Mine Boundary Detection Using Markov Random Field Models

Xia Hua and Jennifer Davidson

Department of Electrical & Computer Engineering
Iowa State University

Noel Cressie

Department of Statistics
Iowa State University

Abstract

Detection of objects in images in an automated fashion is necessary for many applications, including automated target recognition. In this paper, we present results of boundary detection using Markov random fields. Once the boundaries of regions are detected, object recognition can be conducted to classify the regions within the boundaries. Thus, an approach that gives good boundary detection is very important in many automated target recognition systems. Our algorithm for boundary detection combines a Bayesian approach with a histogram specification technique to locate edges of objects that have a closed-loop boundary. The boundary image is modeled by a Markov random field. The method is relatively insensitive to the input parameters required by the user and provides a fairly robust automated detection procedure that produces an image with closed one-pixel-wide boundaries. We apply our method to mine data with very good results.

1. Introduction

This research presents the results of solving the mine detection problem using boundary identification of objects. Our data typically have low signal to noise ratio and there is background clutter and noise. Consequently, the usual engineering approaches to edge detection, followed by thinning and linking, proved to be unsuccessful. These approaches use local pixel intensity information to identify whether a pixel location is part of a boundary. Filters are used to collect local gradient information and, if the magnitude of the local gradient is large enough, the pixel is declared an edge pixel. Unfortunately, such techniques are sensitive to noise and, in addition, global boundary information is not available, which hinders the determination of closed object boundaries. A closed object boundary is one that forms a loop, and has no gaps as the boundary is traced.

In this research, we develop a statistical method to obtain statistically optimal one-pixel wide closed object boundaries in gray-level images. We first apply a segmentation algorithm (Helterbrand *et al.*, 1994a; Hong and Rosenfeld, 1984) to obtain an initial estimate of one-pixel-wide closed boundaries in image data. Next, the iterated conditional modes (ICM) algorithm (Besag, 1986) is applied to produce a maximum-a-posterior (MAP) estimate of the boundary image. At this point in the processing, the boundaries found are used to produce two sets of regions in the image: 1)

targets which are inside the small boundaries, and 2) background, which are the remaining regions. Contrast stretching using histogram specification is performed to separate the gray values on the targets from those of the background. This histogram-specified image is then used as input to the segmentation process as a “better” estimate of the boundaries. ICM is run a second time, producing a final MAP boundary estimate. As applications of this algorithm to real data show, this algorithm is a fairly robust automated detection procedure that locates the boundaries of objects in mine data.

2. Markov Random Field Boundary Models

In this section, we introduce the posterior boundary model by specifying an intensity model and a class of prior boundary models. The posterior boundary model will have support only on a subset of the total boundary configuration space (Helterbrand *et al.*, 1994a).

Let D denote an $n \times m$ rectangular array of pixel locations. Let $Y(s)$ represent a random variable that denotes the observed intensity at pixel location $s \in D$, and let $X(s)$ denote the unobserved, true underlying intensity at pixel location s . It is assumed that $\{X(s) : s \in D\}$ is a random field with realizations that are constant within four-connected regions of D .

We assume an object in the image is a region whose intensity values are (approximately) constant. Thus, based on the premise that objects are defined as regions of statistical homogeneity, it is reasonable to model the intensity probability mass function (pmf) as,

$$P(y \mid \omega) \equiv P(Y(s) = y(s) : s \in D \mid \omega). \quad (1)$$

Here, $y \equiv \{y(s) : s \in D\}$ and $w \equiv \{w(s) : s \in D\}$, where $w(s) = 1$ if a boundary is present at location s , and $= 0$ otherwise.

In Equation (1), we restrict the boundary ω to a subset of all the possible boundary images. Specifically, restrictions relating to the width of the boundaries (one-pixel wide), connectivity properties (eight-connected closed boundaries), and other restrictions are placed on the final properties we want for the boundaries, which we collectively term “permissible” boundaries. We denote the set of permissible boundaries by Ω_p . The set Ω_p allows us to search a much smaller set for good boundary estimates and, at the same time, guarantee that each such estimate has the desired properties of closed-loop and single-pixel width. See Helterbrand *et al.*, (1994a) for more details.

Further, a Gaussian model is assumed for the observed intensities, $Y(s) = X(s) + \epsilon(s)$; $s \in D$, where ϵ is a normal white noise process with zero mean and variance σ^2 , $\epsilon(s) \sim N(0, \sigma^2)$, and the ϵ -process is independent of the X -process. We note that each $w \in \Omega_p$ partitions D into disjoint four-connected regions. Let $d(\omega)$ denote the generic four-connected regions for a particular ω . The number of disjoint connected regions depends on ω and will be denoted by $K(\omega)$. Then, each $w \in \Omega_p$ implies a partition of D into disjoint connected regions $\{d_i(w) : i=1, \dots, K(\omega)\}$, where it is assumed that $X(s)$ is constant on connected regions. That is, $X(s) = X(t)$ if $s, t \in d_i(w)$; $i=1, \dots, K(\omega)$.

Define $\mu_i(\omega) \equiv X(s); s \in d_i(\omega)$. For $\omega \in \Omega_p$, it follows that

$$P(Y(s) : s \in D | \omega) = \prod_{i=1}^{K(\omega)} P(Y(s) : s \in d_i(\omega) | \omega) ,$$

where $Y(s) | \omega \sim N(\mu_i(\omega), \sigma^2); s \in d_i(\omega)$. Thus, upon letting $n_i(\omega)$ denote the number of distinct sites $s_{ij} \in d_i(\omega) \subseteq D$, the observed intensity model is specified as

$$P(Y(s) : s \in D | \omega) = \prod_{i=1}^{K(\omega)} \prod_{j=1}^{n_i(\omega)} \left(2\pi\sigma^2 \right)^{-\frac{1}{2}} \exp \left\{ -(Y(s_{ij}) - \mu_i(\omega))^2 / 2\sigma^2 \right\} ,$$

where $\{\mu_i(\omega) : i = 1, \dots, K(\omega)\}$ and $\sigma^2 > 0$ are parameters.

Next, we describe a class of Markov random field (MRF) prior boundary models. A Gibbs distribution relative to $\{D, N\}$ is defined as a probability measure $P(\omega)$ on Ω_p with the following representation:

$$P(\omega) = \frac{1}{Z_o} \exp(-U_o(\omega)); \quad \omega \in \Omega_p ,$$

where $U_o(\omega) \equiv \sum_{c \in C} V_c(\omega)$, C denotes the set of cliques for the neighborhood system N , and Z_o is the partition function, or normalizing constant, defined by $Z_o \equiv \sum_{\omega \in \Omega_p} \exp(-U_o(\omega))$. The function

$U_o(\omega)$ is referred to as the energy function and $V_c(\omega)$ is called a potential function. Each V_c is a function on Ω_p with the property that $V_c(\omega)$ is a function only of those values of $\omega(s)$ for which $s \in c$.

By Bayes' Theorem, the resulting class of posterior boundary models for $\omega \in \Omega_p$ given the observed intensities $\{Y(s) : s \in D\}$ is,

$$P(\omega | Y(s) : s \in D) \propto \exp \left\{ - \left(U_o(\omega) + \sum_{i=1}^{K(\omega)} \sum_{j=1}^{n_i(\omega)} (Y(s_{ij}) - \mu_i(\omega))^2 / 2\sigma^2 \right) \right\} .$$

Since $\{Y(s) : s \in D\}$ is given, define $U_1 = U_o + \sum_{i=1}^{K(\omega)} \sum_{j=1}^{n_i(\omega)} (Y(s_{ij}) - \mu_i(\omega))^2 / 2\sigma^2$ and we see that the posterior boundary pmf is also a MRF on Ω_p , $P(\omega | Y(s) : s \in D) = \frac{1}{Z_1} \exp(-U_1(\omega)); \quad \omega \in \Omega_p$, and Z_1 is the partition function that ensures the posterior pmf sums to unity. See (Cressie, 1991, Section 7.4) for more discussion on image analysis using MRFs.

3. MRF Boundary Models for Mine Detection

Given a specified posterior boundary pmf on Ω_p , we developed an approach to identify the boundaries of mines and mine-like objects. We performed mine detection as shown in the flowchart in Figure 1. The algorithm begins by using a multi-resolution image segmentation algorithm to

obtain a mostly data-driven initial labeling of boundary and nonboundary pixels. Statistical label models additionally require a priori knowledge of the number of labels to use and their corresponding parameters; however, this information is not known in our case. Thus, the segmentation algorithm output is used to estimate the number of labels and the corresponding model parameters. Given the output from the segmentation process, the statistical labeling algorithm uses iterated conditional modes (ICM) to search for a statistically (local) optimal labeling. ICM provides a suboptimal local solution and, thus, the input to the ICM must be a good initial guess of the boundary image.

The data we used was one of six multi-spectral images of one scene. The scene contained mines and mine-like objects. The bandwidth of the data we used was in the range 375 – 425 nanometers.

We next describe a general overview of the steps of the algorithm as shown in the flowchart. The original subimage was 256 x 256. For pragmatic reasons (explained below), first obtained a 64 x 64 subimage from the original image using local averaging; this is shown in Figure 3. We apply the segmentation algorithm to the image in Figure 3, followed by a step to produce a permissible closed boundary image. See Figure 4 for the permissible image. Then, applying the ICM algorithm to the image in Figure 4, we obtain a maximum-a-posterior (MAP) boundary estimate as shown in Figure 5. This MAP boundary estimate is further refined by removing boundary line segments in between two regions having the same gray values. This cleaned-up image is shown in Figure 6. Next, target and background regions were identified. Target regions are those inside the closed boundaries. Histograms were calculated, one each for these two regions. The histograms are shown in Figures 7 and 8. Next, histogram specification transformations were performed on these two histograms to separate target gray values from background gray values. The purpose of separating target and background pixels is to provide a better input to the segmentation part of the algorithm, as the image is run through the process described above a second time. The histogram-specified image is shown in Figure 9, and its histogram is given in Figure 10. Next, the histogram-specified image is input a second time to the segmentation algorithm, and the entire processing sequence performed again. The output from the second pass is shown in Figure 11. The final result, a MAP boundary estimate, is shown in Figure 12.

Figure 1. Flowchart of data processing.

Figure 2. 480 x 720 original image.

Figure 3. 64 x 64 local-averaged image.

Figure 4. Initial estimate after segmentation and before ICM.

Figure 5. MAP boundary estimate on first pass.

Figure 6. Cleaned MAP boundary estimate.

Figure 7. Histogram of the target regions in the original image.

Figure 8. Histogram of the background regions in the original image.

Figure 9. Histogram-specified image.

Figure 10. Histogram of the transformed image.

Figure 11. Second initial estimate after segmentation and before ICM.

Figure 12. Final MAP boundary estimate.

Next, we describe in detail the steps shown in the flowchart given in Figure 1.

From the right side of the 480 x 720 original image shown as Figure 2, we extracted a 256 x 256 subimage, using the software called Khoros (Konstantinides and Rasure, 1994), and reduced the subimage to a 64 x 64 image using a 4 x 4 local averaging method. The 64 x 64 image is shown in Figure 3. We chose this size of image to make run-times realistic. Larger images would have given even larger run-times than the 13–16 hours we experienced for a 64 x 64 image. The data was run on a Sun Sparcstation I, and we estimate a speedup by a factor of 6 – 8 possible on current workstations, such as a Dec Alpha.

We applied the modified segmentation algorithm to the 64 x 64 image. The segmentation algorithm needs five parameter values that are critical to obtain a good initial estimate to input to the ICM algorithm. Our segmentation algorithm is detailed in Helterbrand *et al.*, (1994b) and is based on earlier work by Hong and Rosenfeld (1984). This algorithm uses a multi-resolution or pyramid image representation. The root node of the pyramid is the original image, that is, the image data at its finest resolution. The succeeding levels of the pyramid are constructed by grouping a 4 x 4 neighborhood of pixels from the preceding level. The main idea in using pyramids is that at each level of reduced resolution, homogeneous regions (regions with approximately constant gray value) become more compact and hence easier to segment. The argument for partitioning the image this way is that region classification is generally more powerful than pixel classification and, at each level in the pyramid, information is collected over neighborhoods and passed up to the next level. "Parent" nodes at higher levels are connected by link strengths to their 16 "children" at the next lower level. The basic idea of the algorithm is to define link strengths between parent/child pairs on adjacent levels of the pyramid, combining information on gray value similarity and spatial proximity of the pairs. Link strengths are computed iteratively, from the root node to the leaf nodes, for the number of iterations specified. Typically, only 8-10 iterations are necessary. See Helterbrand *et al.*, (1994b) for more details.

The five parameters that the user must specify to run the segmentation algorithm are:

1. Variance: The variance between levels in pyramid (modeled as Gaussian).
2. Alone: The "Alone" parameter governs the level at which a region becomes defined, which is directly proportional to the number of pixels in that region.
3. Merge: Merge two regions if the gray value difference between the two regions is less than the "Merge" value.
4. Iteration: The number of passes through the entire pyramid.
5. Delta: A large value means that the region is more compact and less spread out.

In our experiments, we found good response for the following parameter values: Variance = 100, Alone = 0.015, Merge = 14, Iteration = 8, and Delta = 2.5. Then, with minor changes, the boundaries found by the segmentation algorithm were made into permissible closed boundaries, using the algorithm developed by (Helterbrand, 1993). This permissible closed boundary estimate is given in Figure 4.

With this initial estimate, we next applied the ICM algorithm to the image and obtained a MAP boundary estimate; this is shown in Figure 5. We remark that in the context of the MAP closed boundary problem, the ICM algorithm automatically selects the mode of the current candidate

boundaries. Such a choice guarantees that $P(w^{(t)}) \geq P(w^{(t-1)})$, where t represents the number of iterations. However, if the ICM algorithm converges, it is most likely to a local maximum. The user selects several parameters, namely one that governs the total numbers of pixels on the boundaries, another that penalizes small regions (such as 1-pixel in size), another that controls the smoothness of the boundary, and a final one for the Gaussian parameter σ^2 for the intensity model.

The next step was to remove line segments between regions with the same gray values; the result is shown in Figure 6. A line in the image that contains (approximately) the same region on two sides is removed. The effect of removing line segments results in a better estimate, as we are looking for nontrivial closed-loop boundaries.

Next, we separated the gray values of the targets from those of the background. First, we filled in the target regions of the previous image. We combined the filled-in image with the locally averaged 64 x 64 image to separate the background from the filled-in regions, and calculated both target and background histograms, which are shown in Figure 7 and Figure 8, respectively. Based on these histograms, we applied histogram specification to separate the gray values on the targets from the those of the background. In our case, the original gray levels in the range 60-150 were transformed to the range 10-20, and all other values remained same. This method separated the targets from the background, as shown in Figure 9. The histogram of the transformed image is shown in Figure 10. This completes the first step of processing on the image.

Next, we used the histogram-specified image as an input image to a second pass of processing. We performed the segmentation algorithm using the first MAP boundary estimate as a starting value and obtained the second initial estimate, as shown in Figure 11. Comparing this with the estimate shown in Figure 4, we can see that the second one provides a much cleaner estimate of boundaries around the targets. Finally, we applied the ICM algorithm to this estimate and obtained the MAP boundary estimate shown in Figure 12. This final boundary estimate not only indicated the correct positions of the targets but also outlined the shapes of the targets.

4. Conclusions

In this research, we have developed a process that can successfully detect noisy targets in a fairly automated manner, even with a local optimization technique such as the ICM algorithm. The MRF boundary model is a very general one, and we have detected the boundary of mines and minelike objects in real data using these models. We believe that this method is a general and potentially powerful one that warrants further investigations and applications. We are currently working on ways to decrease the processing time from the order of hours down to fractions of hours.

5. Acknowledgments

This work was supported from grants by ONR (N00014-93-1-0001) and NSF (DMS-9204521). The authors gratefully acknowledge the use of J. Helterbrand's code in executing the algorithm

on real data.

Bibliography

- (Besag, 1986) J.E. Besag. On the statistical analysis of dirty pictures. *J. R. Stat. Soc. B*, 48:259–302, 1986.
- (Cressie, 1991) N.A.C. Cressie. *Statistics for Spatial Data*. Wiley and Sons, NY, 1991.
- (Helterbrand *et al.*, 1994a) J. D. Helterbrand, N. Cressie, and J. L. Davidson. A statistical approach to identifying closed object boundaries in images. *Advances in Applied Probability*, pages 831–854, Dec. 1994.
- (Helterbrand *et al.*, 1994b) J. D. Helterbrand, J. L. Davidson, and N. Cressie. Closed boundary identification in gray-scale imagery. *to appear in J. of Mathematical Imaging and Vision*, 1994.
- (Helterbrand, 1993) J. D. Helterbrand. *Spatial Dependence Models and Image Analysis*. PhD thesis, Iowa State University, Ames, IA, 1993.
- (Hong and Rosenfeld, 1984) T. H. Hong and A. Rosenfeld. Compact region extraction using weighted pixel linking in a pyramid. *IEEE Trans. on Patt. Analys. and Mach. Intell.*, PAMI-6:222–229, 1984.
- (Konstantinides and Rasure, 1994) K. Konstantinides and J. R. Rasure. The Khoros software development environment for image and signal processing. *IEEE Trans. on Image Processing*, 3(3):233–242, May 1994.