PERGAMON

Pattern Recognition 32 (1999) 1667–1674

PATTERN RECOGNITION
THE JOURNAL OF THE PATTERN RECOGNITION SOCIETY

On the tractability of estimating the germ process of certain germ–grain random set models and related problems

N.D. Sidiropoulos*

Department of Electrical Engineering, University of Virginia, Charlottesville VA 22903, USA

Received 1 October 1997; received in revised form 10 April 1998; accepted 14 May 1998

Abstract

The germ process embedded in a germ–grain random set is often not merely part of a constructive mathematical model, but has physical significance as well, e.g., in the modeling of cell cultures. In such situations, knowledge of the realization of the germ process associated with a particular random set observation is highly desirable. In this paper, it is shown that, for an interesting class of germ–grain models, maximum likelihood estimation of the germ process on the basis of a random set observation is an NP-complete computational problem. Certain additional computational problems associated with the probabilistic specification of morphologically processed random sets are also discussed. © 1999 Pattern Recognition Society. Published by Elsevier Science Ltd. All rights reserved.

Keywords: Random sets; Germ–grain models; Boolean model; NP-completeness; Mathematical Morphology

1. Introduction

A boolean random closed set (RACS) is constructively defined as follows:

$$X = \bigcup_{i=1,\,2,\ldots} G_{p_i} \oplus \{p_i\},$$

where $P = \{p_1, p_2, \ldots\}$ is a stationary Poisson point process (the points p_i, $i = 1, 2, \ldots$, are called the *germs*, and P is called the *germ process*), $\{G_{p1}, G_{p2}, \ldots\}$ is an i.i.d. sequence of non-empty bounded random closed sets (the *grains*), and $\oplus$ stands for Minkowski set addition. Here, since one of its arguments is a singleton, $\oplus$ simply amounts to a translation of the associated grain. Informally (cf. [1,2]) for a rigorous exposition), one may think of P as a random collection of points scattered over $\mathbb{R}^n$, with the following properties:

(i) *Poisson distribution of point counts:* The number of points in B, a bounded (Borel) subset of $\mathbb{R}^n$, has a Poisson distribution:

$$P(\Phi(B) = m) = \frac{(\lambda|B|)^m e^{-\lambda|B|}}{m!}, \quad m = 0, 1, 2, \ldots,$$

where $\Phi(B)$ is the number of points in B, $|B|$ is the volume of B, and $\lambda > 0$ is a constant parameter known as the *intensity* of the stationary Poisson point process.

(ii) *Independent scattering:* The point counts in k disjoint (Borel) sets constitute k independent random variables.

The grains $\{G_{p_1}, G_{p_2}, \ldots\}$ may be informally thought of as simpler random shapes, e.g., disks of random radius centered at the origin, in which case the Boolean model is a union of such disks, centered at $\{p_1, p_2, \ldots\}$, respectively.

Discrete random sets are defined on a countable $L \subset \mathbb{R}^n$, e.g., a regular sampling lattice. Consider a *Binomial germ–grain* discrete random set, for which the germ process is a homogeneous Bernoulli process on L, of intensity $\lambda \in [0, 1]$, and the grains are random subsets

* Corresponding author. Tel. 001-804-924-6310; fax: 001-804-924-8818.

E-mail address: nikos@virginia.edu (N.D. Sidiropoulos)

PII: S0031-3203(99)00029-1

of L. A very special instance of the Binomial germ–grain model is one in which the grain associated with each point in the lattice is fixed and known a priori; we will call this a *fixed-grain binomial germ–grain* (discrete) random set. The only source of randomness in such a model is the (homogeneous Bernoulli lattice) germ process; it is completely specified via the associated intensity parameter $\lambda \in [0, 1]$: each and every point in L is an element of P with probability λ, independently of all others.

This is, admittedly, a *very* restricted model; however, it will serve our purposes perfectly, for, on one hand, it is "rich enough" to exhibit the fundamental computational limitation associated with ML germ estimation; while, on the other hand it is simple enough to allow us to make a formal intractability argument, by means of NP-completeness. Also note that, even though the general binomial germ–grain model *is not* a Boolean model (Boolean models are Poisson processes on the space of compact sets), the fixed-grain binomial germ–grain model is equivalent to a fixed-grain Boolean model on the lattice.

The Boolean model has received a lot of attention in the random set literature, for a variety of reasons, including central limit theory arguments, its plausibility in certain modeling situations, and its relative tractability. The literature on statistical inference of the Boolean model has focused on *parameter estimation*, i.e., the estimation of the intensity of the germ process and/or the size distribution of the grains [4–10]. The intensity parameter of the germ process is the (ensemble) mean number of germs per unit area. Estimates of the intensity parameter are used to characterize the average density of, e.g., cell populations, across many sample images. Practitioners are often interested in *counting* cells in a particular sample image, i.e., a more refined estimate of the germ process per se. This clearly is more ambitious than estimating the intensity parameter of the germ process, a difficult problem in its own right. It is shown in Section 2 that the pursuit of such an estimate of the germ process is overly ambitious, in the sense that it entails the solution of an NP-complete computational problem.

NP stands for the class of problems solvable in *nondeterministic polynomial* time. There exist deterministic Turing machines that solve these problems, but their run time is at best exponential in the size of the input. P is the class of problems solvable in *polynomial* time by a deterministic Turing machine. By definition, a problem Π is NP-complete if Π belongs to NP, and every other problem in NP can be transformed to Π in polynomial time. A problem in NP can also be shown to be NP-complete by proving that it contains a known NP-complete problem as a special case. This is known as *proof by restriction*. If an algorithm could be developed to solve any given NP-complete problem in deterministic polynomial time, then *every* problem in NP would have been solvable in deterministic polynomial time. Although no one has proven that this is impossible, the current consensus is that the likelihood of such an algorithm ever emerging is very small, as NP contains many computationally very demanding problems which have been under scrutiny for a long time. NP-completeness is a cornerstone of theoretical computer science. A classic reference on the subject is the book by Garey and Johnson [11].

This paper is organized as follows. Section 2 is concerned with the problem of maximum likelihood estimation of the germ process associated with a binomial germ–grain model. It contains the basic NP-completeness result, and discusses its ramifications in terms of the closely related problem of computing *minimal skeletons*. Section 3 is concerned with some intractability issues that arise in the context of probabilistic specification of morphologically processed random sets. Section 4 offers a synopsis of the work presented in this paper.

2. The problem of estimating the germ process of a Binomial germ–grain model is NP-complete

Theorem 1. *Maximum Likelihood estimation of the germ process realization associated with a given binomial germ–grain random set observation is an NP-complete computational problem.*

Proof. The proof is by restriction to a known NP-complete problem. Such a proof consists simply of showing that the problem at hand contains a known NP-complete problem as a special case. The key is to "peel off" all inessential aspects of the problem at hand, until a known NP-complete problem emerges.

Consider a fixed-grain binomial germ–grain random set X, defined on a *finite* lattice L. Since the proof is by *restriction*, we may assume that (i) $\lambda \in [0, .5)$, and (ii) the deterministic grains associated with distinct lattice points are distinct as well, i.e., the association between lattice points and the set of grains is 1-1 and onto. We therefore adopt a slight abuse of notation and use G_p to denote the grain associated with location $p \in L$ (i.e., G_p has absorbed the translation, and is "centered" at p).

Given a realization S of X, maximum likelihood estimation of the germs is equivalent to the following:

$$\text{maximize}: \lambda^{|P|}(1-\lambda)^{|P^c|}$$

$$\text{subject to}: \bigcup_{p \in P \subseteq L} G_p = S, \tag{1}$$

where c stands for complementation with respect to L, i.e., $P^c = L - P$, $|\cdot|$ measures number of elements, and G_p is the grain associated with p. It follows from (i) that

(1) is equivalent to the constrained minimization problem:

$$\text{minimize}: |P|$$

$$\text{subject to}: \bigcup_{p \in P \subseteq L} G_p = S. \tag{2}$$

The problem is clearly in NP (testing all possible subsets will yield the solution). Define a collection C of subsets of S as follows:

$$C = \{G_p, p \in L \,|\, G_p \subseteq S\}$$

and Eq. (2) may be re-written as

$$\text{minimize}: |C'|$$

$$\text{subject to}: \bigcup_{c \in C' \subseteq C} c = S. \tag{3}$$

The most important part of an NP-completeness proof by the method of reduction is identifying a suitable problem from the palette of known NP-complete problems. For our present purpose, a most convenient known NP-complete problem is the so-called *minimum cover* problem, which can be compactly stated as follows [11]:

Problem 1. (Minimum Cover - Decidability Version) **Instance:** A collection C of subsets of a set S, and a positive integer K. **Question.** Does C contain a *cover* for S of size K or less, that is, a subset $C' \subseteq C$ with $|C'| \leqslant K$ and such that $\bigcup_{c \in C'} c = S$?

It follows directly that the associated estimation problem is NP-complete:

Problem 2. ((Size of) Minimum Cover) **Instance:** A collection C of subsets of a set S. **Question.** What is the smallest positive integer K such that C contains a *cover* for S of size K, that is, a subset $C' \subseteq C$ with $|C'| = K$ and such that $\bigcup_{c \in C'} c = S$?

which, in turn, implies NP-completeness of the problem in Eq. (3).

The following is an important Corollary:

Corollary 1. *Maximum likelihood estimation of the germ process realization associated with a given Boolean random set observation is an NP-complete computational problem.*

Proof. Recall that the fixed-grain binomial germ–grain model considered above is equivalent to a special case of a Boolean model; the proof by restriction argument therefore applies. □

It is perhaps appropriate to comment briefly on the meaning of NP-completeness. The fact that a certain computational problem is NP-complete means that (i) algorithms that solve a *general* instance of the problem exist (enumeration of all *possible* solutions is often an obvious such "algorithm"), but (ii) the associated computational complexity (for non-trivial problem sizes) is very high, at least as high as that required to solve any problem from a class of very demanding problems for which an efficient (polynomial-time) algorithm is not known, and is highly unlikely to be found.

However, the key is the word *general*: special instances of the problem may well admit an efficient solution. In our present context, such special cases that immediately come to mind are

- $\lambda > 0.5$: In which case, the problem becomes one of finding a *maximum cover*, which is easy: visit each $p \in L$, and test whether $G_p \subseteq S$; if so, then $p \in \hat{P}_{\mathrm{ML}}$; else $p \notin \hat{P}_{\mathrm{ML}}$, where $\hat{P}_{\mathrm{ML}}$ is the ML estimate of the germ process. This situation is, in a way, less interesting from a practical standpoint. For example, in the modeling of cell cultures, $\lambda > 0.5$ would correspond to very dense cell populations; in the modeling of clutter and in filtering applications, $\lambda > 0.5$ would correspond to severe occlusion/distortion.
- Fixed-grain binomial germ–grain models in 1-D, in which case simple algorithms for estimating the germ process can often be constructed, e.g., based on dynamic programming (DP) or other recursive arguments. The 1-D case is a very special one in many other respects as well, e.g., cf. [12,13].

It is important to stress that these special cases are not counter-examples to the general NP-completeness result. As with all NP-complete problems, *over*-simplifying the problem by imposing a lot of additional structure will eventually render the problem polynomial-time.

2.1. *Ramifications*: *minimal skeletons*

The concept of *skeleton* as a compact description ("sketch") of shape can be traced back to the work of Lantuéjoul [14,15]. Depending on the properties sought, one may define a skeleton in many ways. A common objective is *economy* of representation, primarily due to the emphasis on recognition and coding applications [16–18].

In loose terms, the general idea is as follows. Given a shape, S, find a collection of points $P = \{p_1, p_2, \ldots\}$ (the skeleton), and associated grains $\{G_{p_1}, G_{p_2}, \ldots\}$ (the skeleton *functional*) such that $G_{p_i} \in \mathscr{G}, \forall i$, and

$$S = \bigcup_{i=1,2,\ldots} G_{p_i} \oplus \{p_i\}.$$

Of course, numerous solutions may exist. Depending on the application, additional restrictions may be imposed

on P, to assure, e.g., connectivity or other desirable properties, and these may lead to a unique solution. From a coding perspective, *economy* is the prime concern, and this is reflected, for example, in the *reduced morphological skeleton* proposal of Maragos [16,17]. It is therefore natural to ask for a *minimal* skeleton representation. Define a collection C of subsets of S as follows:

$$C = \{G \oplus \{p\} \subseteq S, G \in \mathcal{G}, p \in L\}$$

and the problem of finding a minimal skeleton representation with respect to $\mathcal{G}$ may be stated as

$$\text{minimize}: |C'|$$

$$\text{subject to}: \bigcup_{c \in C' \subseteq C} c = S, \tag{4}$$

which is NP-complete, as we have seen. This means that the problem of computing minimal skeleton representations is, in general, computationally intractable. The result constitutes an effective argument in support of reduced (but not necessarily minimal) skeleton representations [16–18].

3. Probabilistic Specification of Morphologically Processed Random Sets

The celebrated Choquet–Kendall–Matheron theorem [3,19,20] for random closed sets, states that a RACS X is completely characterized by its *capacity functional*, i.e., the collection of hitting probabilities over a sufficiently rich family of so-called *test sets*. Random set theory is a mature branch of theoretical and applied probability, whose scope is the study of set-valued random variables. There exist numerous references on the subject; cf. [3,20] for foundations, and the recent books [21,4] for many references on various aspects of random set theory. An extensive bibliography on random closed sets and related topics can be found on-line under http://liinwww.ira.uka.de/bibliography/Math/ random.closed.sets.html.

For our present purposes, it suffices to focus attention on finite discrete random sets (DRS), a special class of random sets defined on a finite lattice. A DRS may be thought of as a sampled version of an underlying RACS; cf. [22] for a rigorous analysis of a suitable sampling process, as well as a formal argument which establishes the usefulness of DRS theory. DRSs can be viewed as finite-alphabet random variables, taking values in a finite *partially ordered* set (poset).

One of the "wheels" of discrete probability and combinatorial theory is the principle of *inclusion–exclusion* (e.g., cf. the classic book of Feller [23]). Although this principle applies to numerous problems, it is often hard to make the connection, and realize that it indeed applies to a particular problem at hand. As noted by Rota [24] "It has often taken the combined efforts of many a combinatorial analyst over long periods to recognize an inclusion–exclusion pattern."

As it turns out, the inclusion–exclusion principle is the simplest but also the typical case of a very general principle of enumeration, regarding the inversion of indefinite sums ranging over an arbitrary poset. This principle is known as *Moebius inversion*, and it is the analog of the "Fundamental Theorem of calculus" in the context of enumeration.

For the special case of finite DRSs, a variant of Moebius inversion provides an elementary *constructive* proof of the Choquet–Kendall–Matheron Theorem.

3.1. Preliminaries

The following theorem is a cornerstone of enumeration and combinatorics. Cf. [25] for a general proof.

Theorem 2 (Moebius inversion for Boolean algebras). *Let L be a finite set (i.e., $|L| < \infty$), and $\Sigma(L)$ be its power set. $(\Sigma(L), \subseteq)$ is a complete lattice with unit element L and zero element $\emptyset$. $(\Sigma(L), \subseteq)$ is isomorphic to the Boolean Algebra of (finite) rank $|L|$. Let p be a real-valued functional on $\Sigma(L)$. Define the lower and upper sum functionals, q, and r, respectively, by*

$$q(A) \triangleq \sum_{S \subseteq A} p(S), \quad \forall A \in \Sigma(L),$$

$$r(A) \triangleq \sum_{S \supseteq A} p(S), \quad \forall A \in \Sigma(L)$$

Then, $\forall S \in \Sigma(L)$ (inversion from below)

$$p(S) = \sum_{A \subseteq S} (-1)^{|A|} q(S \cap A^c) = \sum_{A \subseteq S} (-1)^{|S|-|A|} q(A)$$

and, $\forall S \in \Sigma(L)$ (inversion from above)

$$p(S) = \sum_{A \supseteq S} (-1)^{|A^c|} r(S \cup A^c) = \sum_{A \supseteq S} (-1)^{|A|-|S|} r(A),$$

where $A^c = L - A, \forall A \in \Sigma(L)$.

A stand-alone proof involves a convenient change of variables, followed by the application of an enumeration lemma.

We will need the following technical lemma. Although it cannot be found "as-is" in the classic references, [24,26,27] or in the relatively up-to-date book of Aigner [25] the statement has appeared in the context of *Belief function* computation [28] and a proof can be constructed using Moebius inversion from above.

Lemma 1.

$$q(A) = \sum_{S \subseteq A^c} (-1)^{|S|} r(S) = \sum_{S \supseteq A} (-1)^{|S^c|} r(S^c), \quad \forall A \in \Sigma(L)$$

and

$$r(A) = \sum_{S \subseteq A} (-1)^{|S|} q(S^c) = \sum_{S \supseteq A^c} (-1)^{|S^c|} q(S), \quad \forall A \in \Sigma(L).$$

Consider a mapping $\phi : \to \Sigma(L) \mapsto \Sigma(L)$. ϕ is called an *erosion* if it distributes over intersection. ϕ is called a *dilation* if it distributes over union. A pair (ε, δ) of mappings from $\Sigma(L)$ to itself is called an *adjunction* on $\Sigma(L)$ if

$$\delta(A) \subseteq S \Leftrightarrow A \subseteq \epsilon(S), \quad \forall A \in \Sigma(L), \quad \forall S \in \Sigma(L).$$

If (ε, δ) is an adjunction, then ε is necessarily an erosion (i.e., it distributes over intersection), and δ is a dilation (i.e., it distributes over union). For any given erosion, ε, there exists a unique dilation, δ, dubbed the *right adjoint* of ε, such that the pair (ε, δ) is an adjunction. This δ is given by

$$\delta(A) = \cap\{S \in \Sigma(L) | A \subseteq \varepsilon(S)\}.$$

Similarly, for any given dilation, δ, there exists a unique erosion, ε, dubbed the *left adjoint* of δ, such that the pair (ε, δ) is an adjunction. This ε is given by

$$\varepsilon(S) = \cup\{A \in \Sigma(L) | \delta(A) \subseteq S\}.$$

If (ε, δ) is an adjunction, then ε and δ are *adjoint* to each other. A thorough treatment of adjunctions can be found in Ref. [29] Related material can also be found in Ref. [30].

We will need the following elementary Lemma. A proof can be constructed along the lines of the material in pp. 85–86 of Heijmans [29].

Lemma 2. *Any dilation, δ, can be represented as*

$$\delta(A) = \cup_{z \in A} S(z).$$

Given $\delta(A) = \cup_{z \in A} S(z)$ define

$$\varepsilon(A) \triangleq \{z \in L \mid S(z) \subseteq A\}.$$

It can be shown [29] that (ε, δ) is an adjunction. Furthermore, it can be shown (by uniqueness of left adjoint) that

Lemma 3. *Any erosion, ε, can be represented as*

$$\varepsilon(A) = \{z \in L \mid S(z) \subseteq A\}$$

We will need the following Lemma. The proof is elementary.

Lemma 4. *Let (ε, δ) be an adjunction. Define*

$$q_\delta(A) \triangleq \sum_{\delta(S) \subseteq A} p(S)$$

and

$$r_\epsilon(A) \triangleq \sum_{\epsilon(S) \supseteq A} p(S).$$

Then

$$q_\delta(A) = q(\epsilon(A)), \quad \forall A \in \Sigma(L)$$

and

$$r_\epsilon(A) = r(\delta(A)), \quad \forall A \in \Sigma(L).$$

3.2. Finite DRSs

A DRS X is simply a probability measure on $\Sigma(\Sigma(L))$. Denote this by $P_X(\cdot)$. Let $p_X(\cdot)$ denote the restriction of $P_X(\cdot)$ to the atoms, i.e., the elements of $\Sigma(L)$. This is the *probability mass function* of the DRS X. Define the *capacity functional, generating functional, covering functional, and cumulative distribution functional*, $T_X(\cdot), Q_X(\cdot), R_X(\cdot), F_X(\cdot)$, respectively, by

$$T_X(A) \triangleq P_X(X \cap A \neq \emptyset),$$

$$Q_X(A) \triangleq P_X(X \cap A = \emptyset) = 1 - T_X(A) = \sum_{S \subseteq A^c} p_X(S),$$

$$R_X(A) \triangleq P_X(X \supseteq A) = \sum_{S \supseteq A} p_X(S),$$

$$F_X(A) \triangleq 1 - R_X(A).$$

Identify p_X with the functional p of Theorem 2. Then, it becomes clear that

$$Q_X(A) = q(A^c)$$

and

$$R_X(A) = r(A),$$

where q, r are the lower, and upper sum functionals, respectively, of Theorem 2. It now follows from Theorem 2, and Lemma 1, that $\forall S \in \Sigma(L)$

$$p_X(S) = \sum_{A \subseteq S} (-1)^{|A|} Q_X(S^c \cup A) = \sum_{A \subseteq S} (-1)^{|S|-|A|} Q_X(A^c)$$

and

$$p_X(S) = \sum_{A \supseteq S} (-1)^{|A^c|} R_X(S \cup A^c) = \sum_{A \supseteq S} (-1)^{|A|-|S|} R_X(A).$$

Furthermore, $\forall A \in \Sigma(L)$

$$Q_X(A) = \sum_{S \subseteq A} (-1)^{|S|} R_X(S) = \sum_{S \supseteq A^c} (-1)^{|S^c|} R_X(S^c)$$

and

$$R_X(A) = \sum_{S \subseteq A} (-1)^{|S|} Q_X(S) = \sum_{S \supseteq A^c} (-1)^{|S^c|} Q_X(S^c).$$

Since $Q_X(A) = 1 - T_X(A)$ and $F_X(A) = 1 - R_X(A)$, $\forall A \in \Sigma(L)$, we now have identities which relate all five functionals T_X, Q_X, R_X, F_X, and p_X. As a trivial corollary, any one of these functionals is sufficient to specify X.

3.3. Connection with mathematical morphology

Mathematical morphology is an important quantitative shape analysis tool in image processing. Its foundations were laid down by Matheron [31,32], Serra [33,34] and collaborators, during the late 1960s to early 1980s. An excellent recent treatment which unifies several seemingly distinct approaches within a purely algebraic framework can be found in Ref. [29].

Morphological image operators [29] are compositions of two classes of elementary building blocks, namely erosions and dilations. Several special cases of erosions and dilations can be defined; the definition given in Section 3.1 is the most general one within our framework.

Let (ε, δ) be an adjunction. Identify p_X with the functional p of Theorem 2. Then, as before, $Q_X(A) = q(A^c)$ and $R_X(A) = r(A)$. We further observe that $Q_{\delta(X)}(A) = q_\delta(A^c)$, and $R_{\epsilon(X)}(A) = r_\epsilon(A)$, where q_δ, and r_ϵ have been defined in Lemma 4. By applying this latter lemma, we conclude

$$Q_{\delta(X)}(A) = Q_X((\epsilon(A^c))^c), \quad \forall A \in \Sigma(L)$$

and

$$R_{\epsilon(X)}(A) = R_X(\delta(A)), \quad \forall A \in \Sigma(L).$$

A probabilistic proof of the first result for the special case of translation-invariant operators has appeared [35]; similarly, a probabilistic proof of the second result for the special case of translation-invariant operators has appeared [36].

In the previous Section, we have concluded that either one of the functionals T_X, Q_X, R_X, F_X, or p_X, is a sufficient and constructive specification of X. From the latter two identities we now conclude that Q_X (or, equivalently, $T_X = 1 - Q_X$) is the most convenient specification if our interest is in processing X via an operator which distributes over union (i.e., a dilation), whereas R_X (or, equivalently, $F_X = 1 - R_X$) is the most convenient specification if our interest is in processing X via an operator which distributes over intersection (i.e., an erosion). This picture is depicted in Fig. 1. One may obviously substitute T_X for Q_X, and/or F_X for R_X. The choice of Q_X versus R_X depends solely on whether one will apply a dilation, or erosion, respectively, on X. However, in morphological shape analysis and synthesis we sequentially process X using a variety of erosions and dilations. This means that one is forced to use *both* specifications, i.e., move between the two in anticipation of the next operation in line. This movement is made possible by Lemma 1, which is essentially another incarnation of Moebius inversion. As such, it involves a combinatorial computational cost: given all necessary values of the generating functional (or, equivalently, the capacity functional), it takes $O(2^{|A|})$ elementary operations to compute the value of the covering functional (or, equivalently, the cumulative distribution functional) for argument A, and vice-versa.

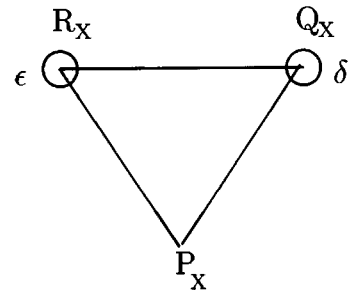

Fig. 1. The three fundamental functionals.

3.4. Fast Moebius transform

If the underlying lattice (base frame) L consists of n points, it appears that one needs $O(2^n 2^n) = O(4^n)$ elementary operations to move between the two basic alternative random set specifications. As it turns out, this is not the case: there exists a "fast Moebius Transform" due to Thoma [28] (also cf. [37]) that allows one to move more efficiently between the two specifications. The basic idea behind this algorithm is simple and quite elegant. Recall that

$$Q_X(A) = \sum_{S \subseteq A^c} p_X(S)$$

and

$$p_X(A) = \sum_{S \subseteq A} (-1)^{|A|-|S|} Q_X(S^c)$$

form a Moebius pair. Thoma's algorithm transforms $p_X(\cdot)$ into $Q_X(\cdot)$ in n steps, each step associated with one element of L. Define $m_0(\cdot) \triangleq p_X(\cdot)$, and think of $m_i(\cdot)$ as a *state* functional. Let $L = \{\theta_1, \ldots, \theta_n\}$. The ith step of the algorithm (associated with θ_i) modifies the state functional as follows:

$$m_i(A) = \begin{cases} m_{i-1}(A) + m_{i-1}(A \backslash \{\theta_i\}), & \theta_i \in A, \\ m_{i-1}(A), & \theta_i \notin A. \end{cases}$$

Thoma [28] proves that, after the completion of the nth step, $m_n(A) = Q_X(A^c)$, $\forall A \subseteq L$. In addition, each step in the algorithm can be inverted; [1,2,28] this provides a similar algorithm to compute the inverse transform:

$$m_{i-1}(A) = \begin{cases} m_i(A) - m_i(A \backslash \{\theta_i\}), & \theta_i \in A, \\ m_i(A), & \theta_i \notin A. \end{cases}$$

Going backwards, $m_0(\cdot) = p_X(\cdot)$ eventually results. Similar algorithms can be constructed for moving between the remaining functionals [28].

It is easy to see that Thoma's algorithm requires $O(n2^n)$ (instead of a brute-force $O(4^n)$) elementary computations to transform a given functional into another functional. This is a *faster* Moebius transform, but unfortunately not *fast enough* for our purposes: still, complexity is exponential in n, and n is in the order of 2^{16}–2^{20} for typical

image processing applications. Even for a 16×16 image ($n = 256$), $n2^n = 10^{73}$ Million elementary operations, which is clearly out of question.

Given, e.g., $Q_X(A), \forall A \subseteq L$, Thoma's algorithm computes $p_X(A), \forall A \subseteq L$. In estimation and detection problems (e.g., hypothesis testing), one is often interested in computing the likelihood of a given observation, $A \subseteq L$, under a postulated data-generating model, i.e., $p_X(A)$ for a *single* A, and given *model* $Q_X(A), \forall A \subseteq L$. Brute-force computation of this quantity would be $O(2^{|A|}) \leqslant O(2^n)$ - thus the fast Moebius transform does not facilitate this computation.

In the context of Belief function computations, certain shortcuts exist that build upon special properties of the functionals involved [28]. For example, suppose that $Q_X(A)$ exhibits a *typical value*, e.g., without loss of generality, suppose that $Q_X(A) = 0$ for all but a few *atypical* subsets of L. Then, clearly, the computation of

$$p_X(A) = \sum_{S \subseteq A} (-1)^{|A|-|S|} Q_X(S^c)$$

can be greatly simplified, by considering only those subsets of A whose complements are atypical. While fruitful in the context of Belief function computations, this idea does not appear to provide any computational savings for an independent germ–grain random set model, since no one of its functionals exhibits a typical value: even for a simple binomial germ - deterministic grain model, $Q_X(A)$ is a function of the *shape* of A.

In short, the computational complexity of Moebius transformation appears to be a significant barrier for applications.

4. Conclusions

This paper focused on some computational (in)tractability issues associated with certain inference problems for random sets. A key result is that maximum likelihood estimation of the germ process embedded in a binomial germ–grain model is an NP-complete computational problem. The implications of this result in the context of computing minimal skeleton representations have been discussed. The second part of the paper has examined computational problems associated with the probabilistic specification of morphologically processed random sets. In this latter direction, recent results on a fast Moebius transformation [28,37] have been examined in the context of statistical inference and morphological filtering of random sets.

Acknowledgements

The author would like to thank Dr. Paul K. Black for providing copies of references [28,37] and a stimulating discussion; Dr. Ilya Molchanov for several insightful and constructive comments on an earlier draft of the paper; and two Anonymous reviewers, whose remarks helped in improving this paper.

References

[1] A.F. Karr, Point Processes and their Statistical Inference, Marcel Dekker, New York, 1990.
[2] D. Stoyan, W.S. Kendall, J. Mecke, Stochastic Geometry and its Applications, Wiley, Berlin, 1987.
[3] G. Matheron, Random Sets and Integral Geometry, Wiley, New York, 1975.
[4] I.S. Molchanov, Statistics of the Boolean Model for Practitioners and Mathematicians, Wiley, Chichester, 1997.
[5] I.S. Molchanov, Estimation of the size distribution of spherical grains in the Boolean model, Biometrical J. 32 (1990) 877–886.
[6] I.S. Molchanov, A consistent estimate of parameters of Boolean models of random closed sets, Theory Probab. Appl. 36 (1991) 600–607.
[7] I.S. Molchanov, On statistical analysis of Boolean models with non-random grains, Scand. J. Statist. 21 (1994) 73–82.
[8] I.S. Molchanov, D. Stoyan, Asymptotic properties of estimators for parameters of the Boolean model, Adv. Appl. Probab. 26 (1994) 301–323.
[9] I.S. Molchanov, Statistics of the Boolean model: from the estimation of means to the estimation of distributions, Adv. Appl. Probab. 27 (1995) 63–86.
[10] N.D. Sidiropoulos, J.S. Baras, C.A. Berenstein, An algebraic analysis of the generating functional for discrete random sets, and statistical inference for intensity in the discrete boolean random set model, J. Math. Imaging Vision 4 (1994) 273–290.
[11] M.R. Garey, D.S. Johnson, Computers and Intractability: A Guide to the Theory of NP-completeness, Freeman and Co., New York, 1979.
[12] E.R. Dougherty, J.C. Handley, Recursive maximum-likelihood estimation in one-dimensional discrete Boolean random set model, Signal Processing 43 (1995) 1–15.
[13] J.C. Handley, E.R. Dougherty, Maximum-likelihood estimation for discrete Boolean models using linear samples, J. Microsc. 182 (1996) 67–78.
[14] C. Lantuéjoul, La Squelettisation et son Application aux Mesures Topologiques des Mosaiques Cristallines, Ph.D. Thesis, Ecole Nationale Supérieure des Mines de Paris, Paris, 1978.
[15] C. Lantuéjoul, Skeletonization in quantitative metallography, in: R.M. Haralick, J.C. Simon (Eds.), Issues in Image Processing, Sijthoff & Noordhof, Alphen, 1980, pp. 107–135.
[16] P. Maragos, R.W. Schafer, Morphological skeleton representation and coding of binary images, IEEE Trans. Acoustics, Speech, Signal Process. 34 (1986) 1228–1244.
[17] P. Maragos, Pattern spectrum and multiscale shape representation, IEEE Trans. Pattern. Anal. Mach. Intell. 11 (7) (1989) 701–716.

[18] J. Goutsias, D. Schonfeld, Morphological representation of discrete and binary images, IEEE Trans. Signal Process. 39 (6) (1991) 1369–1379.
[19] G. Choquet, Theory of capacities, Ann, Inst. Fourier 5 (1953) 131–295.
[20] D.G. Kendall, Foundations of a theory of random sets, in: E.F. Harding, D.G. Kendall (Eds.), Stochastic Geometry, Wiley, London, England, 1974, pp. 322–376.
[21] I.S. Molchanov, Limit theorems for unions of random closed sets, vol. 1561, Lecture Notes in Mathematics, Springer, Berlin, 1993.
[22] K. Sivakumar, J. Goutsias, Binary random fields, random closed sets, and morphological sampling, IEEE Trans. Image Process. 5 (6) (1996) 899–912.
[23] W. Feller, An Introduction to Probability Theory and its Applications, Third ed. (revised printing), Wiley, New York, 1970 (two volumes).
[24] G.-C. Rota, On the foundations of combinatorial theory: I. Theory of moebius functions, Z. Wahrscheinlichkeits Theorie Verw. Gebiete, 2 (1964) 340–368. Also appears in I. Gessel, G-C. Rota (Eds.), Classic Papers in Combinatorics, Birkhauser, Boston, (1987).
[25] M. Aigner, Combinatorial Theory, Springer, New York, 1979.
[26] H.H. Crapo, Moebius inversion in lattices, Arch. Math., 19 (1968) 595–607. Also appears in I. Gessel, G-C. Rota (Eds.), Classic Papers in Combinatorics, Birkhauser, Boston, (1987).
[27] G. Birkhoff, Lattice Theory, (Third ed.), vol. 25, AMS Colloquium Publications, AMS, Providence, RI, 1967.
[28] H. Mathis Thoma, Belief function computations, in: I.R. Goodman, M.M. Gupta, H.T. Nguyen, G.S. Rogers (Eds.), Conditional Logic in Expert Systems, Elsevier Science Publishers B.V, North Holland, Amsterdam, 1991, pp. 269–308.
[29] H.J.A.M. Heijmans, Morphological Image Operators, Academic Press, Boston, 1994.
[30] A.J. Baddeley, H.J.A.M. Heijmans, Incidence and lattice calculus with applications to stochastic geometry and image analysis, Research Report BS-R9213, CWI, 1992.
[31] G. Matheron, Elements Pour Une Theorie Des Milieux Poreux, Masson, Paris, 1967.
[32] G. Matheron, Random Sets and Integral Geometry, Wiley, New York, 1975.
[33] J. Serra, Image Analysis and Mathematical Morphology, Academic, New York, 1982.
[34] J. Serra (Ed.), Image Analysis and Mathematical Morphology, vol. 2, Theoretical Advances, Academic Press, San Diego, 1988.
[35] N.D. Sidiropoulos, Statistical Inference, Filtering, and Modeling of Discrete Random Sets, Ph.D. Thesis, University of Maryland, June, 1992.
[36] J. Goutsias, Morphological Analysis of Discrete Random Shapes, J. Math. Imaging Vision 2 (1992) 193–215.
[37] P. Black, An Examination of Belief Functions and Other Monotone Capacities, Ph.D. Thesis, Carnegie Mellon University, April, 1996.

About the Author—NICHOLAS D. SIDIROPOULOS received the Diploma in Electrical Engineering from the Aristotelian University of Thessaloniki, Greece, and the M.S. and Ph.D. degrees in Electrical Engineering from the University of Maryland at College Park, in 1988, 1990 and 1992, respectively. From September 1992 to June 1994 he served his military service, also working part-time as a member of the technical staff, Systems Integration Division, G-Systems Ltd., Athens, Greece. From 1994 to 1997 he was a Research Scientist at the Institute for Systems Research, University of Maryland. He joined the Department of Electrical Engineering at the University of Virginia in the summer of 1997, where he is currently an Assistant Professor. His research interests include nonlinear filtering, mathematical morphology and random set theory, and nonparametric nonlinear regression; recent work includes bilinear shape-constrained regression problems in chromatographic analysis, flow injection analysis, and multiuser communication; and broadband signal and image compression.