CVGIP: IMAGE UNDERSTANDING
Vol. 60, No. 3, November, pp. 382–391, 1994

NOTE

Statistical Foundation for Hypothesis Testing of Image Data

KENICHI KANATANI*

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

Received November 22, 1991; accepted November 16, 1993

A statistical foundation is given to the problem of hypothesizing and testing geometric properties of image data heuristically derived by Kanatani (*CVGIP: Image Understanding* 54 (1991), 333–348). Points and lines in the image are represented by "N-vectors" and their reliability is evaluated by their "covariance matrices". Under a Gaussian approximation of the distribution, the test takes the form of a χ^2 test. Test criteria are explicitly stated for model matching and testing edge groupings, vanishing points, focuses of expansion, and vanishing lines. © 1994 Academic Press, Inc.

1. INTRODUCTION

Statistical analysis of error behavior is a key to the development of robotics applications of computer vision. Understanding of error behavior often leads to finding techniques for improving accuracy, and even if errors are inevitable, the knowledge of how reliable each computation is is indispensable in guaranteeing performance of the systems that use such computations. Also, statistical reliability estimation is vital if one attempts to enhance the robot performance by using multiple sensors and fusing the resulting multiple data (*sensor fusion*), because in order for multiple data to be fused they must be properly weighted so that reliable data contribute more than unreliable data.

In general, all geometric configurations of points and lines in images ultimately reduce to atomic elements. Hence, the "covariance matrices" of all quantities can be computed by propagating the error behavior *bottom-up* from atomic to primitive to complex configurations. Covariance matrices thus computed indicate the reliability of computation in quantitative terms, and the reliabilities of different computations can be compared on the same basis. A rigorous mathematical foundation has been given to such statistical evaluation by Kanatani [10], who computed covariance matrices of various types of geometric computations from the statistical model of pixels, which he regarded as atoms.

This approach is appropriate *if the true configuration is known*. By this approach, however, one cannot infer the true configuration from a given configuration. However, such inference is indispensable for automatic 3-D image interpretation. For example, a decision must be made as to whether an assumed object model matches the observed image or not, edge segments can be grouped together or not, lines are concurrent or not, points are collinear or not, and so on. In the past, such inference has often been based on arbitrarily set discrepancy measures and thresholds [1, 3, 5, 12–14]. Lacking theoretical grounds, each threshold value must be adjusted to a particular environment empirically.

Kanatani [9] presented a hypothesizing and testing approach to this problem, which is in a sense "dual" to the error estimation process: one considers how a given geometric configuration should be altered into a hypothesized form, then considers how primitives should be altered to achieve that, and so on. In other words, the discrepancy between the observed configuration and its hypothesized form is propagated *top-down* from complex to primitive to atomic quantities. In the end, the "credibility" of a configuration is computed in quantitative terms based on the statistical model of atoms, and the credibilities of different configurations can be compared on the same basis.

Regarding edges as atoms, Kanatani [9] computed *to what extent individual edge segments must be displaced in order to support the hypothesis*. This is a realistic approach because edge detection is usually the first step of image processing for machine vision applications. However, all the testing criteria were derived rather heuristically. As a result, the following issues have remained unsettled:

- Although the geometric meaning of the edge displacement measure introduced in [9] is very clear, other types of measures could be used for the same purpose. Can we justify one particular measure on statistical grounds?
- In [9], estimation and testing were treated separately in the sense that any estimation can be tested by the same procedure. Can we derive an "optimal estimate" based on the hypothesis by theoretical means?
- Although all types of tests were reduced to a single measure of edge displacement with a single threshold, the

*E-mail:kanatani@cs.gunma-u.ac.jp.

1049-9660/94 $6.00

threshold must be empirically adjusted. This is certainly an advance as compared with adjusting problem-dependent thresholds each time, but can we determine the threshold value by a statistical argument?

We now present a theory to answer all these issues. The second issue has already been settled by Kanatani [10]. Here, we focus on the first and third and give them a mathematical solution by employing the statistical analysis of [10]. Adopting the formalism of "computational projective geometry" [8], we represent points and lines in the image by "N-vectors" and evaluate their reliability by their "covariance matrices" based on [10]. Under a Gaussian approximation of the distribution, the test takes the form of a χ^2 test. The test criterion is explicitly stated for model matching and testing edge groupings, vanishing points, focuses of expansion, and vanishing lines; a real image example is given.

Error analysis of vanishing point estimation was discussed by Weiss et al. [16] and Collins and Weiss [4] using heuristically derived error models, and by Brillault-O'Mahony [2] by a rigorous statistical model of line segments expressed in image coordinates. The statistical analysis of Kanatani [10] is similar to that of Brillault-O'Mahony [2] in spirit but is more general, and is more convenient to apply to the hypothesizing and testing approach.

2. PERSPECTIVE PROJECTION AND N-VECTORS

Assume the following camera imaging model [8]. The camera is associated with an XYZ coordinate system with origin O at the center of the lens and Z-axis along the optical axis (Fig. 1). The plane $Z = f$ is identified with the image plane, on which an xy image coordinate system is defined so that the x- and y-axes are parallel to the X- and Y-axes, respectively. Let us call the origin O the *viewpoint* and the constant f the *focal length*.

A point (x,y) on the image plane is represented by the unit vector $\mathbf{m}$ indicating the orientation of the ray starting from the viewpoint O and passing through that point; a line $Ax + By + C = 0$ on the image plane is represented by the unit surface normal $\mathbf{n}$ to the plane passing through the viewpoint O and intersecting the image plane along that line (Fig. 1). Their components are given by

$$\mathbf{m} = \pm N\left[\begin{pmatrix} x \\ y \\ f \end{pmatrix}\right], \quad \mathbf{n} = \pm N\left[\begin{pmatrix} A \\ B \\ C/f \end{pmatrix}\right], \tag{1}$$

where $N[\cdot]$ denotes normalization into a unit vector. We call $\mathbf{m}$ and $\mathbf{n}$ the *N-vectors* of the point and the line [8]. In the following, we adopt the formulation of "computational projective geometry" [8], regarding a unit vector

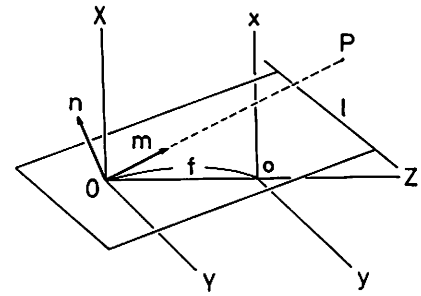

FIG. 1. Camera imaging geometry and N-vectors of a point and a line.

$\mathbf{m}$ whose Z-component is 0 as the N-vector of an *ideal point* (a *point at infinity*) and $\mathbf{n} = (0, 0, \pm 1)$ as the N-vector of the *ideal line* (the *line at infinity*). The following is a brief summary of the facts relevant to this paper.

The N-vector of a point in the scene is defined to be the N-vector of its projection on the image plane. The N-vector of a line in the scene is defined to be the N-vector of its projection on the image plane. In order to avoid the confusion of whether we are referring to a point in the scene or its projection on the image plane, we call a point in the scene a *space point* and a point on the image plane an *image point*. Similarly, we call a line in the scene a *space line* and a line on the image plane an *image line*.

Let $\mathbf{m}$ and $\mathbf{n}$ be the N-vectors of an image point P and an image line l, respectively. It is immediately seen that image point P is *on* image line l, or image line l *passes through* image point P, if and only if

$$(\mathbf{m}, \mathbf{n}) = 0, \tag{2}$$

where $(\cdot, \cdot)$ denotes the inner product of vectors. If this is the case, we say that image point P and image line l are *incident* to each other [8]. We call Eq. (2) the *incidence equation*.

An image point that is on two distinct image lines is called their *intersection;* An image line that passes through two distinct image points is called their *join*. Let $\mathbf{n}_1$ and $\mathbf{n}_2$ be the N-vectors of two distinct image lines. The N-vector $\mathbf{m}$ of their intersection is given by

$$\mathbf{m} = \pm N[\mathbf{n}_1 \times \mathbf{n}_2], \tag{3}$$

because $\mathbf{m}$ must satisfy the incidence equation (2) for both image lines: $(\mathbf{m}, \mathbf{n}_1) = 0$ and $(\mathbf{m}, \mathbf{n}_2) = 0$. Dually, let $\mathbf{m}_1$ and $\mathbf{m}_2$ be the N-vectors of two distinct image points. The N-vector $\mathbf{n}$ of their join is given by

$$\mathbf{n} = \pm N[\mathbf{m}_1 \times \mathbf{m}_2], \tag{4}$$

because $\mathbf{n}$ must satisfy the incidence equation (2) for both image points: $(\mathbf{m}_1, \mathbf{n}) = 0$ and $(\mathbf{m}_2, \mathbf{n}) = 0$.

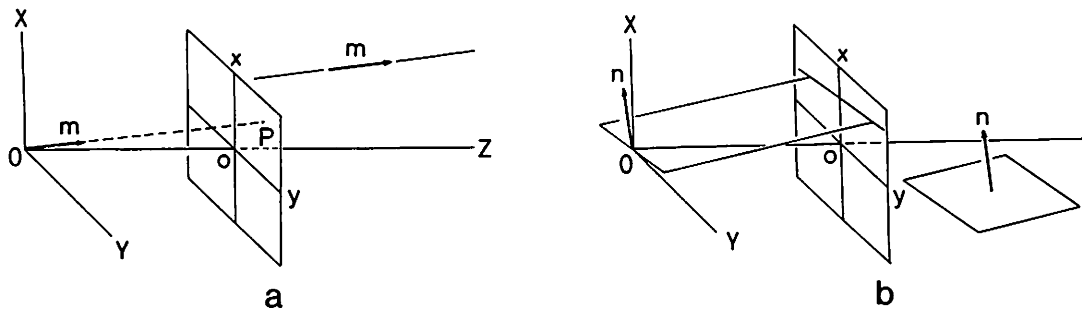

FIG. 2. (a) The vanishing point of a space line. (b) The vanishing line of a planar surface in the scene.

Projections of parallel space lines meet at a common vanishing point on the image plane. Formally, the *vanishing point* of a space line is the limit of the projection of a point that moves along the space line indefinitely in one direction (both directions define the same vanishing point). From Fig. 2a, it is easy to confirm the following theorem [8]:

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

Projections of planar surfaces that are 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 space lines lying on it. From Fig. 2b, it is easy to confirm the following theorem [8]:

THEOREM 2. *A planar surface of unit surface normal* **n** *has, when projected, a vanishing line of N-vector* ±**n**.

3. STATISTICAL MODEL OF EDGE FITTING

In conventional image processing, edges are detected by the Hough transform or an edge operator, and a sequence of edge pixels are obtained after thresholding and thinning processes are applied. Then lines are fitted to them, say, by least squares. According to this scenario, the reliability of subsequent computations reduces to the reliability of line fitting to edges. Thus, we need a statistical model of error behavior for line fitting.

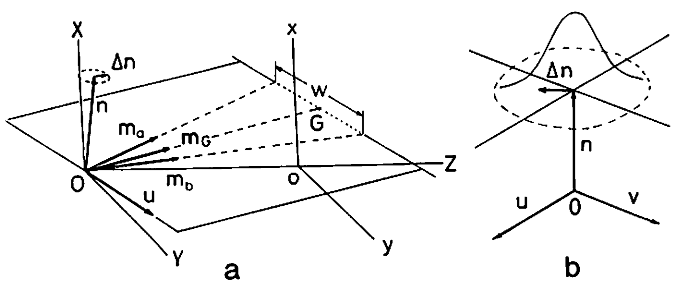

FIG. 3. (a) Line fitting to an edge segment. (b) Gaussian approximation of the distribution of Δm.

Let **n** be the N-vector of the image line fitted to ideal edge pixels with no noise. In the presence of noise, each edge pixel is displaced from its ideal position. Let $\mathbf{n}' = \mathbf{n} + \Delta\mathbf{n}$ be the N-vector of the image line fitted to such displaced edge pixels. Since noise behavior is random, the error $\Delta\mathbf{n}$ is a random-valued vector (Fig. 3a). The *covariance matrix* of **n** is defined by

$$V[\mathbf{n}] = E[\Delta\mathbf{n}\Delta\mathbf{n}^{\mathrm{T}}], \tag{5}$$

where $E[\cdot]$ means expectation and $^{\mathrm{T}}$ denotes transpose [10]. The expression for this covariance matrix is theoretically derived in [10]:

PROPOSITION 1. *The covariance matrix of the N-vector* **n** *of an edge segment of length w in orientation* **u** *is given by*

$$V[\mathbf{n}] \approx \frac{6\kappa}{w^3}\mathbf{u}\mathbf{u}^{\mathrm{T}} + \frac{\kappa}{2f^2 w}\mathbf{m}_G\mathbf{m}_G^{\mathrm{T}}, \tag{6}$$

where $\mathbf{m}_G$ *is the N-vector of the center point of the edge and* κ *is the image resolution.*

Here, the length w is measured in pixels. If $\mathbf{m}_a$ and $\mathbf{m}_b$ are the N-vectors of the end points of the edge segment, the vectors **u** and $\mathbf{m}_G$ are formally defined by

$$\mathbf{u} = \pm N[\mathbf{m}_a - \mathbf{m}_b], \quad \mathbf{m}_G = \pm N[\mathbf{m}_a + \mathbf{m}_b], \tag{7}$$

and the three vectors $\{\mathbf{u}, \mathbf{m}_G, \mathbf{n}\}$ form an orthonormal system of unit eigenvectors of $V[\mathbf{n}]$ for eigenvalues $6\kappa/w^3$, $\kappa/2f^2w$, and 0, respectively [10]. The image resolution κ is defined by

$$\kappa = \frac{\varepsilon^2}{\rho}, \tag{8}$$

where ε is the *image accuracy*, defined as the root mean square of the displacement of each edge pixel, and ρ is the *edge density,* defined as the number of edge pixels per unit pixel length [10].

4. GAUSSIAN APPROXIMATION

Let $V[\mathbf{n}]$ be the covariance matrix of N-vector $\mathbf{n}$. Theoretically, there exist infinitely many distributions that have mean $\mathbf{0}$ and covariance $V[\mathbf{n}]$. Among them, the most natural one is the Gaussian distribution. We assume that the perturbation $\Delta\mathbf{n}$ is sufficiently small as compared with $\mathbf{n}$.

PROPOSITION 2. *If*

$$V[\mathbf{n}] = \sigma_1^2\mathbf{u}\mathbf{u}^{\mathrm{T}} + \sigma_2^2\mathbf{v}\mathbf{v}^{\mathrm{T}}, \quad \sigma_1 \geq \sigma_2 > 0, \tag{9}$$

is the spectral decomposition of the covariance matrix $V[\mathbf{n}]$ *of the N-vector* $\mathbf{n}$*, the Gaussian distribution density of* $\Delta\mathbf{n}$ *with mean* $\mathbf{0}$ *and covariance matrix* $V[\mathbf{n}]$ *is given by*

$$F(\Delta\mathbf{n}) = \frac{1}{2\pi\sigma_1\sigma_2} e^{-(\Delta\mathbf{n}, V[\mathbf{n}]^{-}\Delta\mathbf{n})/2}, \tag{10}$$

where

$$V[\mathbf{n}]^{-} = \frac{\mathbf{u}\mathbf{u}^{\mathrm{T}}}{\sigma_1^2} + \frac{\mathbf{v}\mathbf{v}^{\mathrm{T}}}{\sigma_2^2}. \tag{11}$$

Proof. Since $\mathbf{n}$ is normalized into a unit vector, the perturbation $\Delta\mathbf{n}$ is orthogonal to $\mathbf{n}$ to a first approximation. Hence, if $\Delta\mathbf{n}$ is sufficiently small, it defines a two-dimensional distribution over the plane perpendicular to $\mathbf{n}$ (Fig. 3b). Eq. (9) implies that $\Delta\mathbf{n}$ is most likely to occur in orientation $\mathbf{u}$ and least likely to occur in orientation $\mathbf{v}$. The mean square of projected $\Delta\mathbf{n}$ is σ_1^2 in orientation $\mathbf{u}$, and σ_2^2 in orientation $\mathbf{v}$. Such a Gaussian distribution is given by Eq. (10). ■

The matrix $V[\mathbf{n}]^{-}$ is called the *pseudo-inverse* of $V[\mathbf{n}]$. Given that a particular N-vector $\mathbf{n}' = \mathbf{n} + \Delta\mathbf{n}$ is observed, let us consider whether or not this value can be regarded as a sample from distribution (10). A well known statistical technique is the following χ^2 test. If $\mathbf{n}$ obeys the distribution (10), the quadratic form $(\Delta\mathbf{n}, V[\mathbf{n}]^{-}\Delta\mathbf{n})$ obeys a χ^2 distribution with two degrees of freedom. Hence, we can infer that $\Delta\mathbf{n}$ cannot be regarded as a sample from the distribution (10) with $(100 - a)\%$ confidence if

$$(\Delta\mathbf{n}, V[\mathbf{n}]^{-}\Delta\mathbf{n}) > \chi_{a,2}^2 \tag{12}$$

where $\chi_{a,2}^2$ is the $a\%$ point of the χ^2 distribution with two degrees of freedom. It is easy to show that

$$\chi_{a,2}^2 = -2\log\frac{a}{100}. \tag{13}$$

If we note that $\mathbf{n}$ is the eigenvector of $V[\mathbf{n}]^{-}$ for eigenvalue 0 (hence $V[\mathbf{n}]^{-}\mathbf{n} = \mathbf{0}$), we see that

$$\begin{aligned}(\mathbf{n}', V[\mathbf{n}]^{-}\mathbf{n}') &= (\mathbf{n} + \Delta\mathbf{n}, V[\mathbf{n}]^{-}(\mathbf{n} + \Delta\mathbf{n})) \\ &= (\Delta\mathbf{n}, V[\mathbf{n}]^{-}\Delta\mathbf{n}).\end{aligned} \tag{14}$$

Hence, we obtain the following criterion.

PROPOSITION 3. *If* $V[\mathbf{n}]$ *is the covariance matrix of N-vector* $\mathbf{n}$*, an N-vector* $\mathbf{n}'$ *cannot be regarded as a sample of* $\mathbf{n}$ *with confidence* $(100 - a)\%$ *if*

$$(\mathbf{n}', V[\mathbf{n}]^{-}\mathbf{n}') > \chi_{a,2}^2. \tag{15}$$

A similar criterion was suggested by Collins and Weiss [4] as a Gaussian approximation of the Bingham distribution, which they introduced heuristically. However, our criterion has a distinctive difference from theirs: they replaced the covariance matrix $V[\mathbf{n}]$ by the *sample covariance* computed from a large number of samples. Here, the covariance matrix $V[\mathbf{n}]$ is *not* a sample covariance; it has a *theoretically derived analytical expression* (6). Thus, our criterion can be applied to a *single* sample.

5. MODEL MATCHING

Consider the following problems of model-based object recognition.

- We have multiple candidates of 3-D wireframe models in a database. Given an image, we detect edges and try to match the 3-D models one by one by changing the 3-D position and orientation. Finally, the one that best matches is chosen as the true object in the scene.
- The object in the image is identified but its 3-D position and orientation are not known. So, its 3-D wireframe model is matched to the edge image by changing the parameters of the 3-D position and orientation, and those parameters which yield the best match are determined.

In such processes, it is natural to measure the degree of matching by the discrepancies of the detected edge segments from the supposed line segments of the model (Fig. 4a). In the past, various parameters indicating the discrepancies in position and orientation are heuristically introduced for this purpose [3, 5, 12, 13].

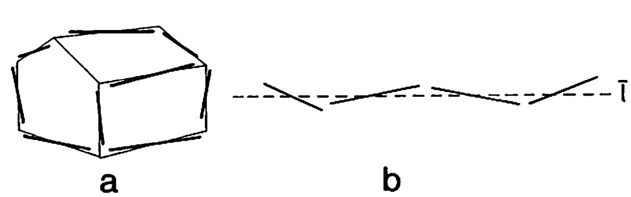

FIG. 4. (a) Model matching for object recognition. (b) How should we group these edge segments together?

Now, Eq. (15) describes to what extent the "line" fitted to an edge segment is likely to deviate from the edge segment, but this can be interpreted to describe to what extent an "edge segment" is likely to deviate from a line on which the edge segment is supposed to lie. Thus, we can decide whether or not an edge segment is unreasonably deviated without introducing any ad hoc parameters. Substituting eq. (6), we obtain the following criterion:

PROPOSITION 4. *Let* $\mathbf{m}_{G\alpha}$, $\mathbf{u}_\alpha$, *and* w_α *be the N-vector of the center point of an edge segment, its orientation, and its length, respectively, and let* $\bar{\mathbf{n}}$ *be the N-vector of a line. If*

$$\frac{w_\alpha^3}{6\kappa}(\bar{\mathbf{n}}, \mathbf{u}_\alpha)^2 + \frac{2f^2 w_\alpha}{\kappa}(\bar{\mathbf{n}}, \mathbf{m}_{G\alpha})^2 > \chi^2_{a,2} \tag{16}$$

is satisfied, the edge segment cannot be regarded as lying on the line with confidence $(100 - a)\%$.

6. TESTING EDGE GROUPINGS

Suppose multiple fragmented edge segments are detected by an edge operator. We want to know if they can be combined together to be fitted by a single line (Fig. 4b). A naive idea is to set thresholds for the discrepancies in the positions and orientations of two consecutive edge segments, replacing them by a single edge segment if the discrepancies are below their thresholds [1, 15]. However, this can cause inconsistencies, as discussed in [9]. The hypothesizing and testing principle was proposed by Kanatani [9], but he did not mention how the estimation should be done optimally or how the test should be done on a statistical basis.

First, consider the hypothesizing stage. It can be shown [10] that the N-vector $\mathbf{n}$ of the line fitted to points of N-vectors $\mathbf{m}_\alpha$, $\alpha = 1, \ldots, N$, is computed optimally as the unit eigenvector of the *moment matrix*

$$\mathbf{M} = \sum_{\alpha=1}^{N} W_\alpha \mathbf{m}_\alpha \mathbf{m}_\alpha^{\mathrm{T}}, \tag{17}$$

where W_α are the *optimal weights* [10] given by

$$W_\alpha = \frac{1}{(\mathbf{n}, V[\mathbf{m}_\alpha]\mathbf{n})}, \tag{18}$$

and $V[\mathbf{m}_\alpha]$ is the covariance matrix of the N-vector $\mathbf{m}_\alpha$ of the αth point. If an edge segment of length w (pixels) is a dense alignment of edge pixels, it can be shown [10] that the moment matrix $\mathbf{M}$ is approximated by

$$\mathbf{M} = \frac{2f^2 w}{\kappa}\mathbf{m}_G \mathbf{m}_G^{\mathrm{T}}, \tag{19}$$

where $\mathbf{m}_G$ is the N-vector of its center point and κ is the image resolution (Fig. 3). Consider N edge segments, and let $\mathbf{m}_{G\alpha}$ be the N-vector of the center point of the αth edge segment, and $\mathbf{u}_\alpha$ its orientation, $\alpha = 1, \ldots, N$. Since the moment matrix $\mathbf{M}_\alpha$ of each edge segment is given in the form of Eq. (19), the total moment matrix M is approximated by

$$M = \sum_{\alpha=1}^{N} \mathbf{M}_\alpha \approx \frac{2f^2}{\kappa}\sum_{\alpha=1}^{N} w_\alpha \mathbf{m}_{G\alpha}\mathbf{m}_{G\alpha}^{\mathrm{T}}. \tag{20}$$

The N-vector $\bar{\mathbf{n}}$ of the line $\bar{l}$ fitted to the N edge segments is given by the unit eigenvector of $\mathbf{M}$ for the smallest eigenvalue. This is equivalent to computing $\bar{\mathbf{n}}$

$$\sum_{\alpha=1}^{N} w_\alpha(\bar{\mathbf{n}}, \mathbf{m}_{G\alpha})^2 \to \min, \tag{21}$$

i.e., fitting a line to the center points of the N edge segments by least squares with lengths w_α as their weights. Thus, the heuristic method suggested in [9] is given a theoretical justification.

Now, consider the second stage of testing. It makes sense to accept the hypothesis that all the edge segments are collinear if the individual edge segments are very close to the line fitted to all the lines. The closeness can be measured by the criterion of Proposition 3. Substituting Eq. (6) and applying the addition theorem of the χ^2 distribution, we obtain the following criterion:

PROPOSITION 5. *Let* $\mathbf{n}_\alpha$, $\mathbf{m}_{G\alpha}$, $\mathbf{u}_\alpha$, *and* w_α *be the N-vector of the* α*th edge segment, the N-vector of its center point, its orientation, and its length, respectively,* $\alpha = 1, \ldots, N$. *If the N-vector* $\bar{\mathbf{n}}$ *of the line fitted to all the edge segments satisfies*

$$\sum_{\alpha=1}^{N}\left(\frac{w_\alpha^3}{6}(\bar{\mathbf{n}}, \mathbf{u}_\alpha)^2 + 2f^2 w_\alpha(\bar{\mathbf{n}}, \mathbf{m}_{G\alpha})^2\right) > \kappa\chi^2_{a,2N}, \tag{22}$$

the N edge segments cannot be regarded as collinear with confidence $(100 - a)\%$.

This criterion is almost the same as the one proposed in [9]. The only difference is that the criterion in [9] includes a threshold to be adjusted empirically while this criterion includes no such threshold: all constants have theoretically well defined values. It is true that the image resolution κ is difficult to estimate for a given image, so in practice we must adjust it empirically, and it may seem that there exists no practical difference. However, there is a significant difference in application potential between simply *assuming* an empirical value and *approximating* a theoretically well defined value empirically.

7. TESTING VANISHING POINTS

Vanishing points provide one of the most important clues to 3-D interpretation, because their N-vectors indicate the 3-D orientations of the corresponding space lines (Theorem 1). Let $\mathbf{n}_\alpha$, $\alpha = 1, \ldots, N$, be the N-vectors of the edge segments. It can be shown [10] that the N-vector $\mathbf{m}$ of the vanishing point is optimally computed as the unit eigenvector of the *moment matrix*

$$\mathbf{N} = \sum_{\alpha=1}^{N} W_\alpha \mathbf{n}_\alpha \mathbf{n}_\alpha^{\mathrm{T}}, \tag{23}$$

where W_α are the *optimal weights* [10] given by

$$W_\alpha = \frac{1}{(\mathbf{m}, V[\mathbf{n}_\alpha]\mathbf{m})}, \tag{24}$$

and $V[\mathbf{n}_\alpha]$ is the covariance matrix of the N-vector $\mathbf{n}_\alpha$ of the αth edge segment. The covariance matrix of the N-vector $\mathbf{m}$ of thus estimated vanishing point can be theoretically computed [10].

There remains, however, one crucial problem. The vanishing point is computed on the assumption that *the edge segments are projections of parallel space lines*. In reality, how can we tell which edge segments are projections of parallel space lines? A well known heuristic is to assume that edge segments are projections of parallel space lines if they are concurrent, i.e., meeting at a common intersection when extended [6, 7, 14]. Theoretically, nonparallel space lines could be projected to concurrent edge segments, but such an exceptional coincidence cannot be expected in general.

Still, a problem remains. Since edge segments are detected by processing real images, they are bound to contain errors. As a result, projections of parallel space lines may not be concurrent (Fig. 5a). How can we decide whether or not such edge segments are concurrent? A naive idea is to set a threshold and regarding them as concurrent if the mutual discrepancies are below the threshold [6, 7]. However, the threshold cannot be fixed, since small displacements of the individual edge segments can cause large deviations of their intersections as discussed in [9].

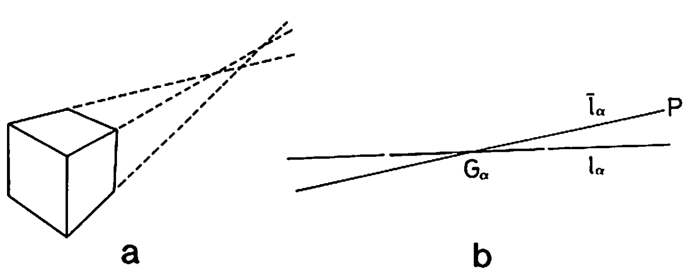

FIG. 5. (a) How can we judge the concurrency of edge segments? (b) Testing the concurrency hypothesis.

Kanatani [9] introduced the hypothesizing and testing principle to this problem by a heuristic method. We now modify his method into the form of a statistical test as follows. We first hypothesize that given edge segments are projections of parallel space lines, and fit a line to each of them optimally by the method mentioned in the preceding section. Then, we optimally estimate their vanishing point as stated above. Let P be the estimated vanishing point, and $\mathbf{m}$ its N-vector. We accept the hypothesis if the individual edge segments are very close to the lines passing through the supposed vanishing point. Let G_α be the center point of the αth edge segment (Fig. 5b). The N-vector of the line $\bar{l}_\alpha$ passing through P and G_α is

$$\bar{\mathbf{n}}_\alpha = \pm N[\mathbf{m} \times \mathbf{m}_{G\alpha}] = \pm \frac{\mathbf{m} \times \mathbf{m}_{G_\alpha}}{1 - (\mathbf{m}, \mathbf{m}_{G\alpha})^2}. \tag{25}$$

According to Proposition 3, if $V[\mathbf{n}_\alpha]$ is the covariance matrix of the N-vector $\mathbf{n}_\alpha$ of the αth line $\bar{l}_\alpha$, the discrepancy of line $\bar{l}_\alpha$ from line l_α is measured by $(\bar{\mathbf{n}}_\alpha, V[\mathbf{n}_\alpha]^- \bar{\mathbf{n}}_a)$. The covariance matrix $V[\mathbf{n}_\alpha]$ of the αth edge segment is given by Eq. (6). Since the second term on the right-hand side is negligible as compared with the first term when computing the vanishing point [10], we obtain the approximation

$$(\bar{\mathbf{n}}_\alpha, V[\mathbf{n}_\alpha]^- \bar{\mathbf{n}}_\alpha) \approx \frac{w_\alpha^3}{6\kappa}(\bar{\mathbf{n}}_\alpha, u_\alpha)^2, \tag{26}$$

where w_α is the length of the αth edge segment. Note that line $\bar{l}$ passes through G_α, so $(\bar{\mathbf{n}}_\alpha, \mathbf{m}_{G\alpha}) = 0$. Since $\mathbf{u}_\alpha = \pm n_\alpha \times \mathbf{m}_{G\alpha}$, we have

$$(\bar{\mathbf{n}}_\alpha, \mathbf{u}_\alpha)^2 = (\bar{\mathbf{n}}_\alpha, \mathbf{n}_\alpha \times \mathbf{m}_{G\alpha})^2 = (\bar{\mathbf{n}}_\alpha \times \mathbf{n}_\alpha, \mathbf{m}_{G\alpha})^2. \tag{27}$$

Since the error $\mathbf{n}_\alpha - \bar{\mathbf{n}}_\alpha$ is most likely to occur in orientation $\mathbf{u}_\alpha$, which is orthogonal to both $\mathbf{n}_\alpha$ and $\mathbf{m}_G\alpha$, vector $\bar{\mathbf{n}}_\alpha \times \mathbf{n}_\alpha$ is nearly parallel to $\mathbf{m}_{G\alpha}$. Hence,

$$(\bar{\mathbf{n}}_\alpha \times \mathbf{n}_\alpha, \mathbf{m}_{G\alpha})^2 \approx \|\bar{\mathbf{n}}_\alpha \times \mathbf{n}_\alpha\|^2 = 1 - (\bar{\mathbf{n}}_\alpha, \mathbf{n}_\alpha)^2. \tag{28}$$

Substituting Eq. (25) into this and applying the addition theorem for the χ^2 distribution, we obtain from Eq. (26) and Proposition 3 the following criterion:

PROPOSITION 6. *Let* $\mathbf{n}_\alpha$, $\mathbf{m}_{G\alpha}$, *and* w_α *be the N-vector of the* α*th edge segment, the N-vector of its center point, and its length, respectively,* $\alpha = 1, \ldots, N$. *A point of N-vector* $\mathbf{m}$ *cannot be regarded as their vanishing point with confidence* $(100 - a)\%$ *if*

$$\sum_{\alpha=1}^{N} w_\alpha^3 \left(1 - \frac{|\mathbf{m}, \mathbf{n}_\alpha, \mathbf{m}_{G\alpha}|^2}{1 - (\mathbf{m}, \mathbf{m}_{G\alpha})^2}\right) > 6\kappa\chi^2_{a,2N}. \quad (29)$$

Again, this criterion is almost the same as the one proposed in [9]—the difference is that the criterion in [9] includes a threshold to be adjusted empirically while this criterion includes no such threshold. As for edge groupings, this is a significant advantage even if we are obliged to empirically adjust the image resolution κ.

8. TESTING FOCUSES OF EXPANSION

If space points are rigidly translating in the scene (or the camera is translating relative to them), their trajectories are parallel in the scene, defining a common intersection on the image plane—known as the *focus of expansion* ([8]; Fig. 6). Since the geometric relationship is the same as for vanishing points, focuses of expansion also provide important clues to 3-D motion: their N-vectors indicate the 3-D orientations of the corresponding 3-D translations (see Theorem 1).

The procedure for testing vanishing points described in the preceding section can also be applied to testing focuses of expansion. This was not mentioned in [9] because all the tests in [9] were based on "edge" displacements, while focuses of expansion are not defined in terms of edges: they are the common intersections of the "trajectories" connecting corresponding feature points in two images taken at different time instants.

Consider a trajectory segment passing through two corresponding feature points. Since the two feature points belong to different images, the error behaviors can be assumed to be independent. Then, we can theoretically derive the covariance matrix of the N-vector n of the resulting trajectory [10].

PROPOSITION 7. *The covariance matrix of the N-vector* $\mathbf{n}$ *of a trajectory of orientation* $\mathbf{u}$ *is given by*

$$V[\mathbf{n}] = \frac{w^2}{\varepsilon^2}\mathbf{u}\mathbf{u}^\mathrm{T} + \frac{\varepsilon^2}{4f^2 - w^2}\mathbf{m}_G\mathbf{m}_G^\mathrm{T}, \quad (30)$$

where w *is the distance between the two feature points that define the trajectory, and* $\mathbf{m}_G$ *the N-vector of their center points.*

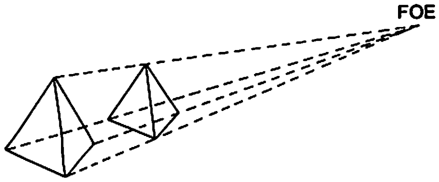

FIG. 6. Focus of expansion.

If $\mathbf{m}_a$ and $\mathbf{m}_b$ are the N-vectors of the end points of the trajectory, the vectors $\mathbf{u}$ and $\mathbf{m}_G$ are formally defined by Eqs. (7).

Let $\mathbf{n}_\alpha$, $\alpha = 1, \ldots, N$, be the N-vectors of N trajectories on the image plane defined by translating space points. It can be shown [10] that the N-vector $\mathbf{m}$ of their focus of expansion is optimally computed as the unit eigenvector of the moment matrix defined by Eq. (23), where the optimal weights are given by Eq. (24) if $V[\mathbf{n}_\alpha]$ is regarded as the covariance matrix of the N-vector of the αth trajectory. The covariance matrix of the N-vector $\mathbf{m}$ of the thus estimated vanishing point can be theoretically computed [10].

Again, a crucial problem is how to find pairs of corresponding feature points. As in the case of vanishing points, we first hypothesize that given pairs of points are projections of translating space points, and optimally estimate their focus of expansion as stated above. Then, we accept the hypothesis if individual pairs are close to the trajectories passing through the estimated focus of expansion.

Let $\mathbf{m}$ be the N-vector of the estimated focus of expansion P. Let $\mathbf{m}_{G\alpha}$ be the N-vector of the center point G_α of the αth pair. The N-vector of the line $\bar{l}_\alpha$ passing through P and G_α is

$$\bar{\mathbf{n}}_\alpha = \pm N[\mathbf{m} \times \mathbf{m}_{G\alpha}] = \pm\frac{\mathbf{m} \times \mathbf{m}_{G\alpha}}{1 - (\mathbf{m}, \mathbf{m}_{G\alpha})^2}. \quad (31)$$

The covariance matrix $V[\mathbf{n}_\alpha]$ of the N-vector of the αth trajectory is given by Eq. (30), and the second term on the right-side hand is negligible as compared with the first term when computing the focus of expansion [10]. According to Proposition 3, the discrepancy of line l_α from line $\bar{l}_\alpha$ is measured by

$$(\bar{\mathbf{n}}_\alpha, V[\mathbf{n}_\alpha]^- \bar{\mathbf{n}}_\alpha) \approx \left(\frac{w_\alpha}{\varepsilon}\right)^2 (\bar{\mathbf{n}}_\alpha, \mathbf{u}_\alpha)^2. \quad (32)$$

Note that line $\bar{l}_\alpha$ passes through point G_α, so $(\bar{\mathbf{n}}_\alpha, \mathbf{m}_{G\alpha}) = 0$. As for testing vanishing points, we have the approximation

$$(\bar{\mathbf{n}}_\alpha, \mathbf{u}_\alpha)^2 = (\bar{\mathbf{n}}_\alpha, \mathbf{n}_\alpha \times \mathbf{m}_{G\alpha})^2 = (\bar{\mathbf{n}}_\alpha \times \mathbf{n}_\alpha, \mathbf{m}_{G\alpha})^2 \approx \|\bar{\mathbf{n}}_\alpha \times \mathbf{n}_\alpha\|^2 = 1 - (\bar{\mathbf{n}}_\alpha, \mathbf{n}_\alpha)^2. \quad (33)$$

Substituting Eq. (31) into this and applying the addition theorem for the χ^2 distribution, we obtain from Eq. (32) and Proposition 3 the following criterion:

PROPOSITION 8. *Let* $\mathbf{n}_\alpha$, $\mathbf{m}_{G\alpha}$, *and* w_α *be the N-vector of the trajectory passing through the* α*th pair, the N-vector of its center point, and the distance between the two feature points, respectively,* $\alpha = 1, \ldots, N$. *A point*

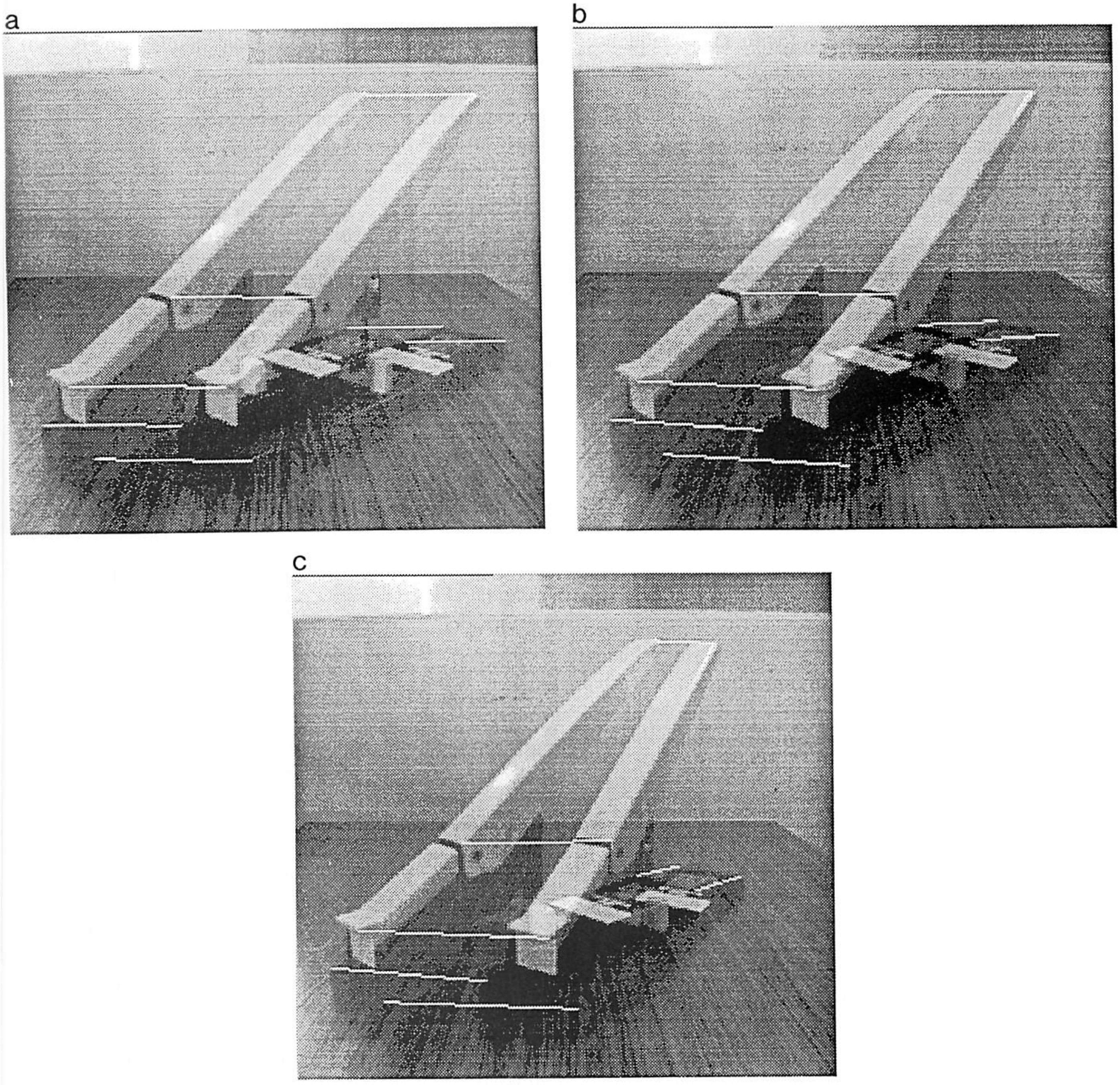

FIG. 7. Superimposed images of a translating stapler. (a) A pure translation. (b) A translation and a small rotation. (c) A translation and a large rotation.

of the N-vector **m** *cannot be regarded as their focus of expansion with confidence* $(100 - a)\%$ *if*

$$\sum_{\alpha=1}^{N} w_\alpha^2 \left(1 - \frac{|\mathbf{m}, \mathbf{n}_\alpha, \mathbf{m}_{G\alpha}|^2}{1 - (\mathbf{m}, \mathbf{m}_{G\alpha})^2}\right) > \varepsilon^2 \chi^2_{a,2N}. \tag{34}$$

EXAMPLE. Figure 7 shows superimpositions of two 512×512-pixel real images, in which a stapler undergoes (a) a pure translation, (b) a translation and a small rotation, and (c) a translation and a large rotation. The focal length is estimated to be $f = 600$ (pixels). Seven feature points are chosen, and their trajectories are defined by connecting the corresponding positions. We hypothesize that all the trajectories are concurrent. The validity of this hypothesis depends on the image accuracy ε (in pixels) with which the feature points are detected and the significance level of the test. If $\varepsilon \leq 2.5$, the hypothesis is accepted for (a) but rejected for (b) and (c) with 95% confidence. If $2.5 < \varepsilon \leq 7$, the hypothesis is accepted for (a) and (b) but rejected for (c) with 95% confidence. If $\varepsilon > 7$, the hypothesis is accepted for (a), (b), and (c) with 95% confidence.

9. TESTING VANISHING LINES

Let $\{P_\alpha\}$, $\alpha = 1, ..., N$, be the vanishing points of N sets of parallel space lines. If the N space lines are all horizontal, the vanishing points $\{P_\alpha\}$ are on the "horizon," i.e., the vanishing line of the horizontal ground in the scene. Conversely, we can apply the heuristic that space lines are parallel to a planar surface in the scene if their vanishing points are collinear on the image plane. In the presence of noise, however, supposedly collinear vanishing points may not be strictly collinear (Fig. 8a). Moreover, since vanishing points are obtained by computation, they can appear in any locations on the image plane—even at infinity. How can we test for collinearity of the vanishing points?

The idea of Kanatani [9] can be restated on the present statistical basis as follows. As in the cases of edge grouping, vanishing points, and focuses of expansion, we first hypothesize that given vanishing points $\{P_\alpha\}$ are collinear, and optimally fit a line $\bar{l}$ by the optimal method described in Section 4. The covariance matrix of each vanishing point is given analytically [10]. Let $\bar{\mathbf{n}}$ be its N-vector. For each P_α, we compute the point $\bar{P}_\alpha$ that is closest to P_α on the fitted line $\bar{l}$ (Fig. 8b). If $\mathbf{m}_\alpha$ is the N-vector of P_α, it can be shown [9] that the N-vector $\bar{\mathbf{m}}_\alpha$ of $\bar{P}_\alpha$ is given by

$$\bar{\mathbf{m}}_\alpha = \pm N\left[\mathbf{k} - (\bar{\mathbf{n}}, \mathbf{k})\bar{\mathbf{n}} - \frac{|\bar{\mathbf{n}}, \mathbf{m}_\alpha, \mathbf{k}|}{(\mathbf{m}_\alpha, \mathbf{k})}\bar{\mathbf{n}} \times \mathbf{k}\right], \quad (35)$$

where $\mathbf{k} = (0, 0, 1)^T$. Points $\bar{P}_\alpha$ are regarded as the "correct" positions of the vanishing points.

Suppose vanishing points $\{P_\alpha\}$ are detected as intersections of concurrent image lines fitted to edge segments. Consider the lines passing through the correct vanishing point $\bar{P}_\alpha$ and the center points of the individual edge segments that define the vanishing point P_α (Fig. 8b). Applying the testing criterion for vanishing points, we can test our hypothesis by checking the discrepancies of individual edge segments from the supposedly correct lines. From Proposition 8, we obtain the following criterion. Let $\mathbf{k} = (0, 0, 1)^T$.

PROPOSITION 9. *Let $\{\mathbf{m}_\alpha\}$, $\alpha = 1, ..., N$, be the N-vectors of the vanishing points to be tested. Let $\mathbf{n}_\beta^{(\alpha)}$, $\mathbf{m}_{G\beta}^{(\alpha)}$, and $w_\beta^{(\alpha)}$ be the N-vector of the βth edge segment that defines the αth vanishing point, the N-vector of its center point, and its length, respectively, $\beta = 1, ..., N^{(\alpha)}$. The vanishing points cannot be regarded as collinear with confidence* $(100 - a)\%$ if

$$\sum_{\alpha=1}^{N}\sum_{\beta=1}^{N^{(\alpha)}} w_\beta^{(\alpha)3}\left(1 - \frac{|\bar{\mathbf{m}}_\alpha, \mathbf{n}_\beta^{(\alpha)}, \mathbf{m}_{G\beta}^{(\alpha)}|^2}{1 - (\bar{\mathbf{m}}_\alpha, \mathbf{m}_{G\beta}^{(\alpha)})^2}\right) > 6\kappa\chi^2_{a,2n}, \quad (36)$$

where $n = \sum_{\alpha=1}^{N} N^{(\alpha)}$ *and*

$$\bar{\mathbf{m}}_\alpha = \pm N\left[\mathbf{k} - (\bar{\mathbf{n}}, \mathbf{k})\bar{\mathbf{n}} - \frac{|\bar{\mathbf{n}}, \mathbf{m}_\alpha, \mathbf{k}|}{(\mathbf{m}_\alpha, \mathbf{k})}\bar{\mathbf{n}} \times \mathbf{k}\right]. \quad (37)$$

10. CONCLUDING REMARKS

A statistical foundation has been given for the process of hypothesizing and testing geometric properties of image data derived by Kanatani [9]. We represented points and lines in the image by "N-vectors" by adopting the formalism of "computational projective geometry" of Kanatani [8], and evaluated the reliability of computation by "covariance matrices" of N-vectors by applying the statistical analysis of Kanatani [10]. Under a Gaussian approximation of the distribution, the test takes the form of a χ^2 test. Test criteria were explicitly stated for model matching, testing edge groupings, vanishing points, focuses of expansion, and vanishing lines. For these problems, it has been customary to introduce ad hoc parameters to be thresholded, but the criteria given here do not involve any ad hoc parameters; they are built on a rigorous statistical basis.

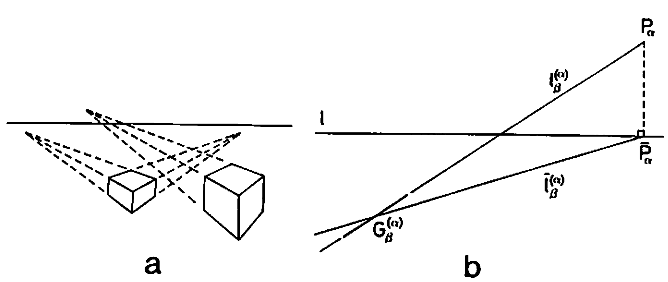

FIG. 8. (a) How can we judge the collinearity of vanishing points? (b) Testing the collinearity hypothesis.

ACKNOWLEDGMENTS

The idea of building a rigorous foundation of geometric hypothesis testing based on a statistical error model was inspired by the criticism of Dr. Bernard Buxton of GEC (U.K.) on the author's previous work [9]. This led the author to establish the comprehensive statistical theory of geometric computation [10]. The choice of testing criteria underwent heated discussions with the author at the University of Oxford (U.K.), the University of Maryland (U.S.A.), and many conferences and meetings. The author thanks all of those who participated in these discussions.

REFERENCES

1. K. M. Andress and A. C. Kak, Evidence accumulation and flow of control in a hierarchical spatial reasoning system, *AI Mag.* 9 (1988), 75–94.
2. B. Brillault-O'Mahony, New method for vanishing point detection, *CVGIP: Image Understanding* **54** (1991), 289–300.
3. R. A. Brooks, *Model-Based Computer Vision*, UMI Res. Press, Ann Arbor, MI, 1984.
4. R. T. Collins and R. S. Weiss, Vanishing point calculation as a statistical data, in *Proceedings, 3rd Int. Conf. Comput. Vision, Osaka, 1990*, pp. 400–403.
5. L. Du, G. D. Sullivan, and K. D. Baker, 3D grouping by viewpoint consistency ascent, *Image Vision Comput.* **10** (1992), 301–307.
6. K. Kanatani, Reconstruction of consistent shape from inconsistent data: Optimization of 2½D sketches, *Int. J. Comput. Vision* 3 (1989), 261–292.
7. K. Kanatani, *Group-Theoretical Methods in Image Understanding*, Springer, Berlin, 1990.
8. K. Kanatani, Computational projective geometry, *CVGIP: Image Understanding* **54** (1991), 333–348.
9. K. Kanatani, Hypothesizing and testing geometric properties of image data, *CVGIP: Image Understanding* **54** (1991), 349–357.
10. K. Kanatani, Statistical analysis of geometric computation, *CVGIP: Image Understanding* **59** (1994), 286–306.
11. K. Kanatani, *Geometric Computation for Machine Vision*, Oxford Univ. Press, Oxford, 1993.
12. D. G. Lowe, *Perceptual Organization and Visual Recognition*, Kluwer Academic, Boston, 1985.
13. D. G. Lowe, Fitting parameterized three-dimensional models to images, *IEEE Trans. Pattern Anal. Mach. Intell.* **13** (1991), 441–450.
14. 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.
15. R. Weiss and M. Boldt, Geometric grouping applied to straight lines, *Proceedings IEEE Conf. Computer Vision Pattern Recognition, Miami Beach, 1986*, pp. 656–661.
16. R. Weiss, H. Nakatani, and E. M. Riseman, An error analysis for surface orientation from vanishing points, *IEEE Trans. Pattern Anal. Mach. Intell.* **12** (1990), 1179–1185.