CVGIP: IMAGE UNDERSTANDING
Vol. 59, No. 3, May, pp. 286–306, 1994

Statistical Analysis of Geometric Computation

KENICHI KANATANI*

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

Received November 12, 1991; accepted October 12, 1993

This paper studies the statistical behavior of errors involved in fundamental geometric computations. We first present a statistical model of noise in terms of the *covariance matrix* of the N-vector. Using this model, we compute the covariance matrices of N-vectors of lines and their intersections. Then, we determine the *optimal weights* for the least-squares optimization and compute the covariance matrix of the resulting optimal estimate. The result is then applied to line fitting to edges and computation of vanishing points and focuses of expansion. We also point out that *statistical biases* exist in such computations and present a scheme called *renormalization*, which iteratively removes the bias by automatically adjusting to noise without knowing noise characteristics. Random number simulations are conducted to confirm our analysis.

1. INTRODUCTION

Traditionally, the research of computer vision has aimed at "recognizing" objects—detecting prominent features in the scene, separating objects from the background, inferring their 3D shapes, identifying or classifying them, and finally obtaining symbolic descriptions of them. Possible applications of this approach include automatic surveillance for airport traffic control, road traffic control, and automatic security inspection. Recently, however, the rapid progress of robotics technology has created a new goal, "sensing" of the environment—locating objects in the scene, inferring their 3D positions and motions, and computing the 3D motion of the camera. Possible applications of this approach include automatic navigation and robot operations in industrial environments, hazardous environments, and outer space.

Statistical analysis of error behavior is a key to the development of such robotics applications, since the reliability of computation is of the utmost importance in such applications. Understanding error behavior often leads to new 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 becomes vital when using multiple sensors and fusing the data (*sensor fusion*), because overall reliability does not increase unless individual data are properly weighted so that reliable data contribute more than unreliable data.

Statistical approaches to image processing and computer vision problems are not new (e.g., [8–10, 24, 27]), and many techniques for improving accuracy have been proposed. The best known is the *Kalman filter*, which is essentially *linearized iterative optimization:* each time new data is added, the solution is modified linearly under Gaussian approximation so that it becomes optimal over the data so far observed. The Kalman filter was originally devised for linear dynamic systems, but its various variations and related ideas have been applied to many types of computer vision problems involving a sequence of image data (e.g., [3–5, 20–22, 26]). However, the main emphasis in such studies is the "techniques" for computing robust and accurate solutions, and not as much attention has been paid to systematic analysis of error behavior.

In this paper, we first present a statistical model of noise in terms of the covariance matrix of the *N-vector* of a point datum. Using this model, we compute the covariance matrices of N-vectors of lines and their intersections. Then, we determine the *optimal weights* for the least-squares optimization and compute the covariance matrix of the resulting optimal estimate. The result is applied to line fitting to edges and computation of vanishing points and focuses of expansion. We also point out that *statistical bias* exists in the optimization computations and present a scheme called *renormalization*, which iteratively removes the bias by automatically adjusting to noise without knowing noise characteristics. Random number simulations are conducted to confirm our analysis.

Error analysis of vanishing points was given by Weiss *et al.* [25], Collins and Weiss [6], and Kanatani [12]. Brillault-O'Mahony [2] presented a detailed analysis of vanishing points by introducing a detailed statistical model of line segments. Errors in focuses of expansion and translational motions were meticulously analyzed by Snyder [23]. Er-

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

1049-9660/94 $6.00

ror analysis was also done for optical flow detection [19] and stereo [1, 7].

In general terms, error analysis means estimating error magnitudes whether in the statistical sense (i.e., means and standard deviations) or in the absolute sense (i.e., worst case bounds). It consists of a "model of errors" that idealizes the error sources and "inference techniques," which should be treated separately from the model of errors. In many existing error analyses, such as those cited above, the distinction is not clear, and the model of errors and the inference techniques are based on the application problem in question. As a result, the error behavior of that problem may be explained well, but it is often very difficult to extend the results to other applications.

The aim of this paper is to give an *abstraction* of error analysis in an *application independent* formalism, although its usefulness is illustrated by applying it to particular (but fundamental) problems. The approach in this paper is to some extent similar to that of Brillault-O'Mahony [2], but our formulation is more general. Based on the general (and also application independent) computational theory of Kanatani [11], which he called *computational projective geometry*, our approach can be applied to a much wider range of computer vision problems [12–15, 17, 18].

2. STATISTICAL MODEL OF NOISE

2.1. *N-Vectors of Points and Lines*

Assume the following camera imaging model [11]. The camera is associated with an XYZ coordinate system with origin O at the center of the lens and the 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. We 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 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

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

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 [11]. If $\mathbf{m}$ and $\mathbf{n}$ are the N-vectors of a point P and a line l, respectively, point P is *on* line l, or line l *passes through* point P, if and only if

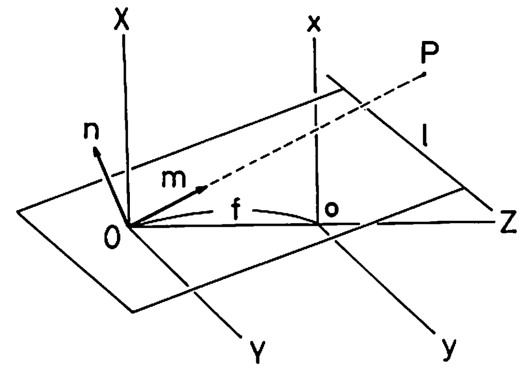

FIG. 1. Camera imaging geometry and N-vectors of a point and a line.

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

where $(\cdot, \cdot)$ denotes the inner product of vectors. If this equation is satisfied, we say that point P and line l are *incident* to each other [11] and call Eq. (2) the *incidence equation*.

A point that is on two distinct lines is called their *intersection;* a line that passes through two distinct points is called their *join*. Let $\mathbf{n}_1$ and $\mathbf{n}_2$ be the N-vectors of two distinct 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 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 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 lines: $(\mathbf{m}_1, \mathbf{n}) = 0$ and $(\mathbf{m}_2, \mathbf{n}) = 0$.

The use of N-vectors for representing points and lines on the image plane is equivalent to using *homogeneous coordinates* known in projective geometry. Although homogeneous coordinates can be multiplied by any nonzero number, computational problems arise if they are too large or too small. So, it is more convenient to normalize the three components into a unit vector, which is precisely the N-vector as defined above. Kanatani [11] reformulated projective geometry from this viewpoint. Rewriting the relationships of projective geometry as "computational procedures," he called his formalism *computational projective geometry*. In this paper, we adopt his formalism, regarding a unit vector $\mathbf{m}$ whose Z-component is 0 as the N-vector of an *ideal point* (a *point at infinity*) and $\mathbf{n} = (0,$

0, ±1) as the N-vector of the *ideal line* (the *line at infinity*). For the details, see [11, 16].

2.2. *Covariance Matrix of an N-Vector*

We now study the effect of noise. In the following, we extend the meaning of the term "noise." A digital image consists of discrete pixels, and the noise in the strict sense affects the electric signal that carries information about the gray levels of the pixels. The signal is then quantized and stored as image data. As a result, point and line data detected by applying image operations are not accurate. Consequently, N-vectors computed from them are not exact. Here, we regard such errors as being caused by "noise." This means that the noise behavior is characterized not only by the camera and the memory frame system but also by the image operations involved—edge operators, thinning algorithms, etc.

Let $\mathbf{m}$ be the N-vector of a point on the image plane when there is no noise. In the presence of noise, a perturbed N-vector $\mathbf{m}' = \mathbf{m} + \Delta\mathbf{m}$ is observed. We regard the "noise" $\Delta\mathbf{m}$ as a random variable. Namely, each observation is regarded as a "sample" from a "statistical ensemble." If the noise $\Delta\mathbf{m}$ is sufficiently small compared with $\mathbf{m}$ itself, the error behavior is characterized by the *covariance matrix*

$$V[\mathbf{m}] = E[\Delta\mathbf{m}\Delta\mathbf{m}^T], \qquad (5)$$

where T denotes transpose and $E[.]$ denotes the expectation over the statistical ensemble. Since $\mathbf{m}$ is a unit vector, the noise $\Delta\mathbf{m}$ is always orthogonal to $\mathbf{m}$ to a first approximation; $\mathbf{m}' = \mathbf{m} + \Delta\mathbf{m}$ is again a unit vector to a first approximation. We immediately observe the following:

- The covariance matrix $V[\mathbf{m}]$ is symmetric and positive semi-definite.
- The covariance matrix $V[\mathbf{m}]$ is singular with $\mathbf{m}$ itself as the unit eigenvector for eigenvalue 0: $V[\mathbf{m}]\mathbf{m} = \mathbf{0}$.
- If σ_1^2, σ_2^2, and 0 ($\sigma_1 \geq \sigma_2 > 0$) are the three eigenvalues and if $\{\mathbf{u}, \mathbf{v}, \mathbf{m}\}$ is the corresponding orthonormal system of eigenvectors, the covariance matrix $V[\mathbf{m}]$ is expressed in the "spectral decomposition" [16]:

$$V[\mathbf{m}] = \sigma_1^2\mathbf{u}\mathbf{u}^T + \sigma_2^2\mathbf{v}\mathbf{v}^T(+0\mathbf{m}\mathbf{m}^T) \qquad (6)$$

- The root mean square of the orthogonal projection of the noise $\Delta\mathbf{m}$ onto orientation $\mathbf{l}$ (unit vector) takes its maximum for $\mathbf{l} = \mathbf{u}$ and its minimum for $\mathbf{l} = \mathbf{v}$. The maximum and minimum values are σ_1 and σ_2, respectively.
- The root-mean-square magnitude of $\Delta\mathbf{m}$ is $\sqrt{\mathrm{tr}\, V[\mathbf{m}]} = \sqrt{\sigma_1^2 + \sigma_2^2}$ ("tr" denotes trace).

In intuitive terms, noise $\Delta\mathbf{m}$ is most likely to occur in orientation $\mathbf{u}$ (= the unit eigenvector of $V[\mathbf{m}]$ for the largest eigenvalue σ_1^2) and least likely to occur in orientation $\mathbf{v}$ (= the unit eigenvector of $V[\mathbf{m}]$ for the second largest eigenvalue σ_2^2). The magnitude $\|\Delta\mathbf{m}\|$ is σ_1 in orientation $\mathbf{u}$ and σ_2 in orientation $\mathbf{v}$ in the sense of root mean square (Fig. 2a).

2.3. *Model of Noise*

We assume that noise (in our extended sense) occurs at each point on the image plane independently with an identical distribution density. Let s_x and s_y be the standard deviations in the x and y orientations, respectively, and s_{xy} the covariance. From Appendix A, we obtain

THEOREM 1. *The covariance matrix $V[\mathbf{m}]$ of the N-vector of a point P on the image plain is given by*

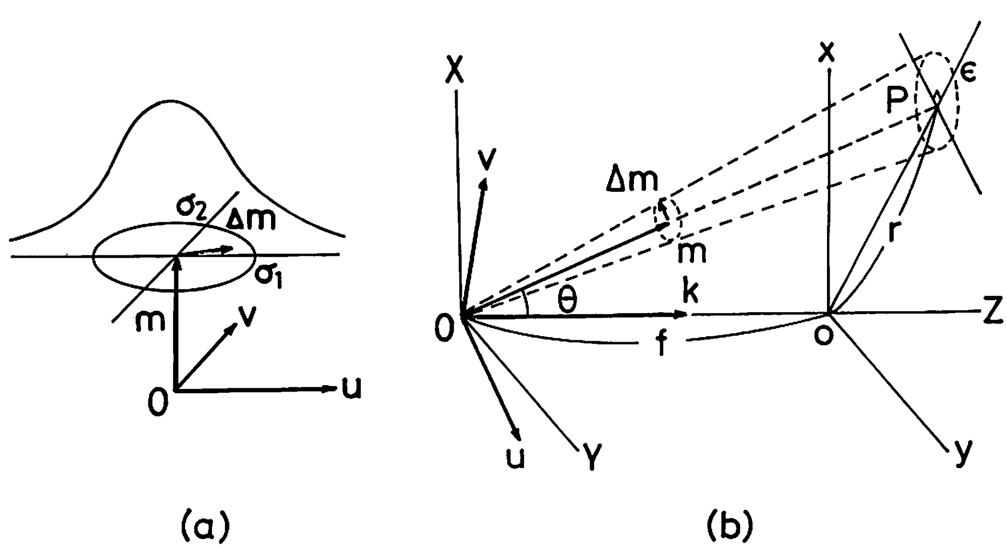

FIG. 2. (a) Perturbation of N-vector m and (b) a model of noise.

$$V[\mathbf{m}] = \frac{1}{|OP|^2} \boldsymbol{P}_\mathrm{m} \begin{pmatrix} s_x^2 & s_{xy} & 0 \\ s_{xy} & s_y^2 & 0 \\ 0 & 0 & 0 \end{pmatrix} \boldsymbol{P}_\mathrm{m}, \tag{7}$$

where $\boldsymbol{P}_\mathrm{m} = \boldsymbol{I} - \mathbf{m}\mathbf{m}^\mathrm{T}$ *is the projection matrix along* $\mathbf{m}$.

Consider the case where the noise distribution is isotropic. This means that $s_x = s_y$ and $s_{xy} = 0$. If we put the root-mean-square magnitude of noise to be

$$\varepsilon = \sqrt{s_x^2 + s_y^2}, \tag{8}$$

then $s_x = s_y = \varepsilon/\sqrt{2}$. Let us call ε (measured in pixels) the *image accuracy*. We obtain the following.

PROPOSITION 1. *If the noise is isotropic, the covariance matrix* $V[\mathbf{m}]$ *of the N-vector of a point at distance r from the image origin is given by*

$$V[\mathbf{m}] = \frac{\varepsilon^2/f^2}{2(1 + r^2/f^2)} \left(\mathbf{u}\mathbf{u}^\mathrm{T} + \frac{\mathbf{v}\mathbf{v}^\mathrm{T}}{1 + r^2/f^2} \right), \tag{9}$$

where

$$\mathbf{u} = \sqrt{1 + \frac{f^2}{r^2}}\, \mathbf{m} \times \mathbf{k}, \qquad \mathbf{v} = \mathbf{u} \times \mathbf{m}, \tag{10}$$

and $\mathbf{k} = (0, 0, 1)^\mathrm{T}$.

Proof. If $\mathbf{u}$ and $\mathbf{v}$ are defined as above, three vectors $\{\mathbf{u}, \mathbf{v}, \mathbf{m}\}$ form an orthonormal system, so $\mathbf{u}\mathbf{u}^\mathrm{T} + \mathbf{v}\mathbf{v}^\mathrm{T} + \mathbf{m}\mathbf{m}^\mathrm{T} = \mathbf{I}$. Hence, $\mathbf{P}_\mathrm{m} = \mathbf{I} - \mathbf{m}\mathbf{m}^\mathrm{T} = \mathbf{u}^\mathrm{T} + \mathbf{v}\mathbf{v}^\mathrm{T}$. Since $(\mathbf{u}, \mathbf{k}) = 0$ from Eq. (10), we obtain from Eq. (7)

$$V[\mathbf{m}] = \frac{\varepsilon^2}{2|OP|^2} (\mathbf{u}\mathbf{u}^\mathrm{T} + \mathbf{v}\mathbf{v}^\mathrm{T})(\mathbf{I} - \mathbf{k}\mathbf{k}^\mathrm{T})(\mathbf{u}\mathbf{u}^\mathrm{T} + \mathbf{v}\mathbf{v}^\mathrm{T})$$

$$= \frac{\varepsilon^2}{2|OP|^2} (\mathbf{u}\mathbf{u}^\mathrm{T} + (1 - (\mathbf{v}, \mathbf{k})^2)\mathbf{v}\mathbf{v}^\mathrm{T}). \tag{11}$$

From Eqs. (0), three unit vectors $\mathbf{k}$, $\mathbf{m}$, and $\mathbf{v}$ are coplanar and $(\mathbf{v}, \mathbf{k}) = 0$. Hence, if θ is the angle between $\mathbf{m}$ and $\mathbf{k}$, we have $1 - (\mathbf{v}, \mathbf{k})^2 = (\mathbf{u}, \mathbf{k}) = \cos\theta$. From Fig. 2b, we see that $\cos\theta = 1/\sqrt{1 + r^2/f^2}$, $\sin\theta = 1/\sqrt{1 + f^2/r^2}$, and $|OP|^2 = f\sqrt{1 + r^2/f^2}$. From these, we obtain Eq. (9). ■

If the size of the image is small compared with the focal length f, we can assume that $r \ll f$ and hence $1/(1 + r^2/f^2) \approx 1$. We call this approximation the *small image approximation*. In this approximation, the covariance matrix $V[\mathbf{m}]$ of Eq. (9) becomes $(\varepsilon/f)^2(\mathbf{u}\mathbf{u}^\mathrm{T} + \mathbf{v}\mathbf{v}^\mathrm{T})/2 = (\varepsilon/f)^2(\mathbf{I} - \mathbf{m}\mathbf{m}^\mathrm{T})/2$. If we put

$$\bar{\varepsilon} = \sqrt{E[\|\Delta\mathbf{m}\|^2]}, \tag{12}$$

then $\bar{\varepsilon}^2 = \mathrm{tr}\, V[\mathbf{m}] = (\varepsilon/f)^2$, and hence

COROLLARY 1. *In the small image approximation, the covariance matrix* $V[\mathbf{m}]$ *is given by*

$$V[\mathbf{m}] = \frac{\bar{\varepsilon}^2}{2} (\mathbf{I} - \mathbf{m}\mathbf{m}^\mathrm{T}). \tag{13}$$

2.4. *Effective Focal Length*

In our model, the true position of a point coincides with its expected position to a first approximation. We now show that this does not hold if higher order effects are considered. According to our model (Fig. 1), the image coordinates (x, y) and the corresponding N-vector $\mathbf{m} = (m_1, m_2, m_3)^\mathrm{T}$ are related in the form

$$x = f\frac{m_1}{m_3}, \qquad y = f\frac{m_2}{m_3}. \tag{14}$$

If noise $\Delta\mathbf{m}$ occurs, it can be shown (Appendix B) that

$$E\left[f\frac{m_1 + \Delta m_1}{m_3 + \Delta m_3}\right] = x\left(1 + \frac{\bar{\varepsilon}^2/2}{1 + r^2/f^2} + O(\Delta\mathbf{m})^3\right),$$
$$E\left[f\frac{m_2 + \Delta m_2}{m_3 + \Delta m_3}\right] = y\left(1 + \frac{\bar{\varepsilon}^2/2}{1 + r^2/f^2} + O(\Delta\mathbf{m})^3\right), \tag{15}$$

where (and hereafter) $O(\cdots)^n$ denotes a term of order n or higher in $\cdots$. This means that the expected image coordinates do not agree with their expected values. One way to remove this statistical bias is to define the N-vector $\mathbf{m}$ not by the first of Eqs. (1) but by

$$\mathbf{m} = N\left[\begin{pmatrix} x \\ y \\ \hat{f} \end{pmatrix}\right], \tag{16}$$

where

$$\hat{f} = f\left(1 + \frac{\bar{\varepsilon}^2/2}{1 + r^2/f^2}\right). \tag{17}$$

We call this $\hat{f}$ the *effective focal length*. The following can be confirmed (Appendix B):

$$E\left[f\frac{m_1 + \Delta m_1}{m_3 + \Delta m_3}\right] = x + O(\Delta\mathbf{m})^3,$$
$$E\left[f\frac{m_2 + \Delta m_2}{m_3 + \Delta m_3}\right] = y + O(\Delta\mathbf{m})^3. \tag{18}$$

Hence, all we need to do is treat the correspondence between points on the image plane and their N-vectors unsymmetrically: when we define the N-vector of a "data point" (x, y), we use Eq. (16), while when we interpret

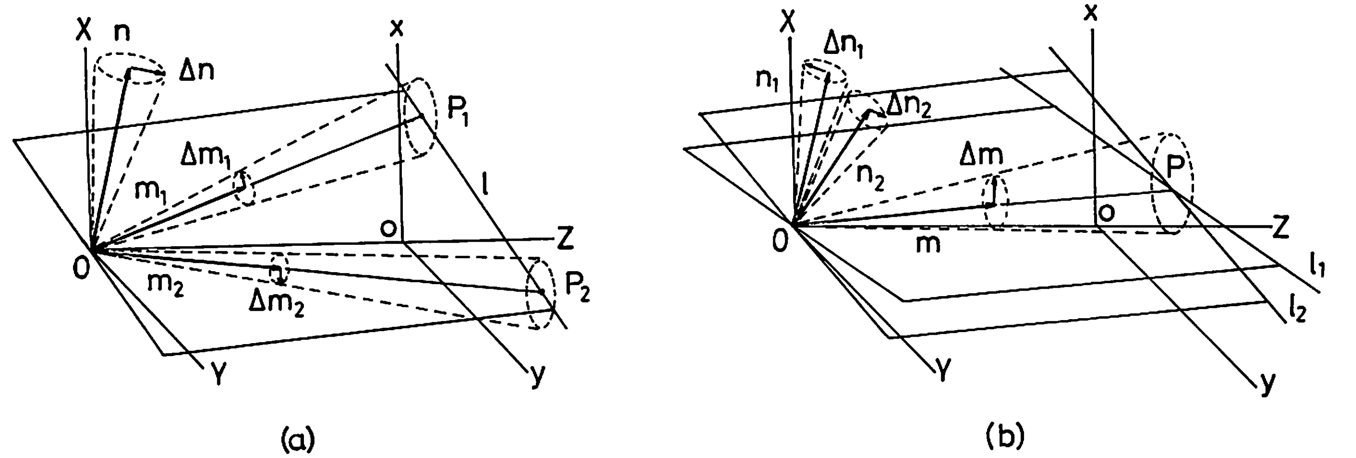

FIG. 3. (a) Statistical behavior of the join of two points and (b) statistical behavior of the intersection of two lines.

a "computed" N-vector $\mathbf{m}$, we use Eqs. (14). In the small image approximation, the effective focal length is approximated by

$$\hat{f} = f\left(1 + \frac{\bar{\varepsilon}^2}{2}\right). \tag{19}$$

The intuitive meaning of the focal length is as follows. If the N-vector $\mathbf{m}$ is defined by the first of Eqs. (1), the error model of Proposition 1 defines an error distribution symmetric with respect to the axes defined by $\mathbf{u}$ and $\mathbf{v}$. This means that when projected onto the image plane, noise is more likely to occur away from the image origin than toward it. Our unsymmetric treatment effectively displaces data points toward the image origin to cancel this bias. However, this bias is a *second order effect* and extremely small. If $\varepsilon = 1 \sim 10$ (pixels) for $f = 1000$ (pixels), for example, we have $\bar{\varepsilon}^2 \approx 10^{-6} \sim 10^{-4}$, which can be negligible compared with 1.

3. COVARIANCE MATRICES OF JOINS AND INTERSECTIONS

Let $\mathbf{u}$ and $\mathbf{v}$ be vectors, and $\mathbf{A}$ a matrix. We define $\mathbf{u} \times \mathbf{A}$ to be the matrix constructed by the vector product of $\mathbf{u}$ and each column of $\mathbf{A}$, and $\mathbf{A} \times \mathbf{v}$ the matrix constructed by the vector product of each row and $\mathbf{v}$. Formally, we define for vectors $\mathbf{u} = (u_i)$, $\mathbf{u} = (v_i)$, and matrix $\mathbf{A} = (A_{ij})$,

$$\mathbf{u} \times \mathbf{A} = \begin{pmatrix} u_2A_{31} - u_3A_{21} & u_2A_{32} - u_3A_{22} & u_2A_{33} - u_3A_{23} \\ u_3A_{11} - u_1A_{31} & u_3A_{12} - u_1A_{32} & u_3A_{13} - u_1A_{33} \\ u_1A_{21} - u_2A_{11} & u_1A_{22} - u_2A_{12} & u_1A_{23} - u_2A_{13} \end{pmatrix}, \tag{20}$$

$$\mathbf{A} \times \mathbf{v} = \begin{pmatrix} A_{12}v_3 - A_{13}v_2 & A_{13}v_3 - A_{11}v_3 & A_{11}v_2 - A_{12}v_1 \\ A_{22}v_3 - A_{23}v_2 & A_{23}v_1 - A_{21}v_3 & A_{21}v_2 - A_{22}v_1 \\ A_{32}v_3 - A_{33}v_2 & A_{33}v_1 - A_{31}v_3 & A_{31}v_2 - A_{32}v_1 \end{pmatrix}. \tag{21}$$

The following identity is easy to confirm:

$$(\mathbf{u} \times \mathbf{A}) \times \mathbf{v} = \mathbf{u} \times (\mathbf{A} \times \mathbf{v}). \tag{22}$$

We simply write this as $\mathbf{u} \times \mathbf{A} \times \mathbf{v}$.

Let $\mathbf{m}_1$ and $\mathbf{m}_2$ be the N-vectors of two points. The N-vector $\mathbf{n}$ of their join is given by Eq. (4). If the N-vectors $\mathbf{m}_1$ and $\mathbf{m}_2$ are perturbed by image noise, the computed N-vector of the join is also perturbed, say by $\Delta\mathbf{n}$ (Fig. 3a). The covariance matrix $V[\mathbf{n}] = E[\Delta\mathbf{n}\Delta\mathbf{n}^T]$ of the join is given as follows.

PROPOSITION 2. *The covariance matrix* $V[\mathbf{n}]$ *of the N-vector* $\mathbf{n}$ *of the join of two points of N-vectors* $\mathbf{m}_1$ *and* $\mathbf{m}_2$ *is*

$$V[\mathbf{n}] = \frac{\mathbf{P}_n(\mathbf{m}_2 \times V[\mathbf{m}_1] \times \mathbf{m}_2 + \mathbf{m}_1 \times V[\mathbf{m}_2] \times \mathbf{m}_1)\mathbf{P}_n}{1 - (\mathbf{m}_1, \mathbf{m}_2)^2}, \tag{23}$$

if the two points are independent,[1] *having covariance matrices* $V[\mathbf{m}_1]$ *and* $V[\mathbf{m}_2]$, *respectively, where* $\mathbf{P}_n = \mathbf{I} - \mathbf{n}\mathbf{n}^T$.

Proof. Put $\mathbf{a} = \mathbf{m}_1 \times \mathbf{m}_2$. Perturbations $\mathbf{m}_1 \to \mathbf{m}_1 + \Delta\mathbf{m}_1$ and $\mathbf{m}_2 \to \mathbf{m}_2 + \Delta\mathbf{m}_2$ cause a perturbation $\Delta\mathbf{a}$ of $\mathbf{a}$ and, to a first approximation,

$$\Delta\mathbf{a} = \Delta\mathbf{m}_1 \times \mathbf{m}_2 + \mathbf{m}_1 \times \Delta\mathbf{m}_2. \tag{24}$$

[1] To be exact, we should say that "the distributions of the image noise for the two points are statistically independent of each other," but for simplicity we use abbreviated expressions like this.

Evidently, $E[\mathbf{a}] = 0$. Put

$$\Delta\mathbf{a}_1 = \Delta\mathbf{m}_1 \times \mathbf{m}_2, \qquad \Delta\mathbf{a}_2 = \mathbf{m}_1 \times \Delta\mathbf{m}_2. \tag{25}$$

Since $\Delta\mathbf{a}_1$ and $\Delta\mathbf{a}_2$ are independent, we obtain

$$E[\Delta\mathbf{a}\Delta\mathbf{a}^T] = E[\Delta\mathbf{a}_1\Delta\mathbf{a}_1^T] + E[\Delta\mathbf{a}_2\Delta\mathbf{a}_2^T]. \tag{26}$$

The first term on the right-hand side becomes

$$\begin{aligned} E[\Delta\mathbf{a}_1\Delta\mathbf{a}_1^T] &= -E[(\mathbf{m}_2 \times \Delta\mathbf{m}_1)(\Delta\mathbf{m}_1 \times \mathbf{m}_2)^T] \\ &= -\mathbf{m}_2 \times E[\Delta\mathbf{m}_1\Delta\mathbf{m}_1^T] \times \mathbf{m}_2 \\ &= -\mathbf{m}_2 \times V[\mathbf{m}_1] \times \mathbf{m}_2. \end{aligned} \tag{27}$$

Similarly, $E[\Delta\mathbf{a}_2\Delta\mathbf{a}_2^T] = -\mathbf{m}_1 \times V[\mathbf{m}_2] \times \mathbf{m}_1$. Hence,

$$V[\mathbf{a}] = -\mathbf{m}_2 \times V[\mathbf{m}_1] \times \mathbf{m}_2 - \mathbf{m}_1 \times V[\mathbf{m}_2] \times \mathbf{m}_1. \tag{28}$$

Since $\mathbf{n} = N[\mathbf{a}]$, we obtain from Appendix A

$$V[\mathbf{n}] = \frac{1}{\|\mathbf{a}\|^2} \boldsymbol{P}_\mathbf{n} V[\mathbf{a}] \boldsymbol{P}_\mathbf{n}. \tag{29}$$

Substituting Eq. (28) into this, together $\|\mathbf{a}\|^2 = \|\mathbf{m}_1 \times \mathbf{m}_2\|^2 = 1 - (\mathbf{m}_1, \mathbf{m}_2)^2$, we obtain Eq. (23). ∎

Corollary 2. *In the small image approximation, the covariance matrix* $V[\mathbf{n}]$ *of the N-vector* $\mathbf{n}$ *of the join of two points of N-vectors* $\mathbf{m}_1$ *and* $\mathbf{m}_2$ *is*

$$V[\mathbf{n}] = \frac{\bar{\varepsilon}^2}{1 - (\mathbf{m}_1, \mathbf{m}_2)^2}\left(\mathbf{I} - \mathbf{n}\mathbf{n}^T - \frac{\mathbf{m}_1\mathbf{m}_1^T + \mathbf{m}_2\mathbf{m}_2^T}{2}\right), \tag{30}$$

if the two points are independent.

Proof. In the small image approximation, $\mathbf{m}_1$ and $\mathbf{m}_2$ have the following covariance matrices (Corollary 1):

$$V[\mathbf{m}_1] = \frac{\bar{\varepsilon}^2}{2}(\mathbf{I} - \mathbf{m}_1\mathbf{m}_1^T), \qquad V[\mathbf{m}_2] = \frac{\bar{\varepsilon}^2}{2}(\mathbf{I} - \mathbf{m}_2\mathbf{m}_2^T). \tag{31}$$

Using the identity $\mathbf{u} \times \mathbf{I} \times \mathbf{u} = \mathbf{u}\mathbf{u}^T - \mathbf{I}$ for unit vector $\mathbf{u}$ and the relation $\mathbf{n} = \pm N[\mathbf{m}_1 \times \mathbf{m}_2] = \pm\mathbf{m}_1 \times \mathbf{m}_2/\|\mathbf{m}_1 \times \mathbf{m}_2\|$, we obtain

$$\mathbf{m}_2 \times V[\mathbf{m}_1] \times \mathbf{m}_2 = \frac{\bar{\varepsilon}^2}{2}(\mathbf{m}_2\mathbf{m}_2^T - \mathbf{I} + \|\mathbf{m}_1 \times \mathbf{m}_2\|^2\mathbf{n}\mathbf{n}^T). \tag{32}$$

Similarly,

$$\mathbf{m}_1 \times V[\mathbf{m}_2] \times \mathbf{m}_1 = \frac{\bar{\varepsilon}^2}{2}(\mathbf{m}_1\mathbf{m}_1^T - \mathbf{I} + \|\mathbf{m}_1 \times \mathbf{m}_2\|^2\mathbf{n}\mathbf{n}^T). \tag{33}$$

Since $\|\mathbf{m}_1 \times \mathbf{m}_2\|^2 = 1 - (\mathbf{m}_1, \mathbf{m}_2)^2$, we obtain

$$\begin{aligned} &\mathbf{m}_2 \times V[\mathbf{m}_1] \times \mathbf{m}_2 + \mathbf{m}_1 \times V[\mathbf{m}_2] \times \mathbf{m}_1 \\ &= \frac{\bar{\varepsilon}^2}{2}(\mathbf{m}_1\mathbf{m}_1^T + \mathbf{m}_2\mathbf{m}_2^T - 2\mathbf{I} + 2(1 - (\mathbf{m}_1, \mathbf{m}_2)^2)\mathbf{n}\mathbf{n}^T). \end{aligned} \tag{34}$$

Hence, from Proposition 2,

$$V[\mathbf{n}] = -\frac{\bar{\varepsilon}^2}{2} \times \frac{\mathbf{P}_\mathbf{n}(\mathbf{m}_1\mathbf{m}_1^T + \mathbf{m}_2\mathbf{m}_2^T - 2\mathbf{I} + 2(1 - (\mathbf{m}_1, \mathbf{m}_2)^2)\mathbf{n}\mathbf{n}^T)\mathbf{P}_\mathbf{n}}{1 - (\mathbf{m}_1, \mathbf{m}_2)^2}. \tag{35}$$

Since $\mathbf{P}_\mathbf{n}$ is the orthogonal projection along the N-vector $\mathbf{n}$, which is orthogonal to $\mathbf{m}_1$ and $\mathbf{m}_2$, it follows that $\mathbf{P}_\mathbf{n}^2 = \mathbf{P}_\mathbf{n}$, $\mathbf{P}_\mathbf{n}\mathbf{n} = \mathbf{0}$, $\mathbf{P}_\mathbf{n}\mathbf{m}_1 = \mathbf{m}_1$, and $\mathbf{P}_\mathbf{n}\mathbf{m}_2 = \mathbf{m}_2$. Using these, we obtain Eq. (30). ∎

Consider the intersection of two lines (Fig. 3b). Let $\mathbf{n}_1$ and $\mathbf{n}_2$ be their N-vectors. The N-vector of their intersection is given by Eq. (3). If noise perturbs $\mathbf{n}_1$ and $\mathbf{n}_2$, the computed $\mathbf{m}$ is also perturbed. Since the formulation is completely dual if expressed in terms of N-vectors [11], the covariance matrix $V[\mathbf{m}]$ is computed in the same way as for the join of two points.

Proposition 3. *The covariance matrix* $V[\mathbf{m}]$ *of the N-vector* $\mathbf{m}$ *of the intersection of two lines of N-vectors* $\mathbf{n}_1$ *and* $\mathbf{n}_2$ *is*

$$V[\mathbf{m}] = -\frac{\mathbf{P}_\mathbf{m}(\mathbf{n}_2 \times V[\mathbf{n}_1] \times \mathbf{n}_2 + \mathbf{n}_1 \times V[\mathbf{n}_2] \times \mathbf{n}_1)\mathbf{P}_\mathbf{m}}{1 - (\mathbf{n}_1, \mathbf{n}_2)^2}, \tag{36}$$

if the two lines are independent, having covariance matrices $V[\mathbf{n}_1]$ *and* $V[\mathbf{n}_2]$, *respectively, where* $\mathbf{P}_\mathbf{m} = \mathbf{I} - \mathbf{m}\mathbf{m}^T$.

4. OPTIMAL LEAST-SQUARES ESTIMATION

4.1. *Optimal Weights and Optimal Estimation*

Points that are on a common line are said to be *collinear*; lines that meet at a common intersection are *concurrent*. If $\mathbf{n}_\alpha = 1, \ldots, N$, are the N-vectors of concurrent lines, and $\mathbf{m}$ is the N-vector of their common intersection, we have the incidence relations $(\mathbf{m}, \mathbf{n}_\alpha) = 0$, $\alpha = 1, \ldots, N$. Hence, if $\mathbf{n}_\alpha$ are given, $\mathbf{m}$ is robustly computed by the least-squares optimization

$$J = \sum_{\alpha=1}^{N} W_\alpha(\mathbf{m}, \mathbf{n}_\alpha)^2 \rightarrow \min, \tag{37}$$

where W_α are positive weights (Fig. 4a). The weights W_α should be determined so that reliable data are given large weights while unreliable data are given small weights.

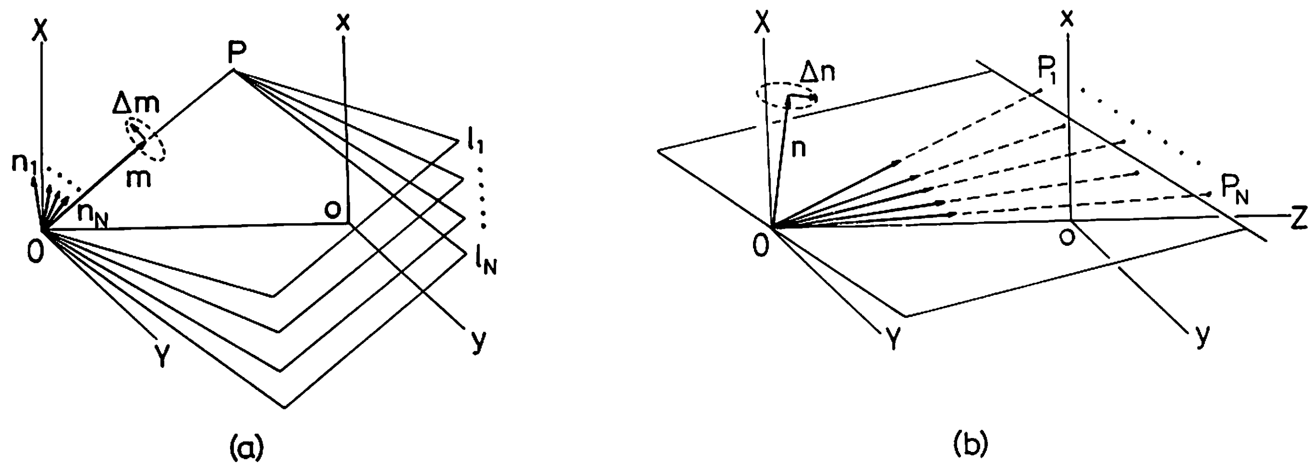

FIG. 4. (a) The common intersection of concurrent lines, and (b) the common line fitted to collinear points.

Since

$$J = \sum_{\alpha=1}^{N} W_\alpha(\mathbf{m}, \mathbf{n}_\alpha)^2 = \left(\mathbf{m}, \left(\sum_{\alpha=1}^{N} W_\alpha \mathbf{n}_\alpha \mathbf{n}_\alpha^{\mathrm{T}}\right)\mathbf{m}\right) \quad (38)$$

is a quadratic form in unit vector $\mathbf{m}$, it is minimized by the unit eigenvector of the *moment matrix*

$$N = \sum_{\alpha=1}^{N} W_\alpha \mathbf{n}_\alpha \mathbf{n}_\alpha^{\mathrm{T}} \quad (39)$$

for the smallest eigenvalue.

Each inner product $(\mathbf{m}, \mathbf{n}_\alpha)$ is precisely 0 if there is no noise. If $\mathbf{n}_\alpha$ is perturbed into $\mathbf{n}'_\alpha = \mathbf{n}_\alpha + \Delta\mathbf{n}_\alpha$, each inner product becomes $\varepsilon_\alpha = (\mathbf{m}, \mathbf{n}'_\alpha) = (\mathbf{m}, \Delta\mathbf{n}_\alpha)$, which is a random variable of mean 0. Its variance is

$$\sigma_\alpha^2 = E[\varepsilon_\alpha^2] = E[(\mathbf{m}, \Delta\mathbf{n}_\alpha)^2] = \mathbf{m}^{\mathrm{T}} E[\Delta\mathbf{n}_\alpha \Delta\mathbf{n}_\alpha^{\mathrm{T}}]\mathbf{m} = (\mathbf{m}, V[\mathbf{n}_\alpha]\mathbf{m}). \quad (40)$$

Hence, $u_\alpha = \varepsilon_\alpha/\sigma_\alpha$ approximately obeys the standard Gaussian distribution of mean 0 and variance 1. If each N-vector $\mathbf{m}_\alpha$ is independent, the joint probability density of $\{u_\alpha\}$, $\alpha = 1, \ldots, N$, is approximated by

$$\prod_{\alpha=1}^{N} \frac{1}{\sqrt{2\pi}} e^{-u_\alpha^2/2} = \left(\frac{1}{\sqrt{2\pi}}\right)^N e^{-\Sigma_{\alpha=1}^{N} u_\alpha^2/2}. \quad (41)$$

It is reasonable to estimate the $\mathbf{m}$ that maximizes this density (the *maximum likelihood estimation*). This is equivalent to

$$\sum_{\alpha=1}^{N} u_\alpha^2 = \sum_{\alpha=1}^{N} \frac{(\mathbf{m}, \mathbf{n}'_\alpha)^2}{(\mathbf{m}, V[\mathbf{n}_\alpha]\mathbf{m})} \to \min. \quad (42)$$

Comparing this with (37), we see that we should choose the weights

$$W_\alpha = \frac{1}{(\mathbf{m}, V[\mathbf{n}_\alpha]\mathbf{m})}. \quad (43)$$

It is easily seen that reliable data have small covariance matrices and are thereby assigned large weights, while unreliable data have large covariance matrices and are thereby assigned small weights.

The above results hold exactly for the common line for collinear points due to the duality of our formulation [11]. Namely, the N-vector $\mathbf{n}$ of the common line for multiple collinear points of N-vectors $\{\mathbf{m}_\alpha\}$, $\alpha = 1, \ldots, N$, is robustly computed by the least-squares optimization

$$\sum_{\alpha=1}^{N} W_\alpha(\mathbf{n}, \mathbf{m}_\alpha)^2 \to \min \quad (44)$$

(Fig. 4b), and the solution is given by the the unit eigenvector of the *moment matrix*

$$M = \sum_{\alpha=1}^{N} W_\alpha \mathbf{m}_\alpha \mathbf{m}_\alpha^{\mathrm{T}} \quad (45)$$

for the smallest eigenvalue. The optimal weights in the sense of the maximum likelihood estimation are given by

$$W_\alpha = \frac{1}{(\mathbf{n}, V[\mathbf{m}_\alpha]\mathbf{n})}. \quad (46)$$

Let us call these weights *optimal weights*, and the moment matrices of Eqs. (39) and (45) with optimal weights the *optimal moment matrices*.

There arises, however, a computational problem: *the optimal weights contain the N-vectors* $\mathbf{m}$ *and* $\mathbf{n}$ *that we* *want to compute*. It seems that this difficulty can be avoided by using estimates of **n** and **m**, say the solution computed by using uniform weights, and iterate the process. It can be shown that if the estimate of **n** (or **m**) has an error $\Delta\mathbf{n}$ (or $\Delta\mathbf{m}$), the solution computed by using approximately optimal weights has an error of $O(\Delta\mathbf{n})^2$ (or $O(\Delta\mathbf{m})^2$) (Appendix C). However, the use of constant weights introduces *statistical bias* however the estimate is chosen. We discuss more about this problem later.

4.2. *Covariance Matrix of Optimal Estimation*

Consider the computation of the common intersection of concurrent lines. If there is no noise, the computed N-vector **m** is exact. If the N-vector of each line is perturbed, the resulting N-vector **m** is also perturbed, say by $\Delta\mathbf{m}$. Its covariance matrix $V[\mathbf{m}] = E[\Delta\mathbf{m}\Delta\mathbf{m}^{\mathrm{T}}]$ is given as follows.

THEOREM 2. *Let* $\{\mathbf{m}, \mathbf{u}, \mathbf{v}\}$ *be the orthonormal system of eigenvectors of the optimal unperturbed moment matrix* **N** *for eigenvalues* 0, λ_u, *and* λ_v, *respectively. If each line is independent and if* $\lambda_u \neq 0$ *and* $\lambda_v \neq 0$, *the covariance matrix* $V[\mathbf{m}]$ *of the N-vector of the optimally estimated common intersection is*

$$V[\mathbf{m}] = \frac{\mathbf{u}\mathbf{u}^{\mathrm{T}}}{\lambda_u} + \frac{\mathbf{v}\mathbf{v}^{\mathrm{T}}}{\lambda_v}. \tag{47}$$

Proof. A perturbation of each N-vector $\mathbf{n}_\alpha \to \mathbf{n}_\alpha + \Delta\mathbf{n}_\alpha$ causes a perturbation $\Delta\mathbf{N}$ of the optimal moment matrix **N** of Eq. (39). To a first approximation,

$$\Delta\mathbf{N} = \sum_{\alpha=1}^{N} W_\alpha(\Delta\mathbf{n}_\alpha\mathbf{n}_\alpha^{\mathrm{T}} + \mathbf{n}_\alpha\Delta\mathbf{n}_\alpha^{\mathrm{T}}). \tag{48}$$

According to the *perturbation theorem* (Appendix D), the unit eigenvector **m** of **N** for the eigenvalue 0 is perturbed by

$$\Delta\mathbf{m} = \frac{(\mathbf{u}, \Delta\mathbf{N}\mathbf{m})}{0 - \lambda_u}\mathbf{u} + \frac{(\mathbf{v}, \Delta\mathbf{N}\mathbf{m})}{0 - \lambda_v}\mathbf{v}. \tag{49}$$

Since $(\mathbf{n}_\alpha, \mathbf{m}) = 0$, we have

$$\begin{aligned}(\mathbf{u}, \Delta\mathbf{N}\mathbf{m}) &= \sum_{\alpha=1}^{N} W_\alpha(\mathbf{u}, (\Delta\mathbf{n}_\alpha\mathbf{n}_\alpha^{\mathrm{T}} + \mathbf{n}_\alpha\Delta\mathbf{n}_\alpha^{\mathrm{T}})\mathbf{m}) \\ &= \sum_{\alpha=1}^{N} W_\alpha(\mathbf{m}, \Delta\mathbf{n}_\alpha)(\mathbf{u}, \mathbf{n}_\alpha).\end{aligned} \tag{50}$$

Similarly, $(\mathbf{v}, \Delta\mathbf{N}\mathbf{m}) = \sum_{\alpha=1}^{N} W_\alpha(\mathbf{m}, \Delta\mathbf{n}_\alpha)(\mathbf{v}, \mathbf{n}_\alpha)$. Hence, if we define

$$\mathbf{s}_\alpha = \frac{(\mathbf{u}, \mathbf{n}_\alpha)}{\lambda_u}\mathbf{u} + \frac{(\mathbf{v}, \mathbf{n}_\alpha)}{\lambda_v}\mathbf{v}, \tag{51}$$

Eq. (49) is written as

$$\Delta\mathbf{m} = -\sum_{\alpha=1}^{N} W_\alpha(\mathbf{m}, \Delta\mathbf{n}_\alpha)\mathbf{s}_\alpha, \tag{52}$$

so the covariance matrix $V[\mathbf{m}]$ is

$$\begin{aligned}V[\mathbf{m}] &= E[\Delta\mathbf{m}\Delta\mathbf{m}^{\mathrm{T}}] \\ &= \sum_{\alpha,\beta=1}^{N} W_\alpha W_\beta E[(\mathbf{m}, \Delta\mathbf{n}_\alpha)(\mathbf{m}, \Delta\mathbf{n}_\beta)]\mathbf{s}_\alpha\mathbf{s}_\beta^{\mathrm{T}}.\end{aligned} \tag{53}$$

Since $\Delta\mathbf{n}_\alpha$ are independent, we have

$$\begin{aligned}E[(\mathbf{m}, \Delta\mathbf{n}_\alpha)(\mathbf{m}, \Delta\mathbf{n}_\beta)] &= \mathbf{m}^{\mathrm{T}}E[\Delta\mathbf{n}_\alpha\Delta\mathbf{n}_\beta^{\mathrm{T}}]\mathbf{m} \\ &= \delta_{\alpha\beta}(\mathbf{m}, V[\mathbf{n}_\alpha]\mathbf{m}) = \frac{\delta_{\alpha\beta}}{W_\alpha},\end{aligned} \tag{54}$$

where Eq. (43) was used, and $\delta_{\alpha\beta}$ is the Kronecker delta taking 1 for $\alpha = \beta$ and 0 otherwise. Substituting this into Eq. (53), we obtain

$$\begin{aligned}V[\mathbf{m}] &= \sum_{\alpha=1}^{N} W_\alpha\mathbf{s}_\alpha\mathbf{s}_\alpha^{\mathrm{T}} = \sum_{\alpha=1}^{N} W_\alpha\left(\frac{\mathbf{u}^{\mathrm{T}}\mathbf{n}_\alpha\mathbf{n}_\alpha^{\mathrm{T}}\mathbf{u}}{\lambda_u^2}\mathbf{u}\mathbf{u}^{\mathrm{T}}\right. \\ &\quad \left. + \frac{\mathbf{u}^{\mathrm{T}}\mathbf{n}_\alpha\mathbf{n}_\alpha^{\mathrm{T}}\mathbf{v}}{\lambda_u\lambda_v}(\mathbf{u}\mathbf{v}^{\mathrm{T}} + \mathbf{v}\mathbf{u}^{\mathrm{T}}) + \frac{\mathbf{u}^{\mathrm{T}}\mathbf{n}_\alpha\mathbf{n}_\alpha^{\mathrm{T}}\mathbf{v}}{\lambda_v^2}\mathbf{v}\mathbf{v}^{\mathrm{T}}\right) \\ &= \frac{(\mathbf{u}, \mathbf{N}\mathbf{u})}{\lambda_u^2}\mathbf{u}\mathbf{u}^{\mathrm{T}} + \frac{(\mathbf{u}, \mathbf{N}\mathbf{v})}{\lambda_u\lambda_v}(\mathbf{u}\mathbf{v}^{\mathrm{T}} + \mathbf{v}\mathbf{u}^{\mathrm{T}}) + \frac{(\mathbf{v}, \mathbf{N}\mathbf{v})}{\lambda_v^2}\mathbf{v}\mathbf{v}^{\mathrm{T}}.\end{aligned} \tag{55}$$

Since **u** and **v** are eigenvectors of **N** for the eigenvalues λ_u and λ_v, respectively, we have $(\mathbf{u}, \mathbf{N}\mathbf{u}) = \lambda_v$, and $(\mathbf{u}, \mathbf{N}\mathbf{v}) = 0$. Hence, we obtain Eq. (47). ■

COROLLARY 3. *If each line is independent, the root mean square of the error* $\Delta\mathbf{m}$ *of the N-vector* **m** *of the optimally estimated intersection is*

$$\sqrt{E[\|\Delta\mathbf{m}\|^2]} = \sqrt{\frac{1}{\lambda_u} + \frac{1}{\lambda_v}}, \tag{56}$$

where λ_u *and* λ_v *are the positive eigenvalues of the optimal unperturbed moment matrix* **N**. *This is* $O(\nu/\sqrt{N})$ *for* $\|\Delta\mathbf{n}_\alpha\| \approx \nu$.

Proof. Note that $E[\|\Delta\mathbf{m}\|^2] = \mathrm{tr}\,V[\mathbf{m}]$. From Eq. (47), we obtain $\mathrm{tr}\,V[\mathbf{m}] = 1/\lambda_u + 1/\lambda_v$, hence Eq. (56). If $\|\Delta\mathbf{n}_\alpha\| \approx \nu$, then $V[\mathbf{n}_\alpha] = O(\nu^2)$. Consequently, $W_\alpha = O(1/\nu^2)$ and hence $\mathbf{N} = O(N/\nu^2)$. It follows that the eigenvalues λ_u and λ_v of **N** are also $O(N/\nu^2)$, meaning that Eq. (56) is $O(\nu/\sqrt{N})$. ■

The same results hold for optimal line fitting due to the duality of our formulation [11].

THEOREM 3. *Let* $\{\mathbf{n}, \mathbf{u}, \mathbf{v}\}$ *be the orthonormal system of eigenvectors of the optimal unperturbed moment matrix* $\mathbf{M}$ *for eigenvalues* 0, λ_u, *and* λ_v, *respectively. If each point is independent and if* $\lambda_u \neq 0$ *and* $\lambda_v \neq 0$, *the covariance matrix* $V[\mathbf{n}]$ *of the N-vector of the optimally fitted line is*

$$V[\mathbf{n}] = \frac{\mathbf{u}\mathbf{u}^T}{\lambda_u} + \frac{\mathbf{v}\mathbf{v}^T}{\lambda_v}. \qquad (57)$$

COROLLARY 4. *If each point is independent, the root mean square of the error* $\Delta\mathbf{n}$ *of the N-vector* $\mathbf{n}$ *of the fitted common line is*

$$\sqrt{E[\|\Delta\mathbf{n}\|^2]} = \sqrt{\frac{1}{\lambda_u} + \frac{1}{\lambda_v}}, \qquad (58)$$

where λ_u *and* λ_v *are the positive eigenvalues of the optimal unperturbed moment matrix* $\mathbf{M}$. *This is* $O(\mu/\sqrt{N})$ *for* $\|\Delta\mathbf{m}_\alpha\| \approx \mu$.

Although these results only apply to "unperturbed" moment matrices, it can be shown that the covariance matrix $\tilde{V}[\mathbf{m}]$ (or $\tilde{V}[\mathbf{n}]$) estimated from N perturbed N-vectors is different from the true covariance matrix $V[\mathbf{m}]$ (or $V[\mathbf{n}]$) by only $O(\nu^3/N\sqrt{N})$ (or $O(\mu^3/N\sqrt{N})$) for $\|\Delta\mathbf{n}_\alpha\| \approx \nu$ (or $\|\Delta\mathbf{m}_\alpha\| \approx \mu$) (Appendix C).

4.3. *Statistical Bias of Optimal Estimation*

An estimate is *statistically unbiased* if the expectation of its error is zero and *statistically biased* otherwise. If we taken the expectation of Eq. (49), we see that $E[\Delta\mathbf{m}] = \mathbf{0}$, so the optimal estimate appears to be statistically unbiased. However, this is so because second order terms are neglected in Eq. (48). The true perturbation of $\mathbf{N}$ is

$$\Delta\mathbf{N} = \sum_{\alpha=1}^{N} W_\alpha(\Delta\mathbf{n}_\alpha\mathbf{n}_\alpha^T + \mathbf{n}_\alpha\Delta\mathbf{n}_\alpha^T + \Delta\mathbf{n}_\alpha\Delta\mathbf{n}_\alpha^T), \qquad (59)$$

and its expectation is not zero:

$$E[\Delta\mathbf{N}] = \sum_{\alpha=1}^{N} W_\alpha E[\Delta\mathbf{n}_\alpha\Delta\mathbf{n}_\alpha^T] = \sum_{\alpha=1}^{N} W_\alpha V[\mathbf{n}_\alpha]. \qquad (60)$$

Hence, the expectation of Eq. (49) becomes

$$E[\Delta\mathbf{m}] = -\left(\frac{\mathbf{u}\mathbf{u}^T}{\lambda_u} + \frac{\mathbf{v}\mathbf{v}^T}{\lambda_v}\right)\left(\sum_{\alpha=1}^{N} W_\alpha V[\mathbf{n}_\alpha]\right)\mathbf{m}. \qquad (61)$$

Thus, the optimal estimate is in general statistically biased. Combining this with Theorem 1, we obtain

THEOREM 4. *If each line has covariance matrix* $V[\mathbf{n}_\alpha]$, *the N-vector* $\mathbf{m}$ *of the optimally estimated intersection is statistically biased by*

$$E[\Delta\mathbf{m}] = -V[\mathbf{m}]\left(\sum_{\alpha=1}^{N} W_\alpha V[\mathbf{n}_\alpha]\right)\mathbf{m}. \qquad (62)$$

THEOREM 5. *If each point has covariance matrix* $V[\mathbf{m}_\alpha]$, *the N-vector* $\mathbf{n}$ *of the optimally fitted line is statistically biased by*

$$E[\Delta\mathbf{n}] = -V[\mathbf{n}]\left(\sum_{\alpha=1}^{N} W_\alpha V[\mathbf{m}_\alpha]\right)\mathbf{n}. \qquad (63)$$

It should be emphasized that this statistical bias is a *second order effect:* if each $\Delta\mathbf{n}_\alpha$ is $O(\nu)$, the bias $E[\Delta\mathbf{m}]$ is $O(\nu^2)$, whereas the root-mean-square error $\sqrt{E[\|\Delta\mathbf{m}\|^2]}$ is $O(\nu)$. Similarly, if each $\Delta\mathbf{m}_\alpha$ is $O(\mu)$, the bias $E[\Delta\mathbf{n}]$ is $O(\mu^2)$, whereas $\sqrt{E[\|\Delta\mathbf{n}\|^2]}$ is $O(\mu)$. Hence, the bias is not very apparent unless the image noise is very large.

From Eq. (60), we can construct the following optimal *unbiased* estimation scheme.

PROPOSITION 4. *The N-vector* $\hat{\mathbf{m}}$ *of the common intersection of lines of N-vectors* $\{\mathbf{n}_\alpha\}$ *with covariance matrices* $V[\mathbf{n}_\alpha]$ *estimated as the unit eigenvector of*

$$\hat{\mathbf{N}} = \sum_{\alpha=1}^{N} W_\alpha(\mathbf{n}_\alpha\mathbf{n}_\alpha^T - V[\mathbf{n}_\alpha]) \qquad (64)$$

for the smallest eigenvalue is statistically unbiased.

PROPOSITION 5. *The N-vector* $\hat{\mathbf{n}}$ *of the common line fitted to points of N-vectors* $\{\mathbf{m}_\alpha\}$ *with covariance matrices* $V[\mathbf{m}_\alpha]$ *estimated as the unit eigenvector of*

$$\hat{M} = \sum_{\alpha=1}^{N} W_\alpha(\mathbf{m}_\alpha\mathbf{m}_\alpha^T - V[\mathbf{m}_\alpha]) \qquad (65)$$

for the smallest eigenvalue is statistically unbiased.

Evidently, $E[\Delta\hat{\mathbf{N}}] = E[\Delta\hat{\mathbf{M}}] = \mathbf{O}$ for the unbiased moment matrices $\mathbf{N}$ and $\mathbf{M}$, and hence $E[\Delta\hat{\mathbf{m}}] = E[\Delta\hat{\mathbf{n}}] = \mathbf{0}$ for the resulting estimates $\hat{\mathbf{m}}$ and $\hat{\mathbf{n}}$. It can be shown, however, that we need not introduce the unbiased moment matrix *if the noise in the N-vectors is isotropic.*

COROLLARY 5. *If noise enters each N-vector isotropically, the optimally estimated intersection and the optimally fitted line are statistically unbiased.*

Proof. If $\Delta\mathbf{n}_\alpha$ is isotropic around $\mathbf{n}_\alpha$, its covariance matrix has the form

$$V[\mathbf{n}_\alpha] = \frac{\sigma^2}{2}(\mathbf{I} - \mathbf{n}_\alpha\mathbf{n}_\alpha^T), \qquad \sigma^2 = E[\|\Delta\mathbf{n}_\alpha\|^2]. \qquad (66)$$

Hence,

$$\hat{\mathbf{N}} = \sum_{\alpha=1}^{N} W_\alpha \left(\mathbf{n}_\alpha \mathbf{n}_\alpha^\top - \frac{\sigma^2}{2}(\mathbf{I} - \mathbf{n}_\alpha \mathbf{n}_\alpha^\top) \right)$$

$$= \left(1 + \frac{\sigma^2}{2}\right)\left(\sum_{\alpha=1}^{N} W_\alpha \mathbf{n}_\alpha \mathbf{n}_\alpha^\top\right) - \frac{\sigma^2}{2}\left(\sum_{\alpha=1}^{N} W_\alpha\right)\mathbf{I}, \quad (67)$$

which has the same eigenvectors as $\mathbf{N} = \sum_{\alpha=1}^{N} W_\alpha \mathbf{n}_\alpha \mathbf{n}_\alpha^\top$. Similarly, $\hat{\mathbf{M}}$ has the same eigenvectors as $\mathbf{M}$. ∎

In particular, optimal line fitting near the center of the image plane under the small image approximation is statistically unbiased.

5. EDGES, VANISHING POINTS, AND FOCUSES OF EXPANSION

5.1. *Error in Line Fitting to Edges*

Let us call a dense alignment of collinear pixels an *edge segment*, and each pixel an *edge pixel*. Let $\mathbf{m}_\alpha$, $\alpha = 1, \ldots, N$, be the N-vectors of the edge pixels of an edge segment aligned in this order. Choose their signs so that they all point toward the image plane. Let us call the angle Ω between $\mathbf{m}_1$ and $\mathbf{m}_N$ the *disparity* of the edge segment. Define

$$\mathbf{u} = \pm N[\mathbf{m}_N - \mathbf{m}_1], \qquad \mathbf{m}_G = \pm N[\mathbf{m}_N + \mathbf{m}_1]. \quad (68)$$

We call the vector $\mathbf{u}$ the *orientation* of the edge segment, and the point of the N-vector $\mathbf{m}_G$ its *center point* (Fig. 5). If each pair of consecutive edge pixels has the same disparity, we can write

$$\mathbf{m}_\alpha = \mathbf{u}\sin\phi_\alpha + \mathbf{m}_G \cos\phi_\alpha,$$
$$\phi_\alpha = \frac{\Omega}{N-1}\left(\alpha - \frac{N+1}{2}\right). \quad (69)$$

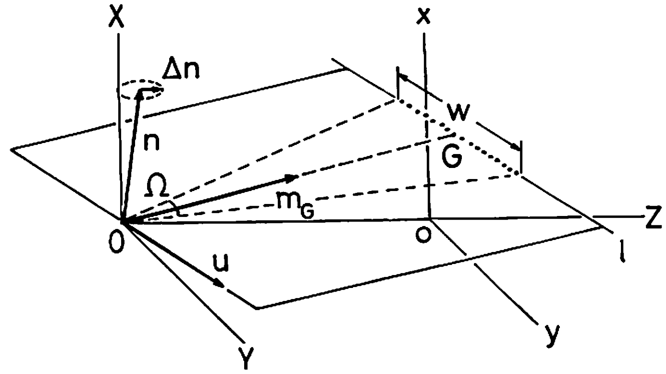

FIG. 5. Line fitting to an edge segment.

Assume that all edge pixels are independent. If the edge segment is short and located near the image origin, the small image approximation $V[\mathbf{m}_\alpha] = \bar{\varepsilon}^2(\mathbf{I} - \mathbf{m}_\alpha\mathbf{m}_\alpha^\top)/2$ (Corollary 2) can be applied, and the optimal weights (46) become

$$W_\alpha = \frac{1}{(\mathbf{n}, V[\mathbf{m}_\alpha]\mathbf{n})} = \frac{2}{\bar{\varepsilon}^2(1 - (\mathbf{n}, \mathbf{m}_\alpha)^2)} = \frac{2}{\bar{\varepsilon}^2}. \quad (70)$$

Hence,the optimal moment matrix (45) becomes

$$\mathbf{M} = \frac{2}{\bar{\varepsilon}^2}\left(\sum_{\alpha=1}^{N} \sin^2\phi_\alpha \mathbf{u}\mathbf{u}^\top + \sum_{\alpha=1}^{N} \cos\phi_\alpha \sin\phi_\alpha(\mathbf{u}\mathbf{m}_G^\top + \mathbf{m}_G\mathbf{u}^\top) + \sum_{\alpha=1}^{N} \cos^2\phi_\alpha \mathbf{m}_G\mathbf{m}_G^\top\right). \quad (71)$$

If the edge pixels are dense, the summations can be approximated by integrations[2]:

$$\sum_{\alpha=1}^{N} \sin^2\phi_\alpha \approx \frac{N}{\Omega}\int_{-\Omega/2}^{\Omega/2} \sin^2\phi \, d\phi = \frac{N}{2}(1 - \operatorname{sinc}\Omega), \quad (72)$$

$$\sum_{\alpha=1}^{N} \sin\phi_\alpha \cos\phi_\alpha \approx \frac{N}{\Omega}\int_{-\Omega/2}^{\Omega/2} \sin\phi\cos\phi \, d\phi = 0, \quad (73)$$

$$\sum_{\alpha=1}^{N} \cos^2\phi_\alpha \approx \frac{N}{\Omega}\int_{-\Omega/2}^{\Omega/2} \cos^2\phi \, d\phi = \frac{N}{2}(1 + \operatorname{sinc}\Omega), \quad (74)$$

where symbol sinc x denotes the function $\operatorname{sinc} x = (\sin x)/x$. Hence,

$$\mathbf{M} = \frac{N}{\bar{\varepsilon}^2}((1 - \operatorname{sinc}\Omega)\mathbf{u}\mathbf{u}^\top + (1 + \operatorname{sinc}\Omega)\mathbf{m}_G\mathbf{m}_G^\top). \quad (75)$$

From Theorem 2, the covariance matrix $V[\mathbf{n}]$ of the optimally fitted line is given by

$$V[\mathbf{n}] = \frac{\bar{\varepsilon}^2}{N}\left(\frac{\mathbf{u}\mathbf{u}^\top}{1 - \operatorname{sinc}\Omega} + \frac{\mathbf{m}_G\mathbf{m}_G^\top}{1 + \operatorname{sinc}\Omega}\right). \quad (76)$$

If the length of the edge segment is w (measured in pixels), the disparity is approximately $\Omega \approx w/f$. If ρ is the *edge density* (i.e., the number of edge pixels per unit pixel length), the total number of edge pixels is $N = \rho w$. If the length w is small compared with f, the disparity Ω is small, so $\operatorname{sinc}\Omega \approx 1 - \Omega^2/6 + \ldots$ Then, we have the following approximation:

[2] Strictly speaking, the summation $\sum_{\alpha=1}^{N}$ should be $\sum_{\alpha=1}^{N-1}$, but the difference can be ignored if N is large.

$$1 - \operatorname{sinc}\Omega \approx \frac{\Omega^2}{6} \approx \frac{w^2}{6f^2}, \qquad 1 + \operatorname{sinc}\Omega \approx 2 - \frac{\Omega^2}{6} \approx 2. \tag{77}$$

Then, the covariance matrix $V[\mathbf{n}]$ is approximated by

$$V[\mathbf{n}] \approx \frac{6\kappa}{w^3}\mathbf{u}\mathbf{u}^{\mathrm{T}} + \frac{\kappa}{2f^2w}\mathbf{m}_G\mathbf{m}_G^{\mathrm{T}}, \qquad \kappa = \frac{\varepsilon^2}{\rho}. \tag{78}$$

The constant κ, which we call the *image resolution*, is small if edge pixels are dense and accurately detected. This is an image property independent of the geometry of the data.[3]

If the length w of the edge segment is very small compared with the focal length f, the right-hand side of Eq. (78) is dominated by the first term, and Eq. (78) is further approximated by

$$V[\mathbf{n}] \approx \frac{6\kappa}{w^3}\mathbf{u}\mathbf{u}^{\mathrm{T}}. \tag{79}$$

From this, we have the following observations:

- The error in the N-vector $\mathbf{n}$ of the optimally fitted line is most likely to occur in the orientation $\mathbf{u}$ of the fitted line.
- The error in the optimally fitted line is approximately proportional to $\kappa^{1/2}$ of the image resolution κ (hence proportional to the image accuracy ε and to $\rho^{-1/2}$ of the edge density ρ).
- The error in the optimally fitted line is approximately proportional to $w^{-3/2}$ of the length w of the edge segment.

From Corollary 5, we also observe that

- The optimally fitted line is statistically unbiased.

Remark. The above observation is obtained with respect to the "N-vector" of the fitted line, which is conceptually infinite, with a view to computing its vanishing point. However, even though the $\mathbf{m}_G\mathbf{m}_G^{\mathrm{T}}$ component is very small as compared with the $\mathbf{u}\mathbf{u}^{\mathrm{T}}$ component, both terms cause discrepancies of edge points from the fitted line by comparable magnitudes. So, if we want to construct a "measure of discrepancy" of an edge segment from its ideal position, we must consider both terms (for the details, see [12, 17]).

5.2. *Error in Vanishing Points*

If lines l_α, $\alpha = 1, \ldots, N$, are projections of parallel lines in the scene, they are concurrent on the image plane.

[3] Strictly speaking, the edge density varies with the orientation of the edge segment: if pixels are aligned in a square grid array, horizontal edge segments are $\sqrt{2}$ times as dense as edge segments of orientation 45°. However, we ignore this anisotropy.

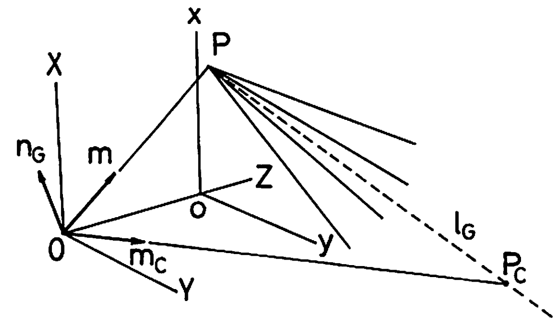

FIG. 6. The center line l_G of concurrent lines. Point P_C is conjugate to the vanishing point P on line l_G.

Their common intersection is their "vanishing point"; its N-vector $\mathbf{m}$ indicates the 3D orientation of the space lines [11]. Let $\mathbf{n}_\alpha$, $\alpha = 1, \ldots, N$, be the N-vectors of the N lines. The N-vector $\mathbf{m}$ of the vanishing point is given as the unit eigenvector for the eigenvalue 0 of the optimal moment matrix in the form of Eq. (39). Let $\mathbf{n}_G$ be the unit eigenvector of $\mathbf{N}$ for the largest eigenvalue. Vector $\mathbf{n}_G$ can be regarded as the N-vector of a hypothetical *center line* l_G of the N lines (Fig. 6). Since the three eigenvectors form an orthonormal system, the unit eigenvector $\mathbf{m}_C$ for the second largest eigenvalue equals $\pm\mathbf{m} \times \mathbf{n}_G$. Vector $\mathbf{m}_C$ is orthogonal to both $\mathbf{n}_G$ and $\mathbf{m}$ and hence can be identified with the N-vector of the point P_C "conjugate" [11] to the vanishing point P on the center line l_G.

The eigenvalue of $\mathbf{N}$ for $\mathbf{n}_G$ is given by

$$(\mathbf{n}_G, \mathbf{N}\mathbf{n}_G) = \sum_{\alpha=1}^{N} W_\alpha(\mathbf{n}_G, \mathbf{n}_\alpha)^2. \tag{80}$$

The eigenvalue for $\mathbf{m}_C$ is

$$(\mathbf{m}_C, \mathbf{N}\mathbf{m}_C) = \sum_{\alpha=1}^{N} W_\alpha(\mathbf{m} \times \mathbf{n}_G, \mathbf{n}_\alpha)^2 = \sum_{\alpha=1}^{N} W_\alpha|\mathbf{m}, \mathbf{n}_G, \mathbf{n}_\alpha|^2, \tag{81}$$

where $|\mathbf{a}, \mathbf{b}, \mathbf{c}|$ $(= (\mathbf{a} \times \mathbf{b}, \mathbf{c}) = (\mathbf{b} \times \mathbf{c}, \mathbf{a}) = (\mathbf{c} \times \mathbf{a}, \mathbf{b}))$ denotes the scalar triple product of vectors $\mathbf{a}$, $\mathbf{b}$, and $\mathbf{c}$. From Theorem 2, the covariance matrix $V[\mathbf{m}]$ of the N-vector $\mathbf{m}$ of the vanishing point is given by

$$V[\mathbf{m}] = \frac{\mathbf{m}_C\mathbf{m}_C^{\mathrm{T}}}{\sum_{\alpha=1}^{N} W_\alpha|\mathbf{m}, \mathbf{n}_G, \mathbf{n}_\alpha|^2} + \frac{\mathbf{n}_G\mathbf{n}_G^{\mathrm{T}}}{\sum_{\alpha=1}^{N} W_\alpha(\mathbf{n}_G, \mathbf{n}_\alpha)^2}. \tag{82}$$

If the separations among the lines are small, the right-hand side is dominated by the first term. If ϕ_α is the angle between $\mathbf{n}_G$ and $\mathbf{n}_\alpha$, we have

$$|\mathbf{m}, \mathbf{n}_G, \mathbf{n}_\alpha|^2 = (\mathbf{m}, \mathbf{n}_G \times \mathbf{n}_\alpha)^2 = \sin^2\phi_\alpha, \tag{83}$$

because $\mathbf{m}$ is parallel to $\mathbf{n}_G \times \mathbf{n}_\alpha$. Let us call ϕ_α the *deviation angle* (from the hypothetical center line). If the covariance matrix $V[\mathbf{n}_\alpha]$ of each line is approximated by Eq. (79), we have

$$V[\mathbf{n}_\alpha] \approx \frac{6\kappa}{w_\alpha^3}\mathbf{u}_\alpha\mathbf{u}_\alpha^{\mathrm{T}} = \frac{6\kappa}{w_\alpha^3}(\mathbf{m}_{G\alpha} \times \mathbf{n}_\alpha)(\mathbf{m}_{G\alpha} \times \mathbf{n}_\alpha)^{\mathrm{T}}, \quad (84)$$

where $\mathbf{u}_\alpha$ is the orientation of the αth edge segment and $\mathbf{m}_{G\alpha}$ is the N-vector of its center point. Since $\mathbf{n}_\alpha = \pm\mathbf{m}_{G\alpha} \times \mathbf{n}_\alpha$, the optimal weight W_α of Eq. (43) is approximated by

$$W_\alpha \approx \frac{w_\alpha^3}{6\kappa(\mathbf{m}, \mathbf{u}_\alpha)^2} = \frac{w_\alpha^3}{6\kappa|\mathbf{m}, \mathbf{m}_{G\alpha}, \mathbf{n}_\alpha|^2}. \quad (85)$$

If θ_α is the angle between $\mathbf{m}_{G\alpha}$ and $\mathbf{m}$, we have

$$|\mathbf{m}, \mathbf{m}_{G\alpha}, \mathbf{n}_\alpha|^2 = (\mathbf{m} \times \mathbf{m}_{G\alpha}, \mathbf{n}_\alpha)^2 = \sin^2\theta_\alpha, \quad (86)$$

because $\mathbf{n}_\alpha$ is parallel to $\mathbf{m} \times \mathbf{m}_{G\alpha}$. The angle θ_α indicates the "disparity" of the vanishing point from the center point of the αth edge segment.

Substituting these into Eq. (82), we obtain the approximation

$$V[\mathbf{m}] \approx \frac{6\kappa\mathbf{m}_C\mathbf{m}_C^{\mathrm{T}}}{\sum_{\alpha=1}^N w_\alpha^3 \sin^2\phi_\alpha/\sin^2\theta_\alpha}. \quad (87)$$

From this, we have the following observations:

- The error in the vanishing point is most likely to occur along the center line.
- The error in the vanishing point is approximately proportional to $\kappa^{1/2}$ of the image resolution κ.
- The error in the vanishing point is approximately proportional to $w_\alpha^{-3/2}$ of the lengths w_α of the individual edge segments.
- The error in the vanishing point is approximately proportional to $1/\sin\phi_\alpha$ of the deviation angles ϕ_α of the individual edge segments from the center line.
- The error in the vanishing point is approximately proportional to $\sin\theta_\alpha$ of the disparities θ_α of the vanishing point from the center points of the individual edge segments.

From Eq. (84), we have $\sum_{\alpha=1}^N W_\alpha V[\mathbf{n}_\alpha] \approx c\mathbf{u}_G\mathbf{u}_G^{\mathrm{T}}$, where c is a positive constant and $\mathbf{u}_G$ is the orientation of the center line. From Eq. (87) and Theorem 4, the statistical bias of the vanishing point is

$$\Delta\mathbf{m} \approx (c'\mathbf{m}_C\mathbf{m}_C^{\mathrm{T}})(c\mathbf{u}_G\mathbf{u}_G^{\mathrm{T}})\mathbf{m} = cc'(\mathbf{m}_C, \mathbf{u}_G)(\mathbf{u}_G, \mathbf{m})\mathbf{m}_C, \quad (88)$$

where c' is a positive constant. If the signs of $\mathbf{m}$ and $\mathbf{m}_C$ are chosen so that they both point toward the image plane, we have $(\mathbf{m}_C, \mathbf{u}_G)(\mathbf{u}_G, \mathbf{m}) < 0$, whichever way $\mathbf{u}_G$ is oriented (Fig. 6). Thus, we observe that

- The vanishing point is statistically biased along the center line *away* from the edge segments.
- The bias is small when the vanishing point is close to the edge segments or infinitely away from them.

5.3. *Error in Focus of Expansion*

When correspondences of feature points are given between two images, we identify their coordinate systems and call the line passing through two corresponding feature points their *trajectory*. As in the case of edge segments, define

$$\mathbf{u} = \pm N[\mathbf{m}' - \mathbf{m}], \qquad \mathbf{m}_G = \pm N[\mathbf{m}' + \mathbf{m}]. \quad (89)$$

We call the vector $\mathbf{u}$ the *orientation* of the trajectory, and the image point of N-vector $\mathbf{m}_G$ its *center point* (Fig. 7).

Under the small image approximation, the covariance matrix $V[\mathbf{n}]$ of the trajectory passing through two feature points of N-vectors $\mathbf{m}$ and $\mathbf{m}'$ is given by Corollary 2. It is easy to confirm by direct calculation that $\mathbf{u}$ and $\mathbf{m}_G$ are both eigenvectors of $V[\mathbf{n}]$:

$$V[\mathbf{n}]\mathbf{u} = \frac{\bar{\varepsilon}^2}{2(1 - (\mathbf{m}, \mathbf{m}'))}\mathbf{u}, \qquad V[\mathbf{n}]\mathbf{m}_G = \frac{\bar{\varepsilon}^2}{2(1 + (\mathbf{m}, \mathbf{m}'))}\mathbf{m}_G. \quad (90)$$

Hence, the covariance matrix $V[\mathbf{n}]$ of Eq. (30) is rewritten as

$$V[\mathbf{n}] = \frac{\bar{\varepsilon}^2}{2}\left(\frac{\mathbf{u}\mathbf{u}^{\mathrm{T}}}{1 - (\mathbf{m}, \mathbf{m}')} + \frac{\mathbf{m}_G\mathbf{m}_G^{\mathrm{T}}}{1 + (\mathbf{m}, \mathbf{m}')}\right). \quad (91)$$

If the distance w between the two feature points (measured in pixels) is small compared with f, the right-hand

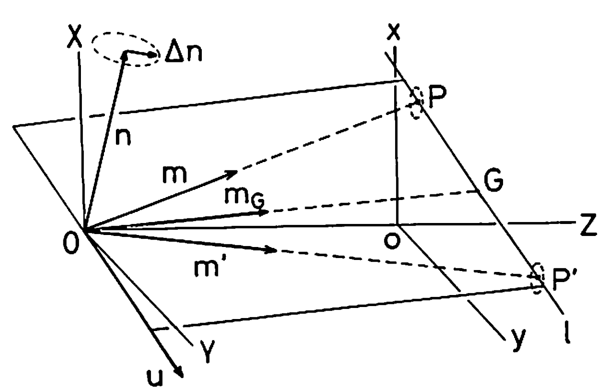

FIG. 7. The orientation u and the center point G of a trajectory.

side is dominated by the first term, and the approximation

$$(\mathbf{m}, \mathbf{m}') \approx \cos\frac{w}{f} \approx 1 - \frac{w^2}{2f^2} \qquad (92)$$

can be applied. Hence, the covariance matrix $V[\mathbf{n}]$ is approximated by[4]

$$V[\mathbf{n}] \approx \left(\frac{\varepsilon}{w}\right)^2 \mathbf{u}\mathbf{u}^{\mathrm{T}}. \qquad (93)$$

The "focus of expansion" is the common intersection of the image trajectories of feature points translating in the scene (or viewed from a translating camera). Its N-vector indicates the 3D orientation of the object (or camera) motion [11]. Since the two images are processed separately, the noise in each image can be regarded as statistically independent.[5] Let $\mathbf{n}_\alpha$ be the N-vector of the αth trajectory. As in the case of vanishing points, the covariance matrix $V[\mathbf{m}]$ of the N-vector $\mathbf{m}$ of the focus of expansion is given by Theorem 2 in the form

$$V[\mathbf{m}] = \frac{\mathbf{m}_C\mathbf{m}_C^{\mathrm{T}}}{\sum_{\alpha=1}^N W_\alpha|\mathbf{m}, \mathbf{n}_G, \mathbf{n}_\alpha|^2} + \frac{\mathbf{n}_G\mathbf{n}_G^{\mathrm{T}}}{\sum_{\alpha=1}^N W_\alpha(\mathbf{n}_G, \mathbf{n}_\alpha)^2}, \qquad (94)$$

where $\mathbf{n}_G$ is the N-vector of the hypothetical "center line" of the N trajectories and $\mathbf{m}_C = \pm\mathbf{m} \times \mathbf{n}_G$ is the N-vector of the point "conjugate" [11] to the focus of expansion on the center line. Again, if the separations among the trajectories are small, the right-hand side is dominated by the first term. If ϕ_α is the angle between $\mathbf{n}_G$ and $\mathbf{n}_\alpha$, we have

$$|\mathbf{m}, \mathbf{n}_G, \mathbf{n}_\alpha|^2 = \sin^2\phi_\alpha, \qquad (95)$$

where ϕ_α is the "deviation angle" from the hypothetical center line. The covariance matrix $V[\mathbf{n}_\alpha]$ of each trajectory is approximated by Eq. (93). Since $\mathbf{u}_\alpha = \pm\mathbf{m}_{G\alpha} \times \mathbf{n}_\alpha$, W_α of Eq. (43) is approximated by

$$W_\alpha \approx \frac{w_\alpha}{\varepsilon^2(\mathbf{m}, \mathbf{u}_\alpha)^2} = \frac{w_\alpha}{\varepsilon^2|\mathbf{m}, \mathbf{m}_{G\alpha}, \mathbf{n}_\alpha|^2}, \qquad (96)$$

where $\mathbf{m}_{G\alpha}$ is the N-vector of the center point of the αth trajectory. If θ_α is the angle between $\mathbf{m}_{G\alpha}$ and $\mathbf{m}$, we have

$$|\mathbf{m}, \mathbf{m}_{G\alpha}, \mathbf{n}_\alpha|^2 = \sin^2\theta_\alpha. \qquad (97)$$

The angle θ_α indicates the "disparity" of the focus of expansion from the center point of the αth trajectory. Substituting these into Eq. (94), we obtain the following approximation:

$$V[\mathbf{m}] \approx \frac{\varepsilon^2\mathbf{m}_C\mathbf{m}_C^{\mathrm{T}}}{\sum_{\alpha=1}^N w_\alpha^2 \sin^2\phi_\alpha/\sin^2\theta_\alpha}. \qquad (98)$$

From this, we have the following observations:

- The error in the focus of expansion is most likely to occur along the center line of the trajectories.
- The error in the focus of expansion is approximately proportional to the image accuracy ε.
- The error in the focus of expansion is approximately proportional to $1/w_\alpha$ of the distance w_α between the corresponding feature points.
- The error in the focus of expansion is approximately proportional to $1/\sin\phi_\alpha$ of the deviation angles ϕ_α of the individual trajectories from the center line.
- The error in the focus of expansion is approximately proportional to $\sin\theta_\alpha$ of the disparities θ_α of the focus of expansion from the center points of the individual trajectories.

As in the case of vanishing points, the statistical bias of the focus of expansion is evaluated from Eq. (93) and Eq. (98) in the form of Eq. (88). Hence, we observe that

- the focus of expansion is statistically biased along the center line *away* from the feature points.
- The bias is small when the focus of expansion is close to the feature points or infinitely away from them.

6. RENORMALIZATION

Although the statistical bias of the vanishing point or the focus of expansion computed as the intersection of concurrent lines can be removed by using the unbiased estimation given by Propositions 4 and 5, two problems arise. First, the optimal weights given by Eq. (43) involve the N-vector that we want to compute. Second, the computation of the unbiased moment matrix requires knowledge of the covariance matrices of the data points or lines, which involve image noise characteristics. However, noise characteristics are difficult to predict a priori for real images in real environments.

The first problem can be resolved by iterations in principle, as pointed out earlier. For the second problem, we take advantage of the fact that, from Eqs. (78) and (91), the covariance matrix $V[\mathbf{n}_\alpha]$ can be expressed in the form

$$V[\mathbf{n}_\alpha] = cV_0[\mathbf{n}_\alpha], \qquad (99)$$

where c is an unknown constant that characterizes the magnitude of image noise, while $V_0[\mathbf{n}_\alpha]$ has a known form.

[4] Again, the remark given at the end of Section 5.1 applies.

[5] Strictly speaking, this is not true if noise is due to camera characteristics such as distortion of the lens or distortion of the raster scanning. For simplicity, we ignore such systematic errors, assuming that noise is introduced randomly and independently by image operations.

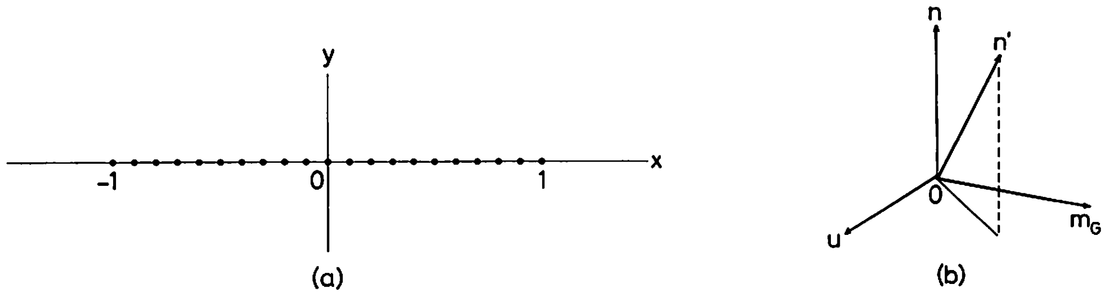

FIG. 8. (a) Simulated edge pixels, and (b) representation of a perturbed N-vector.

Since multiplication of the optimal weights by a constant does not affect the resulting solution, Eq. (43) can be replaced by

$$W_\alpha = \frac{1}{(\mathbf{m}, V_0[\mathbf{n}_\alpha]\mathbf{m})}. \tag{100}$$

Ideally, the constant c should be chosen so that $E[\hat{\mathbf{N}}] = \bar{\mathbf{N}}$, but this is impossible unless the statistical characteristics of the noise are known. On the other hand, if $E[\hat{\mathbf{N}}] = \bar{\mathbf{N}}$, we have

$$E[(\bar{\mathbf{m}}, \hat{\mathbf{N}}\bar{\mathbf{m}})] = (\bar{\mathbf{m}}, E[\hat{\mathbf{N}}]\bar{\mathbf{m}}) = (\bar{\mathbf{m}}, \bar{\mathbf{N}}\bar{\mathbf{m}}) = 0, \tag{101}$$

because from Eq. (37) $J = (\mathbf{m}, \mathbf{N}\mathbf{m})$ takes its absolute minimum 0 for the exact solution $\bar{\mathbf{m}}$ in the absence of noise. This suggests that we require that $(\mathbf{m}, \hat{\mathbf{N}}\mathbf{m}) = 0$ at each iteration step. If $(\mathbf{m}, \hat{\mathbf{N}}\mathbf{m}) \neq 0$ for the current estimates c and $\mathbf{m}$, we define

$$\hat{\mathbf{N}}' = \hat{\mathbf{N}} - \frac{(\mathbf{m}, \hat{\mathbf{N}}\mathbf{m})}{\sum_{\beta=1}^N W_\beta(\mathbf{m}, V_0[\mathbf{n}_\beta]\mathbf{m})} \sum_{\alpha=1}^N W_\alpha V_0[\mathbf{n}_\beta]. \tag{102}$$

Then, $(\mathbf{m}, \hat{\mathbf{N}}'\mathbf{m}) = 0$. Note that $(\mathbf{m}, \hat{\mathbf{N}}\mathbf{m})$ equals the smallest eigenvalue of $\hat{\mathbf{N}}$. From this observation, we obtain the following procedure, which we call *renormalization*:

renormalization $(\{\mathbf{n}_\alpha\}, \{V_0[\mathbf{n}_\alpha]\})$

1. Let $c = 0$ and $W_\alpha = 1$, $\alpha = 1, \ldots, N$.
2. Compute the unit eigenvector $\mathbf{m}$ of

$$\hat{\mathbf{N}} = \sum_{\alpha=1}^N W_\alpha(\mathbf{n}_\alpha\mathbf{n}_\alpha^\top - cV_0[\mathbf{n}_\alpha]) \tag{103}$$

for the smallest eigenvalue, and let λ_m be the smallest eigenvalue.

3. Update c and W_α as follows:

$$c \leftarrow c + \frac{\lambda_m}{\sum_{\alpha=1}^N W_\alpha(\mathbf{m}, V_0[\mathbf{n}_\alpha]\mathbf{m})}, \tag{104}$$

$$W_\alpha \leftarrow \frac{1}{(\mathbf{m}, V_0[\mathbf{n}_\alpha]\mathbf{m})}. \tag{105}$$

4. Return $\mathbf{m}$ if the update has converged; else go back to Step 2.

Iterations of renormalization usually converge very rapidly. For most applications, three or four iterations are sufficient. See Appendix D for further discussions on renormalization.

7. RANDOM NUMBER SIMULATIONS

7.1. *Line Fitting to an Edge*

Figure 8a shows simulated edge pixels. The focal length is set to $f = 20$. To the x and y coordinates of each pixel is added noise obeying an independent normal distribution of mean 0 and standard deviation 0.05. Then, a line is fitted by the optimal least-squares method of Section 4.1 1000 times, each time using different random numbers. The computed N-vector of the fitted line is orthogonally projected onto the plane spanned by $\mathbf{u}$ and $\mathbf{m}_G$ defined by Eqs. 68 (Fig. 8b). Then, the mean and covariance matrix of the 1000 samples are computed. Fig. 9a shows the ellipse centered at the sample mean indicating the root mean square in each direction. The horizontal axis corresponds to the edge orientation $\mathbf{u}$. Figure 9b shows the corresponding theoretical prediction computed by Eq. (78). Thus, Eq. (78) is a very good approximation. As compared with the variation in the direction of $\mathbf{u}$, the variation in the direction of $\mathbf{m}_G$ is very small and hence can be ignored, meaning that the error in $\mathbf{n}$ almost always

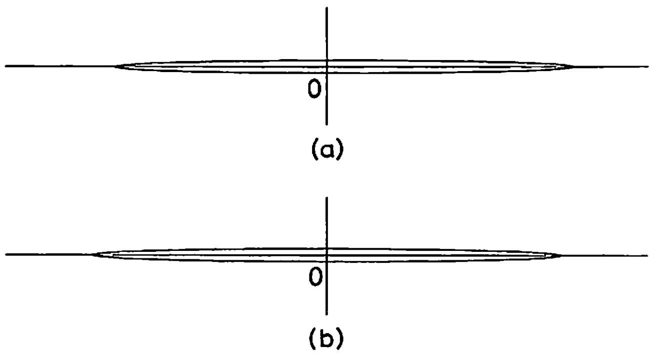

FIG. 9. Line fitting: (a) sample distribution and (b) theoretical distribution.

occurs in the direction of $\mathbf{u}$. Thus, Eq. (79) is also a very good approximation, and the observations made in Section 4.1 are confirmed.

7.2. *Focus of Expansion Estimation*

Since vanishing point estimation and focus of expansion estimation have the same geometric structure (although the source of noise is different), consider the focus of expansion estimation. Figure 10a shows simulated trajectories. The disparity of the focus of expansion from the image origin is $\theta = 45°$, and the deviation angle of the outermost trajectories from the center line is $\phi = 3°$. The focal length is set to $f = 20$. To the x and y coordinates of each endpoint is added noise obeying an independent normal distribution of mean 0 and standard deviation 0.005.

The common intersection is computed by the optimal least-squares method described in Section 4.1 1000 times, each time using different random numbers. The computed N-vector of the estimated focus of expansion is orthogonally projected onto the plane spanned by $\mathbf{m}_C$ and $\mathbf{n}_G$ as defined in Section 5.3 (Fig. 10b). Then, the mean and covariance matrix of the 1000 samples are computed. Figure 11a shows the ellipse centered at the sample mean indicating the root mean square in each direction. The horizontal axis corresponds to the orientation of $\mathbf{m}_C$. Figure 11b shows the corresponding theoretical prediction computed by Eq. (94). Evidently, the statistical bias described exists. The black circle in Fig. 11a indicates the expected bias predicted by Theorem 4. Figure 11c shows the result obtained by applying the renormalization described in Section 6. The computation converges after three or four iterations. The bias is indeed removed without any knowledge of noise characteristics, and the resulting distribution is well characterized by Eq. (94).

As compared with the variation in the direction of $\mathbf{m}_C$, the variation in the direction of $\mathbf{n}_G$ is very small and hence can be ignored, meaning that the error in $\mathbf{m}$ almost always occurs in the direction of $\mathbf{m}_C$. Thus, the approximation of Eq. (98) can be justified.

Figure 12a shows the root mean square error in the N-vector $\mathbf{m}$ for various values of the disparity θ. The solid line indicates the theoretically predicted values computed by Eq. (94), and black circles indicate empirical values, each computed from 1000 samples. It can be confirmed that the error is indeed approximately proportional to $\sin\theta$. Figure 12b plots the root mean square error for various values of the deviation angle ϕ, each computed after 1000 trials. It can be observed that the error is approximately proportional to $1/\sin\phi$, as expected. Thus, the observations in Section 5.3 are confirmed. In Figs. 12a and 12b, white circles indicate the values computed by uniform weights W_α = const. It is clearly seen that the use of optimal weights produces a better result.

In all the experiments, the effects of the effective focal length discussed in Section 2.4 are negligibly small, so we need not introduce the effective focal length, as expected.

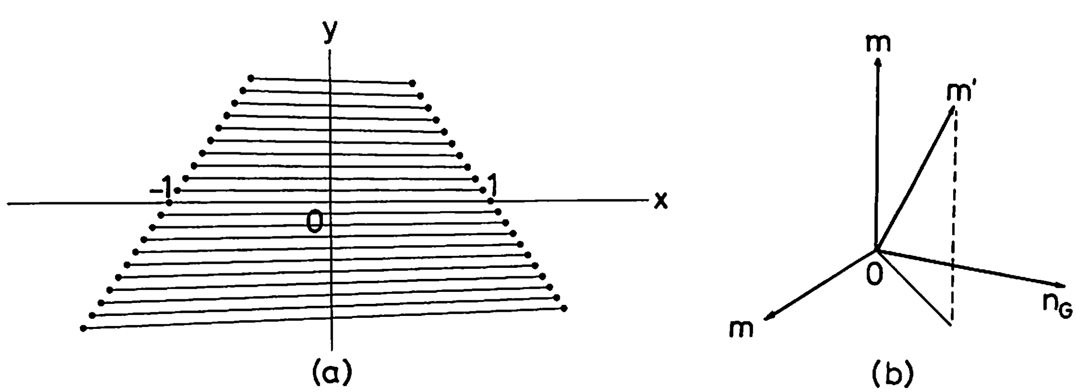

FIG. 10. (a) Simulated trajectories and (b) representation of a perturbed N-vector.

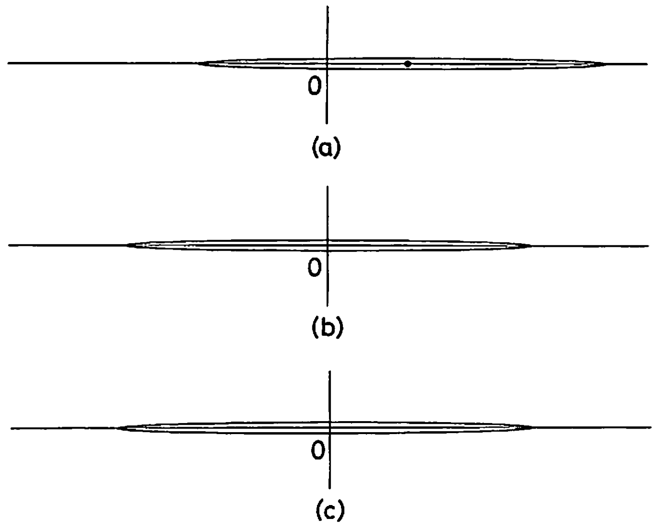

FIG. 11. Focus of expansion estimation: (a) sample distribution and theoretically predicted bias, (b) theoretical distribution, and (c) sample distribution obtained by renormalization.

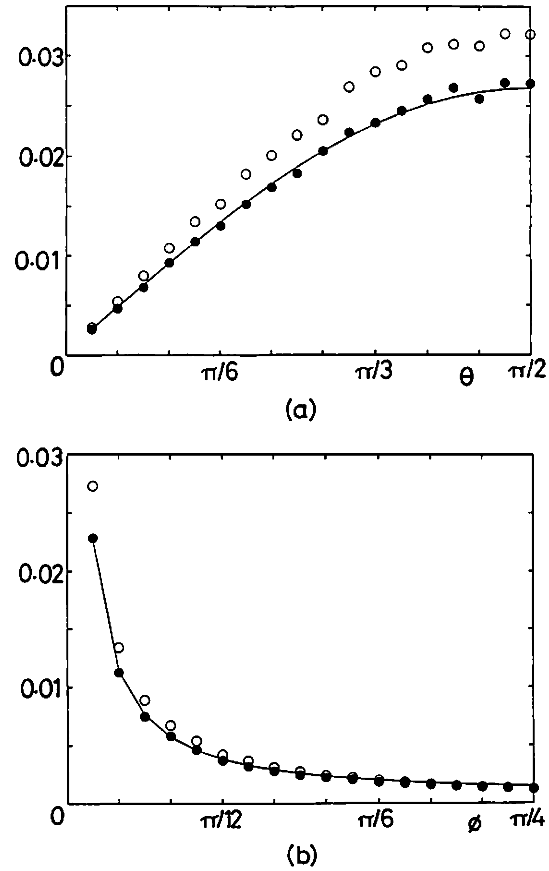

FIG. 12. Root mean square error vs (a) disparity angle θ and (b) deviation angle ϕ.

8. CONCLUDING REMARKS

In this paper, we have studied the statistical behavior of errors involved in fundamental geometric computations in a general "application independent" framework. Following the theory of "computational projective geometry" of Kanatani [11], which is also general and application independent, we first presented a statistical model of noise in terms of the covariance matrix of the N-vector of a pixel. Using this model, we computed the covariance matrices of N-vectors of lines and their intersections. Then, we determined the optimal weights for the least-squares optimization and computed the covariance matrix of the resulting optimal estimate. The result was applied to line fitting and computation of vanishing points and focuses of expansion. In each of these problems, characteristic error behavior was described in detail. We also pointed out that statistical biases exist in such computations and presented a scheme called *renormalization*, which removes the bias by automatically adjusting to noise. Finally, random number simulations were conducted to confirm our analysis.

Our statistical model of noise is idealized in many respects, so it may not always describe noise behavior of real systems completely: images may have a systematic (i.e., highly correlated) noise due to optical or electronic distortions; the edge operators we use may detect edges in systematically biased locations. In general, noise in real systems usually has components that cannot be accounted for by such mathematical idealizations as homogeneity, isotropy, independence, and Gaussian. Hence, predictions based on such an idealized model are usually *underestimations* of errors.

However, such theoretical predictions can be of great value if they are properly used. For example, one can design the system so that theoretically predicted errors are removed or the reliability of computation is maximized, and the analysis is successful if the performance of the resulting system is improved by it. In order to pursue such applications effectively, one must avoid introducing unnecessary complications into the theory. If the improvement of the performance is not sufficient, this very fact sheds light on the role of the sources of noise that are not accounted for by the theory, and a new theory will emerge, and so on. Such practical applications are discussed in [12–15, 17, 18], to which this paper provides a theoretical foundation.

APPENDIX

A. *Covariance Matrices of Normalized Vectors*

Let $\mathbf{h}$ be a unit vector. From Fig. A1, we see that

$$\mathbf{P_h} = \mathbf{I} - \mathbf{h}\mathbf{h}^T \tag{106}$$

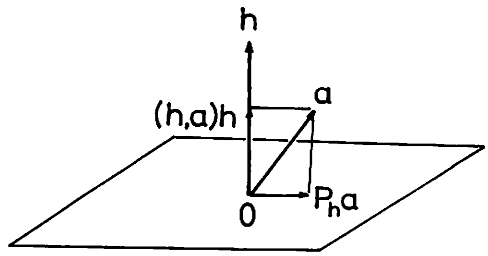

FIG. A1. Projection matrix.

projects a vector $\mathbf{a}$ onto a plane perpendicular to $\mathbf{h}$:

$$\mathbf{P_h a} = \mathbf{a} - (\mathbf{a}, \mathbf{h})\mathbf{h}. \tag{107}$$

The matrix $\mathbf{P_h}$ is called the *projection matrix* along $\mathbf{h}$. It is easy to confirm the following:

$$\mathbf{P}_{\mathbf{h}}^{\mathrm{T}} = \mathbf{P_h}, \qquad \mathbf{P}_{\mathbf{h}}^{2} = \mathbf{P_h}. \tag{108}$$

LEMMA A.1. *To a first approximation in* $\Delta\mathbf{a}$,

$$N[\mathbf{a} + \Delta\mathbf{a}] = N[\mathbf{a}] + \frac{\mathbf{P_a}\Delta\mathbf{a}}{\|\mathbf{a}\|}. \tag{109}$$

Proof. Vector $\Delta\mathbf{a}$ is decomposed into $\Delta\mathbf{a} = \Delta\mathbf{a}_1 + \Delta\mathbf{a}_2$, where $\Delta\mathbf{a}_1$ is parallel to $\mathbf{a}$ and $\Delta\mathbf{a}_2$ is orthogonal to $\mathbf{a}$. Since $\mathbf{a} + \Delta\mathbf{a}_1$ is parallel to $\mathbf{a}$, we have $N[\mathbf{a} + \Delta\mathbf{a}_1] = N[\mathbf{a}]$. Since $\Delta\mathbf{a}_2$ is orthogonal to $\mathbf{a}$, vectors $\mathbf{a} + \Delta\mathbf{a}_2$ and $\mathbf{a}$ have the same norm $\|\mathbf{a} + \Delta\mathbf{a}_2\| = \|\mathbf{a}\|$ to a first approximation. Hence,

$$N[\mathbf{a} + \Delta\mathbf{a}_2] = \frac{\mathbf{a} + \Delta\mathbf{a}_2}{\|\mathbf{a} + \Delta\mathbf{a}_2\|} = \frac{\mathbf{a} + \Delta\mathbf{a}_2}{\|\mathbf{a}\|} = N[\mathbf{a}] + \frac{\Delta\mathbf{a}_2}{\|\mathbf{a}\|}. \tag{110}$$

The perturbation caused by $\Delta\mathbf{a}_1 + \Delta\mathbf{a}_2$ is the sum of the perturbations caused by $\Delta\mathbf{a}_1$ and $\Delta\mathbf{a}_2$ to a first approximation. Hence, we have $N[\mathbf{a} + \Delta\mathbf{a}] = N[\mathbf{a}] + \mathbf{a}_2/\|\mathbf{a}\|$. If we note that $\Delta\mathbf{a}_2 = \mathbf{P_a}\Delta\mathbf{a}$, we obtain the assertion. ■

PROPOSITION A.1. *If* $\mathbf{n} = N[\mathbf{a}]$, *then*

$$V[\mathbf{n}] = \frac{1}{\|\mathbf{a}\|^2}\mathbf{P_n}V[\mathbf{a}]\mathbf{P_n}. \tag{111}$$

Proof. The perturbation $\Delta\mathbf{a}$ causes a perturbation $\Delta\mathbf{n}$ of $\mathbf{n} = N[\mathbf{a}]$. From Eq. (109), we obtain to a first approximation

$$\Delta\mathbf{n} = \frac{\mathbf{P_n}\Delta\mathbf{a}}{\|\mathbf{a}\|}. \tag{112}$$

Hence,

$$V[\mathbf{n}] = E[\Delta\mathbf{n}\Delta\mathbf{n}^{\mathrm{T}}] = \frac{\mathbf{P_n}E[\Delta\mathbf{a}\Delta\mathbf{a}^{\mathrm{T}}]\mathbf{P}_{\mathbf{n}}^{\mathrm{T}}}{\|\mathbf{a}\|^2} = \frac{\mathbf{P_n}V[\mathbf{a}]\mathbf{P}_{\mathbf{n}}^{\mathrm{T}}}{\|\mathbf{a}\|^2}. \tag{113}$$

Since $\mathbf{P}_{\mathbf{n}}^{\mathrm{T}} = \mathbf{P_n}$, we obtain Eq. (23).

B. *Effective Focal Length*

PROPOSITION B.1.

$$E\left[f\frac{m_1 + \Delta m_1}{m_3 + \Delta m_3}\right] = x\left(1 + \frac{1}{2(1 + r^2/f^2)}\left(\frac{\varepsilon}{f}\right)^2 + O(\Delta\mathbf{m})^3\right),$$
$$E\left[f\frac{m_2 + \Delta m_2}{m_3 + \Delta m_3}\right] = y\left(1 + \frac{1}{2(1 + r^2/f^2)}\left(\frac{\varepsilon}{f}\right)^2 + O(\Delta\mathbf{m})^3\right). \tag{114}$$

Proof. Since

$$f\frac{m_1 + \Delta m_1}{m_3 + \Delta m_3} = f\left(\frac{1 + \Delta m_1/m_1}{1 + \Delta m_3/m_3}\right) = x\left(1 + \frac{\Delta m_1}{m_1} - \frac{\Delta m_3}{m_3} + \left(\frac{\Delta m_3}{m_3}\right)^2 - \left(\frac{\Delta m_1}{m_1}\right)\left(\frac{\Delta m_3}{m_3}\right) + O(\Delta\mathbf{m})^3\right), \tag{115}$$

its expectation is

$$E\left[f\frac{m_1 + \Delta m_1}{m_3 + \Delta m_3}\right] = x\left(1 + \frac{E[\Delta m_3^2]}{m_3^2} - \frac{E[\Delta m_1 \Delta m_3]}{m_1 m_3} + O(\Delta\mathbf{m})^3\right). \tag{116}$$

If we put $\mathbf{k} = (0, 0, 1)^{\mathrm{T}}$ and $\mathbf{i} = (1, 0, 0)^{\mathrm{T}}$, we have $m_3 = (\mathbf{m}, \mathbf{k})$ and $m_1 = (\mathbf{m}, \mathbf{i})$. Hence, $E[\Delta m_3^2] = (\mathbf{k}, V[\mathbf{m}]\mathbf{k})$ and $E[\Delta m_1 \Delta m_3] = (\mathbf{i}, V[\mathbf{m}]\mathbf{k})$. Since $\mathbf{u}$ is orthogonal to the Z-axis, we have $(\mathbf{u}, \mathbf{k}) = 0$. From Eq. (9), we have

$$(\mathbf{k}, V[\mathbf{m}]\mathbf{k}) = \frac{\tilde{\varepsilon}^2(\mathbf{v}, \mathbf{k})^2}{2(1 + r^2/f^2)^2},$$
$$(\mathbf{i}, V[\mathbf{m}]\mathbf{k}) = \frac{\tilde{\varepsilon}^2(\mathbf{v}, \mathbf{i})(\mathbf{v}, \mathbf{k})}{2(1 + r^2/f^2)^2}, \tag{117}$$

where we put $\tilde{\varepsilon} = \varepsilon/f$. In Eqs. (10), the signs of $\mathbf{u}$ and $\mathbf{v}$ are irrelevant, so we can choose $\mathbf{u} = \sqrt{1 + f^2/r^2}\mathbf{m} \times \mathbf{k}$ and $\mathbf{v} = \mathbf{u} \times \mathbf{m}$. From Fig. 2, we observe that

$$\cos\theta = m_3 = \frac{1}{\sqrt{1 + r^2/f^2}},$$
$$\sin\theta = \sqrt{1 - m_3^2} = \frac{r/f}{\sqrt{1 + r^2/f^2}}. \tag{118}$$

Hence,

$$(\mathbf{v}, \mathbf{k}) = (\mathbf{u} \times \mathbf{m}, \mathbf{k}) = (\mathbf{m} \times \mathbf{k}, \mathbf{u})$$
$$= \sqrt{1 + \frac{f^2}{r^2}}(\mathbf{m} \times \mathbf{k}, \mathbf{m} \times \mathbf{k})$$
$$= \frac{\sqrt{1 + r^2/f^2}}{r/f}(1 - (\mathbf{m}, \mathbf{k})^2) = \frac{r/f}{\sqrt{1 + r^2/f^2}}$$
$$= \frac{r}{f} m_3, \tag{119}$$

$$(\mathbf{v}, \mathbf{i}) = (\mathbf{u} \times \mathbf{m}, \mathbf{i}) = (\mathbf{m} \times \mathbf{i}, \mathbf{u})$$
$$= \sqrt{1 + \frac{f^2}{r^2}}(\mathbf{m} \times \mathbf{i}, \mathbf{m} \times \mathbf{k})$$
$$= -\frac{\sqrt{1 + r^2/f^2}}{r/f}(\mathbf{m}, \mathbf{k})(\mathbf{m}, \mathbf{i}) = -\frac{f}{r} m_1. \tag{120}$$

Thus,

$$(\mathbf{k}, V[\mathbf{m}]\mathbf{k}) = \frac{\bar{\varepsilon}^2 r^2 m_3^2/f^2}{2(1 + r^2/f^2)^2},$$
$$(\mathbf{i}, V[\mathbf{m}]\mathbf{k}) = -\frac{\bar{\varepsilon}^2 m_1 m_3}{2(1 + r^2/f^2)^2}. \tag{121}$$

Hence,

$$E\left[f\frac{m_1 + \Delta m_1}{m_3 + \Delta m_3}\right] = x\left(1 + \frac{\bar{\varepsilon}^2/2}{1 + r^2/f^2} + O(\Delta\mathbf{m})^3\right). \tag{122}$$

The expression of $E[f(m_2 + \Delta m_2)/(m_3 + \Delta m_3)]$ is obtained similarly. ∎

PROPOSITION B.2. *If*

$$\mathbf{m} = N\left[\begin{pmatrix} x \\ y \\ \hat{f} \end{pmatrix}\right], \qquad \hat{f} = f\left(1 + \frac{\bar{\varepsilon}^2/2}{1 + r^2/f^2}\right), \tag{123}$$

then

$$E\left[f\frac{m_1 + \Delta m_1}{m_3 + \Delta m_3}\right] = x + O(\Delta\mathbf{m})^3,$$
$$E\left[f\frac{m_2 + \Delta m_2}{m_3 + \Delta m_3}\right] = y + O(\Delta\mathbf{m})^3. \tag{124}$$

Proof. Proceeding in exactly the same way as in the proof of Proposition B.1, we obtain

$$E\left[f\frac{m_1 + \Delta m_1}{m_3 + \Delta m_3}\right] = f\frac{m_1}{m_3}\left(1 + \frac{\bar{\varepsilon}^2/2}{1 + r^2/f^2} + O(\Delta\mathbf{m})^3\right)$$
$$= \hat{f}\frac{m_1}{m_3}\left(\frac{f}{\hat{f}}\right)\left(1 + \frac{\bar{\varepsilon}^2/2}{1 + r^2/f^2} + O(\Delta\mathbf{m})^3\right) = x + O(\Delta\mathbf{m})^3. \tag{125}$$

The expression for y is obtained similarly. To be strict, the expression $1 + r^2/f^2$ in the above equation must be $1 + r^2/\hat{f}^2$. However, replacing it with $1 + r^2/f^2$ introduces only a difference of $O(\bar{\varepsilon}^4)$ in the final result. ∎

C. *Approximately Optimal Estimation*

PROPOSITION C.1. *If the N-vector* $\mathbf{m}$ *in the optimal weights is replaced by its estimate* $\mathbf{m} + \Delta\mathbf{m}$, *the resulting perturbation of the unit eigenvector of the moment matrix for the smallest eigenvalue is* $O(\Delta\mathbf{m})^2$.

Proof. If $\mathbf{m}$ is perturbed into $\mathbf{m} + \Delta\mathbf{m}$, the perturbation of the weight $W_\alpha = 1/(\mathbf{m}, V[\mathbf{n}_\alpha]\mathbf{m})$ is, to a first approximation,

$$\delta W_\alpha = -\frac{2(\Delta\mathbf{m}, V[\mathbf{n}_\alpha]\mathbf{m})}{(\mathbf{m}, V[\mathbf{n}_\alpha]\mathbf{m})^2}. \tag{126}$$

If $\|\Delta\mathbf{n}_\alpha\|$ is approximately ν in the sense of root mean square, the covariance matrix $V[\mathbf{n}_\alpha]$ is $O(\nu^2)$. According to Corollary 3, we can assume that $\|\Delta\mathbf{m}\| = O(\nu/\sqrt{N})$. Hence,

$$\delta W_\alpha = O\left(\frac{1}{\nu\sqrt{N}}\right). \tag{127}$$

Let $\mathbf{n}_\alpha + \Delta\mathbf{n}_\alpha$ be the observed N-vectors. Their moment matrix with perturbed weights is

$$\mathbf{N} + \Delta\mathbf{N} = \sum_{\alpha=1}^{N} (W_\alpha + \delta W_\alpha)(\mathbf{n}_\alpha + \Delta\mathbf{n}_\alpha)(\mathbf{n}_\alpha + \Delta\mathbf{n}_\alpha)^\mathrm{T}. \tag{128}$$

Let λ_u, λ_v, and 0 be the eigenvalues of the unperturbed moment matrix $\mathbf{N}$, and $\{\mathbf{u}, \mathbf{v}, \mathbf{m}\}$ be the orthonormal system of the corresponding eigenvectors. According to the perturbation theorem (Appendix D), the unit eigenvector of the perturbed moment matrix $\mathbf{N} + \Delta\mathbf{N}$ for eigenvalue 0 is perturbed in the form

$$\Delta\mathbf{m} = \frac{(\mathbf{u}, \Delta\mathbf{N}\mathbf{m})}{0 - \lambda_u}\mathbf{u} + \frac{(\mathbf{v}, \Delta\mathbf{N}\mathbf{m})}{0 - \lambda_v}\mathbf{v}. \tag{129}$$

Now,

$$\Delta\mathbf{N} = \sum_{\alpha=1}^{N} W_\alpha(\Delta\mathbf{n}_\alpha\mathbf{n}_\alpha^\mathrm{T} + \mathbf{n}_\alpha\Delta\mathbf{n}_\alpha^\mathrm{T}) + \sum_{\alpha=1}^{N} W_\alpha\Delta\mathbf{n}_\alpha\Delta\mathbf{n}_\alpha^\mathrm{T}$$
$$+ \sum_{\alpha=1}^{N} \delta W\alpha\mathbf{n}_\alpha\mathbf{n}_\alpha^\mathrm{T} + \sum_{\alpha=1}^{N} \delta W_\alpha(\Delta\mathbf{n}_\alpha\mathbf{n}_\alpha^\mathrm{T} + \mathbf{n}_\alpha\Delta\mathbf{n}_\alpha^\mathrm{T})$$
$$+ \sum_{\alpha=1}^{N} \delta W_\alpha\Delta\mathbf{n}_\alpha\Delta\mathbf{n}_\alpha^\mathrm{T}. \tag{130}$$

The first and second terms on the right-hand side do not depend on δW_α. Since $(\mathbf{n}_\alpha, \mathbf{m}) = 0$, the third term on the right-hand side does not contribute to $\Delta\mathbf{m}$. Hence, the component of the error of $\Delta\mathbf{m}$ contributed by the perturbation δW_α of the weight W_α is

$$\delta\mathbf{m} = \frac{\sum_{\alpha=1}^N \delta W_\alpha(O(1), \Delta\mathbf{n}_\alpha)}{\lambda_u}\mathbf{u} + \frac{\sum_{\alpha=1}^N \delta W_\alpha(O(1), \Delta\mathbf{n}_\alpha)}{\lambda_v}\mathbf{v}. \tag{131}$$

Since $\Delta\mathbf{n}_\alpha$ are independent and since $\mathbf{u}$ and $\mathbf{v}$ are orthogonal unit vectors, the mean square of $\delta\mathbf{m}$ is

$$E[\|\delta\mathbf{m}\|^2] = \frac{\sum_{\alpha=1}^N (\delta W_\alpha)^2(O(1), V[\mathbf{n}_\alpha]O(1))}{\lambda_u^2} + \frac{\sum_{\alpha=1}^N (\delta W_\alpha)^2(O(1), V[\mathbf{n}_\alpha]O(1))}{\lambda_v^2}. \tag{132}$$

Since $V[\mathbf{n}_\alpha] = O(\nu^2)$, $W_\alpha = O(\nu^{-2})$, $\lambda_u = O(N/\nu^2)$, and $\lambda_v = O(N/\nu^2)$, we have $\|\delta\mathbf{m}\| = O(\nu^3\delta W/\sqrt{N})$ in the sense of root mean square. If Eq. (127) is substituted, we obtain $\|\delta\mathbf{m}\| = O(\nu^2/N) = O(\Delta\mathbf{m})^2$. ∎

PROPOSITION C.2. *Let $V[\mathbf{m}]$ be the covariance matrix computed from the unperturbed moment matrix $\mathbf{N}$, and $\tilde{V}[\mathbf{m}]$ be the covariance matrix computed from a perturbed moment matrix $\tilde{\mathbf{N}}$. If $\Delta\mathbf{n}_\alpha \approx \nu$, then*

$$\tilde{V}[\mathbf{m}] - V[\mathbf{m}] = O\left(\frac{\nu^3}{N\sqrt{N}}\right). \tag{133}$$

Proof. Let $\mathbf{n}_\alpha + \Delta\mathbf{n}_\alpha$ be the observed N-vectors. Their moment matrix is

$$\mathbf{N} + \Delta\mathbf{N} = \sum_{\alpha=1}^N W_\alpha(\mathbf{n}_\alpha + \Delta\mathbf{n}_\alpha)(\mathbf{n}_\alpha + \Delta\mathbf{n}_\alpha)^{\mathrm{T}}. \tag{134}$$

Let λ_u, λ_v, and 0 be the eigenvalues of the unperturbed moment matrix $\mathbf{N}$, and $\{\mathbf{u}, \mathbf{v}, \mathbf{m}\}$ the orthonormal system of the corresponding eigenvectors. If $\mathbf{u} \to \mathbf{u} + \Delta\mathbf{u}$, $\mathbf{v} \to \mathbf{v} + \Delta\mathbf{v}$, $\mathbf{m} \to \mathbf{m} + \Delta\mathbf{m}$, $\lambda_u \to \lambda_u + \Delta\lambda_u$, and $\lambda_v \to \lambda_v + \Delta\lambda_v$ are the perturbations caused by the perturbation $\mathbf{N} \to \mathbf{N} + \Delta\mathbf{N}$, the covariance matrix $V[\mathbf{m}]$ is perturbed into

$$\begin{aligned}\tilde{V}[\mathbf{m}] &= \frac{(\mathbf{u} + \Delta\mathbf{u})(\mathbf{u} + \Delta\mathbf{u})^{\mathrm{T}}}{\lambda_u + \Delta\lambda_u} + \frac{(\mathbf{v} + \Delta\mathbf{v})(\mathbf{v} + \Delta\mathbf{v})^{\mathrm{T}}}{\lambda_v + \Delta\lambda_v} \\ &= V[\mathbf{m}] + \frac{\Delta\mathbf{u}\mathbf{u}^{\mathrm{T}} + \mathbf{u}\Delta\mathbf{u}^{\mathrm{T}}}{\lambda_u} + \frac{\Delta\mathbf{v}\mathbf{v}^{\mathrm{T}} + \mathbf{v}\Delta\mathbf{v}^{\mathrm{T}}}{\lambda_v} \\ &\quad - \frac{\Delta\lambda_u\mathbf{u}\mathbf{u}^{\mathrm{T}}}{\lambda_u^2} - \frac{\Delta\lambda\mathbf{v}\mathbf{v}^{\mathrm{T}}}{\lambda_v^2}\end{aligned} \tag{135}$$

to a first approximation. From the perturbation theorem (Appendix D), we have $\Delta\lambda_u = (\mathbf{u}, \Delta\mathbf{N}\mathbf{u})$, $\Delta\lambda_v = (\mathbf{v}, \Delta\mathbf{N}\mathbf{v})$ to a first approximation. Now, to a first approximation

$$\Delta\mathbf{N} = \sum_{\alpha=1}^N W_\alpha(\mathbf{n}_\alpha\Delta\mathbf{n}_\alpha^{\mathrm{T}} + \Delta\mathbf{n}_\alpha\mathbf{n}_\alpha^{\mathrm{T}}). \tag{136}$$

Hence,

$$\begin{aligned}\Delta\lambda_u &= 2\sum_{\alpha=1}^N W_\alpha(\mathbf{u}, \mathbf{n}_\alpha)(\mathbf{u}, \Delta\mathbf{n}_\alpha), \\ \Delta\lambda_v &= 2\sum_{\alpha=1}^N W_\alpha(\mathbf{v}, \mathbf{n}_\alpha)(\mathbf{v}, \Delta\mathbf{n}_\alpha).\end{aligned} \tag{137}$$

Since $\Delta\mathbf{n}_\alpha$ are independent, their variances are

$$E[\Delta\lambda_u^2] = E[\Delta\lambda_v^2] = \sum_{\alpha=1}^N W_\alpha^2(O(1), V[\mathbf{n}_\alpha]O(1)). \tag{138}$$

If $\|\Delta\mathbf{n}_\alpha\| \approx \nu$ in the sense of the root mean square, we have $V[\mathbf{n}_\alpha] = O(\nu^2)$ and $W_\alpha = O(\nu^{-2})$. Hence,

$$\Delta\lambda_u = O\left(\frac{\sqrt{N}}{\nu}\right), \Delta\lambda_v = O\left(\frac{\sqrt{N}}{\nu}\right) \tag{139}$$

is the sense of the root mean square. We also see from the perturbation theorem (Appendix D) that

$$\Delta\mathbf{u} = \frac{(\mathbf{m}, \Delta\mathbf{N}\mathbf{u})}{\lambda_u - 0}\mathbf{m} + \frac{(\mathbf{v}, \Delta\mathbf{N}\mathbf{u})}{\lambda_u - \lambda_v}\mathbf{v}. \tag{140}$$

From Eq. (136), we obtain

$$\Delta\mathbf{u} = \frac{\sum_{\alpha=1}^N W_\alpha(O(1), \Delta\mathbf{n}_\alpha)}{\lambda_u}\mathbf{m} + \frac{\sum_{\alpha=1}^N W_\alpha(O(1), \Delta\mathbf{n}_\alpha)}{\lambda_u - \lambda_v}\mathbf{v}. \tag{141}$$

Since $\Delta\mathbf{n}_\alpha$ are independent and since $\mathbf{m}$ and $\mathbf{v}$ are orthogonal unit vectors, we have

$$E[\|\Delta\mathbf{u}\|^2] = \frac{\sum_{\alpha=1}^N W_\alpha^2(O(1), V[\mathbf{n}_\alpha]O(1))}{\lambda_u^2} + \frac{\sum_{\alpha=1}^N W_\alpha^2(O(1), V[\mathbf{n}_\alpha]O(1))}{(\lambda_u - \lambda_v)^2}. \tag{142}$$

Since $V[\mathbf{n}_\alpha] = O(\nu^2)$, $W_\alpha = O(\nu^{-2})$, $\lambda_u = O(N/\nu^2)$, and $\lambda_v = O(N/\nu^2)$, the above expression is $O(\nu^2/N)$. The same applies to $\mathbf{v}$. Thus,

$$\|\Delta\mathbf{u}\| = O\left(\frac{\nu}{\sqrt{N}}\right), \qquad \|\Delta\mathbf{v}\| = O\left(\frac{\nu}{\sqrt{N}}\right) \tag{143}$$

in the sense of the root mean square. Substituting Eqs. (139) and (127), $\lambda_u = O(N/\nu^2)$, and $\lambda_v = O(N/\nu^2)$ into Eqs. (135), we obtain Eq. (133). ∎

D. *Convergence of Renormalization*

The following *perturbation theorem* is well known in quantum physics (e.g., see [16]).

PROPOSITION D.1. *Let* $\mathbf{A}$ *be an n-dimensional symmetric matrix having eigenvalues* $\lambda_1, \ldots, \lambda_n$ *with* $\{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ *the corresponding eigenvectors forming an orthonormal system. If matrix* $\mathbf{A}$ *is perturbed into*

$$\mathbf{A}' = \mathbf{A} + \delta\mathbf{A}, \tag{144}$$

each eigenvalue λ_i *is perturbed into*

$$\lambda_i' = \lambda_i + (\mathbf{u}_i, \delta\mathbf{A}\mathbf{u}_i) + O(\delta\mathbf{A})^2. \tag{145}$$

If eigenvalue λ_i *is a simple root, the corresponding eigenvector* $\mathbf{u}_i$ *is perturbed into*

$$\mathbf{u}_i' = \mathbf{u}_i + \sum_{k\neq i} \frac{(\mathbf{u}_k, \delta\mathbf{A}\mathbf{u}_i)}{\lambda_i - \lambda_k}\mathbf{u}_k + O(\delta\mathbf{A})^2. \tag{146}$$

The procedure for renormalization can be expressed in an abstract form as follows. What we want to compute is the unit eigenvector $\bar{\mathbf{u}}_m$ of a positive semi-definite matrix $\bar{\mathbf{A}}$ for eigenvalue 0. The exact value of $\bar{\mathbf{A}}$ is unknown, but from a statistical error analysis we know that

$$\bar{\mathbf{A}} = E[\mathbf{A} - c\mathbf{B}], \tag{147}$$

where $\mathbf{A}$ and $\mathbf{B}$ are symmetric matrices we can compute from image data while c is an unknown constant characterizing the behavior of image noise. Matrices $\mathbf{A}$ and $\mathbf{B}$ are random variables since they are computed from data, while c has a definite value determined by the statistical model of noise. Hence, if we put

$$\hat{\mathbf{A}} = \mathbf{A} - c\mathbf{B}, \tag{148}$$

and if we can choose c such that $E[\hat{\mathbf{A}}] = \bar{\mathbf{A}}$, the unit eigenvector $\mathbf{u}_m$ of $\hat{\mathbf{A}}$ for the smallest eigenvalue is an unbiased estimate of $\bar{\mathbf{u}}_m$. However, we cannot do this unless we know the noise characteristics. On the other hand, if $E[\hat{\mathbf{A}}] = \bar{\mathbf{A}}$, then

$$E[(\bar{\mathbf{u}}_m, \hat{\mathbf{A}}\bar{\mathbf{u}}_m)] = (\bar{\mathbf{u}}_m, E[\hat{\mathbf{A}}]\bar{\mathbf{u}}_m) = (\bar{\mathbf{u}}_m, \bar{\mathbf{A}}\bar{\mathbf{u}}_m) = 0. \tag{149}$$

So, we attempt to compute a unit vector $\mathbf{u}_m$ such that $(\mathbf{u}_m, \hat{\mathbf{A}}\mathbf{u}_m) = 0$ in each iteration step. If $(\mathbf{u}_m, \hat{\mathbf{A}}\mathbf{u}_m) \neq 0$ for the current estimate $\mathbf{u}_m$ and c, we define

$$\hat{\mathbf{A}}' = \hat{\mathbf{A}} - \frac{(\mathbf{u}_m, \hat{\mathbf{A}}\mathbf{u}_m)}{(\mathbf{u}_m, \mathbf{B}\mathbf{u}_m)}\mathbf{B}. \tag{150}$$

Then, $(\mathbf{u}_m, \hat{\mathbf{A}}'\mathbf{u}_m) = 0$. Note that $(\mathbf{u}_m, \hat{\mathbf{A}}'\mathbf{u}_m)$ equals the smallest eigenvalue λ_m of $\hat{\mathbf{A}}$.

If $\mathbf{u}_m$ and c are the converged values, we have $\lambda_m = (\mathbf{u}_m, \hat{\mathbf{A}}\mathbf{u}_m) = 0$. This does not necessarily ensure that $\mathbf{u}_m$ coincides with $\bar{\mathbf{u}}_m$. In other words, $(\mathbf{u}_m, \bar{\mathbf{A}}\mathbf{u}_m)$ is not necessarily 0, because the current image data are not necessarily "typical" (i.e., may not be a good representative of the statistical ensemble). However, we can expect that $\mathbf{u}_m$ is a good approximation with a high probability.

If $\mathbf{u}_m$ is the unit eigenvector of the current $\hat{\mathbf{A}}$ for the smallest eigenvalue λ_m, matrix $\hat{\mathbf{A}}$ is updated at the next step to

$$\hat{\mathbf{A}}' = \hat{\mathbf{A}} - \frac{\lambda_m}{(\mathbf{u}_m, \mathbf{B}\mathbf{u}_m)}\mathbf{B}. \tag{151}$$

According to the perturbation theorem, the smallest eigenvalue λ_m' of $\hat{\mathbf{A}}'$ is

$$\lambda_m' = \lambda_m - \left(\mathbf{u}_m, \left(\frac{\lambda_m}{(\mathbf{u}_m, \mathbf{B}\mathbf{u}_m)}\mathbf{B}\right)\mathbf{u}_m\right) + O\left(\frac{\lambda_m}{(\mathbf{u}_m, \mathbf{B}\mathbf{u}_m)}\mathbf{B}\right)^2 = O(\lambda_m^2). \tag{152}$$

This means that λ_m converges to 0 *quadratically* just like Newton iterations.

Let λ_1 and λ_2 be, respectively, the largest and the second largest eigenvalues of $\hat{\mathbf{A}}$, and $\mathbf{u}_1$ and $\mathbf{u}_2$ the corresponding unit eigenvectors. According to the perturbation theorem, the unit eigenvector $\mathbf{u}_m'$ of $\hat{\mathbf{A}}'$ at the next step for the smallest eigenvalue λ_m' is

$$\mathbf{u}_m' = \mathbf{u}_m + \frac{\lambda_m}{(\mathbf{u}, \mathbf{B}\mathbf{u})}\sum_{i=1,2}\frac{(\mathbf{u}_i, \mathbf{B}\mathbf{u}_m)}{\lambda_m - \lambda_i}\mathbf{u}_i + O(\lambda_m^2) = \mathbf{u}_m + O(\lambda_m). \tag{153}$$

Since λ_m converges to 0 quadratically, the convergence of $\mathbf{u}_m$ is also quadratic. If the optimal weights W_α are computed by using the current eigenvector $\mathbf{u}_m$, the convergence of the W_α is no longer quadratic. However, the convergence is usually very rapid, and three or four iterations are sufficient for most cases.

It should be emphasized that the converged values are not necessarily the *exact* values, because the noise is random and unpredictable: the purpose of renormalization is to remove statistical bias (not completely, though). As shown in Appendix C, if each datum has an independent error of root-mean-square magnitude ν and if the number of data is N, the optimal unbiased estimate has an error of root-mean-square magnitude $O(\nu/\sqrt{N})$, which is the lowest bound that can be achieved.

ACKNOWLEDGMENTS

The author thanks Michael Brady, Andrew Zisserman, and Andrew Blake of the University of Oxford and Azriel Rosenfeld, Yiannis Aloimonos, and Rama Chellappa of the University of Maryland for various discussions on statistical aspects of geometric computation. He also thanks Kouji Urasawa of Oki Electric Industries for helping with the numerical experiments.

REFERENCES

1. S. D. Blostein and T. S. Huang, Error analysis in stereo determination of 3-D point positions, *IEEE Trans. Pattern Anal. Mach. Intell.* **9**, 1987, 752–765. [Corrected in **10**, 1988, 765.]
2. B. Brillault-O'Mahony, New method for vanishing point detection, *CVGIP: Image Understanding* **54**, 1991, 289–300.
3. T. J. Broida, S. Chandrashekhar, and R. Chellappa, Recursive estimation of 3-D motion from a monocular image sequence, *IEEE Trans. Aerospace Electron. Systems* **26**, 1990, 639–656.
4. T. J. Broida and R. Chellappa, Estimation of object motion parameters from noisy images, *IEEE Trans. Pattern Anal. Mach. Intell.* **8**, 1986, 90–99.
5. T. J. Broida and R. Chellappa, Performance bounds for estimating three-dimensional motion parameters from a sequence of noisy images, *J. Opt. Soc. Amer. A* **6**, 1989, 879–998.
6. R. T. Collins and R. S. Weiss, Vanishing point calculation as a statistical data, *Proceedings, Third International Conference on Computer Vision, Osaka, Japan, Dec. 1990*, pp. 400–403.
7. B. Kamgar-Parsi and R. D. Eastman, Calibration of a stereo system with small relative angles, *Comput. Vision, Graphics, Image Process.* **51**, 1990, 1–19.
8. B. Kamgar-Parsi and B. Kamgar-Parsi, Evaluation of quantization error in computer vision, *IEEE Trans. Pattern Anal. Mach. Intell.* **11**, 1989, 929–940.
9. B. Kamgar-Parsi, B. Kamgar-Parsi, and N. Netanyahu, A nonparametric method for fitting a straight line to a noisy image, *IEEE Trans. Pattern Anal. Mach. Intell.* **11**, 1989, 998–1001.
10. K. Kanatani, Errors of the incremental method for curves, *Comput. Vision, Graphics, Image Process.* **26**, 1984, 130–133.
11. K. Kanatani, Computational projective geometry, *CVGIP: Image Understanding* **54**, 1991, 333–448.
12. K. Kanatani, Hypothesizing and testing geometric properties of image data, *CVGIP: Image Understanding* **54**, 1991, 349–357.
13. K. Kanatani, Statistical analysis of focal-length calibration using vanishing points, *IEEE Trans. Robotics Automation* **8**, No. 6, 1992, 767–775.
14. K. Kanatani, Unbiased estimation and statistical analysis of 3-D rigid motion from two views, *IEEE Trans. Pattern Anal. Mach. Intell.* **15**, No. 1 1993, 37–50.
15. K. Kanatani, Statistical bias of conic fitting and renormalization, *IEEE Trans. Pattern Anal. Mach. Intell.*, **16**, No. 3, 1994.
16. K. Kanatani, *Geometric Computation for Machine Vision*, Oxford Univ. Press, Oxford, 1993.
17. K. Kanatani, Statistical foundation for hypothesis testing of image data, *CVGIP: Image Understanding*, to appear.
18. K. Kanatani and Y. Onodera, Anatomy of camera calibration using vanishing points, *IEICE Trans. Inform. and Systems* **74**, No. 10, 1991, 3369–3378.
19. J. K. Kearney, W. B. Thompson, and D. L. Boley, Optical flow estimation: An error analysis of gradient-based methods with local optimization, *IEEE Trans. Pattern Anal. Mach. Intell.* **9**, 1987, 229–244.
20. S. Lee and Y. Kay, A Kalman filter approach for accurate 3D motion estimation from sequence of stereo images, *CVGIP: Image Understanding* **54**, 1991, 244–258.
21. L. Matthies, T. Kanade, and R. Szeliski, Kalman filter-based algorithms for estimating depth from image sequences, *Internat. J. Comput. Vision* **3**, 1989, 209–236.
22. J. Porrill, Fitting ellipses and predicting confidence envelopes using a bias corrected Kalman filter, *Image Vision Comput.* **8**, 1990, 37–41.
23. M. A. Snyder, The precision of 3-D parameters in correspondence based techniques: The case of uniform translation motion in a rigid environment, *IEEE Trans. Pattern Anal. Mach. Intell.* **11**, 1989, 523–528.
24. I. Weiss, Line fitting in a noisy images, *IEEE Trans. Pattern Anal. Mach. Intell.* **11**, 1989, 325–329.
25. 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.
26. G.-S. J. Young and R. Chellappa, 3-D motion estimation using a sequence of noisy stereo images: Models, estimation, and uniqueness results, *IEEE Trans. Pattern Anal. Mach. Intell.* **12**, 1990, 735–759.
27. Y. T. Zhou, V. Venkateswar, and R. Chellappa, Edge detection and linear feature extraction using a 2-D random field model, *IEEE Trans. Pattern Anal. Mach. Intell.* **11**, 1989, 84–95