MARKOV POINT PROCESSES AND THEIR APPLICATIONS IN HIGH-LEVEL IMAGING

M N M van Lieshout[1]

Department of Statistics, University of Warwick
Coventry CV4 7AL, United Kingdom

1 Introduction

The pioneering work by Geman & Geman [21] and Besag [8] stimulated a surge of interest in statistical approaches to image analysis. Until recently, most attention has been given to segmentation or classification tasks, i.e. dividing an image into relatively homogeneous regions of different type. Following [8, 21], a Bayesian approach is usually taken in which a prior Markov random field model is used to impose smoothness on segmentations.

In the computer vision literature, segmentation is considered as a 'low-level' task, calling for local operations on pixel neighbourhoods and converting the input image into another raster image. In 'high-level' problems (including object recognition and scene analysis) we have to interpret the image globally, reducing it to a compact description (e.g. a vector graphics representation) of the scene.

Baddeley & Van Lieshout [2, 3, 4, 30, 31, 32], and Ripley and coauthors [37, 47, 49] have argued that the (continuous) Markov or Gibbs processes studied in stochastic geometry, spatial statistics and statistical physics [6, 42, 45, 46, 48] provide a rich collection of models usable in a broad range of problems involving the higher-level interpretation of spatial and image data. They are of simple mathematical form, interpretable in terms of interaction between objects and amenable to iterative statistical techniques and Markov chain Monte Carlo.

The subclass most widely studied is that of pairwise interaction models [18, 38, 39, 40, 43, 45, 54]. These provide a flexible class for negative association between neighbouring objects, but they do not seem to be able to model clustered patterns and Møller [36] has argued that nearest-neighbour Markov models [6] are better suited to this task.

In support of this claim, Baddeley et al. [7] proved that many cluster processes with bounded non-empty clusters fall within the class of nearest-neighbour Markov point processes. These results suggest that nearest-neighbour Markov processes may be suitable multiple-generation cluster models [29] and help to explain why statistical inference for Poisson cluster processes based on interpoint distances [33] bears so close a resemblance to that for Markov point processes.

On the other hand, the simpler Markov models [48] also allow clustered patterns if interactions between more than two objects are permitted. Baddeley & Van Lieshout [5] proposed a model that can exhibit both clustering and inhibition according to the value of a single parameter. The model has interactions of arbitrary high order and is closely related to the empty space function [18]. A special case is known in the statistical physics literature as the 'penetrable sphere model' [55].

To illustrate the role of Markov spatial processes in image analysis, consider object recognition [2, 4]. Here the aim is to decide whether there are any objects of a specified kind in the image and if so to determine the number of objects and their locations and characteristics. Applications include industrial robot vision, document reading, interpretation of medical scans [12], automated cytology [24], classification of astronomical features [37, 49] and identification of grain structures in materials

[1] includes joint work with A.J. Baddeley and J. Møller

science.

A strong motivation for a prior in this context is 'multiple response' in maximum likelihood solutions, which tend to detect many almost identical objects. This is familiar from computer vision, where it is recommended to select one object from each cluster, but this is similar to a Bayesian approach with an *inhibitive* prior.

Alternative approaches have been described in recent studies [14, 17, 24, 41] on recognising the shape of an interesting object (hand, galaxy, mitochondrion). The shape is described by a flexible template, typically a polygon, with edge lengths and angles governed by a joint prior distribution, typically a Markov chain.

The plan of this paper is as follows. In Section 2 we survey Markov object processes, turning to a discussion of several qualitatively different types of interaction in Section 3. Section 4 briefly describes spatial Markov properties. In Section 5 we discuss the application of Markov processes as a prior distribution in object recognition and the paper closes with a discussion on computational aspects.

2 Survey of Markov spatial models

2.1 Objects

The objects featuring in stochastic geometry range from simple geometrical figures (points, lines, discs) through plane polygons and convex compact sets to completely general closed sets. In particular, the space $\mathcal{F}$ of all closed subsets of $\mathbb{R}^d$ can be made into a locally compact, second countable Hausdorff space (l.c.s. space) so that a random closed set can be defined as a random element of $\mathcal{F}$ [34].

We are mainly concerned with simple geometrical figures that can be specified by the values of a few real-valued parameters giving location, orientation, size etc. For example a disc in $\mathbb{R}^2$ can be specified by its centre (x, y) and radius r. A given class of objects U is treated as a space in its own right, so that objects are regarded as points in U. For the disc example $U = \mathbb{R}^2 \times \mathbb{R}^+$.

It is often useful to represent an object as a 'marked point', that is a pair (s, m) consisting of a point $s \in \mathbb{R}^d$ and a 'mark' $m \in \mathcal{M}$, where $\mathcal{M}$ is an arbitrary l.c.s. space. The point s fixes the location of the object and the mark m contains all other information such as size and shape. A disc in $\mathbb{R}^2$ can be regarded as a point (x, y) marked by a radius r. Objects with additional properties such as colour and surface texture can be represented as marked points by choosing an appropriate mark space $\mathcal{M}$. For example a grey-scale surface texture can be formalised as an upper-semicontinuous function $\mathbb{R}^d \to \mathbb{R}^+$, and the space of all such functions is l.c.s.

2.2 Markov object processes

An *object configuration* is a finite un-ordered set $\mathbf{x} = \{x_1, \ldots, x_n\}$ of objects $x_i \in U$. Writing Ω for the set of all configurations (the exponential space of U), a random object process is a random element of Ω, or equivalently, a point process on U.

The basic reference model is the *Poisson object process* in U. Let μ be a finite non-atomic measure on U. Then, under the Poisson model, the total number of objects has a Poisson distribution with mean $\mu(U)$; given that exactly n objects are present, they are independent and uniformly distributed in U, i.e. $\mathbb{P}(x_i \in B) = \mu(B)/\mu(U)$ for all measurable $B \subseteq U$. In particular, there are no interactions between objects. Further details can be consulted in [54].

To construct spatial processes that do exhibit dependence between 'neighbouring' objects, we specify the probability density of the new process with respect to the Poisson process (thereby restricting attention to processes that are absolutely continuous with respect to the Poisson model). A density

is a measurable and integrable function $p : \Omega \to [0, \infty)$. For the new process, the distribution of the total number of objects is

$$\mathbb{P}(N = n) = \frac{e^{-\mu(U)}}{n!} \int_U \cdots \int_U p(\{x_1, \ldots, x_n\})\, d\mu(x_1) \ldots d\mu(x_n).$$

Writing $q_n = \mathbb{P}(N = n)$, given $N = n$, the n random objects have joint probability density

$$p_n(x_1, \ldots, x_n) = e^{-\mu(U)} \mu(U)^n p(\{x_1, \ldots, x_n\})/(n!\, q_n)$$

with respect to the distribution of n independent uniformly distributed objects in U.

To introduce interactions, let $\sim$ be any symmetric, reflexive relation on U. For instance two objects represented by $u, v \in U$ are 'neighbours' if their intersection is non-empty. A widely used class is that of *pairwise interaction models*

$$p(\mathbf{x}) = \alpha \beta^{n(\mathbf{x})} \prod_{x_i \sim x_j} g(x_i, x_j). \tag{1}$$

Here $\alpha, \beta > 0$ are constants, $n(\mathbf{x})$ is the number of objects in $\mathbf{x}$, and $g : U \times U \to [0, \infty)$ the interaction function. The product is over all pairs of neighbouring objects $x_i \sim x_j$ with $i < j$. Note that for every proposed $g(\cdot, \cdot)$ we have to verify that the model is well-defined and integrable!

The density (1) embraces some interesting special cases. If $g \equiv 1$ then (1) is simply a Poisson process with intensity measure $\beta\mu(\cdot)$; if $g \equiv 0$ it is a *hard object process*.

The intermediate case $g \equiv \gamma$ for a constant $0 < \gamma < 1$ is called a *Strauss object process* and the density can be written

$$p(\mathbf{x}) = \alpha \beta^{n(\mathbf{x})} \gamma^{s(\mathbf{x})} \tag{2}$$

where

$$s(\mathbf{x}) = \sum_{i<j} 1\{x_i \sim x_j\}$$

is the number of pairs of neighbouring objects (e.g. number of overlaps) in the configuration. This process is 'ordered' or 'regular', since $s(\mathbf{x})$ tends to be smaller than under the Poisson model. For $\gamma > 1$ the density (2) is typically not integrable.

Note that if $u \in U, u \notin \mathbf{x}$ with $p(\mathbf{x}) > 0$, the ratio

$$\frac{p(\mathbf{x} \cup \{u\})}{p(\mathbf{x})} = \beta \prod_{x_i \sim u} g(u, x_i), \tag{3}$$

comparable to the local characteristics of a Markov random field, *depends only on u and on the neighbours of u in $\mathbf{x}$*. This important property signifies that all interaction is 'local'.

In the statistical physics interpretation, $-\log p(\mathbf{x} \cup \{u\}) + \log p(\mathbf{x})$ is the energy required to add a new object u to an existing configuration $\mathbf{x}$. In probabilistic terms $p(\mathbf{x} \cup \{u\})/p(\mathbf{x})$ is the Papangelou conditional intensity $\lambda(u; \mathbf{x})$ at u given the rest of the pattern $\mathbf{x}$ on $U \setminus \{u\}$, see [16]. Roughly speaking, $\lambda(u; \mathbf{x})du$ is the conditional probability of an object in the infinitesimal region du centred at u given the configuration agrees with $\mathbf{x}$ outside this region.

The discussion above motivates the following definitions and results taken from Ripley & Kelly [48].

Definition 1 (Ripley and Kelly) *A random object process X with density p is called a* Markov object process *with respect to $\sim$ if for all $\mathbf{x} \in \Omega$*

(a) $p(\mathbf{x}) > 0$ *implies* $p(\mathbf{y}) > 0$ *for all* $\mathbf{y} \subseteq \mathbf{x}$;

(b) *if* $p(\mathbf{x}) > 0$, *then*

$$\frac{p(\mathbf{x} \cup \{u\})}{p(\mathbf{x})} \tag{4}$$

depends only on u *and* $N(\{u\}) \cap \mathbf{x} = \{x_i \in \mathbf{x} : u \sim x_i\}$.

An analogue of the celebrated Hammersley–Clifford theorem also holds [48]: a process with density $p : \Omega \to [0, \infty)$ is Markov iff

$$p(\mathbf{x}) = \prod_{\text{cliques } \mathbf{y} \subseteq \mathbf{x}} q(\mathbf{y}) \tag{5}$$

for all $\mathbf{x} \in \Omega$, and $q : \Omega \to [0, \infty)$ is an (arbitrary) interaction function. The product is over all *cliques*, that is configurations $\mathbf{y} \subseteq \mathbf{x}$ in which all members are neighbours. By convention the empty set and single object configurations are cliques.

2.3 Nearest-neighbour Markov object processes

A further extension due to Baddeley and Møller [6] is to allow interaction behaviour to depend on the realisation of the process. For example, in a one-dimensional renewal process, each point can be said to interact with its nearest neighbours to the left and right, regardless of how far distant these neighbours may be.

In higher dimensions we are particularly interested in the *connected component relation* [6] in which $x_i \underset{\mathbf{x}}{\sim} x_j$ if the two objects x_i, x_j belong to the same connected component in the union set of all objects in $\mathbf{x}$:

$$x_i \underset{\mathbf{x}}{\sim} x_j \quad \text{iff} \quad x_i \sim z_1 \sim \cdots \sim z_n \sim x_j \tag{6}$$

for some path $z_1, \ldots, z_n \in \mathbf{x}$ of mutually overlapping objects.

In general, assume that for each configuration $\mathbf{x}$ we have a symmetric reflexive relation $\underset{\mathbf{x}}{\sim}$ defined on $\mathbf{x}$. The reader might prefer to think of this as a finite graph whose vertices are the objects $x_i \in \mathbf{x}$.

Definition 2 (Baddeley and Møller) *A random object process with density* p *is called a* nearest-neighbour Markov object process *with respect to* $\{\underset{\mathbf{x}}{\sim} : \mathbf{x} \in \Omega\}$ *if, for all* $\mathbf{x}$ *with* $p(\mathbf{x}) > 0$

- $p(\mathbf{y}) > 0$ *for all* $\mathbf{y} \subseteq \mathbf{x}$;
- *the ratio* $\frac{p(\mathbf{x} \cup \{u\})}{p(\mathbf{x})}$ *depends only on* u, *on* $N(\{u\} \mid \mathbf{x} \cup \{u\}) \cap \mathbf{x} = \{x_i \in \mathbf{x} : x_i \underset{\mathbf{x} \cup \{u\}}{\sim} u\}$ *and on the relations* $\underset{\mathbf{x}}{\sim}, \underset{\mathbf{x} \cup \{u\}}{\sim}$ *restricted to* $N(\{u\} \mid \mathbf{x} \cup \{u\}) \cap \mathbf{x}$.

The appropriate Hammersley-Clifford theorem [6] states that a process with density $p(\cdot)$ is nearest-neighbour Markov iff

$$p(\mathbf{x}) = \begin{cases} \prod_{\text{cliques } \mathbf{y} \subseteq \mathbf{x}} q(\mathbf{y}) & \text{if } q(\mathbf{y}) > 0 \text{ for } \textit{all } \mathbf{y} \subseteq \mathbf{x} \\ 0 & \text{otherwise} \end{cases} \tag{7}$$

where $q : \Omega \to \mathbb{R}_+$ satisfies certain regularity conditions. Here a subset $\mathbf{y} \subseteq \mathbf{x}$ is a *clique in* $\mathbf{x}$ if all members of $\mathbf{y}$ are $\mathbf{x}$-neighbours of one another ($u \underset{\mathbf{x}}{\sim} v$ for all $u, v \in \mathbf{y}$).

On a practical note, the normalising constant α in (1) and other Markov processes is usually not available in closed form, thus prohibiting direct sampling and estimation. However, the simple ratios (4) make Markov processes amenable to Markov chain Monte Carlo [9, 53]. The classical approach is via spatial birth-and-death processes [6, 42, 35, 32, 43] but recently Metropolis-Hastings algorithms of various types [22] have been proposed. For an excellent recent survey, see [36]. An overview of parameter estimation techniques can be found in [19] or [45].

3 Inhibition versus attraction

As mentioned in Section 2, for each new object model with density $p(\cdot)$ we have to verify that $p(\cdot)$ is integrable. This is straightforward for two types of models: those imposing an upper bound on the number of objects and those for which all interaction functions are bounded above by 1 (*purely inhibitory models*). The first class includes for instance hard object processes; an example of the second type is the inhibitory Strauss density (2).

For attractive patterns integrability seems more problematic, e.g. the Strauss process is not integrable for values of $\gamma > 1$. Conditioning on the number of objects does yield a well-defined model, but recent simulation experiments [36] suggest an abrupt transition from Poisson-like patterns to tightly clustered patterns rather than exhibiting intermediate, moderately clustered patterns. This confirms theoretical results by Gates and Westcott [20], who showed that partly attractive processes may violate a stability condition, implying that they produce extremely clustered patterns with high probability.

3.1 Cluster processes

The natural model for clustering in stochastic geometry [54] is a two stage process, in which each object ξ in an unobserved parent process $\mathbf{x}$ gives rise to a finite process Z_ξ of daughters. The data $\mathbf{y}$ is the superposition

$$\mathbf{y} = \cup_{x_i \in \mathbf{x}} Z_{x_i}.$$

Interpreted as a marked point process $\{(x_1, Z_{x_1}), \ldots, (x_n, Z_{x_n})\}$, we assume that the marks Z_{x_i} are independent. In addition, we assume that the object space U is equipped with a complete metric d and restrict attention to daughter processes that are absolutely continuous with respect to the Poisson object process (cf. Section 2).

As a simple example, suppose that each parent ξ in a unit rate Poisson process on a compact subset $B \subset \mathbb{R}^d$ generates a Poisson number of offspring with mean ω positioned i.i.d. with probability density $h(\cdot - \xi)$. Let μ be Lebesgue measure restricted to the dilated set $B_{\oplus R} = \{u : d(u, B) \leq R\}$. If $h(\cdot)$ is supported on $B(0, R)$, the density for the total offspring $\mathbf{y} \neq \emptyset$ is

$$f(\mathbf{y}) = \sum_{n=1}^{\infty} \{ \frac{e^{\mu(B_{\oplus R} \setminus B)} \omega^m e^{-\omega n}}{n!} \int_B \cdots \int_B \prod_{j=1}^{m} \left(\sum_{i=1}^{n} h(y_j - x_i) \right) d\mu(x_1) \ldots d\mu(x_n) \}$$

with respect to the distribution of a unit rate Poisson process on $B_{\oplus R}$. This can be factorised as

$$f(\mathbf{y}) = e^{\mu(B_{\oplus R})} \omega^m e^{-m\omega} e^{-\beta} \left[\sum_{C_1, \ldots, C_k} e^{\omega(m-k)} J(\mathbf{y}_{C_1}) \ldots J(\mathbf{y}_{C_k}) \right], \tag{8}$$

where $\beta = (1 - e^{-u})\mu(B)$, the sum is over all *un-ordered* partitions of $\mathbf{y}$ into disjoint *non-empty* subconfigurations, and

$$J(\mathbf{y}_C) = \int_B \prod_{y_j \in \mathbf{y}_C} h(y_j - \xi) d\mu(\xi). \tag{9}$$

Note that $J(\mathbf{y}_C) = 0$ unless $(\mathbf{y}_C)_{\oplus R}$ is connected. In other words, associating with each point y_j a ball of radius R, the process is nearest–neighbour Markov with respect to the connected component relation (6) (assuming a positivity condition holds). Considered as a point process Y on $B_{\oplus R}$, we say that Y is nearest–neighbour Markov with respect to the connected component relation *at distance* $2R$.

More generally, Baddeley et al. [7] proved the following result.

Theorem 1 *Let* $\mathbf{x}$ *be a unit rate Poisson point process on* U *and* $\mathbf{y}$ *a cluster process with parent process* $\mathbf{x}$ *and clusters* $Z_{x_i} \subseteq B(x_i, R)$. *Then* $\mathbf{y}$ *is a nearest-neighbour Markov point process with respect to the connected component relation at distance* $2R$.

The process is in general **not** Markovian in Ripley–Kelly sense. If the parent Poisson process is replaced by a Markov process with respect to

$$u \sim v \text{ iff } d(u, v) \leq r$$

or the connected component relation at range r, the cluster process described above is not nearest-neighbour Markov. Heuristically, this is because parents without offspring can cause interaction by merging of disjoint $\underset{\mathbf{y}}{\sim}$-cliques. If we require each parent to have at least one daughter, the cluster process **is** nearest-neighbour Markov at range $2R + r$ [7].

3.2 Area-interaction processes

As we saw in the previous section, many familiar cluster processes satisfy a connected component Markov property. However, the simpler Ripley–Kelly Markov processes also allow for clustered behaviour [5].

As before, let U be an l.c.s. metric space, equipped with a complete metric d.

Definition 3 (Baddeley and Van Lieshout) *The* area-interaction process *is defined by its density*

$$p(\mathbf{x}) = \alpha \beta^{n(\mathbf{x})} \gamma^{-\nu(S(\mathbf{x}))} \tag{10}$$

with respect to the Poisson model. Here $\beta, \gamma > 0$ *are parameters and* α *is the normalising constant;* ν *is a totally finite regular Borel measure on* U *and*

$$S(\mathbf{x}) = \bigcup_{i=1}^{n} B(x_i, r)$$

is the union of d-balls of radius r *centred at the points of the realisation.*

The model is similar to the Strauss process (2) except that counting the number of neighbour pairs has been replaced by measuring the union set $S(\mathbf{x})$. Hence, considered as an exponential family in β, γ, the sufficient statistic is related to the empty space function rather than to the K-function [18, 44].

Contrary to the Strauss model, however, area-interaction is well-defined (measurable and integrable) for **all** values of $\beta, \gamma > 0$. Even stronger, the model is uniformly absolutely continuous with respect to the distribution of a Poisson process with intensity measure $\beta\mu(\cdot)$, i.e. has uniformly bounded

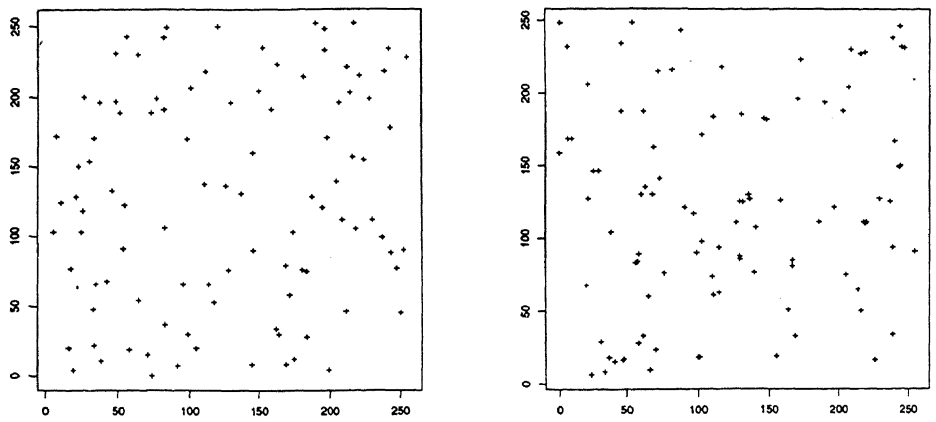

Figure 1: Simulated realisations of an area-interaction process conditional on $n = 100$ points, with $r = 5$ in a window of size 256×256. Left: ordered pattern, $\gamma = 0.9711$, $\gamma^{-25\pi} = 10$; Right: clustered pattern, $\gamma = 1.02975$, $\gamma^{-25\pi} = 0.1$.

Radon-Nikodym derivative (as the measure of $S(\mathbf{x})$ is bounded by $\nu(U)$ uniformly in $\mathbf{x}$). Moreover, (10) satisfies the linear stability condition in [20]. This suggests that the 'singularity' (highly clustered behaviour) of the Strauss model is unlikely. See also Figure 1.

It is intuitively clear that for $0 < \gamma < 1$ the pattern will tend to be 'ordered' and for $\gamma > 1$ 'clustered'. For $\gamma = 1$, (10) reduces to a Poisson process with intensity measure $\beta\mu(\cdot)$. The special case ν is Lebesgue measure and $\gamma > 1$ was introduced by Widom and Rowlinson [55] as a model for liquid-vapour equilibrium in chemical physics. See also [25, 50, 51].

Turning to Markov properties, define $u, v \in U$ to be neighbours whenever $B(u,r) \cap B(v,r) \neq \emptyset$, or equivalently $u \sim v$ iff $d(u,v) \leq 2r$. Then it is readily seen that the area-interaction process (10) is Markov with respect to $\sim$ and exhibits interaction of infinite order:

$$\begin{aligned} q(\emptyset) &= \alpha \\ q(\{a\}) &= \beta\gamma^{-\nu(B(a,r))} \\ q(\{y_1,\ldots,y_k\}) &= \gamma^{(-1)^k \nu(\bigcap_{i=1}^k B(y_i,r))}, \quad k \geq 2. \end{aligned} \tag{11}$$

Various modifications are of interest, for example the balls $B(x_i, r)$ can be replaced by other compact sets $R(x_i)$ depending on x_i [5]. In a parametric statistical model the measure ν and the definition of $R(\cdot)$ might also be allowed to depend on the parameter θ.

Just as the Strauss process is a simple case of pairwise interaction models, area-interaction can be generalised by allowing other interaction 'potentials', for instance,

$$p(\mathbf{x}) = a(\theta) \prod_{i=1}^{n} b(x_i;\theta) \exp\left(-\int_A f(d(\mathbf{x},u))\,du\right)$$

where $d(\mathbf{x},u) = \min_i d(x_i,u)$ and $f : [0,\infty] \to (-\infty,\infty]$. The standard case is $f(t) = \log\gamma\, 1\{t \le r\}$.

Finally, from a more geometrical point of view, replacing 'area' by other fundamental geometrical measurements such as Euler characteristic or perimeter is studied in [1].

4 Spatial Markov properties

A Ripley-Kelly Markov process (Definition 1) has conditional intensities (4) that depend on local information only, a property that can be dubbed 'locally Markov behaviour'. To formulate a 'global' or 'spatial' Markov property, let A be a measurable subset of U. Then the conditional distribution of $X \cap A$ given $X \cap A^c$ depends only on X in the neighbourhood $N(A) \cap A^c = \{u \in A^c : u \sim a \text{ for some } a \in A\}$:

$$\mathcal{L}(X \cap A \mid X \cap A^c) = \mathcal{L}(X \cap A \mid X \cap N(A) \cap A^c). \tag{12}$$

In other words, X restricted to A and X restricted to $N(A)^c$ are conditionally independent given the information in $N(A) \setminus A$ [48].

For generalisations of the spatial Markov property (12) to nearest–neighbour processes (Definition 2), see [28] or [36].

5 Bayesian object recognition

In this Section we would like to indicate how the Markov processes of the previous Sections can be applied in image interpretation problems such as object recognition. The appropriate models here are of the inhibitory type and the exposition below is based on joint work with A.J. Baddeley [2, 3, 4, 30, 31, 32].

Suppose we observe an image of a scene composed of several objects and the task is to locate them. Typical applications are robot vision or the automated reading of documents.

The objects to be recognised are assumed to be representable by a finite number of real parameters that determine size, shape and location. Let U denote the space of possible parameter vector values; a point $u \in U$ represents an object $R(u) \subseteq T$ in the image space T, a (large) array of pixels.

An object configuration $\mathbf{x} = \{x_1, \cdots, x_n\}$, $x_i \in U$ is conveniently modelled as a realisation of an object process on U and induces a set $S(\mathbf{x}) = \bigcup_{i=1}^{n} R(x_i)$ of pixels in image space which we shall call the *silhouette*.

Object recognition can be formulated as a statistical parameter estimation problem by assuming that the image $\mathbf{y}$ depends on the 'true' object pattern $\mathbf{x}$ through a known probability distribution, with density $f(\mathbf{y} \mid \mathbf{x})$. Our image model f consists of a deterministic 'deformation' and a random noise component. Any object configuration $\mathbf{x}$ determines an image $\theta^{(\mathbf{x})}$ in pixel space T, representing the ideal *signal*, which is then corrupted by random noise.

It is convenient to assume that the pixel values y_t are conditionally independent given $\mathbf{x}$. Without loss of generality the conditional distributions of pixel values y_t belong to a family of distributions with densities $\{g(\cdot \mid \theta) : \theta \in \Theta\}$ indexed by a parameter space Θ. Thus

$$f(\mathbf{y} \mid \mathbf{x}) = \prod_{t \in T} g(y_t \mid \theta_t^{(\mathbf{x})}) \tag{13}$$

where $\theta_t^{(\mathbf{x})}$ is the parameter value (signal) at pixel t. This includes additive and multiplicative noise.

As a simple example, consider a 'blur-free' silhouette signal $\theta^{(\mathbf{x})} = S(\mathbf{x})$, in which background pixels are randomly flipped from value 0 to 1 with some fixed probability, and foreground (silhouette)

pixels are unchanged. Let Y be the set of pixels with value 1. The likelihood is nonzero only when $S(\mathbf{x}) \subseteq Y$ and the log likelihood is then a linear function of $|S(\mathbf{x})|$, the total area of the silhouette. Hence, one solution of the maximum likelihood equations is

$$\begin{aligned} \hat{\mathbf{x}}_{\max} &= Y \ominus R \\ &= \{u \in U : R(u) \subseteq Y\}. \end{aligned}$$

Note that this is a generalised erosion operator [52]. This $\hat{\mathbf{x}}_{\max}$ is the largest solution; the other solutions are the subsets $\mathbf{x} \subseteq \hat{\mathbf{x}}_{\max}$ with the same silhouette,

$$S(\hat{\mathbf{x}}) = S(\hat{\mathbf{x}}_{\max}).$$

Thus, performing an erosion to extract features is equivalent to maximum likelihood for a simple noise model.

From this example, we also see that multiple response can occur due to occlusion: maximum likelihood solutions $\mathbf{x}$ tend to contain clusters of almost identical objects. This could well be undesirable and suggests a Bayesian approach with a prior distribution penalising scenes with many overlapping objects. Hence inhibitory Markov models where objects u and v are neighbours whenever their induced objects $R(u)$ and $R(v)$ overlap are natural choices.

Given observation of image $\mathbf{y}$, the posterior probability density for scene $\mathbf{x}$ is

$$p(\mathbf{x} \mid \mathbf{y}) \propto f(\mathbf{y} \mid \mathbf{x}) p(\mathbf{x}). \tag{14}$$

A maximum a posterior (MAP) estimator of the true configuration solves

$$\widehat{\mathbf{x}} = \operatorname{argmax}_{\mathbf{x}} p(\mathbf{x} \mid \mathbf{y}) = \operatorname{argmax}_{\mathbf{x}} f(\mathbf{y} \mid \mathbf{x}) p(\mathbf{x}). \tag{15}$$

Assuming $p(\cdot) > 0$ and taking logarithms, (15) can be rewritten as a penalised maximum likelihood estimation

$$\widehat{\mathbf{x}} = \operatorname{argmax}_{\mathbf{x}} \left[\log f(\mathbf{y} \mid \mathbf{x}) + \log p(\mathbf{x})\right] \tag{16}$$

we interpret $-\log f(\mathbf{y} \mid \mathbf{x})$ as a measure of goodness of fit to the data, and $-\log p(\mathbf{x})$ as a pena for the complexity of the configuration $\mathbf{x}$. For instance, the Strauss process (2) results in a penalt $-\log \beta$ for the presence of each object $x_i \in \mathbf{x}$ and a penalty of $-\log \gamma$ for each pair of neighbou objects (e.g. overlapping objects). Modifications which might be useful in this application are

$$p(\mathbf{x}) = \alpha \prod_{i=1}^{n} \beta^{|R(x_i)|} \prod_{i<j} \gamma^{|R(x_i) \cap R(x_j)|} \tag{17}$$

and, for marked objects, to allow the interaction terms to depend on the marks.

Regarding the choice of parameter values, note that if the raster is made finer (say, quadrupling the number of pixels) then the log likelihood typically increases by the same factor. This suggests that to maintain the balance between f and p in (15)-(16) the parameters $\log \beta$ and $\log \gamma$ of a Strauss model should also be multiplied by this factor. Models such as (17) and (10), with interactions expressed in terms of pixel counts, do not require such adjustment.

5.1 ICM for object recognition

Turning to estimation of the scene $\mathbf{x}$, note that due to their computational complexity the MAP equations (15) cannot be solved analytically and we have to resort to iterative methods.

A simple deterministic technique [2] considers basic changes such as adding an object to the current recognition, slightly modifying the features of an object or deleting one. E.g. an object u will be added to scene $\mathbf{x}$ if

$$\log \frac{f(\mathbf{y} \mid \mathbf{x} \cup \{u\})p(\mathbf{x} \cup \{u\})}{f(\mathbf{y} \mid \mathbf{x})p(\mathbf{x})} > w \tag{18}$$

where $w \geq 0$ is a chosen threshold.

Various strategies to select u are available. For instance, the object space U can be digitised and scanned, resulting in an analogue of Besag's ICM algorithm [8]. Alternatively, one can search for that object whose addition would most increase the posterior likelihood ratio. The latter algorithm is also defined when U is 'continuous' (any l.s.c. space) but the interpretation is more complex: a new object u is added at that position where the Papangelou conditional intensity of the posterior distribution, given the current configuration $\mathbf{x}$ on $U \setminus \{u\}$, is maximal and greater than e^w. Deletions and replacements can be dealt with similarly.

As an example, Figure 2 shows a scanned 128 × 128 image ('pellets') taken from the Brodatz texture album [13]. We treat the pellets as discs of fixed radius 4 pixels but with blurred boundaries. The grey-level histogram has two distinct peaks at value 8 and 172, suggesting that we can regard the background and foreground signal as roughly constant at these values. Assuming additive Gaussian noise, the noise variance was estimated by thresholding the image and taking the sample variance, giving an estimate of 83.1. Blurring was modelled by assuming that the original blur-free signal was subjected to a 3×3 averaging (linear) filter with relative weights 4 for the central pixel, 2 for horizontal and vertical neighbours and 1 for diagonal neighbours.

Figure 3 shows an ICM-approximation to the maximum likelihood estimator computed by steepest ascent from an empty initial configuration. Pellets are correctly identified but there is 'multiple response', i.e. the MLE sometimes contains clusters of objects around the position of each 'true' object. This is alleviated by introducing a prior, in this case a Strauss model (2), as illustrated in Figure 4.

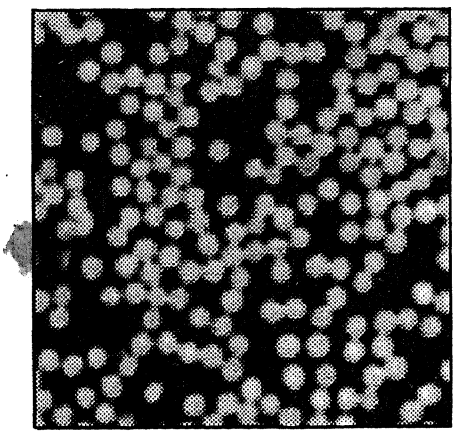

Figure 2: Pellets image taken from Brodatz (1966), digitised on a 128 × 128 square grid.

5.2 Sampling from the posterior

An alternative to ICM is to use sampling techniques [9, 10, 53]. For an up to date account in the context of spatial processes see [36].

The classical approach [6, 35, 43] followed in [4, 31] is based on spatial birth-and-death processes [42]. These are continuous-time pure jump Markov processes whose transitions are either the addition

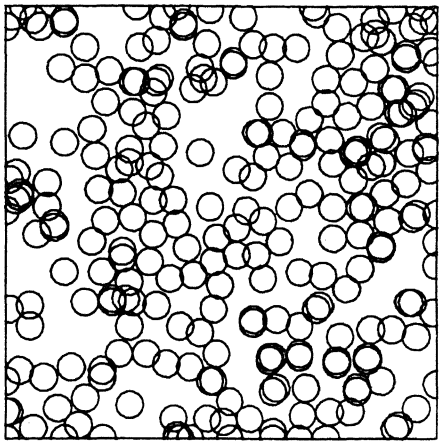

Figure 3: Approximate maximum likelihood reconstruction by steepest ascent of the pellets texture from an empty initial state.

('birth') of a new object or the deletion ('death') of an existing one. A convenient choice is the constant death rate process which has rate $D(\mathbf{x} \setminus \{x_i\}, x_i) = 1$ for a transition from $\mathbf{x}$ to $\mathbf{x} \setminus \{x_i\}$ and birth rate

$$b(\mathbf{x}, u) = \begin{cases} \frac{f(\mathbf{y}|\mathbf{x}\cup\{u\})p(\mathbf{x}\cup\{u\})}{f(\mathbf{y}|\mathbf{x})p(\mathbf{x})} & \text{if } f(\mathbf{y} \mid \mathbf{x})p(\mathbf{x}) > 0 \\ 0 & \text{else} \end{cases} \tag{19}$$

for adding $u \notin \mathbf{x}$ to pattern $\mathbf{x}$. It is easy to see that $b(\cdot,\cdot)$ and $D(\cdot,\cdot)$ satisfy the *detailed balance equations*

$$b(\mathbf{x}, u)p(\mathbf{x} \mid \mathbf{y}) = D(\mathbf{x}, u)p(\mathbf{x} \cup \{u\} \mid \mathbf{y}). \tag{20}$$

Other techniques include jump-diffusion processes [24] or Metropolis-Hastings algorithms [22, 23]. The latter are usually cast as a discrete time Markov chain and operate by proposing a new configuration (birth, death, etc) and accepting it with a probability designed to satisfy detailed balance.

For each of the techniques described, every new application calls for verifying that the Markov process is well-defined and converges to the desired equilibrium distribution $p(\cdot \mid \mathbf{y})$. For instance, the following corollory [31] of Proposition 5.1 and Theorem 7.1 in [42] holds.

Lemma 1 *Let $\mathbf{y}$ be fixed. For any blur-free noise model (13) with $g(\cdot \mid \cdot) > 0$, and any nearest-neighbour Markov object process $p(\cdot)$ with uniformly bounded likelihood ratios*

$$\frac{p(\mathbf{x} \cup \{u\})}{p(\mathbf{x})} \leq \beta < \infty,$$

there exists a unique spatial birth-and-death process with constant death rate 1 and birth rate (19). The process has unique equilibrium distribution p and converges in distribution to p from any initial state.

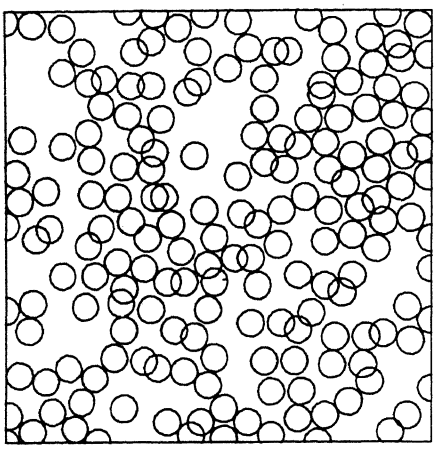

Figure 4: MAP reconstruction of the Brodatz pellets texture by steepest ascent from an empty initial state ($w = 0$) using a Strauss prior with $\log\beta = \log\gamma = -1000$.

Figure 5 shows a sample from the posterior distribution for the pellets texture using the same Strauss prior as for Figure 4.

The main advantage of sampling from the posterior distribution is the ability to estimate any functional of the posterior by taking a sufficient number of independent realisations. Examples of useful functionals are: the distribution (mean, variance) of the number of objects; the probability that there is no object in a given subregion of the image; the distribution of the distance from a given reference point to the nearest object and the first-order intensity [54]. The first-order intensity for the Brodatz example is given in Figure 6.

5.3 Stochastic annealing

A MAP solution can also be found by simulated annealing. Assume the conditions of Lemma 1. For $H > 0$ define

$$p_H(\mathbf{x} \mid \mathbf{y}) \propto \{f(\mathbf{y} \mid \mathbf{x})p(\mathbf{x})\}^{1/H}.$$

This is the density of a nearest-neighbour Markov object process, and the associated spatial birth-and-death process with constant death rate 1 and birth rate (19) exists and converges in distribution to $p_H(\cdot \mid \mathbf{y})$.

As for discrete Markov random fields, H has the interpretation of 'temperature'. If U is discrete then $p_H(\cdot \mid \mathbf{y})$ converges pointwise as $H \to 0$ to a uniform distribution on the set of MAP solutions.

Take a sequence $H_n \searrow 0$ and consider the corresponding family $(X^{(n)})_{n\in\mathbb{N}}$ of spatial birth-and-death processes on $K = \{\mathbf{x} : f(\mathbf{y} \mid \mathbf{x})p(\mathbf{x}) > 0\}$. Let $t_n, n \in \mathbb{N}$ be a sequence satisfying

$$t_n \geq t_0 \left(1 + \frac{\log(\frac{1}{2}(1 - \frac{1}{n}))}{\log(1 - K_n(t_0))}\right).$$

where $K_n(t_0)$ is a certain constant determining the rate of convergence of the n^{th} birth-death process [31, 35]. Construct a time-inhomogeneous Markov process $X_t, t > 0$ whose transition rates are those

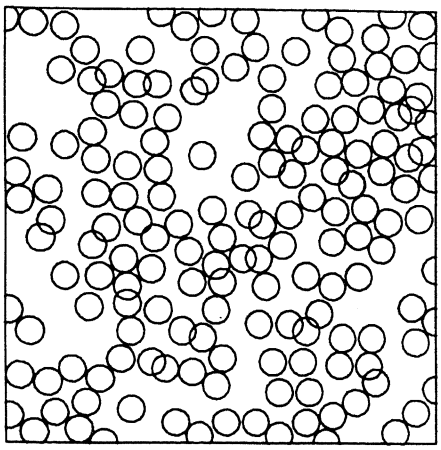

Figure 5: Realisation from the posterior distribution for the Brodatz pellets texture sampled at time 1. The prior distribution is a Strauss process with $\log\beta = \log\gamma = -1000$.

of $X^{(n)}$ during time interval $[s_n, s_{n+1})$ where $s_n = t_0 + \cdots + t_{n-1}$. It can be shown [31] that under certain regularity conditions the sequence of birth-and-death processes constructed this way converges in total variation to a uniform distribution on the set of global maxima of the posterior distribution, regardless of the initial state.

For the Brodatz pellet texture of Figure 2, the results of simulated annealing are very similar to those of posterior sampling (Figure 5).

6 Discussion

In this paper we argued for using Markov spatial processes as a prior distribution in vision. One reason is that, due to the Markov property, ratios of the form $p(A\mathbf{x})/p(\mathbf{x})$ are easy to compute for operators A such as changing the value of a single pixel (for discrete Markov random fields [8, 21]), adding, deleting or modifying an object in a Markov object process or moving vertices in a polygonal model [15].

Focusing on the formulation in Section 5.1, the (forward) log likelihood ratio for a blur-free silhouette signal

$$\log\frac{f(\mathbf{y}\mid\mathbf{x}\cup\{u\})}{f(\mathbf{y}\mid\mathbf{x})} = \sum_{t\in R(u)}\log\frac{g(y_t\mid\theta_1)}{g(y_t\mid\theta_0)} \tag{21}$$

is a sum over pixels in $R(u)$ only (related to the *Hough transform* [26, 27] in computer vision) and depends on the data image only through

$$z_t = \log\frac{g(y_t\mid\theta_1)}{g(y_t\mid\theta_0)}.$$

Since z_t does not depend on the object configuration $\mathbf{x}$, it can be computed only once in the initialisation stage of the algorithm.

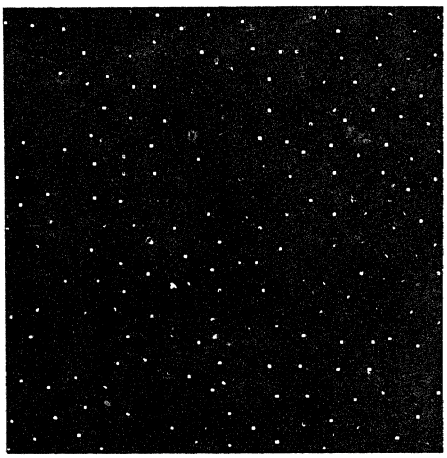

Figure 6: Posterior intensity estimated over 10 time units for the Brodatz pellets texture. The prior distribution is a Strauss process with $\log\beta = \log\gamma = -1000$.

Moreover, after adding or deleting a particular object u the log likelihood ratio (21) requires updating only for v in the region

$$V(u) = \{v \in U : R(v) \cap R(u) \neq \emptyset\}.$$

For example, in a translation model with $U = T = \mathbb{R}^d$ and $R(u) = R_0 + u$ this is the central symmetrisation $V(u) = R_0 \oplus \check{R}_0 + u$.

For the Markov prior, the likelihood ratio for e.g. adding object u is local too and depends only on those objects in the current reconstruction that are neighbours of u (e.g. overlapping $R(u)$).

Monte Carlo samplers can then be build by repeatedly performing these operations. However, MCMC techniques can encounter numerical problems, especially if the temperature parameter H becomes small. Then the birth-and-death process behaves like a deterministic steepest ascent algorithm, suggesting incorporating a search operation [4]. If the dimension of the object space U becomes large, the cost of searching it increases exponentially and multiresolution techniques are useful [4]. Moreover, efficient parametrisation is helpful.

The framework described in this paper is quite general and can be adapted easily to a wide range of problems involving e.g. subpixel resolution of objects, but also stereo pairs, motion tracking [41], edge detection or sketching problems.

Finally, as repeatedly argued by Besag [11] in the context of Markov random fields, it is not necessary to 'believe' that the prior is a model for the underlying scene. In fact, good reconstructions may well have low prior probability. This has been used as an argument to prefer ICM to simulated annealing, since the global optimum may depend too much on undesirable global properties of the prior. Similar remarks hold in the present context.

Acknowledgements

I am grateful to Arnoldo Frigessi for his kind invitation and to many colleagues and collaborators, in particular to Adrian Baddeley, Wilfrid Kendall, Françoise Mathers, Jesper Møller, Ilya Molchanov and Adri Steenbeek.

References

[1] A.J. Baddeley, W.S. Kendall and M.N.M. van Lieshout. Quermass-interaction processes, 1995. Manuscript in preparation.

[2] A.J. Baddeley and M.N.M. van Lieshout. ICM for object recognition. In Yadolah Dodge and Joe Whittaker, editors, *Computational statistics*, volume 2, pages 271–286, Physica/Springer, Heidelberg-New York, 1992.

[3] A.J. Baddeley and M.N.M. van Lieshout. Object recognition using Markov spatial processes. In *Proceedings 11th IAPR International Conference on Pattern Recognition*, pages B 136–139, IEEE Computer Society Press, Los Alamitos, California, 1992.

[4] A.J. Baddeley and M.N.M. van Lieshout. Stochastic geometry models in high-level vison. In K. Mardia and G.K. Kanji, editors, *Statistics and images, Volume 1 Advances in Applied Statistics, a supplement to Journal of Applied Statistics*, volume 20, pages 231–256, Carfax, Oxford, 1993.

[5] A.J. Baddeley and M.N.M. van Lieshout. Area-interaction point processes. *Annals of the Institute of Statistical Mathematics*, 1995. To appear.

[6] A.J. Baddeley and J. Møller. Nearest-neighbour Markov point processes and random sets. *International Statistical Review*, 57:89–121, 1989.

[7] A.J. Baddeley, M.N.M. van Lieshout and J. Møller. Markov properties of cluster processes. Technical Report 278, Department of Theoretical Statistics, University of Aarhus, February 1994.

[8] J. Besag. On the statistical analysis of dirty pictures (with discussion). *Journal of the Royal Statistical Society, Series B*, 48:259–302, 1986.

[9] J. Besag and P.J. Green. Spatial statistics and Bayesian computation. *Journal of the Royal Statistical Society, Series B*, 55:25–37, 1993.

[10] J. Besag, P.J. Green, D. Higdon and K. Mengersen. Bayesian computation and stochastic systems. *Statistical Science*, 1995. To appear.

[11] J. Besag, J. York and A. Mollié. Bayesian image restoration with two applications in spatial statistics. *Annals of the Institute of Statistical Mathematics*, 43:1–59, 1991.

[12] J.A.K. Blokland, A.M. Vossepoel, A.R. Bakker, and E.K.J Pauwels. Automatic assignment of elliptical ROIs: First results in planar scintigrams of the left ventricle. *European Journal of Nuclear Medicine*, 15:87–92, 1989.

[13] P. Brodatz. *Texture: a photographic album for artists and designers*, Dover, New York, 1966.

[14] Y.S. Chow, U. Grenander and D.M. Keenan. *HANDS: a pattern-theoretic study of biological shapes*. Springer–Verlag, New York, 1991.

[15] P. Clifford and G. Nicholls. Simulating polygonal shape models with data, 1995. Manuscript in preparation.

[16] D.J. Daley and D. Vere-Jones. *An introduction to the theory of point processes*. Springer Verlag, New York, 1988.

[17] J. Dengler and M. Guckes. Estimating a global shape model for objects with badly-defined boundaries. In W. Förstner and S. Ruwiedel, editors, *Robust Computer Vision*, pages 122–136, Wichman, Karlsruhe, 1992.

[18] P.J. Diggle. *Statistical analysis of spatial point patterns.* Academic Press, London, 1983.

[19] P.J. Diggle, T. Fiksel, P. Grabarnik, Y. Ogata, D. Stoyan and M. Tanemura. On parameter estimation for pairwise interaction processes. *International Statistical Review*, 62:99–117, 1994.

[20] D.J. Gates and M. Westcott. Clustering estimates for spatial point distributions with unstable potentials. *Annals of the Institute of Statistical Mathematics*, 38:123–135, 1986.

[21] S. Geman and D. Geman. Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6:721–741, 1984.

[22] C.J. Geyer and J. Møller. Simulation procedures and likelihood inference for spatial point processes. Technical Report 260, University of Aarhus, 1993.

[23] P.J. Green. Reversible jump MCMC computation and bayesian model determination, December 1994.

[24] U. Grenander and M.I. Miller. Representations of knowledge in complex systems (with discussion). *Journal of the Royal Statistical Society, Series B*, 56:549–603, 1994.

[25] J.M. Hammersley, J.W.E. Lewis and J.S. Rowlinson. Relationships between the multinomial and poisson models of stochastic processes, and between the canonical and grand canonical ensembles in statistical mechanics, with illustrations and monte carlo methods for the penetrable sphere model of liquid-vapour equilibrium. *Sankhya: The Indian Journal of Statistics, series A*, 37:457–491, 1975.

[26] P.V.C. Hough. Method and means for recognizing complex patterns. US Patent 3069654, 1962.

[27] J. Illingworth and J. Kittler. A survey of the Hough transform. *Computer Vision, Graphics and Image Processing*, 44:87–116, 1988.

[28] W.S. Kendall. A spatial Markov property for nearest-neighbour Markov point processes. *Journal of Applied Probability*, 28:767–778, 1990.

[29] J.F.C. Kingman. Remarks on the spatial distribution of a reproducing population. *Journal of Applied Probability*, 14:577–583, 1977.

[30] M.N.M. van Lieshout. A Bayesian approach to object recognition. In U. Eckhardt, A. Hübler, W. Nagel, and G. Werner, editors, *Geometrical problems of image processing*, volume 4 of *Research in Informatics*, pages 185–190, Proceedings Geobild'91. Akademie Verlag, Berlin, 1991.

[31] M.N.M. van Lieshout. Stochastic annealing for nearest-neighbour point processes with application to object recognition. *Advances in Applied Probability*, 26:281–300, 1994.

[32] M.N.M. van Lieshout. *Stochastic geometry models in image analysis and spatial statistics.* PhD thesis, Vrije Universiteit Amsterdam, 1994.

[33] M.N.M. van Lieshout and A.J. Baddeley. A nonparametric measure of spatial interaction in point patterns. *Statistica Neerlandica*, 1995. To appear.

[34] G. Matheron. *Random sets and integral geometry.* John Wiley and Sons, New York, 1975.

[35] J. Møller. On the rate of convergence of spatial birth-and-death processes. *Annals of the Institute of Statistical Mathematics*, 41:565–581, 1989.

[36] J. Møller. Markov chain Monte Carlo and spatial point processes. Technical Report 293, Department of Theoreticl Statistics, University of Aarhus, 1994.

[37] R. Molina and B.D. Ripley. Using spatial models as priors in astronomical image analysis. *Journal of Applied Statistics*, 16:193–206, 1989.

[38] Y. Ogata and M. Tanemura. Estimation for interaction potentials of spatial point patterns through the maximum likelihood procedure. *Annals of the Institute of Statistical Mathematics*, B 33:315–338, 1981.

[39] Y. Ogata and M. Tanemura. Likelihood analysis of spatial point patterns. *Journal of the Royal Statistical Society, Series B*, 46:496–518, 1984.

[40] A. Penttinen. Modelling interaction in spatial point patterns: parameter estimation by the maximum likelihood method. *Jyvaskyla Studies in Computer Science, Economics and Statistics*, 7:1–105, 1984.

[41] D.B. Philips and A.F.M. Smith. Dynamic image analysis using bayesian shape and texture models. In K. Mardia and G.K. Kanji, editors, *Statistics and images, Volume 1 Advances in Applied Statistics, a supplement to Journal of Applied Statistics*, volume 20, pages 299–322, Carfax, Oxford, 1993.

[42] C.J. Preston. Spatial birth-and-death processes. *Bulletin of the International Statistical Institute*, 46:371–391, 1977.

[43] B.D. Ripley. Modelling spatial patterns (with discussion). *Journal of the Royal Statistical Society, Series B*, 39:172–212, 1977.

[44] B.D. Ripley. *Spatial statistics.* John Wiley and Sons, New York, 1981.

[45] B.D. Ripley. *Statistical inference for spatial processes.* Cambridge University Press, Cambrid 1988.

[46] B.D. Ripley. Gibbsian interaction models. In D.A. Griffiths, editor, *Spatial statistics: p present and future*, pages 1–19. Image, New York, 1989.

[47] B.D. Ripley. *Statistics in the environmental and earth sciences*, chapter Stochastic models fc the distribution of rock types in petroleum reservoirs, 1991.

[48] B.D. Ripley and F.P. Kelly. Markov point processes. *Journal of the London Mathematical Society*, 15:188–192, 1977.

[49] B.D. Ripley and A.I. Sutherland. Finding spiral structures in images of galaxies. *Philosophical Transactions of the Royal Society of London, Series A*, 332:477–485, 1990.

[50] J.S. Rowlinson. Penetrable sphere models of liquid-vapor equilibrium. *Advances in Chemical Physics*, 41, 1980.

[51] J.S. Rowlinson. Probability densities for some one-dimensional problems in statistical mechanics. In G.R. Grimmett and D.J.A. Welsh, editors, *Disorder in physical systems*, Clarendon Press, Oxford, 1990.

[52] J. Serra. *Image analysis and mathematical morpholo$_{gy}$.* Academic Press, London, 1982.

53] A.F. Smith and G.O. Roberts. Bayesian computation via the Gibbs sampler and related Markov chain Monte Carlo methods. *Journal of the Royal Statistical Society, Series B*, 55:3-23, 1993.

[54] D. Stoyan, W.S. Kendall and J. Mecke. *Stochastic geometry and its applications*. Akademie-Verlag, Berlin, 1987.

[55] B. Widom and J.S. Rowlinson. New model for the study of liquid-vapor phase transitions. *The Journal of Chemical Physics*, 52:1670-1684, 1970.

Summary

This paper studies Markov object processes, a class of point process models defined in terms of interactions between 'close' objects. These are widely used to model inhibitory patterns, but were believed to be less suitable for clustering.

We establish a connection with Poisson cluster processes, the standard model for positive association between objects, and discuss a recently proposed class of Markov processes that can exhibit both clustering and inhibition depending on the value of a single parameter.

Finally, we argue that Markov models are natural prior distributions in higher level vision tasks such as object recognition. Maximum likelihood solutions typically contain many similar objects and an inhibitory Markov prior helps to overcome this problem.

Résumé

Dans cet article nous étudions les processus Markov d'objets, classe de modèles ponctuels définie par les interactions d'objets 'proche'. Ces modèles sont souvent utilisés pour les images spaciales pour lesquelles la distribution est plus régulière que celle d'un processus de Poisson, mais que l'on croyait moins appropriée pour les images plus groupées.

Nous montrons un rapport avec des processus de Poisson groupés, les modèles standards pour l'association positive d'objets et présentons une famille de processus Markov proposée recemment qui permet des images spaciales régulières et groupées selon la valeur d'un paramètre.

Finalement, nous exposons que les modèles Markov sont des distributions á priori naturelles pour des problémes de vision de haut niveau tels que l'identification d'objets. Les solutions de vraisemblance maximum contiennent de façon typique beaucoup d'objets similaires et un modèle Markov á priori régulier aide à résoudre ce problème.