

AN APPROXIMATION OF THE \mathcal{K} -FUNCTION FOR THE STUDY OF BINARY IMAGES

GUILLERMO AYALA GALLEGO*
AMELIA SIMÓ VIDAL†

Jensen et al. (1990) gave an exact expression for the \mathcal{K} -function in non-overlapping Boolean models. The present study proposes and evaluates an approximate expression for the \mathcal{K} -function in overlapping isotropic Boolean models based on an approximation of the covariogram of the primary grain. We study the suitability of a Boolean model for two binary images using this approximate expression.

Key words: Boolean model, \mathcal{K} -function, reduced moment measure, image analysis.

1. INTRODUCTION

The Boolean model, a kind of random closed set (Matheron, 1975), has, in recent years, proved to be sufficiently versatile for modelling binary images from very varied sources and is of particular interest for biological and geological images (Stoyan *et al.*, 1987). Its use for modelling three-dimensional objects is also of considerable interest. A recent contribution in this area is Bindrich and Stoyan (1991).

* Guillermo Ayala Gallego. Departamento de Estadística e I.O. e Instituto de Robótica de la Universidad de Valencia. Spain.

† Amelia Simó Vidal. Departamento de Matemáticas. Universidad Jaume I. Castellón. Spain.

-Article rebut el febrer de 1993.

-Acceptat el setembre de 1993.

A **Boolean model** is: Let $\chi = \{x_1, x_2, \dots\}$ be a stationary Poisson point process in \mathbb{R}^d with intensity λ . Let Z_1, Z_2, \dots be a sequence of almost surely compact, convex, independent, random closed sets called grains, identically distributed (as Z_0), distributed independently of χ and satisfying $E\lambda_d(Z_0 + \check{L}) < \infty$ for every compact L in \mathbb{R}^d where $\check{L} = \{-k : k \in L\}$ is the symmetric of L about the origin, $Z_0 + \check{L} = \{z + k : z \in Z_0, k \in \check{L}\}$ the Minkowski addition and λ_d stands for the Lebesgue measure in the Euclidean space \mathbb{R}^d . Then:

$$(1) \quad Z = \bigcup_{n \geq 1} Z_n + z_n$$

is a Boolean model with primary grain Z_0 and intensity λ .

Given a stationary random closed set Φ in \mathbb{R}^d , we can define a random measure which we will call the measure of coverage associated with Φ as: $\nu_\Phi(B) = \lambda_d(\Phi \cap B)$, B being a Borel set in \mathbb{R}^d . In the following, λ_r with $r = 0, \dots, d$ denotes the r -dimensional Hausdorff measure in \mathbb{R}^d . It has been shown that under certain very general conditions ν_Φ , which is obviously determined by Φ , in turn determines the random set Φ , Ayala *et al.* (1991a). In any case, characteristics of ν_Φ provide partial descriptions of Φ which are of practical interest, particularly the \mathcal{K} -function. If $p = P(0 \in \Phi)$ denotes the volume fraction of the random set, then

$$\mathcal{K}(t) = \frac{1}{p} E_0(\nu_\Phi(B(0, t))) = \frac{1}{p} E(\nu_\Phi(B(0, t)) \mid 0 \in \Phi)$$

where E_0 is the expectation with respect to the Palm distribution \mathcal{P}_0 of ν_Φ (Daley and Vere-Jones, 1988) and $B(0, t)$ the ball with the origin as center and with radius t . Intuitively,

$\mathcal{K}(t) = \frac{1}{p}$ (expected Lebesgue measure for the intersection of Φ and the ball with radius t and center at a typical point of Φ).

It can easily be proved (Jensen *et al.*, 1990) that

$$(2) \quad \mathcal{K}(t) = \frac{1}{p^2} \int_{B(0, t)} C(x) dx$$

where $C(x) = \mathcal{P}(0, x \in \Phi)$ is the covariance function of the random set Φ . Assuming an isotropic random set, i.e., distribution invariant under rotations, then the covariance function depends on the modulus of x and (2) is

$$(3) \quad \mathcal{K}(t) = \frac{\Lambda_d}{p^2} \int_0^t r^{d-1} C(r) dr$$

where $\Lambda_d = \lambda_{d-1}(S_0)$, and S_0 is the unit ball surface in \mathbb{R}^d .

An approximation of \mathcal{K} for isotropic Boolean models is proposed in the following section. The third section is devoted to studying its errors. This approximation is used in the fourth section for a goodness-of-fit problem: for two different binary images we test the suitability of the Boolean model.

2. THE APPROXIMATION

For a Boolean model Z in \mathbb{R}^d as described above in (1), it holds (Matheron, 1975) that:

$$(4) \quad C(x) = 2p - 1 + (1 - p)^2 e^{\lambda \gamma_{Z_0}(x)},$$

where $\gamma_{Z_0}(x) = E\lambda_d(Z_0 \cap (Z_0 + x))$, the geometric covariogram of the primary grain Z_0 and $p = 1 - e^{-\lambda V_d}$ with $V_d = E\lambda_d(Z_0)$. Matheron (1975) shows that the function $g(ru) = \lambda_d(L \cap (L + ru))$ with $r > 0$, $u \in S_0$ and L convex and compact, is differentiable from the right at $r = 0$ and its derivative is $\lambda_{d-1}(\Pi_u L)$, where $\Pi_u L$ denotes the projection of L onto the hyperplane whose normal vector is u . Bearing in mind that we assume the primary grain Z_0 to be almost surely convex, compact and isotropic, by applying Cauchy's formula (Santaló, 1976) it follows that:

$$(5) \quad \gamma'_{Z_0}(0) = -\beta_d S_{d-1},$$

where ∂Z_0 denotes the boundary between Z_0 and $S_{d-1} = E\lambda_{d-1}(\partial Z_0)$ and $\beta_d = \frac{1}{2\pi^{1/2}} \frac{\Gamma(d/2)}{\Gamma((d+1)/2)}$. Consequently the first order approximation of the geometric covariogram of the primary grain near $r = 0$ is

$$(6) \quad \gamma_{Z_0}(r) \approx V_d - \beta_d S_{d-1} r.$$

From (4) and (5) it is easy to obtain the first derivative of the covariance function, $C'(0)$.

The approximation which we propose consists in using (3), and performing the Taylor expansion up to the order $d + 1$ of the \mathcal{K} -function. It is verified that this method is equivalent to replacing the covariance function in equation (3) with its development at the origin to order one: $C(r) \simeq p + C'(0)r$. For higher order developments we need to know the second, ... derivatives at the origin of the covariance function whose general expression is unknown. For any random set (not necessarily Boolean) the approximation is as follows:

$$(7) \quad \mathcal{K}(t) \approx \frac{\Lambda_d}{pd} t^d + \frac{\Lambda_d C'(0)}{(d+1)p^2} t^{d+1}.$$

If the random set is a Boolean model then (7) is

$$(8) \quad \mathcal{K}(t) \approx \mathcal{K}_1(t) = \frac{\Lambda_d}{pd} t^d - \lambda \Lambda_d \beta_d S_{d-1} \frac{1-p}{(d+1)p^2} t^{d+1}.$$

For $d = 2$:

$$(9) \quad \mathcal{K}(t) \approx \mathcal{K}_{1,2}(t) = \frac{\pi}{p} t^2 - \frac{2\lambda S_1(1-p)}{3p^2} t^3,$$

S_1 being the mean perimeter of the primary grain.

3. ERRORS

Two different models are considered in order to evaluate the approximation.

1. A 2-D Boolean model: the primary grain is a random disc with uniform radius in $[0, \rho]$.
2. A 3-D Boolean model: the primary grain is a random ball with uniform radius in $[0, \rho]$.

The covariance function is known for both models (Stoyan *et al.*, 1987) and we can evaluate \mathcal{K} using numerical integration. (9) ($d = 2$) and (8) ($d = 3$) provide us with the first approximation for these models. The values chosen for the parameters were $\rho = 0.1$ with three different volume fractions, $p = 0.1, 0.5, 0.9$ (or equivalently, three λ values).

Each volume fraction corresponds to two plots of the same row. In the plot on the left, we have shown the \mathcal{K} -functions (thick line) and \mathcal{K}_1 -functions (fine line) within a row. In the column on the right we have shown $\mathcal{K}/\mathcal{K}_1$. For $r > \rho$, $C(r) = p^2$ as shown Ayala *et al.* (1991b) and so we have evaluated the approximation up to $r = \rho$. Figure 1 corresponds to the 2-D case and Figure 2 to the 3-D one.

Globally, we can say that the approximation proposed, (8), is fairly good, improving, though not greatly, as the volume fraction increases.

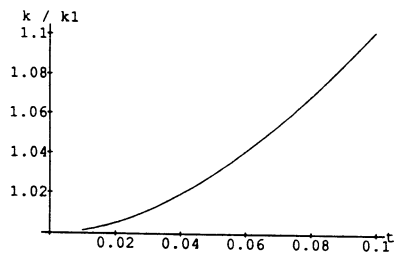
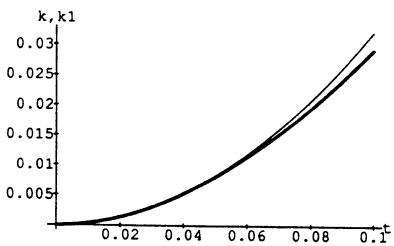
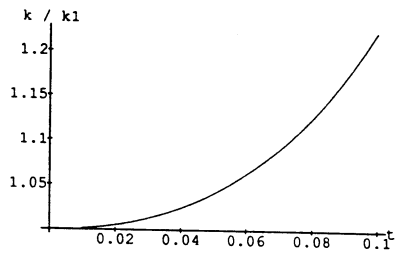
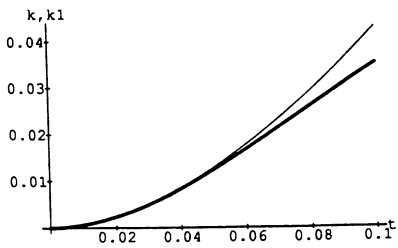
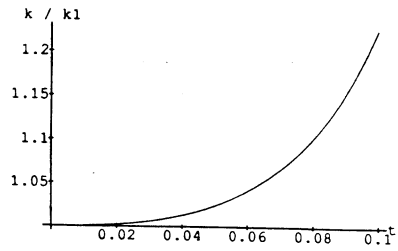
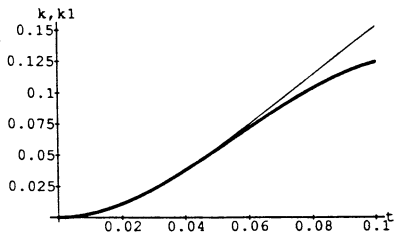


Figure 1

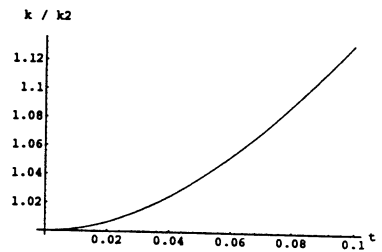
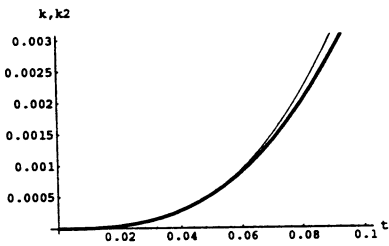
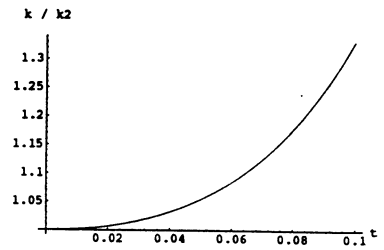
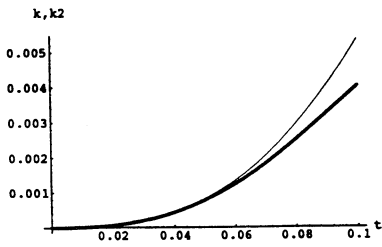
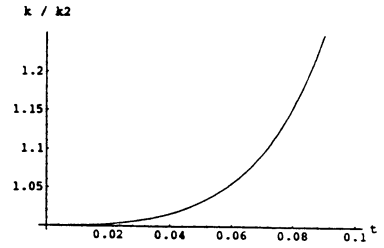
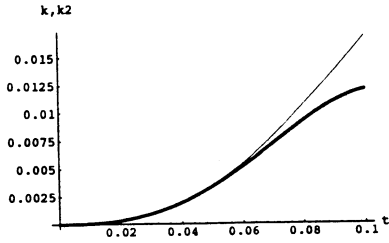


Figure 2

4. SOME APPLICATIONS

We shall analyse two binary images in order to demonstrate the use of the approximation. The first image, Figure 3, represents the distribution of heather in the countryside (it is a centred square from Diggle's image, 1981). The second image, figure 4, is a cross-section of fibres within a nerve fascicle in the sciatic nerve of a male rat (Ruiz, 1986). In order to test the suitability of the Boolean model we use the following method: first, we shall estimate p and S_1 (see (9)) by means of a method not related with the covariance function (we assume a Boolean model.) This provides us with the first approximate version of \mathcal{K} , $\mathcal{K}_{1,2}$. Then using the usual estimator (see Ripley, 1988), we shall estimate the covariance at the points belonging to a grid. In this way, we will have, using (3), a second approximation for \mathcal{K} , $\hat{\mathcal{K}}$. Note that in this second case we are not assuming any hypothesis about the model. We have chosen Kellerer's method (Kellerer, 1983 and 1985) to estimate the parameters in the first case. This method uses only three quantities: area and Euler-Poincaré characteristic of the image and perimeter of the image without its intersection with the edges of the window. From these quantities we can estimate the intensity of the Poisson point process, the mean area and the perimeter of the primary grain. Assuming a unit square as sampling window for both images, the estimated intensity for the heather is 217.36, the mean area is 0.00373 and the mean perimeter is 0.22256. For the nerve fascicle, the values observed were: 272.9242, 0.00138 and 0.13646. From these values we will obtain two versions of $\mathcal{K}_{1,2}$. We have plotted (Figure 5) two curves, one for each binary image, representing $\mathcal{K}_{1,2}$ against $\hat{\mathcal{K}}$, and the line $x = y$. From this figure, we can formulate the following conclusions:

1. Boolean model does not seem suitable in either case. This may, of course, be due to the quality of the approximation, although, as we can see in Figure 5, near the origin there is a clear difference between the estimate based on the real images and the approximations, which argues against this supposition. It may also be due to the estimator used to obtain the approximations. Although the problem has been studied fairly extensively for the point processes, this is not the case for random closed sets. This line of study needs further exploration. (Doguwa and Upton, 1989, and Ripley, 1988).
2. The mean area for the intersection of the heather and a ball with radius t and centre at a typical point of the set is smaller than might be expected for the adjusted Boolean model. The difference decreases with t .
3. The mean area observed for the nerve fascicle is also less than the mean area when we assume a Boolean model, but the opposite is true as t increases.

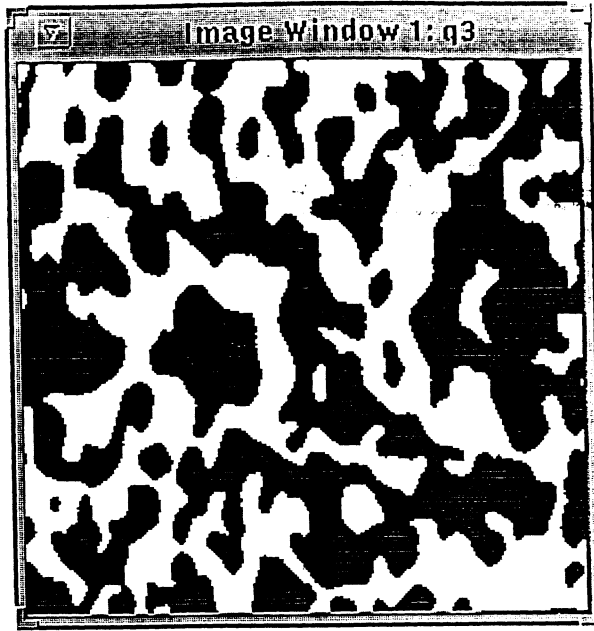


Figure 3

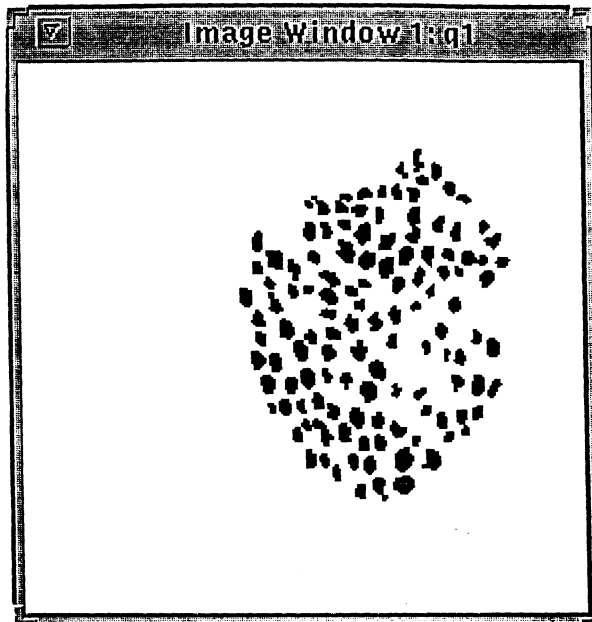


Figure 4

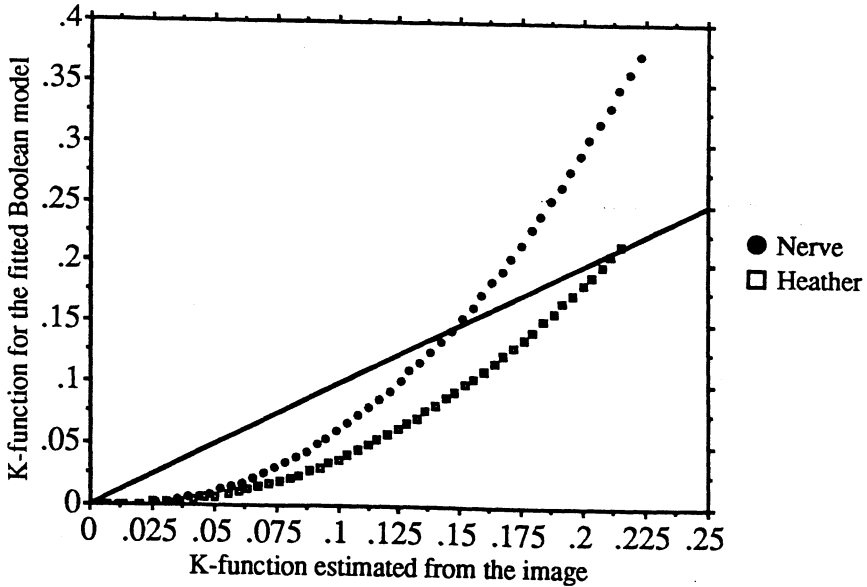


Figure 5

5. FINAL REMARKS

In the foregoing section we have shown an application of the approximation proposed with a view to establishing whether two images can be considered as realizations of an isotropic Boolean model. We believe this might be its main application. If a model of this kind is considered suitable *a priori*, it is clear that our approximations provide fairly good estimators with a much lower computation cost than if we calculated Ripley-type estimators for random sets.

6. ACKNOWLEDGEMENT

For the final version of this paper we are indebted to the interesting and detailed comments of the referee.

REFERENCES

- [1] Ayala, G., J. Ferrandiz and F. Montes (1991a). "Random set and coverage measure". *Adv. Appl. Prob.*, **23**, 972-974.
- [2] Ayala, G., J. Ferrandiz and F. Montes (1991b). "Sparse sampling and maximum likelihood estimation for Boolean models." *Biom. J.*, **33**, 237-245.
- [3] Bindrich, U. and D. Stoyan (1991). "Stereology for pores in wheat bread: statistical analyses for the Boolean model by serial sections". *J. Microsc.*, **162**, 231-239.
- [4] Daley, D.J. and D. Vere-Jones (1988). *An Introduction to the Theory of Point Processes*. Springer-Verlag, New York.
- [5] Diggle, P.J. (1981). "Binary mosaics and the spatial pattern of heather." *Biometrics*, **37**, 531-539.
- [6] Doguwa, S.I. and Upton, G.J.G. (1989). "Edge-corrected estimation for the reduced second moment measures of point processes." *Biom. J.*, **31**, 563-575.
- [7] Jensen, E.B., Kieu, K. and Gundersen, H.J.G. (1990). "On the stereological estimation of reduced moment measures." *Ann. Inst. Statist. Math.*, **42**, 445-461.
- [8] Kellerer, A.M. (1983). "On the number of clumps resulting from the overlap of randomly placed figures in a plane." *JAP*, **20**, 126-135.
- [9] Kellerer, A.M. (1985). "Counting figures in planar random configurations. *JAP*, **22**, 68-81.
- [10] Matheron, G. (1975). *Random sets and Integral Geometry*. Wiley, London.
- [11] Ripley, B. (1988). *Statistical Inference for Spatial Processes*. Cambridge University Press, Cambridge.
- [12] Ruiz, A. (1986). *Estudio de la conducción y morfometría del nervio ciático de la rata albina en un modelo de neuropatía alcohólica experimental*. PhD thesis, Universidad de Valencia.
- [13] Santaló, L.A. (1976). *Integral Geometry and Geometric Probability*. Addison-Wesley, Reading, Mass.
- [14] Stoyan, D., Kendall, W.S. and J. Mecke (1987). *Stochastic Geometry and its applications*. Wiley, Berlin.