SIZE-BIASED RANDOM CLOSED SETS

M. N. M. VAN LIESHOUT
*Centrum voor Wiskunde en Informatica**
POBox 94079, 1090 GB Amsterdam, The Netherlands

Abstract. We indicate how granulometries may be useful in the analysis of random sets. After defining a size distribution function which may be used as a summary statistic in exploratory data analysis, we propose a Hanisch-type estimator and construct new Markov random set models which favour certain sizes above others. The models are illustrated by simulated realisations.

Key words: contact distribution function, empty space function, granulometry, Hanisch-type estimator, Markov random set, size distribution function.

1. Introduction

Stochastic geometry [29] is concerned with the study of random closed sets (rcs). Roughly speaking, an rcs is a mapping X from a probability space into the family of closed subsets of R^d such that $\{X \cap K \neq \emptyset\}$ and $\{X \cap G = \emptyset\}$ are measurable for all compact sets K and all open sets G. The probability distribution of X is completely specified by its *capacity functional*

$$T(K) = P(X \cap K \neq \emptyset) \tag{1}$$

with K ranging over the class of compact sets, which plays a role comparable to that of the distribution function of a real-valued random variable. However, the collection of test sets K is huge, and lower-dimensional summary statistics are called for. Typically these are obtained from the capacity functional (1) by restricting the choice of K. For instance, allowing only singletons results in the coverage probabilities $p(x) = P(x \in X)$, $x \in R^d$.

Below we will assume that the random closed set X is stationary, i.e. its distribution is invariant under translations. In that case, the coverage probabilities $p(x)$ do not depend on the argument x, and $p(x) \equiv p = E|X \cap U|$, the expected volume covered within any set U of unit volume $|U| = 1$. To exclude degenerate cases, it is assumed that $0 < p < 1$.

A summary statistic for assessing the 'size' of pores left open by a stationary rcs is the *empty space function* $F_B(\cdot)$ defined by taking $K = rB$ $(r \geq 0)$ in (1), that is

$$F_B(r) = P(0 \in X \oplus r\check{B}). \tag{2}$$

The related *contact distribution function* is defined as

$$H_B(r) = P(0 \in X \oplus r\check{B} \mid 0 \notin X) = \frac{F_B(r) - F_B(0)}{1 - F_B(0)} \tag{3}$$

* This research was partially carried out while the author was at the University of Warwick, Coventry CV4 7AL, UK.

for $r \geq 0$. Both (2) and (3) depend on a structuring element B, usually assumed to be a compact convex set containing an environment of the origin. Typical choices include balls and squares, although non-isotropic structuring elements may be preferred when investigating directional effects. Under these conditions, $H_B(\cdot)$ is a distribution function but this is not the case for general B (see [29]). Note that $1 - H_B(r)$ can be interpreted as the conditional probability that a copy of rB placed at a test point 0 lies entirely in the empty space left by X given that the test point itself does not fall in X. A similar interpretation holds for the empty space function $F_B(r)$.

The particle size operator underlying (2)–(3) is $\psi_r(Y) = (Y^c \oplus r\check{B})^c = Y \ominus r\check{B}$. However, although under our restrictions on B, $\psi_r(\cdot)$ is anti-extensive and increasing, in general the sieving condition

$$\psi_r(\psi_s(Y)) = \psi_s(\psi_r(Y)) = \psi_{\max(r,s)}(Y)$$

for all $r, s \geq 0$ does not hold and hence $(\psi_r)_{r \geq 0}$ is not a granulometry [21]. Thus, in the remainder of this paper, we will consider replacing the dilation in (2)–(3) by a closing.

2. Size distribution functions

In mathematical morphology, a *size distribution law* of a stationary random closed set X may be defined using a granulometry $\psi_r(X) = X \circ rB$ ($r \geq 0$) based on Euclidean openings with a non-empty convex compact structuring element B. Indeed, set [28, p. 335]

$$G_1(r) = 1 - P(x \in X \circ rB \mid x \in X) = 1 - \frac{P(x \in X \circ rB)}{P(x \in X)}, \quad r \geq 0$$

for the size distribution law of the particles, and similarly

$$G_0(r) = P(x \in X \bullet rB \mid x \notin X) = 1 - \frac{1 - P(x \in X \bullet rB)}{1 - P(x \in X)}, \quad r \geq 0$$

for the pores. By the stationarity of X, the definitions do not depend on the choice of $x \in R^d$. Intuitively, $G_1(r)$ will be the conditional probability that a point in X is eliminated by opening with rB ($r > 0$), i.e. the probability that the B-size of X is less than r. Similarly, $G_0(r)$ is the conditional probability that a point from X^c is included by closing with rB, hence that the B-size of the complement is less than or equal to r. By allowing r to vary over $(-\infty, \infty)$, the joint size distribution law

$$G(r) = \begin{cases} 1 - P(x \in X \circ rB) & r \geq 0 \\ 1 - P(x \in X \bullet |r|B) & r < 0 \end{cases} \tag{4}$$

of pores and particles is obtained.

Size distribution laws have been used for a long time in the empirical sciences, and more recently in the analysis of (binary) images. For instance Serra [28] employs size distribution laws for shape and texture analysis, Maragos [20] uses them for multiscale shape representation, [16, 25] apply size distributions to shape filtering and restoration problems, while Sivakumar [27] gives applications in texture classification and morphological filtering. See also [9].

However, from a probabilistic point of view, $G(\cdot)$ (and $G_1(\cdot)$) is not a proper distribution function, as it is semi-continuous from the left rather than from the right. For this reason we prefer the following definition.

Definition 1 *Let X be a stationary random closed set and B a non-empty convex compact structuring element. Define the* size distribution function *of X by*

$$P_B(r) = \begin{cases} P(x \in X \bullet rB) & r \geq 0 \\ P(x \in X \circ |r|B) & r < 0 \end{cases} . \tag{5}$$

The function $P_B(\cdot)$ is called the *granulométrie bidimensionelle* in metallurgi [10]. It is easily verified that $P_B(0) = p$, the coverage fraction of the stationary random closed set X.

The size distribution function $P_B(\cdot)$ is well-defined and does not depend on the choice of $x \in R^d$. In contrast to $G(\cdot)$, it is a proper distribution function. Compared to the empty-space function (2), note that the latter is absolutely continuous except in 0 (see Hansen et al. [15]) but that this property does not generally hold for the size distribution which may have countably many discontinuities.

Explicit expressions for size distributions may be hard to find, being related to covering probabilities [13]. The contact distribution function is available in closed form for Boolean models [29] (but not for most other random set models!) and, if the primary grain is convex, depends only on the moments of a few functionals of the grain. For instance, in R^2 the mean perimeter determines the whole of $H_B(r)$, which, as pointed out in Ripley [23], may result in poor distinguishing power as an exploratory data analysis tool.

3. Estimation and edge effects

In this section we discuss estimating the size distribution function $P_B(r)$ of a stationary rcs X (cf. Definition 1). Since in practice X is only observed within some compact window W of positive volume $|W|$, due to edge effects caused by parts of X outside W, the volume fraction estimator

$$\hat{P}_B(r) = \begin{cases} \frac{|W \cap (X \bullet rB)|}{|W|} & r \geq 0 \\ \frac{|W \cap (X \circ |r|B)|}{|W|} & r < 0 \end{cases} \tag{6}$$

may be biased. To overcome this problem, a minus sampling estimator [23, 29] has been proposed. Briefly, this is just the volume fraction estimator with W replaced by $W \ominus (rB \oplus r\check{B})$. By the local knowledge principle [28], this estimator is pointwise unbiased. However, it is not necessarily monotone in r, nor is all available information used. More refined techniques based on survival analysis ideas have been suggested by Hansen et al. [15] for deriving a Kaplan-Meier type estimator [1] for the empty space function (2). Chiu and Stoyan [8] showed that the ideas underlying the Kaplan-Meier approach are very similar to those involved in the Hanisch estimator [14]. In the remainder of this section, we will derive a Hanisch-type estimator for the size distribution function (Definition 1).

To do so, we need three local size measures: with respect to X, its empty spaces and the boundary. As before, let B be a non-empty convex compact structuring element and X a stationary random closed set observed in a compact observation window W. Set

$$\rho(x,X) = \begin{cases} \sup\{r \geq 0 : \exists h \text{ such that } x \in (rB)_h \subseteq X\} & x \in X \\ 0 & x \notin X \end{cases} \tag{7}$$

$$\eta(x,X) = \begin{cases} \inf\{r \geq 0 : x \in X \bullet rB\} & x \notin X \\ 0 & x \in X \end{cases} \tag{8}$$

$$\zeta(t,W^c) = \begin{cases} \inf\{r \geq 0 : (rB \oplus r\check{B})_t \cap W^c \neq \emptyset\} & t \in W \\ 0 & t \notin W \end{cases} \tag{9}$$

It is easy to see that $X \circ rB = \{x \in X : \rho(x,X) \geq r\}$, $X \bullet rB = \{x \in R^d : \eta(x,X) \leq r\}$ and $W \ominus (rB \oplus r\check{B}) = \{t \in W : \zeta(t,W^c) \geq r\}$.

Discretising over a sampling grid $T = \{t_i\} \subseteq W$, the minus sampling estimator of $P_B(r)$ can be written in terms of ρ, η, ζ as

$$\hat{P}_B(r) = \begin{cases} \frac{\#\{i:\eta(t_i,X) \leq r \leq \zeta(t_i,W^c)\}}{\#\{i:\zeta(t_i,W^c) \geq r\}} & r \geq 0 \\ 1 - \frac{\#\{i:\rho(t_i,X) < |r|;\zeta(t_i,W^c) \geq |r|\}}{\#\{i:\zeta(t_i,W^c) \geq |r|\}} & r < 0 \end{cases}. \tag{10}$$

Note that (10) does not use all information contained in the data. In particular, if $t_i \notin W \ominus (rB \oplus r\check{B})$, but $\eta(t_i,X) \leq \zeta(t_i,W^c)$ the correct void size at t_i is measured. Using this observation, one can define a Hanisch-type estimator for $P_B(r)$ $(r \geq 0)$.

Definition 2 *Let X be a realisation of a stationary random closed set observed in a compact window W. Then for all $r \geq 0$ with $\#\{i : \zeta(t_i,W^c) \geq r\} > 0$, define*

$$\hat{P}_B^H(r) = \sum_{s \leq r} \frac{\#\{i : \eta(t_i,X) = s \leq \zeta(t_i,W^c)\}}{\#\{i : \zeta(t_i,W^c) \geq s\}} \tag{11}$$

and for $r < 0$ with $\#\{i : \zeta(t_i,W^c) \geq |r|\} > 0$, let

$$\hat{P}_B^H(r) = 1 - \sum_{s < |r|} \frac{\#\{i : \rho(t_i,X) = s \leq \zeta(t_i,W^c)\}}{\#\{i : \zeta(t_i,W^c) \geq s\}} \tag{12}$$

The Hanisch-type estimator $\hat{P}_B^H(r)$ is pointwise unbiased for $P_B(r)$. It is increasing and semi-continuous from the right. However, it may be non-negative and exceed 1. If this is undesirable, one can take $R = \sup\{r > 0 : \#\{i : t_i \in W \ominus (rB \oplus r\check{B})\} > 0\}$ and normalise the summands in (11) and (12) by $\hat{P}_B^H(R)$ and $\hat{P}_B^H(-R)$ respectively. The resulting estimator $\tilde{P}_B^H(r)$ is ratio-unbiased. Details can be found in [19], which also provides examples on the use of $\hat{P}_B^H(\cdot)$ in exploratory data analysis.

4. Size-biased Markov random sets

Except from being helpful statistics in exploratory data analysis, size distributions can be used to define new models for random sets observed in a compact window W. Similar ideas for finite random fields have been suggested by Chen and Kelly [6, 7] and Sivakumar and Goutsias [26].

Since the cardinality of the set X is no longer finite, we have to proceed by specifying a density $p(\cdot)$ with respect to some reference process, eg a Boolean model on W. This is a random closed set defined in two steps: first a Poisson process of *germs* is generated (with intensity $\lambda > 0$); then to each of the germs x_i, a random compact *grain* K_i is assigned according to probability distribution $\mu(\cdot)$, independently of other grains. The union $\bigcup_i(x_i \oplus K_i)$ is called a Boolean model. Details on random set densities can be found in [18].

In analogy to [7, 26], set

$$p(X) = \alpha \exp\left[-\int f(s)d\hat{P}_X(s)\right] \tag{13}$$

where $f : R \to R$ is a bounded (measurable) function, $\hat{P}_X(\cdot)$ is an estimator of the size distribution function (5) based on X and α is the normalising constant ensuring that $p(\cdot)$ integrates to 1. It can be shown that for both the naive estimator (6) and the normalised or unnormalised Hanisch-type estimator (Definition 2), (13) is well-defined.

Statistical inference for complex random set models usually relies on iterative procedures making 'local' changes to the random set. First, consider the case were both the germs and the grains of which the random set is composed are fully observable. In that case, $Y = \{(x_i, K_i)\}$ is a germ-grain process and (13) with $X = \cup_i(x_i \oplus K_i)$ is its density with respect to a Poisson marked point process with intensity measure $\lambda\,\mathrm{leb}(\cdot) \times \mu(\cdot)$ (where leb denotes Lebesgue measure, λ is the point intensity and $\mu(\cdot)$ the mark distribution of the grains). Suppose that addition of a grain K at u is considered and the function $f(\cdot)$ is supported on $[-G, G]$. Then the log likelihood ratio depends on X through

$$-\int_{-G}^{G} f(s)d\hat{P}_{X\cup K_u}(s) + \int_{-G}^{G} f(s)d\hat{P}_X(s), \tag{14}$$

writing $K_u = u \oplus K$.

It can be shown that (14) only depends on those $(x_i, K_i) \in Y$ for which

$$(u, K) \sim (x_i, K_i) \Leftrightarrow (K_i)_{x_i} \oplus (GB \oplus G\check{B}) \cap K_u \oplus (GB \oplus G\check{B}) \neq \emptyset.$$

Hence, seen as a grain-marked point process, Y is Markov [24, 3] with respect to the neighbourhood relation $\sim$.

If grains are not individually observable, note that if $X = X_1 \cup \cdots \cup X_k$ is partitioned into its connected components $X_1, \ldots, X_k$, the opening $X \circ B = \cup\{B_h \subseteq X\}$ also partitions, as the convexity of B implies that B_h must fall entirely in one of the X_i. Thus, the naive and Hanisch-type estimators satisfy $\hat{P}_X(s) = \sum_{i=1}^{k} \hat{P}_{X_i}(s)$

for $s < 0$. Similarly, $X \bullet B$ factorises over the connected components of $W \setminus X$ and hence $p(\cdot)$ is of the form

$$p(X) = \prod_{i=1}^{k} \phi(X_i) \prod_{i=1}^{l} \phi(X_i^c).$$

Thus, altering X will only affect the connected components that are modified, a state-dependent Markov property as introduced by Baddeley and Møller [3]. See also [18, 22].

4.1. Morphologically smooth random sets

Let $f(s) = |W| \log \gamma \, \mathbf{1}_{(-1,0]}(s)$. Then, using the volume fraction estimator (6) yields density

$$p(X) = \alpha \gamma^{-|X \setminus (X \circ B)|}, \tag{15}$$

generalising the Chen-Kelly model [7] for binary random fields. Note that for $\gamma > 1$, the most likely realisations X are open with respect to the structuring element B. Thus sets build of approximately convex components are favoured over those with thin or elongated pieces, sharp edges or small isolated clutter.

By duality, taking $f(s) = |W| \log \gamma \, \mathbf{1}_{(0,1]}(s)$ yields

$$p(X) = \alpha \gamma^{-|(X \bullet B) \setminus X|}, \tag{16}$$

favouring for $\gamma > 1$ sets that are approximately closed with respect to B and discouraging small holes or rough edges.

Both models are well-defined for $\gamma \leq 1$ too, for $\gamma < 1$ encouraging morphological roughness.

Note that in (15)–(16), $|\cdot|$ denotes Lebesgue measure restricted to W and hence $p(\cdot)$ is susceptible to edge effects. This can be alleviated by using the Hanisch-type estimator. The resulting model also influences the morphological smoothness of its realisations.

4.2. Morphological area-interaction random sets

Let $f(s) = |W| \log \gamma \, \mathbf{1}\{s \leq -1\}$ and $\hat{P}_X(\cdot)$ given by (6). Then

$$p(X) = \alpha \gamma^{-|X \circ B|},$$

an opening-smoothed version of the area-interaction process in [2, 18]. Similarly, for $f(s) = -\log \gamma \, \mathbf{1}\{s > 1\}$,

$$p(X) = \alpha \gamma^{1 - |X \bullet B| / |W|},$$

a closing-smoothed area-interaction process. Again, Hanisch-type estimators may be employed to better account for edge effects.

Similar ideas may be used if the area measure in the exponent of γ is replaced by the Euler characteristic or other quermass integral [4]. Moreover, since the closing operator removes small holes, the closing-smoothed Euler-interaction process may be integrable when the non-smoothed version is not.

4.3. SIZE–SYMMETRIC RANDOM SETS

Let f be the indicator function of $(-G, G]$, hence

$$p(X) = \alpha \exp[-\int_{-G}^{G} \gamma d\hat{P}_X^H(s)]. \tag{17}$$

For $\gamma > 0$, particle and pore sizes exceeding G will be favoured, while for $\gamma < 0$ the sizes tend to be smaller than G.

Typical samples from (17) using the Metropolis-Hastings sampler of [11, 12] are displayed in Figure 1. The basic idea is that although the normalising constant α in (17) is not available in closed form, the log likelihood ratios (14) do not depend on the normalisation constant and are 'local'. Hence one can run a Markov chain having equilibrium distribution (17) with transition probabilities based on the likelihood ratio for a sufficiently long time.

For ease of computation, we consider a simple square for our structuring element of three by three pixels. Then $\eta(\cdot, X)$, the size measure of voids, can be computed using the distance transform algorithm [5] for the 'square' metric on R^2 defined by

$$d((p_1, p_2), (q_1, q_2)) = \max\{|p_1 - q_1|, |p_2 - q_2|\}.$$

By duality, $\rho(\cdot, X)$ can be computed by reversing the fore- and background. The reference Boolean model has intensity parameter $\lambda = 1000$ and square primary grains with radius r distributed according to a geometric distribution with parameter $\delta = .2$. Finally, $G = 5$ and $|\gamma| = 2500$ and $\hat{P}_X^H$ is as in Definition 2. It is apparent that the scale in the left image ($\gamma = 2500$) is larger than that in the right image ($\gamma = -2500$).

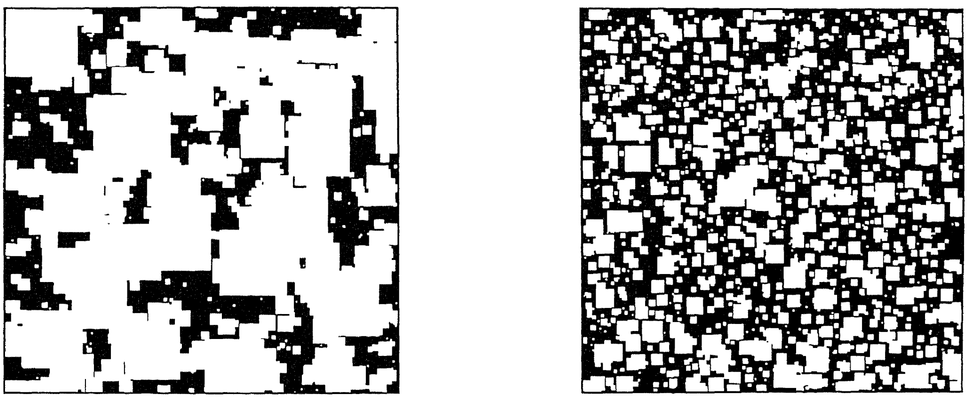

Fig. 1. Samples after 200000 Metropolis-Hastings steps for (17) with $G = 5$ in a 512x512 image for $\gamma = 2500.0$ (left) and $\gamma = -2500.0$ (right).

Exact simulation of size-biased Markov random set models is theoretically possible [17], since its log likelihood ratios are uniformly bounded. However, since (14) is not in general monotone in X (not even for the morphological area-interaction models in Section 4.2), upper and lower bounds based on the current state of the sampler would have to be computed at every iteration. For this reason, we prefer to use the computationally easier Metropolis-Hastings method.

References

1. A.J. Baddeley and R.D Gill. Kaplan-Meier estimators for interpoint distance distributions of spatial point processes. *Annals of Statistics*, 25:263–292, 1997.
2. A.J. Baddeley and M.N.M. van Lieshout. Area-interaction point processes. *Annals of the Institute of Statistical Mathematics*, 47:601–619, 1995.
3. A.J. Baddeley and J. Møller. Nearest-neighbour Markov point processes and random sets. *International Statistical Review*, 57:89–121, 1989.
4. A.J. Baddeley, W.S. Kendall and M.N.M. van Lieshout. Quermass–interaction models. Research report 293, University of Warwick, 1996.
5. G. Borgefors. Distance transformations in digital images. *Computer Vision, Graphics and Image Processing*, 34:344–371, 1986.
6. F. Chen. *Robust and morphologically constrained image segmentation.* PhD thesis, University of Massachusetts at Amherst, 1994.
7. F. Chen and P.A. Kelly. Algorithms for generating and segmenting morphologically smooth binary images. In: *Proceedings of the 26th Conference on Information Sciences*, 1992.
8. S.N. Chiu and D. Stoyan. Estimators of distance distributions for spatial patterns. *Statistica Neerlandica*, 1998. To appear.
9. E.R. Dougherty J.B. Pelz, F. Sand and A. Lent. Morphological image segmentation by local granulometric size distributions. *Journal of Electronic Imaging*, 1:46–60, 1992.
10. A.G. Fabbri. *Image processing of geological data.* Van Nostrand Reinhold, 1984.
11. C.J. Geyer and J. Møller. Simulation procedures and likelihood inference for spatial point processes. *Scandinavian Journal of Statistics*, 21:359–373, 1994.
12. P.J. Green. Reversible jump MCMC computation and Bayesian model determination. *Biometrika*, 82:711-732, 1995.
13. P. Hall. *Introduction to the theory of coverage processes*. John Wiley & Sons, 1988.
14. K.H. Hanisch. Some remarks on estimators of the distribution function of nearest neighbour distance in stationary spatial point processes. *Mathematische Operationsforschung und Statistik, Series Statistik*, 15:409–412, 1984.
15. M.B. Hansen, R.D. Gill and A.J. Baddeley. Kaplan-Meier type estimators for linear contact distributions. *Scandinavian Journal of Statistics*, 23:129–155, 1996.
16. R.M. Haralick, P.J. Katz and E.R. Dougherty. Model based morphology: the opening spectrum. *Graphical Models and Image Processing*, 57:1–12, 1995.
17. W.S. Kendall and J. Møller. Perfect Metropolis-Hastings simulation of locally stable spatial point processes. In preparation.
18. M.N.M. van Lieshout. On likelihoods for Markov random sets and Boolean models. In: *Proceedings of the International Symposium on Advances in Theory and Applications of Random Sets*, D. Jeulin (Ed.), pp. 121–135. World Scientific Publishing, 1997.
19. M.N.M. van Lieshout. Size distributions in stochastic geometry. CWI Research Report, PNA97-15, 1997.
20. P. Maragos. Pattern spectrum and multiscale shape representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11:701–716, 1989.
21. G. Matheron. *Random sets and integral geometry*. John Wiley and Sons, 1975.
22. J. Møller and R. Waagepetersen. Markov connected component fields. Research Report 96-2009, Department of Mathematics and Computer Science, Aalborg University, 1996.
23. B.D. Ripley. *Statistical inference for spatial processes.* Cambridge University Press, 1988.
24. B.D. Ripley and F.P. Kelly. Markov point processes. *Journal of the London Mathematical Society*, 15:188–192, 1977.
25. D. Schonfeld and J. Goutsias. Optimal morphological pattern restoration from noisy binary images. *IEE Transactions on Pattern Analysis and Machine Intelligence,* 13:14–29, 1991.
26. K. Sivakumar and J. Goutsias. Morphologically constrained discrete random sets. In: *Proceedings of the International Symposium on Advances in Theory and Applications of Random Sets*, D. Jeulin (Ed.), pp. 49–66. World Scientific Publishing, 1997.
27. K. Sivakumar. *Morphological analysis of random fields: theory and applications.* PhD thesis, Johns Hopkins University, 1997.
28. J. Serra. *Image analysis and mathematical morphology.* Academic Press, 1982.
29. D. Stoyan, W.S. Kendall, and J. Mecke. *Stochastic geometry and its applications.* Akademie-Verlag, 1987. Second edition, 1995.