A Bayesian Method for Automatic Landmark Detection in Segmented Images

Katarina Domijan DOMIJANK@TCD.IE
Department of Statistics, University of Dublin, Trinity College, Dublin 2, Ireland

Simon Wilson SWILSON@TCD.IE
Department of Statistics, University of Dublin, Trinity College, Dublin 2, Ireland

Abstract

The identification of landmark points of a figure in an image plays an important role in many statistical shape analysis techniques. In certain contexts, manual landmark detection is an impractical task and an automated procedure has to be employed instead. Standard corner detectors can be used for this purpose, but this approach is not always suitable, as the set of landmark points best representing the figure is not necessarily limited to corners. We present a Bayesian approach for automatic landmark detection, where a set of N landmark vertices is fitted to the edge of a segmented region of an image. We propose a likelihood function for the observed segmented region given the vertices and then use a Metropolis sampler to sample landmark vertices given the observed region. Careful consideration has to be given to the selection of a prior for the distribution of the landmarks.

1. Introduction

The shape of an object in a two dimensional image is often characterized by a set of N labelled points and hence is represented by an N $\times$ 2 matrix. This type of shape representation scheme can be extended to R dimensional surfaces and it satisfies the requirements of invariance to translation, scale and rotation. It therefore is a basis for many shape analysis methods. Such data arises in many applications and the corresponding labelled points are commonly called landmarks. In certain applications, for example in biological homology, landmarks are assumed to be uniquely defined locations that are identifiable across a particular class of objects or individuals. In general, it is assumed that a set of landmarks is found in at least two objects and the interest is focused on their relative positions. This abstraction allows shape theory to stand apart from issues of interpretation (Goodall, 1991).

In object recognition context, one does not a priori know the class of objects that the region of interest belongs to. Therefore, this kind of definition is not applicable. For this purpose, we define landmarks to be be a set of coordinate points that best describe a given region. The distinction between landmarks of an object and salient points of an image is that the purpose of salient points is not to summarize the shape contour of an object, but rather to represent a subset of image pixels where the image information is supposed to be most important (Sebe & Lew, 2003).

Manual landmark detection is too time-consuming in content based image retrieval applications where one might be dealing with large databases of images. Arguably it is also too subjective (Brett & Taylor, 2000). In segmented images where a region contour is clearly defined it is possible to use corner detectors such as Harris (Harris & Stephens, 1988), as well as a number of other algorithms. For the purposes of image retrieval, it is interesting to obtain information on the uncertainty of the shape retrieved, which is why a Bayesian approach is useful.

In this paper, presented is a Bayesian method for automatic detection of landmarks in pre-segmented images. The idea is to fit a set of N landmark vertices to the edge of a segmented region of interest, with the aim of describing the shape of that region well. The edge of the region is taken to be the object contour. There is a restriction for the segmented region to be

Appearing in *Proceedings of the workshop Machine Learning Techniques for Processing Multimedia Content*, Bonn, Germany, 2005.

solid, i.e. without holes. In theory, the particular segmented region would represent one object of interest in that image.

The Bayesian framework requires a likelihood function to be proposed for the observed segmented region given the landmark vertices and then a Metropolis sampler is used to sample landmark vertices given the observed region. Hence we obtain a distribution for the set of landmarks given the segmented region from which we can draw inferences on the landmarks set. In the following section of the paper this model is described in more detail. Subsequently, the method was applied to an artificial test example and a pre-segmented image of a painting from the Bridgman Art Library, London.

Two main Bayesian approaches to high level imaging, which involves working with components of an image in such tasks as object recognition are based on pattern theory (Grenander & Miller, 1994) and marked point processes (Baddeley & van Lieshout, 1993); for a recent contribution see (Hurn, 1998). The first approach uses a deformable template to represent the outline of a typical object and the natural variability is often represented by a probability measure on the parameters affecting the deformations. Kent et al. (2000) further consider some statistical aspects of this approach including maximum likelihood based. In the second approach, the images are characterized by processes of simple geometrical figures, each specified by a location and a mark containing information such as the shape and size of the figure. Rue and Hurn (1999) combine these two approaches by imbedding the template models into a marked point process framework. Other work has been done in estimating object boundaries in an image, usually with some prior knowledge of the object shape. However, these approaches differ from the one discussed in this paper in that they seek to obtain a contour of an object, as opposed to selecting a set of points that best represent an already estimated contour shape obtained from a segmentation of the image.

2. The Bayesian Model

The Bayesian approach to the problem of selecting a set of landmark points to best represent the shape of a segmented region in an image could be described as the following: the prior distribution for the scene of interest X, π(x), is combined with the likelihood of the data Y arising from a particular scene X, $\pi(y|x)$. In this particular case, X is a set of ordered N landmark points, where each point is specified by a two coordinate location vector in the image matrix. The data Y is a matrix of pixels in the segmented image, indexed as either belonging to the region of interest or not. Inferences for X are made using the posterior distribution

$$\pi(x|y) \propto \pi(y|x)\pi(x).$$

Hence π(x) is the prior distribution for the locations of landmarks and π(y|x) is the likelihood of the observed shape arising given the landmark points' locations. The rest of this section describes the model choices for π(y|x) and π(x).

2.1. Prior Distribution for X

The set of ordered landmark vertices forms a N-sided landmark polygon. Note that in this model, the number of vertices N is a constant which needs to be set by the user.

To model the fact that the landmark polygon edges are not permitted to cross over, one can specify the prior with the indicator function π(x) $\propto$ I [edges crossing].

The prior distribution does not place a restriction on the points to be on the edge of the segmented region. Also the points need not be equally spaced, as this restriction may not always result in landmarks best describing the segmented region.

2.2. Likelihood

One possible data model is an increasing function of the distance of the pixels from edge of the landmark polygon. So the data model assumed is

$$\begin{aligned}\pi(y|x,\alpha) &= \prod_{pixels(s,t)\in S} \pi(y_{st} \in S|x,\alpha) \times \\ &\times \prod_{pixels(s,t)\notin S} \pi(y_{st} \notin S|x,\alpha)\end{aligned}$$

where

$$\pi(y_{st} \in S|x,\alpha) = \begin{cases} \frac{1}{1+\exp\left(-\alpha\frac{d}{D}\right)} & \text{if } y_{st} \in \mathrm{L} \\ 1 - \frac{1}{1+\exp\left(-\alpha\frac{d}{D}\right)} & \text{if } y_{st} \notin \mathrm{L} \end{cases}$$

and

$$\pi(y_{st} \notin S|x,\alpha) = 1 - \pi(y_{st} \in S|x,\alpha).$$

S is the region of interest in the image, L is the region bounded by the landmark polygon, D is the largest minimum distance between the pixel and each edge in the landmark polygon and d is the smallest minimum distance. The likelihood term contains the unknown parameter α for which a uniform prior between 0 and a large upper bound is used.

Note that this simulation of the likelihood simply models the property that pixels from the polygon edge are

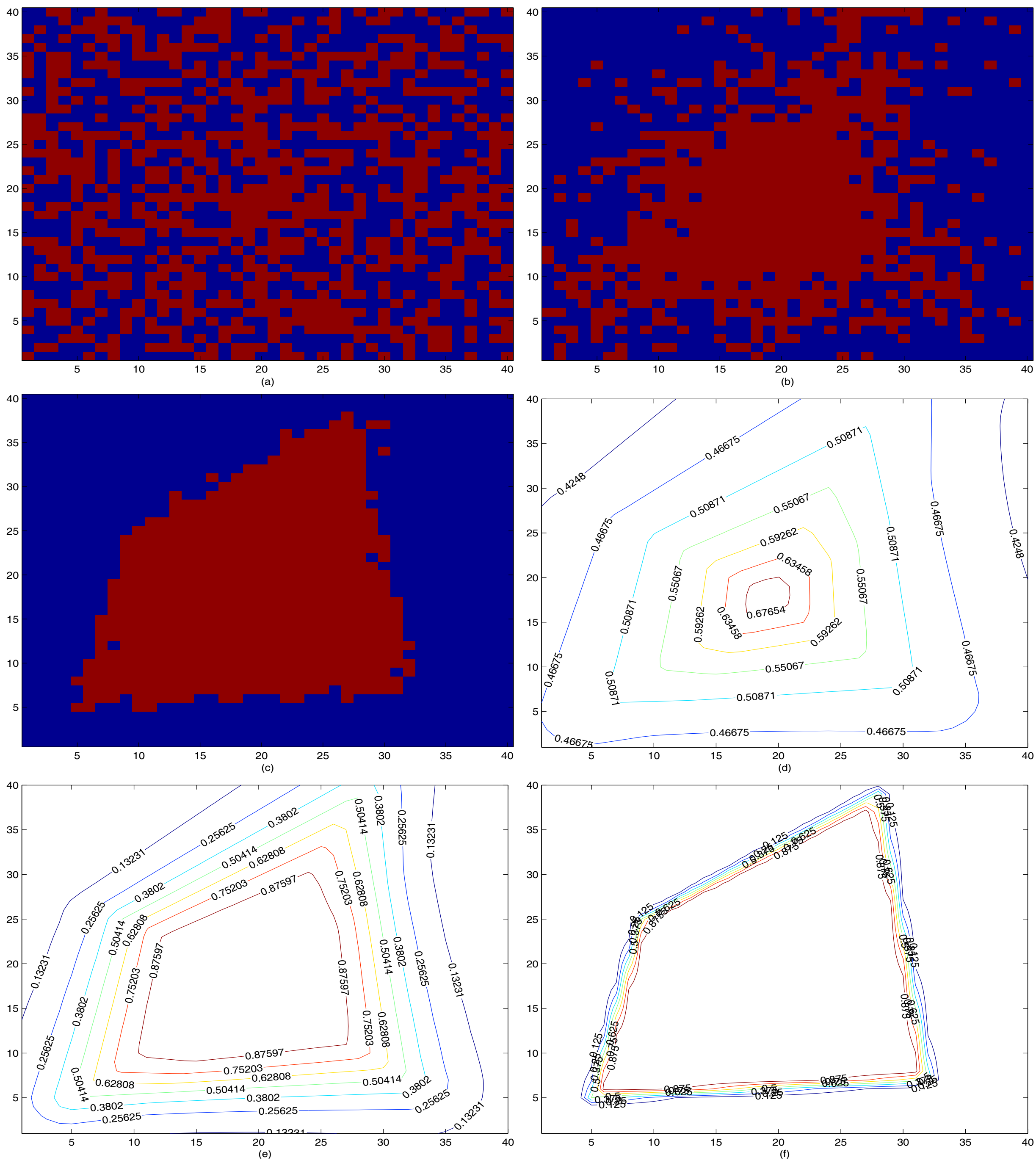

Figure 1. (a) Observed segmented region simulated for a set of landmarks with α=1, (b) α=10 and (c) α=80. (d) observed likelihood function contours for α=1, (e) α=10 and (f) α=80.

less likely to be classifies as being inside the shape. More complex likelihoods do not appear necessary.

Figures 1(a) to (c) show the observed segmented region which was simulated for a given (artificial) set of landmarks with different parameters (α=1, α=10, α=80). Figures 1(d) to (f) show the observed likelihood function contours. Hence the full posterior distribution is

$$\pi(x, \alpha|y) \propto \pi(y|x, \alpha)\pi(x)\pi(\alpha).$$

2.3. Inferences

The Metropolis algorithm (Metropolis et al., 1953) was used to obtain an iterative sequence of {x,α} that converges in distribution to $\pi(x, \alpha|y)$. The approach used was to update x and α one at a time while the other one is held fixed. The conditional distributions for both variables can be derived from the posterior distribution, the distribution of primary concern being $\pi(x|y)$:

$$\pi(x|y, \alpha) \propto \pi(y|x, \alpha)\pi(x).$$

The candidate generating density for x was set to be multivariate normal, where at each iteration of the algorithm the location of only one vertex at a time was perturbed. The vertex to be perturbed was randomly chosen.

The Metropolis algorithm requires initial values to be provided for all the variables. For parameter α, a value greater than zero was randomly chosen. From a segmented image, a starting set of landmark points can be obtained by randomly selecting their locations in the image matrix, or by first using an edge detector to obtain the edge points of the shape of interest and then randomly sampling from the edge point locations to obtain a set of N landmark points. The randomly selected initial landmark points can be reordered by an algorithm such as the nearest neighbour.

3. Results

3.1. Artificial Test Example

An artificial image was created to illustrate the sampling behaviour of the model. The shape Of interest is a simple rectangular region. An initial set of landmarks (with N=4) was obtained by randomly selecting their locations in the image matrix.

A sequence of realisations from the π(x|y) is obtained once the convergence of the algorithm appears to have been reached. In order to assess the convergence four separate simulations were run with overdispersed starting points. Figure 2 shows the sequences for all 6000 iterations for the four simulation runs.

Figure 3(a) shows four starting landmark sets and the object of interest. Figure 3(b) shows the estimates (sample means) from the posteriors of the four landmark sets. Note that the first half of the iterations of the simulation runs was discarded for the purpose of making inferences from the posterior.

3.2. Bridgman Art Library Painting

One segmented region was chosen in a pre-segmented image from the Bridgman art library. The Prewitt edge detector (Prewitt & Mendelsohn, 1966) was used to identify the edge points of the region from which 15 points were randomly selected as the starting set of

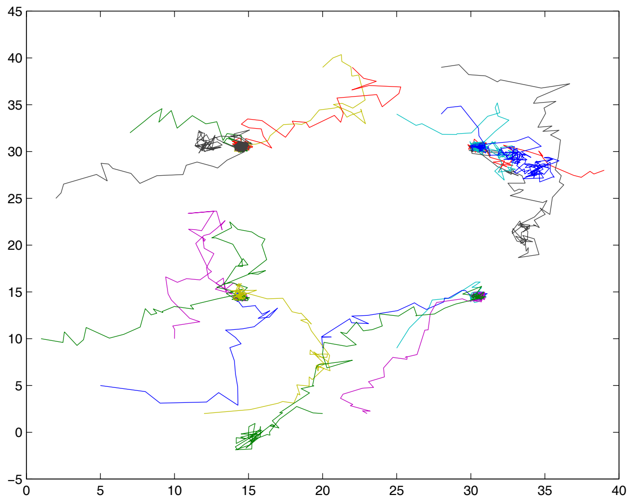

Figure 2. Four independent sequences of the simulations with different starting points. All 6000 iterations are plotted for each sequence and each landmark vertex. The starting points are indicated by crosses.

landmarks. Figure 4(a) shows the starting landmark set superimposed on the region and figure 4(b) shows the estimate of the landmark sets from the posterior distribution. Whereas the Metropolis algorithm seems to converge for the artificial test example, there are still some convergence and mixing problems with the more complicated shape.

4. Discussion

In this study, the problem of automatically generating a set of landmark points to describe the shape of a region of interest in segmented images has been attempted by using a Bayesian framework. The advantage of the Bayesian approach is that it provides information about the uncertainty of the shape, i.e. the uncertainty of how good the landmarks chosen are at summarizing the region of interest. This is particularly useful in content based image retrieval applications, which is the aim of the future research on this topic. This automatic landmark detection method will be implemented to a content based image retrieval application, where given a large database of segmented images, the shapes of segmented regions in different images are compared using Procrustes analysis.

5. Acknowledgements

Some of the images used in the development of this method are courtesy of the Bridgman Art Library London. This work is funded through the European Union

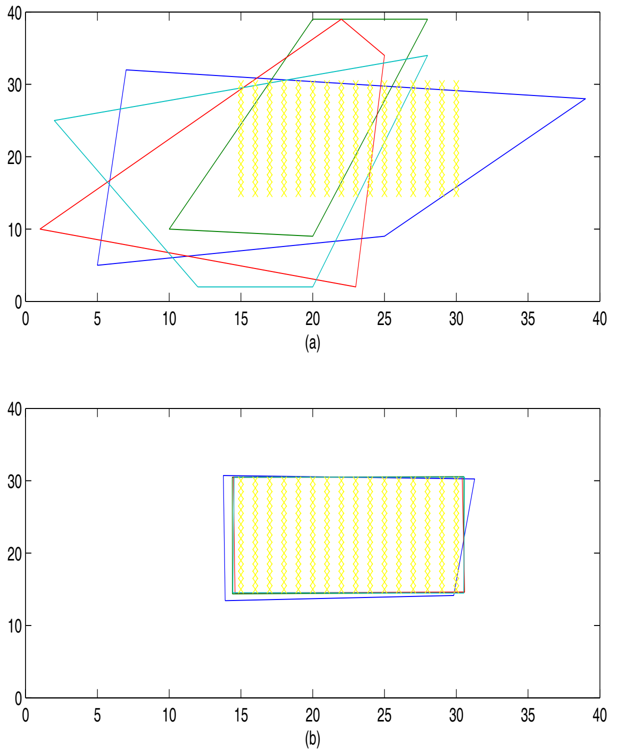

Figure 3. (a) Four starting landmark sets and the shape of interest. (b) Sample mean estimates of the four landmark sets.

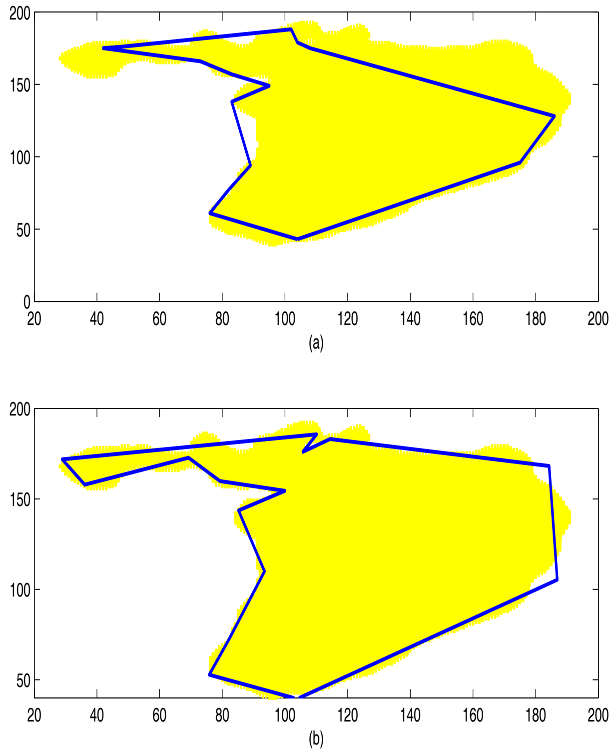

Figure 4. (a) Starting landmark set superimposed on the region. (b) Estimate of the landmark sets from the posterior distribution.

Network of Excellence MUSCLE.

References

Baddeley, A., & van Lieshout, M. (1993). Stochastic geometry models in high-level vision. In K. Mardia (Ed.), *Statistics and images*, 233–258. Abingdon: Carfax.

Brett, D., & Taylor, C. J. (2000). A method of automated landmark generation for automated 3d pdm construction. *Image and Vision Computing*, *18*, 739–748.

Goodall, C. R. (1991). Procrustes methods and the statistical analysis of shape (with discussion). *J. Royal Statistical Soc. B*, *53*, 285–340.

Grenander, U., & Miller, M. I. (1994). Representations of knowledge in complex systems. *J. Roy. Statist. Soc. B*, *56*, 549–603.

Harris, C., & Stephens, M. (1988). A combined corner and edge detector. *Proceedings of the 4th Alvey Vision Conference*, 147–151.

Hurn, M. (1998). Confocal flourescence microscopy of leaf cells: an application of bayesian image analysis. *Appl. Statist.*, *47*, 361–377.

Kent, J. T., Dryden, I. L., & Anderson, C. R. (2000). Using circulant symmetry to model featureless objects. *Biometrika*, *87*, 527–544.

Metropolis, N., Rosenbluth, N., Teller, M., & Teller, E. (1953). Equations of state calculations by fast computing machines. *J. Chemical Physics*, *21*, 1087–1092.

Prewitt, J. M. S., & Mendelsohn, M. L. (1966). The analysis of cell images. *Ann. N. Y. Acad. Sci.*, *128*, 1035–1053.

Rue, H., & Hurn, M. (1999). Bayesian object identification. *Biometrika*, *86*, 649–660.

Sebe, N., & Lew, M. (2003). Comparing salient point detectors. *Pattern Recognition Letters*, *24*, 89–96.