

UNCLASSIFIED
UNLIMITED DISTRIBUTION

3

LEVEL 3

ORIG. RAPPORT 4180/81
DOSSIER: 3632D 007
JANVIER 1981

ORIG. REPORT 4180 81
FILE: 3632D-007
JANUARY 1981

AD A 095864

SEGMENTATION ALGORITHMS FOR
DETECTION OF TARGETS IN IR IMAGERY

L. Sevigny

RECEIVED
1981
JAN 21

DRS
FILE COPY

Centre de Recherches pour la Défense
Defence Research Establishment
Valcartier, Québec

BUREAU RECHERCHE ET DEVELOPPEMENT
MINISTRE DE LA DEFENSE NATIONALE
CANADA

RESEARCH AND DEVELOPMENT BRANCH
DEPARTMENT OF NATIONAL DEFENCE
CANADA

NON CLASSIFIED
DIFFUSION ILLIMITEE

CRDV R-4180/81
DOSSIER: 3632D-007

UNCLASSIFIED

17
DREV-R-4180/81
FILE: 3632D-007

6
SEGMENTATION ALGORITHMS FOR
DETECTION OF TARGETS IN IR IMAGERY

by
10 L. Sévigny (Algorithmes de Segmentation pour la
Detection de Cibles sur
Images IR)

11 11/11/81

12 1/81

Accession For	
NTIS GRA&I	<input checked="" type="checkbox"/>
DTIC TAB	<input type="checkbox"/>
Unannounced	<input type="checkbox"/>
Justification	
By	
Distribution	
Availability Codes	
Dist	Avail and/or Special
A	

CENTRE DE RECHERCHES POUR LA DEFENSE
DEFENCE RESEARCH ESTABLISHMENT
VALCARTIER
Tel. (418) 844-4271

DTIC
SELECTED
1981
D

Quebec, Canada

January/janvier 1981

NON CLASSIFIE

474 745

IR

UNCLASSIFIED

i

RESUME

Ce rapport présente une classe d'algorithmes de segmentation (segmenteurs) spécialisés dans la détection de cibles sur images IR et fondés uniquement sur le principe voulant que la signature thermique d'une cible dépasse en importance celle de tout objet de l'arrière-plan. Ces algorithmes sont le résultat d'efforts visant à améliorer un premier segmenteur, appelé générateur de silhouettes, imaginé en fonction d'images IR du type BOFORS. Le segmenteur en question découpe l'image en deux parties d'après un seuil d'intensité unique. Sa fiche d'extraction est généralement excellente lorsque l'arrière-plan est globalement plat. Lorsque cette condition n'est pas remplie, on peut parfois se tirer d'affaire en utilisant une fonction seuil au lieu d'un seuil fixe. Le générateur de silhouettes et ses diverses variantes tentent de venir à bout de l'arrière-plan simplement en morcelant l'image. Une solution plus prometteuse consiste à redresser l'arrière-plan de façon à réduire son emprise sur l'image. C'est précisément ce que la Technique de Redressement de l'Arrière-Plan (TRAP) fait. Etant donné que TRAP s'applique aux lignes aussi bien qu'aux colonnes d'une image, il en résulte 2 images distinctes: structure fine horizontale et structure fine verticale. Ces images recèlent maintes possibilités quant à la détection de cibles, lesquelles sont en grande partie explicitées dans le rapport. (NC)

ABSTRACT

This report presents a class of segmentation algorithms (segmenters) for detection of targets in IR imagery based on the single assumption that the targets possess a larger thermal signature than the background. This class of algorithms emerged as a result of efforts to improve an early segmenter devised to extract targets from IR BOFORS imagery. This segmenter proceeds according to a Single Intensity Threshold whence the name SIT Generator to designate it. The extraction record of the SIT Generator is generally excellent whenever the background, on a large-scale basis, is relatively uniform. When this condition is not met, one can use a thresholding intensity function in lieu of a fixed threshold. The SIT Generator and its variants try to cope with the background simply by partitioning the image. A more promising avenue consists in levelling the background so as to curb its ascendancy over the image. This is in essence what the Background Elimination Technique (BET) expounded herein does. Since BET can be applied either to the set of lines or columns of an image, it generates 2 images referred to as the Horizontal Fine Structure image and the Vertical Fine Structure image respectively. These images offer many possibilities for detection of targets and several of them are explicitly described in the report. (U)

TABLE OF CONTENTS

RESUME/ABSTRACT	i
1.0 INTRODUCTION	1
2.0 RECURSIVE LOW-PASS DIGITAL BUTTERWORTH FILTER	5
2.1 Z-Transfer Function	5
2.2 Difference Equations	6
2.3 Delay Time	7
2.4 Equivalent FIR Filter	9
3.0 ON THE ESTIMATION OF SOME STATISTICAL PARAMETERS	10
3.1 Mean and Variance	10
3.2 Skewness	11
3.3 Kurtosis	12
3.4 Autocorrelation Function	14
4.0 IR IMAGE MODELLING	16
5.0 SILHOUETTE GENERATOR	17
5.1 Single Intensity Threshold (SIT)	17
5.2 Thresholding Intensity Functions	19
6.0 GROSS STRUCTURE ANALYSIS (GSA)	23
6.1 Gross Structure Statistics	23
7.0 FINE STRUCTURE ANALYSIS (FSA)	28
7.1 Background Elimination Technique (BET)	28
7.2 Fine Structure Statistics	36
7.3 Target Area Delimitation	46
7.4 Experimental results	51
8.0 THRESHOLDING OF FINE STRUCTURE IMAGES	71
9.0 CONCLUSION	73
10.0 ACKNOWLEDGEMENTS	74
11.0 REFERENCES	75

FIGURES 1 to 18

TABLES I to II

1.0 INTRODUCTION

This report presents a class of segmentation algorithms (segmenters) for detection of targets in IR imagery, based on the single assumption, made explicit in Sect. 4, that the targets possess a larger thermal signature than the background. This class of algorithms emerged as a result of efforts to improve an early segmenter devised to detect targets in IR BOFORS imagery as part of an Automatic IR Target Acquisition System (AIRTAS). That particular segmentation algorithm was referred to in previous reports as a silhouette generator. It is a fairly simple algorithm and hence should be easy to implement in real time. Briefly, its defining procedure is: a) partition the image of interest into a certain number of subimages; b) determine the histogram of each subimage; c) estimate the upper gray level of the background and d) threshold the image accordingly. The silhouette generator, then, segments the image according to a single intensity threshold whence the more appropriate name of Single Intensity Threshold Silhouette Generator, or simply SIT Generator, to designate it.

The extraction record of the SIT Generator is generally excellent whenever the background, on a large-scale basis, is relatively uniform. This was indeed the case with the BOFORS imagery, but we should not expect that condition to prevail when images more akin to real-life situations, like those that make up the Alabama Data Base, are considered. In such circumstances the image must be divided into regions of uniform background and the SIT Generator applied to each of them as we would do for distinct images. This procedure amounts to defining several intensity thresholds in relation to the image being

considered. More generally, one can in this way build a thresholding intensity function, that is, a function that assigns a specific threshold to each line of an image. These concepts are examined in Sect. 5 in the wake of a detailed description of the SIT Generator.

The SIT Generator and its variants try to cope with the background simply by partitioning the image. This approach is bound to succeed provided the regions are properly outlined which almost inevitably calls for an adaptive partitioning scheme. It is not easy, however, to devise such a scheme. One possible avenue that we have explored consists in coarsely estimating the position of the targets in order to restrict the search to a smaller area than the image itself. To this end, we performed a gross line-by-line analysis of the image so as to pinpoint the lines carrying a target, based on the values of the following set of parameters (most such statistical quantities are defined in Sect. 3): mean value, median, standard deviation relative to the mean value, standard deviation relative to the median, mean value minus the median, ratio of the standard deviation relative to the mean value over the mean value itself and, finally, a coefficient of bimodality. The results obtained and given in Sect. 6, show that it should indeed be possible to effect an algorithm which will give hints as to where the targets are, thus enabling one to define a target area. However, such an algorithm will be lacking in generality since it implicitly assumes that the background is relatively uniform on a line-by-line basis. In other words, the algorithm will be orientation-dependent which certainly constitutes a major drawback in many practical situations.

The results obtained from the imagery considered through a gross line-by-line analysis prompted us to find a means of rendering the procedure orientation-independent. The problem here stems from the nonuniformity of the background and hence the solution is obviously to curb the ascendancy of the background over the image by levelling it in some way. This is in essence what the Background Elimination Technique (BET) expounded in Sect. 7.1 does. This technique operates on a one-dimensional signal (any given line or column of an image) and uses a narrow bandwidth low-pass filter to assess the general tendency of the background in order to subtract it from the signal itself. Because of its real time implementation potential, we opted for a recursive filter and, more explicitly, for a 4-pole Butterworth filter (Sect. 2 gathers background material related to such filters). Since BET can be applied either to the set of lines or columns of an image, it generates 2 images referred to as the Horizontal Fine Structure (HFS) image and the Vertical Fine Structure (VFS) image respectively. The background of these fine structure images can be considered uniform, on a large-scale basis, although it is highly textured. It is shown in Sect. 7 that one can define from HFS (extent in the y-direction) and VFS (extent in the x-direction) a relatively small-sized target area, and in many instances even pinpoint individual targets, simply by statistical considerations (as was the case of the aforementioned gross line-by-line analysis). This is very interesting since it means we can designate targets without segmenting the image.

BET and the resulting fine structure images offer many possibilities, let alone those we already apprehend. For instance, one can exploit the uniformity of these images to segment them individually

or in some combined form, with the aid of, say, a segmenter like the SIT Generator, which incidentally should be well suited for this task by its very nature. This aspect is emphasized in Sect. 8.

This work was performed at DREV between November 1978 and April 1979 under PCN 32D07 Automatic Target Acquisition.

FPBF FILTER

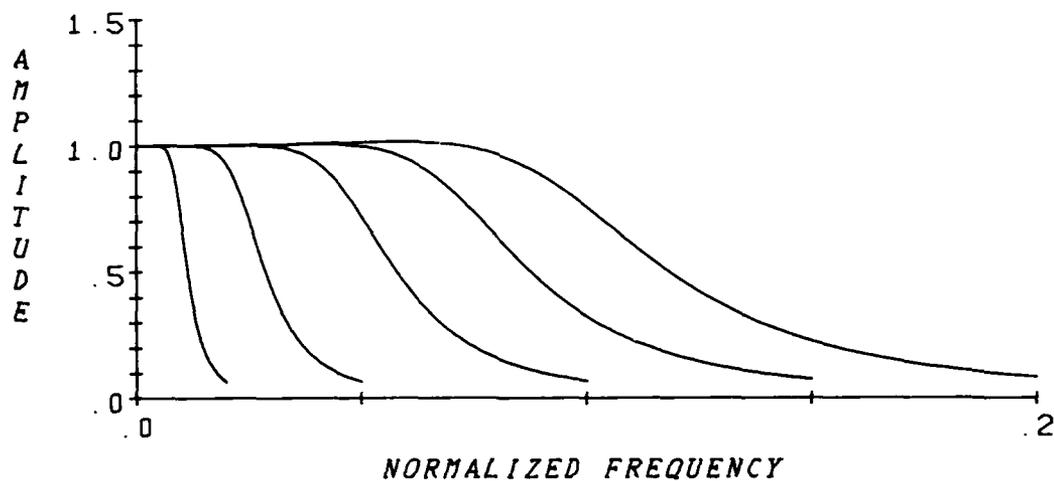


FIGURE 1 - Amplitude-vs.-normalized frequency characteristic of 4-pole low-pass digital Butterworth filters for several values of f_c/f_s (0.01, 0.025, 0.05, 0.075, 0.1)

2.0 RECURSIVE LOW-PASS DIGITAL BUTTERWORTH FILTER

A Butterworth filter approximates a rectangular passband via a monotonic amplitude-vs.-frequency characteristic (Fig. 1). The transition region of such a filter, although gradual, is more or less sharp depending on the number of poles of its transfer function - the greater the number of poles the narrower the transition band is. In this section we give without proof a certain number of results pertaining to a particular digital Butterworth filter, namely a 4-pole filter, that has been used to estimate the background of IR images. Most of this material is drawn from Refs. 1 and 2.

2.1 Z-Transfer Function

The Z-transfer function of a one-dimensional 4-pole low-pass Butterworth filter is given by (Ref. 1):

$$H(Z) = \frac{Z^2}{(Z - Z_1)(Z - Z_1^*)(Z - Z_2)(Z - Z_2^*)} , \quad [1]$$

with

$$Z_1 = \exp[-2\pi f_c (\cos 67.5^\circ - j \sin 67.5^\circ) / f_s] , \quad [2]$$

$$Z_2 = \exp[-2\pi f_c (\cos 22.5^\circ - j \sin 22.5^\circ) / f_s] , \quad [3]$$

where f_c is the 3-dB cutoff frequency of the filter and f_s the sampling frequency of the signal. The asterisk denotes the complex conjugate. When the denominator in [1] is multiplied out and the terms rearranged we obtain:

$$H(Z^{-1}) = b_0 Z^{-2} \times \frac{1}{(1 + b_1 Z^{-1} + b_2 Z^{-2})} \times \frac{1}{(1 + b_3 Z^{-1} + b_4 Z^{-2})} , \quad [4]$$

with

$$\begin{aligned} b_1 &= -(z_1 + z_1^*) , & b_2 &= z_1 z_1^* , \\ b_3 &= -(z_2 + z_2^*) , & b_4 &= z_2 z_2^* , \end{aligned} \quad [5]$$

and where b_0 is a coefficient to adjust the gain at $\omega=0$ to unity:

$$b_0 = \frac{1}{|H(Z)|} \Bigg|_{Z = e^{-j0}} = (1 + b_1 + b_2)(1 + b_3 + b_4) . \quad [6]$$

Figure 1 shows plots of $|H(Z^{-1})|$ versus f/f_s for several values of f_c/f_s .

2.2 Difference Equations

A Z-transfer function is implemented or realized via an mth-order difference equation defining what is called an mth-order digital network (Ref. 2) consisting of delays, multipliers, and summations. A direct realization of a transfer function requires the smallest amount of computation. However, in most instances it proves desirable (Ref. 2) to realize a given network by means of either cascade or parallel combinations of second-order systems because the latter realizations are less sensitive to the adverse effects associated with finite register length. Hence [4] can be written

$$H(Z^{-1}) = b_0 Z^{-2} H_1(Z^{-1}) H_2(Z^{-1}) , \quad [7]$$

that is, as a cascade of two second-order systems. The resulting set of linear difference equations (Fig. 2) is:

$$\begin{aligned}
 f_1(nT) &= x[(n-2)T] \quad , \\
 f_2(nT) &= f_1(nT) - b_1 f_2[(n-1)T] - b_2 f_2[(n-2)T] \quad , \\
 f_3(nT) &= f_2(nT) - b_3 f_3[(n-1)T] - b_4 f_3[(n-2)T] \quad , \\
 y(nT) &= b_0 f_3(nT) \quad .
 \end{aligned}
 \tag{8}$$

To proceed with this set of equations it is necessary to define the initial conditions of x , f_2 and f_3 . These are usually set to zero.

2.3 Delay Time

It can be shown (Ref. 2) that the phase characteristic of the Z-transfer function given in [1] is:

$$\begin{aligned}
 \psi = -2\omega T + \tan^{-1} & \left[\frac{b_1 \sin \omega T + b_2 \sin 2\omega T}{1 + b_1 \cos \omega T + b_2 \cos 2\omega T} \right] \\
 + \tan^{-1} & \left[\frac{b_3 \sin \omega T + b_4 \sin 2\omega T}{1 + b_3 \cos \omega T + b_4 \cos 2\omega T} \right] \quad .
 \end{aligned}
 \tag{9}$$

For small values of ωT , we have

$$\sin \omega T \approx \omega T, \quad \cos \omega T \approx 1 \quad \text{and} \quad \tan^{-1} a \approx a \quad .$$

Substituting these into [9] yields

$$\psi = -\omega T \left[2 - \frac{b_1 + 2b_2}{1 + b_1 + b_2} + \frac{b_3 + 2b_4}{1 + b_3 + b_4} \right] \quad .
 \tag{10}$$

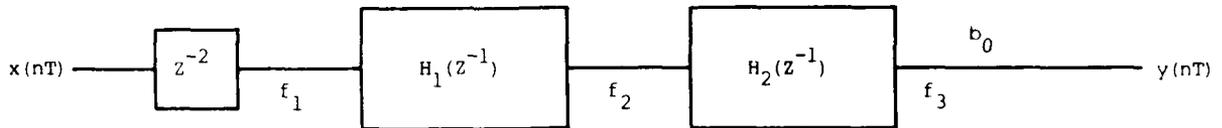


FIGURE 2 - Cascade realization of a 4-pole low-pass digital Butterworth filter; T is the sampling interval.

FPBF FILTER

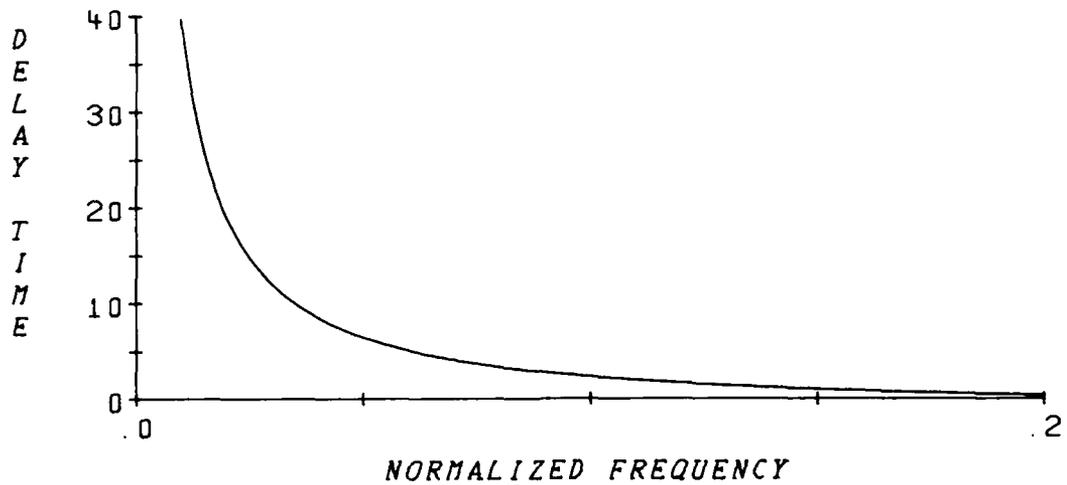


FIGURE 3 - Plot of the delay time of a 4-pole low-pass digital Butterworth filter vs. f_c/f_s

This equation represents a linear phase characteristic meaning that the output signal is shifted right a number of sampling intervals approximately equal to the quantity (delay time) enclosed in square brackets. The value of the delay time versus the normalized cutoff frequency is plotted in Fig. 3.

2.4 Equivalent FIR Filter

It is well known that the transfer function of a rectangular FIR filter is a $\sin x/x$ function. More precisely (Ref. 3)

$$H(f) = \frac{\sin \pi L f/f_s}{\pi L f/f_s} , \quad [11]$$

where L is the number of sampling intervals spanned by the filter impulse response. By definition, the 3-dB cutoff frequency is obtained by solving

$$\frac{\sin \pi L f_c/f_s}{\pi L f_c/f_s} = \frac{1}{\sqrt{2}}$$

from which we readily get

$$f_c/f_s = 0.44/L . \quad [12]$$

As a rule of thumb, we can use $1/2 L$ as the normalized cutoff frequency of any FIR filter of size L .

3.0 ON THE ESTIMATION OF SOME STATISTICAL PARAMETERS

This section gives a certain number of formulas used for estimating various statistical parameters referred to in subsequent sections. In fact, this section is nothing but a precis of statistical signal analysis as it relates to the subject of this report. For a more comprehensive treatment, the reader is directed to the references cited below.

3.1 Mean and Variance

Let $\{x_n\}$; $n=1, 2, \dots, N$ be the data values of a single time (space) history record $x(t)$. It is often desirable to think of physical data in terms of a combination of a static or time-invariant component and a dynamic or fluctuating component. The first component may be described by a mean value which is simply the average of all values (Ref. 4):

$$\bar{x} = \frac{1}{N} \sum x_n \quad . \quad [13]$$

This quantity (unless otherwise stated, all summations are for $n=1$ to N) is an unbiased estimate of the true mean value. The dynamic component may be described by a variance which is simply the mean square value about the mean:

$$s^2 = \overline{x^2} = \frac{1}{N} \sum (x_n - \bar{x})^2 \quad . \quad [14]$$

The positive square root of the variance is called the standard deviation and denoted s . The quantities s and s^2 calculated here are biased estimates of the true standard deviation and variance respectively. However, the bias is negligible for large values of N .

3.2 Skewness

The mean value and the variance are only the first two moments of a probability density function. The third and fourth moments also prove to be useful for describing physical data. The third moment or skewness measures the lack of symmetry in a density function and is defined in the following way (Ref. 5):

$$s = \sum (x_n - \bar{x})^3 / Ns^3 . \quad [15]$$

To grasp the physical meaning of the skewness it is better to write it as indicated below

$$s = \left(\sum_{x_n > \bar{x}} |x_n - \bar{x}|^3 - \sum_{x_n < \bar{x}} |x_n - \bar{x}|^3 \right) / Ns^3 . \quad [16]$$

where the vertical bars denote the absolute value. From [16] we see that in the case of a positively skewed signal the fluctuations that matter occur above the mean value, and conversely for a signal exhibiting a negative skewness.

3.3 Kurtosis

The formula used for the computation of the fourth moment, variously called kurtosis, excess or peakedness, is

$$K = \sum (x_n - \bar{x})^4 / Ns^4 - 3 \quad [17]$$

which includes a corrective factor of -3, the use of which in computing kurtosis has the effect of making both skewness and kurtosis equal to zero for a normal density function. This fact being established, leptokurtic and platykurtic density functions are defined in terms of deviations from the normal density function. Thus, the usual definitions (Ref. 6) are:

Leptokurtic - A density function that is peaked,

$$K > 0, \quad [18]$$

and

Platykurtic - A density function that is flat,

$$K < 0. \quad [19]$$

The exact meaning of the kurtosis statistic is not clear to statisticians (Refs. 6-9), let alone to laymen in this field. It seems that it has long been accepted that a symmetrical platykurtic density function, with $K < 0$, is characterized by a flatter top and more abrupt terminals than the normal curve and that a symmetrical leptokurtic density function, with $K > 0$, has a sharper peak at the mean and more extended tails. However, Chissom (Ref. 6) cautions that it is difficult to determine the shape of a density function from the kurtosis value alone, since almost any density function may have a negative kurtosis

value. Nonetheless, he recognizes that to have a positive kurtosis value the distribution of measures must contain a good number of cases in the tails, i.e. a tailing off effect must be present. Darlington (Ref. 7), for one, reveals another amazing aspect of kurtosis. He wonders if kurtosis is really peakedness, and concludes that a better term for describing it is "bimodality", where the lower the kurtosis, the greater the bimodality. Clearly, the most bimodal of all possible density functions is a symmetric 2-point density, while the least bimodal (or most unimodal) density function is concentrated entirely at one point. It can be shown (Ref.7) that these density functions have respectively lowest and highest kurtosis because in a symmetric 3-point density in which p is the density at the mean,

$$K = 1/(1 - p) - 3 \quad . \quad [20]$$

As p approaches 1 (i.e. as the density approaches being concentrated entirely at its mean), K approaches infinity. On the other hand, when $p=0$ (i.e. when the density is a 2-point, rather than a 3-point, density) K achieves its lowest possible value of -2. But to confuse the issue, Hildebrand (Ref. 8) exhibits a family of density functions that are solidly bimodal, but have kurtosis coefficients ranging from -2 to +3.

In spite of all the trickiness associated with the kurtosis statistic, the inequality $K \geq c$, where c is an appropriate constant, has been used in practice as a test of a normal density against densities with heavier tails or, more generally, for testing light-tailed densities against heavy-tailed ones. Other statistics (Refs. 9-10) used

for the same purpose are:

$$U = \frac{s^2}{\sum |x_n - m|/N}, \quad [21]$$

$$V = \frac{Z/2}{\sum |x_n - m|/N}, \quad [22]$$

$$W = \frac{Z/2}{s^2}, \quad [23]$$

$$a = \frac{\sum |x_n - \bar{x}|/N}{s^2}, \quad [24]$$

where m and Z are respectively the median and the range (high extreme minus low extreme) of the set of data. According to Hogg (Ref. 9), W should be used only when trying to detect if a density function is light-tailed or not. For the normal density function, the value of the ratio defined by [24] is $\sqrt{2/\pi}=0.7979$; this ratio will be higher for platykurtic and lower for leptokurtic density function. The same is true in reverse for the U statistic.

3.4 Autocorrelation Function

The autocorrelation function at the displacement r is defined (Ref. 4) by the formula

$$\hat{R}_r = \hat{R}_x(r) = \frac{1}{N-r} \sum_{n=1}^{N-r} (x_n - \bar{x})(x_{n+r} - \bar{x}) \quad [25]$$

where r is the lag number, and \hat{R}_r is the estimate of the true value R_r at lag r . The autocorrelation function may take on negative as well as positive values. A normalized value for the autocorrelation function is obtained by dividing \hat{R}_r by \hat{R}_0 where

$$\hat{R}_0 = \hat{R}_x(0) = \frac{1}{N} \sum (x_n - \bar{x})^2 = s^2 . \quad [26]$$

when \hat{R}_r is normalized, one obtains the quantity \hat{R}_r / \hat{R}_0 which theoretically will be between plus and minus one, that is,

$$- 1 \leq \hat{R}_r / \hat{R}_0 \leq 1 . \quad [27]$$

The importance of the autocorrelation function for describing physical data stems from the fact that a sharply peaked autocorrelogram which diminished rapidly to zero, is typical of wide-band random data. For the limiting case of hypothetical white noise (random data with energy distributed uniformly over all frequencies), the autocorrelogram is a Dirac delta function at zero displacement.

4.0 IR IMAGE MODELLING

The algorithms developed in this report, for detecting targets in IR imagery, are based on the simple assumption that the targets appear as hot regions within a cooler slowly varying surround. By slowly varying surround we mean that in the absence of any targets the main fluctuations, i.e. large-amplitude fluctuations, are concentrated at the lower end of the spatial frequency spectrum. Superimposed on this continuous background, which accounts for gradations of gray level across the image, there might be (Fig. 4) sharp lines due to relatively small-size targets.

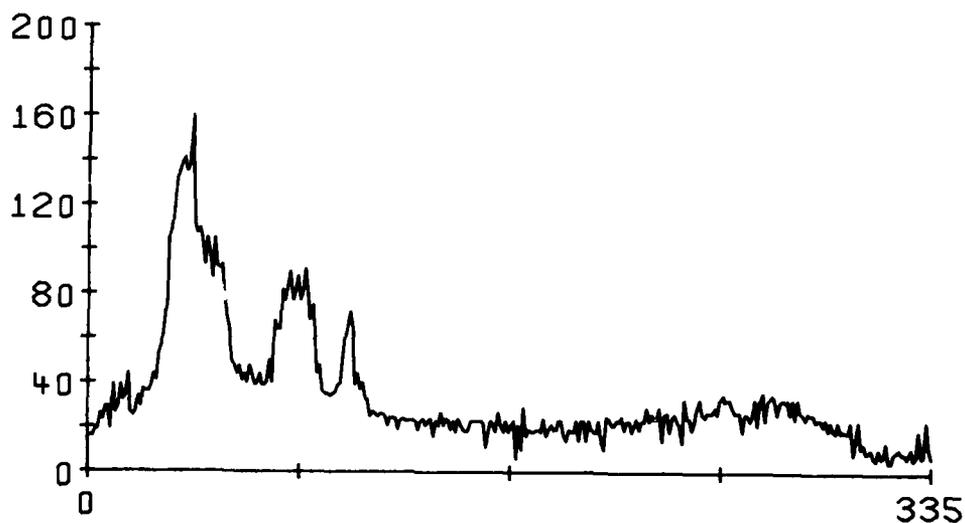
LINE 175, IMAGE ALA 6 3

FIGURE 4 - Line 175 of image 6 (gray level or brightness vs. column number) from the Alabama Data Base. The 3 peaks correspond respectively to a tank, an APC and a jeep. Such a signal can be interpreted as a set of sharp lines superimposed on an otherwise slowly varying background.

5.0 SILHOUETTE GENERATOR

We will first describe an algorithm that has already been used to detect targets (Refs. 11-13) in IR BOFORS imagery. This algorithm is part of a computer simulated Automatic IR Target Acquisition System (AIRTAS) and was previously referred to as a silhouette generator. This potential device starts from partial histograms and attempts to estimate the gray level corresponding to the maximum temperature prevailing in the background. The thermoscopic image (Fig. 5A1) that illustrates the working of the silhouette generator measures 420 x 335 pixels. It is extracted from the Alabama Data Base where it is labeled ALA 6 3 (the last digit specifies the spectral region: 3-5 μm band or 8-14 μm band). Figure 5A2 is an histogram equalized version of this image showing more clearly details of the scene depicted.

5.1 Single Intensity Threshold (SIT)

The defining procedure of the original version of the silhouette generator is:

- 1) Divide the image into a certain number of subimages.

The way a given image must be split should really be determined by experiment. Because of its size (96 x 256 pixels), a BOFORS image was solely divided along the horizontal axis. With a thermoscopic image, on the other hand, we get best results when we divide both the horizontal and vertical axes (Fig. 5A3) into the same

number of regions, namely 4. We generally tend to make the subimages about square although this is not absolutely needed. However, as a rule, at least one subimage should be representative of the background, i.e., should not contain any targets. Moreover, it should be large enough to provide a good estimate of the gray level corresponding to the highest temperature of the background.

- 2) Determine the histogram of each subimage.

This is the main mathematical operation performed by the silhouette generator and since it is a one-pixel-at-a-time process it can be easily implemented, "on the fly", by real time hardware.

- 3) Determine the cutoff gray level of each partial histogram.

The histogram being scanned from the highest bin down, the cutoff gray level (upper gray level of the background) is defined as the gray level of the first bin occupied by at least 3 pixels. One can imagine many variants to this scheme and it might be worthwhile to investigate this point further.

- 4) Choose the smallest cutoff gray level as an intensity threshold for the whole image.

In doing this one should exert some caution because it might well happen that the smallest cutoff gray level will be zero or something very small. To obviate such nonsenses we restrict the choice to those cutoff gray levels greater than the 80th percentile of the whole image. This amounts to assuming that less than 20% of the surface of the image is occupied by targets.

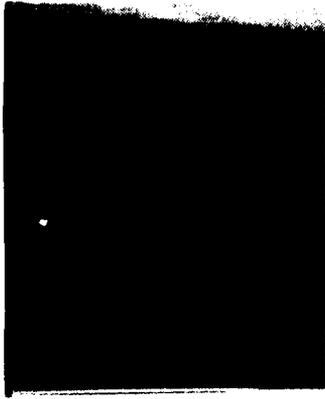
The thermoscopic image of Fig. 5A embraces 3 targets in a row near the center of the image. From left to right, these correspond to a tank, an APC, and a jeep respectively. The result obtained by applying the silhouette generator to this unprocessed raw image is shown in Fig. 5C1 (the pixels whose gray levels are greater than the threshold are saturated while those whose gray levels are less than or equal to the threshold are zeroed). This example demonstrates that under certain circumstances the miss rate of the silhouette generator is unacceptably high, and that the shape of the detected targets might be altered. On the positive side, the segmented image is clean and consists of solid blobs that are relatively easy to interpret - a blob stretching from one side of the image to the other is certainly not a potential target.

5.2 Thresholding Intensity Functions

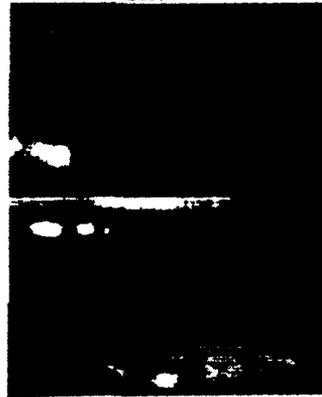
The algorithm just described is best suited to detect the brightest targets. It will inevitably miss faint targets because many background pixels have gray levels in the same range as the targets themselves. In consequence, the 80th percentile is driven much too far in the light portion of the gray scale. Using a lower percentile will

FIGURE 5 - Silhouette generator

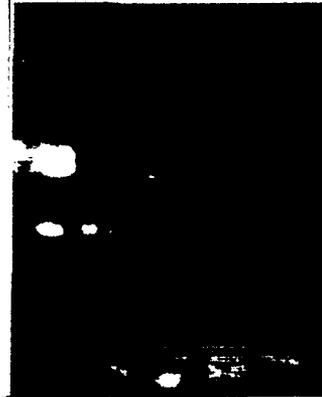
- A) Image ALA 6 3 (1: raw; 2: histogram equalized; 3: subimages delineated);
- B) Thresholding intensity functions;
- C) Segmented images;
 - 1) Single intensity threshold;
 - 2) Staircase intensity threshold;
 - 3) Interpolated staircase intensity threshold.



1



2



3

A

B

C

not generally help because we might end up with targets embodied in a very large blob. The probability of detection of the silhouette generator can, however, be improved if we treat the image as (for the case of Fig. 5A1) 4 vertically shifted images of size 105 x 335 pixels and apply the algorithm to each of them independently. In this way, a new threshold (Fig. 5B2) is derived for each horizontal slice and the final result (Fig. 5C2) is a segmented image where the 3 targets stand out clearly, and where their shape is better preserved. However, artifacts may appear if the thresholds of 2 adjacent slices differ widely, as is obvious in Fig. 5C2. A manifest remedy is to smooth the transition between 2 slices by, say, linearly interpolating the relevant thresholds. What results is a thresholding intensity function (Fig. 5B3), that is, a function attributing a specific intensity threshold to each line of the image. The segmented image (Fig. 5C3) generated by this continuous function is quite similar to the one obtained with a staircase function except that there are no artifacts. The only noticeable flaw seems to be a slight alteration of the shape of the targets. The concept of a thresholding function can be easily extended. However, as far as the silhouette generator is concerned, the crux remains the manner in which the subimages or the slices are defined.

6.0 GROSS STRUCTURE ANALYSIS (GSA)

In the previous section, we have shown that the silhouette generator works fairly well provided the background is relatively uniform. For those situations where this is true, we can rely entirely on the Sit Generator. When this assumption does not hold, as is generally the case with ground scenes, we can circumvent the problem by slicing the image into a certain number of partial images presenting each a uniform background. The unsettled question we will now tackle is the way of defining the slices.

6.1 Gross Structure Statistics

As the targets are small and their number is limited, the image is mostly background. Devising an algorithm that would discard large portions of the image, so that we could restrict the search to a certain target area much smaller in size than the image would help us greatly. If the search area is smaller than the entire image, chances are that the embedded background will be almost uniform. As a first attempt in this direction, one may treat the lines of an image as a collection of one-dimensional random signals and try to flag, by measuring various statistical parameters, the lines that intersect a target. Figure 6 shows plots of 7 statistical parameters computed for the image of Fig. 5A1 (we state again that all computations in this report are performed on original unprocessed raw images but that, for display purpose only, the images are postprocessed using a histogram-equalization technique). These 7 parameters are:

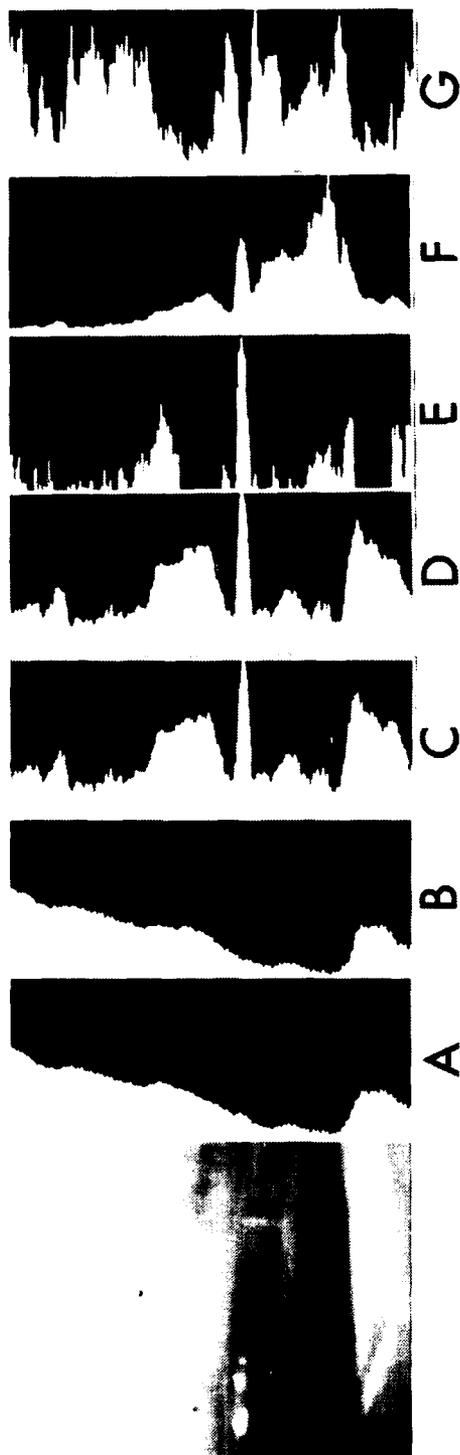


FIGURE 6 - The image on the left is image ALA 6 3. The 7 plots on its right side show the line number vs. the following quantities:

- A) Mean value (scale: 0-255);
- B) Median (scale: 0-255);
- C) Standard deviation relative to the mean value (scale: 0-30);
- D) Standard deviation relative to the median (scale: 0-32);
- E) Mean value minus the median (scale: 0-12);
- F) Ratio of C) over A) (arbitrary scale);
- G) Coefficient of bimodality (scale: 0-140).

- a) mean value,
- b) median,
- c) standard deviation relative to the mean value,
- d) standard deviation relative to the median,
- e) mean value minus the median,
- f) ratio of c) over a),
- g) coefficient of bimodality.

From the plots of these quantities, we draw the following conclusions:

- 1) The mean value and the median are of no great use per se. Nevertheless, they illustrate the fact that the background luminance varies slowly but with an amplitude that can be large.
- 2) The trends of the mean value and of the median are about the same and so are the standard deviations relative to both.
- 3) The standard deviations exhibit a well defined peak whose height is an absolute maximum and whose location matches the position of the targets.
- 4) The absolute maximum of the difference between the mean value and the median also lines up with the 3 targets. The idea of using this difference as an estimator stems from the fact that the mean value is very sensitive to outliers that might be present in a set of data whereas the median

is not. Based on what we said in Sect. 4, we can expect this difference to be positive. Figure 6E confirms that this is indeed the case.

- 5) The ratio of the standard deviation relative to the mean value over the mean value itself is not very informative. Because the mean value can go very low no meaningful peak can be localized.
- 6) The coefficient of bimodality is low for target lines. However, this must be interpreted as a seemingly necessary but insufficient condition. The coefficient of bimodality is defined (Fig. 7) as the number of times a given signal crosses its mean value, where the lower this coefficient, the greater the bimodality.

It should be possible with these findings to build an algorithm that will give hints as to where the targets are and, therefore, enable one to define a target area. However, we did not pursue this line because the results are highly directional (for the case under discussion, for example, processing the columns instead of the lines would be frustrating). The underlying assumption is to the effect that the background is "uniform" for almost each member of the set of signals considered (lines or columns). Although this assumption is much less restrictive than assuming the background is uniform for the whole image, it is nevertheless too restrictive for applications involving ground scenes.

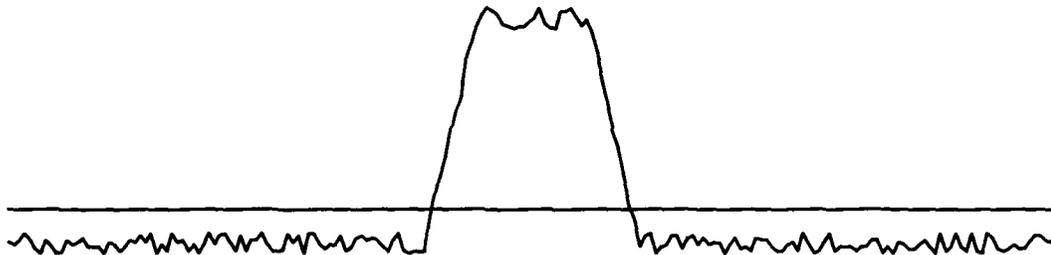
COEFFICIENT OF BIMODALITY

FIGURE 7 - The coefficient of bimodality of the signal illustrated here is 2 since it crosses twice the line corresponding to its mean value. This is typical of a strongly bimodal signal.

7.0 FINE STRUCTURE ANALYSIS (FSA)

The previous sections demonstrate that the background constitutes a stumbling block that is very difficult to circumvent. Why not then, instead of dealing with the image in its integrity, try to eliminate the background, or at the very least to render it more "uniform". This is what we study in this section as a first step to further processing.

7.1 Background Elimination Technique (BET)

In Sect. 4 it was said that a signal (a given line or column of an image) bearing a target can be thought of as composed of a sharp peak superimposed on an otherwise continuous (slowly varying) background. To estimate the background one must then find a way to smooth the signal but without including the peaks that might be part of it. The most straightforward approach is to use a narrow bandwidth low-pass filter to guess the general tendency of the background and then subtract it from the signal. A low-pass digital filter can have a finite impulse response (FIR filter) or an infinite impulse response (IIR filter) and either can be realized recursively or nonrecursively. Because of its real time implementation potential, we opted for a recursive IIR filter and, more explicitly, for a 4-pole Butterworth filter (FPBF). Other digital filters might do as well or better than this one but, since we obtained good results with the FPBF, we did not explore other possibilities.

To illustrate BET we will use the signal of Fig. 4, which corresponds to line 175 of image 6 from the Alabama Data Base. This

signal is fed to a low-pass FPBF digital filter whose 3-dB normalized cutoff frequency (f_c/f_s) is equal to 0.01. The filtered signal generated is shown in Fig. 8A along with the input signal. From [12] we see that to obtain an equivalent result with a rectangular FIR filter, the filter size must be equal to 44. It is obviously advantageous in such a situation to rely on a recursive IIR filter. Two points are worth mentioning about the filtered signal of Fig. 8A:

- a) since we deemed the initial conditions to be zero, there is a droop in the curve at its origin, and
- b) the filtered signal is shifted to the right. This is evidenced by the distance separating the absolute maximums of the 2 curves. Using [10] one can check that the shift spans about 42 sampling intervals.

The first anomaly can be easily corrected by selecting the initial conditions so that there is no transient at the origin. At this point, the filter sees a step function of height H , where H is the value of the signal at $t=0+$. One can then prove that the required initial conditions are (the quantities appearing below are defined in Sect. 2):

$$f_3(nT) = H/b_0$$

and

[28]

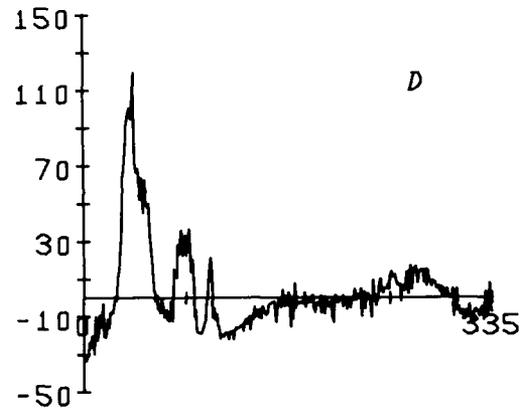
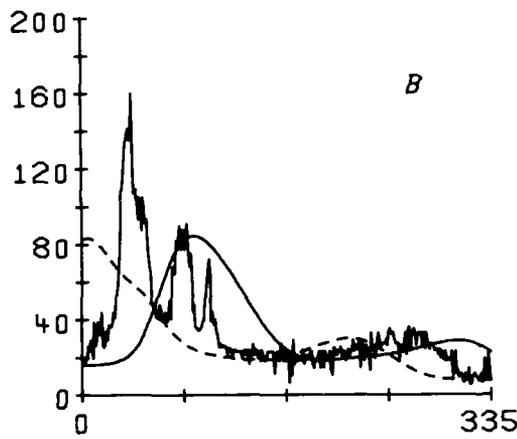
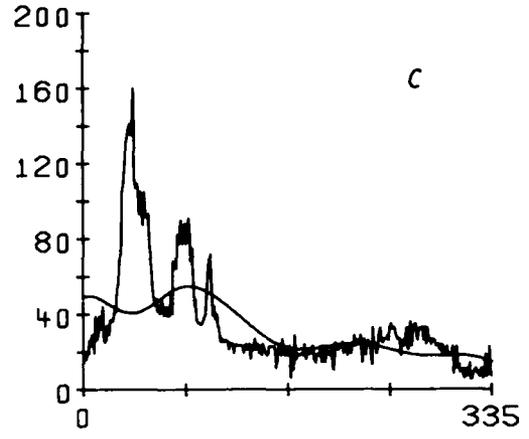
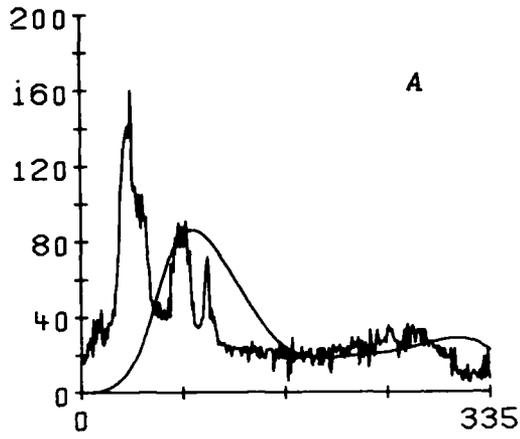
$$f_2(nT) = H/(1 + b_1 + b_2)$$

for $n \leq 0$. These initial conditions are nothing but the asymptotic response of the filter to a step function. Let us notice that one can

FIGURE 8 - Background Elimination Technique (BET); Line 175 of image
ALA 6 3

- A) FPBF filter initially at rest;
- B) FPBF filter with nonzero initial conditions; the solid line is the left filtered signal and the dashed line the right filtered signal;
- C) Arithmetic mean of the two filtered signals;
- D) Fine structure or fluctuating component of the input signal.

The cutoff frequency of the filter is 0.01.



use for H the average of the first 3 or 5 pixels, or anything else, in lieu of the first pixel alone; this may even be necessary if the first pixel manifests a tendency to wildness. Figure 8B shows the filtered signal (solid line) that results when we use these new initial conditions.

The second anomaly can be as easily corrected by shifting the filtered signal to the left but we will take advantage of it to clip the peaks. Let us consider Fig. 8B. The signal is fed to the filter from left to right. Normally, we would expect the filtered signal to peak at, or close to, the position of the main spike in the input signal. Instead, it overshoots to the right. Therefore, had the signal been fed from right to left, the overshoot would have occurred to the left (dashed line in Fig. 8B). By combining both filtered signals in some fashion, we can expect to obtain a curve that will completely bypass the peaks to follow only the broad characteristics of the input signal. The following combinations were tried:

- 1) minimum value,

$$y(t) = \begin{cases} x_L(t), & \text{if } x_L(t) < x_R(t) \\ x_R(t), & \text{otherwise ;} \end{cases} \quad [29]$$

- 2) arithmetic mean,

$$y(t) = (x_L(t) + x_R(t))/2 ; \quad [30]$$

- 3) geometric mean,

$$y(t) = \sqrt{x_L(t) x_R(t)} ; \quad [31]$$

where $x_L(t)$ and $x_R(t)$ are the left and right filtered signals respectively. All things considered, the arithmetic mean (Fig. 8C) was judged most satisfactory. Figure 8D exhibits the fine structure (fluctuating component) of the illustrative signal, that is, what is left of the signal once the estimated trend of the background is removed.

The crux of BET is the choice of the proper bandwidth or, what comes to the same thing, cutoff frequency of the FPBF filter. We attempted to define a procedure (based on Fourier spectra) for selecting it but with no great success. Figures 9 and 10, built on the model of Fig. 8, shed some light on the problem and its possible solution. This signal constitutes a challenge since the target sits right in the middle of a narrow well. If we use a cutoff frequency of 0.01 (Fig. 9), the inertia of the filter is such that the well is bypassed, and consequently shows up again in the fluctuating component (Fig. 9D). We would of course like the filtered signals to follow the well. For this purpose we have to use a wider bandwidth, as in Fig. 10 where the cutoff frequency is 0.05. There the fluctuating component (Fig. 10D) is intuitively more satisfactory. Another aspect of the same question relates to the delay time introduced by the filter. That of a 0.01 cutoff frequency filter is about 40 sampling intervals and that of a 0.05 filter, 6. Therefore, combining the right and left filtered signals, we can say that the first filter is geared to clip peaks 80 sampling intervals wide, whereas the second filter is possibly restricted to much narrower peaks, of the order of 12 sampling intervals. This is well evidenced by Fig. 10B where the filtered signals peak on either side of the target hot spot. Conclusively, BET

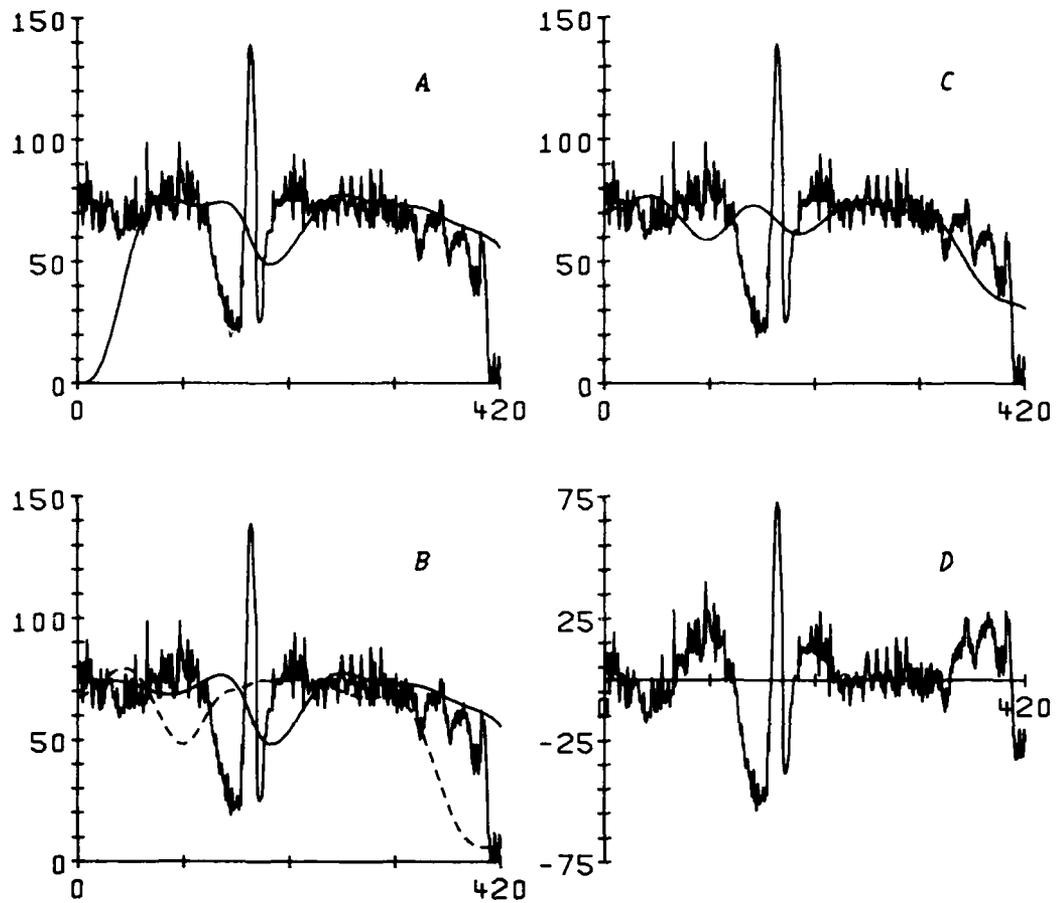


FIGURE 9 - Column 98 of image ALA 4 8; $f_c/f_s = 0.01$

although akin to highpass filtering, differs from it in one important way, the target peaks are saved, i.e., they are not split into 2 peaks respectively for the left and right edges.

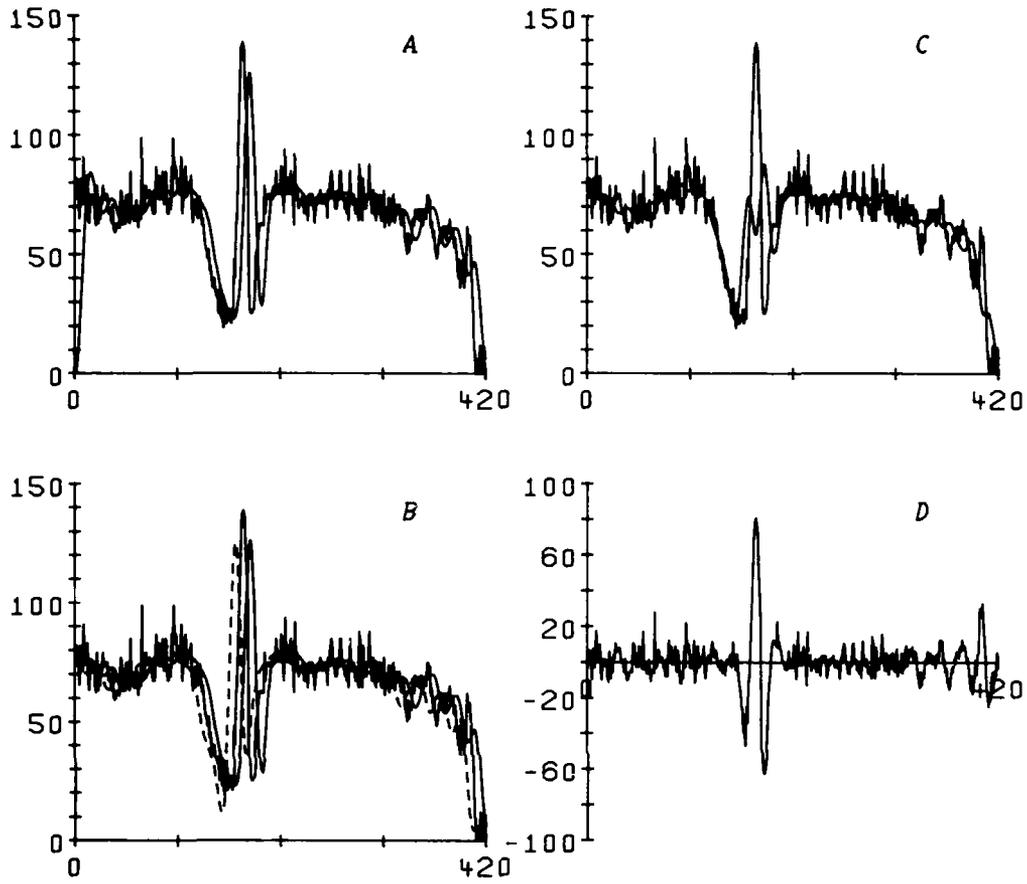


FIGURE 10 - Column 98 of image ALA 4 8; $f_c/f_s = 0.05$

7.2 Fine Structure Statistics

We dropped the analysis of the gross structure of an image (Sect. 6) because the results, for ground scenes at least, are highly directional. There was a great temptation to resume this sort of analysis with the fine structure of the image. We did not resist. However, it was not long before we realized that the set of statistical parameters would have to be enlarged. For one thing, both the mean value and the median are meaningless (both are close to zero) and, of course, so is the difference between these two quantities and the ratio of the standard deviation over the mean value. For another thing, the coefficient of bimodality, as defined in Sect. 6.1, is no longer informative. To make up for these parameters, we threw in the skewness, the kurtosis and the correlation length. We did some exploratory work with the U, V, W and a statistics (Sect. 3) but we abandoned when it became obvious we were heading for a disappointment.

The statistical parameters used to characterize the fine structure of an image are then the variance (hereafter designated by V ; there should be no confusion with V statistics since we will not refer anymore to this one), the skewness (S), the kurtosis (K) and the correlation length. The defining formulas of the first 3 parameters and their physical meaning are given in Sect. 3. The definition of the correlation length is intermingled with that of the normalized autocorrelation function [25-27]. Since full determination of the autocorrelation function is liable to use too much computing time (in prospect of a real time hardware implementation of the ideas put forward here) only the first value (at lag number 1) of this function will be

determined. However, we will assume that the shape of the autocorrelation function matches an exponential curve,

$$C_r = \hat{R}_r / \hat{R}_0 = \exp(-r/L) \quad [32]$$

where L is the correlation length in unit of sampling interval. We are going to use the correlation length as a rejection criteria, that is, if the correlation length of a given signal is less than ℓ , or if the first value, C_1 , of the normalized autocorrelation function is less than $\exp(-1/\ell)$, the signal is discarded as noise. From this standpoint, the postulated shape is rather conservative for two other commonly postulated shapes (straight line and Gaussian) have a greater value (Fig. 11) at $r=1$.

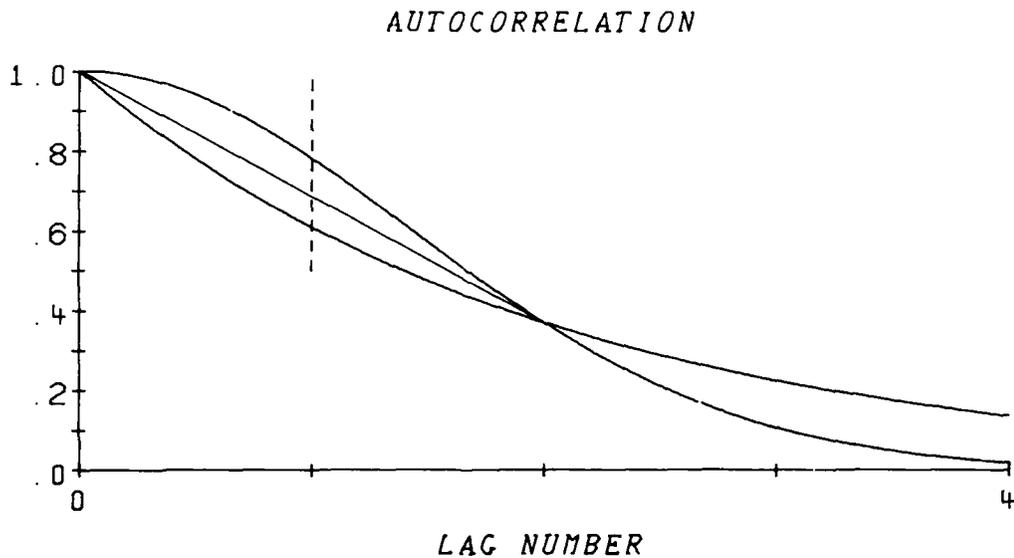


FIGURE 11 - This figure depicts 3 curves having the same correlation length ($L = 2$). These correspond to commonly postulated autocorrelation shapes.

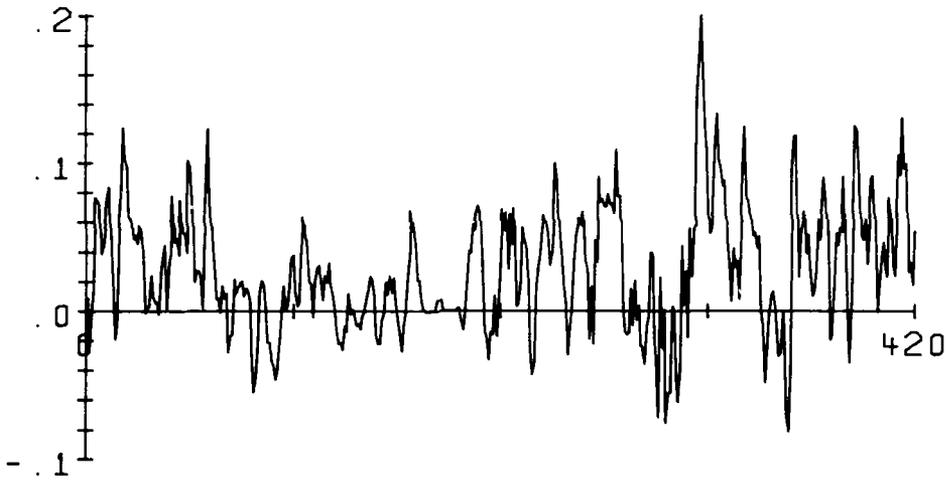
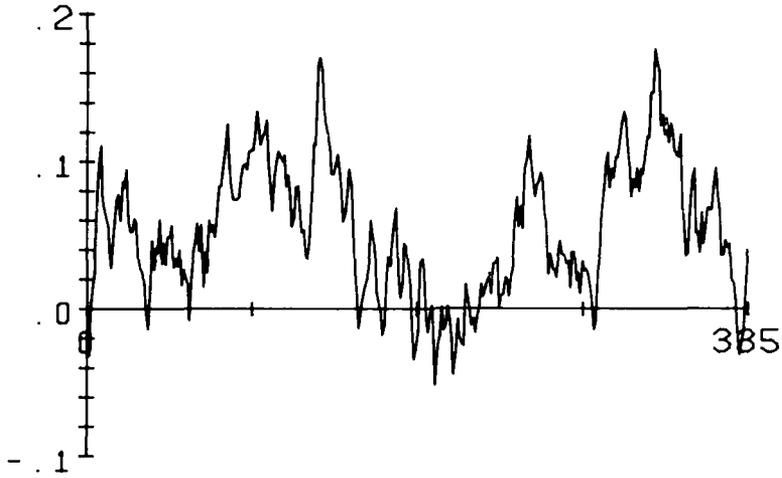
Figure 12 shows smoothed plots (3-point moving average) of the 5 statistical parameters (\bar{x} , V , S , K and C_1) ascribed to each line and column of image ALA 31 8. We recall that to obtain them, the signal at hand (a given line or column of the original raw image) is first deprived of its background by using BET in conjunction with a low-pass FPBF filter whose cutoff frequency is 0.05. The 5 statistical parameters characterizing the resultant fluctuating component are then calculated by using the formulas given in Sect. 3. As mentioned before, the mean value (Fig. 12A) is useless in regard to information content, but since we need it to determine the other parameters we have to calculate it anyhow. The variance, on the contrary, is highly informative. It exhibits well-defined peaks (Fig. 12B), both along the horizontal (top curve) and vertical (bottom curve) axes, whose position corresponds precisely to the position of the targets. Moreover, each of the 3 peaks that are part of the top curve spans a number of columns representative of the width of the underlying target. In the other direction, as the targets lined up, the width of the chief peak matches the height of the largest target. It is to be noted that these are qualitative observations. To do otherwise, one would have to define what is meant by the width of a peak. The third statistical parameter, skewness, displays (Fig. 12C) the same behavior although less convincingly, particularly as regards the vertical axis. Also, one notices a negative spike that can be tracked down in the image as a cold spot. In reality, this is a burn mark reproduced in all the images of the aforementioned data base. Figure 12D gives the value of the kurtosis for every line and column of image ALA 31 8. Here again, the targets are easily localized on the top graph, whereas the bottom curve is misleading. The isolated peak to the right might well be interpreted

as arising from a target hot spot but it may also result from a cold one. To remove the ambiguity we must revert to the skewness: if the skewness is negative it is a cold spot, otherwise it is a hot spot. We recalled from Sect. 3 that a high kurtosis value ascertains the presence of outliers in a set of data, but whether these outliers occur above or below the mean value can only be fixed by the skewness. In spite of its title, Fig. 12E displays the first value of the normalized autocorrelation function. The dashed lines on these graphs correspond to the value of C_1 calculated for an exponential autocorrelation function, [32], whose correlation length is 1.5 ($C_1 = 0.51$), 1.0 ($C_1 = 0.37$) and 0.5 ($C_1 = 0.14$). The idea is to set a lower threshold to discard abnormal or noisy signals. A threshold of 0.51 works well for the case considered but, as a rule, it is too severe. On the opposite side, a threshold of 0.14 does not take a high toll but then one may question its usefulness. The only threshold left ($C_1 = 0.37$) proved, in the light of experimental results, to be unreliable. We will explain in a next section how we managed to solve the problem. However, we did not arrive at a clear-cut solution and the role as well as the usefulness of the normalized autocorrelation function, in the analysis of the fine structure of an image, will have to be reassessed.

UNCLASSIFIED
40

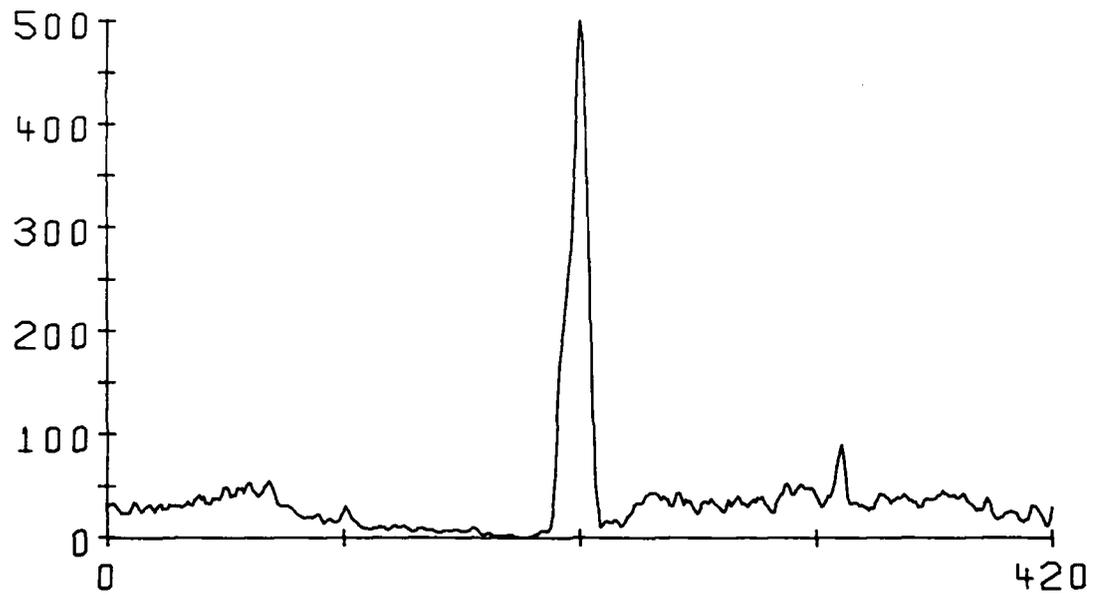
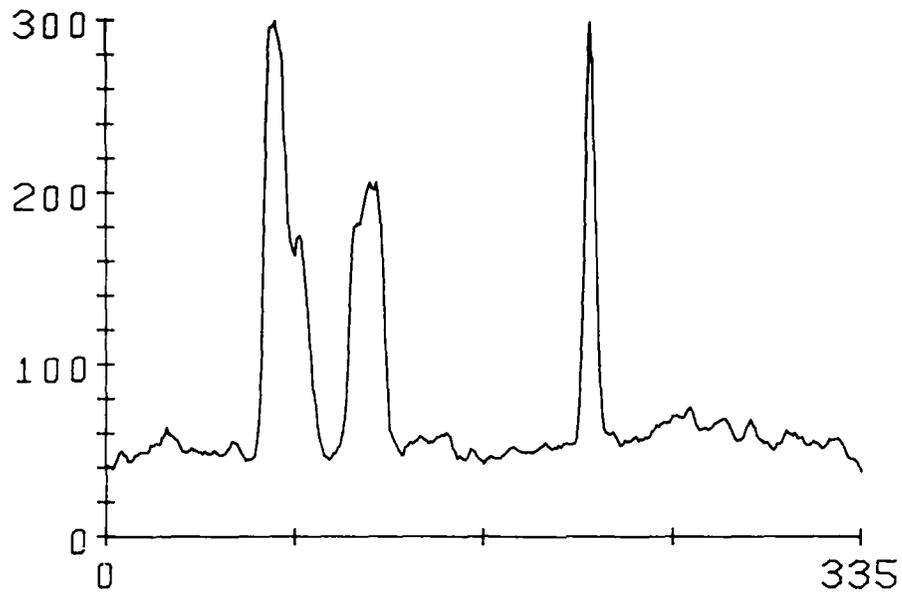
FIGURE 12 - Fine structure statistics of image ALA 31 8. The cutoff frequency of the FPBF filter used in conjunction with BET is 0.05. All these curves were smoothed using a 3-point moving average. Top records correspond to column statistics and bottom records to line statistics.

MEAN (IMAGE 31; F=0.05)

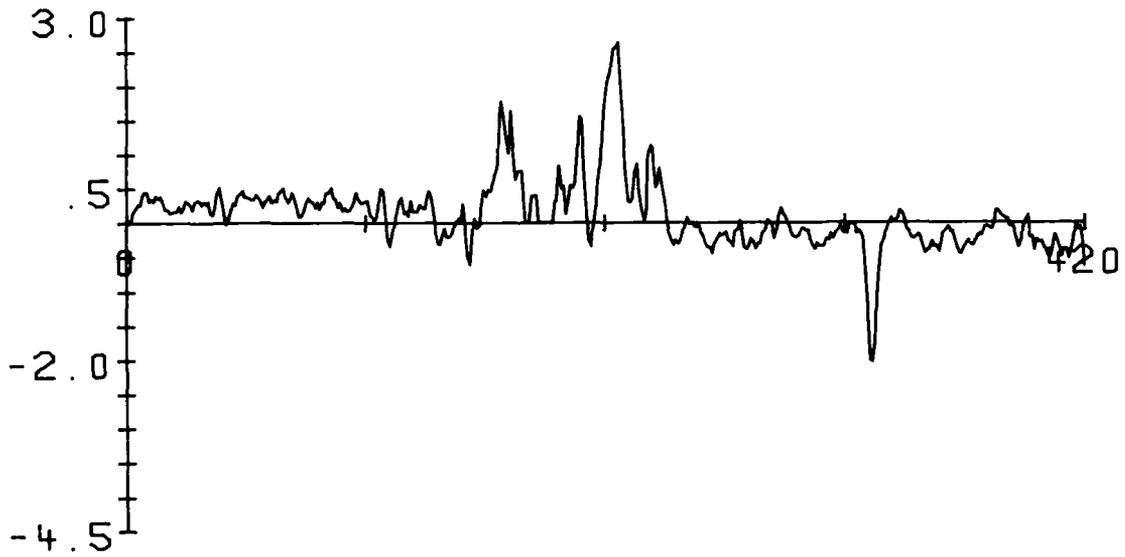
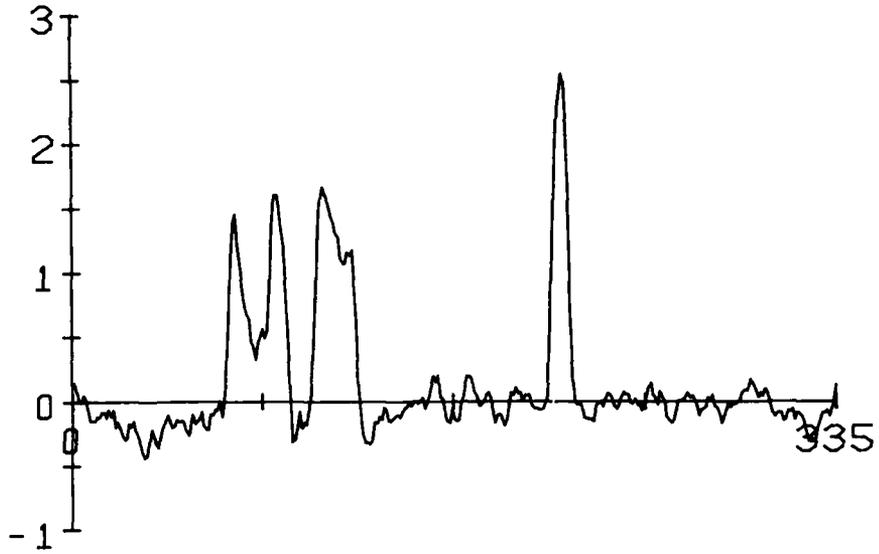


UNCLASSIFIED
42

VARIANCE (IMAGE 31; F=0.05)



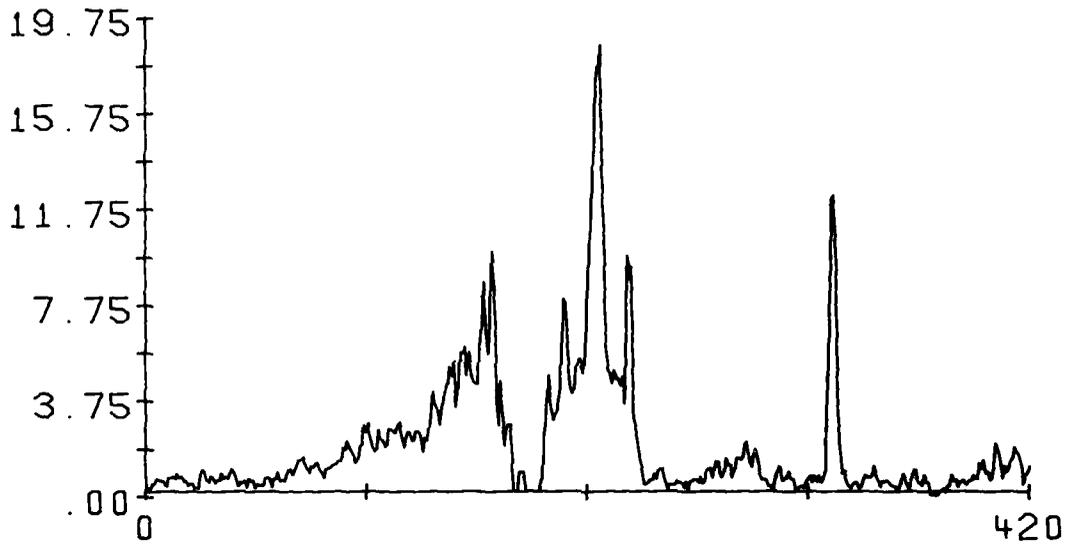
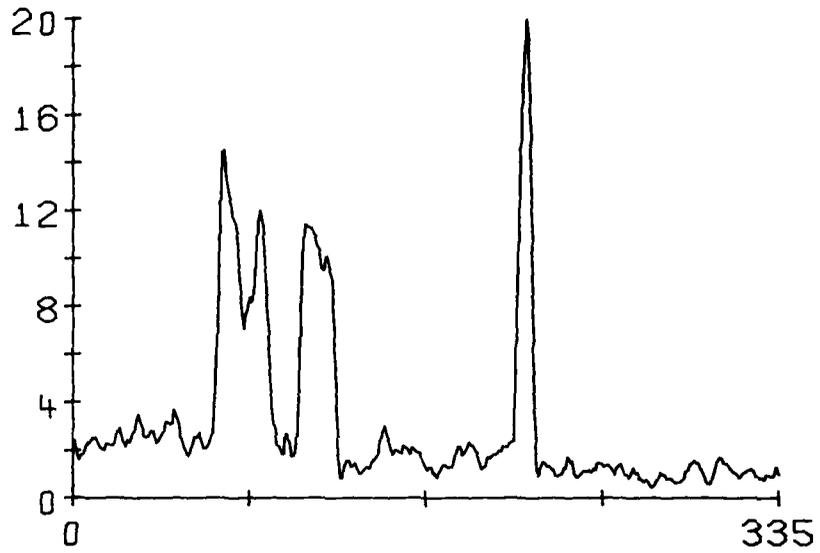
SKEWNESS (IMAGE 31; F=0.05)



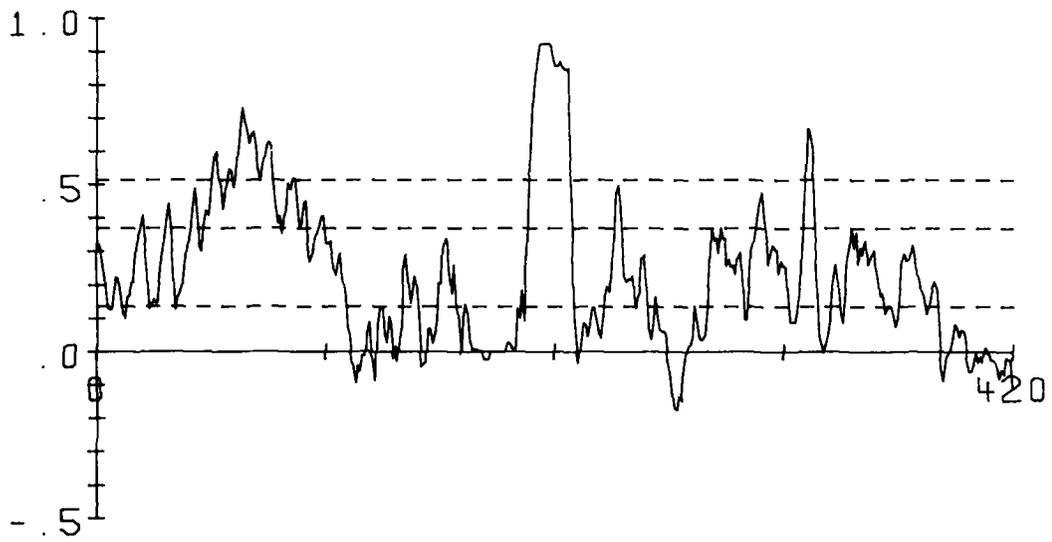
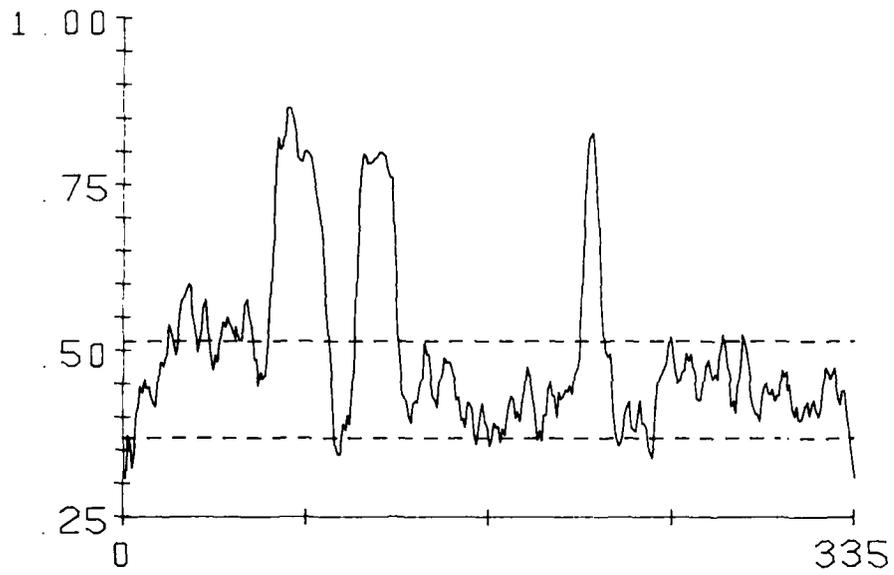
UNCLASSIFIED

44

KURTOSIS (IMAGE 31; F=0.05)



CORR. LENGTH (IMAGE 31; F=0.05)



7.3 Target Area Delimitation

Plots such as those of Fig. 12 contain all the information one needs to pinpoint individual targets, or else to delimit a relatively small-sized target area. We must now extract this information in a way amenable to automation. We tried various schemes that are in fact variations on the same theme - normalization and multiplication of a few parameters coupled with a rejection criteria. After a long trial and error, we decided on the following procedure based on the product of V , the variance, by K , the kurtosis:

- a) starting from raw data, the records of the various statistical parameters are first smoothed using a 3-point moving average;
- b) the variance records are normalized so that their maximum value is 1:

$$\tilde{V} = V/\max(V); \quad [33]$$

- c) the kurtosis records are balanced and the values less than zero clipped prior to normalization:

$$\underline{K} = \max((K-\bar{K}), 0), \quad [34]$$

$$\tilde{K} = \underline{K} / \max(\underline{K}), \quad [35]$$

where K is the mean value of the record at hand;

- d) the skewness records are likewise balanced:

$$\tilde{S} = S - \bar{S}; \quad [36]$$

e) the threshold, q , for abnormal or noisy signals, in relation with the autocorrelation records, is set to 0.51 ($L=1.5$) provided this value does not exceed the upper quartile (p_{75}) of the record. Otherwise, it is set to 0.37 ($L=1$), subject to the same condition, and as a last resort to 0.14 ($L=0.5$):

$$q = \begin{cases} 0.51 & \text{if } p_{75} > 0.51, \text{ otherwise} \\ 0.37 & \text{if } p_{75} > 0.37, \text{ otherwise} \\ 0.14; & \end{cases} \quad [37]$$

- f) we form VK-product records as follows:

$$VK(j) = \begin{cases} \tilde{VK}(j) & \text{if } \tilde{S}(j) > 0 \text{ and } C_1(j) > q \\ 0, & \text{otherwise.} \end{cases} \quad [38]$$

Figure 13 shows the results of this procedure for the case of image ALA 31 8 considered before. From these graphs, we conclude easily that there are 3 targets in this image and that they are ranged in a row right in the middle of it. We are sure that the number of targets is 3 for there is only 1 peak along the vertical direction. However, had 2 peaks been present in the bottom graph of Fig. 13, we would have been

PRODUCT VK(S,L) (IMAGE 31; F=0.05)

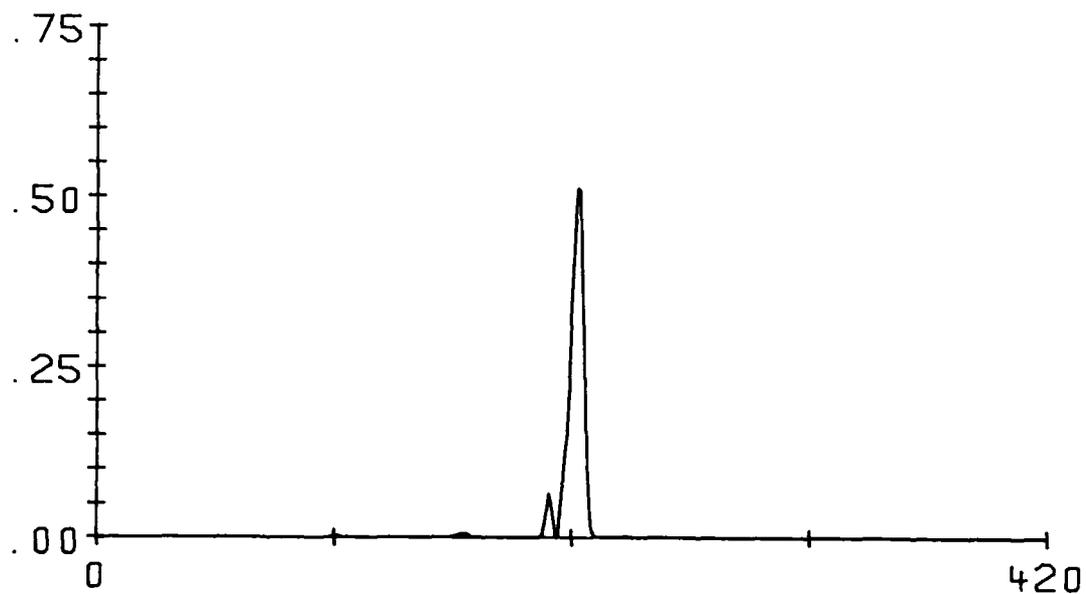
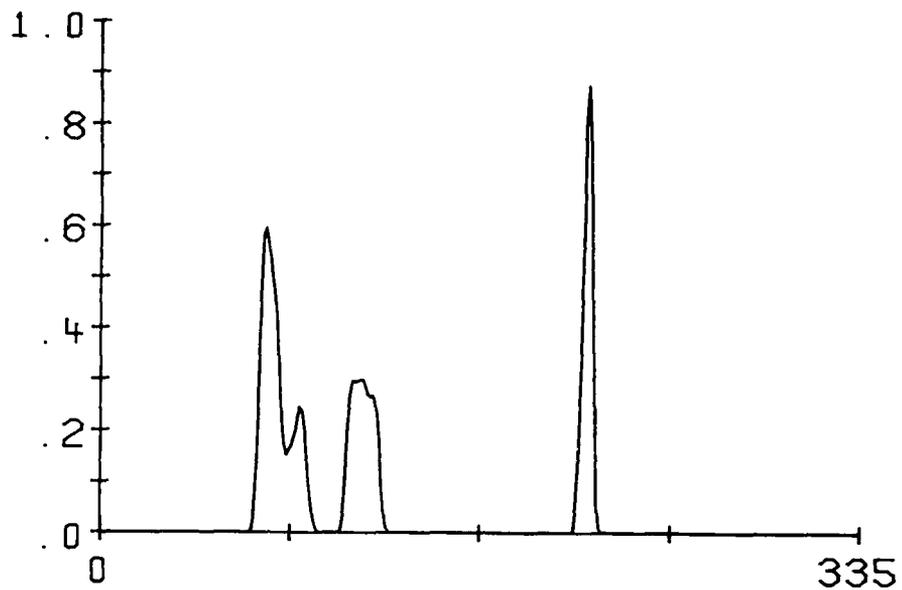


FIGURE 13 - VK-Product records of image ALA 31 8

confined to merely state that the number of targets is at least 3 (greatest number of peaks), and at the very most, 6 (product of the numbers of peaks). Let us look back at the above procedure. Skewness and kurtosis records are balanced to remove any bias that might be present in these records. This is bound to happen because BET acts only in one direction and then, as in a one-dimensional filtering operation, features in a perpendicular direction go unnoticed (a good example of this is Fig. 18C where a crevasse runs across the VFS image, with the result that the skewness is negatively biased).

To reduce VK-product records to numbers specifying the exact location of the targets, we proceed as follows:

- a) firstly, VK-product records are smoothed using a 3-point moving average;
- b) the coordinates of the highest peak are saved;
- c) a threshold is set at 10 percent of the maximum value:

$$q = \max(VK(j))/10 \quad [39]$$

and a new, binary, VK-product record is generated:

$$VK(j) = \begin{cases} 1 & \text{if } VK(j) \geq q \\ 0 & \text{if } VK(j) < q \end{cases} \quad [40]$$

- d) a gap-filling algorithm is then used to join the runs of 1s that are separated by less than three 0s;
- e) the runs of 1s of gap-filled, binary, VK-product records that consist of less than three 1s are discarded;
- f) for each run of 1s, we determine the coordinate of the leading and trailing 1 as well as the length of the run and its middle point.

The results for the VK-records of Fig. 13 are shown in Table I.

TABLE I

Target designation based on the VK-records: image ALA 31 8

Horizontally (Top Record)

Target Range	:	(69, 92)	;	(108, 124)	;	(211, 218)
Target Width	:	24	;	17	;	8
Target Midpoint	:	80	;	116	;	214

Vertically (Bottom Record)

Target Range	:	(206, 217)
Target Height	:	12
Target Midpoint	:	211

Highest Peak: (215, 212)

Figure 13 and Table I depict a clear-cut situation. Figure 14 and Table II, on the contrary, represent one of the worst case encountered. Since there is more than one peak in both directions, we do not know the exact number of target-like hot spots. However, we can ascertain that this number lies somewhere between 5 and 10. Faced with such an ambiguity, it is better to define a target area, that is, an area including all the target-like hot spots detected with the VK-product records. Such a target area can be delimited by using the 2 peaks farther apart in both directions (Fig. 15B), or else, to limit any further search, by using the target range in one direction as the width (height) of the target area, and hence define not one but several (Fig. 15C and 15D) target areas. In the same vein, one can use the coordinates of the highest peak to initiate a search, for experimental results show that these quite often correspond to the position of a real target.

7.4 Experimental Results

The ideas and techniques expounded in this section were extensively tested on a set of 43 thermoscopic images known as the Alabama Data Base. The spectral region of the majority (30 out of 43) of these images corresponds to the 8-14 μm band, and the remainder to the 3-5 μm band. Altogether the images contain 85 targets distributed as follows (detailed ground truth accompanies Fig. 16): 40 tanks, 29 APCs, 15 jeeps and, finally, a bus. The size of the images is 420 x 335 pixels and they are digitized to 256 levels.

PRODUCT VK(S,L) (IMAGE 41; F=0.05)

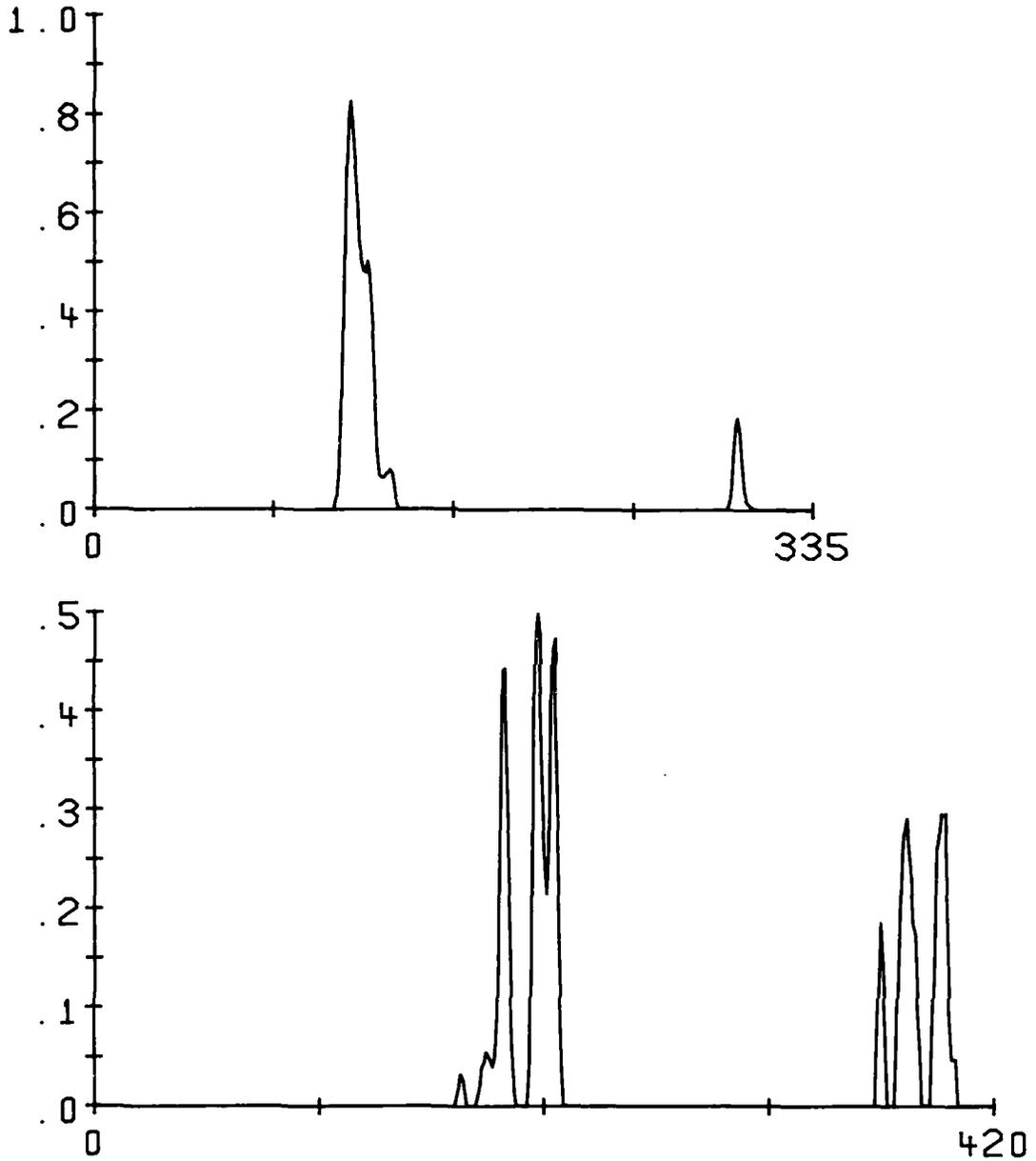


FIGURE 14 - VK-Product records of image ALA 41 3

The analysis of the fine structure of an image was primarily undertaken to improve the performance of the silhouette generator by restricting the search for targets to an area that would ideally be much smaller in extent than the image itself. This was also the purpose of GSA (Sect. 6) though it was discarded here for it turned out to be orientation-dependent. But, as Fig. 16 asserted it, such is not the case of FSA. Moreover, it so happens that in many instances the target area delimited by FSA is tiny enough as to allow to pinpoint individual targets (Table I and Figs. 16-31). This is very interesting since it means one can designate targets without segmenting the image, simply by statistical considerations. However, such pinpointing operations should probably be limited to applications involving one target at a time, although FSA manages well when confronted with several targets arranged in a line (Figs. 16-3,8,12,14 etc.). In this last case, however, a small target might well be obscured by a larger one next to it (Figs. 16-6,33). This phenomenon occurs in a direction parallel to the line formed by the targets, for in a perpendicular direction FSA obviously perceives only one target. Indeed, it can be said that FSA in a way senses the targets as if they were projected on both axes. So, when the projections along one axis partially overlap, the quantities measured (target midpoint; see Tables I and II) do not necessarily fit with all the targets involved. This explains why many crosses in Fig. 16 do not sit right on top a nearby target (Figs. 16-16,19,22,31,36). It is also for the same reason that a group of targets is interpreted as a single target (Fig. 16-2), and that a target is missed in some L-shaped formations (Figs. 16-17,35). On the other hand, whenever it is not possible to unambiguously pinpoint individual targets (Figs. 16-4,7,9,23), one can always define somehow (Fig. 15) a target area and

subject it to further processing, or possibly initiate a search by using the position of the highest peak (Tables I and II) as an initial guess, for experimental results (Figs. 16-28, 29,41) show, as mentioned before, that this peak quite often corresponds to a real target.

TABLE II

Target designation based on VK-records: image ALA 41 3

Horizontally (Top Record)

Target Range : (115, 132) ; (298, 302)
Target Width : 18 ; 5
Target Midpoint : 122 ; 300

Vertically (Bottom Record)

Target Range : (188,195) ; (204,217) ; (365,369) ; (375,385) ; (391,399)
Target Height : 8 ; 14 ; 5 ; 11 ; 9
Target Midpoint : 191 ; 210 ; 367 ; 380 ; 395

Highest Peak: (120, 207)

Target Area: (115, 302 ; 188, 399)

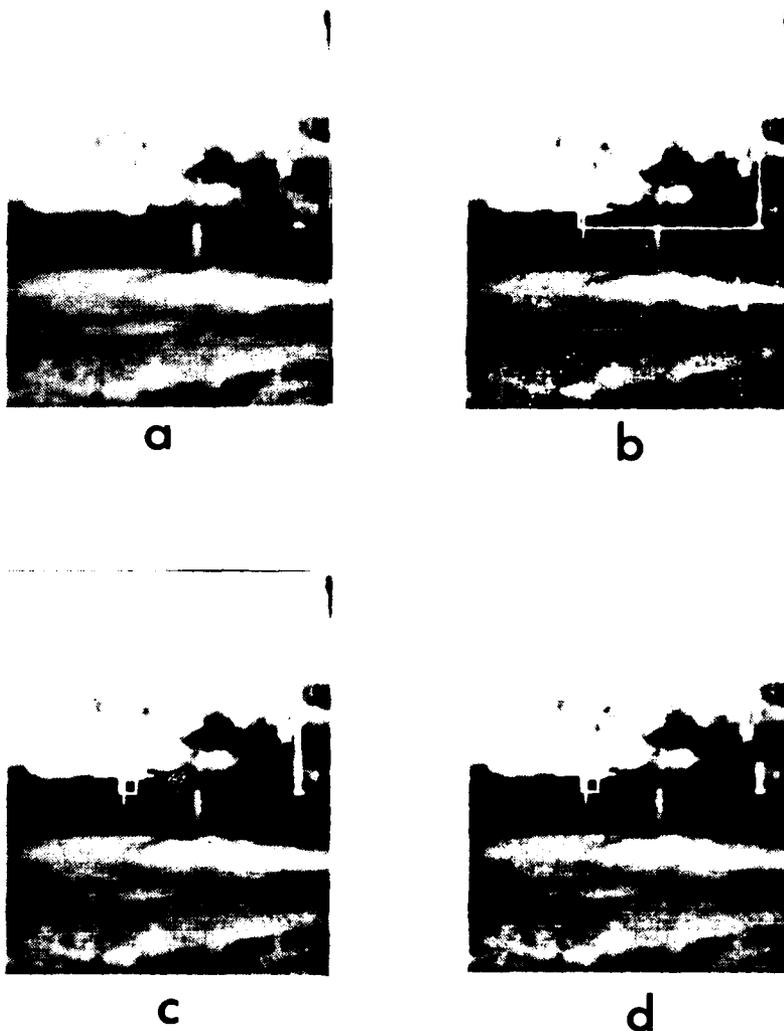


FIGURE 15 - Various ways to represent the data of Table II. The position of the highest peak is marked with a cross in b, c and d, while the square in b represents the target area defined by using the 2 peaks farther apart in both directions.

FIGURE 16 - All the images that constitute the Alabama Data Base have been histogram equalized (top row) and the results, (exemplified by Tables I and II) obtained by statistically analysing the fine structure images mapped into them (bottom row). We recall that the images are histogram equalized for display purpose only. The images that were actually processed are the original, unprocessed, raw images from the aforementioned data base. When one of the following images bears nothing but crosses, these designate the calculated midpoint of the detected targets. On the other hand, a cross that lies within a target area designates the position of the highest VK-product peak.

Alabama Data Base. Ground truth.

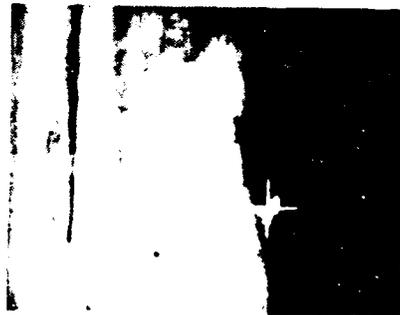
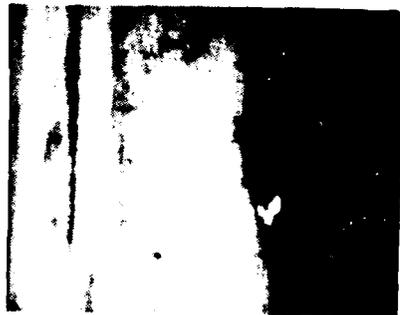
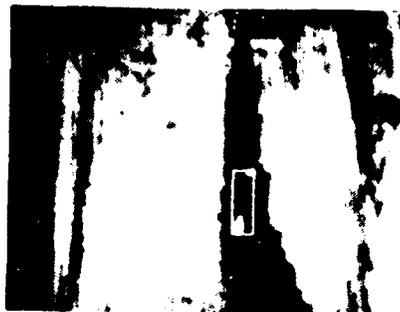
<u>Image No.</u>	<u>Target(s)</u>	<u>Aspect(s)</u>
1	T	S
2	J.A.T	3R.S.S
3	T.A	3F.3R
4	J.T.	S.S
5	T	3F
6	T.A.J	S.S.S
7	T	3R
8	T.A	F.S
9	J.T.A	S.S.S
10	T	3F
11	T	3R
12	T.A	3R.3F
13	J.A.T	S.F.F
14	T.A.J.	S.S.S
15	T	3F
16	A.T	S.S
17	A.J.T	S.3R.S
18	T	3R
19	T.A	3F.S
20	T.J	R.3R
21	A	R
22	T.A.J	3R.3F.S
23	T	S
24	T	F

UNCLASSIFIED
58

<u>Image No.</u>	<u>Target(s)</u>	<u>Aspect(s)</u>
25	T	S
26	T	3R
27	T.A	F.F
28	T.A	S.S
29	T.A	R.R
30	T.A.J	S.S.S
31	T.A.J.	S.S.S
32	T.A.J.	F.F.S
33	B.A	S.F
34	A.J	3F.S
35	J.T.A	S.S.S
36	T.A.J.	S.S.S
37	T.A	3R.3R
38	T.A	3R.3R
39	T.A	3R.3R
40	T	3R
41	T.A	3R.3R
42	T.A	3R.3F
43	T.A	3F.3R

Legend: A = APC, B = bus, J = jeep, T = tank,
S = side, F = front,
R = rear,
3 = 3/4 view

UNCLASSIFIED
59



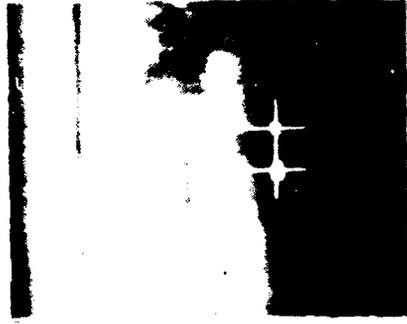
4

3

2

1

UNCLASSIFIED
60



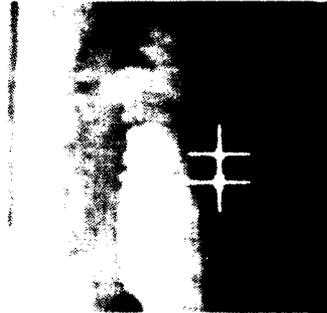
8

7

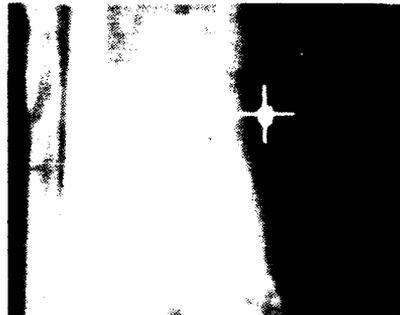
6

5

UNCLASSIFIED
61



12



11

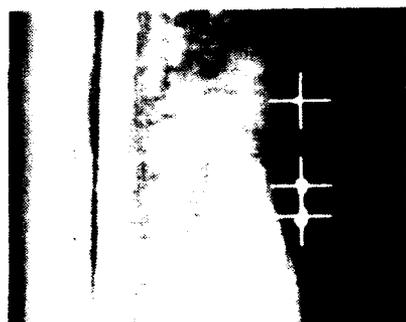
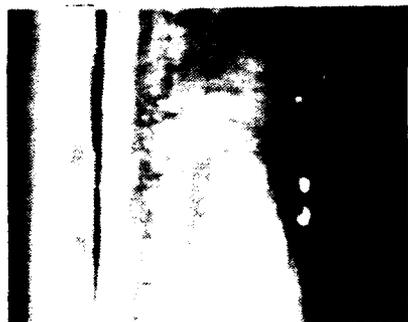
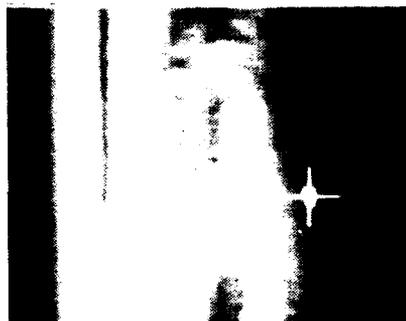
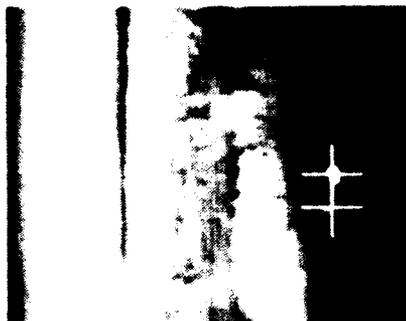
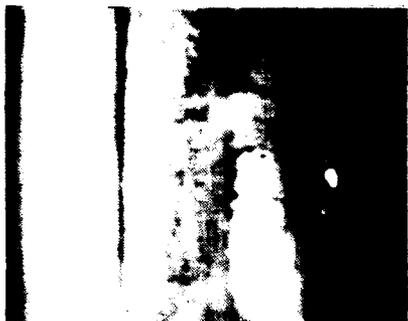


10



9

UNCLASSIFIED
62



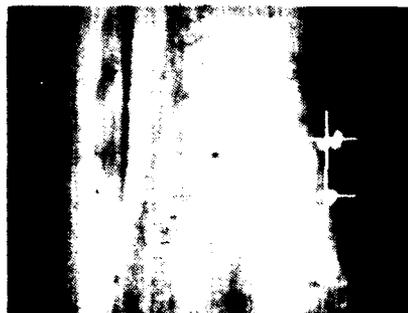
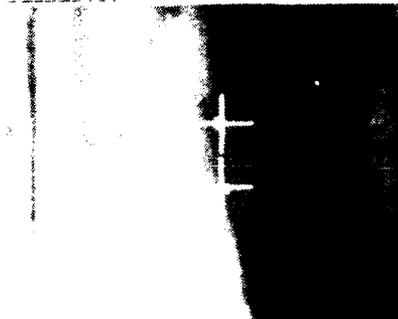
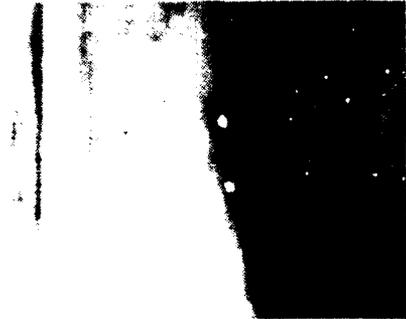
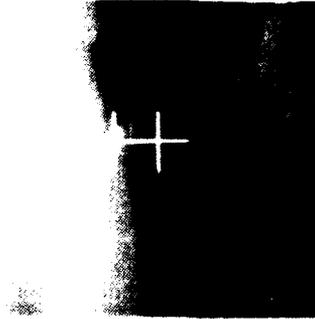
16

15

14

13

UNCLASSIFIED
63



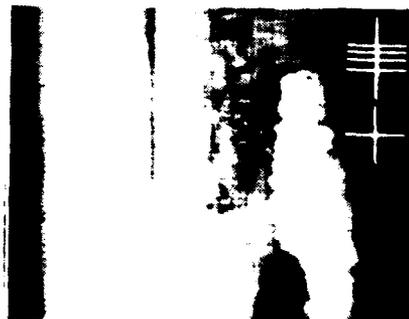
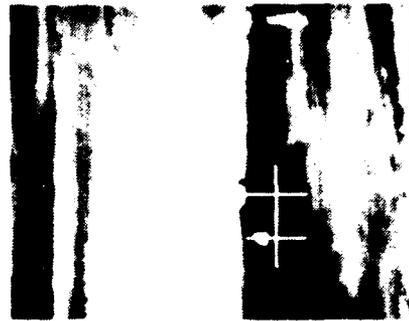
20

19

18

17

UNCLASSIFIED
64



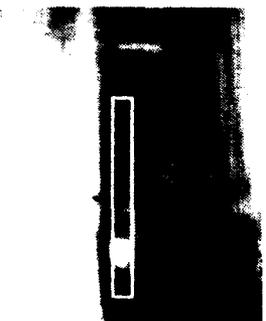
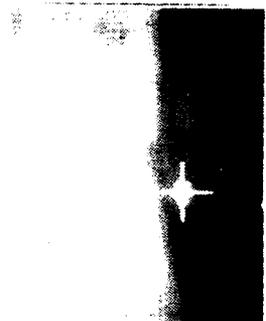
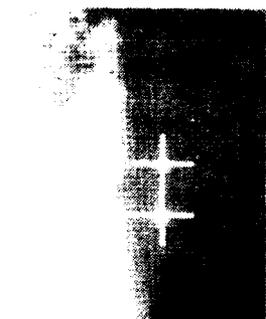
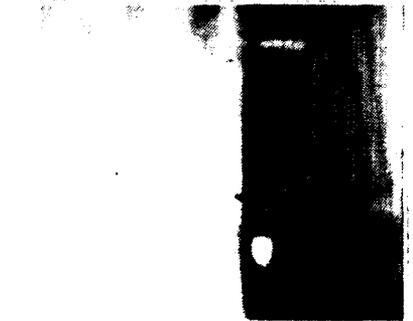
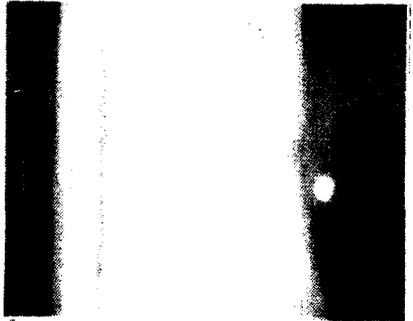
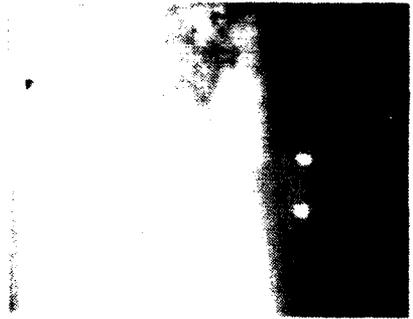
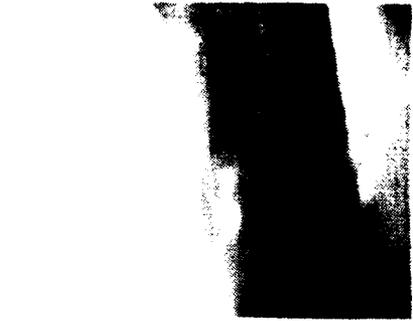
24

23

22

21

UNCLASSIFIED
65



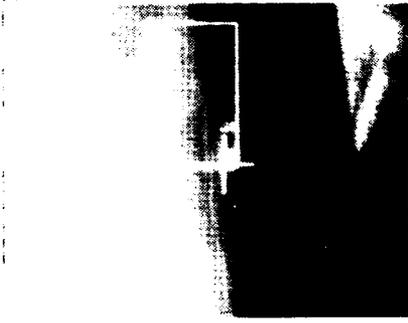
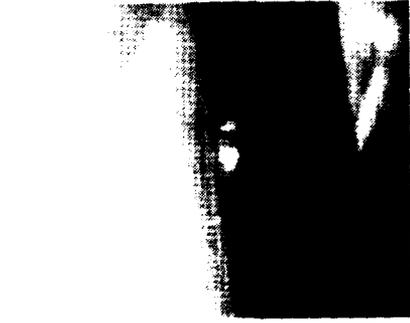
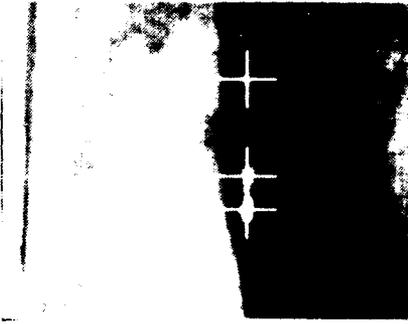
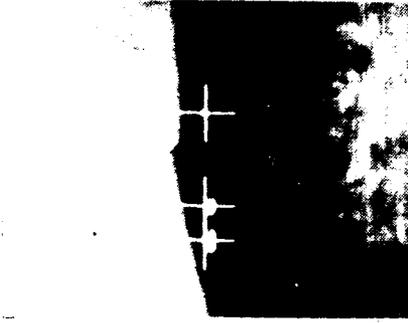
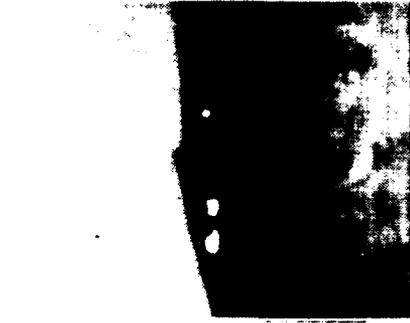
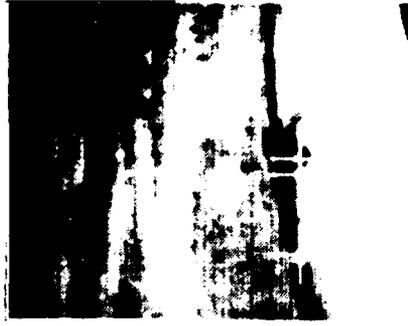
28

27

26

25

UNCLASSIFIED
66



32

31

30

29

UNCLASSIFIED
67



33



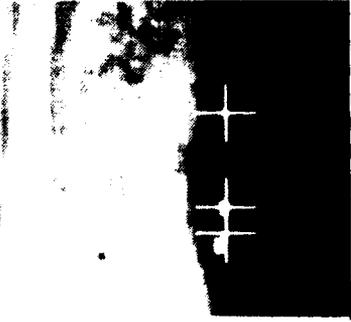
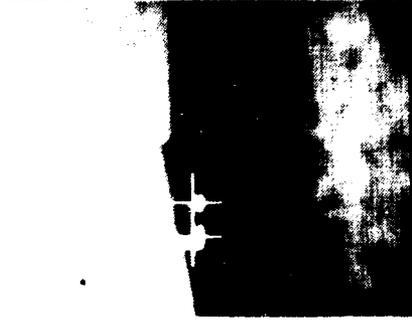
34



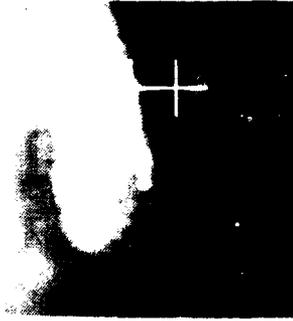
35



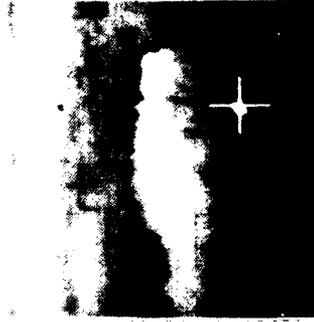
36



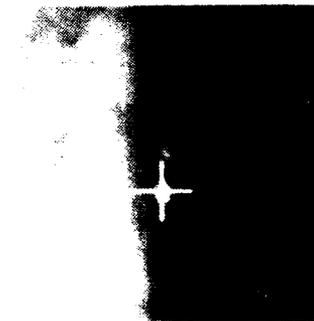
UNCLASSIFIED
68



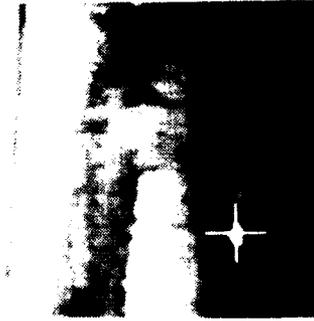
40



39

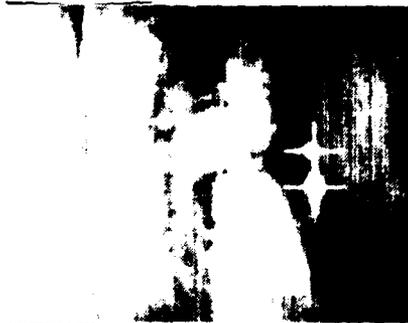


38



37

UNCLASSIFIED
69



43



42



41

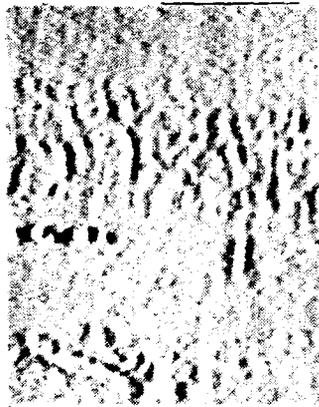
UNCLASSIFIED
70

ORIGINAL



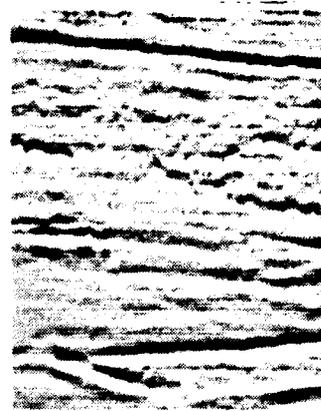
A)

HORIZONTAL



B)

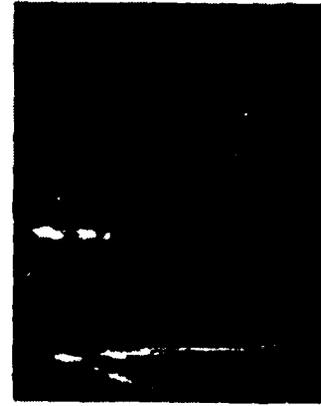
VERTICAL



C)



D)

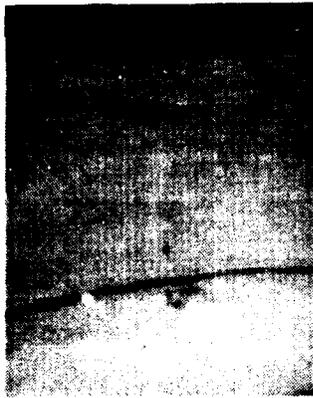


E)

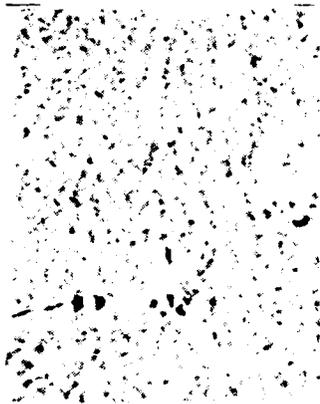
FIGURE 17 - Thresholding of the fine structure images derived from image 6 3 of the Alabama Data Base

8.0 THRESHOLDING OF FINE STRUCTURE IMAGES

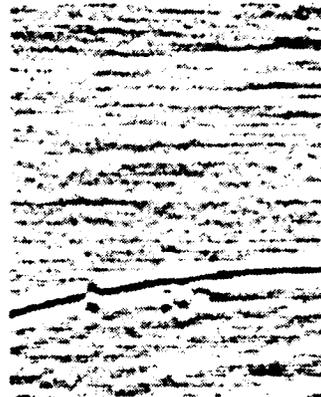
Once the background of an image has been subtracted by using BET, one has the choice between two alternatives: to process the resultant fine structure images along the lines set out in Sect. 7, or to threshold them by using the silhouette generator. In this last case, since the background of the fine structure images is "uniform", the SIT Generator should be well suited for this task. This affirmation is confirmed by the results of Figs. 17 and 18. Figures 17B and 18B show the HFS images derived respectively from images 6 3 and 13 8 of the Alabama Data Base, whereas Figs. 17C and 18C show the corresponding VFS images. These fine structure images were postprocessed, for display purpose, first by adding a constant bias, so as to remove negative gray levels, and then by stretching the gray levels bounded by the 5th and 95th percentiles linearly over the display range. The ideas alluded to in this section are fully developed in Ref. 14.



a



b



c



d



e

FIGURE 18 - Thresholding of the fine structure images derived from image 13 8 of the Alabama Data Base.

9.0 CONCLUSION

The present report laid the foundations of a class of segmentation algorithms (segmenters) for detection of targets in IR imagery. The single basic assumption is that the targets possess a larger thermal signature than other objects embedded into the background. This class of segmentation algorithms emerged as a result of efforts to improve an early segmenter devised to extract targets from IR BOFORS imagery. This first segmenter proceeds according to a single intensity threshold whence the name SIT Generator to designate it. Its extraction record is generally excellent whenever the background, on a large-scale basis, is relatively uniform. When this condition is not met, one can use a thresholding intensity function in lieu of a fixed threshold. The SIT Generator and its variants try to cope with the background simply by partitioning the image. A more promising avenue consists in levelling the background so as to curb its ascendancy over the image. This is in essence what the Background Elimination Technique (BET) expounded in Sect. 7 does. Since BET can be applied either to the set of lines or columns of an image, it generates 2 images referred to as the Horizontal Fine Structure (HFS) image and the Vertical Fine Structure (VFS) image respectively. We have shown that one can pinpoint targets merely by statistically analyzing these fine structure images, without having really to segment them. Nevertheless, this last approach seems very promising and we intend to fully exploit it.

UNCLASSIFIED

74

10.0 ACKNOWLEDGEMENTS

The author thanks Mr G. McLaughlin for helpful discussions in relation to the estimation of some statistical parameters.

11.0 REFERENCES

1. Charles S. Weaver et al., "Digital Filtering with Applications to Electrocardiogram Processing", IEEE Audio Vol. AU-16, No. 3, September 1968, p. 350.
2. B. Gold and C. M. Rader, "Digital Processing of Signals", McGraw-Hill, 1969.
3. A. Papoulis, "The Fourier Integral and its Applications", McGraw-Hill, 1962.
4. J. S. Bendat and A. G. Piersol, "Measurement and Analysis of Random Data", Wiley, New York
5. M. Abramowitz and I. A. Stegun, "Handbook of Mathematical Functions", Dover, 1965.
6. B. S. Chissom, "Interpretation of the Kurtosis Statistic", The American Statistician, October 1970.
7. R. B. Darlington, "Is Kurtosis Really Peakedness?", The American Statistician, April 1970, 24, (2), pp. 19-22.
8. D. K. Hildebrand, "Kurtosis Measures Bimodality?", The American Statistician, February 1971.
9. R. V. Hogg, "More Light on the Kurtosis and Related Statistics", Journal of the American Statistical Association, June 1972, Volume 67, Number 311, Theory and Methods Section.
10. "Biometrika Tables for Statisticians", Vol. 1. Edited by E.S. Pearson and H.O. Hartley, Cambridge University Press, 1956
11. L. Sévigny, "Simulation d'un système d'acquisition automatique d'objectif infrarouge dans un contexte sol-sol", DREV Report 4081/77, June 1977, UNCLASSIFIED.
12. L. Sévigny, "La reconnaissance de forme et l'acquisition d'objectif en infrarouge: nouvel algorithme de détection", DREV Report 4099/78, March 1978, UNCLASSIFIED.
13. L. Sévigny, "Extracteurs séquentiels pour l'acquisition de cibles sur images", DREV Report 4153/79, August 1979, UNCLASSIFIED.
14. L. Sévigny, Evaluation of a class of segmenters for IR Imagery", DREV Report 4172/80, May 1980, UNCLASSIFIED.

CRDV R-4180/80 (NDN CLASSIFIE)

Bureau - Recherche et développement, MDN, Canada.
CRDV, C.P. 880, Courcellette, Qué. G0A 1R0

"Algorithmes de segmentation pour la détection de cibles sur images IR"
par L. Sévigny

Ce rapport présente une classe d'algorithmes de segmentation (segmenteurs) spécialisés dans la détection de cibles sur images IR et fondés uniquement sur le principe voulant que la signature thermique d'une cible dépasse en importance celle de tout objet de l'arrière-plan. Ces algorithmes sont le résultat d'efforts visant à améliorer un premier segmenteur, appelé générateur de silhouettes, imaginé en fonction d'images IR du type BQORS. Le segmenteur en question découpe l'image en deux parties d'après un seuil d'intensité unique. Sa fiche d'extraction est généralement excellente lorsque l'arrière-plan est globalement plat. Lorsque cette condition n'est pas remplie, on peut parfois se tirer d'affaire en utilisant une fonction seuil au lieu d'un seuil fixe. Le générateur de silhouettes et ses diverses variantes tentent de venir à bout de l'arrière-plan simplement en morcelant l'image. Une solution plus prometteuse consiste à redresser l'arrière-plan de façon à réduire son emprise sur l'image. C'est précisément ce que la technique de Redressement de l'Arrière-Plan (IRAP) fait. Étant donné que IRAP s'applique aux lignes aussi bien qu'aux colonnes d'une image, il en résulte 2 images distinctes: structure fine horizontale et structure fine verticale. Ces images reçoivent maintes possibilités quant à la détection de cibles, lesquelles sont en grande partie explicitées dans le rapport. (NC)

CRDV R-4180/80 (NDN CLASSIFIE)

Bureau - Recherche et développement, MDN, Canada.
CRDV, C.P. 880, Courcellette, Qué. G0A 1R0

"Algorithmes de segmentation pour la détection de cibles sur images IR"
par L. Sévigny

Ce rapport présente une classe d'algorithmes de segmentation (segmenteurs) spécialisés dans la détection de cibles sur images IR et fondés uniquement sur le principe voulant que la signature thermique d'une cible dépasse en importance celle de tout objet de l'arrière-plan. Ces algorithmes sont le résultat d'efforts visant à améliorer un premier segmenteur, appelé générateur de silhouettes, imaginé en fonction d'images IR du type BQORS. Le segmenteur en question découpe l'image en deux parties d'après un seuil d'intensité unique. Sa fiche d'extraction est généralement excellente lorsque l'arrière-plan est globalement plat. Lorsque cette condition n'est pas remplie, on peut parfois se tirer d'affaire en utilisant une fonction seuil au lieu d'un seuil fixe. Le générateur de silhouettes et ses diverses variantes tentent de venir à bout de l'arrière-plan simplement en morcelant l'image. Une solution plus prometteuse consiste à redresser l'arrière-plan de façon à réduire son emprise sur l'image. C'est précisément ce que la technique de Redressement de l'Arrière-Plan (IRAP) fait. Étant donné que IRAP s'applique aux lignes aussi bien qu'aux colonnes d'une image, il en résulte 2 images distinctes: structure fine horizontale et structure fine verticale. Ces images reçoivent maintes possibilités quant à la détection de cibles, lesquelles sont en grande partie explicitées dans le rapport. (NC)

CRDV R-4180/80 (NDN CLASSIFIE)

Bureau - Recherche et développement, MDN, Canada.
CRDV, C.P. 880, Courcellette, Qué. G0A 1R0

"Algorithmes de segmentation pour la détection de cibles sur images IR"
par L. Sévigny

Ce rapport présente une classe d'algorithmes de segmentation (segmenteurs) spécialisés dans la détection de cibles sur images IR et fondés uniquement sur le principe voulant que la signature thermique d'une cible dépasse en importance celle de tout objet de l'arrière-plan. Ces algorithmes sont le résultat d'efforts visant à améliorer un premier segmenteur, appelé générateur de silhouettes, imaginé en fonction d'images IR du type BQORS. Le segmenteur en question découpe l'image en deux parties d'après un seuil d'intensité unique. Sa fiche d'extraction est généralement excellente lorsque l'arrière-plan est globalement plat. Lorsque cette condition n'est pas remplie, on peut parfois se tirer d'affaire en utilisant une fonction seuil au lieu d'un seuil fixe. Le générateur de silhouettes et ses diverses variantes tentent de venir à bout de l'arrière-plan simplement en morcelant l'image. Une solution plus prometteuse consiste à redresser l'arrière-plan de façon à réduire son emprise sur l'image. C'est précisément ce que la technique de Redressement de l'Arrière-Plan (IRAP) fait. Étant donné que IRAP s'applique aux lignes aussi bien qu'aux colonnes d'une image, il en résulte 2 images distinctes: structure fine horizontale et structure fine verticale. Ces images reçoivent maintes possibilités quant à la détection de cibles, lesquelles sont en grande partie explicitées dans le rapport. (NC)

CRDV R-4180/80 (NDN CLASSIFIE)

Bureau - Recherche et développement, MDN, Canada.
CRDV, C.P. 880, Courcellette, Qué. G0A 1R0

"Algorithmes de segmentation pour la détection de cibles sur images IR"
par L. Sévigny

Ce rapport présente une classe d'algorithmes de segmentation (segmenteurs) spécialisés dans la détection de cibles sur images IR et fondés uniquement sur le principe voulant que la signature thermique d'une cible dépasse en importance celle de tout objet de l'arrière-plan. Ces algorithmes sont le résultat d'efforts visant à améliorer un premier segmenteur, appelé générateur de silhouettes, imaginé en fonction d'images IR du type BQORS. Le segmenteur en question découpe l'image en deux parties d'après un seuil d'intensité unique. Sa fiche d'extraction est généralement excellente lorsque l'arrière-plan est globalement plat. Lorsque cette condition n'est pas remplie, on peut parfois se tirer d'affaire en utilisant une fonction seuil au lieu d'un seuil fixe. Le générateur de silhouettes et ses diverses variantes tentent de venir à bout de l'arrière-plan simplement en morcelant l'image. Une solution plus prometteuse consiste à redresser l'arrière-plan de façon à réduire son emprise sur l'image. C'est précisément ce que la technique de Redressement de l'Arrière-Plan (IRAP) fait. Étant donné que IRAP s'applique aux lignes aussi bien qu'aux colonnes d'une image, il en résulte 2 images distinctes: structure fine horizontale et structure fine verticale. Ces images reçoivent maintes possibilités quant à la détection de cibles, lesquelles sont en grande partie explicitées dans le rapport. (NC)

DREV R-4180/80 (UNCLASSIFIED)

Research and Development Branch, DND, Canada,
DREV, P.O. Box 880, Courcellette, Que. G0A 1R0

"Segmentation Algorithms for Detection of Targets in IR Imagery"
by L. Sévigny

This report presents a class of segmentation algorithms (segmenters) for detection of targets in IR imagery based on the single assumption that the targets possess a larger thermal signature than the background. This class of algorithms emerged as a result of efforts to improve an early segmenter devised to extract targets from IR 80FDRS imagery. This segmenter proceeds according to a Single Intensity Threshold whence the name SII Generator to designate it. The extraction record of the SII Generator is generally excellent whenever the background, on a large-scale basis, is relatively uniform. When this condition is not met, one can use a thresholding intensity function in lieu of a fixed threshold. The SII Generator and its variants try to cope with the background simply by partitioning the image. A more promising avenue consists in levelling the background so as to curb its ascendancy over the image. This is in essence what the Background Elimination technique (BEI) expounded herein does. Since BEI can be applied either to the set of lines or columns of an image, it generates 2 images referred to as the Horizontal Fine Structure image and the Vertical Fine Structure image respectively. These images offer many possibilities for detection of targets and several of them are explicitly described in the report. (U)

DREV R-4180/80 (UNCLASSIFIED)

Research and Development Branch, DND, Canada,
DREV, P.O. Box 880, Courcellette, Que. G0A 1R0

"Segmentation Algorithms for Detection of Targets in IR Imagery"
by L. Sévigny

This report presents a class of segmentation algorithms (segmenters) for detection of targets in IR imagery based on the single assumption that the targets possess a larger thermal signature than the background. This class of algorithms emerged as a result of efforts to improve an early segmenter devised to extract targets from IR 80FDRS imagery. This segmenter proceeds according to a Single Intensity Threshold whence the name SII Generator to designate it. The extraction record of the SII Generator is generally excellent whenever the background, on a large-scale basis, is relatively uniform. When this condition is not met, one can use a thresholding intensity function in lieu of a fixed threshold. The SII Generator and its variants try to cope with the background simply by partitioning the image. A more promising avenue consists in levelling the background so as to curb its ascendancy over the image. This is in essence what the Background Elimination technique (BEI) expounded herein does. Since BEI can be applied either to the set of lines or columns of an image, it generates 2 images referred to as the Horizontal Fine Structure image and the Vertical Fine Structure image respectively. These images offer many possibilities for detection of targets and several of them are explicitly described in the report. (U)

DREV R-4180/80 (UNCLASSIFIED)

Research and Development Branch, DND, Canada,
DREV, P.O. Box 880, Courcellette, Que. G0A 1R0

"Segmentation Algorithms for Detection of Targets in IR Imagery"
by L. Sévigny

This report presents a class of segmentation algorithms (segmenters) for detection of targets in IR imagery based on the single assumption that the targets possess a larger thermal signature than the background. This class of algorithms emerged as a result of efforts to improve an early segmenter devised to extract targets from IR 80FDRS imagery. This segmenter proceeds according to a Single Intensity Threshold whence the name SII Generator to designate it. The extraction record of the SII Generator is generally excellent whenever the background, on a large-scale basis, is relatively uniform. When this condition is not met, one can use a thresholding intensity function in lieu of a fixed threshold. The SII Generator and its variants try to cope with the background simply by partitioning the image. A more promising avenue consists in levelling the background so as to curb its ascendancy over the image. This is in essence what the Background Elimination technique (BEI) expounded herein does. Since BEI can be applied either to the set of lines or columns of an image, it generates 2 images referred to as the Horizontal Fine Structure image and the Vertical Fine Structure image respectively. These images offer many possibilities for detection of targets and several of them are explicitly described in the report. (U)

DREV R-4180/80 (UNCLASSIFIED)

Research and Development Branch, DND, Canada,
DREV, P.O. Box 880, Courcellette, Que. G0A 1R0

"Segmentation Algorithms for Detection of Targets in IR Imagery"
by L. Sévigny

This report presents a class of segmentation algorithms (segmenters) for detection of targets in IR imagery based on the single assumption that the targets possess a larger thermal signature than the background. This class of algorithms emerged as a result of efforts to improve an early segmenter devised to extract targets from IR 80FDRS imagery. This segmenter proceeds according to a Single Intensity Threshold whence the name SII Generator to designate it. The extraction record of the SII Generator is generally excellent whenever the background, on a large-scale basis, is relatively uniform. When this condition is not met, one can use a thresholding intensity function in lieu of a fixed threshold. The SII Generator and its variants try to cope with the background simply by partitioning the image. A more promising avenue consists in levelling the background so as to curb its ascendancy over the image. This is in essence what the Background Elimination technique (BEI) expounded herein does. Since BEI can be applied either to the set of lines or columns of an image, it generates 2 images referred to as the Horizontal Fine Structure image and the Vertical Fine Structure image respectively. These images offer many possibilities for detection of targets and several of them are explicitly described in the report. (U)