The Discrimination of Water from Shadow Regions on SAR Imagery

Jianzhong Qian

Department of Electrical Engineering
Virginia Polytechnic Institute and State University
Blacksburg, Virginia 24061

Robert M. Haralick

Department of Electrical Engineering
University of Washington
Seattle, Washington 98195

ABSTRACT

Unlike MSS LANDSAT imagery and other photography, the specific characteristics of the intensity of water and shadow in an SAR image make the task of discriminating them extremely difficult. In this paper, we present a scene analysis system for automatically identifying the water regions and shadow regions whose differences appear as subtle differences in tone and texture. In the preprocessing, a region dependent Multi-threshold Adaptive Filter (MTAF) is proposed for texture preserving noise removal. In the low-level labeling, a probabilistic relaxation algorithm with dynamic adaptive compatibility coefficients is provided to extract the initial object regions. In the high-level interpretation, a relational graph model based on the knowledge of water and shadow regions on SAR imagery is constructed. Then, spatial reasoning is carried out according to this relational model. The contextual information from different processing modules, either at the pixel level or region level, is consistently combined to reduce the labeling ambiguities. The experiments shown that more than 85% of shadow regions and water regions on a set of four SAR images can be identified correctly by this system.

1. INTRODUCTION

In areas of high relief, an airborne SAR radar image creates many radar shadow regions which may be confused with the water regions, since both water and shadow regions appear dark and with subtle differences in tone and texture. There has been little work done to automatically discriminate water regions and shadow regions on a SAR radar image. With the increasing need for analyzing large volumes of radar data by computer, however, this becomes an important recognition task.

In this task, the world of objects is known. However, the specific objects in the scene and their location and shapes are unknown. The information about these objects is uncertain. The goals of a desired system should be able to identify the objects, to determine their locations and to validate the presence of specified objects. In this paper, we describe a knowledge-based scene analysis system (see Figure 1) which uses different levels of contextual information to achieve the goal.

This two-level statistical/structural scene analysis system is capable of identifying shadows and water regions on a one-look-direction radar image. First, the statistical textural feature extraction algorithm is used to extract significant features from the gray-tone co-occurrence matrices for discriminating water regions and shadow regions on the radar imagery. A Maximum Likelihood Decision rule, which incorporates both tonal and texture information into the labeling process, is applied to obtain the initial labeling. An adaptive probabilistic relaxation algorithm, which incorporates the contextual information into the labeling process, is then applied to improve the labeling accuracy. Second, according to our prior knowledge of the water regions and shadow regions on SAR imagery, a relational model is constructed and several structural contextual features are measured to constitute a symbolic description. Then a spatial reasoning process using a set of structural decision rules is employed to generate a symbolic image, in which the specified objects (shadows and water) are clearly identified. As a preprocessing step, a texture preserving noise removal filtering technique is applied to enhance the useful features.

A set of two SAR images collected at Huntsville, Alabama on 17 June, 1977 and another set of two collected at Elizabeth City area on 10 October, 1980 have been tested by this system. The experiments done show that the two level combination discriminating algorithm can provide significant capability for discriminating the water regions from the shadow regions. Because of the use of contextual information, the probabilistic labeling relaxation improved the labeling result on the pixel level, and the spatial reasoning method reduced ambiguities on the region.
From the final results of the three test images, we can see that most of the water regions and shadow regions are identified correctly.

2. KNOWLEDGE OF WATER AND SHADOW REGIONS ON SAR IMAGERY

The scene analysis system proposed in this paper is a knowledge-based pattern recognition system. Characteristics of shadow and water regions on SAR images are important knowledge sources for this system. We will analyze them in this section and will organize them into decision rules in the later sections.

Suppose a system uses the same antenna for transmitting and receiving, the return signal intensity from an object is a function of angle of incidence, wavelength, terrain geometry, terrain roughness, moisture content, dielectric properties of the object, etc. [1, 2]. Also the received signal intensity can be sensitive to wind speed and wind direction for floating objects such as water and semifloating objects such as tree and other vegetations. For a given radar image the radar wavelength and the angle of incidence is fixed. The differences of tone and texture between water and shadow regions on SAR imagery mostly depend upon different ground situations. These differences, however, are subtle.

In flat terrain area, low returns are received from surfaces acting as specular reflectors having surface roughness much less than the wavelength of the radar. Therefore, the gray tone values of water regions on radar imagery are relatively low and appear pretty dark. When the wind is still and water is flat, quiet-water surfaces are near perfect specular reflectors. In this case, the return signal will be almost zero. As a result, the gray tone in these regions is completely black. It is just like that in ideal shadow regions on SAR imagery.

Textures of water regions on SAR imagery usually are created by following two factors. In the open water areas which are either standing or flowing water bodies without vegetation covering them, the water surface agitated by wind backscatters a strong radar return which can be called the "sea return". The tonal and textural differences of the open water area on the SAR imagery may indicate the surface wave action. It turns out that the gray tones corresponding to wavy water surface create the texture of water region. Also, there is the case of the floating and standing vegetation which may cover or partially cover the surface of water bodies. This will create another kind of texture in water regions.

On the other hand, the tones and textures in shadow regions on SAR imagery are caused by following factors. When a bright return occurs from hill slopes or high objects facing the radar look direction, a shadowed region on the far sides of the crests follows, since there is no signal returned from the occluded part of the terrain. Unlike shadows in aerial photography which are weakly illuminated by energy scattered by atmosphere, the radar shadow is completely black and sharply defined. If there is no noise, then an ideal shadow region on SAR imagery should not have any texture. Another case, however, should be considered. If the aircraft reaches a position such that the farther side slope of the mountain is no longer behind the crests, instead of the totally black shadow, a very weak signal from the slope faced away from radar will return. The corresponding parts in the radar image appear relatively dark and we also can view them as shadows. An intermediate case may occur between above two situations so that the shadow is made up of a mixture of totally black and relatively dark tones. This causes a variable texture pattern so that it is difficult to separate the shadow from the water regions.

Besides the above tonal and textural information, there are several structural features which are very helpful for discriminating between water and shadows on SAR imagery. The first thing is a complex signal return case known as the "cardinal effect". This is a result caused by a corner reflector formed by the combinations of three flat surfaces at right angle to one another. These adjacent smooth surfaces cause a double reflection that yields very bright "speckles" or lines on the SAR imagery. The second is known as "Far shore brightening effect". This is caused by the smooth water area with a higher beach which is facing the radar look direction. Since this feature usually covers only small areas of the scene, they often appear as bright lines on SAR imagery. For this case, the dark area comes first, following a bright linear feature which is nearly normal to the radar look direction.

Because the oblique illumination of SAR produces strong returns from the sides of ridges and peaks facing the radar antenna, this makes the boundary appear very bright between shadow and shadow making objects where the near range area of the boundary is not flat. The bright linear feature and the dark region follows along the radar look direction. This situation is just opposite to that of water body boundary. Another property of shadows in SAR
image is that the shadow length along the radar look direction is limited to a certain value for a given image, since the height of shadow-making objects is relatively short compared to the flying height of the airplane.

The above analysis shows that there are some tonal, textural features, and several structural contextual dependency features which are useful for discriminating water and shadows on radar imagery. In some cases, however, they are uncertain and ambiguous. The remainder of this paper discusses these issues.

3. TEXTURE PRESERVING AND NOISE REMOVAL

Because both water region and shadow regions in radar image are not rich in texture, the noise will tend to mask whatever texture there is and make the recognition task more difficult. Noise in radar imagery can come from either radar coherent speckle noise[4] or quantization noise. We need to remove this noise while preserving the weak texture and other subtle detail information.

There is a compromise between noise removal and texture preservation. Usually, a filter which has a powerful noise-cleaning capability may remove or spatially distort edge, line and texture information. In contrast, a filter which preserves subtle detail will tend to have low noise-cleaning capability. Generally speaking, a real world image consists of many regions in which local activity varies from region to region. It is difficult to be satisfied with the image using only one simple filter. Based on this consideration, a way in which several simple filters can be combined to form a more efficient and more flexible context dependent filter is desired. Thus, the advantages of simple filters can be preserved, their drawbacks avoided and, at the same time, an optimal effect can be obtained.

A Multiple Threshold Adaptive Filtering (MTAF) has been proposed to achieve this goal. It uses a generalized gradient function and a local variance function, which measures the local contextual information, as evidence to determine the nature of the filtering for each local neighborhood. More other functions may be chosen as evidence of local activities and different determining methods may be applied for different filtering tasks. A more detailed description on the MTAF can be found in [3,16]. Here we only briefly describe this filtering with a knowledge-based presegmentation procedure using the Dempster-Shafer evidence theory [17].

The algorithm is as follows:

A frame of discernment \( \Theta \) for presegmenting the noisy image consists of the following exhaustive and exclusive subsets:

1. Regions which are homogeneous. They will be filtered by a moving average filter.
2. Regions which have relative weak edges or intermediate ramp edges. They will be filtered by the Sigma filter[4].
3. Regions which have sharp edges that are parallel to the direction of moving window. They will be filtered by the eight edge direction weighted filter[5].
4. Regions which have sharp edges that are not parallel to the window moving direction. They will be filtered by the median filter[6].

Each subset of \( \Theta \) corresponds to a hypothesis. The frame of discernment, \( \Theta \), delimits a sample space which contains all possible segmentation situations, only one of which is true at any one time. Let the set of hypotheses be \( \{ H_a, H_s, H_e, H_m \} \), respectively. Then, the frame of discernment consists of \( \{ H_a, H_s, H_e, H_m \} \) which correspond to the above four situations, respectively.

The evidence provided by each measurement \( \omega_k \) is mapped to basic probability assignments (bpa) over the hypotheses discerned by a frame \( \Theta \). The bpa represents the impact of each distinct piece of evidence on the subsets of a frame \( \Theta \). According to the values of the measurements, a set of thresholds can be used to extract the local evidence. Let \( \{ e_A, e_S, e_E, e_M \} \) be the set of local evidence representing “no edges,” “ramp edges,” “sharp edges,” etc., respectively. Let \( T_k \) be the set of thresholds for the \( k \)th measurement, \( k = 1, 2, \ldots N \) the number of measurements, \( s = 1, 2, \ldots S \) the number of thresholds. Only one of them is true at location \( (i,j) \). For example, for the generalized gradient function, \( \omega_1 \), we have

14 / SPIE Vol. 1297 Hybrid Image and Signal Processing II (1990)
\[
\begin{aligned}
\{& e_{A1} = \text{true} \quad \text{if } \omega_1(ij) < T_{1,1}; \\
& e_{S1} = \text{true} \quad \text{if } T_{1,1} \leq \omega_1(ij) \leq T_{1,2}; \\
& e_{E1} = \text{true} \quad \text{if } \omega_1(ij) > T_{1,3} \text{ AND } (\omega_1(ij) = GD); \\
& e_{M1} = \text{true} \quad \text{if } (\omega_1(ij) > T_{1,3}) \text{ AND } (\omega_1(ij) \neq GD). \\
\end{aligned}
\]

where GD is the generalized gradient value at the direction of the window moving direction[3]. These thresholds give the processor a flexible choice for different image types and different processing purposes. Each piece of extracted evidence will be mapped into bpa over the set of hypotheses. Applying Dempster’s rule, the procedure pools the multiple bodies of evidence obtained from the different measurements to get new belief functions. Finally, a single hypothesis among others is verified, if and only if the belief function has a maximum value.

In our experiments, the filtered pixel value \( W_{ij} \) is given by

\[
W_{ij} = \begin{cases} 
A_{ij} & \text{if } H_a \text{ is verified}; \\
S_{ij} & \text{if } H_s \text{ is verified}; \\
E_{ij} & \text{if } H_e \text{ is verified}; \\
M_{ij} & \text{if } H_m \text{ is verified}.
\end{cases}
\]

where \( W_{ij} \) is the filtered pixel value, \( A_{ij} \) is the output of the 5 by 5 averaging filter, \( S_{ij} \) is the output of the 7 by 7 sigma filter, \( E_{ij} \) is the output of the eight edge direction weighted filter, and \( M_{ij} \) is the output of the 3 by 3 median filter.

As a result, the averaging filter is applied only to the very homogeneous regions, which avoids blurring the weak texture. The median filter is applied only to those regions which have sharp edges and these sharp edges are not parallel to the median filter moving direction, which avoids eliminating the lines and small objects and avoids creating artifacts in other regions. The sigma filter is only applied to those regions where preserving the weak texture is more important than just noise smoothing. Thus, a balanced texture preserving and noise removal effect can be simultaneously achieved. To see how the MTAF filter improves the low-level labeling result more obviously, we added Gaussian normal noise \( N(0, \sigma^2) \) with different \( \sigma \) values to the SAR images. Then the same statistical labeling procedure was applied to the noisy one, the filtered one, and the original one, respectively. The relative labeling accuracy is measured by comparing them with that of original image labeling. Our experiments shown that the relative accuracy can be improved from 61.9 percent to 92.2 percent (with noise \( \sigma = 10 \)), even from 65.5 percent to 96.5 percent (with noise \( \sigma = 5 \)) [3].

4. STATISTICAL LABELING AND ADAPTIVE RELAXATION

In this section we describe a statistical labeling algorithm to obtain the initial labeling. This algorithm combines as much as possible contextual information from different sources at the pixel level to reduce the uncertainty and ambiguity.

Texture is an important feature to characterize and discriminate regions in an image. One of the most widely used method for texture analysis is the gray-tone co-occurrence matrices (GTCM) proposed by Haralick et al. [7]. Here we choose two texture features among others that can be used to extract useful texture information from the GTCM. One is the entropy \( F_e \), the other is inverse difference moment \( F_d \). They are defined by

\[
F_e = -\sum_i \sum_j M_{ij} \log(M_{ij})
\]

(1)

\[
F_d = \sum_i \sum_j \frac{1}{1+(i-j)^2} M_{ij}
\]

(2)

and

\[
M_{ij} = \# \{((m_1,n_1),(m_2,n_2))| f(m_1,n_1) = i, f(m_2,n_2) = j, \| (m_2,n_2) - (m_1,n_1) \| = d \}
\]

(3)
where $d$ is a displacement vector, $M_{ij}$ are the unnormalized counts of how many times two neighboring resolution cells, which are spatially separated by $d$, occur on the image $f$, one with gray tone $i$ and the other with gray tone $j$.

The average entropy has a higher value than other texture features for an image with the same gray tone levels [7]. Since this measure is largest for equal $M_{ij}$ and small when they are unequal, the entropy measure is useful to enhance bright linear features and weak texture values. The usefulness of inverse difference moments texture feature for classifying radar image segments was demonstrated by Shanmugan et al. [8].

We examine the histograms of the radar images and conclude as follows:

1. The histogram of the water region clearly overlaps that of the shadow region in any one-dimensional feature space (either the gray tone, entropy, or inverse difference moment); we have to combine the supports of these three measurements together to form a three dimensional vector $X$.

2. Most of the histograms exhibit a Gaussian like shape. Thus, the Gaussian Maximum Likelihood can be used for the initial labeling process. It assigns measurement vector $X$ to class $\omega_r$, if and only if

$$P(X|\omega_r)P(\omega_r) = \max P(X|\omega_p)P(\omega_p)$$

where

$$P(X|\omega_r) = \frac{1}{(2\pi)^{d/2}|\Sigma_r|^{1/2}} \exp\left[-\frac{1}{2}(X - \mu_r)^T\Sigma^{-1}_r(X - \mu_r)\right]$$

is the class conditional density function, $\mu_r = E[X|\omega_r]$ is the class conditional mean vector, $\Sigma_r = E[(X - \mu_r)(X - \mu_r)^T]$ is the class conditional covariance matrix, both are estimated from training samples for class $\omega_r$, and $P(\omega_r)$ is a prior probability for class $\omega_r$.

Probabilistic relaxation labeling is not a new algorithm [9-14]. However, unlike simple object labeling such as in the tetrahedral block world where there exists a small number of legal labelings, pixel labeling in radar image analysis has enormous ambiguity and the contextual information is generally unknown. Furthermore, in pixel labeling the probabilistic relaxation procedure generally shows a degradation after several iterations. Since the number of iterations for minimum error is unknown in advance, it may become worse after several iterations. This is an especially serious problem in the case that the ground truth data are difficult or impossible to get.

For solving this problem, we proposed a method to improve the labeling performance for our system which uses a non-weighted product neighbor function in which the compatibility coefficients are dynamic. Consequently, it results in a fast convergence speed and a better accuracy in the relaxation process [16]. The procedures are as follows.

First, we compute the initial conditional probabilities for each label, which for pixel $i$ are given by

$$P(q_i, 1) = P(\omega_i|X) = \frac{P(X|\omega_r)P(\omega_r)}{\sum_{\omega_p}P(X|\omega_p)P(\omega_p)}$$

The relaxation will iterate on these conditional probabilities. We wish to emphasize the dependence on the iteration and to deemphasize the dependence on the measurement $s$ associated with the iteration so we use the notation $P(q_i, t)$ to denote the class conditional probability for the $i$th pixel at iteration $t$. The set of measurements on which the class conditional probability $P(q_i, t)$ depends is called the $t$th level context for the $i$th pixel.

From a Bayesian point of view, the probabilistic relaxation is given by [9]

$$P(q_i, t + 1) = \frac{P(q_i, t)Q(q_i, t)}{\sum_{q_i}P(q_i, t)Q(q_i, t)}$$
where $P(q_i, t)$ is the conditional probability that unit $i$ takes label $q_i$ given the $t$th level context, and $Q(q_i, t)$ is the $t$th estimate of the neighborhood function which indicates the degree of neighboring support for that conditional probability.

Typically, there are two types of combinations for the support $Q(q_i, t)$ which the neighbors give the current unit. One is the weighted sum of their supports, the other is the product of their supports. The weighted sum neighbor function does not reach the optimum result for our task because the situations in SAR imagery are complicated. There is no guarantee that setting the a single set of weights can achieve all desired effects simultaneously in the whole updating process. We use a non-weighted product neighbor function in which the compatibility coefficients are adaptive and dynamic. That is

$$Q(q_i, t) = \prod_{j \in N(i)} \sum_{q_j} P(q_i, t) J_{ij}(t)$$

where $N(i)$ is the set of neighbors for unit $i$, $J_{ij}(t)$ is the compatibility coefficient $J_{ij}(q_i|q_j)$, which represents the compatibility between the unit $i$ with the label $q_i$ and its neighbor with the label $q_j$ after $t$ iterations. The compatibility coefficient $J_{ij}(q_i|q_j)$ is computed according to the updated $q_i$ and $q_j$. It is the key term that decides the extent of support in neighbor function $Q(q_i, t)$. However, neighboring patterns of every pixel in SAR imagery actually are unknown. The initial probabilistic labeling process only provides an estimate of the identities of neighboring unknown patterns and has a number of errors and ambiguity. If we compute the compatibility coefficients just using the single pixel estimate in the nearest neighbor system, the initial labeling error will be incorporated into the compatibility coefficients. This will cause degradation in relaxation process. Since the majority of the initial estimates in a chosen window are correct and there exist some correlations between the neighbors, we use a local average estimate instead of a single pixel estimate to reduce the risk of using the incorrect contextual information.

We compute the adaptive compatibility function $J_{ij}(q_i|q_j)$ in the following way. By Bayes formula

$$J_{ij}(q_i|q_j) = \frac{P(q_i, q_j)}{P(q_i)} = \frac{P(q_i)P(q_j|q_i)}{P(q_j)}$$

Using the local average estimate, we choose:

$$P(q_j) = \frac{1}{|N(j)|} \sum_{a \in N(j)} P(q_a, 1)$$

where $j \in N_i(i)$ and $N_i(i)$ is the nearest neighbors of unit $i$, and $N(j)$ is the set of neighbors surrounding the neighbor set $N_i(i)$ for unit $i$.

For the same reason we choose

$$P(q_i) = \frac{1}{|N_2(i)|} \sum_{b \in N_2(i)} P(q_b, 1)$$

where $N_2(i)$ is a set of neighbors of unit $i$, however, $N_2(i)$ is the second order neighbor set. Thus, we have extended the unit $i$ to all its neighbors $b$ in order to estimate $P(q_i)$, so we should also consider all $b$’s neighbors $N(b)$ when we estimate $P(q_i|q_i)$. By approximating the conditional probability we choose

$$P(q_i|q_i) = \frac{1}{|N(b)|} \sum_{a \in N(b)} P(q_a, 1)$$

where $N(b)$ is the set of neighbors of unit $b$. This is equivalent to taking the average of all the probabilities calculated in the above step for each neighboring neighbor.
where $N(b)$ is a set of neighbors of neighbor $N_2(i)$ for unit $i$. The relationship between these neighbor sets are shown in Figure 2.

Using the above technique, we compute the compatibility coefficients $J_{ij}(t)$ for every pixel in the SAR imagery. The immediate context, the neighborhood context and the next larger context are measured once and combined into $J_{ij}(t)$. As a result, $J_{ij}(t)$ slowly varies from pixel to pixel and from iteration to iteration. Then a non-weighted product neighbor function $Q(q_i, t)$ is computed. Finally, we assign the pixel to that class which has the highest probability after relaxation.

5. USING A RELATIONAL MODEL TO REDUCE AMBIGUITIES

One important use of the relational model in scene analysis is to help identify an unknown object that has been extracted from a scene. A possible organization for relational models for scene analysis can be found in [15].

Since some objects in scene, such as water regions and shadow regions in SAR image, are very difficult to distinguish between each other completely and the pictorial similarity is not always a reliable criterion for segmenting a scene into regions that completely correspond to the interesting objects, there are inherent ambiguities in the results produced by low level labeling. Obviously, only those labels are valid that can be derived from an arrangement of real objects in scene. Properties of objects and relations between them imply corresponding properties and relations of the SAR image regions that result from these objects. These projected properties and relations contain the possible labelings of regions with object identification. The results that we have from our low level labeling provided better labeling accuracy than that the simple segmentation did. So the goal that we try to achieve here is just to use high level contextual information to verify the objects and reduce the ambiguities. As a result, the relational graph model that we need is much simpler than the usual one.

According to the relational model (Figure 3), a set of decision rules are determined as follows:

Assign a region to shadow, if

1) it is a region which is assigned to shadow by the low level labeling and it is surrounded by other non-water region; or
2) it is a region which is assigned to shadow by the low level labeling and it is adjacent to bright linear features in the radar look direction.

Assign a region to water, if

1) it is a region which is assigned to water by the low level labeling and it is adjacent to bright linear features in the opposite radar look direction; or
2) it is a region which is assigned to water by the low level labeling and it is surrounded by other non-shadow regions; or
3) it is a region which is assigned to water by the low level labeling and it has a length longer than the limitation of shadow length along radar look direction; or
4) it is a region which is assigned to water by the low level labeling and it has some vegetation on it.

Determine a region to be a false shadow region and merge it with its surround, if

1) it is a region which is assigned to shadow by the low level labeling and it is surrounded by a water region; or
2) it is a region which is assigned to shadow by the low level labeling and it has a length longer than the limitation of shadow length along radar look direction;
3) it is a region which is assigned to shadow by the low level labeling and it lies between a water region and other non-shadow regions.

Determine a region to be a false water region and merge it with its surround, if

it is a region is assigned to a water region by the low level labeling and it is surrounded by a shadow region.
For the above relational model, the shapes of water regions and shadow regions are arbitrary, hence classical region attributes representing an object such as medial axis etc. are not helpful in this situation. Also, the representation of regions by their circumscribing boxes is not suitable for our case, because sometimes a spurious adjacent or surrounding relation will hold between boxes of two regions, while in fact it is not true of the regions themselves. We choose several structural measurements for the regions. They are: the size of the region, the relative position of different adjacent regions along the radar look direction, the maximum length of the region along radar look direction, the region state, and the number of region boundary pixels.

The most important thing is how to measure these spatial relations among regions. The algorithm for this is as follows.

First, for the convenience of spatial reasoning, each region in the symbolic image should be identified by a unique region identifier. Therefore, the algorithm defines new region indexes sequentially for the three categories which obtained from low-level processing. For a three digit index, the first digit may represent the original low level labeling. For example, the water region indexes may start from 100 and the shadow region indexes may start from 200 etc. Thus, we still can recognize the initial assigned label from the new sequential index.

Second, it uses a linear geometric transformation to rotate the symbolic image to such a position in which the horizontal scanning line just is parallel to the radar look direction.

Third, it scans the region symbolic image line by line in the radar look direction. If the scanning line meets a new region label, the scanning process will be interrupted and a tracing process which traces the region external boundary will start at this break point. We use a one-pass, depth first boundary tracing procedure using a left first, clockwise directed four connected neighborhood search technique for tracing a region's boundary:

1) Record coordinates of the starting point of the region, keep the region always in the right side of tracing direction and trace the boundary in clockwise.

2) For each successor, detect the next tracing direction by searching the same label from four connected neighbors in the order starting from the left side of previous moving direction, then the front, the right, finally the back (see Figure 4).

3) Record the left side adjacent region label and its starting position and ending position in the tracing process.

4) Count the number of pixels of the region boundary in the tracing process.

5) Mark the region's state, (Mark 0 indicates a closed region which lies completely within the image. Mark 1 indicates that the region touches the image boundary.)

6) If next point = starting point, then stop tracing, go to step 7, else go to step 2.

7) Continue the scanning process from the break point, compute the maximum region length along radar look direction and the region size.

8) Check every label that the scanning line met within a hash table. If the region has been traced, go to step 7, else go to step 1 until the last line.

The high level image analysis techniques require rapid access to region information and to the relations between regions. For every region, therefore, the region attributes and the relations of adjacent regions (between or surround) are stored in a big list. And a hash function is used for fast searching and accessing to their structural information. After that, a property file which contains a list of property values for the above measurements can be created. Then the spatial reasoning process simply becomes a table-look up procedure according to the structural decision rule described previously.

6. EXPERIMENTAL RESULTS

A set of four SAR images was tested using the system described in this paper. These images were collected over Huntsville, Alabama on 17 June, 1977 and over Elizabeth City area on 10 October, 1980. Only two of them are presented here due to space considerations.
In the low level statistical labeling process the Gaussian Maximum Likelihood labeling procedure is applied to the test image (Figure 5a). The feature space for labeling consists of the gray tone value and the texture values. Then the 3-dimensional mean vector and 3 by 3 covariance matrix are estimated from the training sample set for the water region and the shadow region and others, respectively. The texture features, both entropy and inverse difference moment, are extracted from the GTCM which are computed for window size of 12 by 12 and a distance of 6. The initial labeling result is shown in Figure 5b.

To improve the overall labeling accuracy, the adaptive relaxation using contextual information is applied to Figure 5b. The initial statistical labeling is determined by the normalized maximum likelihood estimate. The immediate context, the neighborhood context and the next larger context in the initial labeling are measured once and combined into the compatibility function. Comparing relaxation result Figure 5c with Figure 5b, the effect of reducing the ambiguity in the shadow regions is obvious. However, we can not expect this method to change the labeling result too much, since the initial labeling error can influence the degree of the improvement. Further improvement can be obtained from the use of spatial relations between the structural features.

To examine the performance of the structural recognition algorithm described in Section 5, the symbolic image of Figure 5c were tested. The radar look direction for both images are approximately from North to South (Top to Bottom), so we rotated the images 90 degree before measuring their region attributes and relations. After eliminating small region, defining new region indexes, and using the depth first tracing strategy described in Section 5, the structural contextual features were measured and the corresponding property lists were created (see Table 1).

Table 1. The region attributes and relations list for the symbolic image of Figure 5c. (R1 = the surrounding region if R2=0. R2 = the second adjacent region that is not in the same label as R1.)

<table>
<thead>
<tr>
<th>Region Number</th>
<th>Region Index</th>
<th>Maximum Length</th>
<th>Adjacency R1</th>
<th>Adjacency R2</th>
<th>State Mark</th>
<th>Region Size</th>
<th>Boundary Length</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>100</td>
<td>319</td>
<td>300</td>
<td>204</td>
<td>1</td>
<td>89684</td>
<td>2325</td>
</tr>
<tr>
<td>2</td>
<td>101</td>
<td>21</td>
<td>300</td>
<td>215</td>
<td>1</td>
<td>148</td>
<td>63</td>
</tr>
<tr>
<td>3</td>
<td>200</td>
<td>21</td>
<td>300</td>
<td>100</td>
<td>0</td>
<td>1088</td>
<td>440</td>
</tr>
<tr>
<td>4</td>
<td>201</td>
<td>27</td>
<td>100</td>
<td>300</td>
<td>0</td>
<td>1375</td>
<td>541</td>
</tr>
<tr>
<td>5</td>
<td>202</td>
<td>18</td>
<td>300</td>
<td>0</td>
<td>0</td>
<td>129</td>
<td>57</td>
</tr>
<tr>
<td>6</td>
<td>203</td>
<td>15</td>
<td>300</td>
<td>0</td>
<td>0</td>
<td>107</td>
<td>52</td>
</tr>
<tr>
<td>7</td>
<td>204</td>
<td>35</td>
<td>100</td>
<td>300</td>
<td>1</td>
<td>3070</td>
<td>927</td>
</tr>
<tr>
<td>8</td>
<td>205</td>
<td>24</td>
<td>300</td>
<td>0</td>
<td>0</td>
<td>203</td>
<td>69</td>
</tr>
<tr>
<td>9</td>
<td>206</td>
<td>23</td>
<td>300</td>
<td>0</td>
<td>0</td>
<td>640</td>
<td>227</td>
</tr>
<tr>
<td>10</td>
<td>207</td>
<td>14</td>
<td>300</td>
<td>0</td>
<td>0</td>
<td>180</td>
<td>99</td>
</tr>
<tr>
<td>11</td>
<td>208</td>
<td>24</td>
<td>300</td>
<td>0</td>
<td>0</td>
<td>1220</td>
<td>339</td>
</tr>
<tr>
<td>12</td>
<td>209</td>
<td>16</td>
<td>300</td>
<td>0</td>
<td>0</td>
<td>104</td>
<td>56</td>
</tr>
<tr>
<td>13</td>
<td>210</td>
<td>22</td>
<td>300</td>
<td>0</td>
<td>0</td>
<td>1012</td>
<td>377</td>
</tr>
<tr>
<td>14</td>
<td>211</td>
<td>34</td>
<td>300</td>
<td>0</td>
<td>0</td>
<td>1777</td>
<td>338</td>
</tr>
<tr>
<td>15</td>
<td>212</td>
<td>21</td>
<td>300</td>
<td>0</td>
<td>0</td>
<td>592</td>
<td>201</td>
</tr>
<tr>
<td>16</td>
<td>213</td>
<td>31</td>
<td>300</td>
<td>0</td>
<td>0</td>
<td>2235</td>
<td>438</td>
</tr>
<tr>
<td>17</td>
<td>214</td>
<td>17</td>
<td>300</td>
<td>0</td>
<td>0</td>
<td>131</td>
<td>66</td>
</tr>
<tr>
<td>18</td>
<td>215</td>
<td>32</td>
<td>101</td>
<td>300</td>
<td>1</td>
<td>1946</td>
<td>406</td>
</tr>
<tr>
<td>19</td>
<td>216</td>
<td>22</td>
<td>300</td>
<td>0</td>
<td>0</td>
<td>382</td>
<td>89</td>
</tr>
<tr>
<td>20</td>
<td>217</td>
<td>16</td>
<td>300</td>
<td>0</td>
<td>0</td>
<td>296</td>
<td>82</td>
</tr>
<tr>
<td>21</td>
<td>300</td>
<td>448</td>
<td>200</td>
<td>100</td>
<td>1</td>
<td>154802</td>
<td>2962</td>
</tr>
</tbody>
</table>

Finally, a new labeling image was generated according to the results of spatial reasoning using the property lists. The results are shown in Figure 5d. Compared to the results of low level labeling, the new results are more reasonable and correct.

Another experiment was done in the following way. We use the mean and covariance estimate from test image of Figure 5a to label another image (Figure 5e) collected in the same period and over the same area. The two texture features and the gray tone feature are used as before. Using the same statistical classifier, we obtained the low level initial labeling result which is shown in Figure 5f. The high level final recognition results are shown in Figure 5g.
The result indicates that high level recognition using the relational model can reduce the ambiguity so significantly that it allows the low level recognition procedure to be less accurate.

7. CONCLUSIONS

We presented a scene analysis system for identifying the shadows and water on SAR image in this paper. Experiments show that this system is practical for our purpose. This system was designed so that it used the contextual information sufficiently in the preprocessing, in the pixel level labeling, or in the region level reasoning. In this way, it overcame the specific identification difficulty that the objects to be discriminated were with subtle differences in tone and texture.

In the preprocessing, we use a region dependent Multi-Threshold Adaptive filtering technique for texture preserving noise removal. In the low-level labeling, we discussed a probabilistic relaxation algorithm in which the adaptive compatibility coefficients were computed by local average estimate and were dynamic in the whole updating process. It extracted most contextual information in pixel level and reduced the labeling error in less than 5 iterations and then stabilized. In the high-level spatial reasoning, the relational graph model based on our prior knowledge of water and shadow regions on SAR imagery was constructed. The corresponding structural decision rules were derived from the relational model. Contextual information in the region-level is measured by a one-pass, depth first boundary tracing procedure using a left first, clockwise directed four neighborhood search technique. Using this method, a property file which contained a list of region attributes and region relations was created. Then a spatial reasoning procedure was performed according to the set of structural decision rules. It reduced the errors and ambiguities resulting from low-level labeling significantly. Most false shadow regions and false water regions were identified and corrected.

Although the system is mainly for the discrimination of shadows and water in radar imagery, the principle and algorithms of the system can be used to identify other objects which may not be completely discriminated using a single level recognition scheme.

8. REFERENCES

Figure 1. The block diagram of the system.

Figure 2. The relationship of the neighbor sets.
(1) The 7x7 window and the nearest neighbors N(1).
(2) The second order neighbors N2(i). I is the pixel to be updated and i=JS.
(3) The neighbors N(Jk) of the neighbor N1(i) for the estimate of Jk, where k=1,2,3,4,5.
(4) The neighbors N(Bm) of the neighbors N2(i) and m=1,2,...,9.

Figure 3. The relational graph model from water, shadow, and others.

Figure 4. Previous moving directions and searching order.
Figure 5. (a) The test image; (b) the initial low level labeling of image 1; (c) the relaxation result of image 1; (d) the high level reasoning result of image 1; (e) the test image 2; (f) the low level labeling of image 2 after relaxation; (g) the high level reasoning result; in (b), (c), (d), (f), and (g), Black–Water regions, Gray–Shadow regions, and White–Others.