Coding Theoretic Approach to Segmentation and Robust CFAR-Detection for LADAR Images

Unoma Ndili[a], Richard Baraniuk[a*], Hyeokho Choi[a], Robert Nowak[a] and Mário Figueiredo[b]

[a]Rice University
Department of Electrical and Computer Engineering
Houston, Texas, U.S.A.

[b]Instituto de Telecomunicações, and
Instituto Superior Técnico
1049-001 Lisboa, Portugal

ABSTRACT

In this paper, we present an unsupervised scheme aimed at segmentation of laser radar (LADAR) imagery for Automatic Target Detection. A coding theoretic approach implements Rissanen's concept of Minimum Description Length (MDL) for estimating piecewise homogeneous regions. MDL is used to penalize overly complex segmentations. The intensity data is modeled as a Gaussian random field whose mean and variance functions are piecewise constant across the image. This model is intended to capture variations in both mean value (intensity) and variance (texture). The segmentation algorithm is based on an adaptive rectangular recursive partitioning scheme. We implement a robust constant false alarm rate (CFAR) detector on the segmented intensity image for target detection and compare our results with the conventional cell averaging (CA) CFAR detector.

Keywords: Laser Radar, Segmentation, CFAR, MDL, Automatic Target Detection, Adaptive Recursive Partitioning

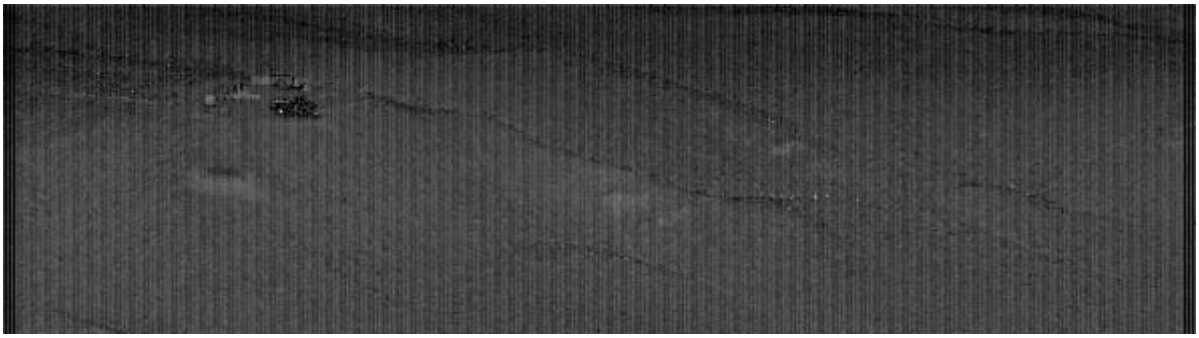
1. INTRODUCTION

Laser Radar (LADAR) is a promising imaging technology because it is able to collect high resolution 2-D and 3-D images by raster scanning a field of view. A LADAR sensor typically collects two channels of data: intensity and range data. An illustrative example of LADAR intensity and range image is shown in Figure 1. The laser radar transmits a series of laser pulses, one for each pixel. Each pulse passes through the atmosphere and reflects off the first opaque objects it contacts. The reflected light returns to the sensor and then undergoes optical heterodyne detection followed by intermediate frequency (IF) filtering and peak detection. The intensity image is a measure of the reflected energy which takes lighting, color variations and textural properties etc. into account. The range image is formed by measuring the time delay between the peaks of the transmitted and the detected waveforms. The range data is a measure of physical distance which unlike intensity, is unaffected by illumination from other sources such as the weather, internal heat, texture etc.[1] This basic difference in the formation of the intensity and range images inspires the difference in the way we process them. LADAR images are degraded by the combined effects of laser speckle and local oscillator shot noise. The former is due to the rough-surfaced nature of encountered objects when compared to the laser wavelength which causes constructive and destructive interference in the reflected signal. The latter is the fundamental noise involved in optical heterodyne detection. Speckle noise degrades LADAR images through anomalies which occur when a deep speckle fade combines with a strong noise peak resulting in measurements which is substantially different from the true value.[2]

The main contribution of this paper is a novel segmentation scheme aimed at automatically identifying regions of important similar statistical characteristics in LADAR images. Specifically, our model enables capturing variations in both mean and variance in one step. Segmentation is a necessary step for autonomous target detection and recognition required in current military practices. An imaged scene may contain man-made objects which are

Correspondence: e-mail[a]: {unondili, richb, choi, nowak}@ece.rice.edu, Web: www.dsp.rice.edu

*This work was supported by the Raytheon grant and the ONR grant N00014-99-1-0813.



(a)

(b)

Figure 1. Original LADAR Images (a) intensity and (b) range

'smooth' and/or natural objects which contain texture information. The goal in segmenting an image of such a scene would then be to identify these objects with different statistical characteristics before target recognition can be performed. Several algorithms proposed in the literature perform segmentation based only on one of these features (intensity or texture)[3,4] or use a multi-stage decoupled scheme.[5,6]

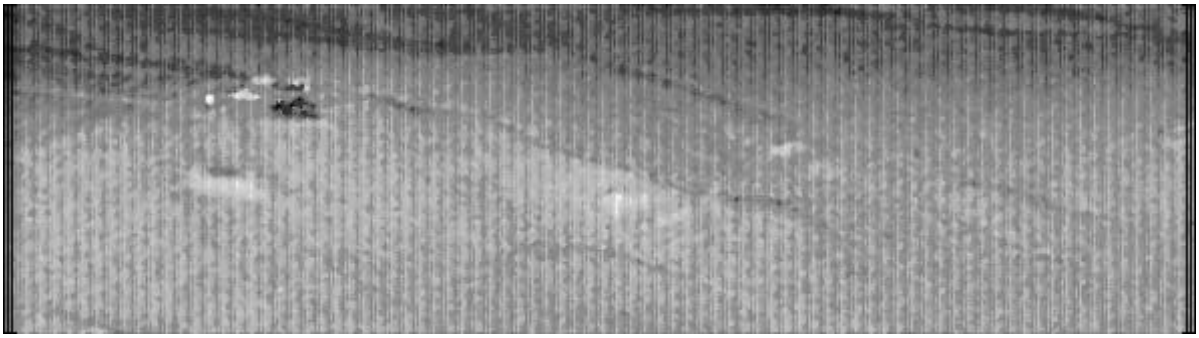
The proposed segmentation algorithm is able to detect changes in intensity and/or variance. To the best of our knowledge, this is the first segmentation scheme that achieves this. Our image model is a Gaussian random field whose mean and variance functions are piecewise constant. That is, we assume that the m-by-n image under study is composed of connected regions of pixels with each pixel independently and identically distributed (i.i.d.) according to a Gaussian density. Our model is:

$$y(i_1, i_2) = f(i_1, i_2) + \sigma(i_1, i_2)\, z(i_1, i_2), \quad 0 \leq i_1, i_2 \leq m, n, \tag{1}$$

where $z(i_1, i_2) \sim$ i.i.d. $\mathcal{N}(0, 1)$ so that $y(i_1, i_2)$ is distributed according to $\mathcal{N}(f(i_1, i_2), \sigma^2(i_1, i_2))$. This is different from the signal plus constant noise models used in most wavelet-based multiscale image analysis procedures.[7] It finds applications in a wide class of images and is appropriate for the LADAR intensity image.

We use a predictive coding theoretic approach based on Rissanen's MDL principle[8–10] to achieve unsupervised segmentation. MDL is the tool used to penalize overly complex models thereby providing a means for data dependent model selection. We propose the **Adaptive Recursive Partitioning** (ARP) algorithm[11] which is a top-down, fast and recursive scheme that achieves rectangular tessellations of the feature space. We explore the performance of our segmentation scheme on real intensity and range LADAR images, leading us to develop a robust Constant False Alarm Rate (CFAR) detector scheme.

Our paper is organised as follows: Section 2 describes the segmentation scheme: The Adaptive Recursive Partitioning algorithm and briefly reviews the MDL principle. Section 3 describes the input LADAR imagery and the



(a)

(b)

Figure 2. Median filtered LADAR images (a) intensity (b) log-range

preprocessing performed on the range data. Section 4 explains the robust CFAR detection algorithm developed based on our segmentation scheme, with results and comparisons with the standard CFAR algorithm used on real LADAR data. We conclude in Section 5.

2. THE ADAPTIVE RECURSIVE PARTITIONING ALGORITHM

The **Adaptive Recursive Partitioning Algorithm** (ARP) is based on a coding theoretic concept. Suppose we want to optimally encode and transmit a finite set of data sequence $X^n = (x_1, \cdots, x_n)$ of length n to some decoder. The ARP determines whether the data is best represented as a homogenous region of common mean and variance under the model class 1 or split into two or four homogenous rectangular regions, each region having a different mean and/or variance, under model class 2. Each subblock is recursively examined and tested in this way until all subblocks are represented as homogenous. Applying a parametric model on the data in a homogenous subblock and adopting a progressive coding scheme, we first transmit sufficient statistics for the estimated parameters (in our case the sample mean and sample variance) of a subblock, and then transmit the data in that subblock based on the parametric model assumed. This progressive scheme enables us take advantage of previously transmitted information since if a block is split into H subblocks say, we only need to transmit sufficient statistics for $H-1$ of the subblocks because given mean and variance of the original block and $H-1$ of its subblocks, it is trivial to derive the statistics of the last subblock.

This decision to either split a region or retain it as homogenous is made by applying the MDL criteria.[7,12,9] If we decide to split a region, the 'best' split point s is again decided based on the MDL principle. We briefly describe the MDL principle in the subsection below.

2.1. MDL Principle

The MDL principle enables data dependent model selection.[13,10,7] Given a probability distribution $p(X^n|\theta)$ on X^n, where θ parameterizes the distribution, the Shannon-optimal code length for prefix codes is given as $-\log p(X^n|\theta)$.[14,15] Now assume we have a set of K competing models to explain our data $\{p_i(X^n|\theta_i)\}_{k=1}^K$, the MDL criteria states that among these K possibilities, the "best" model is the one that minimizes the description length obtained by assuming a *two-part* code,

$$L(X^n) = L(\widehat{\theta}_k) - \log p(X^n|\widehat{\theta}_k), \quad k = 1, \ldots K, \tag{2}$$

where $L(\widehat{\theta}_k)$ is the code length required to describe the maximum likelihood estimate (MLE) $\hat{\theta}_k$ of the parameter θ_k such that the decoder knows the model under which the code for the data was obtained. This is a penalized likelihood with $L(\widehat{\theta}_k)$ as the penalty term. To encode the parameter estimate $\widehat{\theta}_k$ under a Gaussian data assumption, the elements of $\widehat{\theta}_k$ which are real-valued have to be truncated to finite precision in order to yield a finite code-length $L(\widehat{\theta}_k)$. The standard solution is the well known $\frac{1}{2}\log n$ code-length for each component of θ_k, based on asymptotic approximations.[8,9] MDL has been successfully used in several image analysis/segmentation problems.[11,16]

In the MDL approach described above, we first encode and send parameter estimates, then the data itself, coded according to that parameter estimate. The sufficient statistics that have to be collected to obtain estimates of the mean f and variance σ^2 are the sum $\sum_{i=1}^n x_i \equiv t_1$ and sum of squares $\sum_{i=1}^n x_i^2 \equiv t_2$. These constitute the first-part information that is initially sent to the decoder. For the second part, to optimally code the data given that the decoder already knows t_1 and t_2, we require the conditional distribution $p(X^n|\sum_i^n x_i, \sum_i^n x_i^2)$ based on which we build the code for the data according to Shannon's codelength. This conditional distribution is given by[17]:

$$p(X^n|t_1,t_2) = \begin{cases} \frac{\Gamma(\frac{n-1}{2}) r^{2-n}}{2\pi^{\frac{n-1}{2}}} & \Leftarrow \quad X^n \in \mathcal{C}(t_1,t_2) \\ 0 & \Leftarrow \quad X^n \notin \mathcal{C}(t_1,t_2), \end{cases} \tag{3}$$

where $\Gamma(\cdot)$ denotes Euler's gamma function, $C(t_1,t_2) = (X^n : \sum x_i = t_1 \text{and} \sum x_i^2 = t_2)$ is the constraint set and $r = \sqrt{t_2 - \frac{{t_1}^2}{n}}$.

With Eq.(3) used in Eq.(2) for the MDL criteria, we may now derive simple codelengths for each model class in the ARP algorithm. First we define some general notation: Given any region or subblock R of size n_R-by-m_R, let X^R describe the data in region R, i.e., $x \in R$ with common mean f_R and variance σ_R^2. Define $t_1^R \equiv \sum_{x\in R} x$ and $t_2^R \equiv \sum_{x\in R} x^2$. We may split R into H non-overlapping subblocks each represented as R_h for $h = 1, \cdots H$.

2.2. ARP Algorithm

The ARP examines R to determine whether it is best represented as one homogenous block under model class 1, or else split into either two rectangles (horizontally or vertically) or four rectangles with one common vertex, under model class 2. Let R_h ($h = 1, 2$ or $1 \cdots 4$) represent the resulting subblocks after a split. Under model class 2 each subblock is considered homogenous and different from the others.

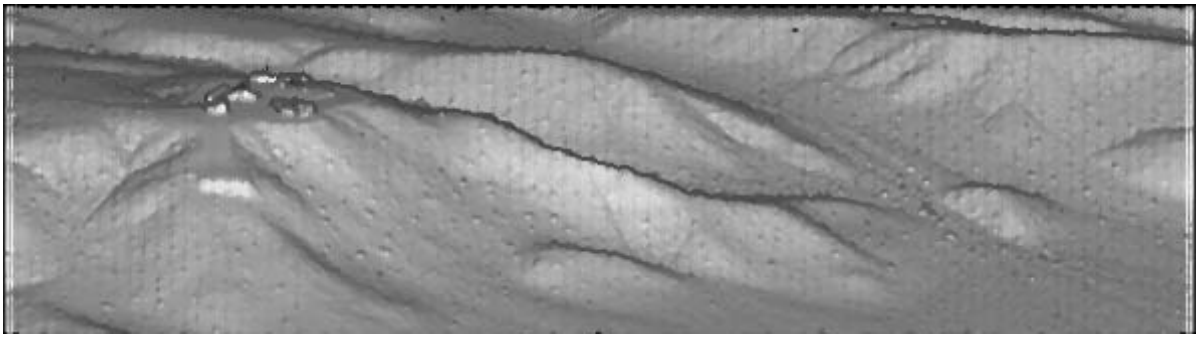
Model Class 1: Assuming sufficient statistics for the model parameters have already been transmitted, we only need to code and transmit the data as a homogeneous block. Given the conditional density in Eq.(3), the description length for the model class is simply:

$$L_1(R) = -\log p(X^R|t_1^R, t_2^R)$$

Partitioning of a subblock is stopped once it is determined to best be described as homogenous.

Model Class 2: Under this model class we may split R into 2 or 4 rectangles. The 'best' split point s is determined using the MDL criteria. We examine all possible split points and decide on the one that gives the minimum codelength. If we decide to split, then we must send the value of s to the decoder in addition to the sum and sum of squares of the resulting regions, i.e., $t_1^{R_h}$ and $t_2^{R_h}$. Since there are $J = (m_R - 1)(n_R - 1)$ possible split positions, we require $\log J$ bits to code s. We emphasize that if we split R into H rectangles ($H = 2$ or 4) we only need to send sufficient statistics for the mean and variance of $H - 1$ of these regions since we already have the statistics for R due to our progressive scheme. As an example, there are J ways to split R into two rectangles $R_1(j)$ and $R_2(j)$, $j = 1, \cdots J$, so that the description length under model class 2 is:

$$L_2\{R(j)\} = -\log p(X^{R_1}|t_1^{R_1(j)}, t_2^{R_1(j)}) - \log p(X^{R_2}|t_1^{R_2(j)}, t_2^{R_2(j)}) + \log J + \log n_R m_R$$



(a)

Figure 3. Gradient image of log range data

Comparing codelengths for the different model classes, if $L_1(R) < \min L_2\{R(j)\}$ for all j, model class 1 is chosen for the subblock and processing is stopped, otherwise we split at the point s which minimizes code lengths obtained from $L_2\{R(j)\}$. Each rectangular region is re-examined and split recursively under model class 2 until either model class 1 is selected or we get to the trivial pixel level. This algorithm was originally developed for Poisson data by Nowak et al.[11] It is a greedy algorithm as it starts from the top with the entire image and proceeds to subdivide downwards until all subblocks are represented as homogenous.

3. PREPROCESSING OF LADAR IMAGE

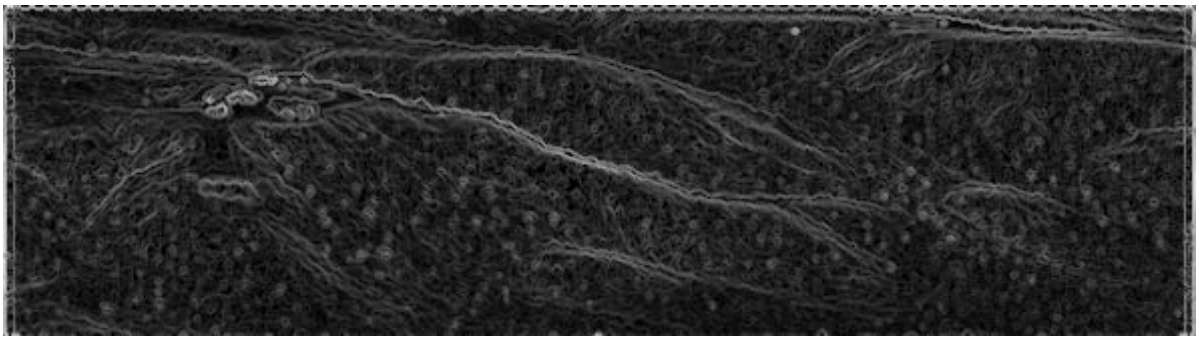
Figures 1(a) and (b) show a typical LADAR intensity and range image. They are real data sets collected from experimental ground-imaging sensors aboard an aircraft. The output has been scaled to 256 gray levels. In the intensity image, strong reflections are displayed as bright whites while negligible reflectance is in dark gray. For the range image, the lighter shades indicate greater distance from the sensor to the scene point. The images are 164×594 pixels containing some targets situated on a hill in a sloping terrain. The pixel spacing is assumed to be large enough so that the measurements are statistically independent. Previous work[2,18] has shown that LADAR images have very large-error pixels which are uniformly distributed, with the remaining pixels having a Gaussian distributed error. Therefore the single-pixel statistical model of the observed $m \times n$ LADAR data $L_i$for$i = 1, \cdots, mn$ given the true values $\overline{L_i}$, has the joint probability density function given as[2]:

$$Pr_{l|\overline{l}}(L|\overline{L}) = \prod_{i}^{mn} \left[[1 - Pr(A)] \frac{\exp\left(-\frac{(L_i - \overline{L}_i)^2}{2(\delta L)^2}\right)}{\sqrt{2\pi(\delta L)^2}} + Pr(A)\frac{1}{\Delta L} \right] \text{ for } L_{min} \leq L_i \leq L_{max}$$

where $Pr(A)$ is the probability of anomaly, i.e., the probability that speckle and shot noise effects combine to give a measurement more than one resolution cell from the true value; ΔL is the width of the laser radar's uncertainty interval and δL is the local measurement accuracy (or precision). The first term equals the probability that the measurement is not anomalous times a Gaussian p.d.f. with mean equal to the true range value. The second term is the probability that the pixel is anomalous times a uniform distribution over the entire uncertainly interval. Our initial preprocessing is aimed at suppressing this second term so that our Gaussian model may be used. We achieve this by a using a simple median filter. Even though more sophisticated methods exist to reduce these sensor dependent effects (e.g., maximum likelihood wavelet based method[2]), this *ad hoc* median filter is sufficient for our application. Figure 2(a) shows the median filtered intensity image.

Previous work on the statistics of LADAR range data[19] show that the logarithm of the range data is shape invariant and easier to work with. This is seen from the difference between the same two range pixels say L_{r_a} and L_{r_b} differing if the distance of the laser radar imager from the imaged scene is varied. However the difference in the log range i.e, $\log(L_{r_a}) - \log(L_{r_b}) = \log \frac{L_{r_a}}{L_{r_b}}$ stays the same. We choose to use the logarithm of the range data instead of the range data directly. The log of the median filtered range image is shown in Fig. 2(b). Also a mesh plot of the

(a)



(b)

Figure 4. (a) Segmented LADAR intensity image. (b) Edge detector on gradient log-range image

log of the range image reveals that it is linearly varying (planar). We therefore use a gradient filter on the logarithm of the range image (shown in Fig. 3) to obtain an approximately piecewise constant image.

4. SEGMENTATION AND ROBUST CFAR DETECTION ON LADAR INTENSITY

We applied the segmentation algorithm discussed in Section 2 (unknown mean and unknown variance) on the intensity image. Figure 4(a) shows the mean plot of the segmented intensity image. The algorithm does a good job of segmenting out homogenous regions while staying robust to outliers. Figure 4(b) shows a simple edge detector used on the gradient log-range image. This is appropriate since the gradient filter basically extracts boundaries where sharp slope changes occur. The segmented intensity plot is very useful on its own as a front end step fed into other processings for target detection and recognition. One such application is the robust CFAR detection scheme described below.

4.1. CFAR Detection

The CFAR detector is often used as a prescreener to identify potential targets in LADAR images on the basis of pixel brightness.[20] The conventional CFAR detector is the cell averaging (CA) CFAR detector that examines a reference window W of samples surrounding the test pixel and returns bright pixels as suspect targets. Every pixel is sequentially examined as a test pixel. If we denote X_t as the amplitude of the test pixel, $\overline{X_s}$ and $\hat{\sigma}_s$ as the estimated mean and the estimated standard deviation of the surrounding samples in W, we may use the following rule:

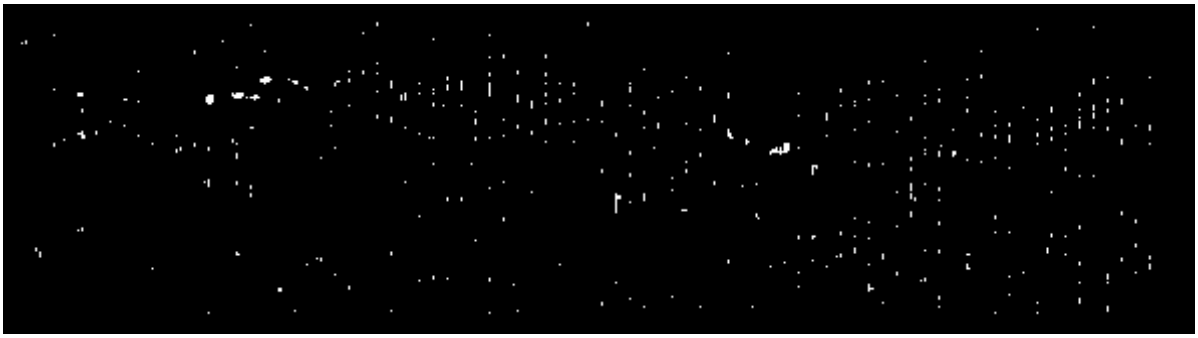


Figure 5. Using the typical 9×9 block CFAR detector on the intensity image

$$\frac{X_t - \overline{X_s}}{\hat{\sigma}_s} > \tau \Longrightarrow \text{declare target}$$

$$\frac{X_t - \overline{X_s}}{\hat{\sigma}_s} < \tau \Longrightarrow \text{declare clutter}$$

The threshold τ is a constant that determines the false alarm rate. This scheme is not adaptive to the varying-size of regions since W is fixed and performs poorly if the background window is not homogenous as in multiple targets and clutter edge situations.[21] Also there are boundary problems associated with this scheme since pixels at the edge of the image do not have sufficient surrounding pixels in W to give reasonable CFAR statistics.

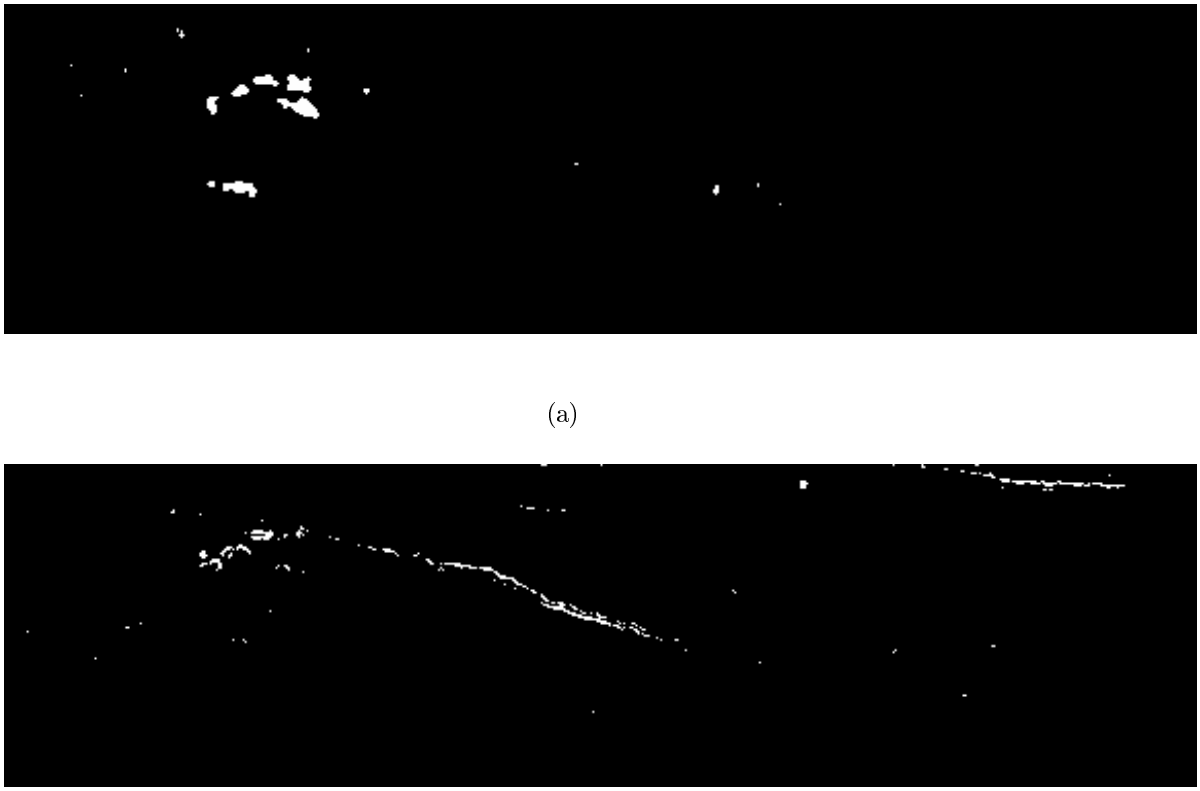
To resolve these issues, our robust CFAR detector uses estimates of the mean $f(i_1, i_2)$ and variance $\sigma^2(i_1, i_2)$ obtained from our segmentation scheme on the intensity image. That is, for each test pixel, $\overline{X_s}$ and $\hat{\sigma}_s$ are the underlying mean and variance values obtained for the subblock into which X_t falls in the segmented image. These statistics are more robust because they adapt to the size of homogenous regions and therefore give more accurate estimates of the parameters. They also do not suffer from clutter boundary edge problems within the image. This makes our scheme less sensitive to outliers that distract the conventional cell-averaging CFAR scheme. We avoid the afore-mentioned image boundary problems since every pixel is part of a region. The edge image obtained from the gradient log-range image is thresholded to detect outstanding edges and can be compared with the results from the robust CFAR detection on the intensity image.

4.2. Results

Figure 5 shows the output from the conventional (CA) CFAR algorithm using a sliding window of 9×9 pixels around the test pixel. We have used a threshold that reflects the 85th percentile of bright pixels in the CFAR image. Several false alarms can be seen and the target regions are not distinct. To improve detection using this algorithm would require a larger window size but with a greater penalty at the image boundary. Figure 6 (a) shows the result of using our robust CFAR algorithm on the intensity image. We are able to get rid of several of the false positives with more distinct 'activity' regions. Figure 6 (b) is the thresholded edge image of the gradient log-range image indicating boundaries where sharp gradient change occurs.

5. CONCLUSION

In this paper, we have investigated a novel approach to unsupervised segmentation of LADAR images using a coding theoretic approach based on the MDL principle. Our model is a random Gaussian field with piecewise constant mean and variance functions. We develop the adaptive recursive partitioning algorithm which is unsupervised, recursive and not restricted to dyadic partitions. The intensity and range data are treated differently because the latter provides location information which is fundamentally different from the energy (reflectance) images obtained from the former.



(a)

(b)

Figure 6. (a) CFAR detector on the intensity image. (b) Thresholded edge-range image.

Preprocessing transforms the images into a suitable form for our assumed image model. Our segmentation scheme is applied to the intensity image. We develop a robust CFAR detector based on our segmentation scheme, to indicate possible target regions. The gradient of the log-range image is thresholded to detect points of change in planar slope. Results from our segmentation scheme and the robust CFAR detector indicate that it is promising for target detection in LADAR imagery because it accurately locates areas of 'activity' in the intensity and range images.

ACKNOWLEDGMENTS

The authors would like to thank Alan Van Nevel and Carey Schwartz of the Naval Air Warfare Center, China Lake, California for providing several real LADAR data sets used for this work.

REFERENCES

1. M. Snorrason, H. Ruda, and A. Caglayan, "Automatic target recognition in laser radar imagery," *Proc. ICASSP* , (Detroit), 1995.
2. A. Koksal, J. Shapiro, and W. Wells, "Model-based object recognition using laser radar range imagery," *Proc SPIE* **3718**, pp. 256–266, 1998.
3. Y. Yu and S. T. Acton, "Polarimetric SAR image segmentation using texture partitioning and statistical analysis," *Proc. ICIP* , (Vancouver), Sept. 2000.
4. A. Jain, T. Newman, and M. Goulish, "Range-intensity histogram for segmenting LADAR images," *Pattern Recognition Letters* **13**, pp. 41–56, Jan. 1992.

5. J. R. Goldschneider and A. Q. Li, "Variational segmentation by piecewise facet models with application to range imagery," *Proc. ICIP* , (Vancouver), 2000.
6. C. Chu, N. Nandhakumar, and J. K. Aggarwal, "Image segmentation using laser radar data," *Pattern Recognition* **23**, pp. 569–581, 1990.
7. N. Saito, "Simultaneous noise suppression and signal compression using a library of orthonormal bases and the minimum description length criterion," *Wavelets in Geophysics* , Foufoula-Georgiou and Kumar (eds.), Academic Press, 1994.
8. J. Rissanen, *Stochastic Complexity in Statistical inquiry*, World Scientific Publishers, Singapore, 1989.
9. M. H. Hansen and B. Yu, "Model selection and the principle of minimum description length," *Journal of the American Statistical Association,* **to appear**, 2001.
10. H. Krim and I. Schick, "Minmax description length for signal denoising and optimal representation," *IEEE Trans. on Information Theory* **45**(3), April. 1999.
11. R. D. Nowak and M. A. T. Figueiredo, "Unsupervised progressive parsing of poisson fields using minimum description length criteria," *Proc. ICIP* , pp. 26–29, (Kobe, Japan), 1999.
12. J. Rissanen and B. Yu, "MDL learning," *Learning and Geometry: Computational Approaches, Progress in Computer Science and Applied Logic* **14**, pp. 3–19, 1995.
13. Y. Leclerc, "Constructing simple stable descriptions for image partitioning," *International Journal of Computer Vision* **3**, pp. 73–102, 1989.
14. J. Rissanen, "Fisher information and stochastic complexity," *IEEE Transaction on Information Theory* **42**, pp. 40–47, 1996.
15. T. Cover and J. A.Thomas, *Elements of Information Theory*, J. Wiley and Sons, New York, 1992.
16. V. Venkatachalam, R. Nowak, R. Baraniuk, and m. Figueiredo, "Unsupervised SAR image segmentation using recursive partitioning," *Proc. SPIE International Symposium on Aerospace/Defense Sensing, Simulation, and Controls, Algorithms for SAR Imagery,* **VII**, (Orlando, Florida), April. 2000.
17. U. Ndili, R. D. Nowak, and M. A. T. Figueiredo, "Coding theoretic approach to image segmentation," *submitted to ICIP* , 2001.
18. J. Kostakis, M. Cooper, T. Green, M. Miller, J. Sullivan, J. Shapiro, and D. Snyder, "Multispectral sensor fusion for ground-based target orientation estimation: FLIR, LADAR, HRR," *Proc of SPIE* **3717**, (Orlando, Florida), April. 1999.
19. J. Huang, A. B. Lee, and D. Mumfurd, "Statistics of range images," *Proc. CVPR* **1**, (South Carolina), June. 2000.
20. J. Tian and J. R. O. Wells, "A CFAR-enhanced image codec for SONAR ATR," *Proc. of SPIE tech. conf on applications in signal and image Processing* **VII**, (Denver), July. 1999.
21. K.-T. Jung and H.-M. Kim, "Performance analysis of generalized order statistic cell averaging CFAR detector with noncoherent integration," *IEEE Trans. Fundamentals* **E81-A**, pp. 1201–1209, June. 1998.