Proceedings
of the
1992 Conference
on
INFORMATION SCIENCES AND SYSTEMS

Volume I

Department of Electrical Engineering
Princeton University
Princeton, New Jersey 08544–5263

Optimal Mask Filtering of Discrete Random Sets under a Union/Intersection Noise Model *

N.D. Sidiropoulos[†] J.S. Baras[§] C.A. Berenstein[¶]

Systems Research Center
University of Maryland
College Park, MD 20742

Abstract

We consider the problem of estimating realizations of highly nonstationary Discrete Random Sets, distorted by a degradation process which can be described by a union/intersection model. We start by presenting some important structural results concerning the probabilistic description of Discrete Random Sets defined on a finite lattice. Then we propose a family of filters which can be viewed as lattice operators, and, for each degradation model, derive the optimal filter by means of minimizing a suitable fidelity criterion.

1 Introduction

An important problem in digital image processing and analysis is the development of optimal filtering procedures which attempt to restore a binary image ("signal") from its noisy version [14, 3, 5, 4]. Here, the noise process usually models the combined effect of two distinct types of degradation, namely, image object obscurations because of clutter, and sensor/channel noise. It is typically assumed that the degraded image can be accurately modeled as the union of the uncorrupted binary image with an independent realization of the noise process, which is a binary image itself [8]. This degradation model is known as the union noise model. Other models exist, such as the intersection noise model, and the union/intersection noise model, which are defined in the obvious fashion.

These models are well justified in practice, because, usually, binary images are obtained by thresholding gray-level images. If the threshold value is set sufficiently low, then the resulting degraded binary images will be well described by a union noise model. Alternatively, if the threshold is set sufficiently high, then the intersection noise model will be appropriate. In between these extreme choices, a union/intersection noise model will be most appropriate. The assumption of independence is crucial for the theoretical analysis of optimal filters, and it is plausible in many practical situations.

*Research partially supported by NSF grant NSFD CDR 8803012, through the Engineering Research Center s Program

[†]Also with the Department of Electrical Engineering

[§]Also with the Department of Electrical Engineering

[¶]Also with the Department of Mathematics

In this paper, we consider some classes of simple yet intuitive spatially varying filters, which we collectively refer to as *Mask Filters*, and, for each degradation model, derive an explicit characterization of the optimal filter, within the class of filters under consideration. These filters are appropriate when the images of the signal and/or the noise are highly nonstationary. A complementary study, dealing with optimal Morphological filters (which are appropriate when the images of signal and noise are stationary) will appear elsewhere.

2 Discrete Random Set Fundamentals

Definition 1 *Let B be a bounded subset of $\mathcal{Z}^2$. Assume that B contains the origin. Let $\Sigma(\Omega)$ denote the σ-algebra on Ω. Let $\Sigma(B)$ denote the power set (i.e. the set of all subsets) of B, and let $\Sigma(\Sigma(B))$ denote the power set of $\Sigma(B)$. A* Discrete Random Set (DRS), *X, on B, is a measurable mapping of a probability space $(\Omega, \Sigma(\Omega), P)$ into the measurable space $(\Sigma(B), \Sigma(\Sigma(B)))$. A DRS X, on B, induces a unique probability measure, P_X, on $\Sigma(\Sigma(B))$.*

Definition 2 *The functional*

$$T_X(K) = P_X(X \cap K \neq \emptyset)$$

is known as the capacity functional *of the DRS X.*

Definition 3 *The functional*

$$Q_X(K) = P_X(X \cap K = \emptyset) = 1 - T_X(K)$$

is known as the generating functional *of the DRS X.*

The following Lemma will be useful. Its proof can be found in the appendix. See [1] for basic Mobius inversion.

Lemma 1 *(Variant of Mobius inversion for Boolean algebras) Let v be a function on $\Sigma(B)$. Then v can be represented as*

$$v(A) = \sum_{S \subseteq A^c} u(S) \qquad \text{"external decomposition"}$$

The function u is uniquely determined by v, namely

$$u(S) = \sum_{C \subseteq S} (-1)^{|C|} v(S^c \cup C)$$

where ^{c} denotes complement with respect to B.

The capacity functional, T_X (or, equivalently, the generating functional, Q_X) carries all the information about a DRS X. This is clearly stated in the following theorem.

Theorem 1
Given $Q_X(K')$, $\forall K' \in \Sigma(B)$, $P_X(A)$, $\forall A \in \Sigma(\Sigma(B))$ is uniquely determined, and, in fact, can be recovered via the measure reconstruction formulas

$$P_X(A) = \sum_{K \in A} P_X(X = K)$$

with

$$P_X(X = K) = \sum_{K' \subseteq K} (-1)^{|K'|} Q_X(K^c \cup K')$$

Proof:
The reconstruction formula for the functional $P_X(X = K)$ in terms of the functional Q_X is a direct consequence of Lemma 1 and the fact that Q_X can be expressed in terms of P_X as

$$Q_X(K) = \sum_{K' \subseteq K^c} P_X(X = K')$$

□

The *uniqueness* part of this theorem is originally due to Choquet [2], and it has been independently introduced in the context of continuous-domain random set theory by Kendall [9] and Matheron [10, 11]. Related results can also be found in Ripley [13]. However, the measure reconstruction formulas are essentially only applicable within a discrete, bounded setting. In the continuous case, the uniqueness result relies heavily on Kolmogorov's extension theorem, which is non-constructive. See [16, 6, 7] for some other interesting results on DRS's.

The capacity functional plays an important role in the study of statistical inference problems for random sets. This is especially true for a class of random set models known as *germ-grain* models, and the Boolean model [15] in particular, whose capacity functional has a simple, tractable form. We will see that the capacity functional has an equally fundamental role in the study of optimal filtering of DRS's.

3 Formulation of the Optimal Filtering Problem

Let X, N, Y be DRS's on B. X models the "signal", whereas N models the noise. Let $g : \Sigma(B) \times \Sigma(B) \mapsto \Sigma(B)$ be a mapping that models the degradation (measurability is automatically satisfied here, since the domain of g is finite). The observed DRS is $Y = g(X, N)$. Let $d : \Sigma(B) \times \Sigma(B) \mapsto Z_+$ be a distance metric between subsets of B. In this context, the optimal filtering problem is to find a mapping $f : \Sigma(B) \mapsto \Sigma(B)$ such that the expected cost (expected error)

$$E(e) \triangleq Ed(X, \widehat{X}), \quad \widehat{X} = f(Y) = f(g(X, N))$$

is minimized, over all possible choices of the mapping ("filter") f. This problem is in general intractable. The main difficulty is the lack of structure on the search space. The family of all mappings $f : \Sigma(B) \mapsto \Sigma(B)$ is a chaotic search space! It is common practice to *impose* structure on the search space, i.e. constrain f to lie in $\mathcal{F}$, a suitably chosen subcollection of *admissible* mappings (family of filters), and optimize within this subcollection. The resulting filter is the best among its peers, but it is not guaranteed to be globally optimal.

We adopt the following distance metric (area of the symmetric set difference)

$$\begin{aligned} d(X, \widehat{X}) &= |(X \backslash \widehat{X}) \cup (\widehat{X} \backslash X)| \\ &= |(X \backslash \widehat{X})| + |(\widehat{X} \backslash X)| \\ &= |(X \cup \widehat{X}) \backslash (X \cap \widehat{X})| \\ &= |(X \cup \widehat{X})| - |(X \cap \widehat{X})| \end{aligned}$$

where $|\;|$ stands for set cardinality, $\backslash$ stands for set difference, i.e. $X \backslash Y = X \cap Y^c$, and ^{c} stands for complementation with respect to the base frame, B. This distance metric is essentially the *Hamming distance* [12] when $X, \widehat{X}$ are viewed as vectors in $\{0,1\}^{|B|}$. Since the component variables are binary, it can also be interpreted as the square of the L_2 distance of vectors in $\{0,1\}^{|B|}$, i.e., with some abuse of notation,

$$d(X, \widehat{X}) = (X - \widehat{X})^T (X - \widehat{X})$$

where on the left hand side symbols are interpreted as sets, while on the right hand side symbols are interpreted as column vectors in $\{0,1\}^{|B|}$, and ^{T} stands for transpose. In this setting, the *sufficiency part* of the Orthogonality Principle (OP) [12] applies. It states that a sufficient condition for the existence of a $f^* \in \mathcal{F}$ such that

$$E\left[(X - f^*(Y))^T (X - f^*(Y))\right] \leq$$

$$E\left[(X - f(Y))^T (X - f(Y))\right], \quad \forall f \in \mathcal{F}$$

is that

$$E\left[(X - f^*(Y))^T (f^*(Y) - f(Y))\right] = 0$$

for all $f \in \mathcal{F}$. However, unlike the case of vectors in R^n, where $\mathcal{F}$ is a vector space over the field of reals (known as the space of *square integrable* estimators), here $\mathcal{F}$ is not a vector space. The proof of the necessity part of the OP strongly depends on $\mathcal{F}$ having a vector space structure. For certain choices of $\mathcal{F}$ it is easy to show that the necessity part of the OP does not hold. At any rate, it is often easier to write down an expression for $Ed(X, X)$, and optimize over $\mathcal{F}$ by brute force.

This choice of distance has many advantages [14], not the least of which is that it enables the derivation of explicit optimality conditions. Even though, technically speaking, $d(X, \widehat{X})$ can be considered as a quadratic distance measure when we view $X, \widehat{X}$ as vectors in $\{0,1\}^{|B|}$, from a set-theoretic point of view $d(X, \widehat{X})$ is clearly not a quadratic distance measure, since it penalizes errors in a linear fashion. However, the squared area of the symmetric set difference (which is a quadratic distance measure in the set-theoretic sense) does not yield useful optimality conditions. This is partly due to the lack of a meaningful and tractable definition for the *expectation of a DRS X*. From an L_2 estimation-theoretic point of view, a proper *formal* definition of the expectation of a DRS X, would be as follows.

$$EX \triangleq arg\ min_{W \in \Sigma(B)} E\left[d(X, W)\right]^2$$

However, there exist several flaws with this formal definition. It can be shown that

$$arg\ min_{W \in \Sigma(B)} E\left[d(X, W)\right]^2 =$$

$$arg\ min_{W \in \Sigma(B)} \left\{ |W|^2 + 2|W| \left(\sum_{z \in W^c} Pr(z \in X) - \sum_{z \in W} Pr(z \in X) \right) - 4 \sum_{z \in W^c} \sum_{y \in W} Pr(z \in X\ ,\ y \in X) \right\}$$

If we assume that $Pr(z \in X) = p$, $\forall z \in B$, and $Pr(z \in X\ ,\ y \in X) = Pr(z \in X) Pr(y \in X) = p^2$, $\forall\ z, y\ s.t.\ z \neq y$, and $p < 0.5$, then it can be shown that $EX = \emptyset$, regardless of the specific value of p. If $p = 0.5$, then *any* $W \in \Sigma(B)$ will do. However, the single most important problem is that, given a specification of the first and second-order statistics of X, it is not clear how to find an explicit solution to the above minimization problem. On the other hand, the *median of a DRS X*, formally defined as

$$MX \triangleq arg\ min_{W \in \Sigma(B)} Ed(X, W)$$

is much easier to deal with. Although the solution to this latter minimization problem is not (in general) unique, it can be forced to be unique by means of a simple regularization. Let $C(z)$ be a Boolean proposition, which, for each point $z \in B$, is either true, or false. Define the *indicator function*

$$1(C(z)) \triangleq \begin{cases} 1 & , if\ C(z)\ is\ true\ at\ z \\ 0 & , otherwise \end{cases}$$

Let *supp* $1(C(z))$ stand for the *support set* of the indicator function $1(C(z))$, i.e. the subset of B where $C(z)$ is true. Then it can be shown that

$$M_R X \triangleq supp\ 1(1 - T_X(\{z\}) < T_X(\{z\}))$$

is the unique minimum cardinality solution to the minimization problem

$$min_{W \in \Sigma(B)} Ed(X, W)$$

These considerations essentially dictate our choice of distance metric. In terms of the degradation, we assume that N is independent of X, and that the mapping g is given by

$$g(X, N) = X \cup N \quad (union\ noise\ model)$$

or,

$$g(X, N) = X \cap N \quad (intersection\ noise\ model)$$

More generally, we can assume that g is a mapping from $\Sigma(B) \times \Sigma(B) \times \Sigma(B)$ to $\Sigma(B)$

$$g(X, N_1, N_2) = (X \cap N_1) \cup N_2$$

(*union/intersection noise model*), where X, N_1, N_2 are assumed to be mutually independent.

4 Optimal Mask Filters

In the case of union noise, a simple yet intuitive class of filters is

$$f(Y) = f_W(Y) = Y \cap W = (X \cup N) \cap W$$

for some $W \in \Sigma(B)$. Similarly, in the case of intersection noise, we can consider the following class of filters

$$f(Y) = f^W(Y) = Y \cup W = (X \cap N) \cup W$$

for some $W \in \Sigma(B)$. Finally, in the case of union/intersection noise, we can consider

$$f(Y) = f_{W_2}^{W_1}(Y) = (Y \cap W_2) \cup W_1 = ((((X \cap N_1) \cup N_2) \cap W_2) \cup W_1)$$

for some W_1, W_2, both in $\Sigma(B)$. We call these filters *mask filters*. They can be viewed as simple operators on the lattice of all subsets of B, i.e. $f_W(Y) = Y \cap W = glb(Y, W)$, $f^W(Y) = Y \cup W = lub(Y, W)$, and $f_{W_2}^{W_1}(Y) = (Y \cap W_2) \cup W_1 = lub(glb(Y, W_2), W_1)$, where glb, lub stand for greatest lower bound and least upper bound, respectively. In the fixed-window case the mask W is fixed; in the adaptive case W is allowed to depend on the observation. We will consider both cases. These relatively simple filters are appropriate when the signal and noise exhibit a highly nonstationary behavior. In this case, traditional shift-invariant neighborhood filtering operators fail to provide adequate filtering, and a simple but optimal spatially varying approach can produce better results. We will further discuss this point later on.

4.1 Optimal fixed-mask filtering

4.1.1 The case of union noise

As before, let $supp\ 1(C(z))$ stand for the subset of B where $C(z)$ is true, and $T_X(\cdot)$ denote the capacity functional of the DRS X. We have the following result.

Proposition 1 *Under the expected symmetric set difference measure, the optimal fixed intersection mask, W, for filtering out independent union noise is given by*

$$W = supp\ 1\,(T_N(\{z\})\,[1 - T_X(\{z\})] \leq T_X(\{z\}))$$

The corresponding minimum expected error achieved by such an optimal choice of W is given by

$$E(e^*) = \sum_{z \in B} min\,(T_X(\{z\}), T_N(\{z\})\,[1 - T_X(\{z\})])$$

The proof is rather straightforward[1]. It is based on the simple observation that $E|X|$ can be expressed as

$$E|X| = \sum_{z \in B} T_X(\{z\})$$

It is interesting to compare the above optimality condition with that of standard Wiener filtering. What it says is that if the effective union noise "power" at z is less than or equal to the signal "power" at z, then filtering should retain the input value at point z, otherwise it should reject it. This is intuitively appealing, and highly reminiscent of a form of binary Wiener filtering. Also notice that all that is required for the design of the optimal W is just the first-order statistics of the signal and the noise, i.e. $T_X(\{z\})$ and $T_N(\{z\})$ for all $z \in B$. These can be efficiently and accurately estimated from training samples of X and N respectively.

If the first-order statistics of the signal and the noise are spatially invariant, then, obviously, the optimal intersection mask is either B ("all pass"), or $\emptyset$ ("reject all"). This case is clearly not interesting. It is exactly when the signal and/or the noise statistics are highly nonstationary (meaning not even first-order stationary) that this filtering approach makes sense. Let us illustrate this point by using a (rather simplistic) artificial example. Consider figure (1). It depicts a realization of a DRS which features a prominent periodic vertical line structure. Figure (2) depicts a degraded version of the same image, obtained by taking the union of the DRS realization of figure (1) (the "signal") with an independent realization of another DRS (the "noise"). Figure (3) depicts the restored image, obtained by intersecting the DRS realization of figure (2), with the "optimal" intersection mask, computed by using the optimality condition of Proposition 1, and substituting *estimates* of the pixel hitting probabilities in place of the true probabilities. These estimates have been obtained by means of simple counting of pixel hitting events over a collection of learning samples of the signal and the noise. In this setting, shift-invariant filtering would not work well. Also, shape-sensitive filtering would not do, because the signal and the noise consist of replicas of the same elementary pattern, namely a square of side 5 pixels. A potentially big gain in quality of restoration rests exactly with proper exploitation of the nonstationary nature of the signal.

[1] Due to space constraints we skip the proofs of Propositions.

4.1.2 The case of intersection noise

This is the "dual" of the case of union noise. One can simply take the complement of all the sets and operations involved, and apply the results of the previous section. This is clear, because

$$d(X, \widehat{X}) = d(X^c, (\widehat{X})^c)$$

and, thus, minimizing $Ed(X, \widehat{X})$ is the same as minimizing $Ed(X^c, (\widehat{X})^c)$, and

$$((X \cap N) \cup W)^c = (X^c \cup N^c) \cap W^c$$

Therefore, employing Proposition 1, and noting that for singletons $\{z\}$

$$T_{X^c}(\{z\}) = 1 - T_X(\{z\})$$

we obtain the following result.

Proposition 2 *Under the expected symmetric set difference measure, the optimal fixed "fill" mask, W, for independent intersection noise is given by*

$$W = supp\ 1\,([1 - T_N(\{z\})]\,T_X(\{z\}) > 1 - T_X(\{z\}))$$

The corresponding minimum expected error achieved by such an optimal choice of W is given by

$$E(e^*) = \sum_{z \in B} min\,(1 - T_X(\{z\}), [1 - T_N(\{z\})]\,T_X(\{z\}))$$

Observe that, once more, the result is intuitively appealing.

4.1.3 The composite problem

Finally, let us consider the case of union/intersection noise. We have the following Proposition.

Proposition 3 *Under the expected symmetric set difference measure, the optimal fixed pair of masks, (W_1, W_2), is given by*

$$W_2 = supp\ 1\,(T_X(\{z\}) > max(T_1(\{z\}), T_2(\{z\})))$$

$$W_1 = supp\ 1\,(T_2(\{z\}) \leq min(T_X(\{z\}), T_1(\{z\})))$$

whereas, the associated minimum expected cost achieved by such an optimal pair of masks is

$$E(e^*) = \sum_{z \in B} min\,(T_X(\{z\}), T_1(\{z\}), T_2(\{z\}))$$

with

$$T_1(\{z\}) = T_X(\{z\})(1 - T_{N_1}(\{z\}))(1 - T_{N_2}(\{z\})) + (1 - T_X(\{z\}))T_{N_2}(\{z\})$$

and

$$T_2(\{z\}) = T_X(\{z\})(1 - T_{N_1}(\{z\})) + (1 - T_X(\{z\}))$$

4.2 Optimal adaptive mask filtering

A drawback of the optimal filters which have been derived up to this point is that they are non-adaptive; no matter what the observation is, the filter is fixed. One would like to improve upon these filters by allowing for an adaptation of the mask using information extracted from the given input. The trade-off is an increase in design/implementation complexity. For simplicity, we only consider union noise. By duality, similar results can be obtained for the case of intersection noise.

Assume that we are presented with a specific input, K, i.e. we are given that $Y = X \cup N = K$. One adaptation strategy is to incorporate this information into the cost function. This is done by considering the *conditional* expectation of $d(X, \widehat{X})$, conditioned on the given information. However, this does not lead to a closed-form solution for the optimal filter. Instead, we can condition on part of the available information. If we condition on the event $X \cup N \subseteq K$, i.e. $(X \cup N) \cap K^c = \emptyset$, then we can work out closed-form expressions for the optimal filter and the associated minimum error.

Proposition 4 *Given that $X \cup N \subseteq K$, the optimal intersection mask, W, for filtering out the noise component, N, is given by the intersection of K with the set*

$$supp\ 1\left([1 - T_X(K^c \cup \{z\})]\,[T_N(K^c \cup \{z\}) - T_N(K^c)] \leq [T_X(K^c \cup \{z\}) - T_X(K^c)]\,[1 - T_N(K^c)]\right)$$

The corresponding minimum cost achieved by such an optimal choice of W is given by

$$E(e^*) = \frac{1}{(1 - T_X(K^c))(1 - T_N(K^c))} \cdot \sum_{z \in K} min\,\{[1 - T_X(K^c \cup \{z\})]\,[T_N(K^c \cup \{z\}) - T_N(K^c)]\,,[T_X(K^c \cup \{z\}) - T_X(K^c)]\,[1 - T_N(K^c)]\}$$

Observe that for $K^c = \emptyset$ (i.e. $K = B$, no information available about the input) the formulas above reduce to the ones for the non-adaptive case, as they should (note that $T_Z(\emptyset) = 0$, for all DRS's Z).

5 Conclusions

Mask filtering is a simple yet intuitive approach to the problem of digital binary image restoration, under a union/intersection degradation model. We have discussed both optimal fixed-mask filtering, and optimal adaptive mask filtering. Although adaptive mask filtering is clearly superior when compared to fixed-mask filtering, it essentially requires knowledge of the capacity functionals of the signal and noise. This is the case, for example, when both the signal, X, and the noise, N, can be modeled as Discrete Boolean Random Sets [16]. On the other hand, fixed-mask filtering only requires knowledge of first-order statistics (pixel hitting probabilities), which can be easily and accurately estimated from training data. Therefore, it provides a simple and robust alternative, when the signal and noise processes are not known in detail. Generally speaking, mask filtering is suitable when the signal and noise DRS's are highly nonstationary, in which case optimal filters turn out being spatially varying, and traditional shift-invariant filters are very hard to optimize, and out of context.

6 Appendix - Proof of Lemma

Uniqueness: Assume that the external decomposition formula holds. Look at the right hand side of the inversion formula.

$$\sum_{C \subseteq S} (-1)^{|C|} v(S^c \cup C) = \sum_{C \subseteq S} (-1)^{|C|} \sum_{D \subseteq S \cap C^c} u(D) =$$

$$\sum_{C \subseteq S} (-1)^{|C|} \sum_{D \subseteq S \setminus C} u(D) = \sum_{C \subseteq S} \sum_{D \subseteq S \setminus C} (-1)^{|C|} u(D) =$$

$$\sum_{D \subseteq S} \sum_{C \subseteq S \setminus D} (-1)^{|C|} u(D) = \sum_{D \subseteq S} u(D) \sum_{C \subseteq S \setminus D} (-1)^{|C|} =$$

$$= u(S)$$

Since

$$\sum_{C \subseteq S} (-1)^{|C|} = \begin{cases} 0 & , S \neq \emptyset \\ 1 & , S = \emptyset \end{cases}$$

Existence: Assume that the inversion formula holds, and look at the right hand side of the external decomposition formula.

$$\sum_{S \subseteq A^c} u(S) = \sum_{S \subseteq A^c} \sum_{C \subseteq S} (-1)^{|C|} v(S^c \cup C) =$$

$$\sum_{S \subseteq A^c} \sum_{C \subseteq S} (-1)^{|C|} v((S \setminus C)^c) =$$

$$\sum_{D \subseteq A^c} \sum_{C \subseteq A^c \setminus D} (-1)^{|C|} v(D^c) =$$

$$\sum_{D \subseteq A^c} v(D^c) \sum_{C \subseteq A^c \setminus D} (-1)^{|C|} = v((A^c)^c) = v(A)$$

As for the uniqueness part. □

References

[1] M. Aigner. *Combinatorial Theory*. Springer-Verlag, New York, 1979.

[2] G. Choquet. Theory of capacities. *Annals of Institute Fourier*, 5:131–295, 1953.

[3] E.R. Dougherty, R.M. Haralick, Y. Chen, B. Li, C. Agerskov, U. Jacobi, and P. H. Sloth. Morphological pattern-spectra-based tau-opening optimization. In *Proc. SPIE Vol. 1606, Boston, Massachusetts.* Society for Optical Engineering, November 1991.

[4] E.R. Dougherty and R.P. Loce. Constrained optimal digital morphological filters. In *Proc. of the Twenty-Fifth Annual Conference on Information Sciences and Systems, The Johns Hopkins University, Baltimore, Maryland*, March 1991.

[5] E.R. Dougherty, A. Mathew, and V. Swarnakar. A conditional-expectation-based implementation of the optimal mean-square binary morphological filter. In *Proc. SPIE Vol. 1451, San Jose, California.* Society for Optical Engineering, February 1991.

[6] John Goutsias and Chuanju Wen. Modeling Discrete Random Shapes: A Random Set Theory Approach. Technical Report JHU/ECE 90-13, The Johns Hopkins University, 1990. Submitted for publication: IEEE Trans. PAMI.

[7] John Goutsias and Chuanju Wen. Discrete random set models for shape synthesis and analysis. In *Proc. SPIE Vol. 1606 Visual Communications and Image Processing '91: Image Processing*, pages 174–185. Society for Optical Engineering, 1991.

[8] R.M. Haralick, E.R. Dougherty, and P.L. Katz. Model-based morphology. In *Proc. SPIE Vol. 1472, Orlando, Florida.* Society for Optical Engineering, April 1991.

[9] D.G. Kendall. Foundations of a theory of random sets. In E.F. Harding and D.G. Kendall, editors, *Stochastic Geometry*, pages 322–376. John Wiley, London, England, 1974.

[10] G. Matheron. *Elements pour une theorie des Milieux Poreux.* Masson, 1967.

[11] G. Matheron. *Random Sets and Integral Geometry.* Wiley, New York, 1975.

[12] H. V. Poor. *An introduction to Signal Detection and Estimation.* Springer-Verlag, New York, 1988.

[13] B.D. Ripley. Locally finite random sets: foundations for point process theory. *Ann. Probab.*, 4:983–994, 1976.

[14] D. Schonfeld and J. Goutsias. Optimal morphological pattern restoration from noisy binary images. *IEEE Transactions on Pattern Anal. Mach. Intell.*, 13(1):14–29, Jan. 1991.

[15] J. Serra. The Boolean Model and Random Sets. *Computer Graphics and Image Processing*, 12:99–126, 1980.

[16] N.D. Sidiropoulos, J. Baras, and C. Berenstein. The structure of divisible discrete random sets and their randomized superpositions. Technical Report TR 91-54, Systems Research Center, University of Maryland, April 1991.

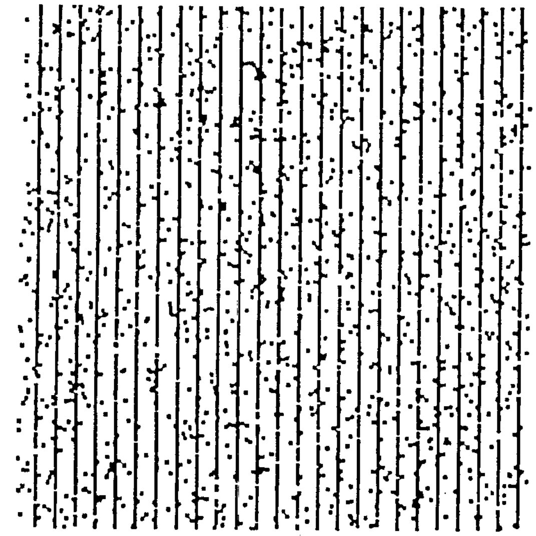

Figure 1: A realization of a non-stationary DRS, X.

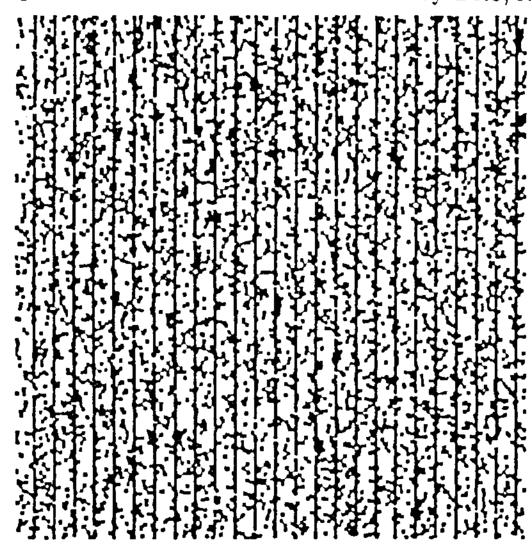

Figure 2: The result of taking the union of the DRS realization of figure (1) with an independent realization of the noise DRS, N.

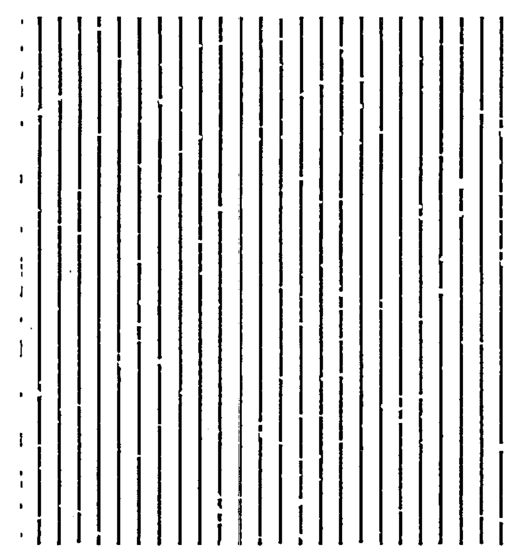

Figure 3: Restored image