

- [3] S.A. Dudani, K.J. Breeding, and R.B. McGhee, "Aircraft identification by moment invariants," *IEEE Trans. Comput.*, vol. C-26, no. 1, pp. 39-46, 1977.
- [4] Z. You and A.K. Jain, "Performance evaluation of shape matching via chord length distribution," *Comput. Vision Graphics Image Processing*, vol. 28, pp. 185-198, 1984.
- [5] R.L. Kasyap and R. Chellappa, "Stochastic models for closed boundary analysis: Representation and reconstruction," *IEEE Trans. Inform. Theory*, vol. IT-27, no. 5, pp. 627-637, 1981.
- [6] S.R. Dubois and F.H. Glanz, "An autoregressive model approach to two-dimensional shape classification," *IEEE Trans. Patt. Anal. Machine Intell.*, vol. PAMI-8, no. 1, pp. 55-66, 1986.
- [7] N. Otsu and T. Kurita, "Proposal of complex autoregressive model as a shape descriptor," *Nat. Convention Rec. IEICE Japan*, vol. D-496, 1989, in Japanese.
- [8] F. Mokhtarian and A. Mackworth, "Scale-based description and recognition of planar curves and two-dimensional shapes," *IEEE Trans. Patt. Anal. Machine Intell.*, vol. PAMI-8, no. 1, pp. 34-43, 1986.
- [9] M. Das, M.J. Paulik, and N.K. Loh, "A bivariate autoregressive modeling technique for analysis and classification of planar shapes," *IEEE Trans. Patt. Anal. Machine Intell.*, vol. 12, no. 1, pp. 97-103, 1990.
- [10] T. Kurita, I. Sekita, and N. Otsu, "Distance measures for planar shape analysis based on a complex autoregressive model," *Trans. IEICE Japan*, vol. J73-D-II, no. 9, pp. 1493-1503, 1990, in Japanese.
- [11] I. Sekita, T. Kurita, and N. Otsu, "Complex autoregressive model and its properties," Tech. rep. ETL, TR-91-12, 1991.
- [12] J. Durbin, "The fitting of time-series models," *Rev. Inst. Int. de Stat.*, vol. 28, no. 3, pp. 233-244, 1960.
- [13] H. Akaike, "A new look at the statistical model identification," *IEEE Trans. Automat. Contr.*, vol. AC-19, no. 6, pp. 716-723, 1974.
- [14] J. Rissanen, "A universal prior for integers and estimation by minimum description length," *Annals Stat.*, vol. 11, no. 2, pp. 416-431, 1983.
- [15] —, "Stochastic complexity and modeling," *Annals Stat.*, vol. 14, no. 3, pp. 1080-1100, 1986.
- [16] R.O. Duda and P.E. Hart, *Pattern Classification and Scene Analysis*, New York: Wiley, 1973.
- [17] G.S. Sebestyen, *Decision-Making Processes in Pattern Recognition*, New York: Macmillan, 1962.
- [18] G.T. Toussaint, "Bibliography on estimation of misclassification," *IEEE Trans. Inform. Theory*, vol. IT-20, no. 4, pp. 472-479, 1974.
- [19] J.D. Markel and A.H. Gray, Jr., *Linear Prediction of Speech*, New York: Springer-Verlag, 1976.
- [20] I. Sekita, T. Kurita, and N. Otsu, "Shape reconstruction from complex autoregressive coefficients," *Nat. Convention Rec. IEICE Japan*, vol. D-258, 1990, in Japanese.

## What's in a Set of Points?

N. Kiryati and A.M. Bruckstein

**Abstract**—The problem of fitting a straight line to a planar set of points is reconsidered. A parameter space computational approach capable of fitting one or more lines to a set of points is presented. The suggested algorithm handles errors in both coordinates of the data points, even when the error variances vary between coordinates and among points and can be readily made robust to outliers. The algorithm is quite general and allows line fitting according to several useful optimality criteria to be performed within a single computational framework. It is observed that certain extensions of the Hough transform can be tuned to be equivalent to well-known  $M$  estimators, thus allowing computationally efficient approximate  $M$  estimation.

**Index Terms**—Hough transform, least squares, line fitting, linear regression,  $M$  estimators, robust regression

### I. INTRODUCTION

Fitting a straight line to a planar set of points is a routine scientific and engineering task. The most popular method employed is probably still fitting the straight line "by eye"; surprisingly, little is known about its performance [14]. The method of least squares is the most common mathematical approach; its estimates of the line parameters are usually chosen to minimize the sum of squared "vertical" fitting errors (residuals).

The conventional least squares technique fails when outliers, i.e., "wild" points, contaminate the data. These tend to produce large residuals that lead to inadequate line fitting by the method of least squares. Since the early 1970's, considerable statistical research effort has been aimed at devising alternative techniques that would be robust to "influential points" in general and outliers in particular. A concise introduction to these techniques can be found in [11]. The texts [6], [12], [17] provide comprehensive treatment and many references to relevant statistical literature.

The so-called "errors in the variables" problem of fitting a straight line in the presence of observation errors in the independent variable ("carrier"), in addition to those in the dependent variable, has received relatively little attention; a classical reference is [13]. It is well known that if the problem can be scaled such that the error variances are equal in both the independent and dependent variables and among all points, then a modified least squares technique, in which fitting errors are measured in the normal direction to the line, should be employed. This specific line fitting problem admits an analytic solution.

The solution of the general problem, in which error variances vary among points and between the dependent and independent variables, is based on minimizing the sum

$$S = \sum_i [W_{x_i}(x_i - X_i)^2 + W_{y_i}(y_i - Y_i)^2] \quad (1)$$

where  $(x_i, y_i)$  are the observed points,  $W_{x_i}$  and  $W_{y_i}$  are "weights" that correspond to the reciprocals of the respective error variances, and  $(X_i, Y_i)$  are the collinear "adjusted" coordinates. This problem

Manuscript received February 13, 1990; revised May 2, 1991. Recommended for acceptance by G.T. Toussaint.

N. Kiryati is with the Department of Electrical Engineering, Technion-Israel Institute of Technology, Haifa, Israel.

A.M. Bruckstein is with the Department of Computer Science, Technion-Israel Institute of Technology, Haifa, Israel.

IEEE Log Number 9102098.

does not admit an analytic solution, hence, numerical techniques have been suggested [20], [16]. Robust line fitting procedures for the "error in the variables" problem are scarce, e.g., [2], and to our knowledge are limited to degenerate cases.

In computer vision, straight line fitting problems arise in the context of feature extraction; roughly collinear patterns in an edge map correspond to straight edges in the digital image from which the edge map was produced. There are several special challenges in this problem. First, just a small fraction of the edge points are associated with the straight line, whereas all others, which are due to noise and to other features, should be treated as outliers. Second, since near real time processing of sequences of images is an important goal, algorithms must be efficient and suitable for parallel implementation. Third, the simultaneous fitting of several straight lines to a single data set is often required.

Although several approaches have been taken, e.g., [10], [19], the mainstream computer vision approach to the line fitting problem is based on the Hough transform [3], [7], which is a computational algorithm for finding large collinear subsets within a planar set of points. The main purpose of this paper is to present a related procedure that is capable of robust or conventional approximate line fitting in the presence of errors in both coordinates, even if the error variances differ between the coordinates and among the data points. Approximate  $M$  estimation is obtained as a special case. The dependence of the number of operations on the number of data points is linear, and inherent parallelism enables fast parallel implementation.

The organization of this paper is as follows: in Section II the Hough transform and a few relevant variants are briefly presented. Section III focuses on the theoretical foundation of the suggested algorithm, and technical issues are considered in Section IV. Examples are provided in Section V.

## II. PARAMETER SPACE APPROACHES TO LINE FITTING

The well-known Hough transform [7] is an efficient computational method for the detection of predefined features in digital images. Consider the Hough transform for straight line detection using normal parameters (which are the distance from the line to the origin and the angle between its normal and the horizontal axis), as suggested by Duda and Hart [3]. Aiming to detect lines through large collinear subsets within a planar set of edge points  $P \triangleq \{(x_i, y_i), i = 1, \dots, M\}$ , each point is regarded as a constraint

$$\rho_i(\theta) = x_i \cos \theta + y_i \sin \theta \quad (2)$$

on the normal parameters  $(\rho, \theta)$  of any line through that point. The intersection of a large number of sinusoids in the  $(\rho, \theta)$  normal parameters plane corresponds to the normal parameters of a straight line through a large collinear subset of  $P$ .

In the standard implementation, (a subset of) the  $(\rho, \theta)$  plane is divided into  $N_\rho \times N_\theta$  rectangular cells and represented by an accumulator array. The algorithm is performed in two stages; the first is an incrementation ("voting") stage, in which for every  $i \in [1, \dots, M]$ , the accumulators corresponding to cells that the sinusoid (2) intersects are incremented. The second stage is an exhaustive search for maxima in the accumulator array. These represent the normal parameters of straight lines through large collinear subsets of points. The execution of the Duda and Hart algorithm requires  $O(M \cdot N_\theta)$  operations in the incrementation stage and  $O(N_\rho \cdot N_\theta)$  operations in the search stage.

Unless special interpolation techniques are employed, the resolution of the parameters of lines detected (fitted) by the Duda and Hart algorithm depends on the quantization of the "continuous" parameter plane. It is clear, however, that several sinusoids intersect at a single

point in the continuous parameter plane only if the corresponding edge points are truly collinear in the image plane. Thus, in the limiting case in which the number of accumulators is very large and each accumulator corresponds to an infinitesimal patch of the parameter plane, the Duda and Hart algorithm would fit lines only through truly collinear subsets of points. Formally, the Duda and Hart algorithm computes

$$\arg \max_{(\rho, \theta)} \sum_{i=1}^M I[r_i(\rho, \theta)] \quad (3)$$

where

$$r_i(\rho, \theta) \triangleq |\rho - \rho_i(\theta)| \quad (4)$$

and the indicator function  $I(r_i)$  is defined by

$$I(r_i) \triangleq \begin{cases} 1 & r_i = 0 \\ 0 & r_i \neq 0 \end{cases} \quad (5)$$

The Duda and Hart algorithm can fit several lines simultaneously to a given data set; furthermore, it is extremely robust to outliers in the data. Yet, since in principle it cannot tolerate small errors in the coordinates of the data points, the Duda and Hart algorithm as described above is quite inadequate for many image analysis problems, where errors in the coordinates of edge points, due to digitization and noise, are usually present. Employing coarse parameter plane quantization was regarded as a solution to the problem, but fundamental theoretical difficulties [9] manifest themselves in significantly degraded performance.

Thrift and Dunn [18] suggested a modification in the Duda and Hart algorithm to make it tolerant to small errors in the coordinates of the data points without giving up its robustness against outliers. The essential improvement is the replacement of the "impulsive" indicator function  $I(r_i)$  in (3) by a smooth, positive, decreasing function  $g(r_i)$  of finite support on  $r_i \geq 0$  that for some  $C_g > 0$  satisfies  $g(r_i) = 0$  for all  $r_i > C_g$ . The only change in implementation with respect to the conventional Duda and Hart algorithm is that sinusoidal bands whose vertical profile is everywhere the symmetric extension of  $g(r_i)$  are accumulated, rather than simple sinusoids. The robustness of the Thrift and Dunn algorithm against outliers follows from the finite support of  $g(r_i)$  on  $r_i$ . It is easily seen [8] that  $r_i$  represents the distance between a point and a straight line in a direction perpendicular to the line. Thus, the Thrift and Dunn algorithm can be regarded as a useful robust algorithm for fitting lines to a set of points in the presence of errors in both coordinates for the special case in which error variances are equal in both coordinates and among all points.

Kiryati and Bruckstein [8] have pointed out that the Thrift and Dunn mechanism can be extended to approximately solve the general problem

$$\arg \min_{(\rho, \theta)} \sum_{i=1}^M C_i[r_i(\rho, \theta)]. \quad (6)$$

By specifying appropriate cost functions  $\{C_i(r_i)\}$ , which in particular need not be monotonic or similar among all data points, the algorithm of [8] and [18] can be used to solve diverse problems. In the context of fitting a straight line to a data set, specifying

$$C_i(r_i) = r_i^2 \quad \forall i \quad (7)$$

tunes the algorithm to compute the least squares line with fitting errors measured in the direction perpendicular to the line. Choosing

$$C_i(r_i) = \min\{r_i^2, a^2\} \quad \forall i \quad (8)$$

yields a robust least squares solution to the same problem. With

$$C_i(r_i) = r_i^2/\sigma_i^2 \quad (9)$$

or, to obtain robustness

$$C_i(r_i) = \min\{r_i^2/\sigma_i^2, a_i^2\} \quad (10)$$

different error variances among the different points can be handled.

An illuminating observation is that by specifying appropriate cost functions, the algorithm of [8] and [18] approximates the well-known robust  $M$  estimators [6], [12] (with fitting errors measured in the direction perpendicular to the line rather than vertically). In particular, when using the cost function (8), the algorithm approximates the so-called *skipped mean  $M$  estimator* [5] (also known as *trimmed mean [1]*) and is a degenerate form of *Hampel's three-part redescending  $M$  estimator* [4].

The algorithm of [8] and [18] is executed in two stages. In the incrementation stage, a discrete approximation of the function

$$C(\rho, \theta) = \sum_{i=1}^M C_i[r_i(\rho, \theta)] \quad (11a)$$

is constructed. In the context of line fitting,  $C(\rho, \theta)$  represents the total error of fit as a function of the normal parameters of lines. In the final stage

$$\arg \min_{(\rho, \theta)} C(\rho, \theta) \quad (11b)$$

is determined by exhaustive search. The major advantage of using normal parameters and measuring fitting errors in the direction perpendicular to the line is that a very systematic incrementation law results. For a given data point, the incrementation procedure reduces to storing a discrete approximation of  $C_i(\cdot)$  in a linear array, and then, for every discrete value of  $\theta_0$ , shifting the linear array by  $\rho_i(\theta_0)$  and adding it to the  $\theta_0$  column of the accumulator array. These operations can be very efficiently implemented and can readily utilize parallel hardware if available.

### III. LINE FITTING WITH ERRORS IN BOTH COORDINATES

Fitting a straight line to a set of points in the presence of observation errors in both coordinates requires, when error variances may differ between the two coordinates and among the data points, the solution of

$$\arg \min_{(\rho, \theta)} C(\rho, \theta) \quad (12a)$$

where

$$C(\rho, \theta) = \sum_{i=1}^M \min_{X_i, Y_i} \left[ \frac{1}{\sigma_{x_i}^2} (X_i - x_i)^2 + \frac{1}{\sigma_{y_i}^2} (Y_i - y_i)^2 \right] \quad (12b)$$

and minimization in (12b) is subject to the constraint that the "adjusted" points  $\{(X_i, Y_i)\}$  satisfy

$$\rho = X_i \cos \theta + Y_i \sin \theta \quad i = 1, \dots, M. \quad (12c)$$

Note that in the special case of equal error variances in both coordinates, i.e.,  $\sigma_{x_i}^2 = \sigma_{y_i}^2 = \sigma_i^2$ , (12) reduces to (11) with  $C_i$  specified by (9). Furthermore, if these variances are also equal among all points, i.e.,  $\sigma_i^2 = \sigma^2 \forall i$ , then (9) reduces to (7).

In this section, a computational solution of (12) via an extension of the algorithm of [8] and [18] is developed. The key to successful implementation remains a systematic incrementation law. A simple modification in the suggested algorithm makes it robust against outliers.

Consider any single term in the summation (12b):

$$C_i(\rho, \theta) = \min_{X_i, Y_i} \left[ \frac{1}{\sigma_{x_i}^2} (X_i - x_i)^2 + \frac{1}{\sigma_{y_i}^2} (Y_i - y_i)^2 \right]. \quad (13)$$

Subject to (12c), (13) is the contribution of a specific data point to the total fitting error as a function of  $(\rho, \theta)$ . The minimization of (13) subject to (12c) can be carried out using Lagrange multipliers. Defining

$$\phi_i = \frac{1}{\sigma_{x_i}^2} (X_i - x_i)^2 + \frac{1}{\sigma_{y_i}^2} (Y_i - y_i)^2 + \lambda(\rho - X_i \cos \theta - Y_i \sin \theta) \quad (14)$$

a solution to the following set of equations is required:

$$\frac{\partial \phi_i}{\partial X_i} = 2(X_i - x_i)/\sigma_{x_i}^2 - \lambda \cos \theta = 0 \quad (15)$$

$$\frac{\partial \phi_i}{\partial Y_i} = 2(Y_i - y_i)/\sigma_{y_i}^2 - \lambda \sin \theta = 0 \quad (16)$$

$$\frac{\partial \phi}{\partial \lambda} = \rho - X_i \cos \theta - Y_i \sin \theta = 0. \quad (17)$$

Using straightforward algebra, it is easy to show that

$$C_i(\rho, \theta) = \frac{r_i^2(\rho, \theta)}{\sigma_{x_i}^2 \cos^2 \theta + \sigma_{y_i}^2 \sin^2 \theta} \quad (18)$$

where  $r_i(\rho, \theta)$  was defined in (4). Defining

$$V_i(\theta) \triangleq \sigma_{x_i}^2 \cos^2 \theta + \sigma_{y_i}^2 \sin^2 \theta \quad (19)$$

yields

$$C_i(\rho, \theta) = r_i^2(\rho, \theta)/V_i(\theta). \quad (20)$$

In the suggested algorithm, (20) guides, for each data point, the incrementation of the accumulator array. It is pleasing to note that in the special case  $\sigma_{x_i}^2 = \sigma_{y_i}^2 = \sigma_i^2$ , it turns out that  $V_i(\theta) = \sigma_i^2$  is independent of  $\theta$ , and the suggested algorithm reduces to that of [8], with  $C_i(r_i)$  defined by (9).

Since even in the general case  $V_i(\theta)$  is independent of  $\rho$ , (20) represents a very systematic incrementation law. For a given data point, the incrementation procedure reduces to storing a discrete approximation of  $C_i(r_i)$  in a linear array, and then, for every discrete value of  $\theta_0$ , shifting the array by  $\rho_i(\theta_0)$ , dividing it by  $V_i(\theta_0)$ , and adding it to the  $\theta_0$  column of the accumulator array. Robustness against outliers merely requires the replacement of (20) by

$$C_i^R(\rho, \theta) = \min\{r_i^2(\rho, \theta)/V_i(\theta), a_i^2\} \quad (21)$$

which has the effect of limiting the contribution of any data point to the total cost to a preset maximum. The additional computation due to the replacement of (20) by (21) is relatively small.

### IV. DISCRETIZATION, RESOLUTION, AND COMPLEXITY

In the previous section, a parameter space approach for line fitting by the computational solution of (12) was presented. In the suggested procedure,  $C(\rho, \theta)$  is approximated by an accumulator array. First, for each data point, accumulation is systematically carried out according to (20) or (21); minima of  $C(\rho, \theta)$  are then determined by searching the array.

Since the line that best fits a set of points obviously intersects the convex hull of the set of points, it is clear that only the subset of the parameter plane that represents the lines that intersect the convex hull need to be approximated by the accumulator array. For improved accuracy given a fixed number of accumulators, the problem should

be translated such that the origin coincides with, say, the center of mass of the data points.

Assume, without loss of generality, that the data points all lie within a circle of unit radius centered at the origin. The relevant subset of the accumulator array is then

$$A = \{(\rho, \theta) : -1 < \rho < 1, 0 \leq \theta < \pi\}. \quad (22)$$

The  $N_\rho \times N_\theta$  accumulator array will contain the values of  $C(\rho, \theta)$  at  $N_\rho \times N_\theta$  sampling points within  $A$ . As described, the execution of the algorithm on a conventional serial machine requires  $O(M \cdot N_\theta \cdot N_\rho)$  operations in the incrementation stage and  $O(N_\theta \cdot N_\rho)$  operations in the search stage. If parallel processing of vector elements is supported by the hardware, the number of operations in the incrementation stage can be reduced to  $O(M \cdot N_\theta)$ , which is similar to that of the conventional Duda and Hart algorithm.

If robust operation of the algorithm is desired, it is of computational advantage to modify (21) to

$$C_i^M(\rho, \theta) = \max \{a_i^2 - r_i^2(\rho, \theta)/V_i(\theta), 0\} \quad (23)$$

and search for the maximum, rather than the minimum, of  $C(\rho, \theta)$ . The limited support of (23) implies that less than  $N_\rho$  accumulators need to be incremented at every value of  $\theta$  for each data point. This feature is inherent in the Duda and Hart algorithm and has been observed and incorporated in the algorithm of Thrift and Dunn.

In [9], it has been shown that in the Duda and Hart algorithm, there exists no finite number  $N_\rho \times N_\theta$  of sampling points that provides sufficient representation of the parameter plane in the Nyquist sense, i.e., that meets the requirements of the sampling theorem. In principle, the same holds for the algorithm just presented. However, since (20), (21), and (23) are "better behaved" than the impulsive indicator function (5), the suggested algorithm is less problematic in this respect than the Duda and Hart algorithm. In practice, reasonable approximation of  $C(\rho, \theta)$  is possible if the number of available accumulators is large with respect to the reciprocals of the error variances.

The available number of accumulators determines the distances between sampling points of  $C(\rho, \theta)$ . Unless careful interpolation (as suggested in [15]) is employed, the intersample distances determine the resolution of the parameters of the fitted lines. Assuming that a reasonable preliminary approximation of the parameters of the line can be obtained using the available number of accumulators, focusing coarse-to-fine techniques [7] can yield superresolution. These are based on allocation of all accumulators to represent a small subset of  $A$  (the neighborhood of the peak) and reexecution of the algorithm.

## V. EXAMPLES

The general line fitting algorithm suggested in this paper has been implemented and executed on a VAX 785 computer. A coarse-to-fine focusing technique using a  $256 \times 256$  accumulator array has been employed.

Consider the planar set of points shown in Fig. 1, to which a straight line should be fitted. Error variances that are equal among all points (but not necessarily in both coordinates) are assumed. Specifying  $\sigma_x = 0$  and  $\sigma_y = 0.1$  for all points (i.e., assuming errors only in the  $y$  coordinate) and using the incrementation rule (20) resulted in line 1 shown in Fig. 2. This is the line that would have been found by the line fitting function available in certain pocket calculators.

If isotropic errors are assumed, e.g.,  $\sigma_x = \sigma_y = 0.05$  for all points and accumulation is again guided by (20), line 2 is found. This is the line that would have been determined by modified least squares procedures in which fitting errors are measured in the direction perpendicular to the line. The effect of the "outlying" data point on the

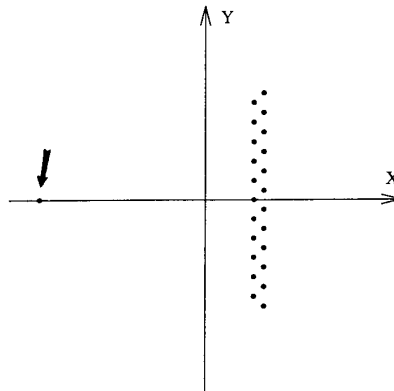


Fig. 1. Planar set of points to which a straight line should be fitted.

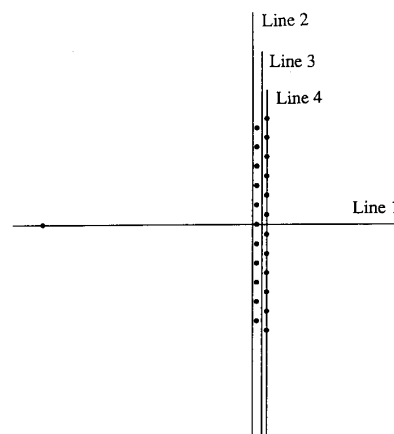


Fig. 2. Straight lines fitted to the data set by the suggested algorithm using different optimality criteria, i.e., different assumptions on the errors. *Line 1:*  $\sigma_x = 0$ ,  $\sigma_y = 0.1$  for all points (errors only in the  $y$  coordinate) and unbounded influence; *line 2:*  $\sigma_x = \sigma_y = 0.05$  for all points (isotropic errors) and unbounded influence; *line 3:*  $\sigma_x = \sigma_y = 0.05$  and maximum influence  $a^2 = 10$ ; *line 4:*  $\sigma_x = \sigma_y = 0.01$  and maximum influence  $a^2 = 10$ .

fitted line is clear. The algorithm can be made robust against outliers by replacing (20) by (21) and specifying a maximum influence  $a^2 = 10$ . This results in line 3, which is a pleasing nontrivial result.

By specifying very small errors ( $\sigma_x = \sigma_y = 0.01$ ) without increasing  $a^2$ , line 4 is obtained. This is the line that would have been fitted to the given set of points by the Duda and Hart algorithm, regarding points that are not precisely on the line as outliers. In this case, a second significant peak exists in the parameter plane, corresponding to the second largest truly collinear subset within the given set of points.

For clarity, examples that involve error variances that are different among the data points have been avoided; such problems, however, can readily be solved by the algorithm.

## ACKNOWLEDGMENT

The authors are grateful to A. Stein for several helpful comments.

## REFERENCES

- [1] V. Barnett and T. Lewis, *Outliers in Statistical Data*. New York: Wiley, 1984, section 4.1.3.
- [2] M. L. Brown, "Robust line estimation with errors in both variables," *J. Amer. Stat. Assoc.*, vol. 77, pp. 71–79, Mar. 1982.
- [3] R. O. Duda and P. E. Hart, "Use of the Hough transformation to detect lines and curves in pictures," *Comm. Assoc. Comput. Machines*, vol. 15, pp. 11–15, Jan. 1972.
- [4] C. Goodall, "M-estimators of location: An outline of the theory," in *Understanding Robust and Exploratory Data Analysis*. (D. C. Hoaglin, F. Mosteller and J. W. Tukey, Eds.). New York: Wiley, 1983.
- [5] F. R. Hampel, E. M. Ronchetti, P. J. Rousseeuw, and A. Stahel, *Robust Statistics: the Approach Based on Influence Functions*. New York: Wiley, 1986.
- [6] P. J. Huber, *Robust Statistics*. New York: Wiley, 1981.
- [7] J. Illingworth and J. Kitter, "A survey of the Hough transform," *Comput. Vision Graphics Image Processing*, vol. 44, pp. 87–116, 1988.
- [8] N. Kiryati and A. M. Bruckstein, "On navigation between friends and foes," *IEEE Trans. Patt. Anal. Machine Intell.*, vol. 13, no. 6, pp. 602–606, June 1991.
- [9] ———, "Antialiasing the Hough transform," in *Proc. IAPR 6th Scand. Conf. Image Anal., CVGIP: Graphical Models Image Processing*, vol. 53, no. 3, pp. 213–222, 1991.
- [10] B. Kamgar-Parsi and B. Kamgar-Parsi, "A nonparametric method for fitting a straight line to a noisy image," CAR-TR-315, Cent. Automat. Res., Univ. Maryland, College Park, Sept. 1987.
- [11] D. Y. Kim, J. J. Kim, P. Meer, D. Mintz, and A. Rosenfeld, "Robust computer vision: A least median of squares based approach." in *Proc. Image Understanding Workshop 1989* (Palo Alto), May 1989.
- [12] G. Li, "Robust regression," in *Exploring Data Tables, Trends and Shapes* (D. C. Hoaglin, F. Mosteller, and J. W. Tukey, Eds.). New York: Wiley, 1985.
- [13] A. Madanski, "The fitting of straight lines when both variables are subject to error," *J. Amer. Stat. Assoc.*, vol. 54, pp. 173–205, Mar. 1959.
- [14] F. Mosteller, A. F. Siegel, E. Trapido, and C. Youtz, "Fitting straight lines by eye," in *Exploring Data Tables, Trends and Shapes* (D. C. Hoaglin, F. Mosteller, and J. W. Tukey, Eds.). New York: Wiley, 1985.
- [15] W. Niblack and D. Petkovic, "On improving the accuracy of the Hough Transform: theory, simulations, and experiments," in *Proc. IEEE Comput. Soc. Conf. CVPR'88*, (Ann Arbor), June 1988, pp. 574–579.
- [16] B. C. Read, "Linear least-squares fits with errors in both coordinates," *Amer. J. Phys.*, vol. 57, no. 7, pp. 642–646, July 1989.
- [17] P. J. Rousseeuw and A. M. Leroy, *Robust Regression and Outlier Detection*. New York: Wiley, 1987.
- [18] P. R. Thrift and S. M. Dunn, "Approximating point-set images by line segments using a variation of the Hough transform," *Comput. Vision Graphics Image Processing*, vol. 21, pp. 383–394, 1983.
- [19] I. Weiss, "Line fitting in a noisy image," *IEEE Trans. Patt. Anal. Machine Intell.*, vol. 11, no. 3, pp. 325–329, Mar. 1989.
- [20] D. York, "Least squares fitting of a straight line," *Can. J. Phys.*, vol. 44, pp. 1079–1086, 1966.