PAPER

Optimal Line Fitting and Reliability Evaluation

Yasushi KANAZAWA† *and* **Kenichi KANATANI**††, *Members*

SUMMARY Introducing a mathematical model of image noise, we formalize the problem of fitting a line to point data as statistical estimation. It is shown that the reliability of the fitted line can be evaluated quantitatively in the form of the covariance matrix of the parameters. We present a numerical scheme called *renormalization* for computing an optimal fit and at the same time evaluating its reliability. We also present a scheme for visualizing the reliability of the fit by means of the *primary deviation pair* and derive an analytical expression for the reliability of a line fitted to an edge segment by using an *asymptotic approximation*. Our method is illustrated by showing simulations and real-image examples.

key words: line fitting, statistical model of noise, optimization technique, image processing, reliability evaluation, edge detection

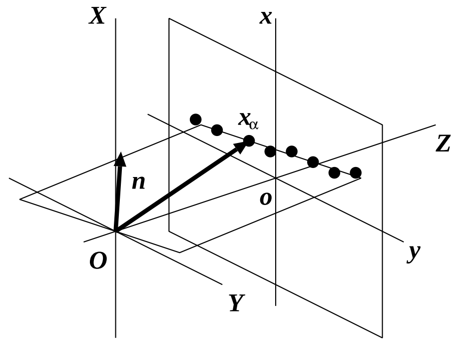


Fig. 1 Line fitting.

1. Introduction

Line fitting is one of the most important processes in image understanding, because boundaries of many objects in indoor environments are projected as straight lines in an image. They can be detected by an edge operator as *edge segments*, i.e., sequences of pixels. By fitting lines to the detected linear edge segments, we can give the objects in the image a 2-D representation (or "line drawing"), which becomes the basis of 3-D interpretation of the scene. For example, the 3-D shapes and locations of objects are inferred from "vanishing points" and "focuses of expansion" computed as the intersections of the fitted lines [5]. In the past, various line fitting techniques have been proposed [3], [13], [15]–[17], but a main concern has been obtaining an *accurate fit*. Equally important, however, is the *evaluation* of the reliability of the resulting fit. Since errors in line fitting propagate to the final 3-D interpretation, its reliability can be evaluated if the reliability of line fitting can be quantitatively evaluated. This type of analysis has been done only in an ad hoc manner in the past [1], [4], [6], [8].

In this paper, we formulate the line fitting problem as *statistical estimation* by introducing a mathematical model of image noise and derive a theoretically optimal fitting scheme in the sense of *maximum likelihood estimation*. Then, we give an *explicit* expression that evaluates the reliability of the computed fit in quantitative terms and propose a scheme for *visualizing* the reliability of the fit by means of the *primary deviation pair*. We also present a simple computational scheme called *renormalization* for computing an optimal fit and at the same time evaluating its reliability. We illustrate our method by showing numerical simulations and real-image examples. Finally, we derive an analytical expression for the reliability of the line fitted to a dense edge pixels by using asymptotic approximation.

2. Line Fitting Problem

Let $\{(x_\alpha, y_\alpha)\}$, $\alpha = 1, ..., N$, be a sequence of points to which a line is to be fitted. In real circumstances, digital images are not ideal, and image processing operations such as edge detection and template matching may not be accurate. We refer to such inaccuracy, irrespective of its sources, collectively as "image noise". Let $(\bar{x}_\alpha, \bar{y}_\alpha)$ be the true position of point (x_α, y_α), i.e., the position that would supposedly be observed if the image were ideal and the detection operation were accurate. We want to obtain a line $Ax + By + C = 0$ that passes through $(\bar{x}_\alpha, \bar{y}_\alpha)$. Namely, we want to compute A, B, and C such that

$$A\bar{x}_\alpha + B\bar{y}_\alpha + C = 0, \quad \alpha = 1, ..., N. \tag{1}$$

Define three-dimensional vectors

$$\boldsymbol{x}_\alpha = \begin{pmatrix} x_\alpha \\ y_\alpha \\ 1 \end{pmatrix}, \quad \boldsymbol{n} = \begin{pmatrix} A \\ B \\ C \end{pmatrix}. \tag{2}$$

Manuscript received Octorber 27, 1995.
Manuscript revised February 28, 1996.

†The author is with the Department of Information and Computer Engineering, Gunma College of Technology, Maebashi-shi, 371 Japan.
††The author is with the Department of Computer Science, Gunma University, Kiryu-shi, 376 Japan.

Since the scale of the vector $\boldsymbol{n}$ is indeterminate, we normalize it to $\|\boldsymbol{n}\| = 1$. Let $\bar{\boldsymbol{x}}_\alpha$ be the true value of $\boldsymbol{x}_\alpha$ obtained by replacing x_α and y_α by $\bar{x}_\alpha$ and $\bar{y}_\alpha$, respectively. The line fitting problem can be formally stated as the following statistical estimation (the inner product of vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ is denoted by $(\boldsymbol{a}, \boldsymbol{b})$):

Problem: Estimate a unit vector $\boldsymbol{n}$ such that

$$(\boldsymbol{n}, \bar{\boldsymbol{x}}_\alpha) = 0, \quad \alpha = 1, ..., N, \tag{3}$$

from the data $\{\boldsymbol{x}_\alpha\}$, $\alpha = 1$, ..., N.

The vectors $\boldsymbol{x}_\alpha$ and $\boldsymbol{n}$ defined by Eqs. (2) can be given the following geometric interpretation. Take an XYZ coordinate system such that the image is at Z=1 as shown in Fig. 1. Then, $\boldsymbol{x}_\alpha$ can be viewed as the vector that starts from the origin O and terminates at (x_α, y_α); vector $\boldsymbol{n}$ can be identified with the unit surface normal to the plane defined by the origin O and the line $Ax+By+C = 0$ in the image. The components of the vectors $\boldsymbol{x}_\alpha$ and $\boldsymbol{n}$ can be interpreted to be the *homogeneous coordinates* of the point (x_α, y_α) and the line $Ax + By + C = 0$, respectively [6].

The configuration described in Fig. 1 can be identified with *perspective projection* [5]. If we want to reconstruct a 3-D structure from an image, we must make this configuration agree with the actual imaging geometry by doing *camera calibration* [4], [8]. Here, however, this configuration is hypothetical and is used merely for the convenience of computation.

3. Optimal Estimator and its Reliability

We write

$$\boldsymbol{x}_\alpha = \bar{\boldsymbol{x}}_\alpha + \Delta\boldsymbol{x}_\alpha, \tag{4}$$

and regard the noise $\Delta\boldsymbol{x}_\alpha$ as a Gaussian random variable of mean $\boldsymbol{0}$, independent for each α, and covariance matrix

$$V[\boldsymbol{x}_\alpha] = E[\Delta\boldsymbol{x}_\alpha\Delta\boldsymbol{x}_\alpha^\top], \tag{5}$$

where $E[\,\cdot\,]$ denotes expectation. The superscript $\top$ denotes transpose. Since the third component of $\Delta\boldsymbol{x}_\alpha$ is always 0, the covariance matrix $V[\boldsymbol{x}_\alpha]$ is generally a singular matrix of rank 2. It can be shown that the optimal estimator of the vector $\boldsymbol{n}$ is obtained by the minimization

$$J[\boldsymbol{n}] = \sum_{\alpha=1}^{N} \frac{(\boldsymbol{n}, \boldsymbol{x}_\alpha)^2}{(\boldsymbol{n}, V[\boldsymbol{x}_\alpha]\boldsymbol{n})} \to \min, \tag{6}$$

under the constraint $\|\boldsymbol{n}\| = 1$ [7].

Let $\hat{\boldsymbol{n}}$ be the resulting estimator, and write

$$\hat{\boldsymbol{n}} = \bar{\boldsymbol{n}} + \Delta\boldsymbol{n}, \tag{7}$$

where $\bar{\boldsymbol{n}}$ is the true value of $\boldsymbol{n}$. The reliability of the estimator $\hat{\boldsymbol{n}}$ is measured by its covariance matrix $V[\hat{\boldsymbol{n}}] = E[\Delta\hat{\boldsymbol{n}}\Delta\hat{\boldsymbol{n}}^\top]$. It can be shown that the covariance matrix has the form

$$V[\hat{\boldsymbol{n}}] = \left(\sum_{\alpha=1}^{N} \frac{\bar{\boldsymbol{x}}_\alpha\bar{\boldsymbol{x}}_\alpha^\top}{(\bar{\boldsymbol{n}}, V[\boldsymbol{x}_\alpha]\bar{\boldsymbol{n}})}\right)^-, \tag{8}$$

where $(\,\cdot\,)^-$ denotes the (*Moore-Penrose*) *generalized inverse* (see Appendix). Equation (8) describes a theoretical bound on the attainable accuracy called the *Cramer-Rao lower bound* [7].

Example: If each coordinate is perturbed independently by Gaussian noise of mean 0 and variance ϵ^2, the covariance matrix $V[\boldsymbol{x}_\alpha]$ has the form

$$V[\boldsymbol{x}_\alpha] = \epsilon^2 \mathrm{diag}(1, 1, 0), \tag{9}$$

where $\mathrm{diag}(\cdots)$ denotes the diagonal matrix whose diagonal elements are $\cdots$ in that order. In this case, the optimization (6) is equivalent to the familiar *least-squares method*

$$\sum_{\alpha=1}^{N} D(p_\alpha, l)^2 \to \min, \tag{10}$$

where $D(p_\alpha, l)$ is the distance between the αth point p_α and the line l to be fitted. The Cramer-Rao lower bound given by Eq. (8) reduces to

$$V[\hat{\boldsymbol{n}}] = \frac{\sigma^2}{1+d^2}\begin{pmatrix} \sum_{\alpha=1}^N \bar{x}_\alpha^2 & \sum_{\alpha=1}^N \bar{x}_\alpha\bar{y}_\alpha & \sum_{\alpha=1}^N \bar{x}_\alpha \\ \sum_{\alpha=1}^N \bar{y}_\alpha\bar{x}_\alpha & \sum_{\alpha=1}^N \bar{y}_\alpha^2 & \sum_{\alpha=1}^N \bar{y}_\alpha \\ \sum_{\alpha=1}^N \bar{x}_\alpha & \sum_{\alpha=1}^N \bar{y}_\alpha & N \end{pmatrix}^-, \tag{11}$$

where d is the distance of the line from the image origin.

4. Least-Squares Approximation

We decompose the covariance matrix $V[\boldsymbol{x}_\alpha]$ into the form

$$V[\boldsymbol{x}_\alpha] = \epsilon^2 V_0[\boldsymbol{x}_\alpha]. \tag{12}$$

The constant ϵ indicates the average magnitude of the image noise; we call it the *noise level*. The matrix $V_0[\boldsymbol{x}_\alpha]$ indicates in which orientation the deviation is likely to occur; we call it the *normalized covariance matrix*. In many circumstances, the qualitative characteristics of image noise, such as homogeneity/inhomogeneity, isotropy/anisotropy, and their relative degrees, can be discerned relatively easily from the characteristics of the imaging device and the image processing algorithm, whereas its absolute magnitude is very difficult to predict a priori. Here, we assume that the covariance matrix $V[\boldsymbol{x}_\alpha]$ is known only *up to scale*: $V_0[\boldsymbol{x}_\alpha]$ is known, but ϵ is unknown.

If the denominator in Eq. (6) is replaced by a constant, we obtain the following *least-squares approximation*:

$$\tilde{J}[\boldsymbol{n}] = (\boldsymbol{n}, \boldsymbol{M}\boldsymbol{n}) \to \min. \tag{13}$$

Here, $\boldsymbol{M}$ is the *moment matrix* defined by

$$\boldsymbol{M} = \frac{1}{N}\sum_{\alpha=1}^{N} W_\alpha \boldsymbol{x}_\alpha \boldsymbol{x}_\alpha^\top, \tag{14}$$

$$W_\alpha = \frac{1}{(\boldsymbol{n}^*, V_0[\boldsymbol{x}_\alpha]\boldsymbol{n}^*)}, \tag{15}$$

where $\boldsymbol{n}^*$ is an appropriate estimate of $\boldsymbol{n}$. The solution of the optimization (13) is obtained as the unit eigenvector of the moment matrix $\boldsymbol{M}$ for the smallest eigenvalue [5].

5. Unbiased Estimation

It appears that the solution of the least-squares approximation can be obtained by guessing the initial value of $\boldsymbol{n}$, substituting it into $\boldsymbol{n}^*$, updating $\boldsymbol{n}$ by the resulting solution, and iterating this process. However, the solution thus obtained is in general statistically biased whatever weights W_α are used. This is reasoned as follows.

Taking the expectation of $\boldsymbol{M}$, we see that

$$\begin{aligned} E[\boldsymbol{M}] &= \frac{1}{N}\sum_{\alpha=1}^{N} W_\alpha E[(\bar{\boldsymbol{x}}_\alpha + \Delta\boldsymbol{x}_\alpha)(\bar{\boldsymbol{x}}_\alpha + \Delta\boldsymbol{x}_\alpha)^\top] \\ &= \bar{\boldsymbol{M}} + \frac{\epsilon^2}{N}\sum_{\alpha=1}^{N} W_\alpha V_0[\boldsymbol{x}_\alpha], \end{aligned} \tag{16}$$

where $\bar{\boldsymbol{M}}$ is the unperturbed moment matrix obtained by replacing $\boldsymbol{x}_\alpha$ by $\bar{\boldsymbol{x}}_\alpha$ in Eq. (14). Equation (16) implies that the solution of the least-squares approximation is statistically biased by $O(\epsilon^2)$ according to the *perturbation theorem* [5].

However, if we define the *unbiased moment matrix*

$$\hat{\boldsymbol{M}} = \boldsymbol{M} - \epsilon^2 \boldsymbol{N}, \tag{17}$$

where

$$\boldsymbol{N} = \frac{1}{N}\sum_{\alpha=1}^{N} W_\alpha V_0[\boldsymbol{x}_\alpha], \tag{18}$$

we have $E[\hat{\boldsymbol{M}}] = \bar{\boldsymbol{M}}$. Hence, an unbiased estimator of $\boldsymbol{n}$ is obtained by the optimization

$$\hat{J}[\boldsymbol{n}] = (\boldsymbol{n}, \hat{\boldsymbol{M}}\boldsymbol{n}) \to \min. \tag{19}$$

The solution is obtained as the unit eigenvector of $\hat{\boldsymbol{M}}$ for the smallest eigenvalue [5].

6. Renormalization

In order to compute the unbiased moment matrix $\hat{\boldsymbol{M}}$ by Eq. (17), we need to know the noise level ϵ precisely, which is very difficult as we mentioned earlier. If ϵ is underestimated, the bias still remains, while if it is overestimated, bias with opposite sign occurs. In order to avoid this difficulty, we apply an iterative scheme called *renormalization* [6], which iteratively estimates ϵ so that the smallest eigenvalue of $\hat{\boldsymbol{M}}$ becomes 0. The procedure for renormalization is as follows [6], [9], [11]:

1. Let c=0 and W_α=1, α=1, ..., N.
2. Compute the matrices $\boldsymbol{M}$ and $\boldsymbol{N}$ by Eqs. (14) and (18), respectively.
3. Compute the smallest eigenvalue λ of the matrix
$$\hat{\boldsymbol{M}} = \boldsymbol{M} - c\boldsymbol{N}, \tag{20}$$
and the corresponding unit eigenvector $\boldsymbol{n}$.
4. If $\lambda \approx 0$, return $\boldsymbol{n}$, c and $\boldsymbol{M}$. Otherwise, update c and W_α as follows:
$$c \leftarrow c + \frac{\lambda}{(\boldsymbol{n}, \boldsymbol{N}\boldsymbol{n})}, \quad W_\alpha \leftarrow \frac{1}{(\boldsymbol{n}, V_0[\boldsymbol{x}_\alpha]\boldsymbol{n})}. \tag{21}$$
5. Go back to Step 2.

Let $\hat{\boldsymbol{n}}$ be the returned value of $\boldsymbol{n}$. An unbiased estimator of the squared noise level ϵ^2 is obtained in the form

$$\hat{\epsilon}^2 = \frac{c}{1 - 2/N}. \tag{22}$$

This is a consequence of the fact that $N\hat{J}[\hat{\boldsymbol{n}}]/\epsilon^2$ is subject to a χ^2 distribution with $N-2$ degrees of freedom in the first order, where $\hat{J}[\hat{\boldsymbol{n}}]$ is the residual of the optimization (19). From Eqs. (8), (17) and (18), we can see that the covariance matrix of the estimator $\hat{\boldsymbol{n}}$ is estimated by

$$V[\hat{\boldsymbol{n}}] = \frac{\hat{\epsilon}^2}{N}\left(\hat{\boldsymbol{M}}\right)_2^-, \tag{23}$$

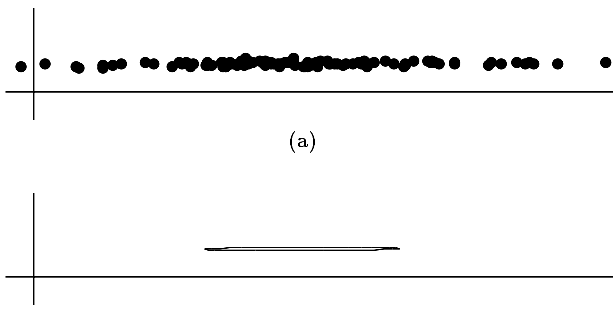
where $(\cdot)_2^-$ denotes the generalized inverse computed after projecting the matrix onto the space of rank 2 by ignoring the smallest eigenvalue [7]. This operation is necessary because the smallest eigenvalue of $\hat{\boldsymbol{M}}$ may not be strictly 0 if the renormalization iterations are prematurely terminated.

7. Primary Deviation Pair

Since $\hat{\boldsymbol{n}}$ is a unit vector, errors in $\hat{\boldsymbol{n}}$ cannot occur in the direction of $\hat{\boldsymbol{n}}$. This means that $V[\hat{\boldsymbol{n}}]$ has the following *spectral decomposition* [5]:

$$V[\hat{\boldsymbol{n}}] = \lambda_1 \boldsymbol{u}\boldsymbol{u}^\top + \lambda_2 \boldsymbol{v}\boldsymbol{v}^\top, \quad \lambda_1 \geq \lambda_2 > 0. \tag{24}$$

Fig. 2 Image points sequence.



(a)

(b)

Fig. 3 (a) Error distribution. (b) Theoretical bound.

Here, λ_1 and λ_2 are the eigenvalues of $V[\hat{\boldsymbol{n}}]$; $\boldsymbol{u}$ and $\boldsymbol{v}$ are the corresponding unit eigenvectors orthogonal to $\hat{\boldsymbol{n}}$. The vector $\boldsymbol{u}$ indicates the orientation of the most likely deviation; λ_1 is the variance in that direction. Hence, the reliability of $\hat{\boldsymbol{n}}$ can be visualized by displaying the two lines represented by

$$\boldsymbol{n}^+ = N[\hat{\boldsymbol{n}} + \sqrt{\lambda_1}\boldsymbol{u}], \quad \boldsymbol{n}^- = N[\hat{\boldsymbol{n}} - \sqrt{\lambda_1}\boldsymbol{u}], \qquad (25)$$

where $N[\cdot]$ denotes normalization into a unit vector. We call the two lines the *primary deviation pair* [9]–[11].

8. Examples

Figure 2 shows eight collinear points in an image. The distance between the two end points is assumed to be 40 pixels. We added Gaussian random noise of mean 0 and standard deviation 3 pixels to the x and y coordinates of each point independently.

Let $(\bar{\boldsymbol{n}}, \boldsymbol{x}) = 0$ and $(\hat{\boldsymbol{n}}, \boldsymbol{x}) = 0$ be the true and the fitted lines, respectively. Since the deviation of $\hat{\boldsymbol{n}}$ from $\bar{\boldsymbol{n}}$ is orthogonal to $\bar{\boldsymbol{n}}$ to a first approximation, we projected the difference $\hat{\boldsymbol{n}} - \bar{\boldsymbol{n}}$ onto the plane perpendicular to $\bar{\boldsymbol{n}}$ for 100 trials, each time using different noise (Fig. 3(a)). The deviation of a fitted line from the true line can be described as a superposition of a parallel displacement and a rotation around a fixed point. In Fig. 3(a), the vertical direction indicates the degree of parallel displacement; the horizontal direction indicates the degree of rotation around the center of the data points (we placed the coordinate origin arbitrarily). We can see that the error distribution is almost linear. In Fig. 3(b), the standard deviation in each orientation predicted by the Cramer-Rao lower bound (8) is indicated by an ellipse, which is nearly degenerated into a line segment; the aspect ratio (height)/(width)

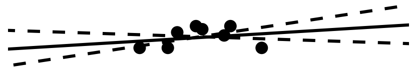


Fig. 4 An optimally fitted line and its primary deviation pair.

(a) (b)

Fig. 5 (a) An edge segment (15 pixels). (b) An optimally fitted line and its primary deviation pair.

is about 1.9×10^{-2}. Comparing this with Fig. 3(a), we can confirm that the theoretical bound on accuracy is almost attained.

Figure 4 shows a sequence of points with noise. A line is fitted by renormalization and drawn in a solid line; its primary deviation pair is drawn in dashed lines. We can confirm that the noise tends to cause a rotation around the center of the data points.

Figure 5(a) is a real image, in which an edge segment of 15 pixels detected by an edge operator is shown. A line is optimally fitted by renormalization and drawn in a solid line in Fig. 5(b); its primary deviation pair is drawn in dashed lines. Figure 6 shows the corresponding lines for a smaller edge segment (9 pixels). Comparing Figs. 5 and 6, we can confirm that the reliability of line fitting decreases as the length of the edge segment becomes shorter.

9. Asymptotic Approximation

Suppose the number N of pixels is an odd number. Let w be the distance between (x_1, y_1) and (x_N, y_N). Put

$$\boldsymbol{x}_C = \begin{pmatrix} \bar{x}_{(1+N)/2} \\ \bar{y}_{(1+N)/2} \\ 1 \end{pmatrix}, \qquad (26)$$

which represents the midpoint. Let $\boldsymbol{u}$ be the unit vector that indicates the orientation of the line $(\boldsymbol{n}, \boldsymbol{x}) = 0$. If $\{(\bar{x}_\alpha, \bar{y}_\alpha)\}$, $\alpha = 1, \ldots, N$, are approximately equidistant, we have

$$\bar{\boldsymbol{x}}_\alpha \approx \boldsymbol{x}_C + \frac{w}{N-1}\left(\alpha - \frac{N+1}{2}\right)\boldsymbol{u}. \qquad (27)$$

Suppose the image noise is homogeneous and isotropic and has the covariance matrix (9). Then, Eq. (8) is

(a) (b)

Fig. 6 (a) An edge segment (9 pixels). (b) An optimally fitted line and its primary deviation pair.

approximated by

$$V[\hat{\boldsymbol{n}}] \approx \frac{\epsilon^2}{N(1+d^2)}\left(\boldsymbol{x}_C\boldsymbol{x}_C^\top + \frac{w^2(N+1)}{12(N-1)}\boldsymbol{u}\boldsymbol{u}^\top\right)^-. \tag{28}$$

Define the *edge density* (the number of edge pixels per unit length) by $\rho = N/w$. If the edge segment is close to the image origin, approximations $\|\boldsymbol{x}_C\| \approx 1$ and $(\boldsymbol{x}_C, \boldsymbol{u}) \approx 0$ hold. In the limit $N \to \infty$, the following asymptotic expression of Eq. (28) is obtained:

$$V[\hat{\boldsymbol{n}}] \approx \frac{12\epsilon^2}{\rho w^3(1+d^2)}\left(\boldsymbol{u}\boldsymbol{u}^\top + \frac{12}{w^2}\boldsymbol{x}_C\boldsymbol{x}_C^\top\right)^-. \tag{29}$$

We assume that the image coordinates are scaled in such a way that the size of the image is of order $O(1)$. Then, $d \ll 1$ and $1 \ll 12/w^2$. Since $\boldsymbol{u}$ and $\boldsymbol{x}_C$ are approximately orthogonal, Eq. (29) is approximated in the following form:

$$V[\hat{\boldsymbol{n}}] \approx \frac{12\epsilon^2}{\rho w^3}\boldsymbol{u}\boldsymbol{u}^\top. \tag{30}$$

The primary deviation pair is approximated by

$$\boldsymbol{n}^+ \approx N[\hat{\boldsymbol{n}} + \frac{2\sqrt{3}\epsilon}{\rho^{1/2}w^{3/2}}\boldsymbol{u}],$$
$$\boldsymbol{n}^- \approx N[\hat{\boldsymbol{n}} - \frac{2\sqrt{3}\epsilon}{\rho^{1/2}w^{3/2}}\boldsymbol{u}]. \tag{31}$$

Eq. (30) plays a fundamental role in evaluating the reliability of camera calibration and 3-D reconstruction based on line fitting [1], [4], [8]. It can be seen from Eqs. (31) that the fitted line is very likely to pass near the midpoint of the edge segment. It can also be seen that the error is approximately proportional to $\rho^{-1/2}$ and $w^{-3/2}$.

10. Concluding Remarks

Introducing a mathematical model of image noise, we have formalized the problem of fitting a line to point data as statistical estimation. We have shown that the reliability of the fitted line can be evaluated quantitatively in the form of the covariance matrix of the parameters and presented a numerical scheme called *renormalization* for computing an optimal fit and at the same time evaluating its reliability. We have also presented a scheme for visualizing the reliability of the fit by means of the *primary deviation pair* and shown simulations and real-image examples to illustrate our method. Finally, we have derived an analytical expression for the reliability of a line fitted to an edge segment by using an *asymptotic approximation*.

Our analysis bears some similarity to what is known as *linear regression* in statistics [2], [14]. Its aim is to discern if one quantity linearly depends on another and, if so, to quantify the dependence. Mathematically, one first expresses the data value in terms of known variables, unknown parameters, and random noise and then infers the unknown parameters from multiple data. In contrast, the line fitting problem we consider here is to infer a *geometry* that constrain the data points, not to estimate a functional relationship. Because of this, the analysis is more difficult than for linear regression.

If the noise distribution is assumed to be homogeneous and isotropic, as we did in our examples, the noise characteristics are symmetric with respect to the line to be fitted. As a result, statistical bias does not occur, so accuracy is no longer improved by renormalization. However, the advantage of renormalization is that an optimal fit and its reliability are obtained by the same computation irrespective of the noise distribution. The basic theory presented here can be applied to curves of other types as well [12]. It can also be applied to fitting a surface to point data in three dimensions [9]–[11]. In this sense, this paper serves as a prototype of a wide range of fitting problems in computer vision and robotics applications.

Acknowledgment

This work was in part supported by the Ministry of Education, Science, Sports and Culture, Japan under a Grant in Aid for Scientific Research B (No. 07458067) and the Okawa Institute of Information and Telecommunication.

References

[1] B. Brillault-O'Mahony, "New method for vanishing point detection," CVGIP: Image Understanding, vol.54, no.2, pp.289–300, 1991.
[2] A. J. Dobson, "An Introduction to Generalized Linear Models," Chapman and Hall, London, 1990.
[3] B. Kamgar-Parsi, B. Kamgar-Parsi and N. Netanyahu, "A nonparametric method for fitting a straight line to a noisy image," IEEE Trans. Patt. Anal. Mach. Intell., vol.11, no.9, pp.998–1001, 1989.
[4] K. Kanatani, "Statistical analysis of focal-length calibration using vanishing points," IEEE Trans. Robotics Automation,

vol.8, no.6, pp.767–775, 1992.
[5] K. Kanatani, "Geometric Computation for Machine Vision," Oxford University Press, Oxford, 1993.
[6] K. Kanatani, "Statistical analysis of geometric computation," CVGIP: Image Understanding, vol.59, no.3, pp.286–306, 1994.
[7] K. Kanatani, "Statistical Optimization for Geometric Computation: Theory and Practice," Elsevier Science, Amsterdam, 1996.
[8] K. Kanatani and T. Maruyama, "Optimal grid pattern for focal length calibration," Advanced Robotics, to appear.
[9] Y. Kanazawa and K. Kanatani, "Direct reconstruction of planar surfaces by stereo vision," IEICE Trans. Inf. & Syst., vol.E78-D, no.7, pp.917–922, 1995.
[10] Y. Kanazawa and K. Kanatani, "Reliability of 3-D reconstruction by stereo vision," IEICE Trans. Inf. & Syst., vol.E78-D, no.10, pp.1301–1306, 1995.
[11] Y. Kanazawa and K. Kanatani, "Reliability of fitting a plane to range data," IEICE Trans. Inf. & Syst., vol.E78-D, no.12, pp.1630–1635, 1995.
[12] Y. Kanazawa and K. Kanatani, "Reliability of conic fitting," Proc. 2nd Asian Conf. Comput. Vision, December 1995, Singapore, vol.3, pp.387–401.
[13] N. Kiryati and A. M. Bruckstein, "What's in a set of points?," IEEE Trans. Patt. Anal. Mach. Intell., vol.14, no.4, pp.496–500, 1992.
[14] P. McCullagh and J. A. Nelder, "Generalized Linear Models," 2nd edition, Chapman and Hall, London, U.K., 1989.
[15] I. Weiss, "Line fitting in a noisy image," IEEE Trans. Patt. Anal. Mach. Intell., vol.11, no.3, pp.325–329, 1989.
[16] M. Werman and Z. Geyzel, "Fitting a second degree curve in the presence of error," IEEE Trans. Patt. Anal. Mach. Intell., vol.17, no.2, pp.207–211, 1989.
[17] Y. T. Zhou, V. Venkateswar and R. Chellappa, "Edge detection and linear feature extraction using a 2-D random field model," IEEE Trans. Patt. Anal. Mach. Intell., vol.11, no.1, pp.84–95, 1989.

Appendix: Bound on Accuracy

Substituting Eqs. (4) and (7) into Eq. (6) and ignoring terms of order three or higher in $\Delta\boldsymbol{n}$ and $\Delta\boldsymbol{x}_\alpha$, we obtain

$$J[\bar{\boldsymbol{n}}+\Delta\boldsymbol{n}] = \sum_{\alpha=1}^{N} \frac{((\Delta\boldsymbol{n},\bar{\boldsymbol{x}}_\alpha)+(\bar{\boldsymbol{n}},\Delta\boldsymbol{x}_\alpha))^2}{(\bar{\boldsymbol{n}},V[\boldsymbol{x}_\alpha]\bar{\boldsymbol{n}})}. \quad (\mathrm{A}\cdot 1)$$

Since $\hat{\boldsymbol{n}}$ is a unit vector, $\Delta\boldsymbol{n}$ must satisfy $(\bar{\boldsymbol{n}},\Delta\boldsymbol{n}) = 0$ to a first approximation. Introducing a Lagrange multiplier λ to this constraint, differentiating $J[\bar{\boldsymbol{n}}+\Delta\boldsymbol{n}]-2\lambda(\bar{\boldsymbol{n}},\Delta\boldsymbol{n})$ with respect to $\Delta\boldsymbol{n}$ and setting the result 0, we obtain

$$\sum_{\alpha=1}^{N} \frac{(\Delta\boldsymbol{n},\bar{\boldsymbol{x}}_\alpha)+(\bar{\boldsymbol{n}},\Delta\boldsymbol{x}_\alpha)}{(\bar{\boldsymbol{n}},V[\boldsymbol{x}_\alpha]\bar{\boldsymbol{n}})}\bar{\boldsymbol{x}}_\alpha = \lambda\bar{\boldsymbol{n}}. \quad (\mathrm{A}\cdot 2)$$

Multiplying matrix $\boldsymbol{P}_{\bar{\boldsymbol{n}}} = \boldsymbol{I}-\bar{\boldsymbol{n}}\bar{\boldsymbol{n}}^\top$ ($\boldsymbol{I}$ is the unit matrix) on both sides and noting that $\boldsymbol{P}_{\bar{\boldsymbol{n}}}\bar{\boldsymbol{x}}_\alpha = \bar{\boldsymbol{x}}_\alpha$ and $\boldsymbol{P}_{\bar{\boldsymbol{n}}}\bar{\boldsymbol{n}} = \boldsymbol{0}$, we have

$$\left(\sum_{\alpha=1}^{N} \frac{\bar{\boldsymbol{x}}_\alpha\bar{\boldsymbol{x}}_\alpha^\top}{(\bar{\boldsymbol{n}},V[\boldsymbol{x}_\alpha]\bar{\boldsymbol{n}})}\right)\Delta\boldsymbol{n} + \sum_{\alpha=1}^{N}\left(\frac{\bar{\boldsymbol{x}}_\alpha\bar{\boldsymbol{n}}^\top}{(\bar{\boldsymbol{n}},V[\boldsymbol{x}_\alpha]\bar{\boldsymbol{n}})}\right)\Delta\boldsymbol{x}_\alpha = 0, \quad (\mathrm{A}\cdot 3)$$

from which we obtain

$$\Delta\boldsymbol{n} = -\boldsymbol{V}\sum_{\alpha=1}^{N}\left(\frac{\bar{\boldsymbol{x}}_\alpha\bar{\boldsymbol{n}}^\top}{(\bar{\boldsymbol{n}},V[\boldsymbol{x}_\alpha]\bar{\boldsymbol{n}})}\right)\Delta\boldsymbol{x}_\alpha, \quad (\mathrm{A}\cdot 4)$$

where

$$\boldsymbol{V} = \left(\sum_{\alpha=1}^{N} \frac{\bar{\boldsymbol{x}}_\alpha\bar{\boldsymbol{x}}_\alpha^\top}{(\bar{\boldsymbol{n}},V[\boldsymbol{x}_\alpha]\bar{\boldsymbol{n}})}\right)^-. \quad (\mathrm{A}\cdot 5)$$

The covariance matrix $V[\hat{\boldsymbol{n}}]$ of the solution $\hat{\boldsymbol{n}}$ is evaluated as follows ($\delta_{\alpha\beta}$ is the *Kronecker delta*):

$$\begin{aligned} V[\hat{\boldsymbol{n}}] &= E[\Delta\boldsymbol{n}\Delta\boldsymbol{n}^\top] \\ &= \boldsymbol{V}\left(\sum_{\alpha,\beta=1}^{N}\frac{\bar{\boldsymbol{x}}_\alpha\bar{\boldsymbol{n}}^\top E[\Delta\boldsymbol{x}_\alpha\Delta\boldsymbol{x}_\beta]\bar{\boldsymbol{n}}\bar{\boldsymbol{x}}_\beta^\top}{(\bar{\boldsymbol{n}},V[\boldsymbol{x}_\alpha]\bar{\boldsymbol{n}})(\bar{\boldsymbol{n}},V[\boldsymbol{x}_\alpha]\bar{\boldsymbol{n}})}\right)\boldsymbol{V} \\ &= \boldsymbol{V}\sum_{\alpha,\beta=1}^{N}\frac{\bar{\boldsymbol{x}}_\alpha\bar{\boldsymbol{n}}^\top\delta_{\alpha\beta}V[\boldsymbol{x}_\alpha]\bar{\boldsymbol{n}}\bar{\boldsymbol{x}}_\beta^\top}{(\bar{\boldsymbol{n}},V[\boldsymbol{x}_\alpha]\bar{\boldsymbol{n}})(\bar{\boldsymbol{n}},V[\boldsymbol{x}_\alpha]\bar{\boldsymbol{n}})}\boldsymbol{V} \\ &= \boldsymbol{V}\sum_{\alpha=1}^{N}\frac{\bar{\boldsymbol{x}}_\alpha(\bar{\boldsymbol{n}},V[\boldsymbol{x}_\alpha]\bar{\boldsymbol{n}})\bar{\boldsymbol{x}}_\alpha^\top}{(\bar{\boldsymbol{n}},V[\boldsymbol{x}_\alpha]\bar{\boldsymbol{n}})^2}\boldsymbol{V} \\ &= \boldsymbol{V}\sum_{\alpha=1}^{N}\frac{\bar{\boldsymbol{x}}_\alpha\bar{\boldsymbol{x}}_\alpha^\top}{(\bar{\boldsymbol{n}},V[\boldsymbol{x}_\alpha]\bar{\boldsymbol{n}})}\boldsymbol{V} = \boldsymbol{V}\boldsymbol{V}^-\boldsymbol{V} \\ &= \boldsymbol{V}. \end{aligned} \quad (\mathrm{A}\cdot 6)$$

Thus, we obtain Eq. (8).

Yasushi Kanazawa received his B.E. and M.E. degrees in information and computer science from Toyohashi University of Technology, Aichi, Japan, in 1985 and 1987, respectively. He is currently Assistant Professor of information and computer engineering at Gunma College of Technology. His research interests include image processing and computer vision.

Kenichi Kanatani received his Ph.D. in applied mathematics from the University of Tokyo in 1979. He is currently Professor of computer science at Gunma University. He is the author of *Group-Theoretical Methods in Image Understanding* (Springer, 1990), *Geometric Computation for Machine Vision* (Oxford University Press, 1993) and *Statistical Optimization for Geometric Computation: Theory and Practice* (Elsevier Science, 1996).