A Focused Target Segmentation Paradigm *

Dinesh Nair and J. K. Aggarwal

Computer and Vision Research Center
The University of Texas at Austin,
Austin, Texas 78712-1084, USA.
email: aggarwaljk@mail.utexas.edu

Abstract. In this paper we present new algorithms for target detection/segmentation in second generation Forward Looking Infra-Red (FLIR) images. An initial detection algorithm that models the background using Weibull functions, is used to identify candidate target locations in the image. A two-stage focused analysis of each candidate target location is then performed to get an accurate representation of the target boundary. A region-growing procedure is used to get an initial estimate of the target region, which is then combined with salient edge information in the image to arrive at a more accurate representation of the target boundary. The region and edge integration is done using a novel method that uses a Bayes' minimum risk classification approach. Finally, to reduce the false alarm rate, a higher level interpretation module is used to classify the detected areas as man-made or natural objects using geometric and FLIR-intensity based features extracted from the target.

1 Introduction

A central problem in computer (machine) vision, and automatic target recognition (ATR) in particular, is obtaining robust descriptions of the *objects of interest* (OOI) in an image. Vital to this problem is the need for good segmentation techniques that identify the regions in an image that are occupied by the objects. In ATR applications, it is important to get an accurate and precise representation of the boundary of the tactical targets [1]. Since targets are usually characterized by their shape and the gray scale representation of the segmented target, segmentation results directly affect the performance of the system.

In this paper, we present a segmentation method which uses low level target segmentation techniques along with higher level interpretation techniques for robust target detection and segmentation. The strategy adopted here is the following: (1) An initial detection algorithm is used to identify regions in the image that are candidate location of objects of interest (targets) by modeling the background using Weibull functions. At this stage all possible object locations in the image are identified, at the expense of a high false alarm rate. Concurrently, a *salient edge image*, which contains all the salient edge information in

* This work was supported by the Army Research Office Contracts DAAH-94-G-0417 and DAAH 049510494.

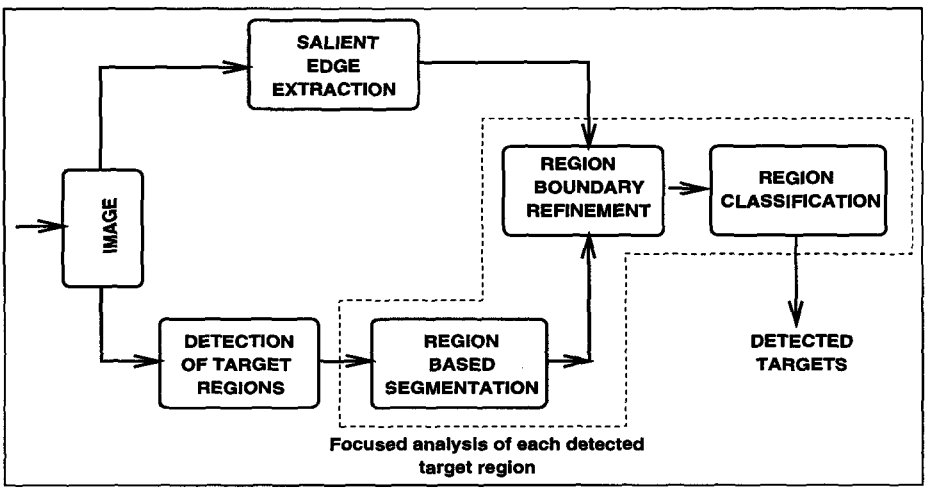

Fig. 1. The general flow of the target segmentation paradigm.

the image, is obtained from the original image. Saliency of edge segments in an image is defined by its length (L), average contrast (C) and smoothness (Δk), as described in [3]. Salient edge segments are those which are more significant than the others in the image. By considering only salient edges, the effect of clutter in an image is reduced.

(2) The initial detection stage is followed by a *focused* analysis of the candidate target areas. The objective at this stage is to get a good representation of the image attributes of the target region under consideration, such as boundary information, size, compactness, etc. The focused analysis consists of a two-stage process. A region-growing procedure is used to get an initial estimate of the target region, which is then combined with the edge information in the corresponding region of the *salient edge image* to get an accurate estimate of the target boundary. (3) In the final stage a higher level interpretation technique is used to identify the OOI from the candidates in the image. A schematic of the overall approach is shown in figure 1.

The rest of the paper is organized as follows: Section 2 describes the algorithm developed to detect the initial candidate locations of the targets in the image. In section 3, the focused segmentation of each candidate target location is presented. Section 4 describes the higher interpretation module and the types of features extracted from the image to discriminate between man-made objects and natural background. Finally, in section 5, a summary of this study is presented.

2 Initial Detection Of Target Locations

In this section, we present a method for better representation of the intensity distribution of the background in second generation FLIR images and a scheme

that uses this information for the initial detection of probable target locations.

To model the background distribution, we have found that the density function of the Weibull distribution (predominantly Rayleigh) approximates the background clusters well (better than a Gaussian distribution in terms of the detection rate). Since the background tends to form more than one cluster, a piecewise approximation of the histogram is done using Weibull functions to arrive at the background model. The approximating Weibull density functions are of the form:

$$\begin{aligned} f(x) &= \frac{A\gamma}{\theta} x^{\gamma-1} \exp\left(-\frac{x^{\gamma}}{\theta}\right) \qquad x \geq 0 \\ &= 0 \qquad\qquad\qquad\qquad\qquad otherwise \end{aligned} \tag{1}$$

where, θ determines the spread of the function, γ determines the shape of the function and A modulates the peak value of the function. For $\gamma = 1$, the function reduces to an exponential form, while for $\gamma = 2$, we obtain the *Rayleigh* form.

Target regions in the image are detected by determining all locations in the image where the gray scale values have a "low" probability of belonging to the background. The probability of each gray scale value of belonging to the background is determined by the background cluster that it is closest to, and is computed in a maximum likelihood sense as follows. Let x represent a gray scale value, where $0 \leq x \leq 255$, and $L(x \in c_k)$ be the *likelihood* of x belonging to cluster c_k, where c_k, $k = 1 \ldots n$, represents the n background clusters. $L(x \in c_k)$ is computed as:

$$L(x \in c_k) = \frac{\frac{A_k \gamma_k}{\theta_k} x^{\gamma_k - 1} \exp\left(-\frac{x^{\gamma_k}}{\theta_k}\right)}{\frac{A_k \gamma_k}{\theta_k} x_{peak,k}^{\gamma_k - 1} \exp\left(-\frac{x_{peak,k}^{\gamma_k}}{\theta_k}\right)}, \tag{2}$$

where, $x_{peak,k}$ represents the gray level value at the peak of cluster k, and the probability that x belongs to the background is obtained as:

$$P(B/x) = max\{L(x \in c_k), k = 1 \ldots n\}, \tag{3}$$

where, B represents the background class, and the probability of x belonging to the target is $(1 - P(B/x))$. The false alarm rate at this stage is directly proportional to the threshold probability chosen as "low". Typically, at this stage all possible target locations are detected, at the cost of a high false alarm rate. Figures 2 and 3 show examples of typical FLIR images, their histograms and the Weibull approximations.

3 Target Segmentation Using Focused Analysis

Once all possible target locations in the image have been identified, a focused analysis approach is used to improve the segmentation around each candidate target location. A region-growing procedure is used to initially segment the target, which is then followed by a *refinement* stage, where the accuracy of the boundary of the segmented target region is improved using the salient edge information (obtained from the *salient edge image*) in and around the target region.

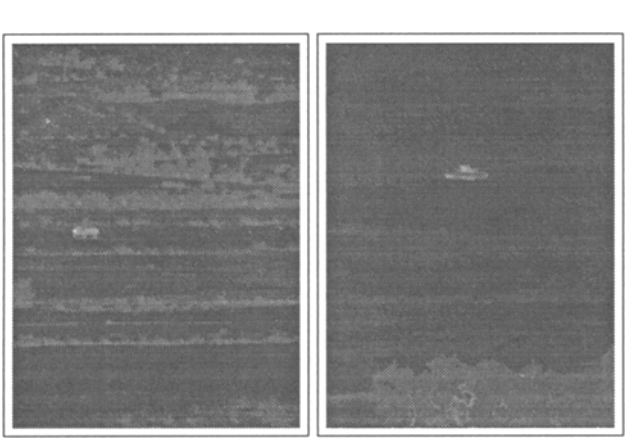

Fig. 2. FLIR images of a truck and a tank.

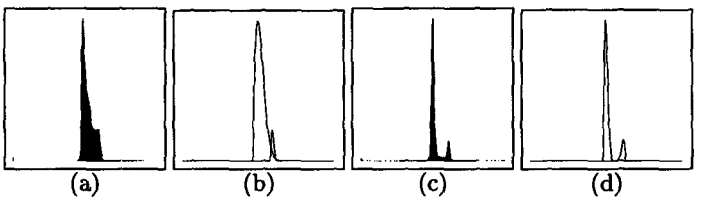

Fig. 3. (a), (c) Histograms of the images in Figure 2. (b), (d) Weibull approximations of the histograms.

3.1 A Local Shape Driven Region Growing Paradigm

A region-growing paradigm is used to segment each candidate target detected. To grow the region around the detected areas, surrounding pixels with a low probability of belonging to the background are added to the region. The process is iteratively continued by adding pixels to the region with higher probabilities of belonging to the background, while the size of the segmented region does not exceed the *a priori* known object size, the compactness of the region does not exceed a predefined maximum, and the boundary of the region does not get too irregular.

The last constraint uses the fact that the targets to be segmented are man-made and hence show some kind of regularity. On the other hand, boundaries of objects that belong to the background (like vegetation, clouds, etc.) are usually characterized by irregular boundaries which exhibit frequent changes in curvature. To determine the percentage of the boundary that is irregular, we first extract the boundary (contour) of the region. The corner points of the boundary are detected. For man-made objects, these corner points can be regarded as the end points of the linear segments that make up its boundary. The contour (boundary) between these corner points should be relatively smooth. The corner points are detected using a cubic B-spline fitting method [2]. Next, the corner

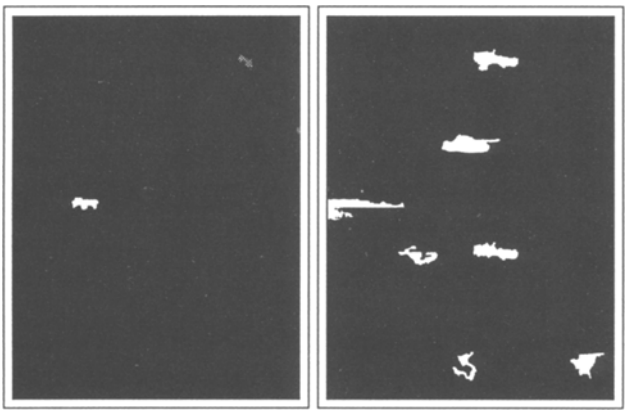

Fig. 4. Region based segmentation of the images in Figure 2.

points are used to split the boundary into segments, where the boundary between two corner points is considered as one segment. Segments that are smaller than some fraction (5%) of the *a priori* known approximate target boundary size are considered as irregular segments. The rest of the segments are then tested for smoothness. For each segment i, a smoothness measure is computed as:

$$S_s^i = \frac{\Delta\kappa_{max} - \Delta\kappa^i}{\Delta\kappa_{max} - \Delta\kappa_{min}}, \tag{4}$$

where, $\Delta\kappa$ is the average change of curvature of the segment, $\Delta\kappa_{min}$ is the smallest average change of curvature for all the segments of the target boundary and $\Delta\kappa_{max}$ is the largest average change of curvature for all the segments. The measure S_s^i satisfies $0 \leq S_s^i \leq 1$.

The smoothness measures are used to order the segments, and segments with a smoothness measure greater than a certain threshold value are considered to be smooth. This threshold value is computed as follows:

$$\Theta = 1 - e^{\frac{-A\mu_{\Delta\kappa}}{\rho_{\Delta\kappa}}}, \tag{5}$$

where, $\mu_{\Delta\kappa}$ and $\rho_{\Delta\kappa}$ are the mean and standard deviation of the average change in curvature of all the segments, and A is a constant. By making Θ a function of the average smoothness of the segments and the deviation of the smoothness of the segments, the threshold value is indicative of the overall smoothness of the boundary. Figure 4 shows the final region-based segmentation for the images in figure 2. A closer look at the segmentation for one of the examples is shown in figure 5.

3.2 Integrating Salient Image Contours with Segmented Regions for Refining Target Boundaries

A new paradigm for the refinement of region-based segmentation results using edge information is presented next. Since segmentation using region-based tech-

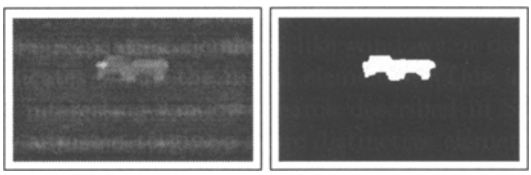

Fig. 5. A closer look at the segmentation of the truck in Figure 2.

niques are inefficient in locating exact target boundaries and tend to miss small parts of the target, the boundaries of the segmented targets are refined using the salient edge segments in the image. An *edge contour* is obtained from an edge as a linked list of edge pixels. An *edge segment* is defined as that part of an *edge contour* between its corner points.

Given the target boundary obtained from the region growing process and the salient edge segments in and around the target region, the boundary refinement problem is stated as:

For every point s on the region boundary, find its new location as a selection from a set of candidate edge element locations $\bar{z} = \{z_j,\ j = 0 \ldots n\}$, *where* $z_0 = s$.

This problem is formulated as a classification problem, where point s will take one of the labels given by $\bar{z}$. Using Bayes decision rule, choose z_j as the new location if

$$p(s|z_j)P(z_j) \geq p(s|z_k)P(z_k) \qquad \forall\, k \neq j, \tag{6}$$

where $p(s|z_j)$ represents the conditional density function of (s, z_j) and $P(z_j)$ is the *prior* probability of z_j. The *prior* probability of each candidate location z_j is estimated as the proximity of the salient edge segment to which z_j belongs to the boundary of the target region. Proximity of an edge segment is defined as the percentage of its segment that is *close* to the region boundary. The *closeness* of a segment pixel to a region boundary is determined by a *Refinement Search Circle* (RSC) of radius λ. If an edge segment pixel lies within a RSC placed on any region boundary point, then the edge segment pixel is considered as *close* to the region boundary. Each salient edge segment $(SE_i, 0 \leq i \leq q$, where q is the number of salient edge segments and $SE_0 = s)$ is assigned a proximity weight determined by the number of its pixels that are close to the target region boundary, and is given by:

$$Prox(SE_i) = \frac{\#\ edge\ elements\ of\ segment\ SE_i\ that\ is\ close\ to\ region\ boundary}{\#\ edge\ elements\ in\ SE_i} \tag{7}$$

Hence, under suitable assumptions [5], the *priors* are computed as:

$$P(z_j) \approx Prox(SE_i^j). \tag{8}$$

All points lying on the same salient edge segment will have the same *prior* probability. The density function $p(s|z_j)$ is assumed to be Gaussian distributed

around z_j with variance σ_j^2 as:

$$p(s|z_j) = \frac{1}{\sqrt{2\pi}\sigma_j} \exp \frac{-(s - z_j)^2}{2\sigma_j^2}. \tag{9}$$

The standard deviation σ_j represents the uncertainty in the location of the edge element z_j and is approximated as:

$$\sigma_j = \frac{Maximum\ strength\ of\ a\ salient\ edge\ element\ in\ the\ image}{Strength\ of\ salient\ edge\ element\ at\ z_j}, \tag{10}$$

where the strength of an edge element is proportional to gradient of the image at the edge element location. Therefore, the edge element with the highest strength has unit variance. Assume that the standard deviation at z_0, $\sigma_0 = \sigma_{max}$, where σ_{max} represents the largest standard deviation among all the salient edge elements in the image. Using the criteria in equation (6) and the density function given in equation (9), z_j is chosen as the candidate location if the following is satisfied:

$$(s - z_j)^2 \leq 2\sigma_j^2 [\log \frac{P(z_j)\sigma_k}{P(z_k)\sigma_j} + \frac{(s - z_k)^2}{2\sigma_k^2}] \qquad \forall\, k \neq j. \tag{11}$$

The computational burden associated with the above criterion is reduced by assigning a *prior* probability $P(z_0) = Prox(SE_0) = \Gamma$, where $0 < \Gamma \leq 1$. Substituting $k = 0$ and $\sigma_0 = \sigma_{max}$ in equation (11) and noting that $z_0 = s$, the winning candidate has to satisfy the condition that:

$$(s - z_j)^2 \leq 2\sigma_j^2 [\log \frac{P(z_j)\sigma_{max}}{\Gamma\sigma_j}]. \tag{12}$$

Thus, by assigning a user defined *prior* probability Γ, the search space is restricted to a circle of radius λ given by the largest value of right hand side in equation (12). The largest value of $|s - z_j|$ in equation (12) is obtained when $P(z_j)$ is at its maximum P_{max}, and $\sigma_j = (P_{max}\sigma_{max}/\Gamma)e^{-0.5}$, and is given by:

$$\lambda = |s - z_j|_{max} = \frac{P_{max}\sigma_{max}}{\Gamma} e^{-0.5}. \tag{13}$$

Since P_{max} is the maximum value $P(z_j)$ can take, which in turn is the maximum proximity weight that is possible (equation (8)), $P_{max} = 1$. Therefore the search radius becomes $\lambda = e^{-0.5}\sigma_{max}/\Gamma$. The radius has been deliberately denoted by λ to emphasize the equivalence of this search radius to the RSC used for finding the *closeness* of the edge segments to the region boundary. From equation (11), it is seen that when choosing from edge locations that belong to the same salient segment, the problem reduces to finding the edge location that is closest to the boundary point.

At the end of the boundary refinement stage, an edge image is obtained which represents the location of the refined target region boundary. To complete breaks in the boundary, with an emphasis on incorporating small parts of the

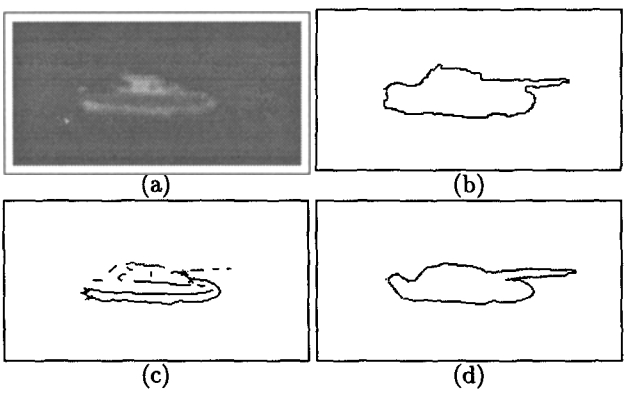

Fig. 6. (a) Target (b) Boundary of target obtained from region-based segmentation (c) Salient edges around the target. (Note: salient edges that belong to a long contour are separated by crosses on the contour.) (d) The refined boundary of the target.

target missed by the region segmentation, the following algorithm is used: (1) Remove isolated edge pixels. These are edge pixels with no neighbors in the refined boundary. (2) For each edge pixel at the site of a break in the boundary, determine if it is part of a salient edge. If it is, incorporate the complete edge segment into the boundary. Repeat this process iteratively until no more salient edge information can be incorporated. (3) Enforce linear connectivity at all the remaining breaks. Figures 6 and 7 show examples of the refinement process.

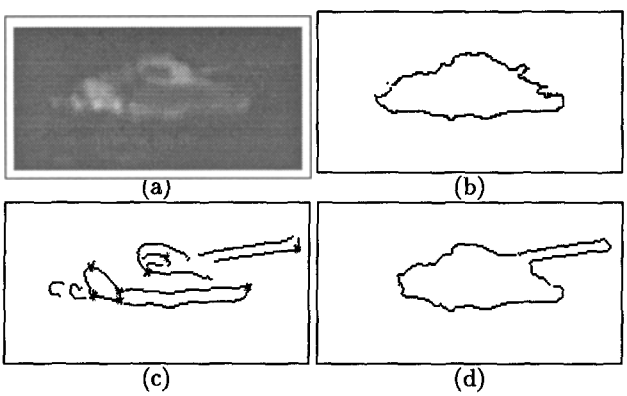

Fig. 7. (a) Target (b) Boundary of target obtained from region-based segmentation (c) Salient edges around the target. (Note: salient edges that belong to a long contour are separated by crosses on the contour.) (d) The refined boundary of the target.

4 Man-Made / Natural Background Categorization

To reduce the false alarm rate, a higher interpretation module is used to identify target regions from background clutter. Image statistics are used to derive a set of feature groups that is used to discriminate between man-made objects and natural backgrounds. Each feature group is modeled by a modular neural network model, the Categorizing and Learning Module (CALM) [4].

4.1 Image Features

Image features extracted, which can be grouped into two categories; (1) Geometric Features, and (2) FLIR-intensity Features, are described below.

Geometric Features Two geometric features are used. These are:

- **Saliency:** This feature group is based on the number of *salient segments* in the target region and consists of two elements:

$$S1 = \frac{\sum \textit{salient edge segments in target region}}{\sum \textit{edge segments in target region}}$$
$$S2 = \frac{\sum \textit{salient edge segments in target region}}{\sum \textit{salient segments in entire image}} \tag{14}$$

- **Boundary Regularity:** A feature group of three elements is used to characterize the smoothness of the target boundary.

$$R1 = \mu_{\Delta K}, \qquad R2 = \sigma_{\Delta K}, \qquad R3 = \frac{\sum_i L_i}{L}, \tag{15}$$

where, $\mu_{\Delta K}$ and σ_{DeltaK} are the same as in equation (5), L_i represents the ith segment on the target boundary that is very small (chosen as 5% of the total target boundary length) and L represents the number of segments on the boundary.

FLIR-Intensity Features Two features are used. These are:

- **Gray level statistics:** This feature group consists of two elements:

$$Mn = \frac{\mu_{target} - \mu_{image}}{\mu_{image}}, \qquad Sd = \frac{\sigma_{target} - \sigma_{image}}{\sigma_{image}}, \tag{16}$$

where, μ represents the mean and σ the variance of a region.

- **Gray level clusters:**

$$CLUS = \textit{No. of distinct gray level clusters in target region}$$
$$SPREAD = \textit{Distance between furthest cluster centroids} \tag{17}$$

4.2 Man-made / Natural Background Classification

Features extracted from the target regions are then classified as man-made or natural using modified CALM networks. Modifications to the CALM were done to incorporate a confidence value with its decisions and to speed up its convergence. Each feature group was modeled by a CALM module and the decision from each of the feature groups was combined to get the final classification result. Details of this can be found in [5].

5 Summary

In this paper we have presented new target detection/segmentation algorithms for use in automatic target recognition systems. By modeling the background by Weibull distributions, a good initial detection of target locations is obtained. A focused analysis of each target location is then performed by a region-growing procedure which uses the underlying background probabilities for checking region homogeneity and local shape characteristics for determining the convergence/stopping criteria. To get better representations of the target boundary, salient edge information in the image is used to refine the boundary obtained by the region-growing method. A novel method to do this using a Bayes' minimum risk classification approach is presented. Finally, the false alarm rate is reduced by using a higher level interpretation module built using modified CALM neural networks, that uses geometric and intensity based features from each detected/segmented region to classify it as man-made or natural. The target detection/segmentation algorithms have been successfully used on second generation FLIR images. A 100% detection rate with a false alarm rate of 5% was obtained when the segmentation method was tested on 200 images from the HULI9306_SIG subset of the COMANCHE data set provided to us by the Night Vision Laboratory.

References

1. B. Bhanu, "Automatic target recognition: State of the art survey," *IEEE Transactions on Aerospace and Electronic Systems*, vol. AES-22, no. 4, pp. 364–379, 1986.
2. G. Medioni and Y. Yasumoto, "Corner detection and curve representation using cubic B-splines," *Computer Vision, Graphics and Image Processing*, vol. 39, pp. 267–278, 1987.
3. K. Rao and J. Liou, "Salient contour extraction for target recognition," *SPIE: Acquisition, Tracking and Pointing*, vol. 1482, pp. 293–306, 1991.
4. J. M. J. Murre, *Learning and Categorization in Modular Neural Networks.* Lawrence Erlbaum Associates, Publishers, 1992.
5. D. Nair and J. K. Aggarwal, "Robust automatic detection/segmentation in 2nd generation FLIR images," *Technical Report, Computer & Vision Research Center, University of Texas at Austin*, TR-96-03-103, 1996.