

Segmentation and Its Place in Machine Vision

*Robert M. Haralick**
Linda G. Shapiro

Department of Electrical Engineering, FT-10
University of Washington
Seattle, WA 98195

ABSTRACT

Segmentation is a basic step in a variety of image processing tasks. This paper indicates how segmentation stands in the grouping phase of vision algorithms by discussing the entire context of the phases of vision algorithms. The main body of the paper reviews common segmentation algorithms and gives examples for most of the algorithms discussed.

1. Overview

To understand what segmentation must accomplish, we must understand its place in the vision toolbox. The machine vision toolbox has five kinds of processes: conditioning, labeling, grouping, extracting, and matching. Segmentation falls in the grouping class. Before we review segmentation itself, we will discuss the five kinds of processes in the vision toolbox.

1.1 Conditioning

Conditioning is based upon a model which suggests that the observed image is composed of an informative pattern modified by uninteresting variations which typically add to or multiply the informative pattern. Conditioning estimates the informative pattern on the basis of the observed image. Thus conditioning suppresses noise which can be thought of as random unpatterned variations affecting all measurements. Conditioning can also perform background normalization by suppressing uninteresting systematic or patterned variations. Conditioning is typically applied uniformly and is context independent.

1.2 Labeling

Labeling is based upon a model which suggests that the informative pattern has structure as a spatial arrangement of events, each spatial event being a set of connected pixels. Labeling determines in what kinds of spatial events each pixel participates. For example, if the interesting spatial events of the informative pattern are only events of high valued pixels and events of low valued pixels, then the thresholding operation can be considered as a labeling operation. Other kinds of labeling operations include edge detection, corner finding, and identifying pixels which participate in varieties of shape primitives.

1.3 Grouping

The labeling operation labels pixels with the kinds of primitive spatial events the pixel participates in. The grouping operation identifies the events by collecting together or identifying maximal connected sets of pixels participating in the same kind of event. If the labels are symbolic then the grouping is really a connected components operation. If the labels are the gray levels, then the grouping operation is what the vision literature calls a segmentation. If the labels are step edges, then the grouping operation constitutes edge linking, etc.

Key Words: Segmentation, spatial clustering, machine vision.

*Contact Person: Robert M. Haralick
Phone No.: 206-545-4974

The grouping operation is the operation in which there is a change of logical data structure. The observed image, the conditioned image, and the labeled image are all digital image data structures. Depending on the implementation, the grouping operation can produce an image data structure in which each pixel is given an index which is associated with the spatial event to which it belongs or the grouping operation can produce a data structure which is a collection of sets. Each set corresponds to a spatial event and contains the pairs of (row, column) positions which participate in the event. In either case, there is a change in the logical data structure. The entities of interest before the grouping step are pixels. The entities of interest after the grouping step are sets of pixels.

1.4 Extracting

The grouping operation determines the new set of entities. But after the grouping step the new entities are naked. The only thing they possess is their identity. The extracting operation computes for each group of pixels a list of its properties. Example properties might include its centroid, its area, its orientation, its spatial moments, its gray tone moments, its spatial-gray tone moments, its circumscribing circle, its inscribing circle, etc. Other properties might depend on whether the group is considered as a region or as an arc. If the group is a region, then number of holes might be a useful property. If the group is an arc, then the average curvature might be a useful property.

Extracting also can measure topological or spatial relationships between two or more groupings. For example, an extracting operation may make explicit that two groupings touch or are spatially close or that one grouping is above another.

1.5 Matching

After the completion of the extracting operation, the events occurring on the image have been identified and measured. But the events in and of themselves have no meaning. The meaning of the observed spatial events occurs when a perceptual organization has occurred in which it is recognized that a specific set of spatial events in the observed spatial organization constitutes an imaged instance of some previously known object such as a "chair" or the letter "A."

It is the matching operation which determines the interpretation of some related set of image events associating these events with some given 3D object or 2D shape. The association determined by matching establishes a correspondence between each spatial event on the image in the related set of events with some spatial event on the 3D object or 2D shape. The association is one which in some sense best matches both the character of the spatial events and their spatial relationships. Thus, after matching, two primitive image events which stand in some spatial relationship will have associated with them two object events which stand in a similar relationship.

There is a wide variety of image operations which are matching operations. The classic one is template matching which is effective only if the variety of instances expected to be encountered is limited. For example, rotation and size variations must be very small. The background must

be nearly uniform. Random shape deformations must be minimal.

Simple shapes will correspond to a primitive spatial event and the property measurement from the primitive spatial event will often be adequate to permit recognition of the shape. In this case, the matching operation amounts to matching the vector of properties measured from the image spatial event with the vector of properties of a prototype representative. Such matching constitutes statistical pattern recognition.

Complex shapes will correspond to a set of primitive spatial events. Here, recognition must proceed by using the property vector of each observed spatial event as well as the spatial relationships between the events. In this case, the matching amounts to determining a relational homomorphism with unary constraints established by the required matching of the property vectors of the observed image events with the property vectors of the prototype primitives. Such a matching is what constitutes structural pattern recognition.

2. Segmentation

What should a good image segmentation be? Regions of an image segmentation should be uniform and homogeneous with respect to some characteristic such as gray tone or texture. Region interiors should be simple and without many small holes. Adjacent regions of a segmentation should have significantly different values with respect to the characteristic on which they are uniform. Boundaries of each segment should be simple, not ragged, and must be spatially accurate.

Achieving all these desired properties is difficult because strictly uniform and homogeneous regions are typically full of small holes and have ragged boundaries. Insisting that adjacent regions have large differences in values can cause regions to merge and boundaries to be lost.

As there is no theory of clustering, there is no theory of image segmentation. Image segmentation techniques are basically ad-hoc and differ precisely in the way they emphasize one or more of the desired properties and in the way they balance and compromise one desired property against another.

Image segmentation techniques can be classified as: measurement space guided spatial clustering, single linkage region growing schemes, hybrid linkage region growing schemes, centroid linkage region growing schemes, spatial clustering schemes, and split and merge schemes. As can be observed from this brief typology, image segmentation can be viewed as a clustering process. The difference between image segmentation and clustering is that in clustering, the grouping is done in measurement space. In image segmentation, the grouping is done on the spatial domain of the image and there is an interplay in the clustering between the (possibly overlapping) groups in measurement space and the mutually exclusive groups of the image segmentation.

The single linkage region growing schemes are the simplest and most prone to the unwanted region merge errors.

The hybrid and centroid region growing schemes are better in this regard. The split and merge technique is not as subject to the unwanted region merge error. However, it suffers from large memory usage and excessively blocky region boundaries. The measurement space guided spatial clustering tends to avoid both the region merge errors and the blocky boundary problems because of its primary reliance on measurement space. But the regions produced are not smoothly bounded, and they often have holes, giving the effect of salt and pepper noise. The spatial clustering schemes may be better in this regard, but they have not been well enough tested. The hybrid linkage schemes appear to offer the best compromise between having smooth boundaries and few unwanted region merges.

The remainder of the paper briefly describes the main ideas behind the major image segmentation techniques and gives example results for several of them. Some of the techniques can produce some very small regions. Additional image segmentation surveys can be found in Zucker (1976), Riseman and Arbib (1977), Kanade (1980), and Fu and Mui (1981).

2.1 Measurement Space Guided Spatial Clustering

This technique for image segmentation uses the measurement space clustering process to define a partition in measurement space. Then each pixel is assigned the label of the cell in the measurement space partition to which it belongs. The image segments are defined as the connected components of the pixels having the same label.

The accuracy of the measurement space clustering image segmentation process depends directly on how well the objects of interest on the image separate into distinct measurement space clusters. Typically the process works well in situations where there are a few kinds of distinct objects having widely different gray tone intensities (or gray tone intensity vectors, for multi-band images) and these objects appear on a near uniform background.

Clustering procedures which use the pixel as a unit and compare each pixel value with every other pixel value can require excessively large computation time because of the large number of pixels in an image. Iterative partition rearrangement schemes, such as ISODATA, have to go through the image data set many times and if done without sampling can also take excessive computation time. Histogram mode seeking, because it requires only one pass through the data, probably involves the least computation time of the measurement space clustering techniques, and it is the one we discuss here.

Histogram mode seeking is a measurement space clustering process in which it is assumed that homogeneous objects on the image manifest themselves as the clusters in measurement space. Image segmentation is accomplished by mapping the clusters back to the image domain where the maximal connected components of the mapped back clusters constitute the image segments. For images which are single band images, calculation of this histogram in an array is direct. The measurement space clustering can be accomplished by determining the valleys in this histogram and declaring the clusters to be the interval of values between valleys. A pixel whose value is in the i th interval is labeled

with index i and the segment it belongs to is one of the connected components of all pixels whose label is i .

Figure 1 illustrates an example image which is the right kind of image for the measurement space clustering image segmentation process. It is an enlarged image of a polished mineral ore section. The width of the field is about 1mm. The ore is from Ducktown, Tennessee and shows subhedral to euhedral pyrite porphyroblests (white) in a matrix of pyrrhotite (gray).

This example image is one which is not ideal for measurement space clustering image segmentation. Figure 2 shows its histogram which has three modes and two valleys, and Figure 3 shows the corresponding segmentation. Notice the multiple boundary area. It is apparent that the boundary between the grain and background is in fact shaded dark and there are many such border regions which show up as dark segments. In this case, we do not desire the edge borders to be separate regions and although the segmentation procedure did exactly as it should have done the results are not what we desired. This illustrates that segmentation into homogeneous regions is not necessarily a good solution to a segmentation problem.

Figure 4 illustrates an image of a section of an F-15 bulkhead. It is clear that the image has distinct parts such as webs and ribs. Figure 5 shows the histogram of this image. It has two well separated modes. The narrow one on the right with a long left tail corresponds to specular reflection points. The main mode has three valleys on its left side and two valleys on its right side. Defining the depth of a valley to be the probability difference between the valley bottom and the lowest valley side and eliminating the two shallowest valleys produces the segmentation shown in Figure 6. The problem in the segmentation is apparent. Since the clustering was done in measurement space, there was no requirement for good spatial continuation and the resulting boundaries are very noisy and busy. Separating the main mode to its two most dominant submodes produces the segmentation of Figure 7. Here the boundary noise is less, the resulting regions more satisfactory, but the detail provided is much less.

Ohlander (1975) refines the clustering idea in a recursive way. He begins by defining a mask selecting all pixels on the image. Given any mask, a histogram of the masked image is computed. Measurement space clustering enables the separation of one mode of the histogram set from another mode. Pixels on the image are then identified with the cluster to which they belong. If there is only one measurement space cluster, then the mask is terminated. If there is more than one cluster, then each connected component of all pixels with the same cluster is, in turn, used to generate a mask which is placed on a mask stack. During successive iterations the next mask in the stack selects pixels in the histogram computation process. Clustering is repeated for each new mask until the stack is empty.

Figure 8 illustrates this process which we call a recursive histogram directed spatial clustering. Figure 9 illustrates a recursive histogram directed spatial clustering technique applied to the bulkhead image of Figure 4. It produces a result with boundaries being somewhat busy and many

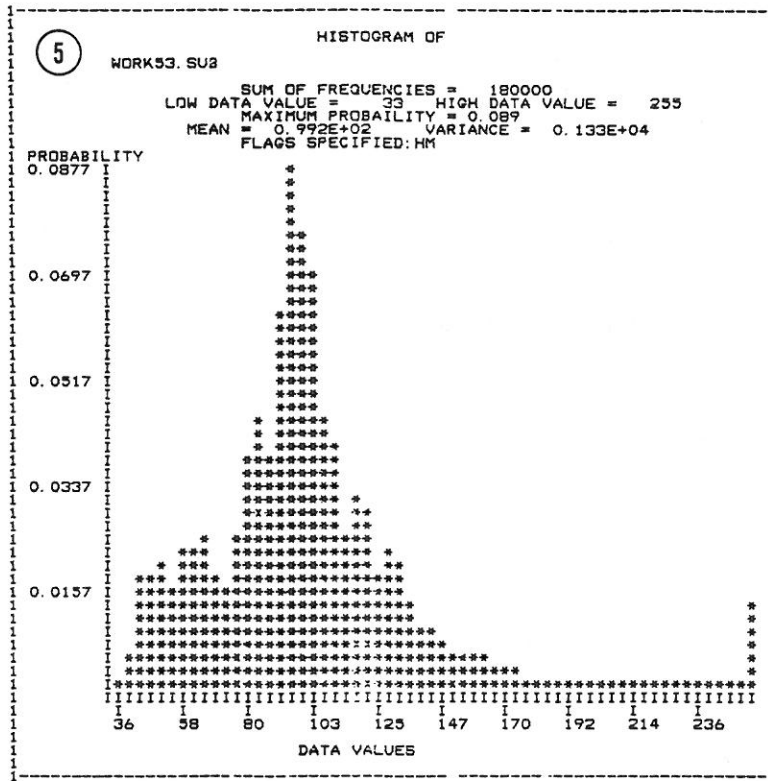
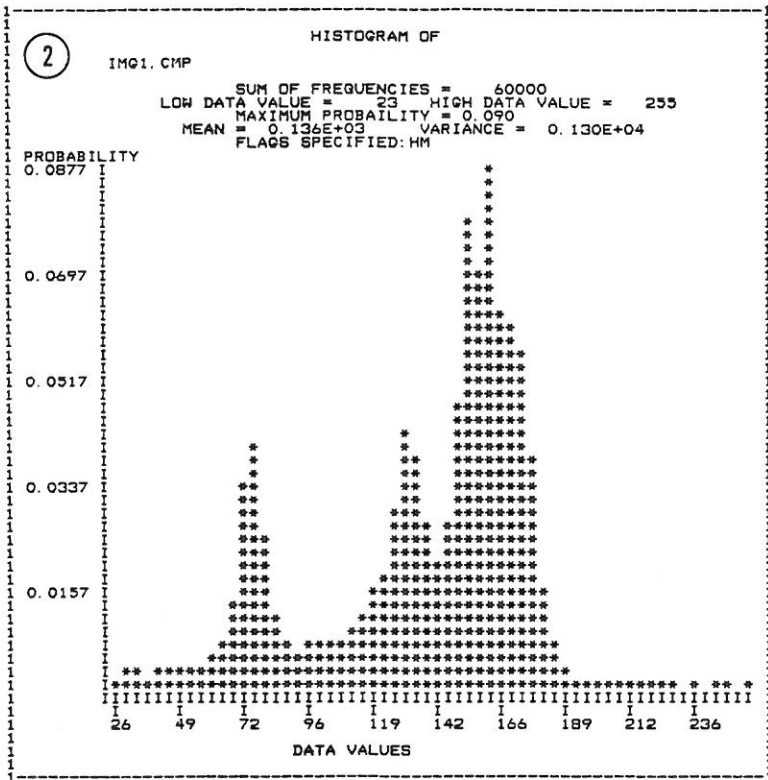
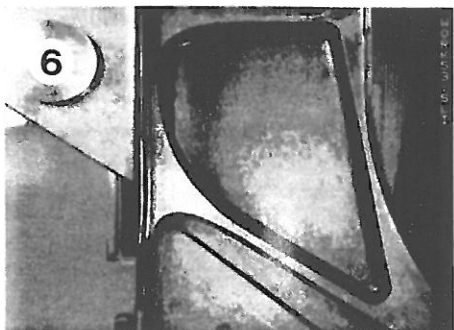
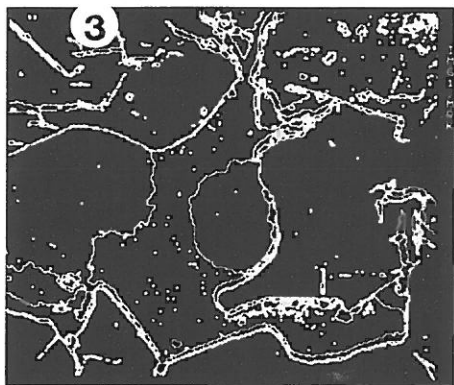
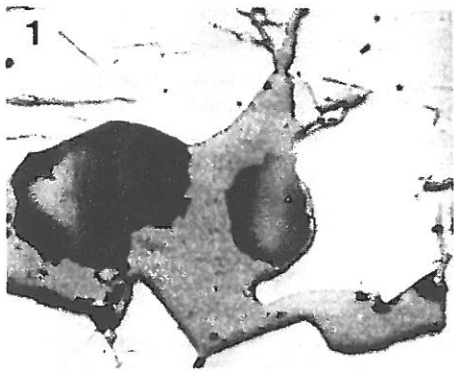


Figure 4 is an image of a section of an F-15 bulkhead.

Figure 5 is a histogram of the bulkhead image of Fig.4

Figure 6 is the segmentation of the bulkhead induced by a measurement space clustering into 5 clusters.

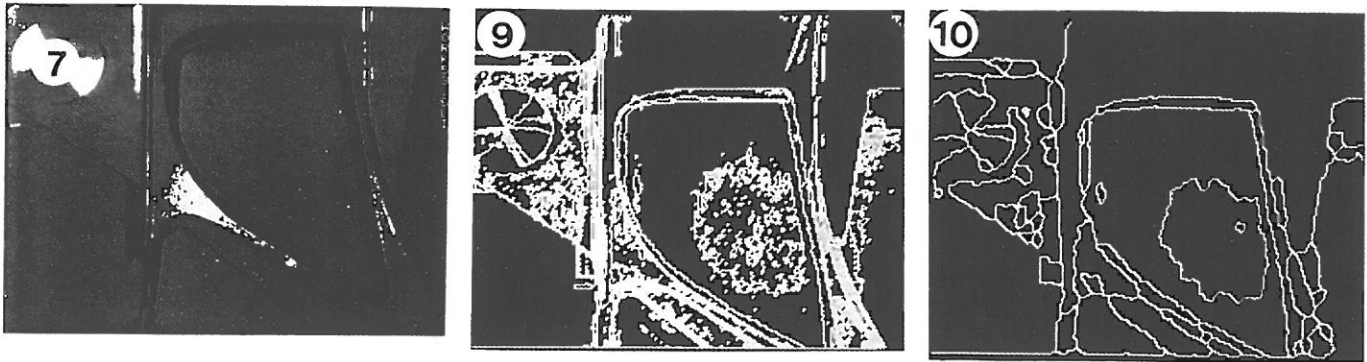


Figure 1 is an enlarged raw mineral ore section; the bright areas are grains of pyrite; the gray areas constitute a matrix of pyrrhotite; the black areas are holes. Because some of the boundaries between regions are shadowed, homogeneous region segmentation may not produce the desired segmentation.

Figure 2 shows the histogram of the image in Fig. 1.

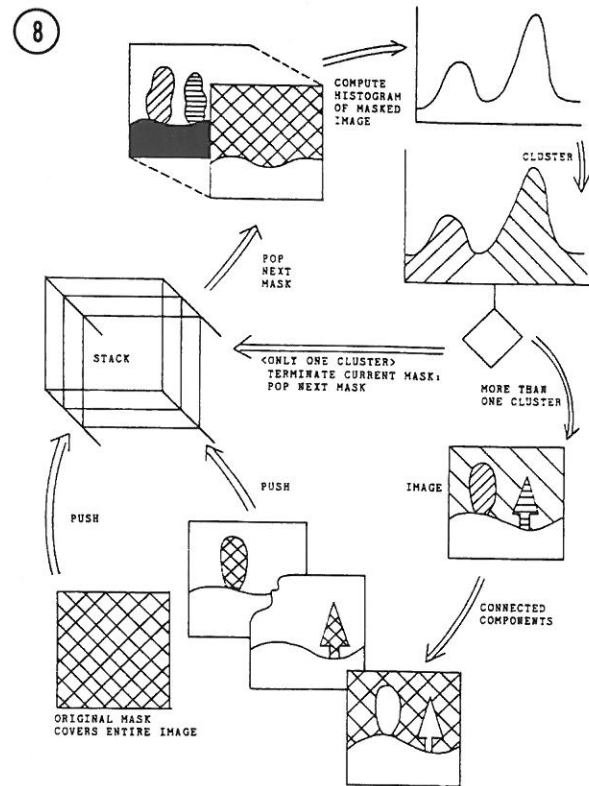
Figure 3 shows the segmentation produced by clustering the histogram of Fig. 2.

Figure 7 shows the segmentation of the bulkhead induced by a measurement space clustering into 3 clusters.

Figure 8 depicts the recursive histogram directed spatial clustering scheme of Ohlander.

Figure 9 shows the results of the histogram directed spatial clustering when applied to the bulkhead image.

Figure 10 shows the results of performing an 8-connected shrink operation followed by a fill operation on the segmentation of Fig. 9.



small regions in areas of specular reflectance. Figure 10 illustrates the results of performing a morphological opening with a 3×3 square structuring element on the segmentation of Figure 9. The tiny regions are removed in this manner, but several important long, thin regions are also lost.

2.2 Thresholding

If the image contains a bright object against a dark background and the measurement space is one-dimensional, measurement space clustering amounts to determining a threshold such that all points smaller than or equal to the threshold are assigned to one cluster and the remaining points are assigned to the second cluster. In the easiest cases a procedure to determine the threshold need only examine the histogram and place the threshold in the valley between the two modes. Unfortunately, it is not always the case that the two modes are nicely separated by a valley. To handle this kind of situation a variety of techniques can be used to combine the spatial information on the image with the gray tone intensity information to help in threshold determination.

Chow and Kaneko (1972) suggest using a threshold which depends on the histogram determined from the spatially local area around the pixel to which the threshold applies. Thus, for example, a neighborhood size of 33 by 33 or 65 by 65 can be used to compute the local histogram. Chow and Kaneko avoided the local histogram computation for each pixel's neighborhood by dividing the image into mutually exclusive blocks, computing the histogram for each block, determining an appropriate threshold for each histogram and then spatially interpolating the threshold values to obtain a spatially adaptive threshold for each pixel.

Weszka, Nagel, and Rosenfeld (1974) suggest determining a histogram for only those pixels having a high Laplacian magnitude. They reason that there will be a shoulder of the gray tone intensity function at each side of the boundary. The shoulder has high Laplacian magnitude. A histogram of all shoulder pixels will be a histogram of all interior pixels

just next to the interior border of the region. It will not involve those pixels in between regions which help make the histogram valley shallow. It will also have a tendency to involve equal numbers of pixels from the object and from the background. This makes the two histogram modes about the same size. Thus the valley seeking method for threshold selection has a chance of working on the new histogram.

Weszka and Rosenfeld (1978) describe one method for segmenting white blobs against a dark background by a threshold selection based on busyness. For any threshold, busyness is the percentage of pixels having a neighbor whose thresholded value is different from their own thresholded value. A good threshold is that point near the histogram valley between the two peaks which minimizes the busyness.

Watanabe (1974) suggests choosing a threshold value which maximizes the sum of gradients taken over all pixels whose gray level equals the threshold value. Kohler (1981) suggests a modification of the Watanabe idea. Instead of choosing a threshold which maximizes the sum of gradient magnitudes taken over all pixels whose gray tone intensity equals the threshold value, Kohler suggests choosing that threshold which detects more high contrast edges and fewer low contrast edges than any other threshold.

Kohler defines the set $E(T)$ of edges detected by a threshold T to be the set of all pairs of neighboring pixels one of whose gray tone intensities is less than or equal to T and one of whose gray tone intensities is greater than T :

$$E(T) = \{((i, j), (k, l)) \mid \begin{aligned} &(1) \text{ pixels } (i, j) \text{ and } (k, l) \text{ are neighbors} \\ &(2) \min \{I(i, j), I(k, l)\} \leq T < \max \{I(i, j), I(k, l)\} \end{aligned}\} \quad (1)$$

The total contrast $C(T)$ of edges detected by threshold T is given by

$$C(T) = \sum_{((i, j), (k, l)) \in E(T)} \min\{|I(i, j) - T|, |I(k, l) - T|\} \quad (2)$$

The average contrast of all edges detected by threshold T is then given by $C(T)/\#E(T)$. The best threshold T_b is determined by that value is maximizes $C(T_b)/\#E(T_b)$.

Milgram and Herman (1979) reason that pixels which are in between regions probably have in between gray tone intensities. If it is these pixels which are the cause of the shallow valleys, then it should be possible to eliminate their effect by only considering pixels having small gradients. They take this idea further and suggest that by examining clusters in the two-dimension measurement space consisting of gray tone intensity and gradient magnitude, it is even possible to determine multiple thresholds when more than one kind of object is present.

Panda and Rosenfeld (1978) suggest a related approach for segmenting a white blob against a dark background. Consider the histogram of gray levels for all pixels which have small gradients. If a pixel has a small gradient, then it is not likely for it to be an edge. If it is not an edge, then it is either a dark background pixel or a bright blob pixel. Hence, the histogram of all pixels having small gradients will be bimodal and for pixels with small gradients, the valley

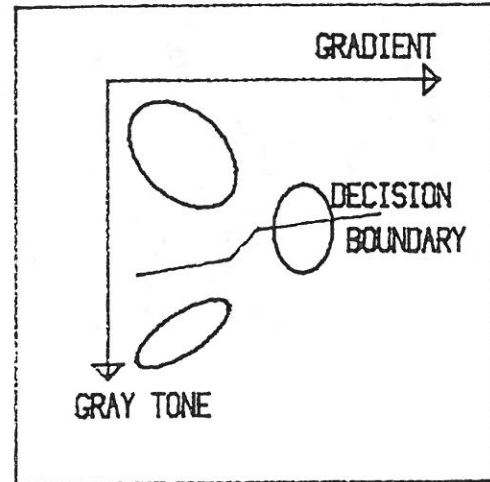


Figure 11. Illustrates how the threshold of the Panda and Rosenfeld technique depends on the gradient magnitude.

between the two modes of the histogram is an appropriate threshold point. Next consider the histogram of gray levels for all pixels which have high gradients. If a pixel has a high gradient, then it is likely for it to be an edge. If it is an edge separating a bright blob against a dark background and if the separating boundary is not sharp but somewhat diffuse, then the histogram will be unimodal, the mean being a good threshold separating the dark background pixels from the bright blob pixels. Thus Panda and Rosenfeld suggest determining two thresholds: one for low gradient pixels and one for high gradient pixels. By this means they perform the clustering in the two dimensional measurement space consisting of gray tone intensity and gradient. The form of the decision boundary in the two dimensional measurement space is shown in Figure 11.

Figure 12 illustrates a FLIR image from the NATO data base which one might think has the right characteristics

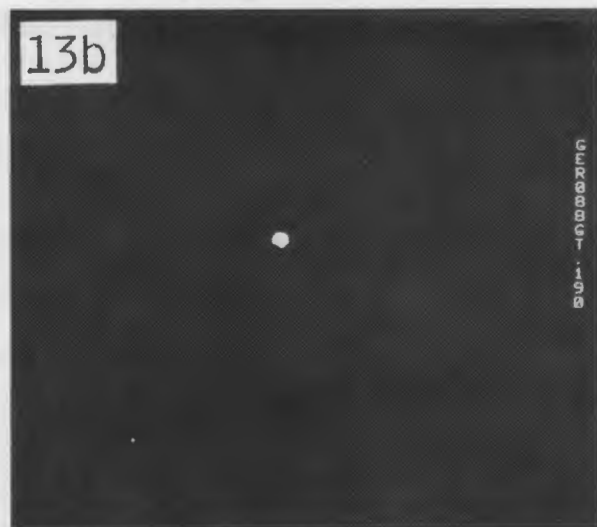
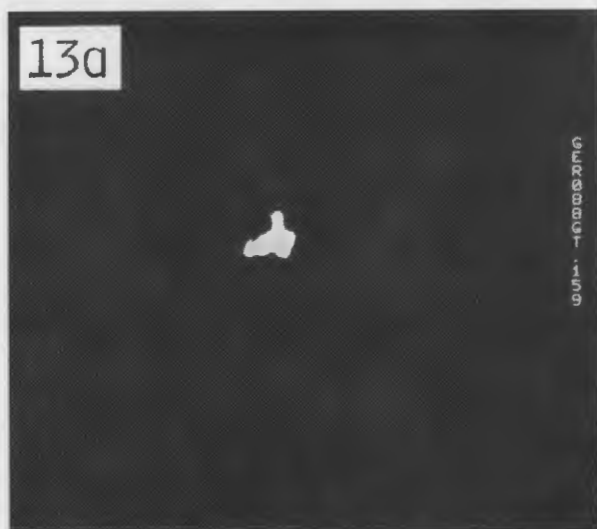
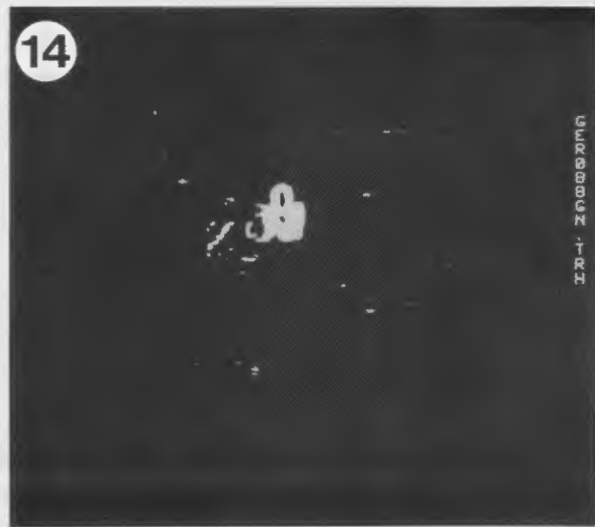
Figure 12 shows a FLIR image from the NATO data base. To reduce noise it was filtered with a Gaussian filter with a sigma of 1.5 and neighborhood size of 15.

Figures 13a & 13b show the FLIR image of Fig. 12 thresholded at gray tone intensity 159(1) and 190 (b).

Figure 14 shows the pixels of the FLIR image having large gradient magnitude.

Figure 15 shows a scattergram of the gray tone intensity-gradient measurement space for the image of Fig. 12. The gray tone intensity is the y axis and the gradient is the x axis. Notice the nicely bimodal gray tone intensity distribution for small gradient magnitude.

Figure 16 shows the segmentation of the image in Fig. 12 using the Panda and Rosenfeld scheme.



for this type of segmentation algorithm. Figures 13a and 13b illustrate the FLIR image thresholded at 159 and 190, respectively. Figure 14 shows the pixels having a large gradient magnitude, where the gradient is computed as the square root of the sum of the squares of the linear coefficients arising from a gray tone intensity cubic fit in a 7×7 window. Figure 15 shows the horseshoe shaped cluster in the two-dimensional gray tone intensity-gradient space where the gray tone intensities and the gradient values have been equal interval quantized.

Figure 16 illustrates the resulting segmentation. Notice that because there is a bright object with a slightly darker appendage on top, the assumption of a homogenous object on a dark background is not met. The result is that only the boundary of the appendage is picked up. A survey of threshold techniques can be found in Weszka (1978).

2.3 Multidimensional Space Clustering

For multiband images such as LANDSAT or Thematic Mapper, determining the histogram in a multi-dimensional space is not feasible. For example, in a six band image where each band has intensities between 0 and 99, the array would have to have $100^6 = 10^{12}$ locations. A large image might be 10,000 pixels per row by 10,000 rows. This only constitutes 10^8 pixels, a sample too small to estimate probabilities in a space of 10^{12} values were it not for some constraints of reality: (1) there is typically a high correlation between the band-to-band pixel values, and (2) there is a large amount of spatial redundancy in image data. Both these factors create a situation in which the 10^8 pixels can be expected to contain only between 10^4 and 10^5 distinct 6-tuples. Based on this fact, the counting required for the histogram is easily done by hashing the 6-tuples into an array.

Clustering using the multidimensional histogram is more difficult than univariate histogram clustering. Goldberg and Shlien (1977, 1978) threshold the multidimensional histogram to select all N -tuples situated on the most prominent modes. Then they perform a measurement space connected components on these N -tuples to collect together all the N -tuples in the top of the most prominent modes. These measurement space connected sets form the cluster cores. The clusters are defined as the set of all N -tuples closest to each cluster core.

An alternate possibility (Narendra and Goldberg, 1977) is to locate peaks in the multi-dimensional measurement space and region grow around them, constantly descending from each peak. The region growing includes all successive neighboring N -tuples whose probability is no higher than the N -tuple from which it is growing. Adjacent mountains meet in their common valleys.

Rather than accomplish the clustering in the full measurement space, it is possible to work in multiple lower order projection spaces and then reflect these clusters back to the full measurement space. Suppose, for example, that the clustering is done on a four band image. If the clustering done in bands 1 and 2 yields clusters c_1, c_2, c_3 and the clustering done in bands 3 and 4 yields clusters c_4 and c_5 then each possible 4-tuple from a pixel can be given a cluster label from the set $\{(c_1, c_4), (c_1, c_5), (c_2, c_4), (c_2, c_5), (c_3, c_4), (c_3, c_5)\}$. A 4-tuple (x_1, x_2, x_3, x_4) gets the cluster label (c_2, c_4) if

(x_1, x_2) is in cluster c_2 and (x_3, x_4) is in cluster c_4 .

3. Region Growing

3.1 Single Linkage Region Growing

Single linkage region growing schemes regard each pixel as a node in a graph. Neighboring pixels whose properties are similar enough are joined by an arc. The image segments are maximal sets of pixels all belonging to the same connected component. Single linkage image segmentation schemes are attractive for their simplicity. They do, however, have a problem with chaining, because it takes only one arc leaking from one region to a neighboring one to cause the regions to merge.

The simplest single linkage scheme defines similar enough by pixel difference. Two neighboring pixels are similar enough if the absolute value of the difference between their gray tone intensity values is small enough. Bryant (1979) defines similar enough by normalizing the difference by the quantity (square root of 2) times the root mean square value of neighboring pixel differences taken over the entire image.

For pixels having vector values, the obvious generalization is to use a vector norm of the pixel difference vector. Instead of using a Euclidean distance, Asano and Yokoya (1981) suggest that two pixels be joined together if the absolute value of their difference is small enough compared to the average absolute value of the center pixel minus neighbor pixel for each of the neighborhoods the pixels belong to. The ease with which unwanted region chaining can occur with this technique limits its potential on complex or noisy data.

3.2 Hybrid Linkage Region Growing

Hybrid single linkage techniques are more powerful than the simple single linkage technique. The hybrid techniques seek to assign a property vector to each pixel where the property vector depends on the $K \times K$ neighborhood of the pixel. Pixels which are similar are so because their neighborhoods in some special sense are similar. Similarity is thus established as a function of neighboring pixel values and this makes the technique better behaved on noisy data.

One hybrid single linkage scheme relies on an edge operator to establish whether two pixels are joined with an arc. Here an edge operator is applied to the image labeling each pixel as edge or non-edge. Neighboring pixels, neither of which are edges, are joined by an arc. The initial segments are the connected components of the non-edge labeled pixels. The edge pixels can either be left assigned edges and be considered as background or they can be assigned to the spatially nearest region having a label.

The quality of this technique is highly dependent on the edge operator used. Simple operators such as the Roberts and Sobel operators may provide too much region linkage, for a region cannot be declared as a segment unless it is completely surrounded by edge pixels. Haralick and Dinstein (1975), however, do report some success using this technique on LANDSAT data. They perform a region

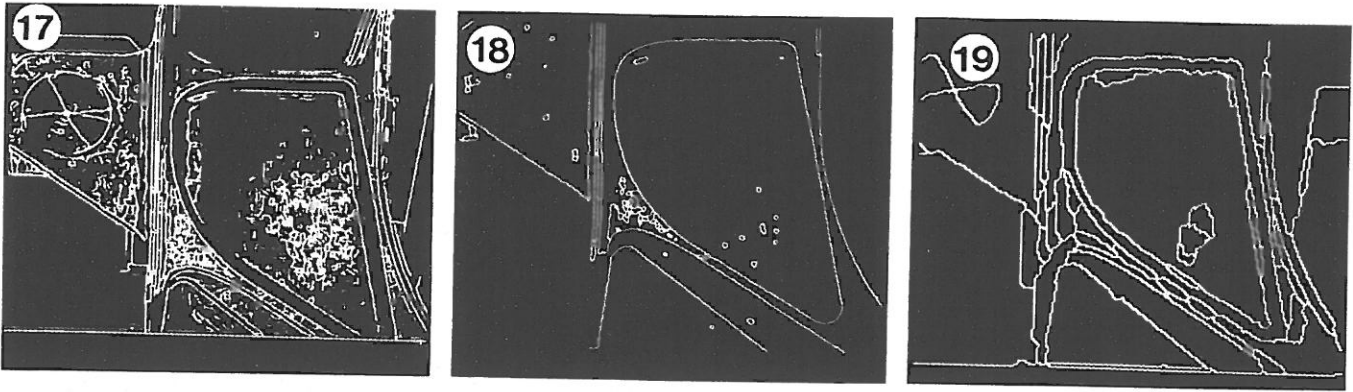


Figure 17 shows the second directional derivative zero-crossing operator using a gradient threshold of 4, a 9×9 neighborhood and a zero-crossing radius of 0.85 applied to the bulkhead image of Fig. 4.

Figure 18 shows a hybrid linkage region growing scheme in which any pair of neighboring pixels, neither of which are edge pixels can link together. The resulting segmentation consists of the connected components of the non-edge pixels and where edge pixels are assigned to their nearest connected component. This result was obtained from the edge image of Fig. 17.

Figure 19 shows the results of overlaying an edge image on the segmentation (Fig. 15) of the bulkhead image of Fig. 4, applying a connected components operator, and removing small regions.

growing of the edge pixels in order to close gaps before performing the connected components operator. Perkins (1980) uses a similar technique.

Haralick (1982, 1984) discusses a very sensitive zero-crossing of second directional derivative edge operator. In this technique, each neighborhood is least squares fitted with a cubic polynomial in two variables. The first and second partial derivatives are easily determined from the polynomial. The first partial derivatives at the center pixel determine the gradient direction. With the direction fixed to be the gradient direction, the second partials determine the second directional derivative. If the gradient is high enough and if in the gradient direction, the second directional derivative has a negatively sloped zero-crossing inside the pixel's area, then an edge is declared in the neighborhood's center pixel.

Figure 17 shows the edges resulting from the second directional derivative zero-crossing operator using a gradient threshold of 4, a 9×9 neighborhood, and a zero-crossing radius of .85. The edges are well placed and a careful examination of pixels on perceived boundaries which are not classified as edge pixels will indicate the step edge pattern to be either non-existent or weak. A connected components of the non-edge pixels accomplishes the initial segmentation. After the connected components operation, the edge pixels are assigned to their spatially closest component by a region filling operation. Figure 18 shows the boundaries

from the region filled image. Obviously, there are some regions which have been merged together. However, those boundaries which are present are placed correctly and they are reasonably smooth. Lowering the gradient threshold of the edge operator could produce an image with more edges and thereby reduce the edge gap problem. But this solution does not really solve the gap problem in general.

Although the connected components of the non-edge pixels of the edge operator may not yield an adequate segmentation, the edges can be used in combination with other segmentation techniques to improve the segmentation derived from those techniques. For example, the Ohlander technique discussed in Section II segments on the basis of measurement space clusters, but can lose some important edges. Such a segmentation will be improved if the edges can be used to refine it. To this end, we started with the Ohlander segmentation and overlaid the binary edge image shown in Fig. 18. This was achieved by starting with a symbolic image of the segmentation (where each pixel's value is its region number) and multiplying this image on a pixel by pixel basis with the binary edge image. The resultant image has pixels with value zero where the edges were, and all the rest of the pixels still show their region numbers. We then applied a connected components operator to the non-edge pixels of this image, and removed small regions. The result is shown in Figure 19. Note that Figure 19 shows a better segmentation than Figure 15 because several important straight lines that were missing have been put in. However, this technique does not always work as well. The same approach applied to the Ohlander segmentation of an image which had more specular reflection was less effective. There were many extraneous edges which were put back into the segmentation.

Yakimovsky (1976) assumes regions are normally distributed and uses a maximum likelihood test to determine edges. Edges are declared to exist between pairs of contiguous and exclusive neighborhoods if the hypothesis that their means are equal and their variances are equal has to be rejected. For any pair of adjacent pixels with mutually exclusive neighborhoods R_1 and R_2 having $N_1 - 1$ and N_2 pixels, respectively, the maximum likelihood technique computes the mean

$$\bar{X}_i = \frac{1}{N_i} \sum_{x \in R_i} \quad (3)$$

and the scatter

$$S_i = \sum_{X \in R_i} (X - \bar{X}_i)^2 \quad (4)$$

as well as the grand mean

$$\bar{X} = \frac{1}{N_1 + N_2} \sum_{X \in R_1 \cup R_2} X \quad (5)$$

and grand scatter

$$S_2 = \sum_{X \in R_1 \cup R_2} (X - \bar{X})^2. \quad (6)$$

The likelihood ratio test statistic T is given by

$$T = \frac{[S^2 / (N_1 + N_2)]^{N_1 + N_2}}{[S_1^2 / N_1]^{N_1} [S_2^2 / N_2]^{N_2}}. \quad (7)$$

Edges are declared between any pair of adjacent pixels when the T statistic from their neighborhoods is high enough. As N_1 and N_2 get large, $2 \log T$ is asymptotically distributed as a chi-squared variate with 2 degrees of freedom.

If it can be assumed that the variances of the two region are identical, then the statistic

$$F = \frac{(N_1 + N_2 - 2)N_1N_2(\bar{X}_1 - \bar{X}_2)^2}{N_1 + N_2(S_1^2 + S_2^2)} \quad (8)$$

has an F distribution with 1 and $N_1 + N_2 - 2$ degrees of freedom under the hypothesis that the means of the regions are equal. For an F value which is sufficiently large, the hypothesis can be rejected and an edge declared to exist between the regions.

Haralick (1980) suggests fitting a plane to the neighborhood around the pixel, and testing the hypothesis that the slope of the plane is zero. Edge pixels correspond to pixels between neighborhoods in which the zero slope hypothesis has to be rejected. To determine a roof or V-shaped edge, Haralick suggests fitting a plane to the neighborhoods on either side of the pixel and testing the hypothesis that the coefficients of fit, referenced to a common framework, are identical.

Another hybrid technique first used by Levine and Leemet (1976) is based on the Jarvis and Patrick (1973) shared nearest neighbor idea. Using any kind of reasonable notion for similarity, each pixel examines its $K \times K$ neighborhood and makes a list of the N pixels in the neighborhood most similar to it. Call this list the similar neighbor list, where we understand neighbor to be any pixel in the $K \times K$ neighborhood. An arc joins any pair of immediately neighboring pixels if each pixel is in the other's shared neighbor list and if there are enough pixels common to their shared neighbor lists, that is, if the number of shared neighbors is high enough.

To make the shared neighbor technique work well each pixel can be associated with a property vector consisting of its own gray tone intensity and a suitable average of the gray tone intensity of pixels in its $K \times K$ neighborhood. For example, we can have (x, a) and (y, b) denote the property vectors for two pixels if x is the gray tone intensity value and a is the average gray tone intensity value in

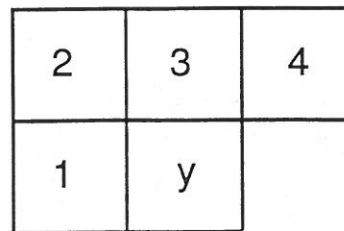


Figure 20 illustrates the region growing geometry for one pass scan left-right, top-bottom region growing. Pixel i belongs to region R_i whose mean is \bar{X}_i , $i = 1, 2, 3$, and 4. Pixel y is added to a region R_j if by a T -test the difference between y and \bar{X}_j is small enough. If for two regions R_i and R_j , the difference is small enough, and if the difference between \bar{X}_i and \bar{X}_j is small enough, regions R_i and R_j are merged together, and y is added to the merged region. If the difference between \bar{X}_i and \bar{X}_j is significantly different, then y is added to the closest region.

the neighborhood of the first pixel and b is the average gray tone intensity value in the neighborhood of the second pixel. Similarity can be established by computing

$$S = w_1(x - y)^2 + w_2(x - b)^2 + w_3(y - a)^2, \quad (9)$$

where w_1, w_2 , and w_3 are non-negative weights. The pixels are called similar enough for small enough values of S .

Pong et al. (1984) suggest an approach to segmentation based on the facet model of images. The procedure starts with an initial segmentation of the image into small regions. Each region with associated property vector is considered a unit. In a series of iterations, the property vector of a region is replaced by a property vector that is a function of its neighboring regions. Then adjacent regions having similar final property vectors are merged. This gives a new segmentation which can then be used as input to the algorithm. Thus a sequence of coarser and coarser segmentations are produced. Useful variations are to prohibit merging across strong edge boundaries or when the variance of the combined region would become too large.

3.3 Centroid Linkage Region Growing

In centroid linking region growing, in contrast with single linkage region growing, pairs of neighboring pixels are not compared for similarity. Rather, the image is scanned in some predetermined manner such as left-right top-bottom. A pixel's value is compared to the mean of an already existing but not necessarily completed neighboring segment. If its value and the segment's mean value are close enough, then the pixel is added to the segment and the segment's mean is updated. If there is more than one region which is close enough, then it is added to the closest region. However, if the means of the two competing regions are close enough, the two regions are merged and the pixel is added to the merged region. If no neighboring region has its mean close enough, then a new segment is established having the given pixel's value as its first member. Figure 20 illustrates the geometry of this scheme.

Keeping track of the means and scatters for all region as they are being determined does not require large amounts of memory space. There cannot be more regions active at one time than the number of pixels in a row of the image. Hence, a hash table mechanism with the space of a small multiple of the number of pixels in a row can work well.

Another possibility is a single band region growing technique using the T -test. Let R be a segment of N pixels neighboring a pixel with gray tone intensity y . Define the mean \bar{X} and scatter S^2 by

$$\bar{X} = \frac{1}{N} \sum_{(r,c) \in R} I(r,c) \quad (10)$$

and

$$S^2 = \sum_{(r,c) \in R} (I(r,c) - \bar{X})^2. \quad (11)$$

Under the assumption that all the pixels in R and the test pixel y are independent and identically distributed normals, the statistic

$$T = \left[\frac{(N-1)N}{(N+1)} (y - \bar{X})^2 / S^2 \right]^{\frac{1}{2}} \quad (12)$$

has a T_{N-1} distribution. If T is small enough y is added to region R and the mean and scatter are updated using y . The new mean and scatter are given by

$$\bar{X}_{\text{new}} \leftarrow (N\bar{X}_{\text{old}} + y) / (N + 1) \quad (13)$$

and

$$S_{\text{new}}^2 \leftarrow S_{\text{old}}^2 + (y - \bar{X})^2 + N(\bar{X}_{\text{new}} - \bar{X}_{\text{old}})^2. \quad (14)$$

If T is too high the value y is not likely to have arisen from the population of pixels in R . If y is different from all of its neighboring regions then it begins its own region. A slightly stricter linking criterion can require that not only must y be close enough to the mean of the neighboring regions, but that a neighboring pixel in that region must have a close enough value to y . This combines a centroid linkage and single linkage criterion. The next section discusses a more powerful combination technique, but first we want to develop the concept of "significantly high."

To give a precise meaning to the notion of too high a difference, we use an α level statistical significance test. The fraction α represents the probability that a T statistic with $N - 1$ degrees of freedom will exceed the value $t_{N-1}(\alpha)$. If the observed T is larger than $t_{N-1}(\alpha)$, then we declare the difference to be significant. If the pixel and the segment really come from the same population, the probability that the test provides an incorrect answer is α .

The significance level α is a user-provided parameter. The value of $t_{N-1}(\alpha)$ is higher for small degrees of freedom and lower for larger degrees of freedom. Thus, region scatters considered to be equal, the larger a region is, the closer a pixel's value has to be to the region's mean in order to merge into the region. This behavior tends to prevent already large regions from attracting to it many other additional pixels and tends to prevent the drift of the region mean as the region gets larger. Figure 21 plots $t_M(\alpha)$

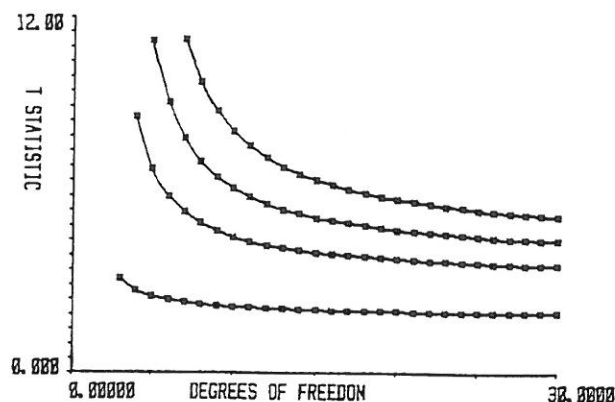


Figure 21 illustrates how the T -test threshold changes as a function of its degrees of freedom for a fixed significance level.

as a function of M , the number of degrees of freedom for a few different values of α .

Note that all regions initially begin as one pixel in size. To avoid the problem of division by 0 (for S^2 is necessarily 0 for one pixel regions and 0 for regions having identically valued pixels) a small positive constant can be added to S^2 . One convenient way of determining the constant is to decide on a prior variance $V > 0$ and an initial segment size N . The initial scatter for a new one-pixel region is then given by NV and the new initial region size is given by N . This mechanism keeps the degrees of freedom of the T -statistic high enough so that a significant difference is not the huge difference required for a T -statistic with a small number of degrees of freedom. Figure 22 illustrates the resulting segmentation of the bulkhead image for a .2% significance level test after all region smaller than 25 pixels have been removed.

Pavlidis (1972) suggests a more general version of this idea. Given an initial segmentation where the regions are approximated by some functional fit guaranteed to have a small enough error, pairs of neighboring regions can be merged if for each region, the sum of the squares of the differences between the fitted coefficients for this region and the corresponding averaged coefficients, averaged over both regions, is small enough. Pavlidis gets his initial segmentation by finding the best way to divide each row of the image into segments with a sufficiently good fit. He also describes a combinatorial tree search algorithm to accomplish the merging which guarantees a best result. Kettig and Landgrebe (1975) successively merge small image blocks using a statistical test. They avoid much of the problem of zero-scatter by considering only cells containing a 2×2 block of pixels.

Gupta, Kettig, Landgrebe, and Wintz (1973) suggest using a T -test based on the absolute value of the difference between the pixel and the nearest region as the measure of dissimilarity. Kettig and Landgrebe (1975) discuss the multi-band situation leading to the F -test and report good success with LANDSAT data.

Nagy and Tolaba (1972) just examine the absolute value

between the pixel's value and the mean of a neighboring region formed already. If this distance is small enough, the pixel is added to the region. If there is more than one region, then the pixel is added to that region with smallest distance.

The Levine and Shaheen scheme (1981) is similar. The difference is that Levine and Shaheen attempt to keep regions more homogeneous and try to keep the region scatter from getting too high. They do this by requiring the differences to be more significant before a merge takes place if the region scatter is high. For a user specified value θ , they define a test statistic T where

$$T = |y - X_{\text{new}}| - (1 - S/\bar{X}_{\text{new}})\theta \quad (15)$$

If $T < 0$ for the neighboring region R in which $|y - \bar{X}|$ is the smallest, then y is added to R . If $T > 0$ for the neighboring region in which $|y - \bar{X}|$ is the smallest, then y begins a new region. Readers of the Levine and Shaheen paper should note that there are misprints in the formulas given for region scatter and region scatter updating.

Brice and Fennema (1970) accomplish the region growing by partitioning the image into initial segments of pixels having identical intensity. They then sequentially merge all pairs of adjacent regions if a significant fraction of their common border has a small enough intensity difference across it.

Simple single-pass approaches which scan the image in a left-right, top-down manner are, of course, unable to make the left and right sides of a V -shaped region belong to the same segment. To be more effective, the single pass must be followed by some kind of connected components merging algorithm in which pairs of neighboring regions having means which are close enough are combined into the same segment. This is easily accomplished by using the two pass label propagation logic of the Lumia et.al. (1983) connected components algorithm.

After the top-bottom, left-right scan, each pixel has already been assigned a region label. In the bottom-up, right-left scan, the means and scatters of each region can be recomputed and can be kept in a hash table. Whenever a pair of pixels from different regions neighbor one another, a T -test can check for the significance of the difference between the region means. If the means are not significant, then they can be merged. A slightly stricter criterion would insist not only that the region means be similar, but also that the neighboring pixels from the different regions must be similar enough. Figure 23 shows the resulting segmentation of the bulkhead image for a .2% significance level after one bottom-up, right-left merging pass and after all regions smaller than 25 pixels have been removed.

One potential problem with region growing schemes is their inherent dependence on the order in which pixels and regions are examined. A left-right, top-down scan does not yield the same initial regions as a right-left, bottom-up scan or for that matter a column major scan. Usually, however, differences caused by scan order are minor.

4. Hybrid Linkage Combinations

The previous section mentioned the simple combination of centroid linkage and single linkage region growing. In this section we discuss the more powerful hybrid linkage combination techniques.

The centroid linkage and the hybrid linkage can be combined in a way which takes advantage of their relative strengths. The strength of the single linkage is that boundaries are placed in a spatially accurate way. Its weakness is that edge gaps result in excessive merging. The strength of centroid linkage is its ability to place boundaries in weak gradient areas. It can do this because it does not depend on a large difference between the pixel and its neighbor to declare a boundary. It depends on a large difference between the pixel and the mean of the neighboring region to declare a boundary.

The combined centroid hybrid linkage technique does the obvious thing. Centroid linkage is only done for non-edge pixels, that is region growing is not permitted across edge pixels or saying it another way edge pixels are not permitted to be assigned to any region and cannot link to any region. Thus if the parameters of centroid linkage were set so that any difference, however large, between pixel value and region mean was considered small enough to permit merging, the two pass hybrid combination technique would produce a connected components of the non edge pixels. As the difference criterion is made more strict, the centroid linkage will produce boundaries in addition to those produced by the edges.

Figure 24 illustrates a one pass scan combined centroid and hybrid linkage segmentation scheme using a significance level test of .2%. Edge pixels are assigned to their closest labeled neighbor and regions having fewer than 25 pixels are eliminated. Notice that the resulting segmentation is much finer than that shown in Figures 22 and 23. Also the dominant boundaries are nicely curved and smooth. Figure 25 illustrates the two pass scan combined centroid and hybrid linkage region growing scheme using a significance level test of .2%. The regions are somewhat simpler because of the merging done in the second pass.

5. Spatial Clustering

It is possible to determine the image segments by simultaneously combining clustering in measurement space with a spatial region growing. We call such a technique spatial clustering. In essence, spatial clustering schemes combine the histogram mode seeking technique with a region growing or a spatial linkage technique.

Haralick and Kelly (1969) suggest the segmentation be done by first locating, in turn, all the peaks in the measurement space histogram, and then determining all pixel locations having a measurement on the peak. Next, beginning with a pixel corresponding to the highest peak not yet processed, both spatial and measurement space region growing are simultaneously performed in the following

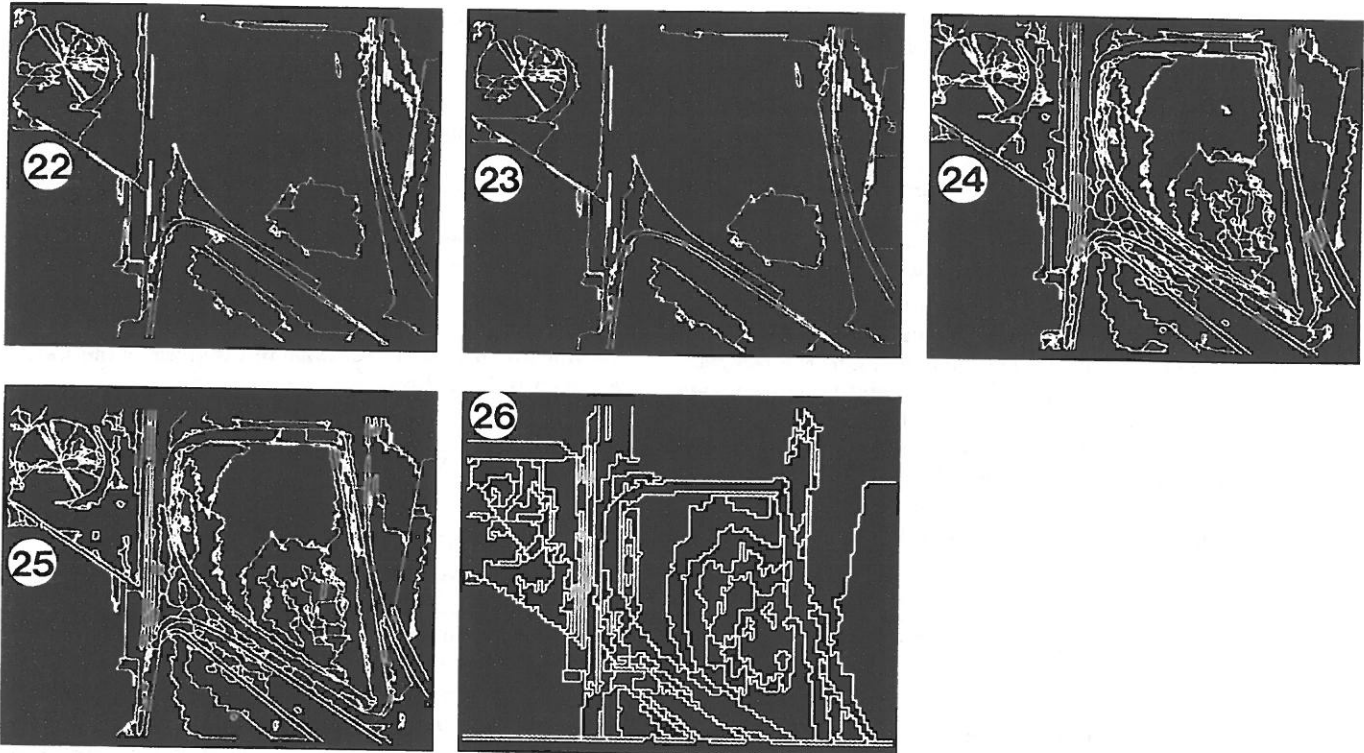


Figure 22 shows the one pass centroid linkage segmentation of the bulkhead image of Fig. 4. A significance level of 0.2% was used.

Figure 23 shows the two pass centroid segmentation of the bulkhead image of Fig. 4. A significance level of 0.2% was used on both passes.

Figure 24 shows the one pass combined centroid and hybrid linkage segmentation of the bulkhead image of Fig. 4. A significance level of 0.2% was used.

Figure 25 shows the two pass combined centroid and hybrid linkage segmentation of the bulkhead image of Fig. 4. A significance level of 0.2% was used on both passes.

Figure 26 shows a split and merge segmentation of the bulkhead image of Fig. 4.

manner. Initially, each segment is the pixel whose value is on the current peak. Consider for possible inclusion into this segment a neighbor of this pixel (in general, the neighbors of the pixel we are growing from) if the neighbor's value (an N -tuple for an N band image) is close enough in measurement space to the pixel's value and if its probability is not larger than the probability of the value of the pixel we are growing from. Matsumoto, Naka, and Yamamoto (1981) discuss a variation on this idea. Milgram (1979) defines a segment for a single band image to be any connected component of pixels, all of whose values lie in some interval I and whose border has a higher coincidence with the border created by an edge operator than for any other interval I . The technique

has the advantage over the Haralick and Kelly technique in that it does not require the difficult measurement space exploring done in climbing down a mountain. However, it does have to try many different intervals for each segment. Extending it to efficient computation in multiband images appears difficult. However, Milgram does report good results of segmenting white blobs against a black background. Milgram and Kahl (1979) discuss embedding this technique into the Ohlander (1975) recursive control structure.

Minor and Sklansky (1981) make more active use of the gradient edge image than Milgram, but restrict themselves to the more constrained situation of small convex-like segments. They begin with an edge image in which each edge pixel contains the direction of the edge. The orientation is so that the higher valued gray tone is to the right of the edge. Then each edge sends out for a limited distance a message to nearby pixels and in a direction orthogonal to the edge direction. The message indicates what is the sender's edge direction. Pixels which pick up these messages from enough different directions must be interior to a segment.

The spoke filter of Minor and Sklansky counts the number of distinct directions appearing in each 3×3 neighborhood. If the count is high enough they mark the center pixel as belonging to an interior of a region. Then the connected components of all marked pixels is obtained. The gradient guided segmentation is then completed by performing a region growing of the components. The region growing must stop at the high gradient pixels, thereby assuring that no undesired boundary placements are made.

Burt, Hong, and Rosenfeld (1981) describe a spatial clustering scheme which is a spatial pyramid constrained

ISODATA kind of clustering. The bottom layer of the pyramid is the original image. Each successive higher layer of the pyramid is an image having half the number of pixels per row and half the number of rows of the image below it. Initial links between layers are established by linking each parent pixel to the spatially corresponding 4×4 block of child pixels. Each pair of adjacent parent pixels has 8 child pixels in common. Each child pixel is linked to a 2×2 block of parent pixels. The iterations proceed by assigning to each parent pixel the average of his child pixels. Then each child pixel compares his value with each of his parent's values and links himself to his closest parent. Each parent's new value is the average of the children to which he is linked, etc. The iterations converge reasonably quickly and for the same reason the ISODATA iterations converge. If the top layer of the pyramid is a 2×2 block of great grandparents, then these are at most 4 segments which are the respective great grandchildren of these 4 great grandparents. Pietikainen and Rosenfeld (1981) extend this technique to segment an image using textural features.

6. Split and Merge

The split method for segmentation begins with the entire image as the initial segment. Then it successively splits each of its current segments into quarters if the segment is not homogeneous enough. Homogeneity can be easily established by determining if the difference between the largest and smallest gray tone intensities is small enough. Algorithms of this type were first suggested by Roberston (1973) and Klinger (1973). Kettig and Landgrebe (1975) try to split all non-uniform 2×2 neighborhoods before beginning the region merging. Fukada (1980) suggests successively splitting a region into quarters until the sample variance is small enough. Efficiency of the split and merge method can be increased by arbitrarily partitioning the image into square regions of a user selected size and then splitting these further if they are not homogeneous.

Because segments are successively divided into quarters, the boundaries produced by the split technique tend to be squarish and slightly artificial. Sometimes adjacent quarters coming from adjacent split segments need to be joined rather than remain separate. Horowitz and Pavlidis (1976) suggest a split and merge strategy to take care of this problem. Muerle and Allen (1968) suggest merging a pair of adjacent regions if their gray tone intensity distributions are similar enough. They recommend the Kolmogorov-Smirnov test. Figure 26 illustrates the result of a Horowitz and Pavlidis type type split and merge segmentation of the bulkhead image.

Chen and Pavlidis (1980) suggest using statistical tests for uniformity rather than a simple examination of the difference between the largest and smallest gray tone intensities in the region under consideration for splitting. The uniformity test requires that there be no significant difference between the mean of the region and each of its quarters. The Chen and Pavlidis tests assume that the variances are equal and known.

Let each quarter have K pixels, X_{ij} be the j th pixel in the i th region, X_i be the mean of the i th quarter and $X_{..}$ be the grand mean of all the pixels in the 4 quarters. Then in order for a region to be considered homogeneous, Chen and Pavlidis require that

$$|X_i - X_{..}| \leq \epsilon, \quad i = 1, 2, 3, 4. \quad (16)$$

We give here the F -test for testing the hypothesis that the mean and variances of the quarters are identical. The value of variance is not assumed known. Under the assumption that the regions are independent and identically distributed normals, the optimal test is given by the statistic F which is defined by

$$F = \frac{K \sum_{i=1}^4 (X_i - X_{..})^2 / 3}{\sum_{i=1}^4 \sum_{k=1}^K (X_{ik} - X_i)^2 / 4(K-1)} \quad (17)$$

It has a $F_{3,4(K-1)}$ distribution. If F is too high the region is declared not uniform.

The data structures required to do a split and merge on images larger than 512×512 are extremely large. Execution of the algorithm on virtual memory computers results in so much paging that the dominant activity may be paging rather than segmentation. Browning and Tanimoto (1982) give a description of a split and merge scheme where the split and merge is first accomplished on mutually exclusive subimage blocks and the resulting segments are then merged between adjacent blocks to take care of the artificial block boundaries.

7. Conclusion

We have briefly surveyed the place of segmentation in vision algorithms as well as common techniques of measurement space clustering, single linkage, hybrid linkage, region growing, spatial clustering, and split and merge used in image segmentation. We have noted that they can be made to be more powerful if they are based on some kind of statistical test for equality of means.

Not discussed as part of image segmentation is the fact that it might be appropriate for some segments to remain apart or to be merged not on the basis of the gray tone distributions, but on the basis of the object sections which they represent. The use of this kind of semantic information in the image segmentation process is essential for the higher level image understanding work. When segmentation problems are specified without the semantic or application information, they tend to be ill-posed problems. Ill-posed means that none of the techniques such as those discussed in this paper are likely to work, because these techniques do not take into account such application specifics. It would be useful for future segmentation work to integrate in a unified way both the syntax of the segmentation with the kinds or general forms of the semantics which are associated with particular applications.

8. References

1. T. Asano and N. Yokoya, Image segmentation schema for low-level computer vision, *Pattern Recognition* **14**, 1981, 267-273.
2. C. Brice and C. Fennema, Scene analysis using regions, *Artificial Intelligence* **1**, 1970, 205-226.
3. J.D. Browning and S.L. Tanimoto, Segmentation of pictures into regions with a tile by tile method, *Pattern Recognition* **15**, 1982, 1-10.
4. P.J. Burt, T.H. Hong, and A. Rosenfeld, Segmentation and estimation of image region properties through cooperative hierarchical computation, *IEEE Trans. Systems Man Cybernet.* SMC-11, 1981.
5. J. Bryant, On the clustering of multidimensional pictorial data, *Pattern Recognition* **11**, 1979, 115-125.
6. P.C. Chen and T. Pavlidis, Image segmentation as an estimation problem, *Comput. Graphics Image Process.* **12**, 1980, 153-172.
7. C.K. Chow and T. Kaneko, Boundary detection of radiographic images by a thresholding method, in *Frontiers of Pattern Recognition* (S. Wanatabe, Ed.), pp. 61-82, Academic Press, New York, 1972.
8. K.S. Fu and J.K. Mui, A survey on image segmentation, *Pattern Recognition* **13**, 1981, 3-16.
9. Y. Fukada, Spatial clustering procedures for region analysis, *Pattern Recognition* **12**, 1980, 395-403.
10. M. Goldberg and S. Shlien, A clustering scheme for multispectral image, *IEEE Trans. Systems Man Cybernet.* SMC-8, 1978, 86-92.
11. M. Goldberg and S. Shlien, A Four-Dimensional Histogram Approach to the Clustering of LAND-SAT Data, in *Machine Processing of Remotely Sensed Data*, IEEE CH 1218-7 MPRSD, Purdue University, West Lafayette, Indiana, June 21-23, 1977, pp. 250-259.
12. J.N. Gupta, R.L. Kettig, D.A. Landgrebe, and P.A. Wintz, Machine boundary finding and sample classification of remotely sensed agricultural data, in *Machine Processing of Remotely Sensed Data*, IEEE 73 CHO 834-2GE, Purdue University, West Lafayette, Indiana, October 16-18, 1973, pp. 4B-25-4B-35.
13. R.M. Haralick and G.L. Kelly, Pattern recognition with measurement space and spatial clustering for multiple image, *Proc. IEEE* **57**, 1969, 654-665.
14. R.M. Haralick and I. Dinstein, A spatial clustering procedure for multi-image data, *IEEE Trans. Circuits and Systems* CAS-22, 1975, 440-450.
15. R. M. Haralick, Edge and region analysis for digital image data, *Comput. Graphics Image Process.* **12**, 1980, 60-73; also in *Image Modeling* (A. Rosenfeld, Ed.), pp. 171-184, Academic Press, New York, 1981.
16. R.M. Haralick, Zero-crossing of second directional derivative edge operator, in *Proceedings of the Society of Photo-Optical Instrumentation Engineers Technical Symposium East*, Arlington, Virginia, Vol. 336, May 3-7, 1982.
17. R.M. Haralick, Digital step edges from zero crossing of second directional derivative, *IEEE Trans. Pattern Anal. Mach. Intell.* PAMI-6, 1984, 58-68.
18. S.L. Horowitz and T. Pavlidis, Picture segmentation by a tree traversal algorithm, *J. Assoc. Comput. Mach.* **23**, 1976, 368-388.
19. R.A. Jarvis and E.A. Patrick, Clustering using a similarity measure based on shared near neighbors, *IEEE Trans. Comput.* C-22, 1973, 1025-1034.
20. T. Kanade, Region segmentation: Signal vs. semantics, *Comput. Graphics Image Process.* **13**, 1980, 279-297.
21. R.L. Kettig and D.A. Landgrebe, Computer classification of multispectral image data by extraction and classification of homogeneous objects, The Laboratory for Application of Remote Sensing, LARS Information Note 050975, Purdue University, West Lafayette, Indiana, 1975.
22. A. Klinger, Data structures and pattern recognition, *Proceedings of the First International Joint Conference on Pattern Recognition*, Washington, D.C., October 1973, pp. 497-498.
23. R. Kohler, A segmentation system based on thresholding, *Comput. Graphics Image Process.* **15**, 1981, 319-338.
24. M.D. Levine and J. Leemet, A method for non-purposive picture segmentation, *Proceedings of the Third International Joint Conference on Pattern Recognition*, 1976.
25. M.D. Levine and S.I. Shaheen, A modular computer vision system for picture segmentation and interpretation, *IEEE Trans. Pattern Anal. Mach. Intell.* PAMI-3, 1981.
26. R. Lumia, L.G. Shapiro, and O. Zemiga, A new connected components algorithm for virtual memory computers, *Comput. Vision Graphics Image Process.* **22**(2), 1983, 287-300.
27. K. Matsumoto, M. Naka, and H. Yanamoto, A new clustering method for LANDSAT images using local maximums of a multidimensional histogram, in *Machine Processing of Remotely Sensed Data*, IEEE CH 1637-8 MPRSD, Purdue University, West Lafayette, Indiana, June 23-26, 1981, pp. 321-325.
28. D.L. Milgram, Region extraction using convergent evidence, *Comput. Graphics Image Process.* **11**, 1979, 1-12.
29. D.L. Milgram and D.J. Kahl, Recursive region extraction, *Comput. Graphics Image Process.* **9**, 1979, 82-88.
30. D.L. Milgram and M. Herman, Clustering edge values for threshold selection, *Comput. Graphics Image Process.* **10**, 1979, 272-280.
31. L.G. Minor and J. Sklansky, The detection and segmentation of blobs in infrared images, *IEEE Trans. Systems Man Cybernet.* SMC-11, 1981, 194-201.

32. J. Muerle and D. Allen, Experimental evaluation of techniques for automatic segmentation of objects in a complex scene, in *Pictorial Pattern Recognition* (G. Cheng *et al.*, Eds.), pp. 3-13, Thompson, Washington, D.C., 1968.
33. G. Nagy and J. Tolaba, Nonsupervised crop classification through airborne multispectral observations, *IBM J. Res. Develop.* **16**, 1972, 138-153.
34. P.M. Narendra and M. Goldberg, A non-parametric clustering scheme, for LANDSAT, *Pattern Recognition* **9**, 1977, 207-215.
35. Ohlander, K. Price, and D.R. Reddy, Picture segmentation using a recursive region splitting method, *Comput. Graphics Image Process.* **8**, 1978, 313-333.
36. D.P. Panda and A. Rosenfeld, Image segmentation by pixel classification in (gray level, edge value) space, *IEEE Trans. Comput.* **C-27**, 1978, 875-879.
37. T. Pavlidis, Segmentation of pictures and maps through functional approximation, *Comput. Graphics Image Process.* **1**, 1972, 360-372.
38. W.A. Perkins, Area segmentation of images using edge points, *IEEE Trans. Pattern Anal. Mach. Intell.* **PAMI-2**, 1980, 8-15.
39. M. Pietikainen and A. Rosenfeld, Image segmentation by texture using pyramid node linking, *IEEE Trans. Systems Man Cybernet.* **SMC-11**, 1981.
40. T.C. Pong, L.G. Shapiro, L.T. Watson, and R.M. Haralick, Experiments in segmentation using a facet model region grower, *Comput. Vision Graphics Image Process.* **25**, 1984.
41. E. Riseman and M. Arbib, Segmentation of static scenes, *Comput. Graphics Image Process.* **6**, 1977, 221-276.
42. T.V. Robertson, Extraction and classification of objects in multispectral images, *Machine Processing of Remotely Sensed Data*, IEEE 73 CHO 837-2GE, Purdue University, West Lafayette, Indiana, October 16-18, 1973, pp. 3B-27-3B-34.
43. S. Wanatabe and the CYBEST group, An automated apparatus for cancer prescreening: CYBEST, *Comput. Graphics Image Process.* **3**, 1974, 350-358.
44. J.S. Weszka, A survey of threshold selection techniques, *Comput. Graphics Image Process.* **7**, 1978, 259-265.
45. J.S. Weszka, R.N. Nagel, and A. Rosenfeld, A threshold selection technique, *IEEE Trans. Comput.* **C-23**, 1974, 1322-1326.
46. J.S. Weszka and A. Rosenfeld, Threshold evaluation techniques, *Trans. Systems Man Cybernet.* **SMC-8**, 1978, 622-629.
47. Y. Yakimovsky, Boundary and object detection in real world image, *J. Assoc. Comput. Mach.* **23**, 1976, 599-618.
48. S. Zucker, Region growing: Childhood and adolescence, *Comput. Graphics Image Process.* **5**, 1976, 382-399.