CLOSING ADDRESS

Bayesian methods for image analysis based on deformable templates

John T Kent
Department of Statistics, University of Leeds, Leeds LS2 9JT, UK

Abstract

The Bayesian paradigm provides a powerful unifying methodology for the identification and description of objects in images. A prior distribution specifies how the shape (plus location, size and orientation) of an object varies by means of a deformable template, and the likelihood specifies how to generate an image given the specification of an object (or objects). The posterior distribution is then used to make inferences about the objects of interest. This paper reviews some recent developments in each of these three areas —prior, likelihood and analysis of the posterior.

1 Introduction

The identification and description of objects in images is one of the major tasks of high level image analysis. In general, this is an area in which much of the progress has been made through clever heuristic suggestions for extracting useful information from images. However, in recent years a unifying framework based on the Bayesian paradigm has achieved wider acceptance. A prior model is specified for the location, size, orientation, and shape of objects in the image, and a likelihood specifies the distribution of either the image itself or of some subset of features in the image, given the prior information.

The purpose of this paper is to review several examples of this strategy emphasizing recent developments in Leeds. Two other papers in this proceedings have substantial review elements which complement the current paper. van Lieshout (1997) covers prior models, especially point processes for the number and spatial arrangement of objects, and deformable template models for shape variability. Titterington (1997) emphasizes low level priors in the form of Markov random fields together with likelihoods involving independent observations, thus giving rise to hidden Markov random fields. In addition, this paper builds on two previous review papers of Kent and Mardia (1994) and Mardia *et al.* (1995); see also Kent (1995, 1997) and Mardia (1996, 1997) .

2 Types of object

In this section we give examples of objects based on landmarks in an image. The term "object" will be used to refer to both physical objects and to more abstract collections of landmarks. Except where stated the main focus is on objects in two dimensions.

(1) Objects with a small number of landmarks in 2 *or* 3 *dimensions.* The simplest geometric examples in two dimensions are triangles or other polygons with labelled vertices. Biological examples include the outline of hands (with landmarks at the tips of the fingers and in the valleys between), or certain bones on which identifiable landmarks can be located. An example based on fish is studied by de Souza *et al.* (1997).

(2) Objects with continuous outlines. In the hand example it is possible to add a large number of equally-spaced "pseudo-landmarks" between the identifiable landmarks, giving a discrete approximation to a continuous curve (Grenander *et al.* 1991; Mardia *et al.* 1991).

(3) Regular grids. In a time sequence of magnetic resonance (MR) images of a two-dimensional slice of the heart through one cycle of a heart beat, it is possible to "tag" the heart magnetically, giving a regular grid of dark horizontal and vertical lines at the start of the cycle. A set of landmarks can be defined by the intersection points of this grid. As the cycle progresses, the heart contracts and the grid deforms.The objective is to identify the grid points from the images, and to follow the deformation through the cycle (Lee *et al.* 1997).

(4) Matching sets of unlabelled landmarks. In electrophoresis, a "blob" of biological matter embedded in a gel is subjected to electric forces, first in the x direction, then in the y direction, to separate different proteins by their physical properties. This process gives rise to an image with a collection of dots (= landmarks), each dot representing a different protein (or noise). The immediate objective is to construct a deformation to match two such images, the deformation being needed to allow for differences in the gels and experimental conditions. The final objective is to identify which proteins are present or not in both images. This problem arises from joint work with Gary Walker, Ian Dryden and Chris Glasbey (Walker 1997).

The examples in this section can be regarded as special cases of the general pattern theory of Grenander (1993). He emphasizes representations which can describe in a structured way deformations of objects into related objects.

3 Prior models

The objects in the previous section can be regarded as examples where an "ideal template" is altered to a "deformed template" which is actually observed in practice. In this section several strategies are given for modelling these deformations. In many cases it is convenient to separate out the location, size, orientation and shape of an object, and to model each one independently. Typically, the location will be given a uniform distribution over the image, and the size will be given a distribution based on the size of objects expected to be observed and capable of being resolved in practice. Orientation (for two-dimensinal objects) might be given a uniform distribution on $[0, 2\pi)$ or restricted to a narrow range of angles, depending on the orientation.

(1) For a small set of landmarks, the most straightforward model for shape is given by a full multivariate normal distribution for the shape of the object. In practice this distribution is constructed either in Bookstein coordinates (Bookstein 1991), or Procrustes tangent coordinates to shape space (Kent 1994; Mardia *et al.* 1997). Training data are needed to estimate the parameters.

(2) With larger numbers of landmarks, a more restricted modelling approach is called for. One useful strategy is to use a singular multivariate normal distribution based on the dominant principal components in Procrustes tangent coordinates for some training data (Kent 1994; Cootes *et al.* 1994). Another approach imposes smoothness on the deformed template with a Markov random field model on the outline (Grenander *et al.* 1991; Kent, Mardia and Walder 1996). A block circulant covariance structure was adopted by Grenander and Miller (1994), where the ideal template was a circle with no identifiable features. Further details can be found in Kent *et al.* (1995). Alternatively, deformations of objects can be achieved through deformations of Euclidean space, using thin-plate splines (Bookstein 1991), or partial differential equations from fluid mechanics (Christiensen *et al.* 1997).
(3) The tagged heart data can be described in terms of landmarks at the intersections of initially horizontal and vertical grid lines, or alternatively, as the centers of the quadrilaterals (called "quads", say) formed by the grid lines. At each time point t during the heart cycle, a Gaussian Markov random field prior can be imposed on the quads to ensure (a) quad locations at time t are close to their positions at time $t-1$ (assumed known for this purpose), and (b) deviations in the positions of quads from their expected positions are similar for neighboring quads in the grid.
(4) For the electrophoresis problem, the parameters specify the deformation and identify the labelling between the two sets of dots in the two images. It is convenient to construct the deformation out of spline-like functions and to ensure its smoothness through derivative-based penalties. When thin-plate splines are used to construct the deformation, the penalty involves a "bending energy" matrix (Bookstein 1991; Mardia *et al.* 1996), giving rise to a an improper Gaussian distribution for the parameters of the deformation. In this application it is more appropriate to build a deformation out of a tensor product of one-dimensional splines due to the preferred roles of the x and y axes.

4 Likelihoods

Suppose the deformable template defines the outline of an object in a two-dimensional image. A simple statistical model for the pixel intensities z_l, say, is for the z_l to be independently normally distributed with constant variance σ^2 and with mean μ_1 for pixels *inside* the object, and mean μ_2 for pixels *outside* the object. Here $l = (l_1, l_2)$ labels the pixels in the image. A variation on this model is to suppose the image of the object is subject to blurring before measurement errors are added.

However, many types of image are not amenable to such straightforward statistical modelling. Variation in lighting effects, clutter in the image, and an incomplete modelling of the objects of interest may make it difficult to devise a straightforward likelihood. In these cases it is often fruitful to pre-process the image first to yield a set of features, and then to model these features. In this section we describe how these considerations can be applied to the examples mentioned above.

One of the earliest examples was the HANDS example of Grenander *et al.* (1991). They used a simple model of different means inside and outside the object, plus white noise.

However, in other examples there is not much difference in intensity between the inside and outside of objects; instead it is the edges which stand out. Thus efforts are needed to model the edges. Cootes *et al.* (1994) estimate from training data a mean profile along a normal line segment to the edge at each landmark and try to match this mean profile to the profile in the image arising from the current position of the deformable template, using a sum of squares matching criterion. It should be noted that along each normal line segment, the average value is subtracted off before matching so that differences in local intensity in the image do not affect the matching. This approach give rise to a "partial likelihood" because the only aspects of the image which enter the likelihood lie along selected normals to the current position of the deformed template. Further the likelihood is rather unusual in that the portion of the image data entering the likelihood changes as the parameters of the deformable template change.

This likelihood has the property that it is higher in value when each edge of the deformed template sits over an edge in the image. However, the peak in the likelihood can be very narrow in some applications, and if the deformable template is a moderate distance from the correct position, there is not much "force" pulling the deformable template to the correct position.

For this reason de Souza *et al.* (1997) created an alternative likelihood whose value diminishes with the distance of an edge in the deformable template to the nearest strong edge in the image. Their approach starts by pre-processing the image with an edge filter to produce a bivariate image giving the smoothed gradient in the x and y directions. Along a line segment of specified length normal to the edge at a landmark on the deformable template, the component of the gradient image along the normal can calculated and its absolute values can be scaled to sum to 1, thus giving a probability distribution. A "partial likelihood" is generated by first choosing a pixel along the normal segment from this distribution, and then by regarding its signed distance from the deformable template edge as having a Gaussian distribution with mean 0. The effect of this likelihood is to "pull" the deformable template towards the "nearest strong edge" in the image.

Next consider more abstract objects such as the tagged heart data and the electrophoresis data. Both of these applications can be described in the following framework. Consider two images on which sets of features $\{x_i\}_1^n \subset \mathbb{R}^2$ and $\{y_j\}_1^m \subset \mathbb{R}^2$ have been identified, respectively. Note that m is not necessarily the same as n and that it is not known how the labels j should be matched to the labels i. The objective is to deform the first image by a mapping Φ, say, so that the points $\Phi(x_i) = x_i'$, say, match as closely as possible to the $\{y_j\}$. The smoothness of Φ is determined by the prior. Given Φ, the $\{y_j\}$ are modelled as independent observations from the mixture of isotropic Gaussian distributions,

$$\sum_{i=1}^{n} \pi_i N_2(x_i', \sigma^2 I_2),$$

to generate the likelihood. The mixing proportions $\pi_i, \quad \sum \pi_i = 1$, are typically parameters to be estimated. It may also be useful to add an additional $(n+1)^{\text{th}}$ category ($i = 0$, say) to cater for values of y_j which do not correspond to any x_i.

In this mixture model there is a natural unobserved variable I_j, say, for each y_j, which identifies the appropriate component i of the mixture distribution. The EM algorithm can be used to help maximize the posterior density and yields posterior estimates $\hat{\pi}_{ji} = P(I_j = i|\text{data})$,

$\sum_i \hat{\pi}_{ji} = 1$. The mode of this distribution for each j is usually clearcut and gives the estimated component i for the feature j. In principle the same i could appear as the same mode for several choices of j, but in practice this will not happen for well-spaced data. Values of i not chosen as a mode for any j correspond to features in the first image not present in the second. Conversely, the modal choice $i = 0$ arises for those values of j which correspond to features in the second image not present in the first.

The tagged heart example is processed sequentially. Suppose that at time $t - 1$ we have estimates of the quad locations $\{x_i\}$. The image at time t is pre-processed to produce a large set of "potential quad" locations $\{y_j\}$ $\quad (m \cong 2n$ in practice), and a perturbation of the $\{x_i\}$ to $\{x_i'\}$ is sought. There is another complication for the heart data. The grid lines disappear more quickly in fluid (blood) than in the heart wall. Thus it is useful to allow quads to "die" over the course of the heart cycle when they can no longer be identified with any potential quads.

For the electrophoresis problem, first pre-process each image, looking *e.g.* for local maxima, to produce a set of dot-like features. Let $\{x_i\}$ denote the features in the first image and $\{y_j\}$ the features in the second. For the purposes of fitting a deformation, it is convenient to regard the $\{x_i\}$ as fixed, and to regard the $\{y_j\}$ as random coming from the mixture density.

Moss and Hancock (1996) give a similar example matching a a radar image to a digital map using features based on small line elements in the two images.

5 Summarizing the posterior distribution

Under most of the above models it is straightforward to write down the posterior density in terms of the prior specification and the likelihood. For many of the applications the posterior density will be very bumpy making deterministic optimization of the posterior density difficult. The reason for the bumpiness is the interaction in the likelihood factor between a typically smooth deformable template and a typically noisy discrete image. Thus, stochastic optimization methods such as Markov Chain Monte Carlo (MCMC) have become popular and successful in recent years.

An exception to the "bumpiness" problem is the mixture density in Section 4 where the EM algorithm can be used to help find a (local) maximum of the posterior density. Further, its use is often more successful than is generally the case in mixture problems because here one is estimating only a small set of deformation parameters rather than a whole set of means and variances.

Summarizing a distribution is relatively straightforward when the number of parameters is fixed and the distribution is reasonably unimodal. However, complications arise when the number of parameters is not fixed. For example, in the heart deformation problem, a single estimate for each quad status and location is needed at time t. That is, for each quad a decision needs to be made on whether it will be alive or dead, and the live quads need to be matched to distinct potential quads. Further, the heart problem is complicated by strong interactions between the labelling of nearby quads – if one quad is assigned to a particular potential quad, the neighbors of the quad cannot be assigned to the same potential quad. Also, a whole section of the grid of quads may be shifted one column to the right, say, without affecting the fit within that section. A heuristic method of summarizing the MCMC simulations is given in Lee *et al.* (1997).

6 Discussion

This paper has emphasized three aspects of Bayesian analysis: the prior, the likelihood, and the summarization of the posterior. There is scope for further development in each of these areas.

The range of prior models for shape variability is the most well-developed aspect of the three, but even here there is still a need for more developments in certain areas, *e.g.* methods for the regularization of high-dimensional shape variability, methods for longitudinal shape change and growth, and methods for the description of shape features involving curvature. For example, ridge curves have become popular features on surfaces in three dimensions such as faces (Kent, Mardia and West 1996). However, ridge curves can change in unpredictable ways when the underlying template is deformed. It does not seem easy to prescribe how a set of ridge curve should be deformed, and then to specify a surface which has the desired new ridge curves.

For the likelihood two questions are raised in this paper. The first is whether to model an original image or whether to pre-process the image first, *e.g.* by an edge filter or a local-maximum-finding filter. The examples of this paper suggest such pre-processing is worthwhile. The second question is whether it is better to attempt to model the whole image given the prior information, or whether it suffices just to estimate some local aspects of the image near the deformed template (*e.g.* as in the "partial likelihood" for the fish example).

Last there is the question of how to summarize the posterior density. Three popular estimates are the posterior mode (MAP estimator), the marginal posterior modes (MPM estimator), and the posterior mean. However, the last two methods do not pay any direct attention to the interactions between the parameters, and the MAP estimator is difficult to calculate given a limited number of MCMC simulations. Rue and Hurn (1997) have suggested a promising new approach using a loss function based on Baddeley's delta metric for comparing two images. More ideas of this sort are needed to take spatial context into account when comparing images. Another difficult question is how to summarize posterior distributions when the number of parameters is not fixed (*e.g.* the live and dead quads in the heart example). Green (1995) has some suggestions for summarizing features of the deformed objects, but more work is needed.

Acknowledgments

This research was supported by the Stochastic Modelling in Science and Technology (SMST) initiative of the EPSRC and by the AFRC. I am grateful to Delman Lee, Kevin de Souza, Gary Walker and Kanti Mardia for helpful discussions.

References

Bookstein, F. L. (1991). *Morphometric Tools for Landmark Data: Geometry and Biology*. Cambridge: Cambridge University Press.

Christiensen, G. E., Rabbitt, R. D., and Miller, M. I. (1997). Deformable templates using large deformation kinematics. *IEEE Transactions on Image Processing*, to appear.

Cootes, T. F., Taylor, C. J., Cooper, D. H., and Graham, J. (1994). Image search using trained flexible shape models. In K. V. Mardia (Ed.), *Statistics and Images, Volume 2*, pp. 111–139. Abingdon, Oxfordshire, UK: Carfax Publishing Co.

Green, P. J. (1995). Reversible jump MCMC computation and Bayesian model determination. *Biometrika* **82**, 711–732.

Grenander, U. (1993). *General Pattern Theory*. Oxford: Clarendon Press.

Grenander, U., Chow, Y., and Keenan, D. M. (1991). *Hands: A Pattern Theoretic Study of Biological Shapes*. Berlin: Springer.

Grenander, U. and Miller, M. I. (1994). Representations of knowledge in complex systems (with discussion). *Journal of the Royal Statistical Society, Series B* **56**, 285–299.

Kent, J. T. (1994). The complex Bingham distribution and shape analysis. *Journal of the Royal Statistical Society, Series B* **56**, 285–299.

Kent, J. T. (1995). Current issues for statistical inference in shape analysis. In K. V. Mardia and C. A. Gill (Eds.), *Proceedings in Current Issues in Statistical Shape Analysis*, pp. 167–175. Leeds University Press.

Kent, J. T. (1997). Data analysis for shapes and images. *Journal of Statistical Planning and Inference* **57**, 181–193.

Kent, J. T., Anderson, C. R., and Dryden, I. L. (1995). Deformations with circulant symmetry. Technical Report STAT 95/17, Department of Statistics, University of Leeds.

Kent, J. T. and Mardia, K. V. (1994). Statistical methodology in image analysis. In O. Ying-Lie, A. Toet, D. Foster, H. J. A. M. Heijmans, and P. Meer (Eds.), *Shape in Picture: Mathematical Description of Shape in Grey-level Images*, Volume 126 of *NATO ASI Series, Series F: Computer and Systems Sciencs*, pp. 443–452. Berlin: Springer-Verlag.

Kent, J. T., Mardia, K. V., and Walder, A. N. (1996). Conditional cyclic Markov random fields. *Advances in Applied Probability (SGSA)* **28**, 1–12.

Kent, J. T., Mardia, K. V., and West, J. M. (1996). Ridge curves and shape analysis. In R. B. Fisher and E. Trucco (Eds.), *Proceedings of the British Machine Vision Conference BMVC'96*, pp. 43–52. British Machine Vision Association.

Lee, D., Kent, J. T., and Mardia, K. V. (1997). A Bayesian approach to tracking tagged cardiac MR images. In this proceedings.

van Lieshout, M. N. M. (1997). Prior distributions for Bayesian image analysis. In this proceedings.

Mardia, K. V. (1996). The art and science of Bayesian object recognition. In K. V. Mardia, C. A. Gill, and I. L. Dryden (Eds.), *Image Fusion and Shape Variability Techniques*, pp. 21–35. Leeds University Press.

Mardia, K. V. (1997). Bayesian image analysis. *Journal of Theoretical Medicine*, to appear.

Mardia, K. V., Kent, J. T., Little, J., and Goodall, C. R. (1996). Kriging and splines with derivative information. *Biometrika* **83**, 207–221.

Mardia, K. V., Kent, J. T., and Walder, A. N. (1991). Statistical shape models in image analysis. In E. M. Keramidas (Ed.), *Computing Science and Statistics: Proceedings of the 23rd Symposium on the Interface*, Fairfax Station, VA. Interface Foundation of North America.

Mardia, K. V., McCullogh, C., Dryden, I. L., and Johnson, V. (1997). Automatic scale-space method of landmark detection. In this proceedings.

Mardia, K. V., Rabe, S., and Kent, J. T. (1995). Statistics, shapes and images. In D. M. Titterington (Ed.), *Complex Stochastic Systems and Engineering*, pp. 85–103. Oxford: Clarendon Press.

Moss, S. and Hancock, E. R. (1996). Registering incomplete radar images using the EM algorithm. In R. B. Fisher and E. Trucco (Eds.), *Proceedings of the British Machine Vision Conference BMVC'96*, pp. 685–694. British Machine Vision Association.

Rue, H. and Hurn, M. A. (1997). Loss functions for Bayesian image analysis. In this proceedings.

de Souza, K. M. A., Kent, J. T., Mardia, K. V., and Tillett, R. D. (1997). Estimation of objects in highly variable images using Markov chain Monte Carlo. Technical Report STAT 97/10, Department of Statistics, University of Leeds.

Titterington, D. M. (1997). Some aspects of Bayesian image analysis. In this proceedings.

Walker, G. (1997). Matching electrophoretic gels. In this proceedings.