The size of objects in natural images

Luis Alvarez[1], Yann Gousseau[2], Jean-Michel Morel[2]

[1] Departamento de Informatica y Sistemas, Universidad de Las Palmas, Campus de Tafira, 35017 Las Palmas, Spain.
[2] Centre de Mathématiques et leurs Applications, ENS Cachan, 61 av. du Président Wilson, 94235 Cachan Cedex, France.

Abstract

This paper introduces a new method for analyzing scaling phenomena in natural images, and draws some consequences as to whether natural images belong to the space of functions with bounded variation. In some sense, our analysis computes the size distribution of objects in an image. By using the dead leaves model, we study the influence of occlusion on size distribution, and prove compatibility with our experimental results.

Contents

1 Introduction

There are many statistical studies of natural images, that is to say of pictures of scenes encountered in the surrounding world. Most of the studies are concerned with first or second order statistics (through the power spectrum, the covariances, the cooccurrences) or with additive decompositions of images. Those statistical studies are partially matched by synthesis models, especially for texture synthesis, which try to compose an image either by imposing some second order statistics, or by adding (wavelet-like) basis functions while maintaining some statistical requirements, as detailed in [1].

Our aim in this paper is first to give an analytical account of those analysis methods, in Section 2. Then, we will try to complete the frame by introducing a new way of analyzing scaling phenomena in an image, in Section 3. We shall show that in natural images, there is a constant form for the size distribution. The definitions of sizes we consider are of several types: area, boundary length, and length of intersections with lines. An experimental program which we performed on many photographs of very diverse natural scenes indicates that the size distribution of homogeneous parts in images has a very peculiar form. We will define in Section 3 what we mean by homogeneous parts, the connected components of image domains where contrast does not exceed a certain threshold. We can roughly define the size distribution of objects in an image as the size distribution of these parts. Our experiments show that the number of homogeneous parts as a function of their size s obeys a law

$$\text{Card}\{\text{Homogeneous regions with size s}\} = \frac{K}{s^\alpha},$$

where K is an image dependent constant. When the size s denotes the area, in most photographs, α is close to 2. As a consequence of the size power law, some information can be obtained about the "natural" function space for images, as will be shown in Section 4: we focus our attention on the space BV of functions with bounded variation. We are in a position to tell when a given image is not in this space, provided the observed size distribution model remains true at smaller (not observable) scales as well. In Section 5, we present a powerful but scarcely used model, the dead leaves model of Matheron, and analyze the effect of occlusion on the size distribution of objects. There is no obvious equivalence between the way we define regions in our experimental program and regions in the dead leaves model. However, we show that, for some notions of size, the power law is preserved by occlusions in the dead leaves model.

2 Statistics of natural images: a review

2.1 Motivations

A digital, gray level image may be seen as the realization of a random vector of size $M \times N$ taking values in a discrete set $V = 1, ..., C$. For typical values like $M = N = C = 256$, the number of possible realizations, $C^{MN} = 2^{524288}$, is huge. Obviously, “natural images”, i.e. digital photographs of natural scenes, only form a small subset of all possible realizations. Looking at random realizations of such vectors is enough to be convinced of this fact. Natural images are highly improbable events. It is therefore interesting to look for statistical characteristics of such images: what are the relationships between gray level values at distant pixels? Is it possible to define a probability law for natural images?

In what follows, I is a digital image, for $x \in M \times N$, $I(x) \in V$ is the gray level value at pixel x, $Pr(A)$ stands for the probability of the event A, and E is the mathematical expectation.

In addition to the legitimate interest in the complicated and fairly unknown structure of images, there are some strong motivations to investigate such statistics. The first one comes from biological vision, and has initiated most of the work on the field. Indeed, it is likely that visual systems are adapted to the type of scenes they deal with (which may be very different from one specie to another). H. Barlow ([3]) suggested that the goal of sensory coding is to reduce redundancies in the input signals, so that studying those redundancies (for example dependency between adjacent pixels) give an insight to the way our brain is extracting informations from the huge amount of data it receives through the eyes. A second motivation arises from the Bayesian framework. Suppose we want to describe an image by a set of parameters $\{p_1, ..., p_n\}$. In order to estimate these parameters, given the image I, we make use of Baye's rule

$$Pr(\{p_i\}|I) = \frac{Pr(I|\{p_i\})Pr(\{p_i\})}{Pr(I)},$$

formula which indeed presupposes the existence of a probability distribution Pr on images. This can for instance be useful for the denoising (through the search of a good approximation of the “true” image knowing a noisy, and possibly convolved image) if we know characteristics of Pr. A third motivation is image compression. Imagine a very simple type of images where each realization is either black or white, then we'll only need one bit to encode any image. More generally, information theory tells us that the compression rate is related to the entropy of the image I,

$$E(I) = -\sum_{all I} Pr(I) \log Pr(I).$$

Shannon's source coding theorem asserts that in order to compress (many) images with almost no loss of information, the number of bits needed is almost equal to the entropy. Moreover trying to improve this bound leads to a drastic loss of information. Codes to achieve such a compression are given by information theory, but are not feasible in practice, and the measure of loss may be poorly adapted to images. However, statistics of images may give an insight to the type of compression schemes that are to be used. We also show in the present paper (see Section 4) that statistics of images may account for the type of functional space images belong to.

The study of the statistics of images is not a new subject. It started with television signals studies ([18], [10]). The subject has gained considerable interest during the last ten years. In all studies, images are assumed to be translation invariant, allowing the relationships between pixels to depend on distance only, even though this assumption is not strictly observed.

When dealing with statistics of natural images, two different approaches are possible: either assuming that a single image is large enough to evaluate parameters on it, by making ergodicity assumptions, or averaging the quantity under study over a large number of images. The first method enables us to study differences between images and may render easier a physical explanation, while the second is more robust and permits the estimation of outliers, which are not numerous enough in a single image.

In order to study natural scenes, one may be interested in studying the light intensity arriving on the lens, and thus in having a linear relationship between gray level and intensity. Moreover one may want to have an histogram centered at 0 and invariant under multiplication of the gray levels by a constant. One way to achieve that is to study $\log(\frac{I}{I_0})$ instead of I, where I_0 is the mean gray level.

2.2 First order statistics

The simplest thing we can study from an image is the gray level histogram, that is to say the distribution of gray levels in any pixel (under translation invariance each pixel is equivalent). Several authors ([29], [14]) found linear tails when plotting log(number of occurrences of gray level) versus log(gray level). They worked on images of logarithm of initial gray level values for the previously mentioned reasons. Those linear tails indicate that the histogram is not Gaussian. Moreover, the distribution is not symmetric, which could be due to the presence of large parts of sky in many images, as suggested in [14].

However, looking at the histogram could be interesting for the purpose of biological studies when dealing with calibrated images, but is not of much use in order to study the morphological structure of images. It is well known that we can strongly modify the histogram of an image without affecting its perception.

2.3 Second order statistics

2.3.1 Covariance and power spectrum

In this section, I is assumed to be a continuous function from $\mathbb{R}^2$ into $\mathbb{R}$. Under the assumption of translation invariance, the covariance of an image I is defined as (assuming the mean is zero)

$$C(x) = E(I(0)I(x)),$$

where 0 is an arbitrary origin and x a location in the image. If $\hat{I}$ is the Fourier transform of the image (I is viewed as a continuous function belonging to L^2

and $\hat{I}$ is a function of the complex variable ω), the power spectrum is defined as

$$P(\omega) = |\hat{I}(\omega)|^2,$$

and we have the Wiener-Khintchin formula

$$P(\omega) = \hat{C}(\omega).$$

It is well known that, when averaged over all directions, the power spectrum of natural images,

$$f(t) = \frac{1}{2\pi}\int_0^{2\pi} P(te^{i\theta})d\theta,$$

has been experimentally found to have the form

$$f(t) \simeq \frac{C}{t^\alpha},$$

with α close to 2, ([10], [12], [5]). This phenomenon is also observed with the logarithm of images ([29]). It should be mentioned that the form of $P(\omega)$ is strongly dependent on the direction ([2]), and that there are large differences in the exponent α of the power law from one image to another ([34]). Moreover, the Fourier transform is a nonlocal operation, and consequently the power spectrum may emphasize a particular aspect of the image, as the presence of an edge. In the case of the presence of a single very sharp edge, for instance at the horizon, the exponent of the power law gives no account of other statistics of the image, see [36].

A very interesting fact about this power law is that it is implied by scale invariance (as stressed by Field, [12]). Indeed, if there is the same energy for frequencies ω between t and t', $t < |\omega| < t'$, and between at and at', $at < |\omega| < at'$, then

$$\int_t^{t'} uf(u) = \int_{at}^{at'} uf(u).$$

But this must be true for all t and t' so that $f(au) = \frac{1}{a^2}f(u)$ for all u, a, so that $f(u) = \frac{C}{u^2}$. It is worth noticing that the same reasoning about sounds leads to $P(t) = \frac{C}{t}$, a property of the so-called $\frac{1}{f}$-noise, which is also observed, see [35].

Assuming this power law for the power spectrum, and because of the Wiener-Khintchine relation, the correlation should be of the form

$$C(x) = A + B\log(|x|).$$

It is worth mentioning that if $P(\omega)$ is a power law with an exponent $\alpha \neq 2$, then the covariance should be of the form

$$C(x) = A + B(|x|^{2-\alpha}),$$

with the signs of A and B ensuring its decay when x increases. Those types of covariances have been observed, at least asymptotically ([27], [28], [2], [14]).

2.3.2 Other second order statistics

The covariances of an image do not characterize the complete second order statistics. Those statistics are fully described by the cooccurrence function:

$$K(i,j,x) = Pr(I(0) = i \;\&\; I(x) = j)$$

for x a site in the image, and i, j gray levels.

Huang and Mumford [14] study the logarithm of the difference of gray levels between two adjacent pixels, working with the logarithm of the images. The observed distribution has a very peculiar form, notably with a high peack in 0 and long tails. Once more this emphasize the fact that images are nonGaussian. The large tails of the distribution denote the presence of edges in the images.

2.4 Linear decomposition of images

Other types of studies deal with the decomposition of images into a linear combination of basis images. An image I is considered as a random vector of size $M \times N$ and is written

$$I = \sum_1^n a_n A_n$$

where $\{a_i\}$ are random variables, and $\{A_i\}$ fixed images. The A_i's are determined by imposing some conditions on the a_i's, and are then used as a statistical description of the images.

In the Principal Components Analysis, the a_i's are requested to be pairwise uncorrelated. The A_i's are obtained by looking for eigenvalues of the covariance matrix. For natural images, they are non-local and resemble a Fourier basis, [26], [4].

In the Independent Components Analysis, the a_i's are requested to be pairwise independent. Such a decomposition may not exist, and the A_i's are approximated by numerical schemes. In [4] the authors used a neural network approach to compute the independent components of natural images (actually on small 12×12 patches extracted from 4 natural images) and obtained vectors which are localized and oriented, and resemble edge detector filters. Those vectors also resemble receptive fields of simple cells in the visual cortex, confirming the hypothesis of Barlow mentioned in the introduction. Another decomposition is studied in [26]. The a_i's should minimize

$$E = \left(I - \sum a_i A_i\right)^2 + \lambda \sum S(\frac{a_i}{\sigma})$$

with $\sigma_i^2 = E(a_i^2)$, λ a constant, and S some increasing real function. The idea is to find A_i's which cost as little as possible with respect to the cost function S. As in the case of independent components analysis the A_i's are localized and oriented.

2.5 Scale invariance in natural images

We do not attempt to give a precise definition of scale invariance, but we call so the property of images to have a zoom-invariant probability law. Scale invariance has been observed in several ways in natural images. First of all, the simple fact of being able to observe many common statistics on images representing arbitrarily chosen scenes (thus seen from arbitrary distance and with any size of objects) is a strong clue to scale invariance. Secondly, as we previously mentioned, the observed form of the power spectrum is the only possible one for scale invariant images. Moreover, many statistics have been found to be preserved when block-averaging images. This is the case of the histogram and covariances ([29], [27]), and also the case for the gray value differences between adjacent pixels ([25]). It must be emphasized that this is the case even though the procedures to obtain smaller images where very brutal: small filtering followed by subsampling. Actually, images are not exactly scale invariant. We already mentioned that the exponent of the power law decay of the power spectrum is not always 2. Moreover, when block-averaging the images, statistics are preserved by slightly renormalizing the image, that is to say by multiplying it by a constant at each blow down, which is a way to measure the departure of images from scale invariance, see [25]. Our results about areas of sections also confirm this scale invariance, as will be heuristically explained in Section 3.

3 Sizes of sections in natural images

As reviewed in the previous section, the known statistical properties of natural images are mainly first or second order statistics, and statistics of additive decompositions of images.

We have a different approach, working in the image domain on items that can have a straightforward visual interpretation. The goal of our experiments is to study the areas of homogeneous regions in digital images. We shall look at the area distribution of the connected components of bilevels of various digital photographs, and realize that those distributions are very similar one another. Indeed, the observed values are very close to functions $\frac{A}{area^\alpha}$, where α is a real number around two, and A an image dependent constant. That is to say :

$$\text{Card}\{\text{Homogeneous regions with area } a\} = \frac{A}{a^\alpha},$$

We shall also study the distribution of the lengths of the boundaries of those connected components of bilevel sets, and also observe a power law, with a different exponent. Eventually, the same result happens to be true for the distribution of the lengths of the intersections of those components with lines of fixed direction. From the first two of those statistics, we are in a position to tell when images are not in the space of functions with bounded variation, BV, as will be shown in Section 4.

3.1 The distribution of areas

We'll now make clearer what we mean by homogeneous region of an image. We begin by equalizing the image histogram, and uniformly quantify it in the following way. We consider a digital image I of size $H \times L$, with G integer gray levels, and write $I(i,j)$ for the gray level at pixel (i,j). Let k be an integer less than G. Let N_1 be the first integer such that more than $\frac{HL}{k}$ pixels have a gray level less than N_1, then N_2 the first integer such that more than $2\frac{HL}{k}$ have a gray level less than N_2, then $N_3, ..., N_k = G$ defined the same way, this sequence being possibly constant at some point. For l varying from 1 to k, let I_l be the binary image with $I_l(i,j) = 1$ if $I(i,j) \in [N_{l-1}, N_l)$ and $I_l(i,j) = 0$ otherwise. We call those images k-bilevels of I. Each bilevel image represents a quantization level of the equalized image.

Next, we look at the area histogram of the connected components of the bilevels. For s an integer varying from 0 to HL, let $f(s)$ be the number of connected components with area s of the set of 1's pixels, in any of the k-bilevels of I. We will both consider 4-connexity (each pixel has 4 neighbors: up, down, right, left) and 8-connexity (we add the diagonal neighbors, so that each pixel has 8 neighbors).

We computed the function f on many digital photographs. We did not attempt to use a single source of images; the digitized images either are scanned photographs or from a digital camera, with diverse optical systems and exposures. Those functions are of the form $f(s) = \frac{C}{s^\alpha}$, with C a constant and α a real number close to two, for values of s in a certain range and reasonable values of k (basically between 4 and 30). The observed fit is excellent, as can be seen from Figure 2, which actually corresponds to one of the worst cases we observed.

For fixed k, we consider the set of points

$$S = \{(\log(s), \log(f(s)), 0 \leq s \leq T_{\max}\},$$

where $T_{max}+1$ is the smallest value of s such that $f(s) = 0$. We perform a linear regression on this set S so as to find the straight line (in the log-log coordinates) $g(\log(i)) = A - \alpha \log(i)$ the closest to S in the least squares sense, and write E for the least squares error. The choice to consider only the data $s \leq T_{max}$ is motivated by the fact that f takes mainly 0 and 1 values at some point, which makes difficult the fit of f to a reference function. Thus we minimize the energy

$$E(\alpha, A) = \sum_{i=1}^{T_{\max}} (\log f(i) - A + \alpha \log i)^2 .$$

3.2 Digital photographs

We present the result for pictures having different scales and textures in Tables 1-4 . The two texture images in Figure 5 are from the Brodatz's album. The value of α seems to be related to the amount of texture in the image; the more

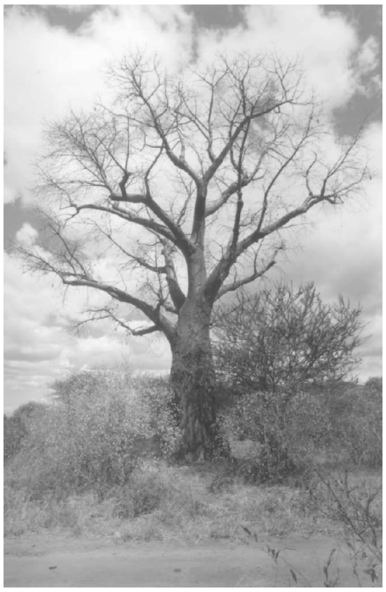

Figure 1: baobab, 665 × 1024, 119 gray levels

textured the image, the bigger the value of α. Typically, for photographs of natural scenes, the value of alpha is between 1.5 and 3 (the values close to 3 being reached for images as the baboon (Figure 3), which present textured areas), whereas for textures it is typically between 2.5 and 3.5.

We also performed the linear regression on sets of points

$$S_{Tmin} = \{(\log(s), \log(f(s)), T_{min} \leq s \leq T_{max}\}$$

for various values of $T_{\min}$ to show that the fit of S to the power law was not forced by small areas only, and moreover that if the contribution of E mainly comes from the large areas, the value of α computed with those large areas was close to the initial value. The results for the image of the baobab are shown in Table 2.

3.3 Other types of images

In order to see whether the power law is in some sense characterizing digital photographs, we computed histograms of areas of bilevels for other types of images. We looked at noise images, white or correlated; text images; and synthesized images consisting of the superposition of simple shapes.

White noise images, that is to say images in which the gray level values at distinct pixels are independent random variables, present a histogram of the form $f(s) = \exp(-Cs)$, with C a constant. We observed this fact on two

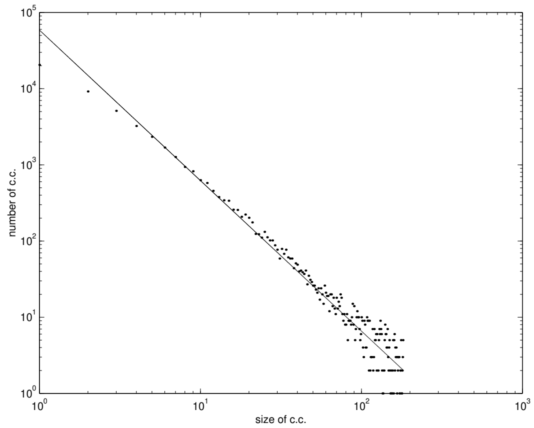

Figure 2: function f (area distribution) for the baobab image (Figure 1), $k = 16$, 8-connexity, $T_{max} = 182$

image	k	α	E	T_{max}	A
baobab	20	1.98	.24	109	11.2
baobab	16	1.98	.45	182	11.0
baobab	12	1.85	.31	111	10.3
baobab	8	1.67	.27	82	9.3

image	k	α	E	T_{max}	A
baobab	20	2.15	.32	146	11.7
baobab	16	2.06	.38	160	11.4
baobab	12	1.96	.30	106	10.8
baobab	8	1.86	.40	133	10.1

Table 1: different values of the quantization number k for the baobab image, left 8-connexity, right 4-connexity. Area distribution is $f(s) = As^{-\alpha}$, T_{max} is the maximal considered area

image	T_{min}	α	E	A
baobab	5	2.04	.44	11.3
baobab	10	2.07	.44	11.4
baobab	20	2.09	.46	11.5
baobab	50	2.08	.50	11.5

Table 2: different values of T_{min} for the baobab image, $k = 16$, 8-connexity

image	k	α	E	T_{max}	A
airport	20	1.95	.43	92	9.7
airport	16	1.93	.47	92	9.6
airport	12	1.79	.38	95	9.0
airport	8	1.83	.35	63	8.7

image	k	α	E	T_{max}	A
baboon	20	2.55	.30	70	11.7
baboon	16	2.38	.33	82	11.3
baboon	12	2.42	.47	78	11.4
baboon	8	2.35	.41	76	11.2

Table 3: different values of k for the baboon and airport images (Figure 3), 8-connexity

Figure 3: baboon and airport images, respectively 512×512 and 510×341

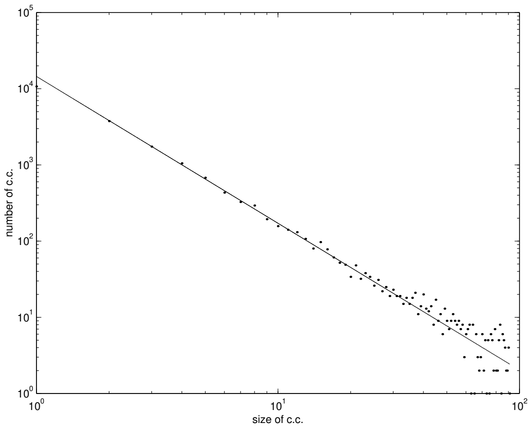

Figure 4: function f for the airport image (Figure 3), $k = 16$, 8-connexity, $T_{max} = 92$

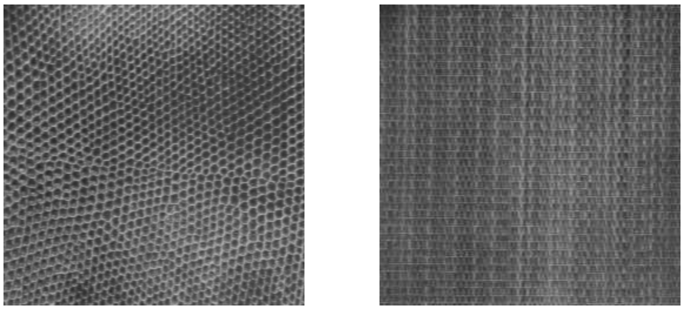

Figure 5: skin and fabric images, 256×256

image	k	α	E	T_{max}	A
skin	20	3.59	.66	19	11.5
skin	16	3.56	.66	26	11.7
skin	12	3.34	.67	31	11.6
skin	8	2.44	.51	25	10.4

image	k	α	E	T_{max}	A
fabric	20	3.42	.43	28	11.3
fabric	16	3.17	.45	29	11.1
fabric	12	2.91	.34	34	10.9
fabric	8	2.61	.49	38	10.7

Table 4: different values of k for the skin and fabric images (Figure 5), 8-connexity

image	α	E	T_{max}	A
uniform noise	5.00	.93	11	13.5
Gaussian noise	4.70	.88	11	13.4

Table 5: $k = 16$ for white noises, 8-connexity

different kinds of white noises: uniform and Gaussian. Thus the value of E is large for those images, and so is α, as presented in Table 5.

Text images produced by text editor do lead, as one would guess, to an histogram consisting of isolated peacks, whose height is not directly related to the value of s.

Concerning the synthesized images, we observed two distinct behaviors. First, if we consider superposition of simple shapes (disks, rectangles, polygons) with various laws of sizes, and possibly made of subobjects, the function $f(s)$ is often well fitted by a power law, but with an α close to one. This is quite difficult to predict by calculations, and we refer to the section presenting the dead leaves model for the mathematical analysis of those phenomena. We present two results. In the first one we considered 50 images produced by the superposition of disks of the same size. Each disk has a random color uniformly drawn between 0 and 255, and its center follows a uniform law in the image (actually in a slightly enlarged image). In the section about the dead leaves model we used the same image to compute average sizes, but with a different analysis to avoid the union of objects and boundary problems. The values of $f(s)$ for each s were then added over all images, and the function f fitted by a power law as before. We then performed the same analysis with disks having an area distributed according to a power law with exponent 1, and a radius between 20 and 80. See Table 6.

Furthermore, if we do the same with more complicated figures, having holes and asperities, the observed law is a power law as well, but with an α close to two. In the example presented, see Figure 6, the basic shape is itself the connected component of the bilevel of a photograph, actually a part of a fly eye. It must be emphasized that discretization effects are very strong in these types of complicated shapes, but that if we perform a morphological closing of size 1 pixel on the basic shapes, the results are not qualitatively affected, which tends to minimize the effect of discretization.

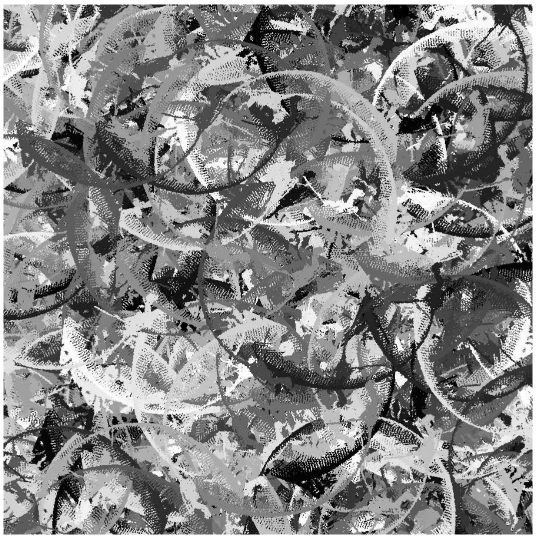

Figure 6: dead leaves model, where all shapes have the same size, 900 $\times$ 900, 256 gray levels

image	α	E	T_{max}	A
disks	0.86	0.29	1000	9.0
disks, power law	0.80	0.39	807	6.9
eyes (Figure 6)	1.66	0.31	686	12.4

Table 6: various synthesized images, 8-connexity, k=16

image convolved with a Gaussian	σ	α	E	T_{max}	A
noise convolved with a Gaussian	0.71	3.88	0.60	26	13.3
noise convolved with a Gaussian	0.85	3.57	0.50	30	13.1
noise convolved with a Gaussian	1.12	3.27	0.55	39	13.0
noise convolved with a Gaussian	1.58	2.87	0.47	59	12.6

Table 7: Uniform white noise image, after convolution with a Gaussian of parameter σ, k=12, 8-connexity

Last but not least we present the results obtained when analyzing a correlated noise. We performed convolutions between white noises and a Gaussian $\exp(-\frac{1}{2\sigma^2}(x^2+y^2))$, where σ is a variable parameter. This was done by multiplication in the frequency domain. Such a convolution can be seen as a crude approximation of the effect of an optical lens. The results we obtain for those images were similar to the ones for digital photographs of textures. We present the results obtained in the case of the white uniform noise in Table 7. We also tested the effect of the convolution with Bessel functions (Fourier transform of disks) and the results were very similar.

Those "non natural images" lead to two remarks about the $1/area^2$ law. First, this law does not characterize natural images, even though a correlated noise looks similar to a natural texture, and the complicated synthesized images may look familiar to us.. Secondly, the size law could be related to the way the optical photographic device captures the image, as suggested by the behavior of noises convolved with a Gaussian. More precisely, we observed that the convolution with a Gaussian increases the value of α for images where the initial α is small (such as text and synthetic images) whereas it tends to decrease its value when it is initially bigger than 2 (noises).

Another, and more satisfactory explanation of this power law relies on the assumption that natural images are scale invariant so that all observed statistics should be scale (zoom) invariant. This simple assumption yields the $1/area^2$ law. Indeed, if we suppose that the total area occupied by regions having an area between A and A' is the same as the total area occupied by regions with area between tA and tA', for all t, A, A', then the same reasoning as presented for the power spectrum (see Section 2.3.1) leads to the power law with exponent 2.

3.4 The distribution of boundary lengths

We performed exactly the same analysis on the boundary lengths of connected components of bilevels as we did before on areas of those components. As a discrete definition of the length of a discrete connected set S (8-connexity), we chose to count the pixels not belonging to S that are neighbors of some pixel of S in the 4-connexity sense. There are many other ways to define discrete boundary length. We tried several other methods that gave basically the same

image	k	β	E	T_{max}	A
baobab	20	2.39	.34	176	13.2
baobab	16	2.37	.35	154	13.0
baobab	12	2.26	.25	125	12.3
baobab	8	2.23	.35	111	11.6

image	k	β	E	T_{max}	A
baboon	20	3.02	.28	82	14.2
baboon	16	2.91	.35	81	13.9
baboon	12	2.93	.41	100	14.0
baboon	8	2.89	.33	96	13.8

Table 8: boundary lengths for the baobab and baboon images, with different values of the quantization number k

image	k	β	E	T_{max}	A
baobab	20	2.72	.42	88	13.3
baobab	16	2.60	.37	77	13.0
baobab	12	2.55	.47	118	13.1
baobab	8	2.33	.44	124	12.6

image	k	β	E	T_{max}	A
baboon	20	3.79	.47	24	12.9
baboon	16	3.58	.35	27	12.8
baboon	12	3.37	.41	29	12.8
baboon	8	3.23	.58	44	13.0

Table 9: intercept lengths for the baobab and baboon images, with different values of the quantization number k

results as the one we detail here. The notations k, E, A, T_{min} and T_{max} refer to the same quantities as before; β now stands for the exponent of the power law whose fit to the boundary length distribution is the best in the least square sense. We chose $T_{min} = 10$, because some small values for the boundary lengths are attained only for regions touching the border of the image. The fit to the power law is again very good, and the exponent β is usually between 2 and 3. We present the results for the two images of the baobab and the baboon in Table 8. We note that $\beta \simeq 2\alpha - 1$ accounts for connected components of bilevel sets satisfying on the average a decent isoperimetric ratio, $c \leq \frac{area^{\frac{1}{2}}}{boundary\ length} \leq C$. This is not the case in general, except for some images of textures.

3.5 The distribution of intercept lengths

We also studied the distribution of the lengths of the intersections of the bilevel sets with lines of fixed direction. This was motivated by the fact that this quantity, in an object-based model of images, is easily amenable to calculations, see Section 5.4. We once more observed a power law, which, according to the previously mentioned model, is compatible with a power law hypothesis for the size distribution of objects, even if the model is not isotropic, See Section 5.5. However, it should be mentioned that the result is strongly varying with the direction. The experiments we present here are concerned with the horizontal direction, see Table 9. The observed values for the exponent of the power law are typically between 2 and 3.5.

4 Size of sections and the BV-norm of natural images

The aim of this section is to give a computational tool to decide whether an image can belong to the space BV of functions with bounded variations. The BV assumption for natural images is far ranging, from image restoration ([30], [31]) to image compression.

4.1 A lower bound for the BV norm

We consider $I \in BV(\Omega)$ a bounded image belonging to the space of functions with bounded variation ([37], [11]) on a domain (e.g. rectangular) $\Omega \subset I\!R^2$. For $\lambda \in I\!R$, define the level set of I with level λ by

$$\chi_\lambda I = \{x, I(x) \geq \lambda\}.$$

Recall that a function is of bounded variation if, for almost every $\lambda \in I\!R$, $\chi_\lambda I$ is a set with finite perimeter and, denoting by $\mathrm{per}(\chi_\lambda I)$ this perimeter (for a precise definition of the perimeter and the essential boundary we refer to [11]),

$$||I||_{BV} = \int_{I\!R} \mathrm{per}(\chi_\lambda I) d\lambda. \tag{1}$$

(By the coarea formula, see [11], we also have $||I||_{BV} = \int_\Omega |DI|$)

In addition, by the classical isoperimetric inequality, we have for every set O with finite perimeter,

$$\mathrm{per}(O) \geq 2\pi^{\frac{1}{2}} \nu(O)^{\frac{1}{2}}, \tag{2}$$

where $\nu(O)$ denotes the Lebesgue measure of O. In the following, we shall consider sections of the image. We always assume that the image I satisfies $0 \leq I(x) \leq C$. We first fix two parameters γ, λ, with $0 \leq \lambda \leq \gamma$. For any $n \in I\!N$, we consider the bilevel sets of I

$$\{x, \lambda + (n-1)\gamma \leq I(x) < \lambda + n\gamma\} = \chi_{\lambda+(n-1)\gamma} I \setminus \chi_{\lambda+n\gamma} I.$$

We call (γ, λ)-section of I any set which is a connected component of a bilevel set $\chi_{\lambda+(n-1)\gamma} I \setminus \chi_{\lambda+n\gamma} I$ for some n. We denote each one of them by $S_{\gamma,\lambda,i}$ for $i \in J(\gamma, \lambda)$, a set of indices. Notice that the (γ, λ)-sections are disjoint and their union is the image domain Ω,

$$\bigcup_{i \in J(\gamma,\lambda)} S_{\gamma,\lambda,i} = \Omega. \tag{3}$$

There are several ways to define the connected components of a set with finite perimeter, since such a set is defined up to a set with zero Lebesgue measure. We denote by H^1 the one-dimensional Hausdorff measure, that is to say the length. In the following, we call Jordan curve a simple closed curve of $\mathbb{R}^2$, i.e. the range of a continuous map $c : [0,1] \to \mathbb{R}^2$, such that $c(s) \neq c(t)$ for all $0 < s < t < 1$, and $c(0) = c(1)$. A Jordan curve defines two and only two connected components (in the usual sense) of $\mathbb{R}^2 \setminus c([0,1])$, one bounded and one unbounded. We shall say that a Jordan curve separates two points x and y if they do not belong to the same connected component of $\mathbb{R}^2 \setminus c([0,1])$. One can prove ([13], [6]) that a definition of connected components for a set with finite perimeter permits the following statements :

Proposition 1 *(and definition)*
Let O be a set with finite perimeter.
(i) The essential boundary of O consists, up to a set of zero H^1-measure, of a countable set of noncrossing simple rectifiable closed curves c_j with finite length such that $per(O) = \sum_j H^1(c_j)$
(ii) Two points are in the same connected component of O if and only if for any representation of the essential boundary by a family of Jordan curves of the preceding kind, c_j, they are not separated by one of the c_j.
(iii) With this definition, the perimeter of a set with finite perimeter is the sum of the perimeters of its connected components.

We denote by $J(n) \subset J(\gamma, \lambda)$ the set of indices of sections which are connected components of $\chi_{\lambda+(n-1)\gamma} I \setminus \chi_{\lambda+n\gamma} I$. As an obvious consequence of Proposition 1, we have

Corollary 1

$$per(\chi_{\lambda+(n-1)\gamma} I \setminus \chi_{\lambda+n\gamma} I) = \sum_{i \in J(n)} per(S_{\lambda,\gamma,i}).$$

When A is a set with finite perimeter, we have ([11])

$$\mathrm{per}(A) = ||\mathbb{1}_A||_{BV}.$$

Lemma 1 *If $B \subset A$ are two sets with finite perimeter, then*

$$per(A \setminus B) \leq per(A) + per(B).$$

Proof Indeed, by the subadditivity of the BV norm, we deduce from

$$\mathbb{1}_{A \setminus B} = \mathbb{1}_A - \mathbb{1}_B$$

that

$$\mathrm{per}(A \setminus B) \leq \mathrm{per}(A) + \mathrm{per}(B).$$

□

In the following theorem, we analyze the statistics of sizes of sections as follows. We fix γ, that is, the overall contrast of considered sections and for each $0 \leq \lambda \leq \gamma$, we count all sections $S_{\gamma,\lambda,i}$ which have an area between s and $s+ds$. In other terms we consider the integer

$$\mathrm{Card}\{i, s \leq |S_{\gamma,\lambda,i}| \leq s+ds\}.$$

We average this number over all λ's in $[0,\gamma]$, and assume that this average number has a density $f(\gamma,s)$ with respect to s. In other terms,

$$\frac{1}{\gamma}\int_0^{\gamma} \mathrm{Card}\{i, s \leq |S_{\gamma,\lambda,i}| \leq s+ds\} d\lambda = f(\gamma,s)ds \tag{4}$$

Theorem 1 *Assume that there exists some $\gamma > 0$ such that (4) holds, i.e. the average number of sections with area s, for $0 \leq \lambda \leq \gamma$, has a density $f(\gamma,s)$. Then there is a constant c, not depending on I, such that*

$$||I||_{BV} \geq c \int_0^{\nu(\Omega)} s^{\frac{1}{2}} f(\gamma,s) ds. \tag{5}$$

Proof Applying Corollary 1 and Lemma 1

$$\begin{aligned}
||I||_{BV} &= \int_{\mathbb{R}} \mathrm{per}\{x, I(x) \geq \lambda\} d\lambda \\
&= \frac{1}{2}\left(\int_{\mathbb{R}} \mathrm{per}\{x, I(x) \geq \lambda\} d\lambda + \int_{\mathbb{R}} \mathrm{per}\{x, I(x) \geq \lambda - \gamma\} d\lambda\right) \\
&\geq \frac{1}{2}\int_{\mathbb{R}} \mathrm{per}(\chi_{\lambda-\gamma} I \setminus \chi_{\lambda} I) d\lambda \\
&= \frac{1}{2}\sum_{n\in\mathbf{Z}} \int_{n\gamma}^{(n+1)\gamma} \mathrm{per}(\chi_{\lambda-\gamma} I \setminus \chi_{\lambda} I) d\lambda \\
&= \frac{1}{2}\int_0^{\gamma} \sum_{n\in\mathbf{Z}} \mathrm{per}(\chi_{\lambda+(n-1)\gamma} I \setminus \chi_{\lambda+n\gamma} I) d\lambda \\
&= \frac{1}{2}\int_0^{\gamma} \sum_{i\in J(\gamma,\lambda)} \mathrm{per}(S_{\gamma,\lambda,i}) d\lambda.
\end{aligned}$$

By isoperimetric inequality (2), we therefore obtain

$$||I||_{BV} \geq \pi^{\frac{1}{2}} \int_0^{\gamma} \sum_{i\in J(\gamma,\lambda)} |S_{\gamma,\lambda,i}|^{\frac{1}{2}} d\lambda.$$

Applying Fubini-Tonelli Theorem, some slicing and the assumption (4), we get

$$
\begin{aligned}
||I||_{BV} &\geq \pi^{\frac{1}{2}} \int_0^{\gamma} d\lambda \int_0^{\nu(\Omega)} \mathrm{Card}\{i \in J(\gamma,\lambda),\ s \leq |S_{\gamma,\lambda,i}| \leq s+ds\} s^{\frac{1}{2}} \\
&= \pi^{\frac{1}{2}} \int_0^{\nu(\Omega)} \int_0^{\gamma} d\lambda \mathrm{Card}\{i \in J(\gamma,\lambda),\ s \leq |S_{\gamma,\lambda,i}| \leq s+ds\} s^{\frac{1}{2}} \\
&= \gamma \pi^{\frac{1}{2}} \int_0^{\nu(\Omega)} s^{\frac{1}{2}} f(\gamma,s) ds.
\end{aligned}
$$

□

We can repeat the preceding analysis by assuming now that

$$
\mathrm{Card}\{i, p \leq \mathrm{per}(S_{\gamma,\lambda,i}) \leq p+dp\}
$$

has an average density $g(\alpha, p)$ with respect to p. That is to say

$$
\frac{1}{\gamma} \int_0^{\gamma} \mathrm{Card}\{i, p \leq \mathrm{per}(S_{\gamma,\lambda,i}) \leq p+dp\} d\lambda = g(\gamma,p)dp. \tag{6}
$$

Then we have the analog of Theorem 1 for the perimeters of sections:

Theorem 2 *Assume that there exists some $\gamma > 0$ such that (6) holds, i.e. the average number of sections with perimeter s, for $0 \leq \lambda \leq \gamma$, has a density $g(\gamma, p)$. Then*

$$
||I||_{BV} \geq \frac{1}{2} \int_0^{+\infty} p g(\gamma,p) dp. \tag{7}
$$

Proof By proceeding as in the proof of Theorem 1 we have again:

$$
||I||_{BV} \geq \frac{1}{2} \int_0^{\gamma} \sum_{i \in J(\gamma,\lambda)} \mathrm{per}(S_{\gamma,\lambda,i}) d\lambda.
$$

Applying again some slicing and assumption (6) we get

$$
\begin{aligned}
||I||_{BV} &\geq \frac{1}{2} \int_0^{\gamma} d\lambda \int_0^{+\infty} \mathrm{Card}\{i \in J(\gamma,\lambda),\ p \leq \mathrm{per}(S_{\gamma,\lambda,i}) \leq p+dp\} p \\
&= \frac{\gamma}{2} \int_0^{+\infty} p g(\gamma,p) dp.
\end{aligned}
$$

□

4.2 Application to natural images

In this section, we draw the consequences of Theorems 1 and 2 for the images analyzed in Section 3. According to the results of this section, we can assume that the considered images satisfy

$$f(\gamma, s) = \frac{C}{s^\alpha} \tag{8}$$

$$g(\gamma, p) = \frac{C}{p^\beta} \tag{9}$$

for some constants $\alpha > 0$, $\beta > 0$. This law has been experimentally checked for several values of $\gamma = \frac{256}{k}$, k ranging from 8 to 20. We also checked that the value of α was almost not modified when the bilevels were not defined from gray level 0, but from some gray level less that $\frac{256}{k}$ (that is to say, in the continuous model, for different values of λ). By Theorem 1 we have

$$||I||_{BV} \geq c \int_0^{\nu(\Omega)} \frac{Cs^{\frac{1}{2}}}{s^\alpha} ds = +\infty \text{ if } \alpha > \frac{3}{2}$$

and in the same way,

$$||I||_{BV} \geq c \int_0^{+\infty} \frac{Cp}{p^\beta} dp = +\infty \text{ if } \beta > 2.$$

thus if we admit that (8) and (9) indeed hold for natural images when $s \to 0$, as is indicated by the experiments of section 4.1, we obtain that the considered images are not in BV if $\alpha > \frac{3}{2}$, or $\beta > 2$. Notice, however, that $\alpha > 2$, which happens for several of the considered images, is not compatible with a finite image area, since then $\int \frac{sds}{s^\alpha} = +\infty$. As suggested to us by Vicent Caselles and Stéphane Mallat, this raises the question of whether the area is correctly measured by covering pixels. In fact, if a region is very ragged, the cardinality of covering pixels may be related to its perimeter as well, in which case the estimate of $g(\gamma, s)$ is more reliable.

We here point out that wavelet coefficients (see [24], [20] for an introduction to wavelet decompositions) also give a way to decide whether or not an image belongs to the space BV. Let (c_k) be the wavelets coefficients of the image I, ordered in a nonincreasing sequence. Let us suppose that the wavelets have compact supports. We say that the c_k's are in l^1 if $\sum |c_k| < +\infty$, and that they are in weak-l^1 if there exists a constant C such that $c_k \leq \frac{C}{k}$. Obviously l^1 is included in weak-l^1. It is quite easy to prove that if the c_k are in l^1, then I is in BV. In the other direction, Cohen and al., [7], recently proved that if I is in BV, then the c_k's are in weak-l^1. Thus it is possible to decide whether an image belongs or not to BV by looking at its wavelet coefficients decay, except if they decrease as $\frac{C}{k}$, which happens to be often the case ([19]). Moreover, it is worth noticing that the wavelet coefficients produced by the characteristic function of a simple shape already decrease as $\frac{1}{k}$. We do not present here a precise comparison between the two criteria. Let us just mention that in the case of the baboon image (Figure 3), both methods agree: this image is not in BV. For the well-known image of Lena, our approach gives an α of 1.9 (for $k = 16$), which suggests Lena being out of BV, whereas from the wavelet approach, the image is in BV.

5 The dead leaves model

A tridimensional scene may be modeled as a group of solid objects occluding each other partially or fully, according to their respective positions and their position with respect to the observer. How can then the observer make inferences on the size of these partially hidden objects ? And which ratio can we assume between the size of visible parts of objects and the size of the objects themselves ? This question can be addressed in a statistical framework : given a size distribution for the objects before occlusion, how can we compute the size distribution of visible parts ? In order to do so, we shall use a random shape model developed by Matheron ([21], [22]) in the framework of his theory, Mathematical Morphology. He called this model dead leaves model. Its formalization requires some explanations, definitions and mathematical results which are given in Section 5.1.

This study has two motivations. The first one comes from phenomenology, which states that the occlusion phenomenon, so essential in human or animal life, does not seem to affect too much our perception of objects. In particular, Kanisza [16] studied several perceptual strategies to detect occlusions, but also to get rid of them by mental completion. This leads to the notion that with only the visible parts of objects, simply by interpolating their hidden parts, a layered view can be mentally reconstructed. Even though gestaltists do not seem to have addressed the question, it is plain that such an interpolation (which Kanisza called amodal completion) can be performed only if a significant part of each object is under sight, large and significant enough to make extrapolation possible. Amodal completion might be in many cases related to our prior knowledge of objects. However, Kanisza and his school proved that the completion takes place fairly independently from this prior knowledge, always giving a preference to constraints of a geometric nature, may these lead or not to a common sense contradicting completion. The second motivation is the experimental program detailed in Section 3. Even though the link between our definition of objects and the real visible parts of objects is not obvious, we can, at least in one case, derive a power law with the help of the model. The dead leaves model permits to study the influence of occlusion on the size distribution of a set of objects, with some restrictions, however:

- We shall only address stationary models, which do not take into account perspective for instance. It is possible to include perspective in the framework of the dead leaves model, but this will not be presented here.

- Objects will be assumed to spread out independently, which is of course an undue restriction for a real scene where objects interact physically.
- Most of the theoretical results won't deal directly with the area of visible objects, but rather with the size distribution of their intersections with a straight line.

5.1 Random closed and compact sets of $\mathbb{R}^n$

This section contains the prerequisites for the understanding of the dead leaves model. The interested reader should consult [23]. For an introduction to the methods and applications of stochastic geometry, we refer to [33]. We denote by $\mathcal{F}$, $\mathcal{G}$, $\mathcal{K}$, respectively the sets of all closed, open and compact sets of $\mathbb{R}^n$. In the following, we define a topology on $\mathcal{F}$ and $\mathcal{K}$ and its associated σ-algebra of Borel sets. The choice for the topology is motivated by the following considerations :

- The set spaces thus obtained are separable, Hausdorff and compact.
- The basic sets (open sets of sets, etc.) for this topology have an easy interpretation.
- The topology induced on $\mathcal{K} \setminus \{\emptyset\}$ is metrizable with the Hausdorff metric.

Let us set for any subset A of $\mathbb{R}^n$,

$$\mathcal{F}^A = \{F \in \mathcal{F} / F \cap A = \emptyset\},$$

$$\mathcal{F}_A = \{F \in \mathcal{F} / F \cap A \neq \emptyset\}.$$

Define in the same way subsets of $\mathcal{K}$, $\mathcal{K}^A$ and $\mathcal{K}_A$. We define a topology $T_{\mathcal{F}}$ on $\mathcal{F}$, generated by the basis of open sets

$$(\mathcal{F}^K, K \in \mathcal{K}; \mathcal{F}_G, G \in \mathcal{G}).$$

It can be shown that $\mathcal{F}$, endowed with the topology $T_{\mathcal{F}}$, is separable, Hausdorff, and compact. We can also endow $\mathcal{K}$ with a topology $T_{\mathcal{K}}$, a basis of which is

$$(\mathcal{K}^F, F \in \mathcal{F}; \mathcal{K}_G, G \in \mathcal{G}).$$

This topology is strictly finer than the topology induced on $\mathcal{K}$ by $T_{\mathcal{F}}$. Define on $\mathcal{K}' = \mathcal{K} \setminus \{\emptyset\}$ the Hausdorff distance d_H by

$$d_H(A, B) = inf\{\delta > 0 / A \subset B^\delta \text{ and} B \subset A^\delta\},$$

where

$$A^\delta = \{x \in \mathbb{R}^n / d(x, A) \leq \delta\}.$$

It can be shown that the topology $T_{\mathcal{K}}$ on $\mathcal{K}'$ and the topology induced by d_H on $\mathcal{K}'$ are equal ("point" $\emptyset$ is isolated in $\mathcal{K}$). We denote by $B_{\mathcal{F}}$ and $B_{\mathcal{K}}$ the Borel σ-algebras generated by $T_{\mathcal{F}}$ et $T_{\mathcal{K}}$ respectively.

Definition 1 *A random closed set (respectively a random compact set) of* $\mathbb{R}^n$ *is a measurable function from a probability space* (Ω, T, P) *into* $(\mathcal{F}, B_{\mathcal{F}})$ *(respectively into* $(\mathcal{K}, B_{\mathcal{K}})$*)*

We associate with a random closed set F its range measure μ_F, defined on Borel sets $B \in B_{\mathcal{F}}$ by $\mu_F(B) = P(F^{-1}(B))$. We write $Pr(F \cap A = \emptyset) = \mu_F(\mathcal{F}_A)$, $Pr(F \cap A \neq \emptyset) = \mu_F(\mathcal{F}^A)$, and so on. A first basic result for the study of random sets is the following:

Theorem 3 *Any random closed set* F *is characterized by its capacity function defined for every compact set* K *by* $G_F(K) = Pr(F \cap K = \emptyset)$.

Proof : Assume F_1 and F_2 are two closed random sets with the same capacity functions G_{F_1} and G_{F_2}. Let us define

$$\mathcal{I} = \{\mathcal{F}^K_{G_1,\ldots,G_n} = \mathcal{F}^K \cap \mathcal{F}_{G_1} \ldots \cap \mathcal{F}_{G_n}, K \in \mathcal{K}, G_1 \ldots G_n \in \mathcal{G}\}.$$

We first remark that $\mathcal{I}$ is an intersection stable family that generates the Borel σ-algebra associated with the topology $T_{\mathcal{F}}$. This is because $\mathcal{F}^{K_1} \cap \mathcal{F}^{K_2} = \mathcal{F}^{K_1 \cup K_2}$, and because $\mathcal{I}$ contains the $\mathcal{F}^K = \mathcal{F}^K_\emptyset$ and $\mathcal{F}_G = \mathcal{F}^\emptyset_G$.

Therefore, it is enough, in order to show that μ_{F_1} and μ_{F_2} are equal, to show that they agree on $\mathcal{I}$. Now, for any integer n, and a random closed set F, we have the following formula

$$\mu_F(\mathcal{F}^K_{G_1,\ldots,G_n}) = \mu_F(\mathcal{F}^K_{G_1,\ldots,G_{n-1}}) - \mu_F(\mathcal{F}^{K \cup G_n}_{G_1,\ldots,G_{n-1}}),$$

so that an induction shows that equality of μ_{F_1} and μ_{F_2} on $\mathcal{I}$ follows from their equality on the sets $\mathcal{F}^{K \cup G}$, for any compact K and open G (sets $\mathcal{F}^{K \cup G}$ are defined the same way as $\mathcal{F}^K$).

Now, for any $G \in \mathcal{G}$, there exists a sequence of compact sets, (K_n) such that $K_n \uparrow G$, so that $K \cup K_n \uparrow K \cup G$ which entails $\mathcal{F}^{K \cup K_n} \downarrow \mathcal{F}^{K \cup G}$. By continuity of the probability measure with respect to monotone sequences, we then have $\mu_{F_1}(\mathcal{F}^{K \cup K_n}) \downarrow \mu_{F_1}(\mathcal{F}^{K \cup G})$, which completes the proof of the equality of μ_{F_1} and μ_{F_2}, since they agree on $\mathcal{F}^{K \cup K_n}$ by hypothesis. $\square$

Definition 2 *We shall say that a random closed set with capacity* G *is stationary if for every compact set* K *and every point* x *of* $\mathbb{R}^n$ *one has* $G(K) = G(K + x)$. *We say that* F *is isotropic if it is stationary and for every rotation* r, $G(K) = G(r(K))$

We end this section with some notations that will be of use in the following. For any sets A and B, we set

$$\begin{aligned} \check{A} &= \{-x, x \in A\} \\ A \ominus B &= \{x \in \mathbb{R}^2 / x + \check{B} \subset A\} \\ A \oplus B &= \{x + y, x \in A, y \in B\}, \end{aligned}$$

$A \ominus B$ is called erosion of A by B, and $A \oplus B$ dilation of A by B.

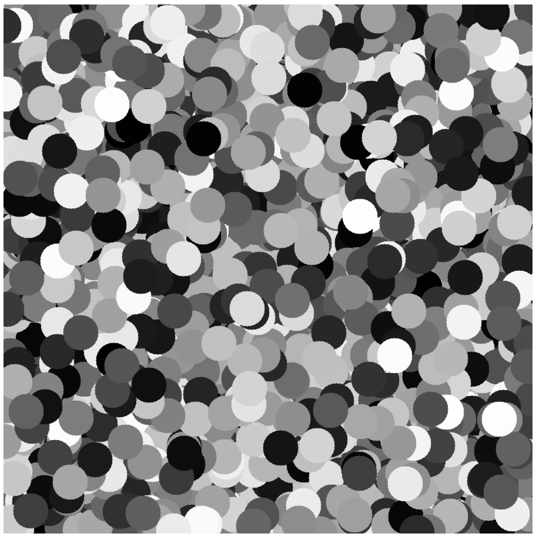

Figure 7: simulation of a dead leaves model, where the grain X_0 is a disk with constant radius

5.2 The dead leaves model

The model we present here is a slight variation of a model introduced by Matheron [22]. The original model was defined on all of $\mathbb{R}^2$, whereas we define it only on a window, but the result is locally identical. In Appendix, we present the dead leaves model as introduced originally by Matheron.

We consider a closed bounded domain Ω in $\mathbb{R}^2$ (a rectangle for instance). Let $X_i, i \in -\mathbb{N}$ ($-\mathbb{N}$ is the set of negative integers) be a family of independent equally distributed random compact sets X_i. We assume that X_0 is almost surely uniformly bounded, i.e. there exists $R > 0$ such that $X_0 \subset B(0, R)$ almost surely. We also assume, for reasons that will become clear soon, that $E\nu(X_0 \ominus D) > 0$ for some disk D. We then consider a family $x_i, i \in -\mathbb{N}$ of random variables which are independently uniformly distributed in $\Omega^R = \Omega + B(0, R)$, independently of the X_i's. We define a sequence of dead leaves in the following way.

For each $i \in -\mathbb{N}$ we call dead leaf at time i the random set $x_i + X_i$. We then call visible part of the dead leaf i the set

$$x_i + X_i \setminus (\bigcup_{i<j\leq 0} (x_j + int(X_j))),$$

where $int(X)$ is the interior of X. We then define the dead leaves model M to be the collection of all the visible part of the dead leaves i for $i \in -\mathbb{N}$. We also define a cell of M to be any connected component of a visible part. In the

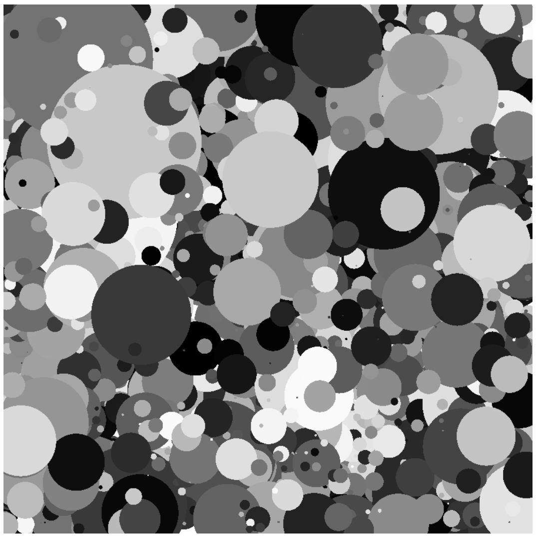

Figure 8: simulation of a dead leaves model, where the grain X_0 is a disk with an exponentially distributed radius

following, X_0 will often be referred to as the grain of the model.

In Figures 7, 8 and 9, we present some experimental realizations of dead-leaves models, for different choices of grain X_0. To simulate the model, we associate with each $x_i + X_i$ a different gray-level uniformly distributed between 0 and 255. We already see from those examples that the behavior of the dead leaves model is strongly dependent on the shape of the leaves, as well as on their size distribution. One also notices that the 3D perception is much stronger in the case of elongated rectangles than in the case of disks.

Lemma 2 *Let K be a compact set of Ω. Then the probability that K be contained in a given dead leaf $x_i + X_i$ is*

$$\frac{E\nu(X_0 \ominus \check{K})}{\nu(\Omega^R)}.$$

where $\nu()$ is the Lebesgue measure.

Proof K is in the dead leaf X_i if $K \subset x_i + X_i$, which is equivalent to $x_i + \check{K} \subset \check{X}_i$, that is, $x_i \in \check{X}_i \ominus K$. It is easily checked that $\check{X}_i \ominus K \subset \Omega^R$. Since x_i is

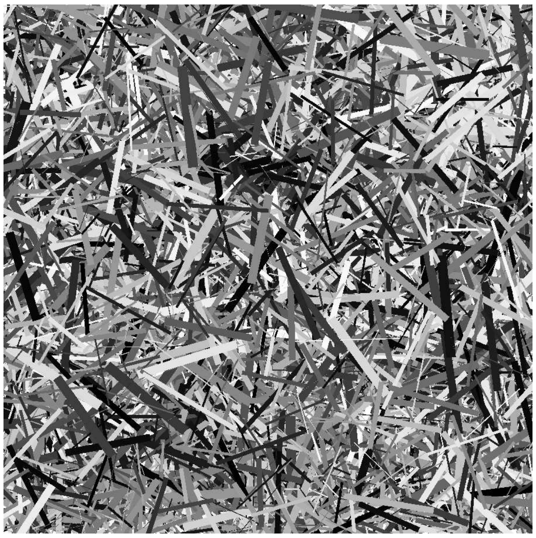

Figure 9: simulation of a dead leaves model, where the grain X_0 is a rectangle with a rotation uniformly distributed in $[0, 2\pi]$, and with width and length following a uniform law between extremal values

uniformly distributed in Ω^R, we deduce that

$$\begin{aligned} Pr(K \subset x_i + X_i) &= Pr(x_i \in \check{X}_i \ominus K) \\ &= \frac{1}{\nu(\Omega^R)} \int_{\Omega^R} Pr(x \in \check{X}_i \ominus K) dx \\ &= \frac{1}{\nu(\Omega^R)} \int_{\Omega^R} E(\mathbb{1}_{\check{X}_i \ominus K}(x)) dx \\ &= \frac{1}{\nu(\Omega^R)} E \int_{\Omega^R} \mathbb{1}_{\check{X}_i \ominus K}(x) dx \\ &= \frac{E\nu(\check{X}_0 \ominus K)}{\nu(\Omega^R)} = \frac{E\nu(X_0 \ominus \check{K})}{\nu(\Omega^R)}. \end{aligned} \tag{10}$$

□

Corollary 2 *If $K \subset \Omega$, K measurable, is such that $E\nu(X_0 \ominus \check{K}) > 0$, then K is covered by some dead leaf with probability 1. Since for some disk D, we have $E\nu(X_0 \ominus D) > 0$, then Ω is almost surely covered by a finite number of dead leaves.*

Proof Let $a = \frac{E\nu(X_0 \ominus \check{K})}{\nu(\Omega^R)} > 0$. Then K is not completely covered by a leaf $x_i + X_i$ with probability $(1 - a) < 1$. Thus, since leaves are independent, K is not covered by any leaf with probability 0. Moreover, it is possible to cover Ω with a finite number of the sets $Y_i = (y_i + D) \cap \Omega$, with y_i points of Ω. Then each Y_i is covered by some dead leaf at time t_i with probability one, and thus at $T = max\{t_i\}$, Ω is covered by a finite number of leaves. $\square$

We are interested in the relation between the size of the grain X_0 and the size of the visible parts of the dead leaves model, or the size of the cells. Next, we are going to prove a fundamental result which yields the probability for a compact set $K \subset \Omega$ to be included in a visible part of the dead leaves model.

Theorem 4 *Let M be the dead leaves model associated with X_0 as before. Then for any compact set $K \subset \Omega$, the probability for K to be included in a visible part of M is given by:*

$$Q(K) = \frac{E\nu(X_0 \ominus \check{K})}{E\nu(X_0 \oplus \check{K})}. \tag{11}$$

Clearly, if K is connected, $Q(K)$ also denotes the probability of K to be included in a cell of the model.

Proof : First, we note that the probability $Q(K)$ corresponds to the probability that for some i, K is included in the dead leaf $x_i + X$ and for $i < j \leq 0$, K does not intercept any of the dead leaves $x_j + X$. As the proof of Lemma 2 shows, in order to compute $Pr(K \cap (x_j + X_j) = \emptyset)$, we can work with $X_j = X$ deterministic, x_j uniformly distributed, and then take the expectation of the resulting formulas. The probability that the compact set K does not intercept the dead leaf $x_j + X$ is equal to the probability of $-x_j \notin X \oplus \check{K}$. We remark that since $X \subset D(0, R)$ and $\check{K} \subset (-\Omega)$, $(X \oplus \check{K}) \subset (-\Omega^R)$. We deduce that the probability in x_j (for fixed X) that K does not meet $x_j + X$ is equal to

$$Pr(K \cap (x_j + X) = \emptyset) = 1 - Pr(-x_j \in X \oplus \check{K}) = 1 - \frac{\nu(X \oplus \check{K})}{\nu(\Omega^R)}.$$

Now, if X_j is a random compact set distributed as X_0, taking the expectation yields

$$Pr(K \cap (x_j + X_j) = \emptyset) = 1 - \frac{E\nu(X_0 \oplus \check{K})}{\nu(\Omega^R)}. \tag{12}$$

We now consider the following disjoint events E_i, for $i \in -\mathbb{N}$: K is included in the dead leaf $x_i + X_i$ and for $i < j \leq 0$ the compact set K does not intercept the dead leaves $x_j + X_j$. Since the leaves are independent, we have, using Lemma 2,

$$Pr(E_i) = (1 - \frac{E\nu(X_0 \oplus \check{K})}{\nu(\Omega^R)})^{-i} \frac{E\nu(X_0 \ominus \check{K})}{\nu(\Omega^R)}. \tag{13}$$

Making the sum over $i \in -\mathbb{N}$ in (13) we obtain that the probability that K is contained in a visible part of the dead leaves covering is

$$Q(K) = \sum_{i \in -\mathbb{N}} (1 - \frac{E\nu(X_0 \ \oplus \check{K})}{\nu(\Omega_R)})^{-i} \frac{E\nu(X_0 \ominus \check{K})}{\nu(\Omega^R)},$$

which yields the announced result by summing the geometric sum,

$$Q(K) = \frac{E\nu(X_0 \ominus \check{K})}{E\nu(X_0 \oplus \check{K})}. \tag{14}$$

□

We end the section by remarking that if each leaf is assigned a value (color) in $[0, C]$, C real, then a realization of the model can be seen as a function from Ω to $[0, C]$. We may then wonder which functional space such a function belongs to. The characteristic function of a set with finite boundary length in Ω is in the space of functions with bounded variation, $BV(\Omega)$, and this will also be the case for almost every realization of the dead leaves model. Indeed, we saw previously that for almost every realization there exists a $i_0 < 0$ such that the union of the leaves with indices $i > i_0$ covers Ω. This is also the case for the corresponding cells. Moreover, the sum of the length of the boundaries of those cells is obviously smaller than the sum of the lengths of the boundaries of their parent leaves. Now, if the parent leaves have boundaries whose length is bounded by a constant value B, the BV norm is finite and bounded by $-i_0 * B$ for almost all realization. This in particular is the case if the leaves are convex since they are included in $B(0, R)$. We have thus proved:

Proposition 2 *If X_0 has uniformly bounded boundary length, then the dead leaves model M is in $BV(\Omega)$ almost surely.*

The meaning of the result is strong and general. Even though we can have arbitrarily small realizations of X_0, these small leaves do not succeed in creating boundaries with infinite length; the reason being that the small leaves are steadily covered by larger ones.

The next section introduces a tool in the study of compact sets that will be of use in the evaluation of the sizes of the cells of the dead leaves model.

5.3 The geometric covariogram

The geometric covariogram of a convex and compact set C of $\mathbb{R}^2$ is the function γ_C from $\mathbb{R}^2$ to $\mathbb{R}$ defined by

$$\gamma_C(l, \theta) = \nu(C \cap C_{l,\theta}) = \gamma_C(-l, \theta),$$

where $C_{l,\theta}$ is the set $C + x$, x being the point with polar coordinates (l, θ), $l > 0$, $\theta \in [0, \pi)$. The γ_C function has compact support and is equal to zero if $l \geq \nu_1(p_{\theta^\perp}(C))$ where p_θ is the orthogonal projection on a line perpendicular to the direction θ, $\theta^\perp = \theta + \pi/2$, and ν_1 is the Lebesgue measure in $\mathbb{R}$. In the following we assume that $\nu_1(p_\theta(C)) > 0$ for all θ (this will be the case with the leaves of M which have non-empty interiors).

We are interested in the size distribution of the intersection of C with all lines with direction θ meeting C. We assume that the intersection of these random lines with a line perpendicular to θ is uniformly distributed in $p_\theta(C)$. We associate with each $s \in p_\theta(C)$ a line L_s of direction θ and meeting C. We denote by $F_C(l,\theta)$ the distribution function of the size of the segments, defined by

$$F_C(l,\theta) = \frac{\nu_1(\{s, \nu_1(L_s \cap C) \geq l\}}{\nu_1(p_\theta(C))}.$$

Lemma 3

$$F_C(l,\theta) = \frac{\nu_1(p_\theta(C \cap C_{l,\theta}))}{\nu_1(p_\theta(C))}.$$

Proof It is easily seen that

$$L_s \cap C \cap C_{(l,\theta)} \neq \emptyset \text{ iff } \nu_1(L_s \cap C) \geq l,$$

which yields the result. □

The function F_C is related to the function γ_C by the next theorem:

Theorem 5 *For any compact and convex set C in $\mathbb{R}^2$, $\theta \in [0,\pi)$, $\gamma_C(.,\theta)$ is differentiable on $[0, p_{\theta^\perp}(C))$ and*

$$\frac{\partial}{\partial l}\gamma_C(l,\theta) = -\nu_1(p_\theta(C \cap C_{l,\theta})) = -\nu_1(p_\theta(C))F_C(l,\theta). \tag{15}$$

Proof : ([23])

for any fixed θ,

$$\begin{aligned} \epsilon\nu_1(p_\theta(C \cap C_{l,\theta})) &\geq \nu(C \cap C_{l,\theta} \setminus C \cap C_{l+\epsilon,\theta}) = \gamma_C(l,\theta) - \gamma_C(l+\epsilon,\theta) \\ &\geq \epsilon\nu_1(p_\theta(C \cap C_{l+\epsilon,\theta})), \end{aligned} \tag{16}$$

and

$$\begin{aligned} \epsilon\nu_1(p_\theta(C \cap C_{l-\epsilon,\theta})) &\geq \nu(C \cap C_{l-\epsilon,\theta} \setminus C \cap C_{l,\theta}) = \gamma_C(l-\epsilon,\theta) - \gamma_C(l,\theta) \\ &\geq \epsilon\nu_1(p_\theta(C \cap C_{l,\theta})). \end{aligned} \tag{17}$$

In order to establish the theorem, it is enough to prove the continuity of the function $f(l) = \nu_1(p_\theta(C \cap C_{l,\theta}))$. For any $l > 0$, $\gamma_C(.,\theta)$ is continuous from the left. Indeed,

$$l_n \uparrow l \Rightarrow C \cap C_{l_n,\theta} \downarrow C \cap C_{l,\theta},$$

(since the interception is a s.c.s. function) so that

$$l_n \uparrow l \Rightarrow p_\theta(C \cap C_{l_n,\theta}) \downarrow p_\theta(C \cap C_{l,\theta}),$$

(because p_θ is continuous from $\mathcal{K}'$ to $\mathcal{K}'$ ([23], pp69)). Therefore, we obtain $\nu_1(p_\theta(C \cap C_{l_n,\theta})) \downarrow \nu_1(p_\theta(C \cap C_{l,\theta}))$ following the monotone continuity of ν_1, so the left-continuity is proved. It is easily seen that $f(l)$ is continuous from the right for any l such that $\gamma_C(l,\theta) > 0$, which ends the proof. $\square$

We note that, in particular, the theorem states that γ_C is a convex function. In what follows, we denote by γ'_C the derivative of γ_C with respect to l. We can generalize the definition of the γ_C function to a random compact set, using the expectation $E\gamma_C(l,\theta)$ with respect to the probability law of C, when such a definition makes sense. In particular, we can define such an expectation when $E\nu(C) < +\infty$, which is the case for the grain of the dead leaves model. For instance, we can consider a family of random compact sets $\{C^r\}_{r\in I}$, depending among other parameters on $r > 0$ in such a way that C^r is a rescaling of C, $C^r = rC$, with a probability distribution $f(r)$. We then define the covariogram as

$$\gamma(l,\theta) = E\int_0^{+\infty} \gamma_{C^r}(l,\theta) f(r) dr.$$

We are interested in the lengths of the intercepts of C, that is to say its intersections with lines of direction θ. We define

$$F_C(l,\theta) = \frac{E\gamma'_C(l,\theta)}{E\gamma'_C(0,\theta)},$$

when this makes sense, that is to say when $E\gamma'_C(0,\theta)$ is finite and nonzero. As an easy consequence of Theorem 5, we have

Corollary 3 *Assume that*

$$0 < -E\gamma'_C(0,\theta) < +\infty.$$

Then for any $l \in [0, p_{\theta^\perp}(C))$,

$$(E\gamma_C(l,\theta))' = E\gamma'_C(l,\theta),$$

so that

$$F_C(l) = \frac{(E\gamma_C(l,\theta))'}{(E\gamma_C(0,\theta))'}. \tag{18}$$

Proof We take the expectation with respect to the probability law of C of all members of relations ((16), (17)). We notice that we can divide by ε and pass to the limit as $\epsilon \to 0$. Indeed, we just have to prove that $E\nu_1(p_\theta(C \cap C_l))$ is continuous with respect to l. Now, we have seen in the proof of Theorem 5 that $\nu_1(p_\theta(C \cap C_l))$ is continuous with respect to l. In addition, we obviously have $\nu_1(p_\theta(C \cap C_l)) \le \nu_1(p_\theta(C))$ which has a finite expectation by assumption. This proves the asked continuity by Lebesgue theorem. So we get by dividing by ε and letting $\varepsilon \to 0$ the analogue of (15), that is

$$\frac{\partial}{\partial l} E\gamma_C(l,\theta) = -E\nu_1(p_\theta(C \cap C_{l,\theta}))) = -E\nu_1(p_\theta(C))F_C(l,\theta) \tag{19}$$

□

Thus in the following, where we will only consider covariograms for random convex sets, we will write γ_C for $E\gamma_C$, and γ'_C for its derivative with respect to l.

5.4 The distribution of the intercepts in the dead leaves model

We are interested in the relation between the size of the convex set X_0, the grain in the dead leaves model, and the size of the connected components of the partition of the plane provided by the dead leaves model M, the cells of M. In the following, as a notion of size of a set, we choose the set of sizes of its intersections with lines of direction θ, the so-called intercepts. We fix θ, and we will sometimes omit it. Let x_0 be the center of Ω, the window in which the model has been defined. Let D_θ be the straight line of direction θ passing by x_0. Let $F^{x_0}_{tess}(l)$ be the distribution function of the length of the component of $M \cap D_\theta$ containing x_0. We define

$$F_{tess}(l) = C \int_l^{+\infty} \frac{-dF^{x_0}_{tess}(u)}{u},$$

(the integral has actually finite bounds, since the model is bounded), according to the usual length biased sampling. The preceding integral is a slight abuse. In fact $\frac{-dF^{x_0}_{tess}}{u} = \mu$ is a locally bounded nonnegative Radon measure on $(0, +\infty]$, since $F^{x_0}_{tess}(l)$ is nonincreasing, and the integral means $F_{tess}(l) = C\mu([l, +\infty])$. Roughly, this definition means that we are interested in the length repartition of all the components of the intersection of the model with a straight line, whereas it is more likely that x_0 be in a large component that in a small one. Besides, in order to avoid edge effects, we assume that $x_0 + X_0 \subset \Omega$ almost surely (we consider a large enough window).

As we are going to see, $F_{tess}(l, \theta)$ is related to $Q(K, \theta)$, the probability that a segment $K \in \Omega$ on a line with direction θ and length l be included in a connected component of M (in this case K is connected, so that the probability that it is included in a visible part or a cell are the same). We write $Q(l, \theta) = Q(K, \theta)$. We will sometimes write $Q(l)$, $F_{tess}(l)$, when there is no ambiguity about the value of θ.

Theorem 6 *Let M be a dead leaves model associated with a convex X_0 such that there exists a point x_0 such that $x_0 + X_0 \subset \Omega$ almost surely. Then for any $\theta \in [0, \pi)$*

$$Q(l, \theta) = \int_l^{+\infty} \frac{y - l}{y} dF^{x_0}_{tess}(y, \theta), \tag{20}$$

so that

$$F_{tess}(l, \theta) = \frac{Q'(l, \theta)}{Q'(0, \theta)}. \tag{21}$$

Proof : Consider the following event $E(y,dy)$: a segment contained in $M \cap D_\theta$ with length $y \leq \lambda \leq y+dy$ contains x_0. The probability of $E(y,dy)$ is $-dF^{x_0}_{tess}(y)$. On the other hand, since M is stationary[1], the conditional probability that the segment of $M \cap D_\theta$ containing x_0 also contains the whole segment $[x_0 - \frac{l}{2}, x_0 + \frac{l}{2}]$, knowing its length is y, is equal to the probability that a segment with length l and uniformly distributed center $x \in [-\frac{y}{2}, \frac{y}{2}]$ be wholly contained in $[-\frac{y}{2}, \frac{y}{2}]$. Obviously, this last probability is $\frac{y-l}{y}$. Since the events $E(y,dy)$ are disjoint, we conclude that

$$Q(l) = \int_l^{+\infty} \frac{l-y}{y} dF^{x_0}_{tess}(y).$$

This implies that $F^{x_0}_{tess}(l) = Q(l) - lQ'(l)$ and therefore by an easy calculation using (20) that $F_{tess}(l) = -CQ'(l)$. But $F_{tess}(0) = 1$, and the theorem is proved. □

Next, we are going to formulate the relation between $Q(l)$, the probability for a segment of size l to be included in M, and the covariogram $\gamma(l)$ associated with the grain X_0. We will write $[0,l]$ for the segment of length l and direction θ, when θ is fixed.

Theorem 7 *Consider a dead leaves model whose grain X_0 is convex. Then*

$$Q(l,\theta) = \frac{\gamma(l,\theta)}{\gamma(0,\theta) - l\gamma'(0,\theta)}. \tag{22}$$

Proof According to Formula (11), we have

$$Q(l) = \frac{E\nu(X_0 \ominus [0,l])}{E\nu(X_0 \oplus [0,l])},$$

and

$$E\nu(X \ominus [0,l]) = E\nu(X_0 \cap (X_0)_{-l}) = \gamma(l),$$

because X is convex, and, again by convexity,

$$E\nu(X_0 \oplus [0,l]) = E\nu(X_0) + lE\nu_1(p_\theta(X_0)),$$

which is a directional version of Steiner formula. By Formula (18), we obtain

$$E\nu(X_0 \oplus [0,l]) = \gamma(0) - l\gamma'(0), \tag{23}$$

and the theorem follows. □

Theorem 8 *Consider a dead leaves model whose grain X_0 is convex. Then the size repartition $F_{tess}(l) = F_{tess}(l,\theta)$ of the connected visible intervals of the dead*

[1] M is not strictly speaking stationary, but we are only considering what happens inside the window Ω

leaves model in direction θ is related to the size repartition $F_C(l) = F_C(l,\theta)$ of the leaves in direction θ by

$$\int_h^{+\infty} F_{tess}(l)dl = \frac{\frac{1}{2}}{1+ah}\int_h^{+\infty} F_C(l)dl, \tag{24}$$

where

$$a = \frac{-\gamma'(0,\theta)}{\gamma(0,\theta)}.$$

Proof This is an easy consequence of Formulae (18), (21) and (22). □

Corollary 4 *The average size of the intercepts for the visible parts of leaves is twice as small as the for the leaves themselves. That is to say*

$$\int_0^{+\infty} l dF_{tess}(l,\theta) = \frac{1}{2}\int_0^{+\infty} l dF_C(l).$$

Proof By Theorem 8,

$$\int_0^{+\infty} F_{tess}(l)dl = \frac{1}{2}\int_0^{+\infty} F_C(l)dl, \tag{25}$$

and the result follows by integration by parts. □

If, in addition, we can make observations on the tessellation, then we can estimate F_{tess} and therefore deduce the value of a since, by (21) and (22),

$$\int_0^{+\infty} F_{tess}(l)dl = \frac{-1}{Q'(0)} = \frac{\gamma(0)^2}{2\gamma'(0)\gamma(0)} = \frac{1}{2a}.$$

Then Formula (24) permits to deduce F_C. *This relation also implies that there is no invariant distribution function, that is to say no distribution of the grains such that $F_{tess} = F_C$.* We also get

$$\int_0^{+\infty} F_C = \int_0^{+\infty} \frac{\gamma'(l)}{\gamma'(0)} = \frac{\gamma(0)}{\gamma'(0)} = \frac{E\nu(X_0)}{\pi EU(X_0)},$$

because, by Cauchy formula,

$$E\int_0^{2\pi} p_\theta(X_0)d\theta = \pi EU(X_0).$$

It seems nonetheless difficult to recover other characteristics of the dead leaves. Conversely, if we know the leaves model, then we can compute a and therefore F_{tess}. Notice, however, that the cells of the model are not necessarily convex. Thus the connected components of the intersection of a given straight-line with the tessellation are not independent : two intersections may belong to the same original leaf, or even to the same cell.

5.5 The stability of the power law

In this section, we prove that two important quantities approximatively follow a power law if the size distribution of objects does. The first one is the probability that two points at distance x apart are in the same visible part of an object, and the second one is the length distribution of connected intercepts, introduced in the previous section. The first of these results was proved, with a different formalism, by Ruderman, [28], who also experimentally observed the power law for the probability of two points belonging to the same object in natural images, the objects he considered being obtained by segmenting the image by hand.

We assume that realizations of the grain X_0 are homothetics $X^r = rX$ of a random compact X, according to the density function $f(r)$, and that X is almost surely convex and included in $B(1)$, a ball of radius 1. Moreover, we assume $f(r) = 0$ for $r < R_{min}$ or $r > R_{max}$, two positive constants. Without loss of generality we assume $R_{min} = 1$. Then we set, for any $\theta \in [0, \pi)$,

$$\gamma(l, r, \theta) = \gamma_{X^r}(l, \theta) = E\nu(X^r \cap X^r_{l,\theta}),$$

where the expectation is taken at fixed r, thus on realizations of X. First of all, we remark that

$$\gamma(l, r, \theta) = E\nu(X^r \cap X^r_{l,\theta}) = r^2 E\nu(X^1 \cap X^1_{\frac{l}{r},\theta}) = r^2\gamma(\frac{l}{r}, 1, \theta),$$

so that for any nonzero l, the geometric covariogram of X_0 is given by:

$$\gamma_{X_0}(l, \theta) = \int_1^{R_{max}} \gamma(l, r, \theta) f(r) dr = l^3 \int_{\frac{1}{l}}^{\frac{R_{max}}{l}} r^2 \gamma(\frac{1}{r}, 1, \theta) f(lr) dr.$$

If we now assume that $f(r) = Cr^{-\delta}$, that is to say that the size of objects follows a power law, we get

$$\gamma_{X_0}(l, \theta) = l^{3-\delta} A(l, \theta),$$

with

$$A(l, \theta) = \int_{\frac{1}{l}}^{\frac{R_{max}}{l}} r^2 \gamma(\frac{1}{r}, 1, \theta) f(r) dr.$$

Now, for $l \geq 1$, the previous integral is from 1 to $\frac{R_{max}}{l}$, since $\gamma(\frac{1}{r}, 1, \theta) = 0$ for $r < 1$, because $X \subset B(1)$. Eventually, if we suppose $\frac{R_{max}}{l}$ large enough and $\delta > 3$, we get

$$A(l, \theta) \approx A(\theta) = \int_1^{+\infty} r^2 \gamma(\frac{1}{r}, 1, \theta) f(r) dr,$$

so that

$$\gamma_{X_0}(l, \theta) \approx A(\theta) l^{3-\delta}, \tag{26}$$

with $A(\theta)$ independent of l. We also notice that if we allow objects to be unbounded, for instance using Matheron's original model (see Appendix 6), then the previous approximation becomes an equality for $l \geq 1$.

The probability $V(l,\theta)$ that two points at distance l apart in direction θ are included in the same visible part of the model is, according to Formula (11),

$$\begin{aligned} V(l,\theta) &= \frac{E\nu(X_0 \ominus \{x, x+l\})}{E\nu(X_0 \oplus \{x, x+l\})} = \frac{E\nu(X_0 \cap (X_0)_{l,\theta})}{E\nu(X \cup (X_0)_{l,\theta})} \\ &= \frac{\gamma_{X_0}(l,\theta)}{2\gamma_{X_0}(0) - \gamma_{X_0}(l,\theta)} \end{aligned}$$

Using Formula (26) we get, for large values of l (small enough compared to R_{max}), and $\delta > 3$,

$$V(l,\theta) \approx C(\theta) l^{3-\delta}. \tag{27}$$

By considering $\gamma(l)$, the average of $\gamma(l,\theta)$ on every direction, we get the probability that two points at distance l are in the same visible part: $V(l) \approx Cl^{3-\delta}$. It is then easy to show that, if each leaf is given a different color, the covariances of the model also follow a power law. This result was already obtained by Ruderman [28], using exact formulae for disks.

We now come to the study of $F_{tess}(l)$ the repartition function of intercepts that was defined in Section 5.4. We recall that, following Formulae (21) and (22),

$$F_{tess}(l,\theta) = K \left(\frac{-\gamma_{X_0}(l,\theta)}{\gamma_{X_0}(l,\theta) - l\gamma'_{X_0}(0,\theta)} \right)',$$

with $K = Q'(0,\theta) = \frac{2\gamma'(0)}{\gamma(0)}$, a negative constant. Now, following the same reasoning as before, it is easy to see that if $f(r)$ follows a power law $Cr^{-\delta}$, and if $\delta > 3$, then $\gamma'_{X_0}(l) \approx -cl^{2-\delta}$, with c a positive constant (as before either for fixed θ or on average in all directions). Indeed, we have

$$\gamma'(l,r,\theta) = E\nu_1(p_\theta(X^r \cap X^r_{l,\theta})) = rE\nu_1 p_\theta((X^1 \cap X^1_{\frac{l}{r},\theta})) = r\gamma'(\frac{l}{r}, 1, \theta),$$

and then

$$\gamma'_{X_0}(l,\theta) = \int_1^{R_{max}} \gamma'(l,r,\theta) f(r) dr = l^2 \int_{\frac{1}{l}}^{\frac{R_{max}}{l}} r\gamma'(\frac{1}{r}, 1, \theta) f(lr) dr.$$

Therefore, for values of $l \geq 1$ and $l \ll R_{max}$ (as before, this last hypothesis can be removed if objects are unbounded) we get

$$F_{tess}(l) = \frac{cl^{2-\delta} + Cl^{3-\delta}}{(A + Bl)^2},$$

with A, B constants so that, if $1 \ll l \ll R_{max}$,

$$F_{tess}(l) \approx Cl^{1-\delta}. \tag{28}$$

We now come to the link between the results about the occlusion model, and the experimental facts of Section 3. It must be noticed that, relating theoretical predictions for the occlusion model and experimental observations, Ruderman, in [28], observes a power law for the covariances on objects that were isolated in the images by hand. It was then argued that scaling phenomena in images could be due to a power law distribution of objects sizes. In the first part of this thesis, we studied the distribution of connected components of bilevels. We experimentally observed that the length distribution of intersections of those components with lines of fixed direction follows a power law (see Section 3). Therefore, eventhough the correspondence between our notion of objects and real objects is far from beeing one to one, our observations confirm the same hypothesis on the power law distribution of objects sizes, which is compatible with the experiments. It must be noticed, however, that the independance assumption on objects disribution is in contradiction with physical interactions between objects.

Moreover, our experiments on natural images gave values of δ around 3 (the experimental density functions for the distribution of intercepts have an exponent γ slightly below 3, see Sections 3.5), for many images, as long as they are not too textured (this is not the case for the baboon image for instance, for which the α exponent was already clearly bigger than 2). This is once more related to scale invariance. Indeed, the limiting case $\delta = 3$ corresponds to a log behavior of the function V (still assuming R_{max}/l is large), which, as we saw in Section 2.3.1, corresponds to the scale invariant case. We must admit, however, that in this case, the experimental value of γ is varying much more from images to images than the ones of α the exponent of the area distribution.

5.6 The size distribution of the relief grains

In this section we are interested in the size distribution of the grains which are completely visible after the occlusion process. More formally, those grains are the $x_i + X_i$ such that

$$x_i + X_i \cap x_j + X_j = \emptyset, i < j \leq 0.$$

Following G. Matheron, those grains will be called relief grains. As before, we consider a dead leaves model on a window Ω, with convex grains X_0. We further assume that $X_0 \subset D(0, R)$ with probability one, and to avoid edge effects we will suppose that the centers are uniformly picked in $\Omega^{2R} = \Omega \oplus D(0, 2R)$. Thus the result of Lemma 2 is valid for any compact K such that $K \ominus D(0, R) = \emptyset$ and $K \cap \Omega \neq \emptyset$, which in particular is the case for any leaf that touches Ω. Moreover, we suppose that X_0 has area A and perimeter U following the density $f(A, U)$. Eventually, we assume that X_0 is isotropic, which will enable us to apply Steiner formula.

Theorem 9 *For a dead leaves model satisfying the preceding assumptions, thus with grains X_0 having area A and perimeter U with density $f(A, U)$, the relief grains touching Ω follows the density*

$$f_r(A, U) = C \frac{f(A, U)}{E\nu(X_0) + A + \frac{EU(X_0)U}{2\pi}},$$

where C is a normalizing constant.

Proof The probability that X_i is centered in Ω, has area between A and $A+dA$, perimeter between U and $U+dU$, and is not touched by any of the following leaves is

$$\frac{\nu(\Omega)}{\nu(\Omega^{2R})}f(A,U)dAdU\left(1-\frac{E(\nu(X_i\oplus X_0)|X_i \text{ has area } A \text{ and perimeter } U)}{\nu(\Omega^{2R})}\right)^{-i},$$

which by Steiner formula is

$$\frac{\nu(\Omega)}{\nu(\Omega^{2R})}f(A,U)dAdU\left(1-\frac{E\nu(X_0)+A+\frac{UEU(X_0)}{2\pi}}{\nu(\Omega^{2R})}\right)^{-i}.$$

We sum this relation over all i's and get the density function of the relief grains up to a normalizing constant, which proves the theorem. □

Those results about relief grains are computable because we know how to get the probability for a compact to be avoided by a leave. On the other hand, we don't know how to get the probability for a compact to be completely overlapped by the subsequent leaves, since this may be achieved in several stages, and thus don't know how to compute the proportion of leaves that disappear completely.

5.7 The average area after occlusion: numerical experiments

As we saw on the three examples of Figures 10-12, both the shape of the grain and its size distribution have a strong influence on the visual flavour of the dead leave model. In particular, the isoperimetric ratio of the grain seems to be of great importance. Contrarily to the average intercept, which is reduced to its half by occlusion whatever the parameters of the grain are, the average area after occlusion seems to follow more complicated rules. This quantity appeared to us as essential, since it gives a partial answer to the question: "how much do we see of objects?". In this section, we present some experimental results concerning the ratio between the average area of the the grain X_0, and the same quantity for the cells of the dead leaves model. We also checked, in a sense, the influence of discretization in our simulations by checking that the average intercept length for the cells was twice as small as the one for the grain X_0. This experimental quantity is also of interest when the shape is not convex, in which case we do not have a theoretical result. We denote

$$R_S=\frac{\text{average area of the cells of } M}{\text{average area of } X_0},$$

and

$$R_L=\frac{\text{average length of intercepts for the cells of } M}{\text{average length of intercepts for } X_0}.$$

experiment	R_S	R_L
disks, same radius	0.257	0.51
disks, uniform radius	0.283	0.51
disks, exponential radius	0.339	0.52
horizontal squares	0.275	0.51
rotated squares	0.23	0.51
elongated rectangles	0.08	0.53
arrows	0.10	0.52
T's	0.061	0.54

Table 10: R_S and R_L in various cases. R_S is the ratio of the average area after and before occlusion. R_L is the ratio of the average intercept length after and before occlusion (predicted by the theory to be 0.5)

The simulations of the model consist in sequentially superposing discrete shapes whose size follows a certain probability law, and whose centers are uniformly drawn in the image, numerous enough to cover all of it. We worked on 900×900 images. The averages for the grain X_0 where computed in the discrete case. In order to reduce boundary influence in the computations of the ratios, we only take into account cells falling in a small enough subimage. We only studied isotropic models (except for the horizontal squares), and compute only segments in the horizontal direction. Moreover, each experiment consists of 50 such images over which the quantity under study is averaged.

We present the results in the following cases:

- disks with radius of 20 pixels (diameter of 41 pixels),
- disks with radius uniformly distributed between 15 and 40 pixels,
- disks with radius exponentially distributed between 5 and 50 pixels,
- squares with sides of 41 pixels parallel to the border of the image,
- squares with sides of 41 pixels, uniformly rotated,
- elongated rectangles of size 108×12, uniformly rotated,
- arrows of fixed size, uniformly rotated,
- T's, uniformly rotated.

We present, in Table 10, R_L and R_S for those experiments.

We notice that the above experimental results confirm the expected value of the ratio $R_L = 0.5$ for the average length of segments, and therefore the validity of our protocol. Moreover, it seems that the result remains true for nonconvex shapes. Concerning the average area ratio R_S, and contrarily to the constant value of R_L, the values are strongly dependent on two parameters: the size probability law, and the shape of the grain. We basically observe two phenomena. The first one is that at constant shape for the grain, the more numerous the small objects, the bigger R_S. This is illustrated in the case of disks with three different size laws, and has been confirmed by other experiments. Secondly, the value of the isoperimetric constant for the grains seems to have a strong influence on R_S. It seems that elongated or nonconvex shapes lead to a small value of R_S. We claim that there must be a link between the decay of the average area by occlusion, and the decay of the average perimeter. This interaction between area and perimeter has already been observed in the case of the relief grains studied in the preceding section.

6 Appendix: Matheron's Dead Leaves Model

In this appendix we present the dead leaves model as it was originally developed by G. Matheron ([22]), on the whole of $\mathbb{R}^2$. The main difference with our simplified model is that it is not possible to draw a single point randomly in $\mathbb{R}^2$, so that we need to draw infinitely many points at each stage. On the other hand, if we draw several leaves at a time, we need them to be disjoint to define the occlusion process. This may be achieved using an infinitesimal point process.

6.1 Poisson point processes

Define P the set of all locally finite sequences of points $\mathbf{x} = (x_i)_{i\in\mathbb{N}}$ of $\mathbb{R}^n$. With each $\mathbf{x} \in P$ we associate the map $B \to \mathrm{Card}(\mathbf{x} \cap B)$, defined from the set of Borel sets of $\mathbb{R}^n$ into $\mathbb{N}$. We then define N, the smallest σ-algebra such that $\mathbf{x} \to \mathrm{Card}(\mathbf{x} \cap B)$ is measurable for each Borel set B. We define a point process as any measurable map into (P, N). Elements of P are closed sets of $\mathbb{R}^n$, since they are discrete sets. It is easily shown that point processes are in fact random closed sets. Indeed, for every compact set K,

$$\mathbf{x}^{-1}(\mathcal{F}_K) = \{\omega \in \Omega / Card(\mathbf{x}(\omega) \cap K) = 0\},$$

which proves that $\mathbf{x}^{-1}(F_K)$ is measurable and $\mathbf{x}$ a random closed set. Therefore such a process is characterized by its capacity function. The interested reader may consult [8], [9].

Definition 3 *Let μ be a measure on $\mathbb{R}^2$. We call Poisson process (p.p.p.) associated with μ any point process $\mathbf{x}$ satisfying*

1. *If $A_1 ... A_n$ are μ-measurable disjoint sets, then $\phi(A_1) .... \phi(A_n)$ are independent random variables.*
2. *For every bounded Borel set B, $Card(B \cap \mathbf{x})$ obeys a Poisson law with parameter $\mu(B)$, i.e. for any k in $\mathbb{N}$,*

$$Pr(Card(\mathbf{x} \cap B)) = k) = \frac{\mu(B)^k}{k!} \exp(-\mu(B)). \tag{29}$$

Theorem 10 *Let μ be a nonatomic measure (that is to say $\mu(x) = 0$ for x in $I\!R^2$), finite on all bounded sets. Then there exists a Poisson point process associated with μ.*

We omit the proof of this result, and refer to [17].We now list some properties of Poisson processes.

- If a p.p.p. is stationary, then $\mathbf{x}$ is associated with the measure $\lambda\nu$ where λ is a positive real number and ν denotes the Lebesgue measure. $\mathbf{x}$ is called Poisson point process with parameter λ. In that case, Property 2 of Definition 3 is entailed by Property 1.

- Let $\mathbf{x}_1$ and $\mathbf{x}_2$ be two p.p.p. associated respectively with μ_1 and μ_2. Then $\mathbf{x}_1 \cup \mathbf{x}_2$ is a p.p.p. associated with $\mu_1 + \mu_2$. This result follows from the following result, whose proof will be omitted for too technical: for every compact set K, $\mathbf{x}_1 \cap \mathbf{x}_2 \cap K$ is empty with probability 1. It is plain that $\mathbf{x}_1 \cup \mathbf{x}_2$ satisfies Properties 1 and 2 of Definition 3, and the sum of two Poisson laws with parameters α and β is a Poisson law with parameter $\alpha + \beta$.

- Let $\mathbf{x}$ be a p.p.p. associated with μ and p a μ-measurable real function defined on $I\!R^2$ such that $0 \leq p(x) \leq 1$ for every x, then we define a new p.p.p. $\mathbf{x}'$ as the set of points of $\mathbf{x}$ surviving with probability $p(x)$, independently of each other. Then $\mathbf{x}'$ is a p.p.p. associated with the new measure

  $$\mu'(B) = \int_B p(x) d\mu(x).$$

  We omit the proof of this result, and refer to [17].

6.2 The boolean model

We consider a particular instance of random closed set of $I\!R^n$. Let $\mathbf{x} = (x_i)_{i \in I\!N}$ be a stationary p.p.p. with parameter λ and $(X_i)_{i \in I\!N}$ a sequence of independent identically distributed closed random sets which are independent of the (x_i)'s. In addition, we assume

$$E(\nu(X_0 \oplus K)) < +\infty, \text{ for every} K \in \mathcal{K}, \tag{30}$$

where ν is the Lebesgue measure in $I\!R^2$ and E is the expectation with respect to the image measure by X_0. This implies, among other properties, that X_0 is almost surely bounded.

Definition 4 *We define a boolean model with intensity λ and grain X_0 and denote by $B(\lambda, X_0)$ the closed random set $\cup_i (x_i + X_i)$*

The closedness of $B(\lambda, X_0)$ is a consequence of Condition (30), which ensures that each $x_i + X_i$ only meets a finite amount of other $x_j + X_j$'s, with probability 1. The capacity function associated with B is remarkably simple.

Theorem 11 *Let $X = B(\lambda, X_0)$ be a boolean model and G its capacity function, then for any compact set K:*

$$G(K) = Pr(X \cap K = \emptyset) = \exp\left(-\lambda E\nu(X_0 \oplus \check{K})\right). \tag{31}$$

Proof Let $K \in \mathcal{K}$ and $\mathbf{x}'$ defined by $\mathbf{x}' = \{x_i \in \mathbf{x} / (x_i + X_i) \cap K \neq \emptyset\}$. Then $\mathbf{x}'$ also is a Poisson process associated with $\mu'(B) = \lambda \int_B Pr(x + X_0 \cap K \neq \emptyset) dx$, by Formula (6.1). Then

$$\begin{aligned}
Pr(X \cap K = \emptyset) &= Pr(\mathbf{x}' = \emptyset) \\
&= \exp\left(-\lambda \int_{\mathbb{R}^2} Pr(x + X_0 \cap K \neq \emptyset) dx\right) \\
&= \exp\left(-\lambda \int_{\mathbb{R}^2} Pr(x \in \check{X}_0 \oplus K) dx\right) \\
&= \exp\left(-\lambda \int_{\mathbb{R}^2} E(\mathbb{1}_{\check{X}_0 \oplus K}(x)) dx\right) \\
&= \exp\left(-\lambda E \int_{\mathbb{R}^2} \mathbb{1}_{\check{X}_0 \oplus K}(x) dx\right) \\
&= \exp\left(-\lambda E\nu(\check{X}_0 \oplus K)\right),
\end{aligned}$$

where $\mathbb{1}_C(x)$ is 1 if $x \in C$ and 0 otherwise. Since $\check{X}_0 \oplus K = -X_0 \oplus \check{K}$, we obtain the announced result. $\square$

Remark: The quantity $Pr(K \subset X)$ is not known in general. It is even difficult to obtain approximations.

6.3 The dead leaves model in $\mathbb{R}^2$

Let us consider $(X_t)_{t \in (-\infty, 0]}$ a family of independent random compact sets satisfying Condition (30) and a continuous function $\lambda(t)$. For $-\infty < t \leq 0$, we consider the infinitesimal boolean model $B(\lambda(t)dt, X_t)$. We superpose the boolean models $B(\lambda(t)dt, X_t)$ from $-\infty$ to 0, and define the dead leaves model $F(\lambda(t), (X_t))$ as the boundaries of the sets obtained by such a superposition procedure. That is to say we keep the boundaries of $F(\lambda(t), (X_t))$ that do not intersect any of the $F(\lambda(t'), (X_{t'}))$ for $t' \in (t, 0]$. For the sake of simplicity, we assume that $\lambda(t) = \lambda$ and $X_t = X$. We remark that in view of tridimensional scenes modeling, the t parameter permits to represent perspective effects. In addition we assume that X has almost surely non empty interior. It can be proved that $F(\lambda, X)$ is a closed random set, and that the superposition of the boolean models from $-\infty$ to 0 covers the whole space $\mathbb{R}^2$. Indeed, any small enough compact set K is included with probability 1 in the superposition of the boolean models: if K is a compact set such that $E\nu(X \ominus \check{K})$ is not zero (which will be the case for sufficiently small K if X has non empty interior), then the probability that K is included in some $x_i + X_i$ at some time t' in $[t, t + dt]$ is $\lambda dt E\nu(X \ominus \check{K})$, (as is proved below). Therefore, as this probability is nonzero, K is included in a grain of the model for some time t with probability one.

Thus, the model is a closed set because it is a locally finite union of closed sets. As in the bounded case, we have:

Theorem 12 *Let $\lambda \in I\!R$ and X be a random compact set satisfying Condition (30). Let us define $Q(K)$ as the probability of the compact set K to be included in a connected component of the partition of $I\!R^2$ provided by the dead leaves model. Then we have:*

$$Q(K) = \frac{E\nu(X \ominus \check{K})}{E\nu(X \oplus \check{K})}. \tag{32}$$

Sketch of proof: K is included in a connected component if for some time $t \in (-\infty, 0)$ K is included in a grain of $B(\lambda dt, X)$, and K does not intercept any of the grains $B(\lambda dt, X)$, from t to 0. If the boolean model at time t is given by $\cup(x_i + X_i)$ then

$$Pr((x_i + X_i) \cap (x_j + X_j) \neq \emptyset, i \neq j) = \circ(dt),$$

and (31) provides

$$\begin{aligned} Pr(K \subset K(\lambda dt, X)) &= Pr(K \subset (x_i + X_i), \text{ for some i}) + \circ(dt) \\ &= \lambda dt E\nu(X \ominus \check{K}) + \circ(dt). \end{aligned}$$

Moreover, the probability that K does not intercept the boolean model from t to 0 is $\exp(-\lambda t E\nu(X \oplus \check{K}))$. Indeed, the union of several boolean models is a boolean model whose parameter is the sum of the parameters of the initial boolean models. Therefore, as the events described above are independent, we obtain

$$Q(K) = \int_{-\infty}^{0} \lambda E\nu(X \ominus \check{K}) \exp(-\lambda t E\nu(X \oplus \check{K}) dt,$$

and therefore

$$Q(K) = \frac{E\nu(X \ominus \check{K})}{E\nu(X \oplus \check{K})}. \tag{33}$$

□

Acknowledgments

We would like to thank Vicent Caselles, Stuart Geman, Stéphane Mallat and David Mumford for several helpful comments and discussions. We also thank Les Treilles foundation which made some of these discussions possible.

References

[1] L. Alvarez, Y. Gousseau, and J.-M. Morel. The size of objects in natural and artificial images. to be published in Advances in Imaging and Electron Physics, 1999.

[2] R. J. Baddeley. The correlational structure of natural images and the calibration of spatial representations. *Cognitive Science*, 21:351–372, 1997.

[3] H. B. Barlow. The coding of sensory messages. In W. H. Thorpe and O. L. Zangwill, editors, *Current Problems in Animal Behavior*, pages 331–360. Cambridge University Press, 1961.

[4] A. J. Bell and T. J. Sejnowski. The "independent components" of natural scenes are edge filters. *Vision Research*, 37:3327–3338, 1997.

[5] G. J. Burton and I. R. Moorhead. Color and spatial structure in natural scenes. *Applied Optics*, 26:157–170, 1987.

[6] V. Caselles and J.-M. Morel. The connected components of sets of finite perimeters in the plane. Preprint, 1998.

[7] A. Cohen, R. DeVore, P. Petrushev, and H. Xu. Non linear approximation and the space $BV(\mathbb{R}^2)$. Preprint, submitted to Amer. J. Math., 1998.

[8] D. R. Cox and V. Isham. *Point Processes.* Chapman and Hall, 1980.

[9] D. J. Daley and D. Vere-Jones. *An Introduction to the Theory of Point Processes.* Springler Verlag, 1988.

[10] N. G. Deriugin. The power spectrum and the correlation function of the television signal. *Telecomunications*, 1:1–12, 1956.

[11] L. C. Evans and R. F. Gariepy. *Measure Theory and Fine Properties of Functions.* Studies in Advanced Math. CRC Press, 1992.

[12] D. J. Field. Relations between the statistics of natural images and the response properties of cortical cells. *Jour. Opt. Soc. Amer.*, A4:2379–2394, 1987.

[13] E. De Giorgi. Su una teoria generale della misura $(r-1)$-dimensionale in uno spazio ad r dimensioni. *Ann. Mat. Pura ed Appl., IV. Ser.*, 36:191–213, 1954.

[14] J. Huang and D. Mumford. Statistics of natural images and models. *To be published*, 1998.

[15] D. Jeulin. Dead leaves models: From space tesselation to random functions. In D. Jeulin, editor, *Advances in Theory and Applications of Random Sets*, 1996.

[16] G. Kanisza. *Organization in Vision: Essays on Gestalt Perception*. Praeger, 1979.

[17] J. F. C. Kingman. *Poisson Processes.* Oxford Studies in Probability. Clarendon Press, 1993.

[18] E. R. Kretzmer. Statistics of television signals. *Bell Syst. Tech. Journal*, 31:751–763, 1952.

[19] S. Mallat. Personal communication.

[20] S Mallat. *A Wavelet Tour of Signal Processing.* Academic Press, 1997.

[21] G. Matheron. *Eléments pour une Théorie des Milieux Poreux.* Masson, Paris, 1967.

[22] G. Matheron. Modèle séquentiel de partition aléatoire. Technical report, CMM, 1968.

[23] G. Matheron. *Random Sets and Integral Geometry.* John Wiley & Sons, New York, 1975.

[24] Y. Meyer. *Wavelets: Algorithms and Applications.* SIAM, 1993.

[25] D. Mumford, S. C. Zhu, and B. Gidas. Stochastic models for generic images. Working draft, 1997.

[26] B. A. Olshausen and D. J. Field. Natural images statistics and efficient image coding. *Network: Computation in Neural Systems*, 7:333–339, 1996.

[27] D. L. Ruderman. The statistics of natural images. *Network: Computation in Neural Systems*, 5:517–548, 1994.

[28] D. L. Ruderman. Origins of scaling in natural images. *Vision Research*, 37(23):3385–3398, 1997.

[29] D. L. Ruderman and W. Bialek. Statistics of natural images: Scaling in the woods. *Physical Review Letters*, pages 814–817, 1994.

[30] L. Rudin. *Images, Numerical Analysis of Singularities and Shock Filters.* PhD thesis, California Institute of Technology, 1987.

[31] L. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D*, 60:259–268, 1992.

[32] J. Serra. *Image Analysis and Mathematical Morphology.* Academic Press, London, 1982.

[33] D. Stoyan, W. S. Kendall, and J. Mecke. *Stochastic Geometry and its Applications.* Wiley Series in Probability and Statistics. Wiley, 2nd edition, 1995.

[34] D. Tolhurst, Y. Tadmor, and T. Chao. Amplitude spectra of natural images. *Ophthal. Physiol. Opt.*, 12:229–232, 1992.

[35] R. F. Voss and J. Clarke. $\frac{1}{f}$ noise in music and speech. *Nature*, 258:317–318, 1975.

[36] R. F. Voss and J. C. Y. Wyatt. Multifractals and the local connected fractal dimension: Classification of early chinese landscape painting. In H. Jones A.J. Crilly, R.A. Earnshaw, editor, *Applications of Fractals and Chaos*, pages 171–184. Springer-Verlag, 1993.

[37] W. P. Ziemer. *Weakly Differentiable Functions.* Springer Verlag, 1989.