An Automated Target Recognition Technique For Image Segmentation and Scene Analysis

Chris W. Baumgart and Christopher A. Ciarcia
EG&G Energy Measurements, Inc.
Los Alamos Operations
Los Alamos, NM 87544
(505)-667 2083
Internet: baumgart_cw@laoops.lanl.gov

ABSTRACT

Automated target recognition software has been designed to perform image segmentation and scene analysis. Specifically, this software was developed as a package for the Army's Minefield and Reconnaissance and Detector (MIRADOR) program. MIRADOR is an on/off road, remote control, multi-sensor system designed to detect buried and surface-emplaced metallic and non-metallic anti-tank mines.

The basic requirements for this ATR software were: (1) an ability to separate target objects from the background in low S/N conditions; (2) an ability to handle a relatively high dynamic range in imaging light levels; (3) the ability to compensate for or remove light source effects such as shadows; and (4) the ability to identify target objects as mines.

The image segmentation and target evaluation was performed utilizing an integrated and parallel processing approach. Three basic techniques (texture analysis, edge enhancement, and contrast enhancement) were used collectively to extract all potential mine target shapes from the basic image. Target evaluation was then performed using a combination of size, geometrical, and fractal characteristics which resulted in a calculated probability for each target shape.

Overall results with this algorithm were quite good, though there is a trade-off between detection confidence and the number of false alarms. This technology also has applications in the areas of hazardous waste site remediation, archaeology, and law enforcement.

1.0 Introduction

Automatic target recognition (ATR) software has been designed to perform image segmentation and scene analysis. Specifically, this software was developed as a package for

MASTER

EP

the Army's Minefield Reconnaissance and Detector (MIRADOR) program developed for the U.S. Army Belvoir Research, Development, and Engineering Center (BRDEC). The final MIRADOR system was designed to be an on/off road, remote control, multi-sensor system designed to detect buried or surface-emplaced metallic and non-metallic mines. Within this context, our purpose was to address the detection of surface emplaced mines by developing image processing techniques and procedures to detect and identify mines within a two-dimensional image derived from a forward-looking video camera. The software utilizes a fusion of several image processing techniques and analyses to extract objects of interest from video camera frames and identify those objects which have a high probability of being surface-emplaced land mines.

The basic requirements for this ATR software were: (1) an ability to separate target objects from the background in low S/N conditions; (2) an ability to handle a relatively high dynamic range in imaging light levels; and (3), the ability to compensate for or remove light source effects such as shadows. Optimization of the algorithm's processing speed was not a requirement under this contract. However, all the image processing techniques in the ATR algorithm are composed of individual pixel operations, making the algorithm readily adaptable to future, parallel processing applications.

2.0 System Description and Data Acquisition

In order to create test images for code development that were as realistic as possible, five real mines (an M80 anti-tank, an M20, an East German anti-personnel, a Swedish influence, and a Chinese box) were used to create 17 test images, one of which is shown in Figure 1. Eight of these images were acquired on a desert scrub-like background, five were taken on a gravel background, and four were acquired on a low cut grass background. Objects such as stones, sticks, and other naturally occurring clutter were included to simulate a true operational environment.

Each of the seventeen test images was acquired with a Pulnix TM-840 CCD camera in conjunction with a Fujinon zoom lens. A focal length of 12.5 mm was used in acquiring all images. Camera height above the ground was approximately 62 inches producing an approximate field-of-view of 40 inches across. The acquired images were then digitized into a 512 x 512 pixel, 8-bit gray scale format. The resolution of the lens/camera/image processing system was measured to be approximately 1 line-pair per centimeter with a detection threshold of approximately 3 millimeters.

The test images were also designed to emphasize several major processing difficulties inherent in visual (and infrared) imaging. These are:

1. Inability to separate the target mine from the background when there is a low signal-to-noise ratio. This inability is the result of low contrast variation between the target object and the background.
2. Variation in imaging light level as a result of the time of data acquisition

within a day-night cycle or during overcast imaging conditions. This variation results in a very dynamic and highly variable range of target gray level configurations. The mine can appear to be bright or dark on a bright or dark background. As a result, the segmentation technique must be capable of handling such a dynamic range.

3. Light source effects such as those produced by a spotlight during night imaging or by sun illumination during the day. These effects can produce shadows of fairly uniform gray level distributions within the image which must be differentiated from objects of interest within the same field-of-view.

3.0 Target Segmentation and Evaluation

3.1 Segmentation

The goal of target segmentation is to isolate the images of the land mines from the background. To accomplish this, procedures were devised to detect specific details of the mine image that would allow the mines to be isolated from the background. The specific characteristics used in developing these procedures included the geometric shape of the mine (man-made objects are of a fairly uniform "euclidean" geometry; i.e., square or round); the texture of the surface of the mine (the manufactured surfaces of these mines are typically smooth compared to the background); and the brightness of the mine relative to the background (the mines are typically lighter than or darker than the background).

A variety of independent image processing procedures were initially developed that successfully isolated the mine shapes within the two-dimensional image field. Their success, however, hinged strongly on their diversity of technique and what can be called "their region of overlap". Each process had to complement the others in order to incorporate as many image content variations as possible.

This integrated diversity of technique was accomplished in a three-fold fashion defined as the TEC analysis procedure. The method used was to analyze an input image three different ways in a parallel fashion while extracting all possible mine target shapes. The technique "T", for texture analysis, was used to scan an image for regions of relative smoothness. The technique "E", for edge enhancement, was used to emphasize all edges within the image, assuming that the Euclidean mine shapes would be easily distinguishable from background objects. Finally, the brightness or relative darkness of an object within an image was parameterized using the contrast stretching technique "C". Here, the contrast of an image was stretched in both a linear and non-linear fashion to enhance the signal-to-noise ratio between a potential mine shape and the background. Figure 2 describes the TEC analysis procedure.

Additionally, there was no way to initially determine if an imaged mine was brighter or darker than the ambient background. It was therefore necessary to process each image in both a positive (original image) and negative (inverted) format. In this way, both dark and bright objects could be treated in the same fashion. And, since both images (positive and

negative) were processed according to the original three-fold TEC analysis procedure, it became necessary to use a parallel process of six individual analysis lines. Once a potential target shape was isolated in the image, its basic geometrical features were measured. Since the imaging system and its field-of-view were known quantities, a mine's absolute area could be estimated. An initial discrimination was then made based on this measured area. Any segmented object with an area greater than or less than 20% of any of the areas listed in a library of known mine characteristics was eliminated from further processing.

3.2 Target Shape Evaluation

Following the identification of possible mines shapes in the image and the preliminary area discrimination, the shapes of the remaining objects were evaluated and the probability (confidence level value) of the object being an actual mine was evaluated. This probability was determined by using the following shape characteristics: (1) The relative error in gross target geometrical properties; (2) the contortion of the target edge as characterized by the line fractal dimension; and (3) the target shape asymmetry.

The relative error in gross target geometrical properties was derived by comparing the nth target with the geometrical characteristics of the one mine in the mine characteristics library closest to it in area. These geometrical characteristics were the percentage difference in target area (dA_n), perimeter (dP_n), and degree of edge contortion (df_n). The total relative error between the target and the library mine was defined as the sum of these characteristics; the smaller this error, the higher the probability that the target was a mine.

The target edge contortion, as characterized by the line fractal dimension, was based on an analysis of the edge of the potential mine target. For two-dimensional shapes, the line fractal dimension is a number ranging from 1.0 to 2.0 and is a measure of the degree of edge contortion. Since it was assumed that man-made objects are more likely to have straight, smooth edges, this parameter served as a useful feature in differentiating between mines (smooth edges) and background clutter (rough edges). However, this feature was not found to be sufficiently unique to serve as the only discriminator for mines. There can exist geometries within an operational environment with smooth, regular edge functions (soda cans, hubcaps, etc.) which are not mines.

Finally, since the images of the test mines were highly symmetric about an axis through their centers, a measure of the asymmetry of the target shape was used as yet a third discriminator. If the object was circular, rectangular, or it demonstrated reflection symmetry as did the test mines, its measure of asymmetry would be low.

These three shape characteristics were combined into a composite function that was used to evaluate the mine detection probability (P_n) of each segmented shape in the field-of-view. The form of this function is shown at the bottom of Figure 2. It has been normalized and scaled to produce a probability ranging from 0 to 10, with 10 being a 100 percent confidence detection.

4.0 Results

The results of the performance testing of this ATR software are summarized in Table 1 and Figure 3. In all, a total of 17 images containing a total of 22 mines were analyzed. These results show that all of the mines were detected at a confidence threshold of 64 percent; however, 30 false alarms were also included. At a confidence threshold of 80%, 26 targets were detected of which 19 were actual mines and 7 were false alarms.

Table 1
Confidence Level Thresholds and False Alarm Rates

Confidence Threshold (%)	Number Of Detections	Number Of False Alarms	Number Of Misses
80	26	7	3
75	33	13	2
74	37	16	1
73	40	19	1
72	40	19	1
71	42	21	1
70	43	22	1
65	50	29	1
64	52	30	0
60	57	35	0
50	70	48	0
40	79	57	0
30	85	63	0
20	94	72	0
10	98	76	0
0	98	76	0

5.0 Conclusions

Results from the application of the ATR software developed in this project to data acquired with a forward-looking video-imaging system demonstrated that the methods selected were very successful. Overall identification accuracy was very good, validating both the software design and approach. All mines in the seventeen test images analyzed were identified at a confidence level of 64 percent or greater, with 86 percent of those mines being identified with a confidence level of 80 percent or greater. The high false alarm rates evidenced at low detection thresholds can most likely be reduced with adjustments and refinements of the TEC analysis technique.

The primary goal of this project was not to demonstrate processing speed but to demonstrate detection capability. Recognizing that this processing system will eventually be best utilized on a fast moving sensor platform, all the processing techniques employed here are composed of individual pixel operations, making each of the TEC processing steps easily adaptable to highly parallel and/or pipeline processing. This approach allows the image segmentation/target recognition techniques employed here to be performed in near-real time.

The techniques employed in this current project do not guarantee absolute recognition of a mine within the imaging system's field-of-view. Any man-made object with a smooth contour and an area close to that of one of the library mines will likely be identified as a mine. Target discrimination and identification can be unique only when this software is fused with the output of other mine detection sensors such as ground penetrating radar and metal detectors. Just such a fused system is currently being developed by EG&G Special Projects and EG&G Energy Measurements under contract to BRDEC.

↙ Figure 1

Figure 86. Target locations for $P \geq 80\%$ for 2AS.DS.

Figure 87. Target locations for $P \geq 80\%$ for AB.DS.

Figure 88. Target locations for $P \geq 80\%$ for M8.DS.

TEC ANALYSIS

NUMBER OF DETECTIONS VS THRESHOLD

END

**DATE
FILMED**

4 / 7 / 94