

# Object Recognition using Markov Spatial Processes

A. J. Baddeley

Centre for Mathematics and  
Computer Science (CWI)  
Amsterdam  
The Netherlands

M. N. M. van Lieshout

Department of Mathematics  
and Computer Science  
Free University of Amsterdam  
The Netherlands

## Abstract

*The Bayesian approach to image processing based on Markov random fields is adapted to image analysis problems such as object recognition and edge detection. In this context the prior models are Markov point processes and random object patterns from stochastic geometry. We develop analogues of Besag's ICM algorithm. The erosion operator of mathematical morphology turns out to be a maximum likelihood estimator for a simple noise model. We show that the Hough transform can be interpreted as a likelihood ratio test statistic.*

## 1 Introduction

This paper studies object recognition problems where a scene composed of several 'objects' is observed in the presence of blur and noise, and the task is to locate them. Bayesian techniques based on Markov random fields, which have been successful in segmentation and tomographic reconstruction [2, 4], are adapted to this new context where the desired output is not a pixel image but a graphical pattern such as a line drawing or list of filled polygons. The role of Markov random fields is taken over by Markov random spatial patterns [1, 3, 9, 10, 13]. Our work considers patterns of several objects, whereas several recent studies [6, 7, 11] develop algorithms to recognise the shape of a *single*, deformable template object.

## 2 Definitions

The observed image  $y$  is digitized on a finite pixel lattice  $T$  ('**image space**'), and  $y_t$  denotes the value at pixel  $t \in T$ .

The objects to be recognized 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 ('**object space**'), so that a point  $u \in U$  represents an object  $R(u) \subseteq T$ . An object configuration is an unordered list of objects,

$$x = \{x_1, \dots, x_n\}, \quad x_i \in U, \quad i = 1, \dots, n, \quad n \geq 0.$$

Note that  $n$  is variable, and the empty list  $\emptyset$  is allowed.

## 3 Maximum likelihood approach

### 3.1 General formulation

Analogously to [12] we formulate object recognition as a statistical parameter estimation problem. Assume the image  $y$  depends on the 'true' object pattern  $x$  through a known probability distribution, with density  $f(y|x)$ . The *maximum likelihood estimator* (MLE) of  $x$  is

$$\hat{x} = \operatorname{argmax}_x f(y|x). \quad (1)$$

Following [5] our stochastic image model  $f(y|x)$  consists of a deterministic 'deformation' and a random noise component. Any object configuration  $x$  determines an image  $\theta^{(x)}$  in pixel space  $T$ , representing the ideal signal, which is then combined with random noise.

We assume the pixel values  $y_t$  are conditionally stochastically independent given  $x$ , and the distribution of  $y_t$  depends only on the value of the signal at that same pixel:

$$f(y|x) = \prod_{t \in T} g(y_t | \theta^{(x)}(t)) \quad (2)$$

where  $g(\cdot|\theta)$  is a known probability density for each value of  $\theta$ . For example this includes additive and multiplicative noise.

For brevity here we assume a blur-free model in which  $\theta^{(x)}(t)$  takes either a foreground or background value  $\theta_1, \theta_0$  according as pixel  $t$  lies inside or outside the ‘silhouette’ formed by the union of the objects,

$$S(x) = \bigcup_{i=1}^n R(x_i).$$

In other words the simplified model assumes the objects are first ‘painted’, without blur, onto the scene which is then corrupted with independent random noise.

Figure 1 shows a realization of the simple model with additive (pixelwise independent) Gaussian noise.

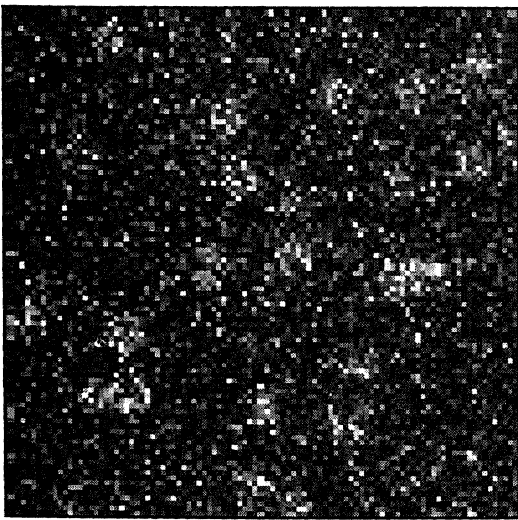


Figure 1: Realisation from simplified model with additive Gaussian noise, SNR = 0.25.

### 3.2 Relation to mathematical morphology

In the case of a binary data image, assume a model in which background pixels are randomly flipped from value 0 to 1 with probability  $q$  independently, and foreground (silhouette) pixels are unchanged. Let  $Y$  be the set of pixels with value 1. The likelihood is nonzero only when  $S(x) \subseteq Y$  and the loglikelihood is then a linear function of  $|S(x)|$ , the total area of the silhouette. Hence, one solution for the MLE is

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

which is a form of the erosion operator of mathematical morphology. This  $\hat{x}_{\max}$  is the largest solution; the

other solutions are the subsets  $x \subseteq \hat{x}_{\max}$  with the same silhouette,

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

In other words, the erosion technique is equivalent to assuming random binary noise in the background and taking the maximal MLE.

### 3.3 Iterative maximum likelihood algorithms

Usually the MLE cannot be determined directly and we need to resort to iterative optimization techniques. The simplest form of iterative adjustment is to add or delete objects. Recursively add an object  $u \in U$  to the current list  $x$ , yielding  $x \cup u$ , if the log likelihood ratio

$$L(x \cup u; y) - L(x; y) = \log \frac{f(y|x \cup u)}{f(y|x)} \quad (3)$$

is sufficiently large; and delete one of the existing objects  $x_i \in x$  to yield  $x \setminus x_i$  if

$$L(x \setminus x_i; y) - L(x; y) = \log \frac{f(y|x \setminus x_i)}{f(y|x)} \quad (4)$$

is sufficiently large. Two possible algorithms are

**coordinatewise optimization:** Visit all points  $u \in U$  in a prearranged order. If  $u$  is not in the current pattern  $x$ , then add  $u$  to  $x$  if (3) exceeds a threshold  $w \geq 0$ . If  $u = x_i$  is in the current pattern, then delete it from  $x$  if (4) exceeds  $w$ . Repeat.

**steepest ascent:** Find the maximum value of (3) over all possible new objects  $u$ , and the maximum of (4) over all existing objects  $x_i$  in  $x$ . If one of these exceeds a threshold  $w \geq 0$ , add (or delete) the corresponding object. Repeat.

Both algorithms increase the likelihood at each step,  $f(y|x^{(k+1)}) \geq f(y|x^{(k)})$ , and in practical cases convergence is guaranteed (at worst there is cycling between images of equal likelihood). However these algorithms do not necessarily yield the global maximum likelihood solution, and the local maximum obtained will depend on the initial configuration  $x^{(0)}$ .

### 3.4 Connection with Hough transform

For our simple model

$$L(x \cup u; y) - L(x; y) = \sum_{t \in R(u) \setminus S(x)} h(y_t, \theta_0, \theta_1) \quad (5)$$

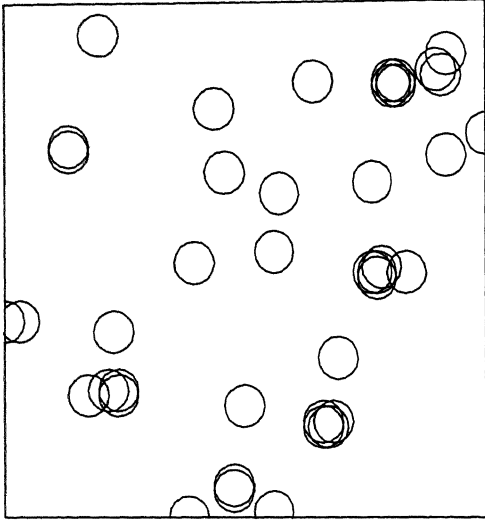


Figure 2: Object recognition using MLE, steepest ascent.

where

$$h(y_t, \theta_0, \theta_1) = \log \frac{g(y_t | \theta_1)}{g(y_t | \theta_0)}$$

is the difference in ‘goodness of fit’ at pixel  $t$ . In particular the log likelihood ratio of a single object  $u$  against an empty scene  $\emptyset$  is

$$L(\{u\}; y) - L(\emptyset; y) = \sum_{t \in R(u)} h(y_t, \theta_0, \theta_1). \quad (6)$$

Now (6) is exactly analogous to the Hough transform. It has the same interpretation of a ‘sum of votes’ from all pixels belonging to the object  $R(u)$ , with the modification that fractional and negative votes are allowed. Thus the Hough transform can be interpreted as a likelihood ratio test statistic.

When the reference scene  $x$  is not empty, the log likelihood ratio takes the form (5) which is a generalization of the Hough transform calling for summation inside the complementary mask  $T \setminus S(x)$ .

For the additive Gaussian noise model, (6) is exactly the classical Hough transform. One can also show that preprocessing the data image before applying the Hough transform is equivalent to assuming a different noise model.

Figure 2 shows the result of steepest ascent on the synthetic data of Figure 1.

## 4 Bayesian approach

### 4.1 General

Maximum likelihood solutions  $x$  tend to contain clusters of almost identical objects where the reality is one single object (see Figure 2). To suppress this instability and multiple response, one can adopt a Bayesian approach analogous to [2, 4, 7, 6]. In addition to the noise model  $f(y|x)$ , the true image  $x$  is assumed to have been generated at random from a prior probability distribution with density  $p(x)$ . The posterior probability distribution for  $x$  after observing data  $y$  is found by Bayes’ formula  $p(x|y) \propto f(y|x)p(x)$  and the *maximum a posteriori* (MAP) estimator of  $x$  is

$$\begin{aligned} \tilde{x} &= \operatorname{argmax}_x p(x|y) \\ &= \operatorname{argmax}_x f(y|x)p(x). \end{aligned} \quad (7)$$

Thus  $p(x)$  can also be regarded as a smoothing penalty attached to the optimization of  $f$ , and  $\tilde{x}$  as a penalized maximum likelihood estimator.

### 4.2 Markov prior model

Since  $x$  is no longer a discretized image but a variable-length list  $x = \{x_1, \dots, x_n\}$  of parameter points in a continuous space, the natural stochastic models come from stochastic geometry and spatial statistics [3, 8, 13]. The counterparts of pixel-based Markov random fields [2, 4] are nearest-neighbour Markov point processes [1, 10].

For brevity we consider only the *Strauss overlap-ping object process* [1] with probability density

$$p(x) = \alpha \beta^{n(x)} \gamma^{r(x)} \quad (8)$$

where  $\beta > 0$  and  $0 \leq \gamma \leq 1$ . Here  $n(x)$  is the number of objects in  $x$  and  $r(x)$  the number of pairs of overlapping objects. Interaction between objects is controlled by  $\gamma$ . If  $\gamma < 1$ , objects tend to avoid each other, in the sense that configurations with many overlapping objects have low probability; indeed, if  $\gamma = 0$ , overlap is not permitted. If  $\gamma = 1$  we get a completely random (Poisson) process.

### 4.3 MAP estimation

Determination of the MAP estimator requires optimization (7) over variable-length lists  $x$  of objects. If (8) is adopted as the prior, we incur a multiplicative penalty  $\beta$  for the presence of each object  $u$ , and a penalty  $\gamma$  for each pair of overlapping objects. These

control the tradeoff between goodness-of-fit to the data and ‘complexity’ of the solution  $x$ . If  $\beta = \gamma = 1$  the MAP estimator is just the MLE; when  $\gamma = 0$  the MAP estimator maximizes the likelihood subject to the constraint that no objects overlap.

Exact solution of (7) is usually impossible and we use iterative algorithms similar to those developed for the MLE. Thus e.g. we iteratively add object  $u$  to list  $x$  iff

$$\log \frac{f(y|x \cup u)p(x \cup u)}{f(y|x)p(x)} > w \quad (9)$$

where  $w \geq 0$  is a chosen threshold. This is analogous to Besag’s ICM algorithm [2].

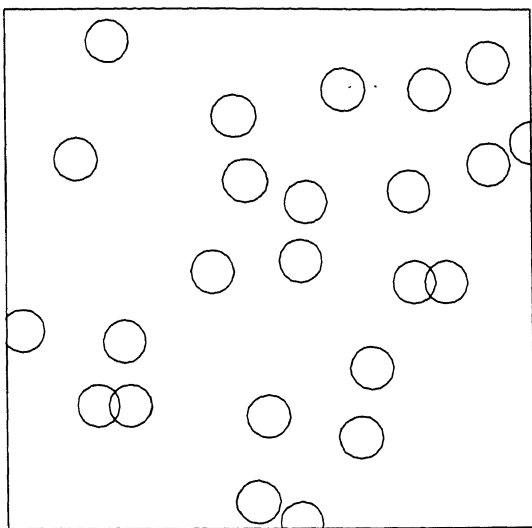


Figure 3: Object recognition using MAP, steepest ascent

Figure 3 shows the result of the steepest ascent MAP algorithm on our synthetic data using a Strauss prior with  $\beta = .0025$ ,  $\gamma = 0.25$  and  $w = 0$ .

## Conclusion

The ultimate objective of this study is to devise object recognition techniques which outperform existing ones. This is not yet demonstrated although the current algorithms perform creditably well in synthetic tests, and the evidence supports the claim that Bayesian methods are appropriate. The most interesting finding is a mathematical connection between existing ad hoc methods for object recognition (Hough transform, erosion) and general theoretical techniques such as maximum likelihood.

## References

- [1] A. Baddeley and J. Møller. Nearest-neighbour Markov point processes and random sets. *International Statistical Review*, 57:89–121, 1989.
- [2] J. Besag. On the statistical analysis of dirty pictures (with discussion). *Journal of the Royal Statistical Society, Series B*, 48:259 – 302, 1986.
- [3] P.J. Diggle. *Statistical analysis of spatial point patterns*. Academic Press, London, 1983.
- [4] 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.
- [5] U. Grenander. *Lectures on Pattern Theory (3 volumes)*. Applied Mathematical Sciences vols. 18, 24, 33. Springer-Verlag, New York-Berlin, 1976, 1978, 1981.
- [6] U. Grenander and D.M. Keenan. A computer experiment in pattern theory. *Communications in Statistics - Stochastic Models*, 5:531 – 553, 1989.
- [7] R. Molina and B.D. Ripley. Using spatial models as priors in astronomical image analysis. *Journal of Applied Statistics*, 16:193 – 206, 1989.
- [8] B.D. Ripley. *Spatial statistics*. John Wiley and Sons, New York, 1981.
- [9] B.D. Ripley. *Statistical inference for spatial processes*. John Wiley and Sons, New York, 1988.
- [10] B.D. Ripley and F.P. Kelly. Markov point processes. *Journal of the London Mathematical Society*, 15:188 – 192, 1977.
- [11] 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.
- [12] L.A. Shepp and Y. Vardi. Maximum likelihood reconstruction for emission tomography. *IEEE Transactions on Medical Imaging*, 1:113 – 122., 1982.
- [13] D. Stoyan, W.S. Kendall, and Mecke J. *Stochastic geometry and its applications*. John Wiley and Sons, Chichester, 1987.