PROCEEDINGS OF THE

31st IEEE CONFERENCE ON DECISION AND CONTROL

DECEMBER 16-18, 1992
WESTIN LA PALOMA
TUCSON, ARIZONA, USA

IEEE
Control
Systems
Society

VOLUME 4 OF 4

92CH3229-2

Proceedings of the 31st Conference
on Decision and Control
Tucson, Arizona • December 1992

Optimal Filtering of Digital Binary Images Corrupted by Union/Intersection Noise *†

N.D. Sidiropoulos‡ J.S. Baras§ C.A. Berenstein¶

Systems Research Center
University of Maryland
College Park, MD 20742
nikos@src.umd.edu

Abstract

We model digital binary image data as realizations of a uniformly bounded discrete random set (or discrete random set, for short), a mathematical object which can be directly defined on a finite lattice. We consider the problem of estimating realizations of discrete random sets distorted by a degradation process which can be described by a union/intersection noise model. We start by providing some theoretical justification of the popularity of certain Morphological filters, namely Morphological openings, closings, unions of openings, and intersections of closings. In particular, we prove that if the signal is "smooth", then these filters are optimal (in the sense of providing the MAP estimate of the signal) under reasonable worst-case statistical scenarios. Then we consider a class of filters which arises quite naturally from the set-theoretic analysis of optimal filters. We call this the class of *mask filters*. We consider both fixed and adaptive mask filters, and derive explicit formulas for the optimal mask filter under quite general assumptions on the signal and the degradation process.

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 degraded version [17, 5]. Here, the degradation mechanism usually models the combined effect of two distinct types of distortion, 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 [10]. 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. The assumption of independence is crucial for the theoretical analysis of optimal filters, and it is plausible in many practical situations. These models are rather general, in that they can be tailored to describe most popular types of signal-independent noise, e.g. salt-and-pepper noise (also known as Binary Symmetric Channel, (BSC) transmission noise), burst channel errors, noise with a geometric structure [10], occlusion, etc.

This research has been largely motivated by the works of Haralick, Dougherty, and Katz [10], and Schonfeld and Goutsias [17]. Their approach is model-based, in that they assume specific probabilistic/geometrical models that govern the behavior of both signal and noise "patterns", i.e. the elementary geometrical primitives from which the signal and noise images are constructed. This work focuses on a different viewpoint. As it turns out, by restricting our attention to suitable classes of filtering operations, and uniformly bounded discrete random sets (defined below), we can obtain optimal filtering results, *under considerably milder assumptions on the signal and noise patterns*, i.e. results that are applicable for all signal and noise models, under the assumption of mutual independence of the signal and the noise. Specifically, one need not assume that signal and noise patterns are "non-interfering". Furthermore, it is possible to obtain simple, closed characterizations of the optimal filter. The resulting formulas are intuitively appealing, and directly amenable to design and implementation.

2 Discrete Random Set Fundamentals

Definition 1 *Let B be a bounded subset of $\mathbb{Z}^2$. 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 Uniformly Bounded Discrete Random Set, or, for brevity,* **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*

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

‡Also with the Department of Electrical Engineering

§Martin-Marietta Chair in Systems Engineering, also with the Department of Electrical Engineering

¶Also with the Department of Mathematics

$$T_X(K) = P_X(X \cap K \neq \emptyset), \; K \in \Sigma(B)$$

is called the **capacity functional** *of the DRS X.*

Definition 3 *The functional*

$$Q_X(K) = P_X(X \cap K = \emptyset) = 1 - T_X(K), \; K \in \Sigma(B)$$

is called the **generating functional** *of the DRS X.*

In the context of DRS's, the generating functional plays a role analogous to the one played by the cumulative distribution function (cdf) in the context of scalar discrete random variables. This is the subject of 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

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

Remark: For a proof of this, and all remaining results, see [19] or [21, 20, 22].

The *uniqueness* part of this theorem is originally due to Choquet [1], and it has been independently introduced in the context of continuous-domain random set theory by Kendall [11] and Matheron [13, 14]. Related results can also be found in Ripley [16]. However, the measure reconstruction formulas are essentially only applicable within a uniformly bounded discrete random set setting. In the case of (uncountably or countably) infinite observation sites, the uniqueness result relies heavily on Kolmogorov's extension theorem, which is non-constructive.

3 Some results on constrained optimality, or, why Morphology is popular

The theory of *Mathematical Morphology* has been developed mainly by Serra [18, 8], Matheron [14], and their collaborators, during the 70's and early 80's. Since then, Mathematical Morphology and its applications have become very popular. The theory is concerned with the quantitative analysis of shape with an emphasis on geometric structure. It is founded on certain elementary set-to-set mappings, namely set dilation/erosion, which are inherently non-linear. These mappings are defined in terms of a *structuring element*, a "small" primitive shape (set of points) which interacts with the input image to transform it, and, in the process, extract useful information about its geometrical and topological structure. Let

$$W^s \triangleq \left\{ z \in \mathbf{Z}^2 \mid -z \in W \right\}$$

The dilation of a set $Y \subset \mathbf{Z}^2$ by a structuring element W is defined as[1]

$$Y \oplus W^s = \left\{ z \in \mathbf{Z}^2 \mid W_z \cap Y \neq \emptyset \right\}$$

whereas the erosion of a set $Y \subset \mathbf{Z}^2$ by a structuring element W is defined as

$$Y \ominus W^s = \left\{ z \in \mathbf{Z}^2 \mid W_z \subseteq Y \right\}$$

Erosion and dilation are *dual* operators, in the sense that $Y \ominus W^s = (Y^c \oplus W^s)^c$, where here ^{c} stands for complementation with respect to $\mathbf{Z}^2$. Two fundamental composite Morphological operators are opening and closing. The opening, $Y \circ W$, of a set $Y \subset \mathbf{Z}^2$ by a structuring element W, is defined as

$$Y \circ W \triangleq (Y \ominus W^s) \oplus W = \bigcup_{z \in \mathbf{Z}^2 \mid W_z \subseteq Y} W_z$$

Similarly, the closing, $Y \bullet W$, of a set $Y \subset \mathbf{Z}^2$ by a structuring element W, is defined as

$$Y \bullet W \triangleq (Y \oplus W^s) \ominus W$$

By duality of erosion/dilation it follows that opening and closing are dual operators. Both can be viewed as nonlinear smoothing operators. Opening and closing are *idempotent (stable)* operators in the sense that $(Y \circ W) \circ W = Y \circ W$, and $(Y \bullet W) \bullet W = Y \bullet W$. A set Y is said to be (Morphologically) *open* (*closed*) with respect to the structuring element W iff $Y \circ W = Y$ ($Y \bullet W = Y$). We shall say that a set Y is *smooth with respect to* W iff Y can be expressed as a union of shifted replicas of W. Y is open with respect to W, iff Y is smooth with respect to W. Y is closed with respect to W iff Y^c is smooth with respect to W.

Morphological filters are very flexible, mainly because of the freedom to choose the structuring element(s), to meet specified criteria. Among other things, Morphological filters have been widely used to filter out certain kinds of impulsive noise, such as the so-called salt-and-pepper noise, in both binary and gray scale images [17, 4, 7, 6, 2, 3, 23]. For example, it is widely believed that opening is suitable under a union noise model, while closing is suitable under an intersection noise model. Indeed, these filters are used extensively, and they deliver adequate filtering in a variety of noisy environments. The natural question, then, is whether we can provide some sort of theoretical justification for their use. As it turns out, these filters are indeed optimal under a reasonable worst-case scenario. In particular, if we assume that the signal, X, is sufficiently smooth, and the noise is i.i.d., then these filters provide the Maximum A Posteriori (MAP) estimate of X, on the basis of the observation Y. We have the following results.

Theorem 2 *Let $O_W(B)$ denote the collection of all W-open subsets of B. Assume that the signal DRS, X,*

[1] Here we follow the original definitions of Serra [18]. In his work the symbol $\oplus$ stands for Minkowski set addition, and the symbol $\ominus$ stands for Minkowski set subtraction.

on B, induces the following probability mass function on $\Sigma(B)$:

$$P_X(X = K) = \begin{cases} \frac{1}{|O_W(B)|} & if\ K \in O_W(B) \\ 0 & otherwise \end{cases}$$

where $|\ |$ stands for set cardinality. Furthermore, assume that the observable DRS is $Y = X \cup N$, where N is a homogeneous Bernoulli lattice process of intensity $r \in [0,1)$ (i.e. each point $z \in B$ is included in N with probability r, independently of all other points), which is independent of X. Then $Y \circ W$ is the unique MAP estimate of X on the basis of Y, regardless of the specific value of r.

The proof (see [19, 22]) *crucially* depends on $|B|$ being finite. Indeed, theorem 2, as well as the three theorems that follow, do not make sense when the lattice extends to infinity. Thus, a uniformly bounded discrete random set approach offers a fresh statistical perspective of Morphological filtering, one which is not apparent within other formulations. The suppositions of the theorem indeed correspond to a worst-case statistical scenario: if all that is known about the signal is that it is almost surely (a.s.) smooth (open) with respect to W, then it is reasonable to model this knowledge using a uniform distribution over the set of all W-open subsets of B, to reflect the fact that the signal exhibits no other (known) probabilistic structure. Also, i.i.d. noise is the worst kind of noise, in the sense of maximizing the Shannon entropy of the noise DRS N. Both these suppositions are plausible in practice, and this explains why the opening filter is successful under a union noise model. It is worth noting that the MAP estimate does not depend on the noise level, r. The following theorem is a straightforward extension of the above theorem.

Theorem 3 *Let $O_{W_1,\cdots,W_M}(B)$ denote the collection of all subsets K of B which can be written as*

$$K = \bigcup_{i=1,\cdots,M} K_i, \quad K_i \in O_{W_i}(B),\ i = 1,\cdots,M$$

Assume that the signal DRS, X, on B, induces the following probability mass function on $\Sigma(B)$:

$$P_X(X = K) = \begin{cases} \frac{1}{|O_{W_1,\cdots,W_M}(B)|} & if\ K \in O_{W_1,\cdots,W_M}(B) \\ 0 & otherwise \end{cases}$$

Furthermore, assume that the observable DRS is $Y = X \cup N$, where N is a homogeneous Bernoulli lattice process of intensity $r \in [0,1)$, which is independent of X. Then

$$\widehat{X}_{MAP}(Y) = \bigcup_{i=1,\cdots,M} Y \circ W_i$$

By duality with respect to complementation, we can obtain the following result.

Theorem 4 *Let $C_{W_1,\cdots,W_M}(B)$ denote the collection of all subsets K of B which can be written as*

$$K = \bigcap_{i=1,\cdots,M} K_i, \quad K_i \in C_{W_i}(B),\ i = 1,\cdots,M$$

Assume that the signal DRS, X, on B, induces the following probability mass function on $\Sigma(B)$:

$$P_X(X = K) = \begin{cases} \frac{1}{|C_{W_1,\cdots,W_M}(B)|} & if\ K \in C_{W_1,\cdots,W_M}(B) \\ 0 & otherwise \end{cases}$$

Furthermore, assume that the observable DRS is $Y = X \cap N$, where N is a homogeneous Bernoulli lattice process of intensity $r \in [0,1)$, which is independent of X. Then

$$\widehat{X}_{MAP}(Y) = \bigcap_{i=1,\cdots,M} Y \bullet W_i$$

A natural question which arises is what happens if we loose the uniform probability structure over the collection of "smooth" realizations. The answer is that the MAP estimate will typically be intractable. However, we can still claim that the proposed estimate in any of the above theorems is the Maximum Likelihood (ML) estimate of X on the basis of Y [19].

In general, if we assume that X satisfies some arbitrary (not necessarily Morphological) smoothness conditions, i.e. *ome* $\in \mathcal{S}$, a class of smooth subsets of B, and that X is uniformly distributed over $\mathcal{S}$, then under an i.i.d. symmetric (Binary Symmetric Channel, BSC) noise model of pixel inversion probability $r < 0.5$, it is easy to see that

$$\widehat{X}_{MAP}(Y) = arg\ min_{K \in \mathcal{S}} d(Y, K)$$

where $d(Y,K)$ is the area of the symmetric set difference distance between Y and K (cf. section 4). In other words, $\widehat{X}_{MAP}(Y)$ is the "projection" of the data Y onto $\mathcal{S}$. However, it is not clear how to compute this projection under general smoothness conditions. Furthermore, quite often the noise is not i.i.d., and the signal is nonsmooth, or only approximately smooth. The lack of a rigorous DRS-theoretic optimization approach for this general case has been evident in the literature. Our programme is to develop such an approach. For this, we need to abandon the MAP rule, and define optimality via the minimization of an appropriate cost function.

4 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 $d : \Sigma(B) \times \Sigma(B) \mapsto \mathbf{Z}_+$ 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 just too big. 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* [15] 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 terms of the degradation, we assume that N is independent of X, and that the mapping g is given by

$$g(X, N) = X \cup N$$

(*union noise model*)

or,

$$g(X, N) = X \cap N$$

(*intersection noise model*)

Although we shall be mainly concerned with either union *or* intersection noise, on one occasion we will allow g to be a mapping from $\Sigma(B) \times \Sigma(B) \times \Sigma(B)$ to $\Sigma(B)$

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

(*combined union/intersection noise model*)

where X, N_1, N_2 will be assumed to be mutually independent.

5 Optimal Mask Filters

In the case of union noise, we can assume, without loss of generality, that the optimal filter is of the form:

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

for some $W \in \Sigma(B)$. Similarly, in the case of intersection noise, we can assume that the optimal filter is of the form:

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

for some $W \in \Sigma(B)$. Finally, in the case of combined union/intersection noise, we can assume that the optimal filter is of the form:

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

$$((((X \cap N_1) \cup N_2) \cap W_2) \cup W_1),$$

for some W_1, W_2, both in $\Sigma(B)$. We will collectively refer to these filters as *mask filters*. For example, in the case of union noise, the optimal filter should retain a subset of the observation points and reject the rest; this should be done in some sort of statistically optimal fashion. This is achieved by intersecting the observation with a suitably chosen "mask", W, which, in general, depends on the observation.

As a first step, we might want to investigate how much we can achieve using a fixed mask W, i.e. one which does not depend on the observation, and optimize the choice of this fixed mask over all possible observations. We will call the resulting constrained optimal filter the optimal *fixed mask filter*. The second step would be to allow W to depend on the observation, via some suitable adaptation strategy. The ideal situation would be to optimize the mask pointwise, i.e. construct a mapping $W(\cdot) : \Sigma(B) \mapsto \Sigma(B)$, which takes an observation and maps it to the best mask for the given observation. However, it seems that, in general, this optimization is intractable. Furthermore, the implementation of such a pointwise optimal strategy requires a realization of the mapping $W(\cdot)$, which seems impractical. Nevertheless, we will show that explicit optimization is possible under some restrictions on the adaptation strategy. We will call the resulting constrained optimal filter the optimal *adaptive mask filter*.

Let us first consider fixed mask filtering. Here, we only work with the combined union/intersection noise model. The other two noise models are special cases. We have the following proposition.

Proposition 1 *Under the expected symmetric set difference measure, an optimal fixed pair of masks, (W_1, W_2), is given by*[2]

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

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

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

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

[2] Here, $supp\ 1(BE)$ stands for the *support* set of the indicator function, i.e., the set of points at which the Boolean expression BE is true.

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

and

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

An obvious drawback of fixed mask filtering is that it does not exploit the autocorrelation structure of the signal and the noise. Furthermore, it is non-adaptive. Whenever higher-order statistics are available, we would like to use them. We would also like to allow for an adaptation of the mask using information extracted from the given input. Adaptive mask filtering fits both bills. The trade-off is an increase in design/implementation complexity.

Let us first consider the case of union 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. The reason is that the minimization of this conditional expectation requires the explicit computation of a pseudo-convolution of likelihoods on the lattice of realizations. This computation is in general intractable. Instead, we can condition on part of the available information. This corresponds to minimizing the expected error over a wider collection of events than what is necessary (and optimal). The trade-off is in terms of tractability. 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. In what follows E denotes conditional expectation, conditioned on $(X \cup N) \cap K^c = \emptyset$.

Proposition 2 *Given that $X \cup N \subseteq K$, an optimal intersection mask, W, for filtering out the noise component, N, is given by the intersection of K with the set*

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

The corresponding minimum cost achieved by such an optimal choice of W is given by[3]

$$E(e^*) = \frac{1}{(1 - T_X(K^c))(1 - T_N(K^c))} \times$$

$$\sum_{z \in K} min\{[1 - T_X(K^c \cup \{z\})][T_N(K^c \cup \{z\}) - T_N(K^c)], [T_X(K^c \cup \{z\}) - T_X(K^c)][1 - T_N(K^c)]\}$$

Observe how information about the higher-order statistics of the signal and the noise is incorporated into the filter structure, by means of the capacity functionals of the signal and the noise. Note that the minimum cost achieved by an optimal choice of W is not necessarily increasing in K; in the expression for the minimum cost, we can show that the sum is increasing in K, but the normalizing factor, $1/((1 - T_X(K^c))(1 - T_N(K^c)))$, is decreasing in K. Thus, the tightest possible K (i.e. the observation itself) is not necessarily the best choice. However, we have experimented with this particular choice with satisfactory results. For example, when the signal and the noise can both be modeled as simple Boolean DRS's [19, 20], then, under what essentially amounts to a high S/N ratio condition, the optimal adaptive mask filter can be shown to reduce to a simple Morphological opening filter.

The case of intersection noise can be addressed by appealing to duality. One can simply take the complement of all the sets and operations involved, and apply the results which have been obtained for the case of union noise [19, 21].

[3] We assume that $Pr(X \cup N \subseteq K) > 0$. Note that this, in turn, implies that $T_X(K^c) < 1$, and $T_N(K^c) < 1$.

6 Conclusions

We have employed a DRS-theoretic approach to the problem of digital binary image restoration under a union or intersection noise model. This has allowed us to prove that certain popular Morphological filters are indeed optimal (in the MAP sense), under reasonable worst-case statistical scenarios. Mask filters arise quite naturally from the set-theoretic analysis of optimal filters. We have derived explicit formulas for the optimal mask filter, under quite general assumptions on the signal and the degradation process.

References

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

[2] E. Dougherty. Optimal Mean Square N-Observation Digital Morphological Filters - I. Optimal Binary Filters. *Computer Vision, Graphics, and Image Processing: Image Understanding*, 55(1):36–54, January 1992.

[3] E. Dougherty. Optimal Mean Square N-Observation Digital Morphological Filters - II. Optimal Gray - Scale Filters. *Computer Vision, Graphics, and Image Processing: Image Understanding*, 55(1):55–72, January 1992.

[4] E. Dougherty and C. Giardina. *Morphological Methods in Image and Signal Processing.* Prentice-Hall, Englewood Cliff, 1988.

[5] 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.

[6] E.R. Dougherty and R.P. Loce. Constrained optimal digital morphological filters. In *Proc. of the 25th Annual Conference on Information Sciences and Systems, The Johns Hopkins University, Baltimore, Maryland*, March 1991.

[7] 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.

[8] J. Serra Ed. *Image Analysis and Mathematical Morphology, vol. 2, Theoretical Advances.* Academic, San Diego, 1988.

[9] J. Goutsias and C. Wen. Modeling Discrete Random Shapes: A Random Set Theory Approach. Technical Report JHU/ECE 90-13, The Johns Hopkins University, 1990.

[10] 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.

[11] 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.

[12] T.Y. Kong and A. Rosenfeld. Digital topology: Introduction and survey. *Computer Vision, Graphics and Image Processing*, 48:357–393, 1989.

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

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

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

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

[17] 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.

[18] J. Serra. *Image Analysis and Mathematical Morphology.* Academic, New York, 1982.

[19] N.D. Sidiropoulos. *Statistical Inference, Filtering, and Modeling of Discrete Random Sets.* PhD thesis, University of Maryland, June 1992.

[20] N.D. Sidiropoulos, J.S. Baras, and C.A. Berenstein. Discrete Random Sets: an Inverse Problem, plus tools for the Statistical Inference of the Discrete Boolean model. In P.D. Gader, E.R. Dougherty, and J. Serra, editors, *Proc. of Conference on Image Algebra and Morphological Image Processing III, SPIE vol. 1769, San Diego, CA.* Society for Optical Engineering, July 1992.

[21] N.D. Sidiropoulos, J.S. Baras, and C.A. Berenstein. Optimal Mask Filtering of Discrete Random Sets under a Union / Intersection Noise Model. In *Proc. 26th Annual Conference on Information Sciences and Systems*, Princeton University, Princeton, N.J., March 1992.

[22] N.D. Sidiropoulos, J.S. Baras, and C.A. Berenstein. Optimal Morphological Filters for Discrete Random Sets under a Union or Intersection Noise Model. In *Proc. of Conference on Visual Communications and Image Processing, SPIE vol. 1818, Boston, Mass.* Society for Optical Engineering, November 1992.

[23] J. Song and E.J. Delp. A Study of the Generalized Morphological Filter. *Circuits, Systems and Signal Processing*, 11(1):229–252, January 1992.

[24] D. Stoyan, W.S. Kendall, and J. Mecke. *Stochastic Geometry and its Applications.* Wiley, Berlin, 1987.