10A
Probabilistic image modelling

M. N. M. van Lieshout
Centrum voor Wiskunde en Informatica, Amsterdam, The Netherlands

1 Introduction

Since the pioneering work of Besag, Cross, and Jain, the Gemans, Ripley, and others (cf. the seminal papers reprinted in the monograph edited by Mardia and Kanji (1993)), Markov random fields have played a prominent role in statistical image analysis. At first, most attention was paid to tackling low-level problems like classification and segmentation with simple categorical or continuously valued pairwise interaction priors or Gaussian autoregression models with small neighbourhood systems. Gradually though it was realized that the above-mentioned models, despite their appealing simplicity and their relatively plausible local characteristics, which form the basic ingredient of most low-level algorithms, may have quite unrealistic global properties. Indeed, ensembles of connected geometrical shapes with smooth boundaries are unlikely to arise as realizations of such discrete random fields.

Accordingly, there has been a surge of interest in the construction of stochastic models with a more satisfying global behaviour. Hurn *et al.* mention several examples, notably the higher-order models of Tjelmeland and Besag (1998), and the Gaussian Markov random fields studied in Rue (2001) and Rue and Tjelmeland (2002). In Section 2 below, we present some more recent advances in this direction.

Another interesting development in the last decade has been the shift in emphasis towards high-level vision tasks that aim to understand and interpret an image in terms of the objects it contains. By their very nature, local Markov random fields are not well-suited as prior distributions for object scenes; far more natural are the Markov marked point processes suggested by Baddeley and Van Lieshout (1992); see also the papers collected in Mardia and Kanji (1993). In Section 3 we provide further details and point out some analogies with the region-based random fields discussed in Section 2.

2 Pixel-level modelling

Hurn *et al.* note that the Potts model and many other classical categorical models 'do not provide good prior models for real scenes, and estimates of attributes such as the number of connected components may be poor'. To tackle these problems, Møller and Waagepetersen (1998) introduce so-called Markov connected component fields, the probability density of which factorizes into terms associated with the connected components in the image rather than with cliques of neighbouring pixels. They prove that under an additional homogeneity and Markov condition with respect to horizontal and vertical neighbours, the probability density can be written as a product of two terms, one governing the area, the other the perimeter of the image components. For the second-order neighbourhood system, terms for the Euler–Poincaré characteristic and the numbers of corners and discontinuities must be added.

Another way to control the size and shape of regions is by means of morphological operators (Serra 1982). For instance, in the binary case, Chen and Kelly (1992) suggest a probabilistic model to favour images that are morphologically smooth in the sense that its foreground set does not feature narrow isthmuses, small islands, or sharp capes. A similar random field for the background pixels penalizes small holes. It can be shown that both models are Markov with respect to a neighbourhood system that depends on the underlying morphological operator. More generally, Sivakumar and Goutsias (1997) use a combination of operators to favour certain shapes and sizes of the foreground and background regions over others.

In summary, the higher-order models in Tjelmeland and Besag (1998), Markov connected component fields, and morphologically constrained random fields all offer some control over the global appearance of likely images, without sacrificing the local character of the full conditionals.

3 High-level modelling

As indicated in Section 1, the class of marked point processes provides a natural framework for object scenes. In such a setup (Baddeley and Van Lieshout 1992, 1993), a point represents the position of an object in the image, its mark captures other attributes. The latter may be a real-valued vector of a few size, shape, and texture parameters as in Descombes *et al.* (1999b); for more complex objects a deformable template mark may be more appropriate (Grenander and Miller 1994).

Hurn *et al.* discuss in detail a range of stochastic shape models, but pay little attention to the interactions between objects. As in low-level image analysis, it is highly desirable from a computational point of view that the conditional dependence structure be local. To quantify this notion, Baddeley and Van Lieshout (1992) assign to each marked point its 'silhouette' in the image, which may be thought of as the discrete representation of the actual object, and define two marked points to be neighbours if their silhouettes overlap. If the conditional intensity of finding an object at some given location with a given mark depends only on this object's neighbours, then the model is called a Markov overlapping object process. A simple, but nevertheless very useful, example is the hard core process (Baddeley and Van Lieshout 1993; Hurn 1998; Rue and Syversveen 1998) consisting of a Poisson number of independent objects conditioned on the event that all silhouettes are disjoint. Alternatively, the amount of overlap may be taken into account as in the area-interaction process (Baddeley and Van Lieshout

1995) defined by an unnormalized probability density that is exponential in the total area occupied by the silhouettes. Occlusion can be formalized by ordering scenes according to which objects lie on top of others (Mardia *et al.* 1997).

Various generalizations have been proposed recently. Considering the fact that images are discrete reflections of a continuous reality, it is especially gratifying to note that the models discussed in Section 2 have analogues in the class of Markov overlapping object processes. For instance, the quermass interaction processes (Mecke 1996; Kendall *et al.* 1999) form an exponential family in the plane with the area, perimeter, and Euler–Poincare characteristic as canonical sufficient statistics. In contrast to connected component fields (Møller and Waagepetersen 1998), in the continuous case some care has to be taken to ensure the model is well-defined (Kendall *et al.* 1999). The point process counterparts of morphologically constrained random fields (Chen and Kelly 1992; Sivakumar and Goutsias 1997) are studied in Van Lieshout (1999). Since the density of both models mentioned above depends on the silhouette image only, a corresponding random set is readily defined (Van Lieshout 2000).

4 Concluding remarks

In this note we have concentrated on models. However, no real progress would have been possible without the simultaneous development of efficient sampling and parameter estimation schemes. As remarked by Hurn *et al.*, perhaps the 'most notable development ... has been the use of Markov chain Monte Carlo maximum likelihood (MCMCML)' (Geyer and Thompson 1992) at the expense of the pseudo-likelihood method, Monte Carlo Newton–Raphson techniques, and stochastic approximation algorithms; see Geyer (1999) for a comprehensive overview. The MCMCML approach focuses on the full likelihood surface, thus allowing for the computation of statistics such as the Fisher information without extra computational effort. The method is easily embedded in a fully Bayesian inference scheme, and an asymptotic theory is available for the Monte Carlo error with respect to the 'true' maximum likelihood estimator. Recently, similar asymptotics have been developed for a combination of the EM algorithm and stochastic approximation (Delyon *et al.* 1999). However, careful tuning of the discount factors involved in the approximation part remains necessary.

Another exciting development is that of exact (or perfect) simulation, following the seminal paper by Propp and Wilson (1996). In contrast to classical MCMC techniques, an exact simulation algorithm outputs an unbiased sample from the target distribution, and neither requires a burn-in time nor sufficiently large time lags between sub-sampled states.

Finally, I would like to congratulate Merrilee Hurn, Oddvar Husby, and Håvard Rue on an interesting overview of recent developments in image analysis, and express the hope that their article will act as a stimulus to further progress.

10B
Prospects in Bayesian image analysis

Alain Trubuil
Institut National de la Recherche Agronomique, Jouy-en-Josas, France

1 Introduction

The ideas presented by Hurn *et al.* offer an excellent review of the impres work carried out by the statisticial community on the Bayesian framework its application to image analysis. Necessarily the presentation could not exhaustive and with respect to applications, a consensus has guided the with modelling illustrated mainly for restoration problems. My aim is to c attention to detection, labelling, and eventually three-dimensional (3D) segr tation. This will be addressed in Section 3. The first part of the discus gives a very short presentation of the work done by Zhu *et al.* (1997–2000 texture modelling and estimation. In Section 2, I will illustrate the usefulne intensity-level curves as very important clues for image interpretation.

2 Texture modelling

In relating some aspects of the work of Zhu *et al.* on texture modelling, my is to provide the reader with recent information on generic texture mode and parameter estimation that complements Section 2.1 of the article.

Given a class of images (e.g. outdoor images) and a set of filters (e.g. S filters and Gabor filters, the impulse response of which is a complex sinu centred at a frequency and modulated by a Gaussian envelope), the aim identify a probability distribution p, associated with elements of this class. estimator of this distribution is obtained from a learning set of M images. togram statistics of filter responses are computed as

$$\mu^l_{\text{obs}}(z) = \frac{1}{M}\sum_{m=1}^{M} \#\{(x,y) \in \Lambda;\ \ F^l(I_m)(x,y) = z\},$$

where $(x,y) \in \Lambda$ denotes a pixel of the lattice Λ underlying the images; $F^l(I_m)(x,y)$ is the filter response of the lth filter at position (x,y) when ap to the mth image of the learning set. If one looks for an estimator of maxi entropy with similar histogram statistics:

$$\mathrm{E}_{\hat{p}}\left(\sum_{(x,y)} \mathbf{1}\{F^l(I)(x,y) = z\}\right) = \mu^l_{\text{obs}}(z), \qquad l = 1,\ldots,L,\ \ z \in Z^l,$$

then $\hat{p}$ is a Gibbsian distribution: $\hat{p}(I) \propto \exp\{-\sum_l \sum_z \lambda^l(z) H^l(I)(z)\}$ with $H^l(I)(z) = \int\int \mathbf{1}\{F^l(I)(x,y) = z\}\, \mathrm{d}x\, \mathrm{d}y$.

Of course, the potential functions λ^l (analogous to Lagrange multipliers) depend upon the observed statistics μ^l_{obs}. First of all, two distinct potential functions are identified. At fine scales, V-shape potential functions with rather flat tails occur with derivative filters. At coarse scales, these curves turn 'upside down'. The Gibbs distribution learned from images has energy that one can attempt to minimize by gradient descent. This leads to a non-linear PDE equation of the diffusion–reaction type with two terms. Starting from an input image $I(x, y; 0)$, the first term diffuses the image whereas the second term forms patterns.

An interesting application of texture learning through this framework is clutter removal. Here the problem is eliminating or attenuating the presence of a complex background (e.g. trees, landscape) in an image displaying objects from a class of interest (e.g. buildings). This difficult problem is handled in three stages in Zhu *et al.* (1997):

(a) learning the **b**ackground images class: $p(I^{\mathbf{b}}) \propto \exp\{-U^{\mathbf{b}}(I^{\mathbf{b}})\}$;

(b) learning the **o**bjects of interest images class: $p(I^{\mathbf{o}}) \propto \exp\{-U^{\mathbf{o}}(I^{\mathbf{o}})\}$;

(c) assuming that $I^{\text{obs}} = I^{\mathbf{o}} + I^{\mathbf{b}}$, looking for a MAP estimator of $I^{\mathbf{o}}$ according to $p(I^{\mathbf{o}}|I^{\text{obs}}) \propto \exp\{-U^{\mathbf{b}}(I^{\text{obs}} - I^{\mathbf{o}}) + U^{\mathbf{o}}(I^{\mathbf{o}})\}$.

In order to gain accuracy and speed in the learning process, Zhu *et al.* realize that filter selection and parameter estimation can be separated. They address the selection problem in Zhu *et al.* (1999) and estimation in Zhu *et al.* (2000). In these papers, the reader can find new ideas for parameter estimation in Gibbsian models.

3 Level curves and segmentation

My aim in this section is to draw attention to level curves of filter response (mainly intensity filter) as a carrier of very useful information for image representation and segmentation. This is related to Section 2.2 devoted to intermediate-level modelling and illustrated with tessellations obtained through the simulation of a posterior distribution combining grey-level homogeneity and number or length of edges between regions. The supports for labels of regions are triangles. During the simulation process triangles are merged, split, or distorted with the help of different proposals of a Metropolis–Hastings algorithm. The posterior energy is related to the Mumford and Shah functional widely used in image segmentation. Triangles have a long history in computer science and approximation (e.g. in finite element techniques for partial differential equations) but in image representation other geometric primitives could be considered. Hence, it appears that in many cases, level curves of intensity function are good candidates for boundaries between objects. Actually, as pointed out by Kervrann *et al.* (2000), a subset of level curves is the optimal solution with respect to the energy:

$$U_\lambda(f, \Omega_1, \ldots, \Omega_P) = \sum_{i=1}^{P} \int_{\Omega_i} (f(x) - \bar{f}_{\Omega_i})^2 \, \mathrm{d}x + \lambda \sum_{i=1}^{P-1} |\Omega_i|^q, \tag{3.1}$$

where f denotes the image intensity function, $\bar{f}_{\Omega_i}$ the mean grey-level over object Ω_i with area $|\Omega_i|$, P the unknown number of objects including the background, and $q = -1, 0, 1, 2$. The first term of (3.1) controls homogeneity whereas the second term controls the number of objects when $q = 0$. It is not possible to penalize directly edge lengths while keeping solutions inside the set of solutions whose object boundaries are intensity level curves. However, it is clear that quite often the triangulation and the level curve approach should be able to provide representations similar to each other. An efficient algorithm has been designed for minimizing (3.1). It relies on a coarse discretization of the grey levels. An object-guided smoothing can be incorporated with the help of anisotropic diffusion to restore surfaces and boundaries of objects. The advantage of the approach, described by Nicholls (1997) over the level curve approach, lies in the direct inclusion of edge lengths which is not possible in the level curve framework, even if more or less controllable through penalization and the smoothing trick. The advantage of the level curve approach is mainly due to fast estimation and concordance in many situations of level lines of original or smoothed images with perceptible edges in images. Within the level curve selection approach described herein, different configurations of objects are obtained depending on the grey-level quantitation chosen. My feeling is that this family of configurations could be used by colouring schemes like the one described by Nicholls. This could make it possible to incorporate a priori some geometric on objects.

4 Applications

From a practical point of view, the relevance of the Bayesian framework has been demonstrated by Hurn *et al.* in restoration and object-based segmentation of complex ultrasound images. I would also like to draw attention to an important issue for imaging in biology: 3D microscopy of biological samples.

Let me consider the fine analysis of complex structures (e.g. antennal lobes) or tissues (e.g. intestinal crypts) for which the ultimate objective is comparing samples. Very often these structures or tissues are made of dozens to hundreds of smaller sub-structures (e.g. glomerulii, cells). The structures are partially observed and often very close to each other in a 3D domain, see Fig. 1 displaying an intestinal crypt. By formulating the problem in the right way, all the tricky algorithms designed for probability distributions may be used to obtain some hints on the spatial organization of sub-structures. Let me illustrate this on nuclei detection in a volume observed through an optical section in confocal microscopy. The nuclei detection process we have used can be divided into four stages:

(a) detection of seeds inside objects,

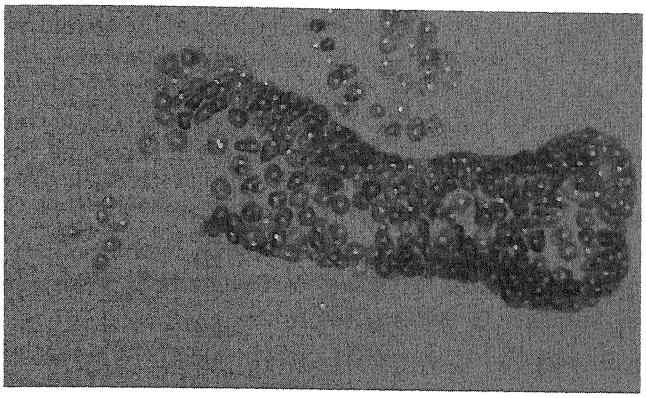

FIG. 1. Volume rendering of nuclei detection of an intestinal crypt.

(b) sub-sampling of seeds,

(c) seed aggregation,

(d) validation.

At the end of stage (b), we have a set $\mathcal{S} = \{s_i,\ i = 1, \dots, N\}$ of seeds in 3D space. The objective is to associate a unique label, l_i, to each seed according to its membership to a nucleus. Here the number of objects (nuclei) is unknown but assumed bounded by N. Moreover, we expect there is at least one seed per object and few false seeds. We define a graph whose nodes are the seeds. An edge is created between two nodes when they are neighbours. In practice, the a priori knowledge about nucleus size is used to define neighbours, ν_i, of a seed, s_i, ν_i consisting of seeds that are no more than a given distance from the considered seed. The probability distribution of labels is chosen as

$$p(L = (l_1, \dots, l_N)) \propto \exp\left\{ -\sum_{i=1}^{N} \sum_{j \in \nu_i} U_{i,j}(l_i, l_j) \right\},$$

with $U_{i,j}(l_i, l_j) = (\exp\{-Q(s_i, s_j)\} - e^{-1} \exp Q(s_i, s_j))\mathbf{1}\{l_i = l_j\}$. The quantity $Q(s_i, s_j)$ is related to image content in a local region between s_i and s_j. If one knows that objects are convex, then a low value of Q is associated with the presence of a rupture in the grey-level intensity profile between s_i and s_j. So the potential Q expresses the relationship between parameters and data. This model is similar to an inhomogeneous Potts model on a graph. An estimate of a mode of p can be searched for by a relaxation algorithm or a more efficient algorithm like the one proposed by Boykov *et al.* (1999). Moreover, constraints can be introduced which may reduce energy computation towards plausible configurations. This has been done and is illustrated in Fig. 1 which depicts a volume rendering visualization of the crypt with incrustation of small spheres for each centre of inertia of homogeneous label subsets. Even if these results are encouraging, they are not completely satisfactory. Indeed, validation is a hard task when hundreds of labels have to be examined from 2D slices of a 3D volume. My feeling is that efficient simulation tools can provide us with relevant validation clues, detecting regions of interest in the data where visual inspection should be encouraged.

Additional references in discussion

Baddeley, A. J. and Van Lieshout, M. N. M. (1992). ICM for object recognition. In *Computational Statistics*, No. 2 (eds Y. Dodge and J. Whittaker), pp. 271–86. Physica/Springer, Heidelberg/New York.

Baddeley, A. J. and Van Lieshout, M. N. M. (1995). Area-interaction point processes. *Annals of the Institute of Statistical Mathematics*, **47**, 601–19.

Boykov, Y., Veksler, O., and Zabih, R. (1999). Fast approximate energy minimization via graph cuts. *Proceedings of the Seventh IEEE International Conference on Computer Vision*, Kerkyra, Greece.

Chen, F. and Kelly, P. A. (1992). Algorithms for generating and segmenting morphologically smooth binary images. *Proceedings of the Twenty-sixth Conference on Information Sciences*, Princeton.

Delyon, B., Lavielle, M., and Moulines, E. (1999). Convergence of a stochastic approximation version of the EM algorithm. *Annals of Statistics*, **27**, 94–128.

Geyer, C. J. (1999). Likelihood inference for spatial point processes. In *Proceedings Seminaire Europèen de Statistique, Stochastic Geometry, Likelihood, and Computation* (eds O. Barndorff-Nielsen, W. S. Kendall, and M. N. M. Van Lieshout), pp. 79–140. CRC Press/Chapman & Hall, Boca Raton.

Kendall, W. S., Van Lieshout, M. N. M., and Baddeley, A. J. (1999). Quermass-interaction processes: conditions for stability. *Advances in Applied Probability (SGSA)*, **31**, 315–42.

Kervrann, C., Hoebeke, M., and Trubuil, A. (2000). Level lines as global minimizers of energy functionals in image segmentation. In *European Conference on Computer Vision*, Lecture Notes in Computer Science, No. 1843, pp. 241–56. IEEE Computer Society.

Van Lieshout, M. N. M. (1999). Size-biased random closed sets. *Pattern Recognition*, **32**, 1631–44.

Van Lieshout, M. N. M. (2000). *Markov Point Processes and their Applications.* Imperial College Press/World Scientific Publishing, London/Singapore.

Mardia, K. V. and Kanji, G. K. (eds) (1993). *Statistics and Images*, Vol. 1. Carfax Publishing, Abingdon.

Mardia, K. V., Qian, W., Shah, D., and de Souza, K. (1997). Deformable template recognition of multiple occluded objects. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **19**, 1036–42.

Mecke, K. R. (1996). Morphological model for complex fluids. *Journal of Physics: Condensed Matter*, **8**, 9663.

Møller, J. and Waagepetersen, R. (1998). Markov connected component fields. *Advances in Applied Probability (SGSA)*, **30**, 1–35.

Propp, J. G. and Wilson, D. B. (1996). Exact sampling with coupled Markov chains and applications to statistical mechanics. *Random Structures and Algorithms*, **9**, 223–52.

Serra, J. (1982). *Image Analysis and Mathematical Morphology.* Academic Press, London.

Sivakumar, K. and Goutsias, J. (1997). Morphologically constrained discrete random sets. In *Advances in Theory and Applications of Random Sets* (ed. D. Jeulin), pp. 49–66. World Scientific Publishing, Singapore.

Zhu, S., Liu, X., and Wu, Y. (1999). Statistics matching and model pursuit by efficient MCMC – A conclusion to Julesz's texture quest? *International Conference on Computer Vision*, Kerkyra, Greece.

Zhu, S. and Liu, X. (2000). Learning in Gibbsian fields: how accurate and how fast can it be? *Conference on Computer Vision and Pattern Recognition*, Hilton Head Island, South Carolina.