

**MULTIDIMENSIONAL ALGORITHM  
FOR FINDING DISCONTINUITIES OF  
SIGNALS FROM NOISY DISCRETE DATA**

A. I. KATSEVICH, A. G. RAMM

ABSTRACT. Algorithms are given for finding discontinuity surfaces of a signal specified with random errors at a regular grid in  $\mathbb{R}^p$ ,  $p \geq 1$ . For each position of a sliding window one arranges points from it in ascending order of the measured values of the signal and divides them into two groups. One group contains points with larger values than the other. The division is made according to the nature of the particular pattern recognition problem. In the case of edge detection these groups have equal number of points. In the case of thin line detection one of the groups have a number of points proportional to the width of a line one is looking for. Then one calculates a functional which measures spatial separation of the above two groups. If the spatial separation is larger than the threshold, a discontinuity is deemed present. The threshold value is computed so that the probability of false alarm equals given parameter  $\epsilon$ . The above approach is generalized to yield a computational scheme applicable for other pattern recognition problems. The consistency of the algorithms is proved. The results of testing of the algorithms are presented.

1. INTRODUCTION

In this paper<sup>1</sup> we propose new algorithms for finding discontinuities of a signal  $s(x)$  specified with random errors at a regular grid in  $\mathbb{R}^p$ ,  $p \geq 1$ . The following assumptions are used: 1) one is given the radius  $R$  with the property  $|s(x) - s(x_0)| \ll 1$  for any  $x$ ,  $|x - x_0| \leq R$ , such that the line segment  $[x, x_0]$  does not intersect a discontinuity surface  $\Gamma$  of the signal  $s(x)$ ; 2) the errors are independent, identically distributed random variables with common (unknown) distribution function  $F(x)$  satisfying the condition:  $\exists F^{-1}(\gamma)$  for any  $\gamma$ ,  $0 < \gamma < 1$ ; 3) the desired probability  $\epsilon$  of the "false alarm" is given. Two variants of the general problem of finding discontinuities of a signal from its noisy measurements are considered.

*Variant I.* Suppose additionally that the discontinuity surface  $\Gamma$  can be accurately approximated by a plane in the neighborhood  $|x - x_0| \leq R$  of any point  $x_0 \in \Gamma$ . The algorithm operates as follows. In a scanning fashion one moves the center of the ball with radius  $R$  (a sliding window) over the grid and for each position of the ball the decision is made: whether  $\tilde{x}_j \in \Gamma$  or  $\tilde{x}_j \notin \Gamma$ , where  $\tilde{x}_j$  is the current center of the ball. To choose between the two hypotheses one arranges the observed values from the window in ascending order of magnitude and computes centroid  $x_j^\dagger$  of the

---

1991 *Mathematics Subject Classification.* Primary 62G05.

*Key words and phrases.* Finding discontinuities, edge detection, thin line detection, noisy data. The authors thank ONR for support. AGR thanks NSF for support.

<sup>1</sup>This paper was published in *Mathl. Comput. Modelling* Vol. 18, pp. 89–107, 1993

second half of the points, i.e. using the half which contains the points with larger values of the signal. Then the distance  $|\tilde{x}_j - x_j^+|$  is compared with the threshold  $A_0$ . If  $|\tilde{x}_j - x_j^+| \leq A_0$ , one decides that  $\tilde{x}_j \notin \Gamma$ . If  $|\tilde{x}_j - x_j^+| > A_0$ , one decides that  $\tilde{x}_j \in \Gamma$ . The grounds for such a decision rule are quite clear. If  $\tilde{x}_j \notin \Gamma$ , the points with small and large values of the signal are almost uniformly mixed inside the ball, hence the centroid  $x_j^+$  is close to the center  $\tilde{x}_j$ . If  $\tilde{x}_j \in \Gamma$ , the points with smaller values of the signal (on one side of  $\Gamma$ ) and the points with larger values of the signal (on the other side of  $\Gamma$ ) are spatially separated (although the noise can blur this separation to a certain extent). Hence the distance  $|\tilde{x}_j - x_j^+|$  is large. The threshold value  $A_0$  is computed so that the probability of the "false alarm" (one decides that  $\tilde{x}_j \in \Gamma$  when  $\tilde{x}_j \notin \Gamma$ ) equals  $\epsilon$ . If the decision " $\tilde{x}_j \in \Gamma$ " has been made, then the orientation of  $\Gamma$  at the point  $x_j$  is determined by the normal vector to  $\Gamma$ , namely the vector joining the points  $\tilde{x}_j$  and  $x_j^+$ .

*Variant II.* In this case the assumption is that the discontinuity surface  $\Gamma$  is the union of two smooth surfaces  $\Gamma_1$  and  $\Gamma_2$  situated close to each other at a constant distance  $2d < 2R$ . Suppose for example that the values of the signal between  $\Gamma_1$  and  $\Gamma_2$  are larger than in the neighborhoods adjoint to  $\Gamma_1$  and  $\Gamma_2$ . Let  $U$  be the domain between  $\Gamma_1$  and  $\Gamma_2$ ,  $\Gamma_0$  be the surface situated at the middle between  $\Gamma_1$  and  $\Gamma_2$ , and  $M_0$  be the number of grid points inside the set  $B_R(\tilde{x}_j) \cap U$ ,  $\tilde{x}_j \in \Gamma_0$ . The algorithm operates as follows. One slides a ball  $B_R(\tilde{x}_j)$  over the grid and for each  $\tilde{x}_j$  decides between the two hypotheses " $\tilde{x}_j \in \Gamma_0$ " and " $\tilde{x}_j \notin \Gamma_0$ ". To this end one arranges the observed values from the window in ascending order. If  $\tilde{x}_j \in \Gamma_0$ , it is likely that the majority of the last  $M_0$  points  $\{x_{i_k}\}_{k=2M-M_0+1}^{2M}$  (i.e. the points with the largest values of the signal) belong to  $B_R(\tilde{x}_j) \cap U$ , hence they are close to the diameter plane  $B_R(\tilde{x}_j) \cap \Gamma_0$  and the minimal eigenvalue  $\lambda_{min}$  of the matrix

$$B_{M_0} := (b_{mn}), \quad b_{mn} = h \sum_{k=2M-M_0+1}^{2M} (x_{i_k}^{(m)} - \tilde{x}_j^{(m)}) \cdot (x_{i_k}^{(n)} - \tilde{x}_j^{(n)}), \quad 1 \leq m, n \leq p.$$

is small. Here  $x_{i_k}^{(m)}$  is the  $m$ -th coordinate of the point  $x_{i_k}$ ,  $m = 1, \dots, p$ ;  $h := h_1 \cdot \dots \cdot h_p$ , where  $h_i$  is the grid step size along the  $i$ -th direction,  $i = 1, \dots, p$ . If  $B_R(\tilde{x}_j) \cap U = \emptyset$ , the points  $\{x_{i_k}\}_{k=2M-M_0+1}^{2M}$  are likely to be almost uniformly distributed over  $B_R(\tilde{x}_j)$  and  $\lambda_{min}$  is large. Thus to make a decision, one compares  $\lambda_{min}$  with the threshold  $\Lambda_0$ . If  $\lambda_{min} < \Lambda_0$ , the hypothesis " $\tilde{x}_j \in \Gamma_0$ " is accepted. If  $\lambda_{min} > \Lambda_0$ , the hypothesis " $\tilde{x}_j \notin \Gamma_0$ " is accepted. The threshold value  $\Lambda_0$  is computed so that the probability of "false alarm" equals  $\epsilon$ . If the decision " $\tilde{x}_j \in \Gamma_0$ " has been made, the orientation of  $\Gamma_0$  at the point  $\tilde{x}_j$  is determined by the normal vector to  $\Gamma_0$ , namely the eigenvector corresponding to  $\lambda_{min}$ . This algorithm can be easily modified for the case when the values of the signal inside  $U$  are smaller than in its neighborhood or for the case when they are larger than values from one neighborhood (say, adjoint to  $\Gamma_1$ ) and smaller than values from the other neighborhood (adjoint to  $\Gamma_2$ ).

The problem of finding discontinuities of a function is very important in many applications. In image analysis Variant I is known as edge detection, Variant II is known as thin line detection [10, pp. 491–496], [12, pp. 84–85]. A large number of edge and thin line detectors have been proposed (see [1], [10, pp. 497–556], [12, pp. 85–129] and references cited therein), but the majority of them ignore

the possible distortion of images by random errors. At the same time smoothing for noise suppression can blur weak edges and thin lines [4]. In [6]–[8] the idea of dividing the pixels in the window into two groups according to their intensity levels was used. The new points in our paper are 1) justification of the proposed algorithms, and 2) a general approach we suggest. It allows us to develop very simple and numerically stable algorithms. Since the terminology of image analysis is very convenient, in what follows Variant I will be called "edge detection", Variant II will be called "thin line detection".

The paper is organized as follows. In Section 2 the edge detector is described. In Section 3 the thin line detector is described. In Section 4 the above two algorithms are generalized, so that the resulting scheme can be applied for the solution of other pattern recognition problems. The spot detection problem is considered as an example. In Section 5 we prove the consistency of the proposed edge detector: the probability of missing the points from the discontinuity hypersurface  $\Gamma$  and the error of the recovery of its orientation go to zero as the grid step size goes to zero. In Section 6 similar results are proved for the proposed thin line detector. In Section 7 we prove the consistency of the general scheme. In Section 8 the results of testing of the proposed algorithms are presented.

## 2. EDGE DETECTION ALGORITHM

Let us consider the following model. Let the signal  $s(x)$ ,  $x \in \mathbb{R}^p$ ,  $p \geq 1$  have a discontinuity hypersurface  $\Gamma$  (which might consist of several disjoint components) and be a smooth function at other points ( $x \notin \Gamma$ ). The known data are the values

$$u_i = s_i + n_i, \quad s_i := s(x_i), \quad (2.1)$$

where  $x_i := (x_i^{(1)}, \dots, x_i^{(p)})$ . The observation grid is supposed to be rectangular with the step size  $h_k$  along the  $k$ -th direction,  $k = 1, \dots, p$ . The random errors  $n_i$  are assumed to be independent, identically distributed. Let us assume that we are given the radius  $R$  such that the signal  $s(x)$  is approximately constant inside any ball  $B_R(x_0) := \{x : |x - x_0| \leq R\}$  not intersecting  $\Gamma$ , that is  $|\nabla s(x)|R \ll 1$ . Let us also assume that the discontinuity surface  $\Gamma$  is sufficiently smooth so that inside any ball  $B_R(x_0)$ ,  $x_0 \in \Gamma$ , it can be approximated by a hyperplane. In this case we say that  $\Gamma$  is locally flat.

Consider the point  $\tilde{x}_j := x_j + (h_1/2, \dots, h_p/2) \in \Gamma$ . Let  $x_{i_k}$ ,  $k = 1, \dots, 2M$ , be the set of grid points which are situated inside  $B_R(\tilde{x}_j)$  and  $s_{i_k}$ ,  $k = 1, \dots, 2M$ , be the corresponding values of the signal. Assume that these values have been arranged in ascending order  $s_{i_1} \leq s_{i_2} \leq \dots \leq s_{i_{2M}}$ . Since  $\Gamma$  is locally flat in the neighborhood of  $\tilde{x}_j \in \Gamma$ , one sees that  $\Gamma$  divides  $B_R(\tilde{x}_j)$  into two equal halfballs  $B_R^-(\tilde{x}_j)$  and  $B_R^+(\tilde{x}_j)$ . Furthermore, one has  $x_{i_1}, \dots, x_{i_M} \in B_R^-(\tilde{x}_j)$ ,  $x_{i_{M+1}}, \dots, x_{i_{2M}} \in B_R^+(\tilde{x}_j)$ . Thus, if one calculates the centroid of the points  $x_{i_{M+1}}, \dots, x_{i_{2M}}$ , namely  $x_j^+ = (x_{i_{M+1}} + \dots + x_{i_{2M}})/M$ , then the straight line joining  $\tilde{x}_j$  and  $x_j^+$  is perpendicular to  $\Gamma$  at the point  $\tilde{x}_j$  (this line contains also the centroid of the points  $x_{i_1}, \dots, x_{i_M}$ ). Since the two sets  $\{x_{i_k}\}_{k=1}^M$  and  $\{x_{i_k}\}_{k=M+1}^{2M}$  are not spatially mixed ( $\{x_{i_k}\}_{k=1}^M \in B_R^-(\tilde{x}_j)$ ,  $\{x_{i_k}\}_{k=M+1}^{2M} \in B_R^+(\tilde{x}_j)$ ,  $B_R^-(\tilde{x}_j) \cap B_R^+(\tilde{x}_j) = \emptyset$ ), the distance  $|\tilde{x}_j - x_j^+|$  is the largest possible (compared with the case when the sets  $\{x_{i_k}\}_{k=1}^M$  and  $\{x_{i_k}\}_{k=M+1}^{2M}$  are mixed). Hence the following algorithm is natural:

*Step 1.* Fix a point  $\tilde{x}_j$ . Arrange the observed values from  $B_R(\tilde{x}_j)$  in ascending order  $u_{i_1} \leq u_{i_2} \leq \dots \leq u_{i_{2M}}$  with  $u_{i_k} := u(x_{i_k})$  and calculate the centroid  $x_j^+$  of the set  $\{x_{i_k}\}_{k=M+1}^{2M}$ .

Now, consider any point  $\tilde{x}_j$  such that  $B_R(\tilde{x}_j) \cap \Gamma = \emptyset$  and apply Step 1 to this  $B_R(\tilde{x}_j)$ . Since the noise and signal are independent and the signal is approximately constant inside  $B_R(\tilde{x}_j)$ , one sees that the sets  $\{x_{i_k}\}_{k=1}^M$  and  $\{x_{i_k}\}_{k=M+1}^{2M}$  are almost uniformly mixed and the distance  $|\tilde{x}_j - x_j^+|$  is close to zero. Therefore one can formulate the following decision rule:

*Step 2.* Select the threshold  $A_0 > 0$ , as described in Step 0 below. Compare the distance  $|\tilde{x}_j - x_j^+|$  with the threshold  $A_0$ . If  $|\tilde{x}_j - x_j^+| \leq A_0$ , the hypothesis " $\tilde{x}_j \notin \Gamma$ " is accepted. If  $|\tilde{x}_j - x_j^+| > A_0$ , the hypothesis " $\tilde{x}_j \in \Gamma$ " is accepted and the orientation of  $\Gamma$  inside  $B_R(\tilde{x}_j)$  is determined by the vector  $\tilde{x}_j - x_j^+$ .

The threshold  $A_0$  is determined from:

$$P\{|\tilde{x}_j - x_j^+| > A_0 | B_R(\tilde{x}_j) \cap \Gamma = \emptyset\} = \epsilon, \quad (2.2)$$

where  $P\{U|V\}$  is the conditional probability of  $U$ , given  $V$ ,  $\epsilon$ ,  $0 < \epsilon < 1$ , is an a priori fixed number. A convenient way to compute  $A_0$  from (2.2) is by using the Monte Carlo method [5, pp. 35–37] as described in Step 0:

*Step 0.* Fix an arbitrary ball  $B_R(\tilde{x}_j)$ . Model a sufficiently large number of the samples of  $u_i$  with  $s_i \equiv 0$  and an arbitrary fixed common probability density function (PDF) for  $n_i$ . Using Step 1, compute the distance  $|\tilde{x}_j - x_j^+|$  for each sample. Select  $A_0$  so that  $|\tilde{x}_j - x_j^+| > A_0$  in  $100\epsilon\%$  cases.

### 3. THIN LINE DETECTION ALGORITHM

In this Section we consider the particular case when the discontinuity surface  $\Gamma$  is the union of two surfaces  $\Gamma_1$  and  $\Gamma_2$  situated close to each other at a constant distance  $\text{dist}(x_1, \Gamma_2) = \text{dist}(x_2, \Gamma_1) = 2d < 2R$  for any  $x_i \in \Gamma_i$ ,  $i = 1, 2$ . Assume that  $d$  is known. The distance here is defined in the usual way  $\text{dist}(x, \Gamma) := \inf_{y \in \Gamma} |x - y|$ . We also assume that the surfaces  $\Gamma_i$ ,  $i = 1, 2$ , are sufficiently smooth so that they can be accurately approximated by a plane inside any ball  $B_R(x_0)$ ,  $x_0 \in \Gamma_i$ . Assume also that the values of the signal at the points between  $\Gamma_1$  and  $\Gamma_2$  are greater than in the neighborhoods adjoint to  $\Gamma_1$  and  $\Gamma_2$ . The problem is to find  $\Gamma_1$  and  $\Gamma_2$ .

Let  $\Gamma_0$  be the surface situated at the middle between  $\Gamma_1$  and  $\Gamma_2$ , that is the midpoint of every interval  $[x_1, x_2]$ ,  $x_1 \in \Gamma_1$ ,  $x_2 \in \Gamma_2$ , perpendicular to  $\Gamma_1$  and  $\Gamma_2$  belongs to  $\Gamma_0$ . Since the distance between  $\Gamma_1$  and  $\Gamma_2$  is known, it is sufficient to find the surface  $\Gamma_0$ . To solve this problem, consider an arbitrary ball  $B_R(\tilde{x}_j)$ ,  $\tilde{x}_j \in \Gamma_0$ . Let  $M_0$  be the number of points from  $B_R(\tilde{x}_j)$  situated between  $\Gamma_1$  and  $\Gamma_2$ . Clearly,  $M_0$  is the same for any ball  $B_R(\tilde{x}_j)$ ,  $\tilde{x}_j \in \Gamma_0$ . Thus if one arranges the observed values of the signal from  $B_R(\tilde{x}_j)$  in ascending order  $u_{i_1} \leq u_{i_2} \leq \dots \leq u_{i_{2M}}$ , it is likely that the last  $M_0$  points  $\{x_{i_k}\}_{k=2M-M_0+1}^{2M}$  lie near the diameter plane  $S_j^{(0)} = B_R(\tilde{x}_j) \cap \Gamma_0$ . This means that the quantity

$$\Delta(\{x_{i_k}\}_{k=2M-M_0+1}^{2M}, S_j^{(0)}) := h \sum_{k=2M-M_0+1}^{2M} \text{dist}^2(x_{i_k}, S_j^{(0)}), \quad h := h_1 \dots h_p \quad (3.1)$$

is small compared to the case when the points  $\{x_{i_k}\}_{k=2M-M_0+1}^{2M}$  are uniformly distributed over  $B_R(\tilde{x}_j)$ . Since the actual orientation of  $S_j^{(0)}$  is not known, the natural measure of the closeness of the points  $\{x_{i_k}\}_{k=2M-M_0+1}^{2M}$  to a plane is given by

$$\Delta(\{x_{i_k}\}_{k=2M-M_0+1}^{2M}) := \min_{S; \tilde{x}_j \in S} \Delta(\{x_{i_k}\}_{k=2M-M_0+1}^{2M}, S). \quad (3.2)$$

The minimum on the right-hand side of (3.2) is computed over all planes containing the point  $\tilde{x}_j$ . Let  $\theta$  be a unit vector perpendicular to a plane  $S$ . Then one obtains from (3.1) and (3.2)

$$\Delta(\{x_{i_k}\}_{k=2M-M_0+1}^{2M}) := \min_{\theta \in S^{p-1}} h \sum_{k=2M-M_0+1}^{2M} \langle x_{i_k} - \tilde{x}_j, \theta \rangle^2, \quad (3.3)$$

where  $S^{p-1}$  is the unit sphere in  $\mathbb{R}^p$ ,  $\langle x, y \rangle = \sum_{i=1}^p x_i y_i$  is the scalar product in  $\mathbb{R}^p$ . Define the  $p \times p$  matrix  $B_{M_0}$  by

$$B_{M_0} := (b_{mn}), \quad b_{mn} = h \sum_{k=2M-M_0+1}^{2M} (x_{i_k}^{(m)} - \tilde{x}_j^{(m)}) \cdot (x_{i_k}^{(n)} - \tilde{x}_j^{(n)}), \quad 1 \leq m, n \leq p. \quad (3.4)$$

Note that  $B_{M_0}$  is a positive semi-definite symmetric matrix with real-valued entries. Hence (3.3) can be rewritten as

$$\Delta(\{x_{i_k}\}_{k=2M-M_0+1}^{2M}) = \min_{\theta \in S^{p-1}} (B_{M_0} \theta, \theta).$$

The solution to (3.3) is  $\Delta(\{x_{i_k}\}_{k=2M-M_0+1}^{2M}) = \lambda_{\min}(B_{M_0})$ , where  $\lambda_{\min}(B_{M_0})$  is the minimal eigenvalue of the matrix  $B_{M_0}$ . The corresponding eigenvector  $\theta_0$  can be used as the estimate of the orientation of  $\Gamma_0$  at  $\tilde{x}_j$ . From (3.1) – (3.4) it follows that if  $\lambda_{\min}(B_{M_0})$  is small, the points  $\{x_{i_k}\}_{k=2M-M_0+1}^{2M}$  lie close to some plane, and, similarly, if  $\lambda_{\min}(B_{M_0})$  is large, the points  $\{x_{i_k}\}_{k=2M-M_0+1}^{2M}$  are distributed uniformly over  $B_R(\tilde{x}_j)$ . As a result we came to the following algorithm.

*Step 1'*. Fix a point  $\tilde{x}_j$ . Arrange the observed values from  $B_R(\tilde{x}_j)$  in ascending order  $u_{i_1} \leq u_{i_2} \leq \dots \leq u_{i_{2M}}$  with  $u_{i_k} := u(x_{i_k})$  and calculate the minimal eigenvalue  $\lambda_{\min}(B_{M_0})$ , where the matrix  $B_{M_0}$  is given by (3.4).

*Step 2'*. Select the threshold  $\Lambda_0 > 0$ , as described in Step 0' below. Compare  $\lambda_{\min}(B_{M_0})$  with the threshold  $\Lambda_0$ . If  $\lambda_{\min}(B_{M_0}) \geq \Lambda_0$ , the hypothesis " $\tilde{x}_j \notin \Gamma_0$ " is accepted. If  $\lambda_{\min}(B_{M_0}) < \Lambda_0$ , the hypothesis " $\tilde{x}_j \in \Gamma_0$ " is accepted and the orientation of  $\Gamma_0$  inside  $B_R(\tilde{x}_j)$  is given by the eigenvector corresponding to  $\lambda_{\min}(B_{M_0})$ .

Similar to Section 2, the threshold  $\Lambda_0$  is determined from the relation:

$$P\{\lambda_{\min}(B_{M_0}) < \Lambda_0 \mid B_R(\tilde{x}_j) \cap \Gamma_0 = \emptyset\} = \epsilon$$

by the Monte Carlo method according to Step 0'.

*Step 0'*. Fix an arbitrary ball  $B_R(\tilde{x}_j)$ . Model a sufficiently large number of the samples of  $u_i$  with  $s_i \equiv 0$  and an arbitrary fixed common PDF for  $n_i$ . Using Step

1' compute  $\lambda_{min}(B_{M_0})$  for each sample. Select  $\Lambda_0$  so that  $\lambda_{min}(B_{M_0}) < \Lambda_0$  in  $100\epsilon\%$  cases.

It is easy to see that the above algorithm (Steps 0', 1', 2') is also applicable in two other cases: a) the values of the signal between  $\Gamma_1$  and  $\Gamma_2$  are smaller than the values from the neighborhoods adjoint to  $\Gamma_1$  and  $\Gamma_2$ . In this case one calculates matrix  $B_{M_0}$  using the first  $M_0$  points  $\{x_{i_k}\}_{k=1}^{M_0}$  from the window (with smallest values of the signal); b) the values of the signal between  $\Gamma_1$  and  $\Gamma_2$  are smaller than the values from one neighborhood (say, adjoint to  $\Gamma_1$ ) and larger than values from another neighborhood (adjoint to  $\Gamma_2$ ). In this case one uses  $M_0$  middle points  $\{x_{i_k}\}_{k=N-M_0/2+1}^{N+M_0/2}$  in (3.4).

#### 4. GENERALIZATION OF THE ALGORITHMS

The algorithms described in Sections 2 and 3 can be generalized in the following way. Suppose one wants to recover a domain  $U$  which satisfies the following properties:

- 1) in a neighborhood of  $U$  one has  $s(x_1) > s(x_2)$  if  $x_1 \in U$ ,  $x_2 \notin U$ , i.e. the boundary of  $U$  is the discontinuity surface (or a part of it);
- 2) there exists a set of points  $\Gamma_0 \subset U$  such that for any  $\tilde{x}_{j_1}, \tilde{x}_{j_2} \in \Gamma_0$  the sets  $B_R(\tilde{x}_{j_1}) \cap U$  and  $B_R(\tilde{x}_{j_2}) \cap U$  can be obtained from each other by translating and rotating;
- 3)  $B_R(\tilde{x}_j) \setminus U \neq \emptyset$  for any  $\tilde{x}_j \in \Gamma_0$ ;
- 4) the set  $U$  can be uniquely determined from  $\Gamma_0$ .

Let us illustrate the geometrical meaning of assumptions 1) – 4) by examples. In the case of edge detection let  $U$  be the domain such that the discontinuity surface  $\Gamma$  is its boundary and let  $\Gamma_0 = \Gamma$ . Clearly properties 1 and 3 are satisfied. Property 2 also holds if  $\Gamma$  is locally flat. Furthermore the set  $U$  is uniquely determined by its boundary  $\Gamma (= \Gamma_0)$ . In the case of thin line detection let  $U$  be the domain between  $\Gamma_1$  and  $\Gamma_2$ , and  $\Gamma_0$  be the "middle surface" (see Section 3). As above, property 1 holds. Property 3 also holds since  $d < R$ . Property 2 is satisfied because we assumed that  $\Gamma_1$  and  $\Gamma_2$  are locally flat. Property 4 is satisfied because  $d$  and  $\Gamma_0$  determine  $\Gamma_1, \Gamma_2$  and, hence,  $U$  uniquely.

Thus for each  $\tilde{x}_j$  one wishes to have a decision rule for choosing between the two hypotheses: " $\tilde{x}_j \in \Gamma_0$ " and " $\tilde{x}_j \notin \Gamma_0$ ". The decision can be made using the following argument. Fix an arbitrary  $\tilde{x}_j$  and arrange the observed values of the signal from  $B_R(\tilde{x}_j)$  in ascending order  $u_{i_1} \leq u_{i_2} \leq \dots \leq u_{i_{2M}}$ . If  $\tilde{x}_j \in \Gamma_0$ , then, by properties 1, 2 and 3, the set of points  $\{x_{i_k}\}_{k=2M-M_0+1}^{2M}$  ( $M_0$  is the number of the grid points inside  $B_R(\tilde{x}_j) \cap U$ ) is likely to exhibit some nonuniformity. Namely, most of these points belong to  $B_R(\tilde{x}_j) \cap U$ , and only a small part of them lie outside  $B_R(\tilde{x}_j) \cap U$ . If  $B_R(\tilde{x}_j) \cap U = \emptyset$ , the points  $\{x_{i_k}\}_{k=2M-M_0+1}^{2M}$  are uniformly distributed inside  $B_R(\tilde{x}_j)$ . Let  $N$  be the number of points from  $\{x_{i_k}\}_{k=2M-M_0+1}^{2M}$  lying outside  $B_R(\tilde{x}_j) \cap U$  (or outside any fixed set obtained from  $B_R(\tilde{x}_j) \cap U$  by rotation around  $\tilde{x}_j$ ). Suppose there exists a continuous functional  $\lambda(\{x_{i_k}\}_{k=2M-M_0+1}^{2M}, \tilde{x}_j)$  with the properties

$$\lambda(\{x_{i_k}\}_{k=2M-M_0+1}^{2M}, \tilde{x}_j) \Big|_{N=N_1} > \lambda(\{x_{i_k}\}_{k=2M-M_0+1}^{2M}, \tilde{x}_j) \Big|_{N=N_2} \quad \text{if } N_1 < N_2, \quad (4.1)$$

$$\exists \lim_{M \rightarrow \infty} \lambda(\{x_{i_k}\}_{k=2M-M_0+1}^{2M}, \tilde{x}_j) := \hat{\lambda}(\gamma, \tilde{x}_j) \quad \text{if } \frac{N}{M_0} \rightarrow \gamma, \quad 0 \leq \gamma \leq 1, \quad (4.2)$$

$$\hat{\lambda}(\gamma_1, \tilde{x}_j) > \hat{\lambda}(\gamma_2, \tilde{x}_j) \quad \text{if } 0 \leq \gamma_1 < \gamma_2 \leq 1. \quad (4.3)$$

Then, on the basis of (4.1), one constructs the following algorithm.

*Step 1''.* Fix a point  $\tilde{x}_j$ . Arrange the observed values from  $B_R(\tilde{x}_j)$  in ascending order  $u_{i_1} \leq \dots \leq u_{i_{2M}}$  with  $u_{i_k} := u(x_{i_k})$  and calculate the functional  $\lambda_j = \lambda(\{x_{i_k}\}_{k=2M-M_0+1}^{2M}, \tilde{x}_j)$ .

*Step 2''.* Select the threshold  $\Lambda > 0$ , as described in Step 0'' below. Compare  $\lambda_j$  with the threshold  $\Lambda$ . If  $\lambda_j \leq \Lambda$ , the hypothesis " $\tilde{x}_j \notin \Gamma_0$ " is accepted. If  $\lambda_j > \Lambda$ , the hypothesis " $\tilde{x}_j \in \Gamma_0$ " is accepted and, if necessary, the orientation of  $\Gamma_0$  at  $\tilde{x}_j$  is calculated according to the nature of the functional  $\lambda(\{x_{i_k}\}_{k=2M-M_0+1}^{2M}, \tilde{x}_j)$ .

Similarly, the threshold  $\Lambda$  is determined from:

$$P\{\lambda(\{x_{i_k}\}_{k=2M-M_0+1}^{2M}, \tilde{x}_j) > \Lambda \mid B_R(\tilde{x}_j) \cap \Gamma_0 = \emptyset\} = \epsilon$$

by the Monte Carlo method:

*Step 0''.* Fix an arbitrary ball  $B_R(\tilde{x}_j)$ . Model a sufficiently large number of the samples of  $u_i$  with  $s_i \equiv 0$  and an arbitrary fixed common PDF for  $n_i$ . Using Step 1'' compute  $\lambda(\{x_{i_k}\}_{k=2M-M_0+1}^{2M}, \tilde{x}_j)$  for each sample. Select  $\Lambda$  so that  $\lambda(\{x_{i_k}\}_{k=2M-M_0+1}^{2M}, \tilde{x}_j) > \Lambda$  in 100% cases.

It is clear that the algorithms from Sections 2 and 3 fit well into the above scheme. In this Section we present one more example which shows that this scheme is sufficiently general and can be used for the solution of other pattern recognition problems. Let us consider the spot detection problem. Suppose one is looking for a "spot" which can be represented by a ball  $B_d(\tilde{x}_{j_0})$  with known radius  $d < R$ , such that the values of the signal inside  $B_d(\tilde{x}_{j_0})$  are larger than in its neighborhood. Then one sees that the properties 1–4 are satisfied. Here  $U = B_d(\tilde{x}_{j_0})$ ,  $M_0$  is the number of grid points inside  $B_d(\tilde{x}_{j_0})$  and  $\Gamma_0 = \tilde{x}_{j_0}$ . The functional  $\lambda(\{x_{i_k}\}_{k=2M-M_0+1}^{2M}, \tilde{x}_j)$  can be easily constructed:

$$\lambda(\{x_{i_k}\}_{k=2M-M_0+1}^{2M}, \tilde{x}_j) := -h \sum_{k=2M-M_0+1}^{2M} |x_{i_k} - \tilde{x}_j|^2. \quad (4.4)$$

It is easy to see that the functional (4.4) satisfies (4.1)–(4.3). We put the minus sign in front of the summation in (4.4) so that the value of the functional is the largest when  $\tilde{x}_j = \tilde{x}_{j_0}$ . Hence, using Steps 0'', 1'', 2'' with the functional defined by (4.4), one finds the position of the spot  $\tilde{x}_{j_0}$ .

Clearly, the algorithm given by Steps 0'', 1'', 2'' can be modified for the case when in a neighborhood of  $U$  one has  $s(x_1) < s(x_2)$  if  $x_1 \in U$ ,  $x_2 \notin U$ . The only difference is in the Step 1'', where one arranges points in the descending order.

## 5. JUSTIFICATION OF THE EDGE DETECTION ALGORITHM

In this Section we prove the asymptotic consistency of the edge detection algorithm as  $\max(h_1, \dots, h_p) \rightarrow 0$  ( $M \rightarrow \infty$ ).

**Theorem 1.** Fix any  $\epsilon, \delta, D > 0$  and consider  $\tilde{x}_j \in \Gamma$ . Then:

- a) the probability of the decision " $\tilde{x}_j \notin \Gamma$ " goes to zero as  $M \rightarrow \infty$ ;  
b) the probability of an angular error larger than  $\delta$  in determining the orientation of  $\Gamma$  at the point  $\tilde{x}_j$  (by the algorithm described in Steps 1 and 2) goes to zero as  $M \rightarrow \infty$ .

For the proof of the theorem we need some auxiliary results which are formulated in Lemmas 1 and 2 below. Then the proof of Theorem 1 is given.

Consider an arbitrary point  $\tilde{x}_j \in \Gamma$ . In this case  $B_R(\tilde{x}_j)$  is split by  $\Gamma$  into two equal parts  $B_R^-(\tilde{x}_j)$  and  $B_R^+(\tilde{x}_j)$ . At the same time, since the signal is corrupted by noise, some of the points from  $B_R^-(\tilde{x}_j)$  will be erroneously classified as belonging to the set  $\{x_{i_k}\}_{k=M+1}^{2M}$ , and the equal number of points from  $B_R^+(\tilde{x}_j)$  will be erroneously classified as belonging to the set  $\{x_{i_k}\}_{k=1}^M$ . Consequently, the orientation of  $\Gamma \cap B_R(\tilde{x}_j)$  will be determined with some error. The points which belong to  $B_R^-(\tilde{x}_j)$  ( $B_R^+(\tilde{x}_j)$ ) and which are erroneously classified as being from  $\{x_{i_k}\}_{k=M+1}^{2M}$  ( $\{x_{i_k}\}_{k=1}^M$ ) will be called the misclassified points. To find the probability  $P(N, M)$  of misclassifying  $N$  or more points, let us order the points from each of the halfballs separately:

$$\begin{aligned} v_{i_1} \leq v_{i_2} \leq \dots \leq v_{i_M}, \quad v_{i_{M+1}} \leq v_{i_{M+2}} \leq \dots \leq v_{i_{2M}}, \\ x_{i_1}, \dots, x_{i_M} \in B_R^-(\tilde{x}_j), \quad x_{i_{M+1}}, \dots, x_{i_{2M}} \in B_R^+(\tilde{x}_j). \end{aligned} \quad (5.1)$$

Here we used the notation  $v_{i_k}$  for the data instead of  $u_{i_k}$  to stress the fact that the ranking (5.1) is different from the ranking used in Step 1. Note that now we *do not* have the inequality  $v_{i_M} \leq v_{i_{M+1}}$ . It is easy to see that there will be  $N$  or more misclassified points in each halfball if and only if

$$v_{i_{M+N}} < v_{i_{M-N+1}}. \quad (5.2)$$

Let  $s^- \approx s(x)$ ,  $x \in B_R^-(\tilde{x}_j)$ ,  $s^+ \approx s(x)$ ,  $x \in B_R^+(\tilde{x}_j)$ ,  $D := s^+ - s^-$ , that is  $D$  is the magnitude of the jump of the signal at the point  $\tilde{x}_j$ . Let  $f(\tau)$  be the common PDF of random errors  $n_i$ . Then, as it follows from (2.1) and (5.1),  $v_{i_k}$ ,  $1 \leq k \leq M$ , is the  $k$ -th order statistic of the random sample of the size  $M$  from a distribution with the PDF  $f(\tau - s^-)$ . Similarly,  $v_{i_{M+k}}$ ,  $1 \leq k \leq M$ , is the  $k$ -th order statistic of the random sample of the size  $M$  from a distribution with the PDF  $f(\tau - s^+)$ . Using the known formula for the distribution of order statistics [3], one obtains from (5.2):

$$P(N, M) := P\{v_{i_{M+N}} < v_{i_{M-N+1}}\} = \int F_N(t - D) dF_{M-N+1}(t), \quad (5.3)$$

where  $\int := \int_{-\infty}^{+\infty}$  and

$$\begin{aligned} F_k(t) &= \int_{-\infty}^t f_k(\tau) d\tau, \quad f_k(\tau) = \frac{M!}{(k-1)!(M-k)!} [F(\tau)]^{k-1} [1 - F(\tau)]^{M-k} f(\tau), \\ F(\tau) &= \int_{-\infty}^{\tau} f(\xi) d\xi. \end{aligned} \quad (5.4)$$

In (5.3) we used the independence of the random variables  $v_{i_{M+N}}$  and  $v_{i_{M-N+1}}$ . To find the asymptotics of  $P(N, M)$  as  $M \rightarrow \infty$ , fix  $\gamma$ ,  $0 < \gamma < 1$ , and let  $N = \gamma M$ . Using the known asymptotics of the order statistics [2], [3]:

$$\begin{aligned} \left| F_N(t-D) - \Phi_1(t) \right| &\leq \frac{3}{\sqrt{M\gamma(1-\gamma)}} := C_M, \quad \Phi_1(t) := \Phi \left( \sqrt{M} \frac{F(t-D) - \gamma}{\sqrt{\gamma(1-\gamma)}} \right), \\ \left| F_{M-N+1}(t) - \Phi_2(t) \right| &\leq C_M, \quad \Phi_2(t) := \Phi \left( \sqrt{M} \frac{F(t) - (1-\gamma)}{\sqrt{\gamma(1-\gamma)}} \right), \end{aligned} \quad (5.5)$$

where  $\Phi(x) = 1/\sqrt{2\pi} \int_{-\infty}^x \exp(-t^2/2) dt$  is the standard normal distribution, one obtains from (5.3) and (5.5)

$$\begin{aligned} P(N, M) &= \int F_N(t-D) dF_{M-N+1}(t) \leq \int (\Phi_1(t) + C_M) dF_{M-N+1}(t) = \\ &\int \Phi_1(t) dF_{M-N+1}(t) + C_M = \Phi_1(t) F_{M-N+1}(t) \Big|_{t=-\infty}^{+\infty} - \int F_{M-N+1}(t) d\Phi_1(t) + \\ &C_M \leq \Phi \left( \sqrt{M} \frac{1-\gamma}{\gamma} \right) - \int (\Phi_2(t) - C_M) d\Phi_1(t) + C_M \leq \Phi \left( \sqrt{M} \frac{1-\gamma}{\gamma} \right) - \\ &\int \Phi_2(t) d\Phi_1(t) + 2C_M \leq \int \Phi_1(t) d\Phi_2(t) + 2C_M + o(\exp(-cM)), \end{aligned} \quad (5.6)$$

where  $c = \text{const} > 0$  (see (A.1)). It follows from (5.5) that  $C_M \rightarrow 0$  as  $M \rightarrow \infty$ . To find the asymptotics of the integral on the right-hand side of (5.6), introduce the new variable  $y = \sqrt{M}(F(t) - (1-\gamma))/\sqrt{\gamma(1-\gamma)}$ . This yields

$$\begin{aligned} \int \Phi_1(t) d\Phi_2(t) &= \frac{1}{\sqrt{2\pi}} \int_{-\sqrt{M\frac{1-\gamma}{\gamma}}}^{\sqrt{M\frac{1-\gamma}{\gamma}}} \Phi \left( \sqrt{\frac{M}{\gamma(1-\gamma)}} \left\{ F \left[ F^{-1} \left( 1 - \gamma + \right. \right. \right. \right. \\ &\quad \left. \left. \left. y \sqrt{\frac{\gamma(1-\gamma)}{M}} \right) - D \right] - \gamma \right\} \right) \exp(-y^2/2) dy \rightarrow \\ &\frac{1}{\sqrt{2\pi}} \int \Phi \left( \sqrt{\frac{M}{\gamma(1-\gamma)}} \{ F(\xi_0) - \gamma \} \right) \exp(-y^2/2) dy \xrightarrow{M \rightarrow \infty} \begin{cases} 0, & F(\xi_0) < \gamma \\ 1, & F(\xi_0) > \gamma, \end{cases} \end{aligned} \quad (5.7)$$

where

$$\xi_0 := F^{-1}(1-\gamma) - D. \quad (5.8)$$

The rate of convergence in (5.7) is estimated in the Appendix. Substituting (5.8) into (5.7), one gets after some calculations

$$P(\gamma M, M) \rightarrow \theta \{ [F^{-1}(1-\gamma) - F^{-1}(\gamma)] - D \} \quad \text{as } M \rightarrow \infty, \quad (5.9)$$

where  $\theta(t) = \begin{cases} 0, & t < 0, \\ 1, & t > 0. \end{cases}$  Define  $\gamma_0$  by

$$\gamma_0 := \inf_{\substack{0 < \gamma < 1 \\ F^{-1}(1-\gamma) - F^{-1}(\gamma) < D}} \gamma. \quad (5.10)$$

Since the function  $F^{-1}(1 - \gamma) - F^{-1}(\gamma)$  is strictly decreasing for  $0 < \gamma < 1$  and  $D > 0$ , one obtains from (5.9) and (5.10)

$$P(\gamma M, M) \rightarrow \theta(\gamma_0 - \gamma) \quad \text{as } M \rightarrow \infty \quad \text{with } 0 \leq \gamma_0 < 1/2. \quad (5.11)$$

Now take any point  $\tilde{x}_j$  such that  $B_R(\tilde{x}_j) \cap \Gamma = \emptyset$  and divide  $B_R(\tilde{x}_j)$  by an arbitrary hyperplane (passing through  $\tilde{x}_j$ ) into two equal parts  $B_R^-(\tilde{x}_j)$  and  $B_R^+(\tilde{x}_j)$ . In this case the asymptotics of the probability of misclassifying  $\gamma M$  points can be obtained by putting  $D = 0$  in (5.9):

$$P(\gamma M, M) \rightarrow \theta(F^{-1}(1 - \gamma) - F^{-1}(\gamma)) \quad \text{as } M \rightarrow \infty. \quad (5.12)$$

From (5.10) it follows that in the case  $D = 0$  one has  $\gamma_0 = 1/2$  and (5.12) reduces to

$$P(\gamma M, M) \rightarrow \theta(1/2 - \gamma) \quad \text{as } M \rightarrow \infty. \quad (5.13)$$

Recall that  $N$  is the number of the misclassified points. It is not hard to see that (5.11) and (5.13) imply

**Lemma 1.** *Fix an arbitrary  $\delta > 0$ . If  $\tilde{x}_j \in \Gamma$ , there exists  $\gamma_0$ ,  $0 \leq \gamma_0 < 1/2$ , such that*

$$P\left\{\left|\frac{N}{M} - \gamma_0\right| > \delta\right\} \rightarrow 0 \quad \text{as } \max(h_1, \dots, h_p) \rightarrow 0 \quad (M \rightarrow \infty).$$

*If  $B_R(\tilde{x}_j) \cap \Gamma = \emptyset$ , divide  $B_R(\tilde{x}_j)$  arbitrarily into two parts  $B_R^-(\tilde{x}_j)$  and  $B_R^+(\tilde{x}_j)$  having equal number of points and let  $N$  be the number of points from the set  $\{x_{i_k}\}_{k=M+1}^{2M}$  lying in  $B_R^-(\tilde{x}_j)$ . Then one has*

$$P\left\{\left|\frac{N}{M} - 1/2\right| > \delta\right\} \rightarrow 0 \quad \text{as } \max(h_1, \dots, h_p) \rightarrow 0 \quad (M \rightarrow \infty).$$

Now let us consider the halfball  $B_R^-(\tilde{x}_j)$  and assume that  $0 < \gamma_0 \leq 1/2$ . Fix an arbitrary nonempty open set  $U \subset B_R^-(\tilde{x}_j)$  and let  $M_U$  be the number of the grid points in  $U$ . It is clear that each grid point can be either misclassified or correctly classified, in other words, there can not be two misclassified points at any given grid node, which is exactly the assumption of the Fermi-Dirac statistics [9, pp. 31–36]. Hence the probability of the event "there are  $N_U$  misclassified points out of the total number  $N$  inside the set  $U$ " is given by

$$P(N_U, N; M_U, M) = \frac{C_{N_U}^{M_U} C_{N-N_U}^{M-M_U}}{C_N^M}. \quad (5.14)$$

As a direct consequence of the Stirling formula  $n! \sim \sqrt{2\pi n} n^n e^{-n}$ , one has

$$C_{sn}^n \sim \frac{1}{\sqrt{2\pi n(1-s)s}} \frac{1}{((1-s)^{1-s} s^s)^n} \quad \text{as } n \rightarrow \infty, \quad (5.15)$$

where  $s$ ,  $0 < s < 1$ , is an arbitrary fixed number. To find the asymptotics of the probability (5.14) as  $M \rightarrow \infty$ , assume that

$$N_U = \alpha_U N, \quad M_U = \beta_U M, \quad N = \gamma_0 M. \quad (5.16)$$

Substituting (5.16) into (5.14) and using (5.15), one obtains

$$P(\alpha_U \gamma_0 M, \gamma_0 M; \beta_U M, M) \rightarrow \frac{c}{\sqrt{M}} r^M \quad \text{as } M \rightarrow \infty, \quad (5.17)$$

where

$$r = \frac{(1 - \gamma_0)^{1 - \gamma_0} \gamma_0^{\gamma_0}}{[(1 - s_1)^{1 - s_1} s_1^{s_1}]^{\beta_U} [(1 - s_2)^{1 - s_2} s_2^{s_2}]^{1 - \beta_U}}, \quad s_1 = \frac{\alpha_U \gamma_0}{\beta_U}, \quad s_2 = \frac{(1 - \alpha_U) \gamma_0}{1 - \beta_U}. \quad (5.18)$$

Here  $c > 0$  is some constant which does not depend on  $M$ . From (5.18) one has

$$\ln r = \beta_U H(s_1) + (1 - \beta_U) H(s_2) - H(\beta_U s_1 + (1 - \beta_U) s_2), \quad (5.19)$$

where  $H(s) = -s \ln s - (1 - s) \ln(1 - s)$  is the entropy [9, pp. 49–53]. In (5.19) we used the identity  $\gamma_0 = \beta_U s_1 + (1 - \beta_U) s_2$ . Since  $-H(s)$  is a strictly convex function, one has  $r = 1$  if and only if  $s_1 = s_2$  ( $\beta_U \neq 0$ ,  $\beta_U \neq 1$ , because  $U$  was chosen to be a proper nonempty subset of  $B_R^-(\tilde{x}_j)$ ). In all other cases  $r < 1$ . From (5.18) one sees that the condition  $s_1 = s_2$  is equivalent to the condition  $\alpha_U = \beta_U$ . This, together with (5.17), almost immediately yields

**Lemma 2.** *Fix an arbitrary  $\delta > 0$  and consider an arbitrary nonempty open set  $U \subset B_R^-(\tilde{x}_j)$ . Then one has*

$$P \left\{ \left| \frac{N_U}{N} - \frac{\text{Vol}(U)}{\text{Vol}(B_R^-(\tilde{x}_j))} \right| > \delta \right\} \rightarrow 0 \quad \text{as } M \rightarrow \infty, \quad (5.20)$$

where  $\text{Vol}(A)$  is the volume of  $A$ .

Note that Lemma 2 guarantees the asymptotically uniform distribution of the misclassified points over the corresponding halfball  $B_R^-(\tilde{x}_j)$  or  $B_R^+(\tilde{x}_j)$ . Hence the distribution of correctly classified points is also asymptotically uniform.

*Proof of Theorem 1.* Let  $M$  be sufficiently large and consider the misclassified points from the set  $\{x_{i_k}\}_{k=M+1}^{2M}$ , i.e. the points which lie in  $B_R^-(\tilde{x}_j)$ . Let  $N$  be the number of such points and  $K = \{k_1, \dots, k_N\}$  be the set of the subindices corresponding to them. By Lemma 1,  $N/M \rightarrow \gamma_0 < 1/2$ , and, by Lemma 2, these points are uniformly distributed over  $B_R^-(\tilde{x}_j)$ . Similarly, the number of correctly classified points  $M - N$  satisfies  $(M - N)/M \rightarrow 1 - \gamma_0 > 1/2$ , and they are uniformly distributed over  $B_R^+(\tilde{x}_j)$ , therefore the vector  $\tilde{x}_j - x_j^+$  must have only one nonzero component, which is perpendicular to  $\Gamma$  at  $\tilde{x}_j$ . Hence the distance between  $\tilde{x}_j$  and

the centroid  $x_j^+$  (with  $M \rightarrow \infty$ ) can be calculated as follows

$$\begin{aligned}
|\tilde{x}_j - x_j^+| &= \left| \frac{\sum_{k \in K} (x_{i_k}^{(1)} - \tilde{x}_j^{(1)}) + \sum_{\substack{k=1 \\ k \notin K}}^M (x_{i_k}^{(1)} - \tilde{x}_j^{(1)})}{M} \right| \rightarrow \\
&\left| \frac{\gamma_0 \sum_{k \in K} (x_{i_k}^{(1)} - \tilde{x}_j^{(1)}) \frac{hM}{N} + (1 - \gamma_0) \sum_{\substack{k=1 \\ k \notin K}}^M (x_{i_k}^{(1)} - \tilde{x}_j^{(1)}) \frac{hM}{M-N}}{hM} \right| \rightarrow \\
&\left| \frac{\gamma_0 \int_{\substack{x \in B_R(\tilde{x}_j) \\ x^{(1)} \leq \tilde{x}_j^{(1)}}} (x^{(1)} - \tilde{x}_j^{(1)}) d\hat{x} + (1 - \gamma_0) \int_{\substack{x \in B_R(\tilde{x}_j) \\ x^{(1)} \geq \tilde{x}_j^{(1)}}} (x^{(1)} - \tilde{x}_j^{(1)}) d\hat{x}}{\text{Vol}(B_R(\tilde{x}_j))} \right| = \\
&\frac{\gamma_0 \int_{-R}^0 t \omega_{p-1}(R^2 - t^2)^{\frac{p-1}{2}} dt + (1 - \gamma_0) \int_0^R t \omega_{p-1}(R^2 - t^2)^{\frac{p-1}{2}} dt}{\int_0^R \omega_{p-1}(R^2 - t^2)^{\frac{p-1}{2}} dt} = \\
&(1 - 2\gamma_0) \frac{\int_0^1 t(1 - t^2)^{\frac{p-1}{2}} dt}{\int_0^1 (1 - t^2)^{\frac{p-1}{2}} dt} R = (1 - 2\gamma_0) \frac{2}{(p+1)B(\frac{p+1}{2}, \frac{1}{2})} R > 0 \tag{5.21}
\end{aligned}$$

since  $\gamma_0 < 1/2$ , where  $h = h_1 \cdot \dots \cdot h_p$ ,  $d\hat{x} := dx^{(1)} \dots dx^{(p)}$ ,  $B(a, b)$  is the beta-function and  $\omega_{p-1} = \pi^{(p-1)/2} / \Gamma((p+1)/2)$  is the volume of the unit ball in  $\mathbb{R}^{p-1}$ . The third line in (5.21) follows from the second one by observing that the elementary volume per point in  $B_R^-(\tilde{x}_j)$  is  $hM/N$ , while in  $B_R^+(\tilde{x}_j)$  it is  $hM/(M-N)$ . In (5.21) we assumed that the axis  $x^{(1)}$  is perpendicular to  $\Gamma$ . At the same time, since  $\gamma_0 \rightarrow 1/2$  for arbitrary  $\tilde{x}_k$ , such that  $B_R(\tilde{x}_k) \cap \Gamma = \emptyset$ , one has  $|x_k^+ - \tilde{x}_k| \rightarrow 0$  as  $M \rightarrow \infty$ . Thus  $A_0$  (the threshold value determined in Step 0) goes to zero as  $M \rightarrow \infty$ . This and (5.21) yield the first assertion of Theorem 1.

The second assertion of Theorem 1 follows immediately from Lemma 2 and (5.21), because if the distributions of the misclassified and correctly classified points are uniform, the line through  $\tilde{x}_j$  and the centroid  $x_j^+$  is perpendicular to  $\Gamma$ .

## 6. JUSTIFICATION OF THE ALGORITHM FOR THIN LINE DETECTION

Let us call the points from  $\{x_{i_k}\}_{k=2M-M_0+1}^{2M}$  ( $\{x_{i_k}\}_{k=1}^{2M-M_0}$ ) which belong to  $B_R(\tilde{x}_j) \setminus U$  ( $B_R(\tilde{x}_j) \cap U$ ),  $\tilde{x}_j \in \Gamma_0$  "the misclassified points". Here  $U$  is the domain between  $\Gamma_1$  and  $\Gamma_2$ . Denote  $V_1 := \text{Vol}(B_R(\tilde{x}_j) \cap U)$ ,  $V_2 := \text{Vol}(B_R(\tilde{x}_j) \setminus U) = \text{Vol}(B_R(\tilde{x}_j)) - V_1$ ,  $\eta_0 := V_2 / \text{Vol}(B_R(\tilde{x}_j))$  and let  $N$  be the number of the misclassified points. Let us assume for simplicity that for each pair of neighboring points  $(x_1, x_2)$ ,  $x_1 \in \Gamma_1$ ,  $x_2 \in \Gamma_2$ , one has  $D(x_1) = D(x_2)$ , where  $D(x_i)$  is the jump magnitude at the point  $x_i$ . Two points  $x_1 \in \Gamma_1$ ,  $x_2 \in \Gamma_2$  are called neighboring if  $x_2 - x_1 \perp \Gamma_1$ ,  $x_2 - x_1 \perp \Gamma_2$ . Following the argument used in the proof of Lemma 1, one can prove

**Lemma 3.** Fix an arbitrary  $\delta > 0$ . If  $\tilde{x}_j \in \Gamma_0$ , there exists  $\gamma_0$ ,  $0 \leq \gamma_0 < \eta_0$ , such that

$$P \left\{ \left| \frac{N}{M_0} - \gamma_0 \right| > \delta \right\} \rightarrow 0 \quad \text{as} \quad \max(h_1, \dots, h_p) \rightarrow 0 \quad (M \rightarrow \infty). \quad (6.1)$$

If  $B_R(\tilde{x}_j) \cap U = \emptyset$ , divide  $B_R(\tilde{x}_j)$  arbitrarily into two parts  $U'$  and  $B_R(\tilde{x}_j) \setminus U'$  having  $M_0$  and  $2M - M_0$  grid points correspondingly ( $\text{Vol}(U') = V_1$ ). Let  $N$  be the number of points from  $\{x_{i_k}\}_{k=2M-M_0+1}^{2M}$  lying in  $B_R(\tilde{x}_j) \setminus U'$ . Then one has

$$P \left\{ \left| \frac{N}{M_0} - \eta_0 \right| > \delta \right\} \rightarrow 0 \quad \text{as} \quad \max(h_1, \dots, h_p) \rightarrow 0 \quad (M \rightarrow \infty). \quad (6.2)$$

Recall that  $\epsilon$  is the probability of "false alarm", which is used in Step 0'.

**Theorem 2.** Fix any pair of neighboring points  $(x_1, x_2)$ ,  $x_1 \in \Gamma_1$ ,  $x_2 \in \Gamma_2$ ,  $\epsilon, \delta > 0$  and consider  $\tilde{x}_j = (x_1, x_2) \cap \Gamma_0$ . Then, for any  $D(x_1) = D(x_2) = D > 0$ , one has: a) the probability of the decision " $\tilde{x}_j \notin \Gamma_0$ " goes to zero as  $M \rightarrow \infty$ ; b) the probability of an angular error larger than  $\delta$  in determining the orientation of  $\Gamma_0$  at the point  $\tilde{x}_j$  (by the algorithm described in Steps 1' and 2') goes to zero as  $M \rightarrow \infty$ .

*Proof.* In what follows we assume that  $M$  is sufficiently large and  $\max(h_1, \dots, h_p) \ll R$ . Without loss of generality one can make the following assumptions:  $\Gamma_1 = \{x : x^{(1)} = -d\}$ ,  $\Gamma_2 = \{x : x^{(1)} = +d\}$ ,  $\tilde{x}_j = 0$ . Hence  $\Gamma_0 = \{x : x^{(1)} = 0\}$ . Let  $\theta = (\theta^{(1)}, \dots, \theta^{(p)})$  and recall that  $d\hat{x} = dx^{(1)} \dots dx^{(p)}$ . Using Lemmas 2 and 3 and replacing the sum in (3.3) by the integral, one obtains

$$\begin{aligned} h \sum_{k=2M-M_0+1}^{2M} \langle x_{i_k} - \tilde{x}_j, \theta \rangle^2 &= \\ (1 - \gamma_0) \int_{B_R(\tilde{x}_j) \cap U} \langle x - \tilde{x}_j, \theta \rangle^2 d\hat{x} + \gamma_0 \frac{V_1}{V_2} \int_{B_R(\tilde{x}_j) \setminus U} \langle x - \tilde{x}_j, \theta \rangle^2 d\hat{x} &= \\ (1 - \gamma_0) \int_{\substack{|x| \leq R \\ |x^{(1)}| \leq d}} \left( \sum_{i=1}^p x^{(i)} \theta^{(i)} \right)^2 d\hat{x} + \gamma_0 \frac{V_1}{V_2} \int_{\substack{|x| \leq R \\ |x^{(1)}| > d}} \left( \sum_{i=1}^p x^{(i)} \theta^{(i)} \right)^2 d\hat{x} &= \\ \sum_{i=1}^p (\theta^{(i)})^2 \left\{ (1 - \gamma_0) \int_{\substack{|x| \leq R \\ |x^{(1)}| \leq d}} (x^{(i)})^2 d\hat{x} + \gamma_0 \frac{V_1}{V_2} \int_{\substack{|x| \leq R \\ |x^{(1)}| > d}} (x^{(i)})^2 d\hat{x} \right\} &= \\ \sum_{i=1}^p (\theta^{(i)})^2 \left\{ (1 - \gamma_0 - \gamma_0 \frac{V_1}{V_2}) \int_{\substack{|x| \leq R \\ |x^{(1)}| \leq d}} (x^{(i)})^2 d\hat{x} \right\} + \gamma_0 \frac{V_1}{V_2} \int_{|x| \leq R} \langle x, \theta \rangle^2 d\hat{x}, & \quad (6.3) \end{aligned}$$

where we have used several times that the integrals of the terms proportional to  $x^{(i)}x^{(j)}$ ,  $i \neq j$ , vanish because of the symmetry. Clearly, the second integral in (6.3) does not depend on  $\theta$ . Since  $|\theta| = 1$  and  $\gamma_0 < \eta_0 = V_2/(V_1 + V_2)$  (as it is stated in Lemma 3), it can be easily seen that the right-hand side of (6.3) attains

minimum when  $\theta = \theta_0 := (1, 0, \dots, 0)$ , that is, when  $\theta_0 \perp \Gamma_0$ . This proves the second assertion of Theorem 2. Furthermore, one sees that to prove the first assertion of this theorem it is sufficient to show that  $\Delta(\{x_{i_k}\}_{k=2M-M_0+1}^{2M})|_{\gamma=\gamma_0} < \Delta(\{x_{i_k}\}_{k=2M-M_0+1}^{2M})|_{\gamma=\eta_0}$ , i.e.

$$(1 - \gamma_0) \int_{\substack{|x| \leq R \\ |x^{(1)}| \leq d}} (x^{(1)})^2 d\hat{x} + \gamma_0 \frac{V_1}{V_2} \int_{\substack{|x| \leq R \\ |x^{(1)}| > d}} (x^{(1)})^2 d\hat{x} < \frac{V_1}{V_1 + V_2} \int_{|x| \leq R} (x^{(1)})^2 d\hat{x}. \quad (6.4)$$

Denoting

$$U_1 := \int_{\substack{|x| \leq R \\ |x^{(1)}| \leq d}} (x^{(1)})^2 d\hat{x}, \quad U_2 := \int_{\substack{|x| \leq R \\ |x^{(1)}| > d}} (x^{(1)})^2 d\hat{x}, \quad (6.5)$$

one gets that (6.4) is equivalent to

$$(1 - \gamma_0) \frac{U_1}{V_1} + \gamma_0 \frac{U_2}{V_2} < \frac{U_1 + U_2}{V_1 + V_2}.$$

Rewrite the last inequality as

$$(V_2(1 - \gamma_0) - V_1\gamma_0) \frac{U_2V_1 - U_1V_2}{V_1V_2(V_1 + V_2)} > 0.$$

Since  $\gamma_0 < V_2/(V_1 + V_2)$ , it is sufficient to prove that  $U_2V_1 - U_1V_2 > 0$  for  $0 < d < R$ . Similar to (5.21), one has

$$V_1 = 2 \int_0^d \omega_{p-1}(R^2 - t^2)^{\frac{p-1}{2}} dt, \quad V_2 = 2 \int_d^R \omega_{p-1}(R^2 - t^2)^{\frac{p-1}{2}} dt. \quad (6.6)$$

Using (6.5) and (6.6), introduce the function

$$\begin{aligned} \varphi(d) := U_2V_1 - U_1V_2 &= 2 \int_d^R t^2 \omega_{p-1}(R^2 - t^2)^{\frac{p-1}{2}} dt \cdot 2 \int_0^d \omega_{p-1}(R^2 - t^2)^{\frac{p-1}{2}} dt - \\ &2 \int_0^d t^2 \omega_{p-1}(R^2 - t^2)^{\frac{p-1}{2}} dt \cdot 2 \int_d^R \omega_{p-1}(R^2 - t^2)^{\frac{p-1}{2}} dt, \end{aligned} \quad (6.7)$$

where we simplified (6.5) by taking the integral over the variables  $x^{(2)}, \dots, x^{(p)}$ . One has

$$\varphi'(d) = 4\omega_{p-1}^2(R^2 - d^2)^{\frac{p-1}{2}} \left\{ \int_0^R t^2 (R^2 - t^2)^{\frac{p-1}{2}} dt - d^2 \int_0^R (R^2 - t^2)^{\frac{p-1}{2}} dt \right\}. \quad (6.8)$$

From (6.8) it follows that there exists  $d^*$ ,  $0 < d^* < R$ , such that

$$\varphi'(d) > 0 \text{ for } 0 \leq d < d^*, \quad \varphi'(d) \leq 0 \text{ for } d^* \leq d \leq R.$$

This, together with  $\varphi(0) = \varphi(R) = 0$ , shows that  $\varphi(d) > 0$  for  $0 < d < R$ . Theorem 2 is proved.

## 7. JUSTIFICATION OF THE GENERAL SCHEME

The general scheme described in Section 4 can be justified using the arguments similar to that of Sections 5 and 6. Let us suppose that for each  $\tilde{x}_j \in \Gamma_0$  the function  $s(x)$  is approximately constant inside  $B_R(\tilde{x}_j) \cap U$  and  $B_R(\tilde{x}_j) \setminus U$ :

$$s_i \approx s(x), x \in B_R(\tilde{x}_j) \cap U \quad \text{and} \quad s_o \approx s(x), x \in B_R(\tilde{x}_j) \setminus U$$

with  $D := s_i - s_o$  being the jump magnitude in the neighborhood of  $\tilde{x}_j$ .

**Theorem 3.** *Fix any  $\epsilon > 0$  and consider  $\tilde{x}_j \in \Gamma_0$ . If properties 1 – 3 and (4.1) – (4.3) are satisfied, the probability of the decision  $\tilde{x}_j \notin \Gamma_0$  goes to zero as  $M \rightarrow \infty$  for any  $D > 0$ .*

*Proof.* Let us note that Lemma 3 holds for arbitrary shape of the set  $U$ , provided that properties 1 – 3 from Section 4 are satisfied. Consider the case  $B_R(\tilde{x}_j) \cap U = \emptyset$ . Using (4.2) and (6.2), one gets

$$P\{|\lambda(\{x_{i_k}\}_{k=2M-M_0+1}, \tilde{x}_j) - \hat{\lambda}(\eta_0, \tilde{x}_j)| > \delta\} \rightarrow 0 \quad \text{as} \quad M \rightarrow \infty \quad (7.1)$$

for any fixed  $\delta > 0$ . According to the choice of the threshold  $\Lambda$  (see Step 0''), one obtains from (7.1)

$$P\{|\Lambda - \hat{\lambda}(\eta_0, \tilde{x}_j)| > \delta\} \rightarrow 0 \quad \text{as} \quad M \rightarrow \infty. \quad (7.2)$$

Now consider the case  $\tilde{x}_j \in \Gamma_0$ . Using (4.2) and (6.1), one obtains

$$P\{|\lambda(\{x_{i_k}\}_{k=2M-M_0+1}, \tilde{x}_j) - \hat{\lambda}(\gamma_0, \tilde{x}_j)| > \delta\} \rightarrow 0 \quad \text{as} \quad M \rightarrow \infty. \quad (7.3)$$

Since  $\gamma_0 < \eta_0$ , one proves Theorem 3 using (4.3), (7.2), (7.3) and the triangle inequality.

## 8. NUMERICAL EXPERIMENTS

Figure 1a represents a synthetic image of square and circle edges with the jump magnitude  $D = 1.5$  specified at a square  $101 \times 101$  grid with the step size  $h_0$  along each direction. The image is corrupted by noise with the uniform distribution on the interval  $[-\sqrt{3}/2, \sqrt{3}/2]$  (thus the variance is  $\sigma^2 = 0.25$ ). The radius of the window has been chosen  $R = 3.5h_0$ , the probability of false alarm has been  $\epsilon = 0.002$ . Figure 1b represents the image of detected edges of Fig. 1a.

To test the algorithm on a natural image we used the black and white photo of a woman's head, which has been scanned with the grey level range 0 – 255 (a value of 0 means black, a value of 255 means white). Figure 2a shows the  $248 \times 230$  image corrupted by discrete noise which takes values 0,  $\pm 1, \dots, \pm 10$  with equal probability. The image of detected edges of Fig. 2a is presented in Figure 2b. White pixels correspond to "no edge" area, grey pixels correspond to places where the edges have been detected. The intensity of grey level at each point is proportional to the estimated jump value at this point. The radius of the window has been also  $R = 3.5h_0$ , the probability of false alarm has been  $\epsilon = 0.001$ .

Figure 3a represents a synthetic image of thin line edges with the jump magnitude  $D = 1.5$  specified at a square  $101 \times 101$  grid. The line edges are two node thick.

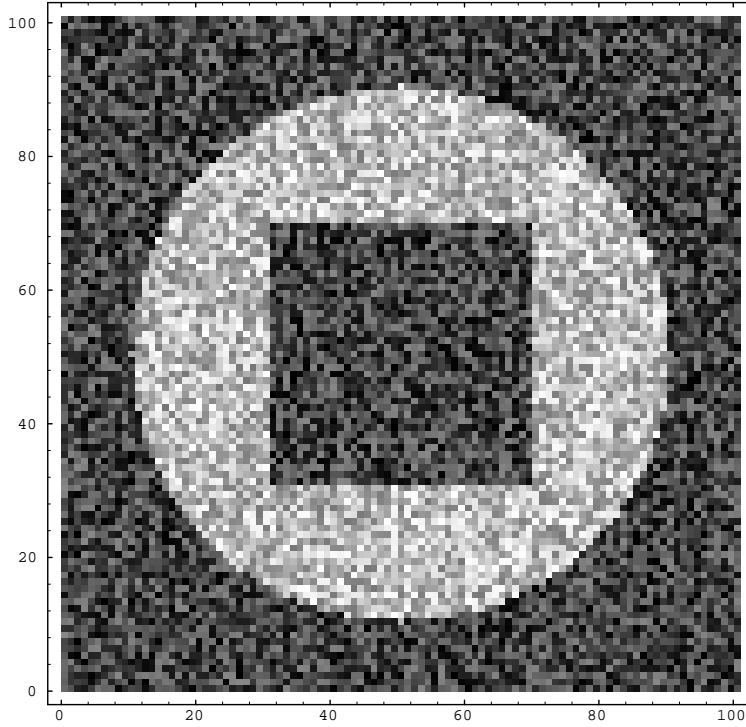


FIGURE 1A. A  $101 \times 101$  synthetic image of square and circle step edges corrupted by noise.

The image is corrupted by the noise with the same parameters as in the first model. The parameters  $R$ ,  $\epsilon$  also are the same as in the first model. Figure 3b represents the image of detected edges of Fig. 3a.

Note that to suppress spurious edge elements in the above examples we used the following general approach [11]: if a detected edge element does not have a single edge element in its neighborhood, it is likely to be a false one. In our case this means that if the decision " $\tilde{x}_j \in \Gamma$ " (or " $\tilde{x}_j \in \Gamma_0$ ") has been made, but there are no other edge points in the ball  $B_R(\tilde{x}_j)$ , we assume that  $\tilde{x}_j \notin \Gamma$  ( $\tilde{x}_j \notin \Gamma_0$ ).

In conclusion let us discuss the recommendations for the choice of  $R$  in numerical experiments. From the theory developed in Sections 5–7 it follows that  $R$  should not be too small, so that there would be enough grid points for the asymptotic statistical properties (Lemmas 1–3, Theorems 1–3) to be satisfied with sufficient accuracy. Thus, one can write  $R \geq R_{min}$ , where the particular value  $R_{min}$  depends on the allowable level of deviation from the limiting values of the statistics. The behavior of those statistics for small  $R$  (and, hence, for small  $M$ ) is difficult to study theoretically and such study requires extensive numerical experiments. The limitation on  $R$  from above  $R \leq R_{max}$  comes from the following assumptions: 1) the signal does not change much within balls with radius  $R$  not intersecting  $\Gamma$ , more precisely:  $\max_{B_R(\tilde{x}_j)} s(x) - \min_{B_R(\tilde{x}_j)} s(x) \ll \sigma$ ,  $B_R(\tilde{x}_j) \cap \Gamma = \emptyset$ , where  $\sigma$  is the standard deviation of noise, and 2) the discontinuity surface  $\Gamma$  can be accurately

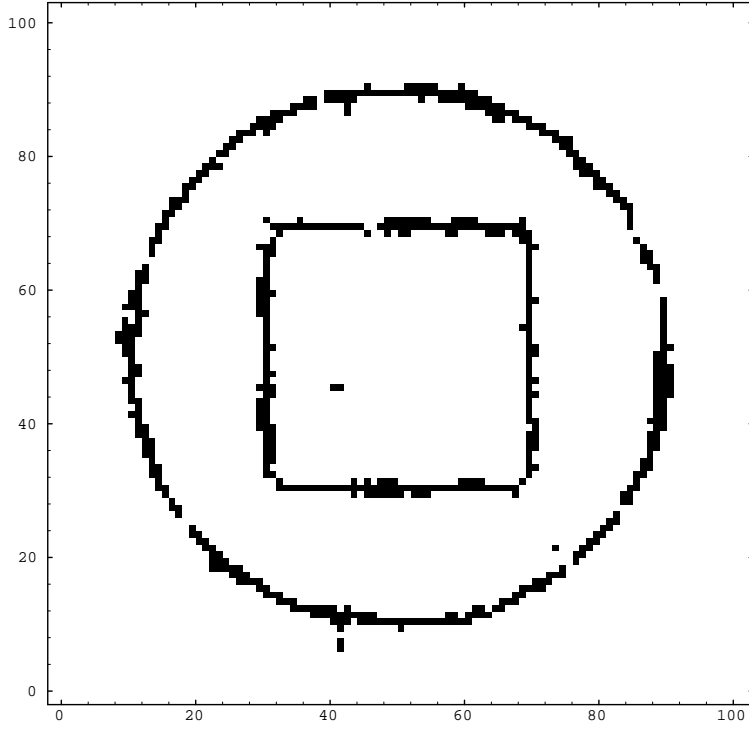


FIGURE 1B. Detected step edges of Fig. 1a.

approximated by a tangent plane within the balls  $B_R(\tilde{x}_j)$ ,  $\tilde{x}_j \in \Gamma$ . Thus, the upper bound  $R_{max}$  depends on the particular problem under consideration and should be specified by the researcher.

## APPENDIX

Let us denote  $\sqrt{M\gamma/(1-\gamma)} := t_M$ ,  $\sqrt{M(1-\gamma)/\gamma} := q_M$ ,  $\sqrt{M/(\gamma(1-\gamma))} := \tau_M$ . Then, using the inequalities  $0 \leq \Phi(t) \leq 1$  and

$$\int_a^\infty \exp(-x^2) dx \leq \int_a^\infty \exp(-ax) dx = a^{-1} \exp(-a^2), \quad a > 0, \quad (\text{A.1})$$

one gets

$$\left| \int_{t_M}^\infty \Phi \left( \tau_M \left\{ F[F^{-1}(1-\gamma+y\tau_M^{-1})-D] - \gamma \right\} \right) \exp(-y^2/2) dy \right| \leq \frac{2}{t_M} \exp\left(-\frac{t_M^2}{2}\right),$$

$$\left| \int_{-\infty}^{-q_M} \Phi \left( \tau_M \left\{ F[F^{-1}(1-\gamma+y\tau_M^{-1})-D] - \gamma \right\} \right) \exp(-y^2/2) dy \right| \leq \frac{2}{q_M} \exp\left(-\frac{q_M^2}{2}\right). \quad (\text{A.2})$$



FIGURE 2A. A natural  $248 \times 230$  image corrupted by noise.

Furthermore, using (A.1) again, one obtains with  $0 < t < \min(t_M, q_M)$ :

$$\int_{-q_M}^{t_M} \Phi\left(\tau_M \left\{ F[F^{-1}(1 - \gamma + y\tau_M^{-1}) - D] - \gamma \right\}\right) \exp(-y^2/2) dy =$$

$$\int_{-t}^t \Phi\left(\tau_M \left\{ F[F^{-1}(1 - \gamma + y\tau_M^{-1}) - D] - \gamma \right\}\right) \exp(-y^2/2) dy + O(t^{-1} \exp(-\frac{t^2}{2})). \quad (\text{A.3})$$

On the interval  $(-t, t)$  one has, with  $1 - \gamma < \xi_1 < 1 - \gamma + y\tau_M^{-1}$ ,  $\xi_0 < \xi_2 < \xi_0 + y\tau_M^{-1}/f(\xi_1)$  and  $\xi_0 = F^{-1}(1 - \gamma) - D$ ,

$$\Phi\left(\tau_M \left\{ F[F^{-1}(1 - \gamma + y\tau_M^{-1}) - D] - \gamma \right\}\right) = \Phi\left(\tau_M \left\{ F\left[\xi_0 + y\tau_M^{-1}/f(\xi_1)\right] - \gamma \right\}\right)$$

$$= \Phi\left(\tau_M \left\{ F(\xi_0) - \gamma + (y\tau_M^{-1}/f(\xi_1))f(\xi_2) \right\}\right). \quad (\text{A.4})$$

Thus, with  $b := F(\xi_0) - \gamma$  and  $\delta_M := (y\tau_M^{-1}/f(\xi_1))f(\xi_2)$ ,  $-t \leq y \leq t$ , (A.1) implies

$$\left| \Phi(\tau_M \{b + \delta_M\}) - \theta(b) \right| \leq 2\tau_M^{-1} \exp(-\tau_M^2/2). \quad (\text{A.5})$$



FIGURE 2B. Detected step edges of Fig. 2a.

Estimate (A.5) holds if  $t = t(M)$ , where  $t(M)\tau_M^{-1} \rightarrow 0$  as  $M \rightarrow \infty$ , since in this case  $\xi_0 + y\tau_M^{-1}/f(\xi_1) \rightarrow \xi_0$  and  $1 - \gamma + y\tau_M^{-1} \rightarrow 1 - \gamma$ . Thus

$$\begin{aligned} & \frac{1}{\sqrt{2\pi}} \int_{-t(M)}^{t(M)} \Phi\left(\tau_M \left\{ F[F^{-1}(1 - \gamma + y\tau_M^{-1}) - D] - \gamma \right\}\right) \exp(-y^2/2) dy = \\ & \frac{1}{\sqrt{2\pi}} \int_{-t(M)}^{t(M)} \left\{ \theta(F(\xi_0) - \gamma) + O(\tau_M^{-1} \exp(-\tau_M^2/2)) \right\} \exp(-y^2/2) dy = \\ & \theta(F(\xi_0) - \gamma) + O\{\tau_M^{-1} \exp(-\tau_M^2/2)\} + O\{\exp(-t^2(M)/2)/t(M)\}. \end{aligned} \quad (\text{A.6})$$

Select arbitrarily  $\alpha$ ,  $0 < \alpha < 1/2$ , and let  $t(M) = M^\alpha$ . Then

$$O\{\tau_M^{-1} \exp(-\tau_M^2/2)\} + O\{\exp(-t^2(M)/2)/t(M)\} = O\{M^{-\alpha} \exp(-M^{2\alpha}/2)\}. \quad (\text{A.7})$$

Also

$$\begin{aligned} & \left( \int_{-q_M}^{-t(M)} + \int_{t(M)}^{t_M} \right) \Phi\left(\tau_M \left\{ F[F^{-1}(1 - \gamma + y\tau_M^{-1}) - D] - \gamma \right\}\right) \exp(-y^2/2) dy \leq \\ & 2 \int_{t(M)}^{\max(t_M, q_M)} \exp(-y^2/2) dy = O\{\exp(-t^2(M)/2)/t(M)\} = \end{aligned}$$

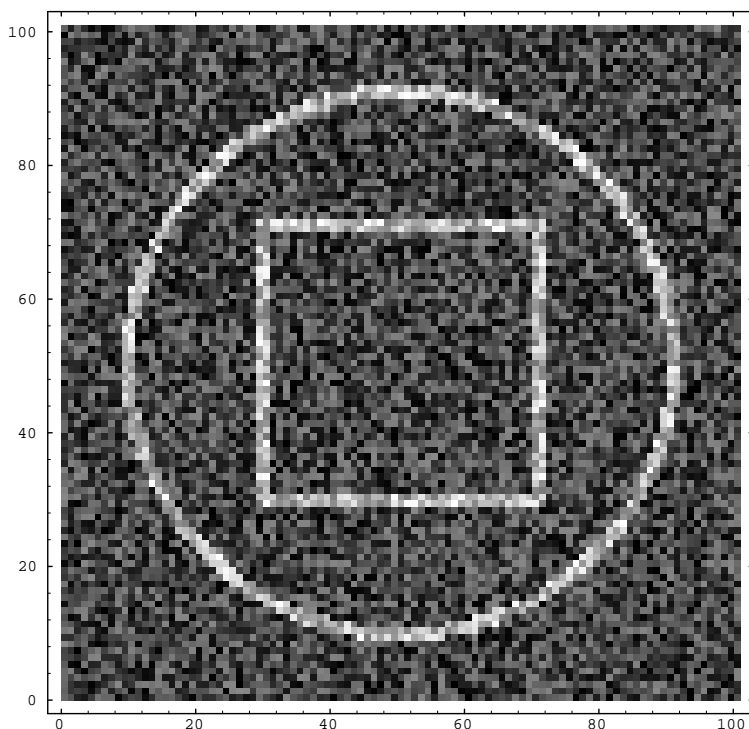


FIGURE 3A. A  $101 \times 101$  synthetic image of two node thick square and circle line edges corrupted by noise.

$$= O\{M^{-\alpha} \exp(-M^{2\alpha}/2)\}. \quad (\text{A.8})$$

Since  $\alpha$ ,  $0 < \alpha < 1/2$ , was arbitrary, one obtains from (A.6) – (A.8)

$$\left| \int_{-q_M}^{t_M} \Phi \left( \tau_M \left\{ F[F^{-1}(1 - \gamma + y\tau_M^{-1}) - D] - \gamma \right\} \right) \exp(-y^2/2) dy - \theta(F(\xi_0) - \gamma) \right| \\ = o\{M^{-\alpha} \exp(-M^{2\alpha}/2)\} \quad \text{for any } \alpha, 0 < \alpha < 1/2.$$

#### REFERENCES

1. L. S. Davis, *A survey of edge detection techniques*, Comput. Graphics Image Processing **4** (1975), 248–270.
2. G. Englund, *Remainder term estimates for the asymptotic normality of order statistics*, Scand. J. Stat. **7** (1980), 197–202.
3. J. Galambos, *Order statistics*, Handbook of Statistics, vol.4. Nonparametric methods (P. R. Krishnaiah and P.K. Sen, eds.), North-Holland, Amsterdam, 1984, pp. 359–382.
4. R. M. Haralick and J. Lee, *Context dependent edge detector*, Proc. of CVPR-88, 1988, pp. 223–228, Ann Arbor, Michigan.
5. M. H. Kalos and P. A. Whitlock, *Monte Carlo methods. Volume I: Basics*, Wiley, New York, 1986.

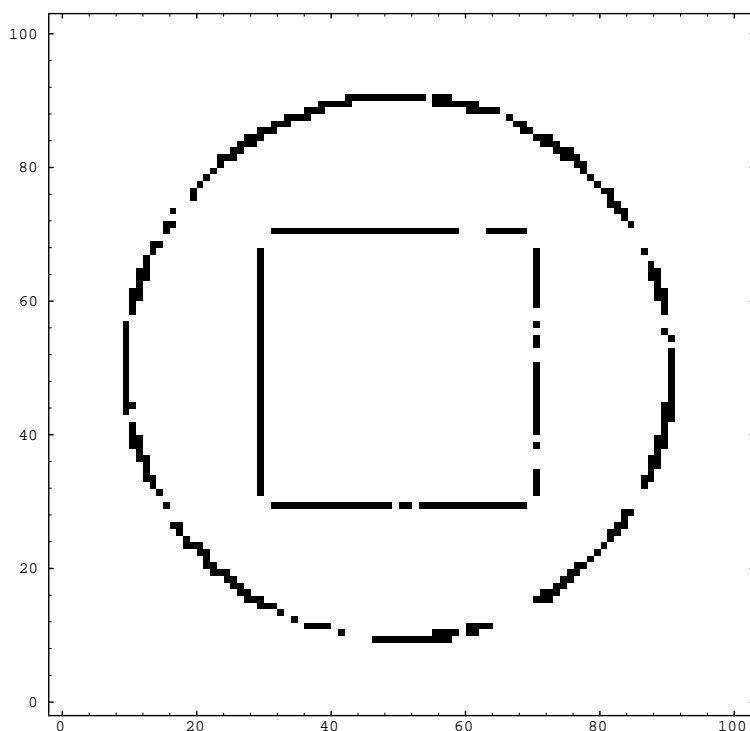


FIGURE 3B. Detected thin line edges of Fig. 3a.

6. A. Kundu, *Robust edge detection*, Proc. of IEEE Comp. Soc. Conf. on CVPR, IEEE Press, New York, 1989, pp. 11–18.
7. A. Kundu and S. K. Mitra, *Image edge extraction using a statistical classifier approach*, IEEE Trans. on PAMI **PAMI-9** (1987), 569–577.
8. J.-S. Lee, *Edge detection by partitioning*, Statistical image processing and graphics (E. J. Wegman and D. J. DePriest, eds.), Marcel Dekker, New York, 1986, pp. 59–69.
9. P. A. P. Moran, *An introduction to probability theory*, Clarendon Press, Oxford, 1968.
10. W. K. Pratt, *Digital Image Processing*, 2nd ed. Wiley, New York, 1991.
11. G. S. Robinson, *Detection and coding of edges using directional masks*, Proc. SPIE Conf. on Advances in Image Transmission Techniques, 1976, San Diego, August
12. A. Rosenfeld and A. C. Kak, *Digital Picture Processing*, Vol. 2, 2nd ed. Academic Press, New York, 1982.

MATHEMATICS DEPARTMENT, UNIVERSITY OF CENTRAL FLORIDA, ORLANDO, FL 32816-1364

*E-mail address:* akatsevi@pegasus.cc.ucf.edu

MATHEMATICS DEPARTMENT, KANSAS STATE UNIVERSITY, MANHATTAN, KS 66506-2602

*E-mail address:* ramm@math.ksu.edu