Target detection in multispectral images using the spectral co-occurrence matrix and entropy thresholding

Mark L. G. Althouse, MEMBER SPIE
U.S. Army Edgewood Research
   Development and Engineering Center
SCBRD-RTE, Bldg. E5554
Aberdeen Proving Ground, Maryland 21010
E-mail: mlalthou@cbda8.apgea.army.mil

Chein-I. Chang, MEMBER SPIE
University of Maryland
Electrical Engineering Department
Baltimore, Maryland 21228

Abstract. Relative entropy thresholding techniques have been used for segmentation of objects from background in gray-level images. These techniques are related to entropy-based segmentations computed for the statistics of a spatial co-occurrence matrix. For detection of spectrally active targets such as chemical vapor clouds in multispectral or hyperspectral imagery, a spectral co-occurrence matrix is employed. Using the entropy of various regions of the matrix, thresholds can be derived that will segment an image family based on the spectral characteristics of the intended target. Experiments are presented that show the detection of a chemical vapor cloud in multispectral thermal imagery. Several manners of dividing the co-occurrence matrix into regions are explored. Thresholds are determined on both a local and global basis and compared. Locally generated thresholds are treated as a distribution and separated into classes. The point of class separation is used as a global threshold with improved results.

Subject terms: infrared; multispectral; detection; vapor; co-occurrence; chemical; entropy; segmentation.

Optical Engineering 34(7), 2135–2148 (July 1995).

1 Introduction

Considerable use is being made of multispectral and hyperspectral data for many applications. As imaging spectrometer sensors become more available and collect larger image cubes, the task of processing the data can become time consuming and complex. Techniques to segment targets from background based on the relative entropy of the probability distribution of a spatial co-occurrence matrix of the original image versus the segmented binary image were introduced in Ref. 1. The techniques are based on the principles of information theory and follow from entropy thresholding work in Ref. 2. The multispectral data used here are collected by an imaging system designed to detect and map multiple organic vapor compounds. Normally, the target vapor will have one or more spectral features appearing in one or more of the imager's spectral bands. Likelihood ratio testing methods of detection have been developed for this type of data and have worked well,[3] but require considerable computing power and are based on assumptions that may not always hold. The search for less complex although possibly less robust but operationally functional methods has led to this work.

Because their unique character is spectral rather than spatial, some spectral resolution in the data is crucial to detecting chemical vapors. Because organic chemical vapors have a very precise spectral footprint in the IR, vapors and airborne interferants can be differentiated. Additionally this spectral characteristic can be used to enhance the contrast of the vapor component of an image relative to the usually spectrally flat background. In the case of the very weak signal, unique spectral features can be used to design a noise-suppressing filter. Here, we extend the use of the co-occurrence matrix, but rather than the spatial transitions previously used,[1] spectral transitions are employed. The temporal co-occurrence matrix, also based on two images, has been used as a description of a moving target.[4-7] Although the temporal transitions are not investigated here, note that the techniques generated for spectral transitions can apply to temporal transitions as well.

2 Spectral Co-occurrence Matrix

In as far as can be found, the spectral co-occurrence matrix is a new representation. A similar representation, the spectral scatter plot[8] has been used, but not in a manner like the co-occurrence matrix is employed here. In place of the single gray-scale image $I(x,y)$, which was segmented to separate an object from the background,[1] we now have a family of gray-scale images of the same scene $I(x,y,\lambda)$. These images are

pixel registered, that is, each pixel (x,y) has the same field of view (FOV) for all values of wavelength λ. The wavelength λ represents the center value of the optical bandwidth of the image. To be of most use, the spectral bands should be independent, that is, without overlap in the optical bandwidths. Independence implies no spectral information will be redundant between bands. Spatial information, on the other hand, is highly correlated between spectral bands for all features except the chemical vapors. We use the spatial correlation to our great advantage.

Given a family of digital images I of spatial dimension $M \times N$ with L gray levels $G = \{1, 2, \ldots, L\}$, and with spectral dimension Λ, the gray level of spatial location (x,y) and spectral band λ is denoted by $I(x,y,\lambda) \in G$. Thus the family of images may be represented by the 3-D array or image cube $I = [I(x,y,\lambda)]_{M \times N \times \Lambda}$. The spectral co-occurrence of I is an $L \times L$ matrix $W = [f_{ij}]_{L \times L}$, which contains the frequency of transitions from one gray level to another associated with the pixel (x,y) in two spectral bands λ_a and λ_b. For a pair of gray levels i and j, the (i,j)'th entry of the co-occurrence matrix W, f_{ij}, is defined as follows

$$f_{ij} = \sum_{m=1}^{M} \sum_{n=1}^{N} \delta(m,n) \ , \tag{1}$$

where

$$\delta(m,n) = 1, \quad \text{if } I(m,n,\lambda_a) = i \text{ and } I(m,n,\lambda_b) = j$$

$$\delta(m,n) = 0, \quad \text{otherwise} \ . \tag{2}$$

To obtain the transition probabilities, one normalizes by the total number of transitions in the co-occurrence matrix,

$$p(i,j) = \frac{f_{i,j}}{\sum_{i=1}^{L} \sum_{j=1}^{L} f_{i,j}} \ . \tag{3}$$

Although the spectral co-occurrence matrix (shown in Fig. 1) looks very much like the spatial co-occurrence matrix, it cannot be divided into quadrants by a gray-level threshold as in the spatial case[1] because it now comprises components from two separate images. Performing a gray-level threshold on either or both of the input images based on the spectral co-occurrence matrix would be inappropriate because the spectral information is only partially contained in a single gray-level image. Rather, we must look at the diagonal versus the nondiagonal elements of the spectral co-occurrence matrix. The diagonal, where $i = j$, contains those pixels that have a constant gray level in both images. Pixels on the diagonal are considered spectrally "flat," that is, devoid of spectral features. If the correct pair of spectral bands has been chosen, the target cloud will have significantly higher or lower gray levels in one of the images. Target pixels will lie at some distance from the diagonal, with the distance off diagonal equal to the gray-level differences. We then choose a threshold that separates the diagonal from the nondiagonal elements of the co-occurrence matrix and generate a binary image based on these two classes of pixels.

Figure 2 illustrates this concept. The co-occurrence matrix is shown, with the main diagonal as a dotted line. A single
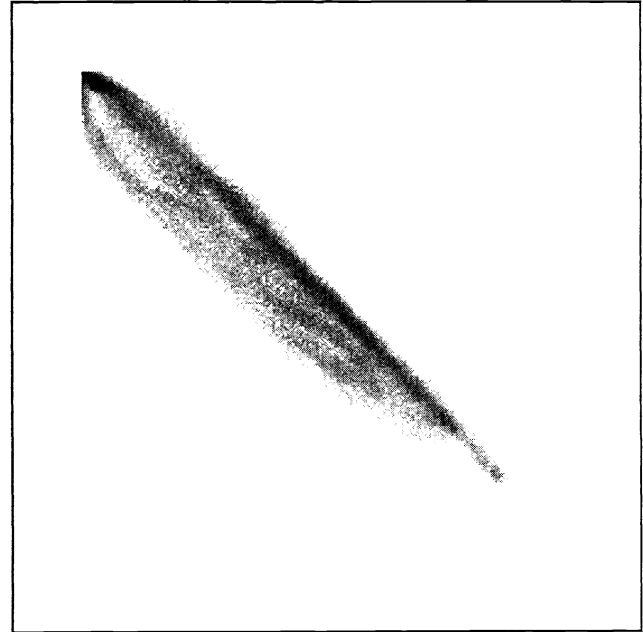

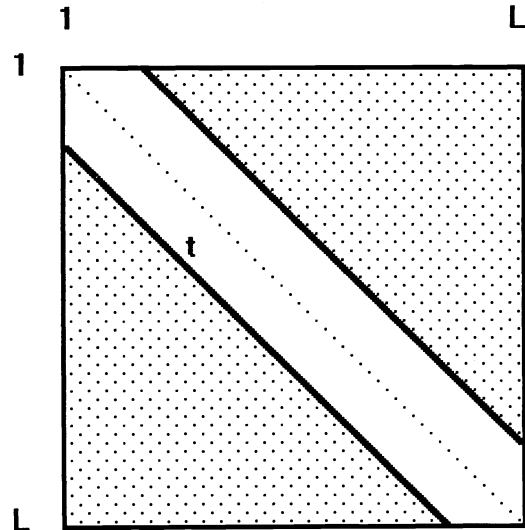**Fig. 1** Spectral co-occurrence matrix.

Fig. 2 Segmentation of a spectral co-occurrence matrix showing main diagonal (dotted line), threshold lines, background region (clear area), and target region (shaded area).

threshold lies parallel to the main diagonal both above and below it at a selected distance from the main diagonal. Thus the co-occurrence matrix is separated into two regions, the diagonal region and the shaded nondiagonal regions. The two nondiagonal regions, while both target or object regions, are treated as one for threshold computation, but separately considered because they represent pixels with either a dark-to-light or light-to-dark transition. Each of these transition types occupies a different shaded triangular region. With this in mind, each of the two nondiagonal regions is viewed as a separate target or object region and a binary image formed from each.

To proceed further with the threshold selection, we need the probabilities associated with the diagonal (D) and non-diagonal (ND) cells. They are defined by

$$P'_D = \sum_{j=1}^{L} \sum_{i=j-t+1}^{j+t-1} \frac{f_{ij}}{\sum_{j=1}^{L}\sum_{i=1}^{L} f_{ij}} = \sum_{j=1}^{L} \sum_{i=j-t+1}^{j+t-1} p(i,j),$$

$$P'_{ND} = \sum_{j=1}^{L} \sum_{i=1}^{j-t} p(i,j) + \sum_{j=1}^{L} \sum_{i=j+t}^{L} p(i,j) . \qquad (4)$$

Within the cell, the individual element probabilities are then

$$p'_D(i,j) = \frac{p(i,j)}{P'_D} ,$$

$$p'_{ND}(i,j) = \frac{p(i,j)}{P'_{ND}} . \qquad (5)$$

Based on these spectral transition probabilities, entropy methods can be used to find the threshold t.

Shown in Fig. 3 are the six frames taken for one scene by a modified FSI 2000 thermal imager. The instrument and some testing of it are described in Ref. 9. There is quite a range of image quality in these frames. Table 1 lists the spectral band center frequencies and bandwidths. The scene is looking across a field, with a tree about 100 yd distant in the center of the frame. The tree is clearly visible in bands 1, 2, and 5 and somewhat visible in band 6. The band 3 image contains mostly vapor, but the top of the tree is visible. Band 4 has no identifiable features; this is caused by a severely degraded optical filter. It is included as representative of some bands of multispectral images that have substantial atmospheric absorption. Given that the frames are fairly well matched or normalized for background scene brightness, the one or ones containing the cloud should have some differences in the pixel gray levels that exhibit a spatial character. In practice, the differences are very subtle and not visible to the operator in the scene unless the vapor concentration is very high.

3 Entropy-Based Thresholding

Entropy is a measure of the information contained in a probability distribution. If a particular event γ happens with prob-
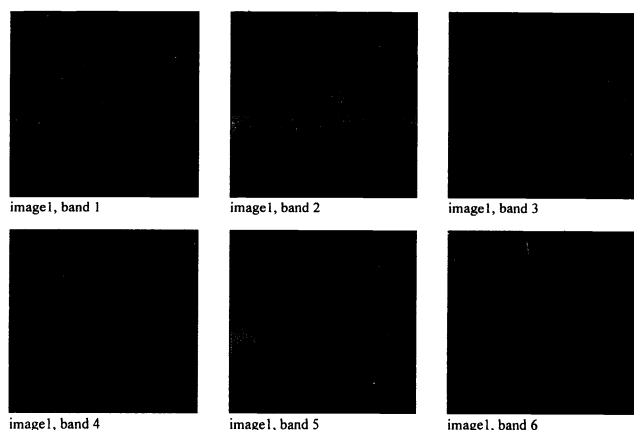
Table 1 Spectral frequency band characteristics.

Band	Center Frequency (μm)	Range
1	8.7	8.5-8.9
2	9.8	9.6-10.0
3	10.6	10.45-10.75
4	11.6	11.45-11.7
5	open	7.6-12.0
6	7.8	7.6-8.1

ability $P(\gamma)$, then the information derived from the occurrence of γ is

$$I(\gamma) = \log\frac{1}{P(\gamma)} , \qquad (6)$$

and $I(\gamma)$ is defined as the self-information of the event γ. Entropy is the average self-information of all the individual outcomes in a probability distribution. Suppose that the random variable Γ can take any of the values $\{\gamma_1, \gamma_2, ..., \gamma_{n-1}, \gamma_n\}$, then the probability that Γ takes the value γ_i is

$$p_\Gamma(\gamma_i) = P(\Gamma = \gamma_i) , \qquad (7)$$

such that

$$\sum_{i=1}^{n} p_\Gamma(\gamma_i) = 1 , \qquad (8)$$

and p_Γ is the discrete probability distribution of Γ. The entropy of Γ, $H(\Gamma)$, is defined by

$$H(\Gamma) = \sum_{i=1}^{n} p_\Gamma(\gamma_i) \log\frac{1}{p_\Gamma(\gamma_i)} . \qquad (9)$$

Entropy is the average self-information of the elements of a probability distribution. It is the average information or uncertainty of a distribution.

Several methods of dividing the co-occurrence matrix by diagonal thresholds were examined. They produced two regions (2R), four regions (4R), and three regions (3R) within the matrix. The 2R method is as shown in Fig. 2, with a single threshold value applied both above and below the main diagonal producing D and ND regions. Each portion (upper and lower) of the ND region is then used to create a binary image. Only one image of the pair will contain the target. The two nondiagonal regions are separate areas and are, in fact, representative of different phenomena. That is, the presence of a lighter than background object in the lower nondiagonal region and the presence of a darker than background object in the upper nondiagonal region. It would make intuitive sense to treat the two regions independently for the calculation of the thresholds. If the two images used for the spectral co-occurrence matrix had no differing features, all of the elements of the co-occurrence matrix would lie on the main diagonal. We assume that noise within the image sequence is responsible for the width of the pixel mapping along the diagonal. This noise is a combination of instrument noise, which is generally quite small, and fluctuation within the viewed scene resulting from the changing atmospheric con-

image1, band 1	image1, band 2	image1, band 3
image1, band 4	image1, band 5	image1, band 6

Fig. 3 Image set with six spectral bands (image 1).

ditions. The noise that would affect the data used here is random over the relatively short data collection time. A collection time of many minutes would be needed to see a permanent shift in the thermal scene caused by a change in ambient conditions. Thus, we expect the elements of the co-occurrence matrix on and in close proximity to the main diagonal to be evenly distributed above and below the diagonal. Based on this reasoning, we can split the co-occurrence matrix into two triangular regions, each including the main diagonal, and determine a threshold for each triangular region independently. This arrangement is illustrated in Fig. 4. In this manner, the thresholds that will segment the lighter or darker objects will not be influenced and biased by the other nondiagonal region. This calculation will now entail four regions (4R); they are the upper diagonal (UD), the lower diagonal (LD), and upper nondiagonal (UND), and the lower nondiagonal (LND). A threshold is calculated independently for the upper and lower D and ND region pairs. Having separate thresholds will be more important for the case of multiple targets of different types, such as a vapor cloud in an image that also has a spectrally active solid target. Another example is the case where the cloud spans the horizon; the cloud is warmer than the sky background, but cooler than the terrain background. Each nondiagonal region would contain part of the vapor cloud image and the binaries could be fused to obtain the complete target image.

A third arrangement of the co-occurrence matrix would again have two thresholds, but not divide the diagonal region into two portions for the calculation. Figure 5 shows the three regions (3R) that this approach would produce. The threshold $t1$ in the figure is labeled t_U in the equations and $t2$ corresponds to t_L. There is some intuitive superiority to this arrangement because we do not need the assumption of an even distribution of background pixels about the diagonal, and the full background class is used to calculate each threshold. In practice, if the filters used are not perfectly matched, there will be some (hopefully linear) shift in the population of background pixels to a position above or below the main diagonal. Thus an arbitrary division of the background class, such as is done in 4R, could yield erroneous results.

The mathematical approach followed here is similar to that used in Ref. 1. Although the 3R method is detailed in
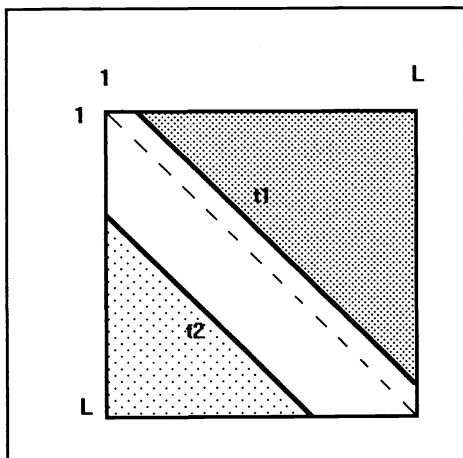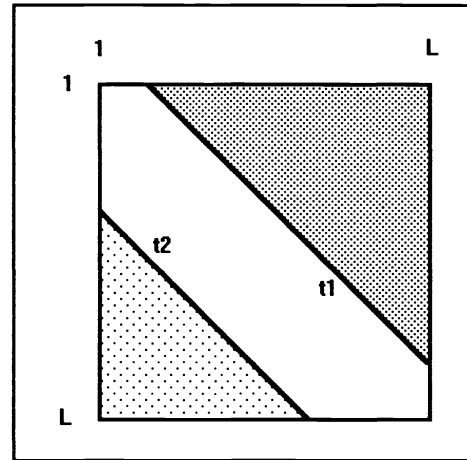

Fig. 5 Spectral co-occurrence matrix showing two thresholds and three regions.

the equations to follow, the other methods are similarly obtained by choice of appropriate limits in the summations.

Following from Eqs. (4) and (5), the region and element probabilities for the ND regions are

$$P^t_{\text{UND}} = \sum_{j=1}^{L} \sum_{i=j+t_U}^{L} p(i,j) \ ,$$

$$p^t_{\text{UND}}(i,j) = \frac{p(i,j)}{P^t_{\text{UND}}} \ , \tag{10}$$

$$P^t_{\text{LND}} = \sum_{j=1}^{L} \sum_{i=1}^{j-t_L} p(i,j) \ ,$$

$$p^t_{\text{LND}}(i,j) = \frac{p(i,j)}{P^t_{\text{LND}}} \ . \tag{11}$$

The diagonal probability distribution is defined by

$$P^{tt}_D = \sum_{j=1}^{L} \sum_{i=j-t_L+1}^{j+t_U-1} p(i,j) \ ,$$

$$p^{tt}_D(i,j) = \frac{p(i,j)}{P^{tt}_D} \ . \tag{12}$$

Note that the tt superscript indicates the dependence on two threshold variables. With these probabilities, the entropies for the regions can now be computed by

$$H_D(t_U,t_L) = - \sum_{j=1}^{L} \sum_{i=j-t_L+1}^{j+t_U-1} p^{tt}_D(i,j) \ \log p^{tt}_D(i,j) \tag{13}$$

for the diagonal region, and

$$H_{\text{UND}}(t_U) = - \sum_{j=1}^{L} \sum_{i=j+t_U}^{L} p^t_{\text{UND}}(i,j) \ \log p^t_{\text{UND}}(i,j) \ ,$$

$$H_{\text{LND}}(t_L) = - \sum_{j=1}^{L} \sum_{i=1}^{j-t_L} p^t_{\text{LND}}(i,j) \ \log p^t_{\text{LND}}(i,j) \ , \tag{14}$$

for the nondiagonal regions.

Fig. 4 Diagram of regions of co-occurrence matrix for the four region method.

The entropy of the diagonal region is a 2-D function. In the 2R method in this section there was only one threshold and in the 4R there were two thresholds but each of the regions was dependent only on one of them. In the 2R and 4R methods, the maxima is found over a 1-D function and for 3R the maxima will have to found over a surface. The entropy of the co-occurrence matrix $\mathbf{H}(t_U, t_L)$ is

$$\mathbf{H}(t_U, t_L) = \frac{H_D(t_U, t_L) + H_{\mathrm{UND}}(t_U) + H_{\mathrm{LND}}(t_L)}{3} , \qquad (15)$$

where the denominator of 3 is simply to keep the entropy magnitudes on the same scale. The 3R entropy $\mathbf{H}(t_U, t_L)$ requires a more lengthy computation time than the other two methods, but gives full and equal weight to the three regions. A division of the spectral co-occurrence into three regions is most satisfying from a system standpoint in that it fits the best the system ''model'' of having a spectrally flat background and an object or target spectrally brighter or darker than the background. The optimal threshold selection, that which maximizes the entropy, is formally stated by

$$(t_U, t_L) = \arg\left[\max_{1 \leq t_U, t_L \leq L} H(t_U, t_L)\right] . \qquad (16)$$

The results of this operation on the image 1 pairs (bands 1 and 2 and bands 3 and 6) are shown in Figs. 6 to 9. A binary image is formed by mapping those pixels with transitions lying in the nondiagonal region to gray level 1 and those lying within the diagonal region to gray level 256. In Fig. 7 the cloud shape can clearly be seen along with a small number of pixels that appear to be clutter, but may indeed be a low vapor concentration close to the ground. On the day these images were collected, the wind was brisk (approximately 5 m/s) and the main cloud plume moved about in the field of view. This cloud image comes from the lower nondiagonal region; the upper nondiagonal region (Fig. 6) had no cloud target. Because a threshold is chosen on the information present, background features are segmented in the absence of a true target. Based on study of the background statistics, a minimum threshold level could be chosen for a target presence decision. The cloud image in Fig. 9 is different from that in Fig. 7 because the spectral bands are collected sequentially. Therefore, the cloud had moved. The
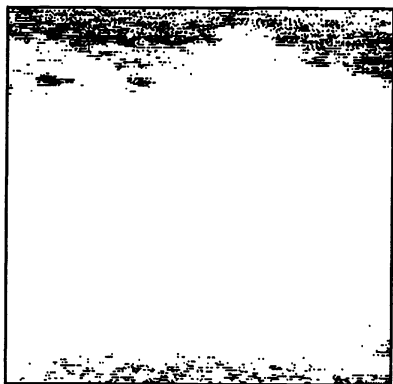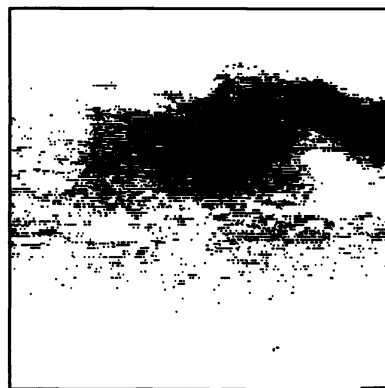

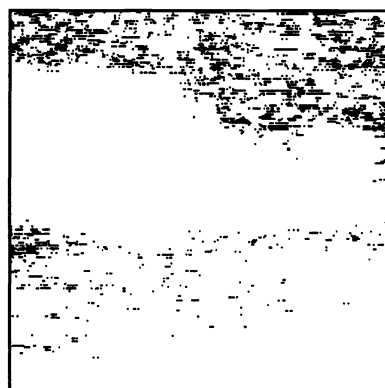Fig. 7 Binary image from lower threshold 19 from bands 2 and 1 of image 1.

Fig. 8 Binary image from upper threshold 3 from bands 3 and 6 of image 1.
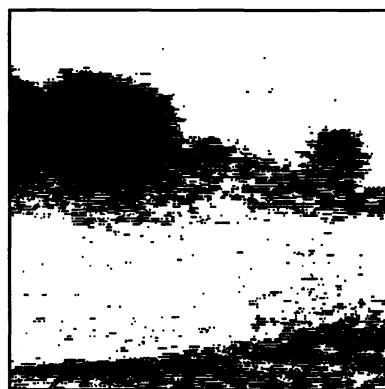

Fig. 9 Binary image from lower threshold 12 from bands 3 and 6 of image 1.

scene background remains static in general over many seconds so it is not improper to analyze an image pair as long as only one image has a nonstationary object such as the vapor cloud. In addition to the vapor cloud, Fig. 9 contains the grassy foreground at the bottom of the image and some clutter in the middle ground. This incomplete segmentation could be due to the narrower distribution of gray levels in bands 3 and 6 and the lack of contrast as a result of the poor condition of the optical filters. Although there is as great a transition length (the gray-level difference of the transition)

Fig. 6 Binary image from upper threshold 1 from bands 2 and 1 of image 1.

in the 1,2 pairing as in the 3,6 pairing the bulk of the distribution lies much closer to the central diagonal in the band 3 and 6 co-occurrence matrix. It might be easier to demonstrate the advantages if a data set with more bands and many targets, such as AVIRIS (Airborne Visible/InfraRed Imaging Spectrometer) type image cubes with 224 bands[10] could be used. It is our hope that high-spectral-resolution imagery of vapor clouds will be collected in the near future.

Data from a second scene (image 2 in Fig. 10) collected by the same imager is now examined. The vapor type and six bands are the same as before, but the scene contains a greater variation of gray levels in the background. One can see some foliage to the left, and electric pole with a warm transformer in the upper center, a hot steam pipe in the lower part of the image (visible in all bands), and glass in a truck reflecting the cold sky at the center right of the image. Using again the pairings of 1, 2 and 3, 6 because the experimental conditions of scene 2 are the same as those of scene 1 except for the FOV, we obtain the resulting binary images in Figs. 11 to 14. Visible in Figure 12 is the vapor cloud, rising from the right-hand side of the image, as well as a bit of the near horizon and some bloom around the steam pipe. The vapor dissemination point is within the scene, but because the vapor is being released from a bottle under pressure it is cool and then warms quickly to the ambient temperature, at which it is warmer than the background. Figures 13 and 14 present an interesting case: Fig. 14 shows the cloud along with the truck windows and some foreground and Fig. 13 shows only the steam pipe, the vapor cloud is not well segmented. These two features (pipe and windows) are the temperature extremes in the scene. Although the range of possible thresholds (1 to 116) is larger than in the image 1 example (1 to 43), the threshold, $t_H = 10$ or 15, selected by the entropy method is not out of the range of previous cases. This indicates that a very large gray-level variation lies in the rather spatially small features segmented and that although their numbers are not great, the co-occurrence matrix elements that relate to these transitions dominated the calculation by their magnitude. Features of temperature extreme can cause the threshold to miss the intended vapor target because of the relatively large transition variations possible.
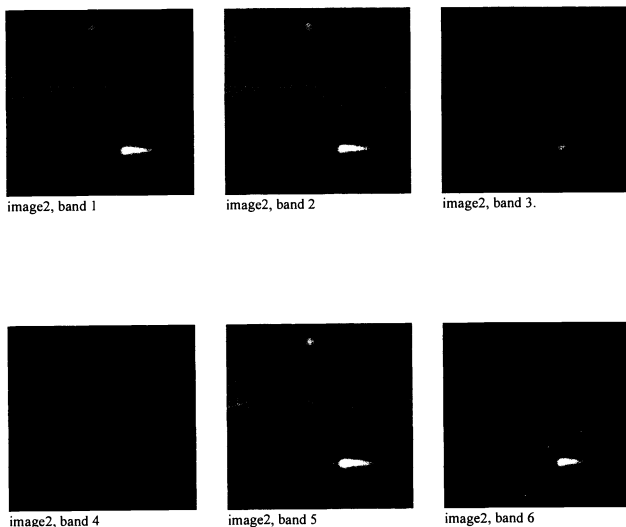

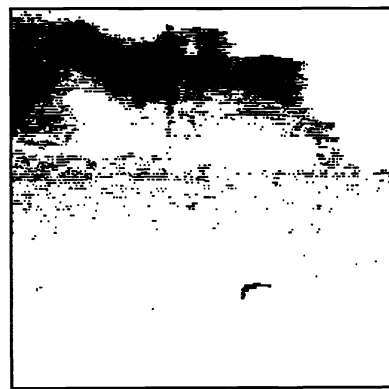Fig. 11 Binary image from upper threshold 1 from bands 1 and 2 of image 2.

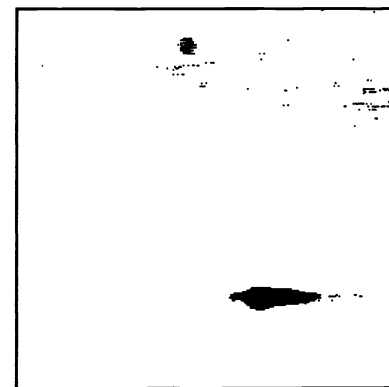Fig. 12 Binary image from lower threshold 19 from bands 1 and 2 of image 2.

Fig. 13 Binary image from upper threshold 10 from bands 3 and 6 of image 2.

image2, band 1   image2, band 2   image2, band 3.

image2, band 4   image2, band 5   image2, band 6

Fig. 10 Six spectral band images for scene 2.

Because the entropy for the 3R method is a surface that must be searched for the maximum value, the computation load is the square of that required for the 2R and 4R methods. In the case of the data studied in this paper, all of the entropy surfaces were smooth and had only one maxima (entropy is a convex function). Thus, an assent algorithm was used to greatly reduce the computational load. The entropy calculation is normally limited to combinations of nonzero region probabilities, which generally reduced the iterations to about
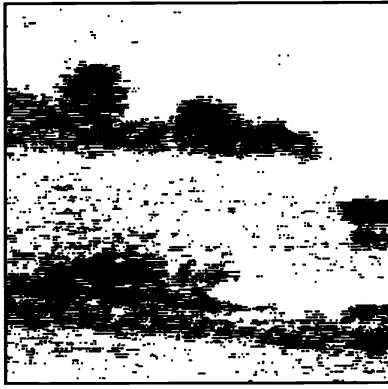
Fig. 14 Binary image from lower threshold 15 from bands 3 and 6 of image 2.

1000 from a possible 16,384 (by limiting the threshold to a value ≤ 128, $128^2 = 16,384$). The asset begins by computing $H(1,1)$, $H(1,2)$, $H(2,1)$, and $H(2,2)$. The largest of these is selected as the peak and the adjacent values (the eight nearest neighbors) are computed. If one of the neighbors is larger, it becomes the peak and the process continues until the maxima is found. For the data used here, the assent to the global maximum required only from 40 to 60 iterations or entropy calculations, a 95% reduction.

4 Simultaneous Detection of Two Targets

To assess the spectral target detection methods, a synthetic cloud has been inserted into band 1 of image 1. Because band 2 of image 1 contains a true vapor cloud, this will permit the investigation of the two-target case and also, because the parameters of the synthetic cloud can be manipulated, some estimation of the algorithms vapor concentration sensitivity may be made.

The synthetic cloud will be a 2-D zero-mean Gaussian concentration distribution. The x and y dimensions are as-

sumed to be independent, and thus the total distribution can be a product. An insertion location in the target image is chosen and defined as point (0,0) for the purpose of the synthetic cloud distribution. The standard deviations σ_X and σ_Y are chosen as well to set the cloud size. Lastly, a peak cloud concentration in gray level ρ is selected. With the parameters chosen, the individual axis density functions $f_X(x)$ and $f_Y(y)$ are computed and then normalized by $f_X(0)$ and $f_Y(0)$ respectively to give a center value of 1. Combining all the terms yields

$$f(x,y) = \rho\left(\frac{f_X(x)}{f_X(0)}\right)\left(\frac{f_Y(y)}{f_Y(0)}\right) . \tag{17}$$

The function $f(x,y)$ must then be quantized before being added to the digital target image. Quantization is accomplished by the Fortran rounding function NINT. By observing the length and width of the quantized cloud distribution, one knows the dimensions of the cloud being inserted. Figure 15 contains quantized curves of $f(x) = \rho f_X(x)/f_X(0)$ for the ρ values of 5, 10, and 20 and also an unquantized curve for comparison. Note that the cloud width changes relatively little because the standard deviation is constant at 10. This feature will help in the assessment. These sizes are used to assess how much of the cloud is resolvable in the segmented image and from that, a measure of the vapor detection threshold. The contrast resolvable in an imager is a function of the instrument parameters, for which general values are assumed, the temperature difference (Δt) between cloud and background, and the vapor concentration and absorption coefficient. Because the cloud is synthetic, some realistic values are selected. A detection threshold for a particular vapor can then be derived as a function of Δt.

A set of synthetic clouds were generated, the parameters of which are contained in Table 2. The x and y lengths are the overall dimensions of the synthetic cloud, i.e., the spans
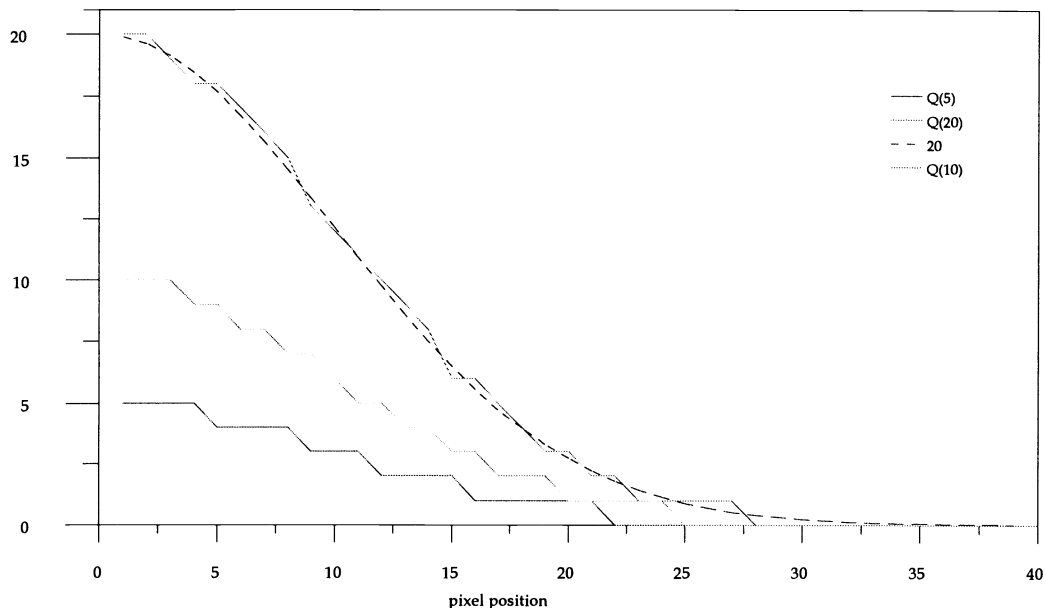

Fig. 15 These curves show half of the synthetic vapor cloud concentration function. The dashed curve is the function before quantization and the solid curves are quantized with peaks of 5, 10, and 20.

Table 2 Synthetic cloud parameters.

image	AC1	AC2	AC4
ρ	10	20	40
σ_x	30	30	30
σ_y	15	15	15
x length	117	137	153
y length	59	69	77
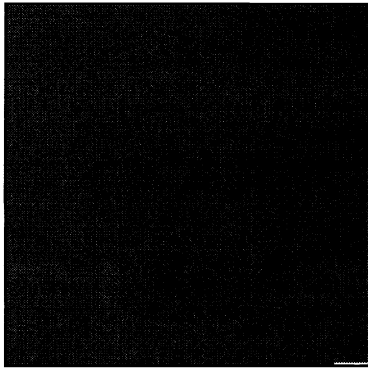
Fig. 16 Synthetic cloud, $\rho = 10$.

Fig. 17 Synthetic cloud, $\rho = 20$.

Fig. 18 Synthetic cloud, $\rho = 40$.

Fig. 19 Band 1 with AC1 inserted.

of pixels with a gray level greater than zero. These clouds are inserted in band 1 of image 1 centered at location (130,200). Figures 16 to 18 are of the synthetic clouds alone (the cloud is darker), and Figs. 19 to 21 show the resulting band 1 images from the insertion (the cloud is lighter). Observe that the synthetic cloud in the second two-target image, with $\rho = 20$, to the eye is quite faint, and the $\rho = 10$ cloud is not visible at all. These two target images were then processed by the three spectral co-occurrence matrix methods with band 2 as the second image of the pair.

Not all of the resulting images are included here. For instance, the two-region method produced a blank image for the smallest synthetic cloud because it uses only one threshold, which is more influenced by the significantly larger actual vapor cloud in band 2. Figures 22 through 28 are arranged left to right by method (2R, 4R, 3R), and top to bottom by synthetic vapor cloud concentration ($\rho = 10$, 20, 40). There are no images for the 2R method for the two lower concentrations as mentioned before because the images are blank. In the rest, it is easy to see the increasing detection moving left to right. The 3R co-occurrence matrix method does the best job of resolving the target cloud, although it also picks up the most sky. A human operator might readily recognize the sky portion as sky, but an autonomous system might have to employ some morphological processing such as region growing[11] to discriminate homogeneous areas of detection from the clutter, which might also be picked up by the co-occurrence matrix or use more than two spectral bands. For higher concentration clouds, there is only slight difference between the result of the 4R and 3R methods, but in the low concentration case (Figs. 22 and 23) the difference is significant. To get an objective measure of the detection or segmentation power of the three methods, signal-to-noise ratio (SNR) is used.
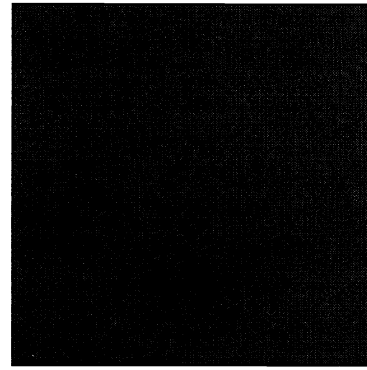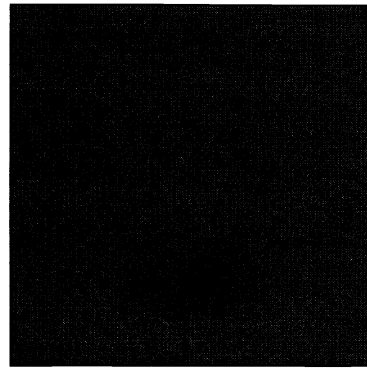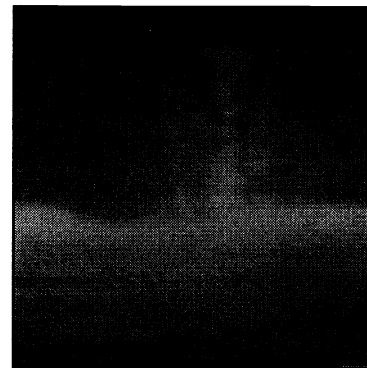
Comparing the cloud "energy" with the noise "energy" will lead us to an SNR value. Computing SNR for a series of images in which the cloud strength or concentration has been varied and subjecting these to the spectral co-occurrence processing methods will result in a series of binary images in which the observer can determine cloud presence by subjective means. This is a difficult procedure to quantify due to the subjective final decision, but it can be made less variant when a temporal sequence is used as well because human perception is very sensitive to the dynamics of the cloud as it moves in time.

In computing the SNR, the difference image with no synthetic cloud, $e(x,y) = I_1(x,y) - I_2(x,y)$, is averaged squared over the local region occupied by the cloud. A ratio is made
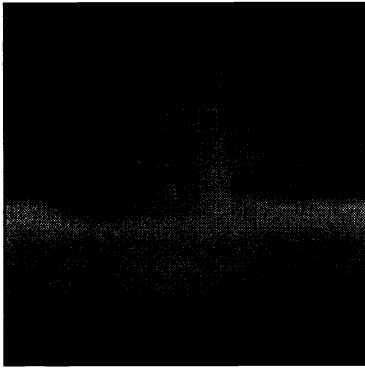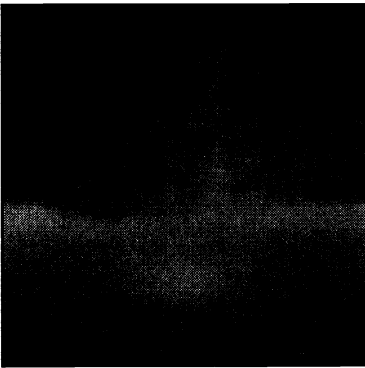
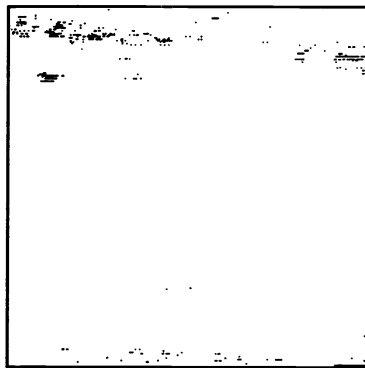**Fig. 20** Band 1 with AC2 inserted.

Fig. 21 Band 1 with AC4 inserted.

Fig. 22 Result of upper threshold 5, four-region method, $\rho = 10$.

Fig. 23 Result of upper threshold 1 of three-region method.

Fig. 24 Result of upper threshold 5, four regions, $\rho = 20$.

Fig. 25 Result of upper threshold 4, three regions.

Fig. 26 Result of threshold 18, two regions, $\rho = 40$.

with a similar measurement of the cloud containing difference frame, $g(x,y) = I'_1(x,y) - I_2(x,y)$, where I' signifies the presence of the synthetic cloud. The ratio R is

$$R = \frac{\sum_{\mathbb{R}} \sum g^2(x,y)}{\sum_{\mathbb{R}} \sum e^2(x,y)} \; , \tag{18}$$

in which $\mathbb{R}$ is the local region under consideration. Gonzalez and Wintz[12] assert that the variance is a measure of contrast, thus the SNR will serve as an indicator of contrast and signal strength. The terms in Eq. (18) may not look like variances, but Ref. 12 leaves off the squared mean term of the variance and when checked here, the inclusion of the mean had very
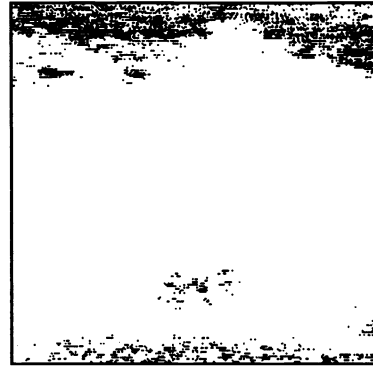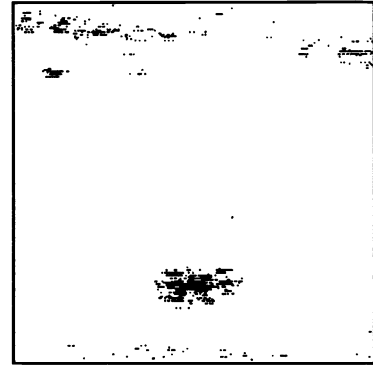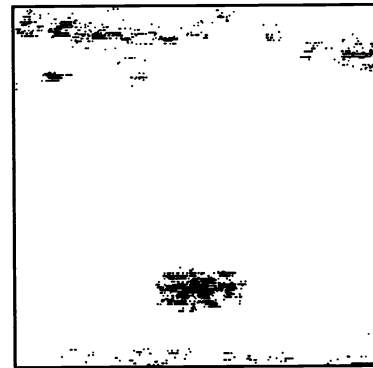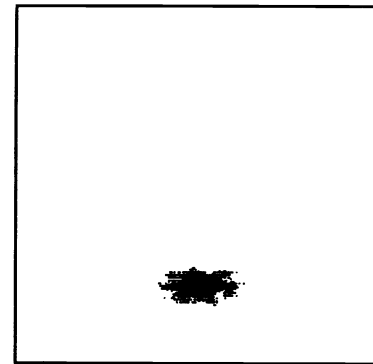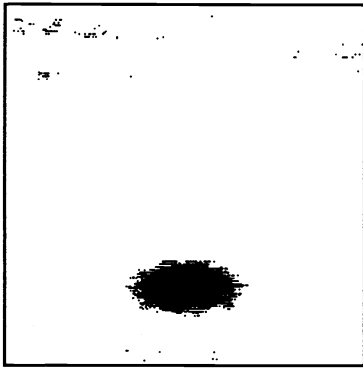
Fig. 27 Result of upper threshold 7, four regions.
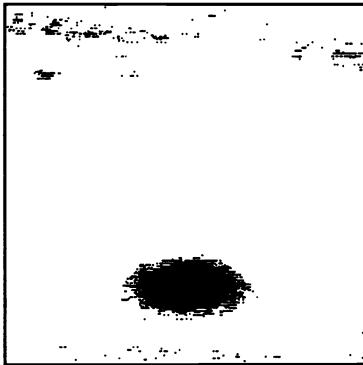

Fig. 28 Result of upper threshold 5, three regions.

little effect on R. Of course, the normalization by the area of the region cancels in the ratio. The SNR is found by

$$SNR = 10 \, \log_{10}(R) \ . \tag{19}$$

For the three synthetic clouds used in the images shown and for the additional values of $\rho = 5$ and $\rho = 80$, the SNR results are plotted in Fig. 29. The curve labeled "raw SNR" is the SNR of the band 1 image with the cloud versus band 1 without the cloud and the curve labeled "diff SNR" is from the difference of bands 1 and 2 with the cloud versus the difference of bands 1 and 2 without the cloud. From an observation of the spectral co-occurrence thresholding results and using the diff SNR curve, the 2R detection limit is between 0.8 and 1.3, the 4R detection limit is about 0.8, and the 3R detection limit is between 0.5 and 0.8. More data might narrow these ranges, but the examples given are sufficient to illustrate the results.

5 Local Thresholding

Local thresholding, the practice of computing a threshold for a window or portion of an image rather than for the image as a whole (global thresholding), has been attempted with varying degrees of success with many of the thresholding techniques developed. The utility of the local implementation of a thresholding method depends to a great part on the nature of the image and the target or object in that image which the threshold is supposed to detect or segment. If either the noise or background gray level varies in a nonuniform manner in the image, for instance, if the image scene is illuminated from a side so as to cause shadows, a bimodal histogram will show
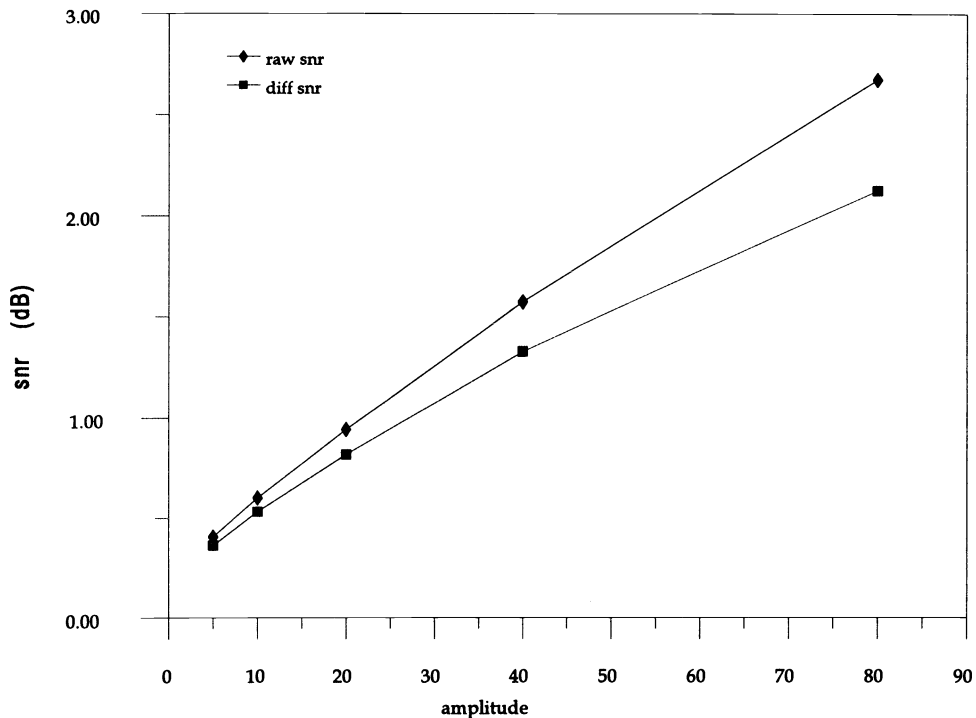

Fig. 29 SNR for the synthetic clouds in the two target images as a function of peak concentration or amplitude.
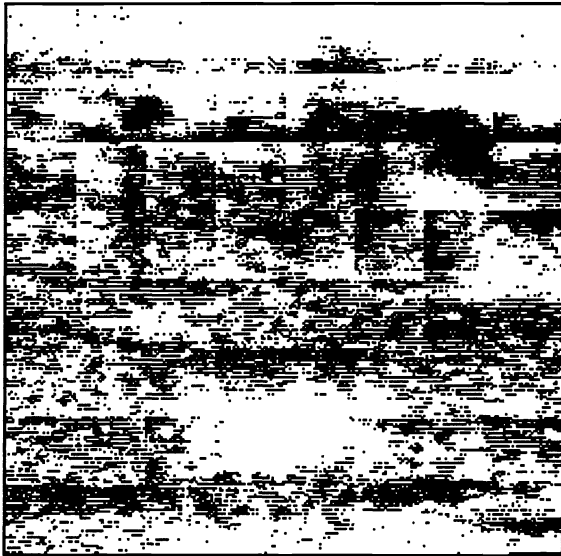
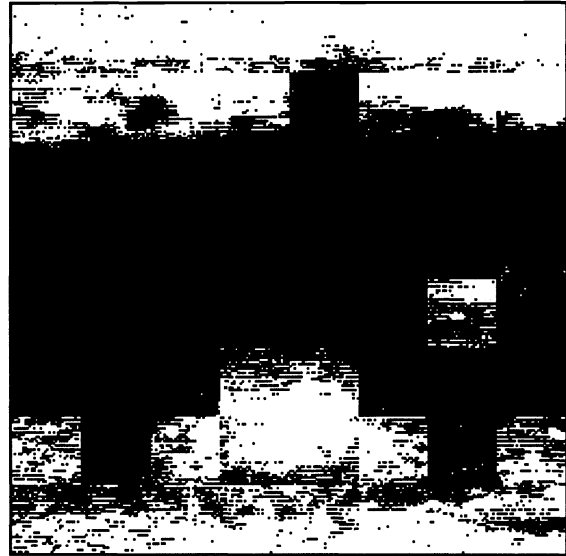**Fig. 30** Image 1, 4R, bands 1 and 2 with two targets, 32×32 window, lower.

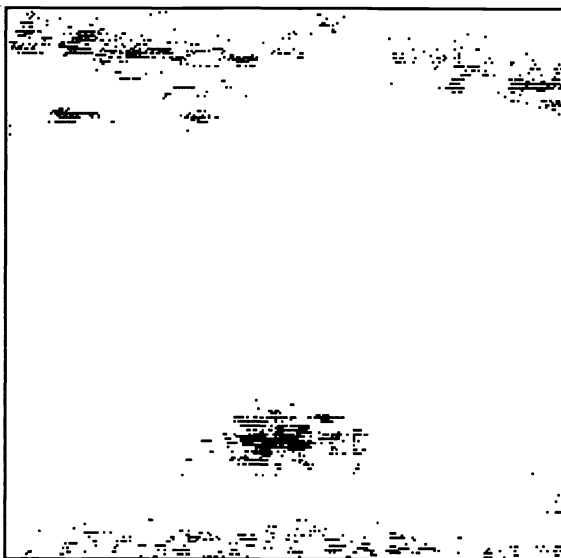**Fig. 32** Image 1, 3R, bands 1 and 2, 32×32 window, two targets, lower.

Fig. 31 Image 1, 4R, bands 1 and 2, two targets, 32×32 window, upper, small cloud has medium concentration (AC2).

a significant population in the valley between the modes. The shadow or uneven shading will cause an error in the segmentation. Alternately, if the object is unevenly lit or has large gray-level variations, these will possibly result in errors.

Medical images are one data type that has benefited from local thresholding. Chow and Kaneko[13] successfully used local methods on sequences of cineangiograms. Local versus global methods are also discussed in Refs. 14 and 15.

The thresholding algorithm in its entirety is executed on each subimage or window of a chosen size of the original image to generate a threshold which is used to segment the gray-scale subimage into a binary subimage. The subimages are then reassembled into a complete binary image.

The window size itself has considerable effect on the magnitude of edge discontinuities and should be chosen with some characteristic of the input image in mind. Fitting char-

acteristics could be the spatial size of variations in the image, general target size, or textural frequency. In field applications, the expected size of chemical vapor clouds has been extensively modeled and an appropriate window could be chosen for one of a few scenario classes. A smaller window means more computations are required, so the window should be no smaller than needed. The target should occupy a statistically significant portion of the image, a fraction of 0.25 to 0.5 will give a significant edge feature. If the target covers the whole window there is no background and the entropy methods will segment variations in the target itself. Because this work is with a limited data set, windows will be chosen by *a priori* knowledge of the cloud size.

Modeling of tactics, delivery systems, and requirements yields an expected cloud width of 100 m at a range of 3000 m from the detector. The cloud would then occupy 1.91 deg of the FOV of 30 deg horizontally and have a width of 32.6 pixels in the digital scene. Because the placement of the cloud in the scene is random, the expected fraction of a 32×32 window that the cloud would occupy is 0.26. This is found by considering the edge of the cloud to be uniformly distributed over the 32 pixels of the window. The mean of a uniform distribution is 0.5, thus 0.5 (32.6 pixels) = 16.3 pixels of the cloud will be in a window or 0.51 of the width of the window is occupied. The height of a cloud is dependant on more variables, so for simplicity, we assume it occupies the same relative vertical fraction as it does the horizontal fraction. Then we have $(0.51)^2 = 0.26$ for the total fraction.

Thresholds computed on the basis of spectral information can be applied to a local spatial area in much the same way as in the previous section. A pair of images of the same scene but with different spectral bands are split into subimages and the spectral co-occurrence matrix entropy thresholding methods of Sec. 3 are applied to each subimage. The binary subimages are reassembled as a last step. This process was carried out for both the 4R and 3R methods and the resulting binary images are shown in Figs. 30 through 33. The large vapor cloud is unresolved by both methods, and the small synthetic
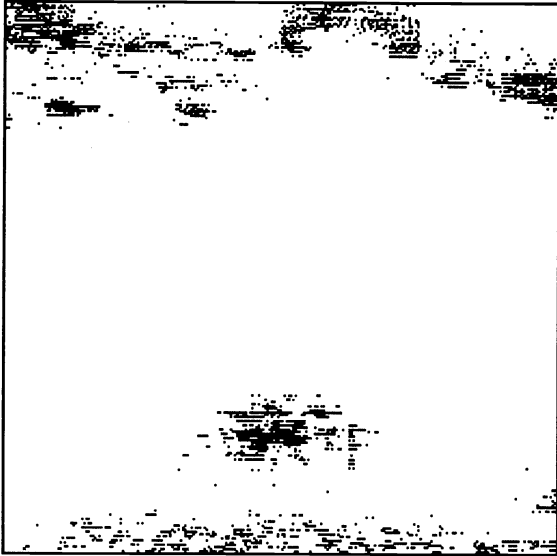
Fig. 33 Image 1, 3R, bands 1 and 2, two targets, 32 × 32 window, upper, small cloud has medium concentration (AC2).

cloud is resolved less well than by the globally computed threshold. As discussed in the previous section, this results from the fact that the chosen method will do its best to segment the local area, even if the only "signal" in the subimage is noise. Similarly, if the subimage signal is primarily target, then the method will segment the target into two regions. Thus, the point is again worth making that the subimage size must be sufficient to contain a fair representation of both target and noise or background objects. The vapor clouds in image 1 and image 2 are quite large, occupying of the order of 25% of the total area. The synthetic clouds, however, fit better the criteria presented for window size selection and the result is more favorable.

If we consider the threshold computed for each of the subimages as a statistical representation of the image, we may treat the distribution of thresholds to arrive at a more acceptable result. Suppose we have an image of size $M \times M$, and select a window of size $A \times A$, then the matrix of thresholds can be $[Th(i,j)]_{B \times B}$, where $B = M/A$ and M is a multiple of A. The histogram of the threshold values

$$H_{Th}(k) = \frac{1}{B^2} \sum_{i=1}^{B} \sum_{j=1}^{B} \delta$$

where $\delta = \begin{cases} 1 & \text{if } Th\ (i,j) = k \\ 0 & \text{otherwise} \end{cases}$. (20)

If the image contains a target, $H_{Th}(k)$ should contain a bimodal distribution, one peak corresponding to the subimages containing the target and the other peak corresponding to subimages containing only noise or background. Because $H_{Th}(k)$ is a distribution, we can apply to it segmentation methods designed to find the best segmentation point for a distribution that is the mixture of two Gaussians. This procedure was implemented with the 3R method computed locally followed by the Otsu[16] method to segment the histogram of local thresholds. That optimal threshold t_{Th} is then applied globally to segment the original image.

Otsu[16] describes a simple gray-level thresholding method which operates on the histogram and has been found to perform rather well in most situations. The premise is that the histogram can be represented as a mixture distribution of two Gaussian distributions. Otsu develops three equivalent criterion functions, which can be reduced to a function of the zeroth- and first-order statistics of the histogram. The criteria can be used to select a threshold that maximizes the between class separation on the basis of the class variances. For an image $I(x,y)$ with pixels having L gray levels, the histogram or gray-level probability distribution is formed by summing the number n_i of each of the L gray levels found in the image. The elements of the histogram are then made a probability distribution by dividing by the total number of pixels in the image N; $p_i = n_i/N$. A threshold t will divide the histogram into two classes, C_1 and C_2.

Each class will have a probability

$$Pr(C_1) = \sum_{i=1}^{t} p_i = \omega(t) \ ,$$

$$Pr(C_2) = \sum_{i=t+1}^{L} p_i = 1 - \omega(t) \ , \tag{21}$$

and a class mean

$$\mu_1 = \frac{\sum_{i=1}^{t} i p_i}{\omega(t)} = \frac{\mu(t)}{\omega(t)} \ , \tag{22}$$

$$\mu_2 = \frac{\mu(L) - \mu(t)}{1 - \omega(t)} \ , \tag{23}$$

where $\mu(L)$ is the mean of the whole image. Three discriminant criterion functions are developed, all of which result in the same threshold, and one is reduced to the function $\sigma_B^2(t)$, which is the between class variance and is defined as

$$\sigma_B^2(t) = \frac{[\mu(L)\omega(t) - \mu(t)]^2}{\omega(t)[1 - \omega(t)]} \ , \tag{24}$$

and the optimal threshold t^* is

$$t^* = \arg \left[\begin{array}{c} \max \sigma_B^2(t) \\ 1 \leq t \leq L \end{array} \right] \ . \tag{25}$$

Even though visually very dim and indistinct in the original images, the target is successfully segmented by the global spectral co-occurrence methods. Viewing the co-occurrence matrix itself, for a target containing pair of images, the off-diagonal population representative of target containing pixels is rarely well separated from the main diagonal cluster. In contrast, the histogram curves exhibit good separation of the bimodal or multimodal peaks, as in Figs. 34 and 35. The segmented images follow in Figs. 36 to 39. Note that the large vapor cloud is more fully segmented, but at the penalty of greater clutter in the vicinity. The clutter could be reduced by a spatial filter of the appropriate size. For the synthetic cloud, the medium concentration cloud has better definition than the result of either the global or simple local methods, whereas the global could be said to be slightly better for the
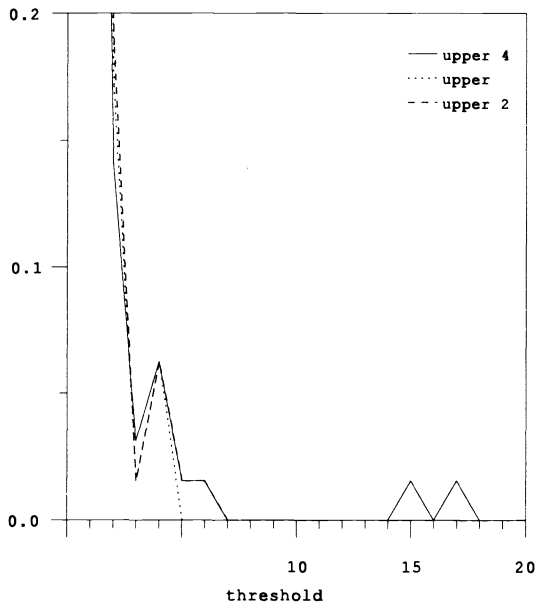
Fig. 34 Histograms of the upper (synthetic cloud) thresholds for three concentrations, AC, AC2, and AC4.
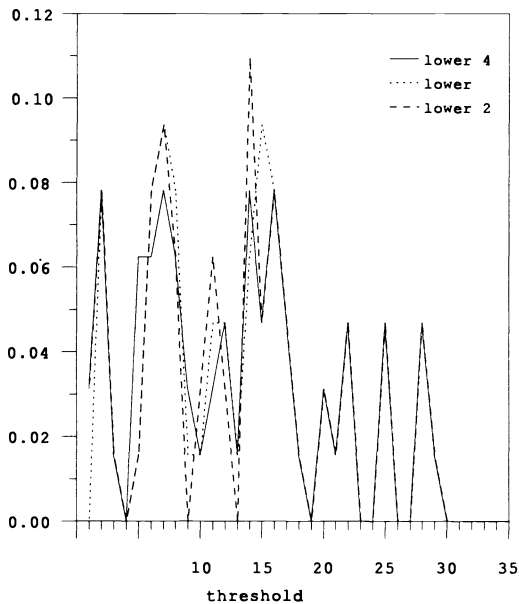

Fig. 35 Histograms of the lower (vapor cloud) thresholds for three synthetic cloud concentrations.

case of the low-concentration cloud. The reason can be seen in the curves in Figure 34, where the thresholds are selected at the first valley. For the high concentration, the two-step local method selected a threshold of 6 and the global a threshold of 5. There is negligible difference in the segmented cloud.

Although the threshold histograms do illustrate the mixture distribution, the curves are not very dramatic, especially for the case of the real vapor cloud. This can be attributed to the large cloud size relative to the window, the concentration gradient within the cloud, the low signal level, and the small number of windows in the entire image (64). With a higher resolution image and optimal window size, the pop-
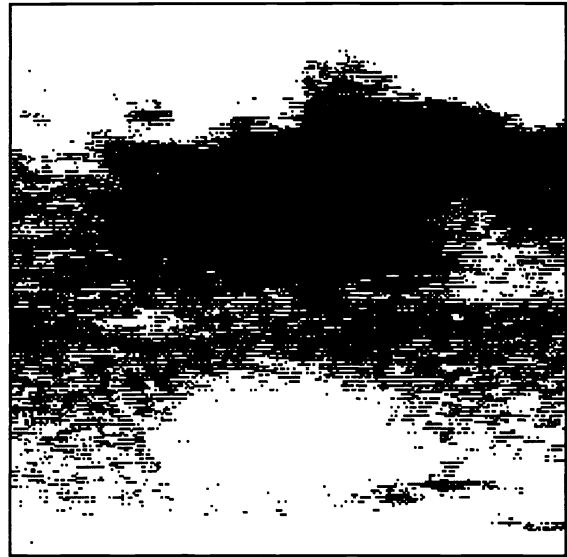

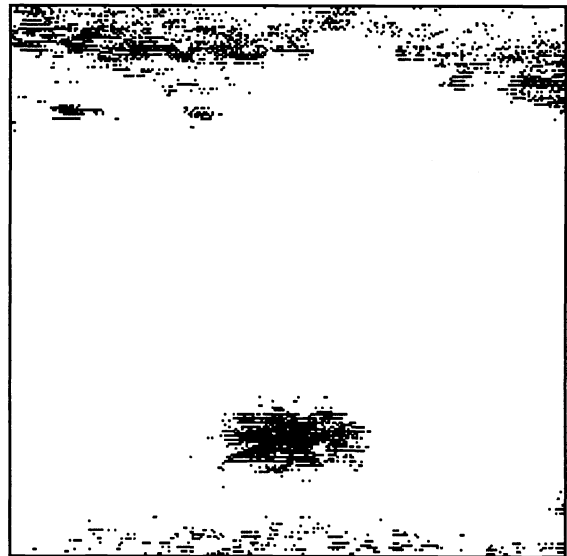Fig. 36 Method 3R locally at 32×32, followed by Otsu, threshold at 12.

Fig. 37 Method 3R, locally at 32×32, followed by Otsu, threshold at 2, medium concentration cloud (AC2).

ulations should be better resolved. As discussed earlier, the 32×32 window is best suited for expected field target vapor clouds, of which the real cloud in these images is not representative, but the synthetic cloud is. Therefore, the synthetic cloud results are a preferable example.

6 Conclusions

A method of chemical vapor cloud detection and target segmentation has been demonstrated to achieve good results when operating on multispectral infrared image sets. Three methods of co-occurrence matrix division were studied and the three-region division with two thresholds was found to be the best. Multiple targets can be detected and segmented if they have distinct spectral signatures. Computation of thresholds locally followed by class separation and global
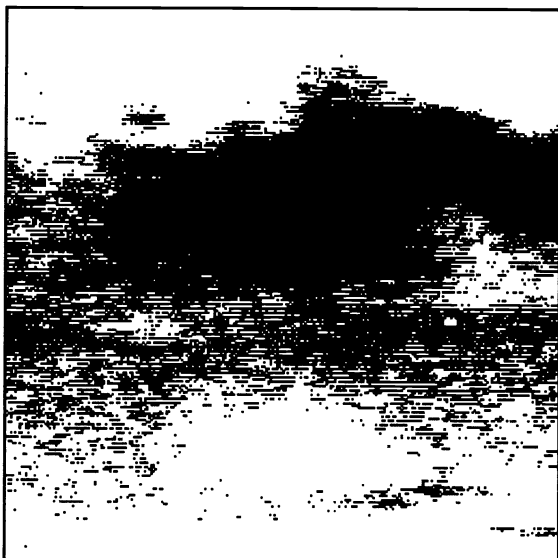
Fig. 38 Method 3R, locally at 32×32, followed by Otsu, threshold at 13.
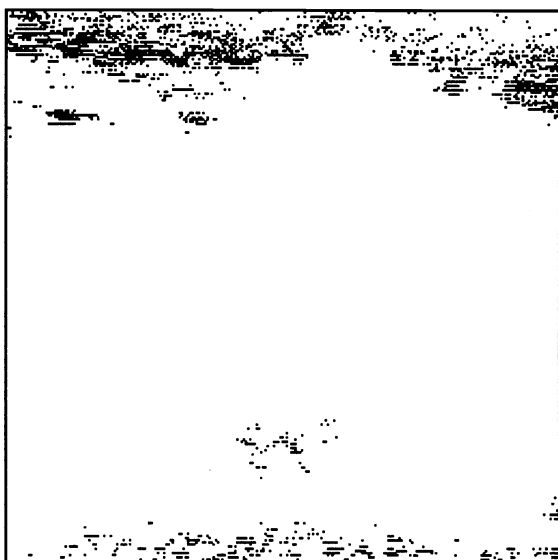

Fig. 39 Method 3R, locally at 32×32, followed by Otsu, threshold at 2, low concentration vapor cloud (AC).

threshold application improves the result over the simple global calculation. An assent algorithm reduced computation time by about 95%. The assent algorithm is possible because of the observed smoothness of the entropy functions.

References

1. C.-I. Chang, K. Chen, J. Wang, and M. L. G. Althouse, "A relative entropy-based approach to image thresholding," *Pattern Recog.* **27**(9), 1275–1289 (1994).
2. N. K. Pal and S. K. Pal, "Entropic thresholding," *Signal Process.* **16**, 97–108 (1989).
3. M. L. G. Althouse and C.-I. Chang, "Chemical vapor detection with a multispectral thermal imager," *Opt. Eng.* **30**(11), 1725–1733 (1991).
4. B. Chanda, B. B. Chaudhuri, and D. Dutta Majumder, "On image en-
hancement and threshold selection using the greylevel co-occurrence matrix," *Pattern Recog. Lett.* **3**, 243–251 (1985).
5. V. Seferadis and M. Ghanbari, "Use of co-occurrence matrices in the temporal domain," *Electron. Lett.* **26**(15), 1116–1118 (1990).
6. V. Seferadis and M. Ghanbari, "Image sequence coding using temporal co-occurrence matrices," *Image Commun.* **4**, 185–193 (1992).
7. J. E. Auborn, J. M. Fuller, and H. M. McCauley, "Target detection by co-occurence matrix segmentation and its hardware implementation," *Opt. Eng.* **32**(11), 2698–2702 (1993).
8. P. J. Whitebread, "Covariance matrix matching for multispectral image classification," in *Proc. ICASSP*, pp. 385–388, Adelaide, Australia (1994).
9. L. Carr, L. Fletcher, P. Holland, J. Leonelli, D. McPherrin, and M. Althouse, "Characterization of filtered FLIR systems designed for chemical vapor detection and mapping," *Proc. SPIE* **1309**, 90–103 (1990).
10. J. Harsanyi and C.-I. Chang, "Hyperspectral image classification and dimensionality reduction: an orthogonal subspace projection approach," *IEEE Trans. Geosci. Remote Sens.* **32**(4), 779–785 (1994).
11. D. H. Ballard and C. M. Brown, *Computer Vision*, Prentice-Hall, Englewood Cliffs, NJ (1982).
12. R. C. Gonzalez and P. Wintz, *Digital Image Processing*, 2nd ed., Addison-Wesley, Reading, MA (1987).
13. C. K. Chow and T. Kanedo, "Automatic boundary detection of the left ventricle from cineangiograms," *Comput. Biomed. Res.* **5**, 388–410 (1972).
14. R. M. Haralick and L. G. Shapiro, "Image segmentation techniques," *Comput. Vis. Graph. Image Process.* **29**, 100–132 (1985).
15. P. K. Sahoo, S. Soltani, and A. K. C. Wong, "A survey of thresholding techniques," *Comput. Vis. Graph. Image Process.* **41**, 233–260 (1988).
16. N. Otsu, "A threshold selection method from grey-level histograms," *IEEE Trans. Syst. Man Cybernet.* **SMC-9**(1), 62–66 (1979).

Mark L. G. Althouse received the BS in physics from the Pennsylvania State University in 1981, the MS from Johns Hopkins University in 1988, and the PhD from the University of Maryland, Baltimore, in 1994, both in electrical engineering. Since 1981 he has been employed by the U.S. Army Chemical and Biological Defense Command at Aberdeen Proving Ground, Maryland, working on the remote detection of chemical vapors and biological aerosols. From 1985 to 1986 he was an exchange scientist at the German NBC Defense Laboratory in Munster, Germany. His current research interests include signal/image processing, signal detection and estimation, pattern recognition, neural networks, and multi- and hyperspectral sensors. Dr. Althouse is a member of OSA, SPIE, IEEE, Tau Beta Pi, and Sigma Xi.
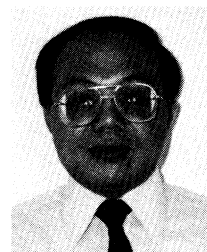
Chein-I. Chang received the BS, MS, and MA degrees from Soochow University, Taipei, Taiwan, in 1973; National Tsing Hua University, Hsinchu, Taiwan, in 1975; and SUNY at Stony Brook in 1977, respetively, all in mathematics, and the MSEE degree from the University of Illinois at Unbana-Champaign in 1982 and the PhD degree in electrical engineering from the University of Maryland, College Park, in 1987. From 1973 to 1977 he was a teaching assistant at Soochow University, National Tsing Hua University, and SUNY at Stony Brook. From 1977 to 1982 he was an instructor of mathematics at the University of Illinois, and was also a research assistant at the University of Maryland from 1982 to 1987. He was a visiting assistant professor from January 1987 to August 1987, asisstant professor from 1987 to 1993, and is currently an associate professor in the Department of Electrical Engineering at the University of Maryland, Baltimore County. His research interests include information theory and coding, signal detection and estimation, biomedical signal/image processing, multispectral/hyperspectral signal and image processing, sensor/data fusion, neural networks, computer vision, and pattern recognition. Dr. Chang is a senior member of the IEEE, SPIE, INNS, Phi Kappa Phi, and Eta Kappa Nu.