b]. \tag{A.3}$$

The N-vector $\mathbf{m}_a$ of the vanishing point P_a indicates the 3-D orientation of lines l_{12} and l_{34}, and the N-vector $\mathbf{m}_b$ of the vanishing point P_b indicates the 3-D orientation of lines l_{23} and l_{41} (Theorem 1). The N-vector $\mathbf{n}_{ab}$ of the vanishing line l_{ab} indicates the unit surface normal to the rectangle (Theorem 1').

Since adjacent edges of a rectangle are perpendicular to each other, the vanishing points P_a and P_b are mutually conjugate (Theorem 5), and their N-vectors are mutually orthogonal:

$$(\mathbf{m}_a, \mathbf{m}_b) = 0. \tag{A.4}$$

If the rectangle we are viewing is, in addition, known to be a square in the scene, an additional constraint is obtained. Let l_{13} be the diagonal passing through P_1 and P_3, and l_{24} the diagonal passing through P_2 and P_4 (Fig. A(b)). Their N-vectors are

$$\mathbf{n}_{13} = \pm N[\mathbf{m}_1 \times \mathbf{m}_3], \quad \mathbf{n}_{24} = \pm N[\mathbf{m}_2 \times \mathbf{m}_4], \tag{A.5}$$

respectively (Theorem 2').

It follows that the intersection P_c of the line l_{13} with the vanishing line l_{ab} is the vanishing point of l_{13}. Similarly, the intersection P_d of the line l_{24} with the vanishing line l_{ab} is the vanishing point of l_{24}. Their N-vectors are

$$\mathbf{m}_c = \pm N[\mathbf{n}_{13} \times \mathbf{n}_{ab}], \quad \mathbf{m}_d = \pm N[\mathbf{n}_{24} \times \mathbf{n}_{ab}], \tag{A.6}$$

respectively (Theorem 2).

Since they indicate the 3-D orientations of the diagonals l_{13} and l_{24} (Theorem 1), and they are orthogonal to each other, the vanishing points P_c and P_d are mutually conjugate (Theorem 5), and their N-vectors are mutually orthogonal:

$$(\mathbf{m}_c, \mathbf{m}_d) = 0. \tag{A.7}$$

Note that these constraints can be expressed in terms of the original N-vectors $\mathbf{m}_1$, $\mathbf{m}_2$, $\mathbf{m}_3$, and $\mathbf{m}_4$ *irrespective of the positions of the vanishing points P_a, P_b, P_c, and P_d on the image plane*. In other words, these vanishing points need not appear within the original image itself. These facts can be used as *camera calibration* of the focal length f. Namely, the value of f is adjusted to satisfy these conjugacy constraints. The camera position and orientation can also be calibrated by using a similar technique [14]. For various techniques of camera calibration, see [2, 4, 18, 19].

B. INTERPRETATION OF A ROAD

Today, research on autonomous land vehicles is conducted all across the world [6, 17, 27, 30, 31]. If the vehicle is to move along an arbitrary road in an uncontrolled environment, it requires sophisticated modules including an *image analysis module*, a *geometric reasoning module*, a *path planning module*, and a *navigation module*.

Consider the geometric reasoning module. 3-D data can be obtained by a direct measurement using stereo or range sensors for the part near the vehicle. However, if an appropriate *model* of the road is available, the 3-D road shape can be computed over a very long range from a single image [15, 16, 20, 23].

Suppose the road has a "constant width." This means that there exists a one-to-one correspondence between the two road boundaries in the scene such that every line segment connecting two corresponding points—let us call such a line segment a *cross segment*—meets the two road boundaries perpendicularly; the *width* of the road is defined as the length of the cross segment. Furthermore,

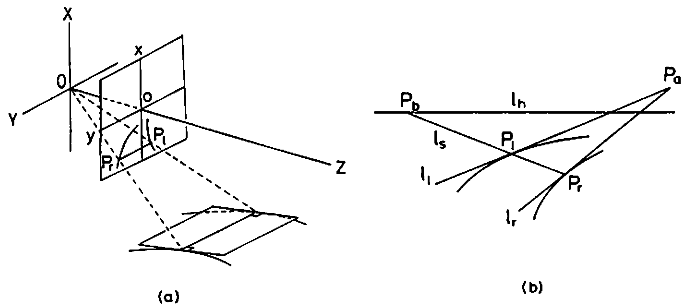

FIG. B. (a) Perspective projection of a road and its cross-segment. (b) The conjugacy constraint of a road image.

suppose the road is such that all cross segments are horizontal, and the tangents to the road boundary at corresponding points are parallel in the scene (Fig. B(a)). This *local flatness approximation* is a very good approximation for usual well constructed roads [23].

Consider a projection image of a road. We assume that smooth spline curves have already been fitted to the two road boundaries. Let P_r and P_l be the projections of two corresponding points on the image plane, and $\mathbf{m}_l$ and $\mathbf{m}_r$ their respective N-vectors (Fig. B(b)). Let l_l and l_r be the tangents to the road boundary image at P_l and P_r, respectively, and $\mathbf{n}_l$ and $\mathbf{n}_r$ their respective N-vectors. Since these two lines are parallel in the scene, their intersection P_a is their vanishing point; its N-vector is

$$\mathbf{m}_a = \pm N[\mathbf{n}_l \times \mathbf{n}_r] \qquad \text{(B.1)}$$

(Theorem 2), which indicates the 3-D orientation of the tangents in the scene (Theorem 1).

Let l_h be the image of the "horizon"—the vanishing line of a horizontal surface in the scene. Its N-vector $\mathbf{n}_h$ indicates the vertical orientation in the scene (Theorem 1'). The horizon l_s can be computed from the camera orientation in the scene; the horizon itself need not appear in the image.

The N-vector of the line l_s connecting P_l and P_r is

$$\mathbf{n}_s = \pm N[\mathbf{m}_l \times \mathbf{m}_r] \qquad \text{(B.2)}$$

(Theorem 2'). Since all cross segments are horizontal, the intersection P_b of l_s with the horizon l_h is the vanishing point of l_s. Its N-vector is

$$\mathbf{m}_b = \pm N[\mathbf{n}_s \times \mathbf{n}_h] \qquad \text{(B.3)}$$

(Theorem 2), which indicates the 3-D orientation of the cross segment (Theorem 1).

Since the cross segment and the road tangents at its endpoints are orthogonal in the scene, the vanishing points P_a and P_b are mutually conjugate (Theorem 5), and their N-vectors are mutually orthogonal:

$$(\mathbf{m}_a, \mathbf{m}_b) = 0. \qquad \text{(B.4)}$$

If Eqs. (B.1)–(B.3) are substituted into this, we have

$$(\mathbf{m}_h, \mathbf{m}_r)|\mathbf{m}_l, \mathbf{n}_l, \mathbf{n}_r| + (\mathbf{m}_h, \mathbf{m}_l)|\mathbf{m}_r, \mathbf{n}_r, \mathbf{n}_l| = 0. \qquad \text{(B.5)}$$

Using this constraint, we can determine the correspondence between the two road boundaries [16, 23]: given a point P_l on one boundary, the corresponding point on the other boundary is sought in such a way that this constraint is satisfied. Once the correspondence is established, it is easy to reconstruct the 3-D shape of the road (see [16] for the details of this procedure).

C. INTERPRETATION OF PLANAR SURFACE MOTION

As mentioned in Section 10, computing 3-D motion from two images—*shape from motion*—is one of the central problems of computer vision. Among various possible formulations, motions of a planar surface in the scene are important both theoretically and practically [7, 10, 21, 25, 28, 29].

Consider the image transformation induced on the image plane by the projection of a motion of a planar surface in the scene. Evidently, collinear points are mapped to collinear points, concurrent lines are mapped to concurrent lines, and the incidence is preserved. Hence, the image transformation is a *collineation* (Section 6; see Fig. C1(a)). This means that if a point of N-vector $\mathbf{m}$ and a line of N-vector $\mathbf{n}$ move, after a motion, to a point of N-vector $\mathbf{m}'$ and a line of N-vector $\mathbf{n}'$, respectively, there exists a three-dimensional nonsingular matrix $\mathbf{A}$ such that the transformation is given by

$$\mathbf{m}' = \pm N[\mathbf{A}^{\mathrm{T}}\mathbf{m}], \quad \mathbf{n}' = \pm N[A^{-1}\mathbf{n}]. \qquad \text{(C.1)}$$

In terms of image coordinates, the first equation is rewritten as

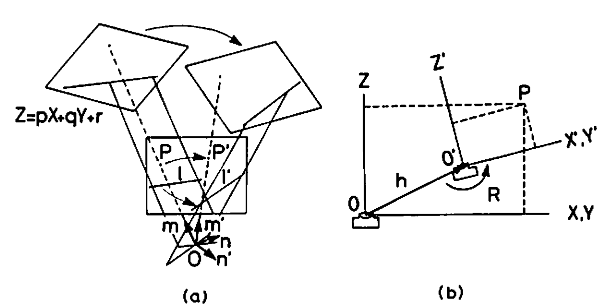

FIG. C1. (a) The collineation induced by the motion of a planar surface. (b) Camera motion and the motion parameters.

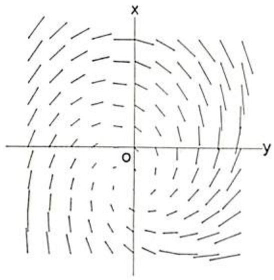

FIG. C2. Optical flow.

$$x' = f\frac{A_{11}x + A_{21}y + A_{31}f}{A_{13}x + A_{23}y + A_{33}f},$$
$$y' = f\frac{A_{12}x + A_{22}y + A_{32}f}{A_{13}x + A_{23}y + A_{33}f}. \tag{C.2}$$

Since the motion of an object is equivalent to the motion of the camera relative to the object, consider camera motion. Suppose the camera is first rotated by matrix $\mathbf{R}$ and then displaced by vector $\mathbf{h}$, where $\mathbf{R}$ and $\mathbf{h}$ are defined with respect to the initial camera (Fig. C1(b)). Let us call $\{\mathbf{R}, \mathbf{h}\}$ the *motion parameters*.

Let the equation of the planar surface be $n_1X + n_2Y + n_3Z = d$, where $n_1^2 + n_2^2 + n_3^2 = 1$ and $d > 0$. The unit vector $\mathbf{n} = (n_1, n_2, n_3)^T$ is the surface normal pointing away from the viewpoint O, and d is the distance of the surface from the viewpoint O. We exclude the case where $d = 0$ (the surface passes through the viewpoint O and hence is invisible). Let us call $\{\mathbf{n}, d\}$ the *surface parameters*.

We omit the proof, but it can be proved that the matrix $\mathbf{A}$ of collineation (6.1) is given by

$$A = \frac{1}{k}[\mathbf{I} - \mathbf{n}\tilde{\mathbf{h}}^T]\mathbf{R}, \tag{C.3}$$

where $\tilde{\mathbf{h}} = \mathbf{h}/d$, and k is an arbitrary nonzero constant.

Suppose we have determined the matrix $\mathbf{A}$ up to a scale factor. If we multiply it by an appropriate constant so that det $\mathbf{A} = 1$, the motion parameters $\{\mathbf{R}, \mathbf{h}\}$ and the surface parameters $\{\mathbf{n}, d\}$ are computed as follows (the depth r is indeterminate; this solution is a refinement of the one given in [28, 29]):

1. Let $\sigma_1^2, \sigma_2^2, \sigma_3^2$ ($\sigma_1 \geq \sigma_2 \geq \sigma_3$) be the eigenvalues of the symmetric matrix $\mathbf{AA}^T$.[9] Let $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ be an orthonormal system of corresponding eigenvectors.

2. If $\sigma_1 = \sigma_2 = \sigma_3$ ($= 1$), the motion parameters are

$$\mathbf{h} = \mathbf{0}, \quad \mathbf{R} = \mathbf{A}, \tag{C.4}$$

and the unit surface normal $\mathbf{n}$ is indeterminate. Otherwise:

3. The unit surface normal $\mathbf{n}$ is given by

$$\mathbf{n} = \frac{\varepsilon}{\sqrt{\sigma_1^2 - \sigma_3^2}}[\pm\sqrt{\sigma_1^2 - \sigma_2^2}\mathbf{u}_1 + \sqrt{\sigma_2^2 - \sigma_3^2}\mathbf{u}_3], \tag{C.5}$$

and the scaled translation vector $\tilde{\mathbf{h}}$ is given by

$$\tilde{\mathbf{h}} = \frac{\varepsilon}{\sigma_2^2}\sqrt{\frac{\sigma_1 + \sigma_3}{\sigma_1 - \sigma_3}}[\mp\sigma_3\sqrt{\sigma_1^2 - \sigma_2^2}\mathbf{u}_1 + \sigma_1\sqrt{\sigma_2^2 - \sigma_3^2}\mathbf{u}_3], \tag{C.6}$$

where the double sign in Eq. (C.6) corresponds to that in Eq. (C.5). Here, $\varepsilon = \pm 1$, and from the two signs, the one that makes $n_3 > 0$ is chosen if $n_3 \neq 0$. If $n_3 = 0$, both signs give the solution.

4. The rotation matrix $\mathbf{R}$ is given by

$$R = \frac{1}{\sigma_2}[\mathbf{I} + \sigma_2^3\mathbf{n}\tilde{\mathbf{h}}^T]A. \tag{C.7}$$

If the motion is instantaneous, we observe an *optical flow* (Fig. C2). In terms of the *N-velocity* (Section 10), it can be proved (we omit the proof) that the flow has the form[10]

$$\dot{\mathbf{m}} = \mathbf{W}^T\mathbf{m} - (\mathbf{m}, \mathbf{W}^T\mathbf{m})\mathbf{m},$$
$$\dot{\mathbf{n}} = -\mathbf{W}\mathbf{n} + (\mathbf{n}, \mathbf{W}\mathbf{n})\mathbf{n}, \tag{C.8}$$

where $\mathbf{W}$ is a matrix of trace 0. In terms of image coordinates, the first of Eqs. (C.8) is rewritten as

$$\dot{x} = fW_{31} + (W_{11} - W_{33})x + W_{21}y - \frac{1}{f}(W_{13}x + W_{23}y)x,$$
$$\dot{y} = fW_{32} + W_{12}x + (W_{22} - W_{33})y - \frac{1}{f}(W_{13}x + W_{23}y)y. \tag{C.9}$$

Suppose the center of the camera is translating by translational velocity $\boldsymbol{\nu}$ and the camera is rotating around it by angular velocity $\boldsymbol{\omega} = (\omega_1, \omega_2, \omega_3)^T$. Let us call $\{\boldsymbol{\nu}, \boldsymbol{\omega}\}$ the (instantaneous) *motion parameters*. It can be proved (we omit the proof) that the matrix $\mathbf{W}$ is written as

[9] The three values σ_1, σ_2, and σ_3 are called the *singular values* of matrix $\mathbf{A}$.

[10] The second one describes the "optical flow of lines."

$$W = \frac{1}{3}(\mathbf{n}, \boldsymbol{\nu}) - \mathbf{n}\tilde{\boldsymbol{\nu}}^{\mathrm{T}} + \boldsymbol{\Omega}, \quad \boldsymbol{\Omega} = \begin{bmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{bmatrix}, \tag{C.10}$$

where $\tilde{\nu} = \nu/d$.

If the matrix **W** is estimated by fitting Eqs. (C.9) to (at least four) points or (at least four) lines moving on the image plane, the motion parameters $\{\boldsymbol{\nu}, \boldsymbol{\omega}\}$ and the surface parameters $\{\mathbf{n}, d\}$ can be computed from **W**. The following solution is a refinement of the one given in [21] (for alternative solutions, see [7, 12, 25]).

1. Compute the symmetric matrix $\mathbf{W}_s$ and the vector **w** by

$$\mathbf{W}_s = \frac{1}{2}(\mathbf{W} + \mathbf{W}^{\mathrm{T}}), \quad \mathbf{w} = \begin{bmatrix} -(W_{23} - W_{32})/2 \\ -(W_{31} - W_{13})/2 \\ -(W_{12} - W_{21})/2 \end{bmatrix}. \tag{C.11}$$

2. If $\mathbf{W}_s = \mathbf{O}$, then $\boldsymbol{\nu} = \mathbf{0}$ and $\boldsymbol{\omega} = \mathbf{w}$, while the surface unit normal **n** is indeterminate. Otherwise:

3. Let $\sigma_1 \geq \sigma_2 \geq \sigma_3$ be the eigenvalues of the symmetric matrix $\mathbf{W}_s$, and let $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ be an orthonormal system of corresponding eigenvectors.

4. The unit surface normal **n** is given by

$$\mathbf{n} = \frac{\varepsilon}{\sqrt{\sigma_1 - \sigma_3}}[\pm\sqrt{\sigma_1 - \sigma_2}\mathbf{u}_1 + \sqrt{\sigma_2 - \sigma_3}\mathbf{u}_3], \tag{C.12}$$

and the scaled translation velocity $\tilde{\boldsymbol{\nu}}$ is given by

$$\tilde{\boldsymbol{\nu}} = \varepsilon\sqrt{\sigma_1 - \sigma_3}[\mp\sqrt{\sigma_1 - \sigma_2}\mathbf{u}_1 + \sqrt{\sigma_2 - \sigma_3}\mathbf{u}_3], \tag{C.13}$$

where the double sign in Eq. (C.13) corresponds to that of Eq. (C.12). Here, $\varepsilon = \pm 1$, and from the two signs, the one that makes $n_3 > 0$ is chosen if $n_3 \neq 0$. If $n_3 = 0$, both signs give the solution.

5. The angular velocity is given by

$$\boldsymbol{\omega} = \mathbf{w} - \frac{1}{2}\mathbf{n} \times \tilde{\boldsymbol{\nu}}. \tag{C.14}$$

REFERENCES

1. J. (Y.) Aloimonos, I. Weiss, and A. Bandyopadhyay, Active vision, *Int. J. Comput. Vision* **1** (1988), 333–356.
2. B. Caprile and V. Torre, Using vanishing points for camera calibration, *Int. J. Comput. Vision* **4** (1990), 127–140.
3. O. D. Faugeras and S. Maybank, Motion from point matches: Multiplicity of solutions, *Int. J. Comput. Vision* **4** (1990), 225–246.
4. W. I. Grosky and L. A. Tamburino, A unified approach to the linear camera calibration problem, *IEEE Trans. Pattern Anal. Mach. Intelligence* **PAMI-12** (1990), 663–671.
5. R. Haralick, Using perspective transformations in scene analysis, *Comput. Graphics Image Process.* **13** (1980), 191–221.
6. S. Ishikawa, H. Kuwamoto, and S. Ozawa, Visual navigation of an autonomous vehicle using white line recognition, *IEEE Trans. Pattern Anal. Mach. Intelligence* **PAMI-10** (1988), 743–749.
7. K. Kanatani, Structure and motion from optical flow under perspective projection, *Comput. Vision Graphics Image Process.* **38** (1987), 122–146.
8. K. Kanatani, Camera rotation invariance of image characteristics, *Comput. Vision Graphics Image Process.* **39** (1987), 328–354.
9. K. Kanatani, Constraints on length and angle, *Comput. Vision Graphics Image Process.* **41** (1988), 28–42.
10. K. Kanatani, Transformation of optical flow by camera rotation, *IEEE Trans. Pattern Anal. Mach. Intelligence* **PAMI-10** (1988), 131–143.
11. K. Kanatani, Reconstruction of consistent shape from inconsistent data: Optimization of 2½D sketches, *Int. J. Comput. Vision* **3** (1989), 261–292.
12. K. Kanatani, *Group-Theoretical Methods in Image Understanding*, Springer-Verlag, Berlin, 1990.
13. K. Kanatani, Hypothesizing and testing geometric properties of image data, *CVGIP: Image Understanding*, **54** (1991) 349–357.
14. K. Kanatani and Y. Onodera, Anatomy of camera calibration using vanishing points, *IEICE Trans. Inf. Syst.* **74** (1991), to appear.
15. K. Kanatani and K. Watanabe, Reconstruction of 3-D road geometry from images for autonomous land vehicles, *IEEE Trans. Robotics Automation* **RA-5** (1990), 127–132.
16. K. Kanatani and K. Watanabe, Road shape reconstruction by local flatness approximation, *Advanced Robotics*, to appear.
17. D. Kuan, G. Phipps, and A.-C. Hsueh, Autonomous robotic vehicle road following, *IEEE Trans. Pattern Anal. Mach. Intelligence* **PAMI-10** (1988), 648–654.
18. R. K. Lenz and R. Y. Tsai, Techniques for calibration of the scale factor and image center for high-accuracy 3-D machine vision metrology, *IEEE Trans. Pattern Anal. Mach. Intelligence* **PAMI-10** (1988), 713–720.
19. R. K. Lenz and R. Y. Tsai, Calibrating a Cartesian robot with eye-on-hand configuration independent of eye-to-hand relationship, *IEEE Trans. Pattern Anal. Mach. Intelligence* **PAMI-11** (1989), 916–928.
20. S.-P. Liou and R. C. Jain, Road following using vanishing points, *Comput. Vision Graphics Image Process* **39** (1987), 116–130.
21. H. C. Longuet-Higgins, The visual ambiguity of a moving plane, *Proc. R. Soc. London B* **223** (1984), 165–175.
22. M. J. Magee and J. K. Aggarwal, Determining vanishing points from perspective images, *Comput. Vision Graphics Image Process* **26** (1984), 256–267.
23. D. G. Morgenthaler, S. Hennessy and D. DeMenthon, Range-video fusion and comparison of inverse perspective algorithms in static images, *IEEE Trans. Syst. Man Cybernet.* **SMC-20** (1990), 1301–1312.
24. P. G. Mulgaonkar, L. G. Shapiro, and R. M. Haralick, Shape from perspective: A rule-based approach, *Comput. Vision Graphics Image Process.* **36** (1986), 298–320.
25. M. Subbarao and A. M. Waxman, Closed form solution to image flow equations for planar surfaces in motion, *Comput. Vision Graphics Image Process.* **36** (1986), 208–228.
26. K. Sugihara, *Machine Interpretation of Line Drawings*, MIT Press, Cambridge, MA, 1986.

27. C. Thorp, M. H. Hebert, T. Kanade, and S. A. Shafer, Vision and navigation of the Carnegie-Mellon Navlab, *IEEE Trans. Pattern Anal. Mach. Intelligence* **PAMI-10** (1988), 362–373.

28. R. Y. Tsai and T. S. Huang, Estimating three-dimensional motion parameters of a rigid planar patch, *IEEE Trans. Acoust. Speech Signal Process.* **ASSP-29** (1981), 1147–1152.

29. R. Y. Tsai, T. S. Huang, and W.-L. Zhu, Estimating three-dimensional motion parameters of a rigid planar patch. II. Singular value decomposition, *IEEE Trans. Acoust. Speech Signal Process.* **ASSP-30** (1982), 525–534.

30. M. A. Turk, D. G. Morgenthaler, K. D. Gremban, and M. Marra, VITS—A vision system for autonomous land vehicle navigation, *IEEE Trans. Pattern Anal. Mach. Intelligence* **PAMI-10** (1988), 342–361.

31. A. M. Waxman, J. LeMoigne, L. S. Davis, B. Srinivasan, T. Kushner, E. Liang, and T. Siddalingaiah, A visual navigation system for autonomous land vehicles, *IEEE J. Robotics Automation* **RA-13** (1987), 124–141.

32. R. Weiss, H. Nakatani, and E. M. Riseman, An error analysis for surface orientation from vanishing points, *IEEE Trans. Pattern Anal. Mach. Intelligence* **PAMI-12** (1990), 1179–1185.