EXTRACTION AND CLASSIFICATION OF
OBJECTS IN MULTISPECTRAL IMAGES*

Thomas V. Robertson

Laboratory for Applications of Remote Sensing
Purdue University, West Lafayette, Indiana

I. ABSTRACT

Presented here is an algorithm that partitions a
digitized multispectral image into parts that correspond
to objects in the scene being sensed. The algorithm
partitions an image into successively smaller rectangles
and produces a partition that tends to minimize a crit-
erion function.

Supervised and unsupervised classification tech-
niques can be applied to partitioned images. This part-
ition-then-classify approach is used to process images
sensed from aircraft and the ERTS-1 satellite, and the
method is shown to give relatively accurate results in
classifying agricultural areas and extracting urban
areas.

II. INTRODUCTION

The classification of a multispectral image involves labeling areas of interest
in the image. These areas of interest are groups of image points that have been
produced by the sensing of objects such as agricultural fields, bodies of water, and
cities. One approach to machine classification of images has been to classify each
image point separately. This approach uses the reflectance of each point in various
spectral bands (channels) to classify the point. Classification algorithms using
point-by-point classification methods have been successful in many applications, but
in some cases classification accuracy has been undesirably low.

Human photointerpreters use spatial properties such as texture, size, and shape
in image interpretation. The presence of this spatial information in multispectral
images suggests that machine classification of multispectral images may be improved
if spatial as well as spectral information is used in the classification algorithms.

The classification method presented in this paper is a two step procedure.
First, an image is partitioned into blocks or sets of image points. The image part-
itioning algorithm is designed so that it is likely that each block contains image
points from a single object of interest. In the second step of the procedure, the
blocks are classified. Classifying blocks instead of individual image points allows
the measurement and use of texture and other spatial characteristics of objects
that are not apparent when single points are classified separately.

* This research was sponsored by NASA contract NAS 52-1773.
** The author is presently with Bell Laboratories, Holmdel, New Jersey.
III. PARTITIONING ALGORITHM

The partitioning algorithm divides an image into disjoint rectangles (blocks) such that each area of interest (object) is approximated by a union of blocks. The basic characteristics of the algorithm are given in the following sections.

A. NOTATION AND DEFINITIONS

An image $I$ is a set of points in a plane that is surrounded by a closed curve $C$ of finite length. In our discussion we will assume that the image points of $I$ are defined by all the intersections surrounded by $C$ of a set of equally-spaced horizontal and vertical lines in the plane. A subimage of $I$ is an image $J$ such that $J \subseteq I$.

A partition $P$ of an image $I$ is a finite set of images $(I_1, I_2, \ldots, I_L)$ such that

$$I = \bigcup_{i=1}^{L} I_i$$

and for $j \neq 1$,

$$I_j \cap I_1 = \emptyset$$

where $\emptyset$ is the empty set. Each $I_j \in P$ will be called a block of $P$.

The area of an image $J$ will be denoted $|J|$. The size of $J$ is the minimum of the horizontal and vertical extent of $J$.

A gray-level function $g(\cdot)$ is a function whose domain is an image and whose range is a bounded interval on the real line. We use $g(X)$ to stand for the gray level at a point $X \in I$. For a given $X$, $g(X)$ will be considered a random variable whose distribution depends on $X$. A gray-level vector $G(\cdot)$ is a vector of gray-level functions:

$$G(X) = (g_1(X), g_2(X), \ldots, g_N(X))$$

where each $g_i(\cdot)$ is a gray-level function.

Consider an image $J$. Let $E(\cdot)$ be expected value. We will use the following notation:

$$M_{g_1}(J) = E(g_1(X) | X \in J)$$

$$M_{g_2}(J)$$

$$\vdots$$

$$M_{g_N}(J)$$

$$M_G(J) = \begin{bmatrix} M_{g_1}(J) \\ M_{g_2}(J) \\ \vdots \\ M_{g_N}(J) \end{bmatrix}$$

We call $M_G(J)$ the mean vector of $J$. Also let

$$S^2_{g_1}(J) = E((g_1(X) - M_{g_1}(X))^2 | X \in J)$$

$$S^2_{g_2}(J) = E((g_2(X))^2 | X \in J)$$

An image $J$ is $G$-regular if for any subimage $K \subseteq J$, $M_G(K) = M_G(J)$. A $G$-regular image is "homogeneous" with respect to $G$ in the sense that the mean values of the gray-level functions $(g_i(\cdot), i=1, 2, \ldots, N)$ are constant throughout the image.

A subimage $J$ of $I$ is $G$-distinct if $J$ is $G$-regular, and if for any subimage $K \subseteq I$ that is adjacent to $J$, $K \cap J$ is not $G$-regular. In other words, a $G$-distinct subimage is surrounded by subimages with different mean values of the $N$ gray-level functions of $G$.

A partition $P$ is $G$-regular if every block of $P$ is $G$-regular; $P$ is called $G$-optimal if every block in $P$ is also $G$-distinct. Note that a $G$-optimal partition is necessarily $G$-regular, but a $G$-regular partition is not $G$-optimal if some pair of adjacent blocks have the same mean vectors.
The mean test to determine the G-regularity of an image $J$ is carried out as follows: First $J$ is partitioned into two subimages $J_1$ and $J_2$. $J$ is determined to be G-regular if and only if $H_G(J_1)=H_G(J_2)$. In Ref. 1 we show that this test makes no errors if the number of image points per unit area is infinite. We also show in Ref. 1 that the G-optimal partition $F^*$ is unique.

B. PARTITION CRITERION

We assume that the blocks in the G-optimal partition $F^*$ of $I$, $F^*=(O_1, O_2, \ldots, O_N)$, correspond to the objects in $I$. Therefore a good partition of $I$ is one that closely approximates $F$. Now present a criterion function that is minimized by good partitions.

Consider an arbitrary partition of $I$, $F=(I_1, I_2, \ldots, I_N)$, and a gray-level function $g(*)$. We first define a criterion $V_g(P)$ for the single gray-level function $g(*)$:

$$V_g(P) = \frac{1}{N} \sum_{i=1}^{N} |I_i| S_g^2(I_i)$$

Recall that the $S_g^2(I_i)$'s are the variances of the blocks in the partition $P$. A block variance tends to be small if the block contains a single object, but a block that overlaps an object boundary contains several objects will have relatively high variance. Since in the criterion function block variances are weighted by the block area, $V_g(P)$ will tend to be small when most of the largest blocks contain only a single object; in other words, when $P$ is approximately g-regular. For a gray-level vector $G(*)$ we define

$$V_G(P) = \sum_{j=1}^{N} V_{g_j}(P)$$

We also define a partition error

$$\Delta V_g(P) = V_g(P) - V_g(P^*)$$

and

$$\Delta V_G(P) = V_G(P) - V_G(P^*)$$

$$= \sum_{j=1}^{N} \Delta V_{g_j}(P)$$

In Ref. 1 we show that $V_G(P)$ is a minimum if and only if $P$ is a G-regular partition.

C. THE ALGORITHM

Figure 1 shows a flow chart of the REcursive PARtitioning algorithm, which we will call RIMPAR. RIMPAR continues to subdivide blocks until the block under consideration is either too small or G-regular. The question of G-regularity is decided by the mean test discussed earlier. The specification of which block sizes are too small is handled by a parameter MINSIZE. In Ref. 1 we prove the following result: Assuming no errors are made in determining G-regularity, for any $\varepsilon > 0$, there are MINSIZE values for which $\Delta V_G(P_\varepsilon) < \varepsilon$, where $P_\varepsilon$ is a partition of $I$ produced by RIMPAR in a finite number of steps, and $I$ is assumed to have an infinite number of points per unit area.

In practice MINSIZE is useful in resolving ambiguities in object definition. The user of RIMPAR can use MINSIZE to specify whether he wants certain target areas to be considered large textured objects or sets of small, relatively homogeneous objects.

To implement the mean test, several partitions of $J$ are tried. These trial partitions are generated by $(k_0-1)$ horizontal and $(k_0-1)$ vertical, equally spaced lines. Here $k_0$ is an integer greater than 1. The trial partition $P=(J_1, J_2)$ that yields the most improvement in an estimate of the partition criterion function $V_G(*)$ is used to carry out an approximate version of the mean test. In this approximate mean test we use the multivariate $T^2$ statistical hypothesis test (Ref. 2) that assumes the gray levels in $J_1$ and $J_2$ are normally distributed, and tests the hypothesis that $H_G(J_1)=H_G(J_2)$.
A. CLASSIFICATION METHODS

In the following experiments we investigate classifying partitioned images and compare this method to classifying the individual points of images. The classification algorithms used are all based on the assumption that the data are characterized by multivariate normal distributions.

In the supervised classification of partitioned images, a statistical distance measure (the Bhattacharyya distance) is used to determine the distances between the estimated distributions of the gray levels in partition blocks and the estimated distributions of the gray levels of subimages of known classification. This technique is compared to supervised per-point classification in which a Bayesian maximum likelihood classifier is used to classify individual image points by comparing point gray levels to the estimated distributions of the gray levels of subimages of known classification.

Unsupervised classification is carried out using a standard clustering algorithm, which can be thought of as following these steps:

1. An initial number $M$ of classes is specified, and the initial distributions of these classes are estimated using an arbitrary subset of the data to be clustered.

2. The partition blocks or image points are then classified using supervised classification techniques and the current estimates of the $M$ class distributions.

3. If the class membership of the partition blocks or image points is unchanged from the previous iteration, the algorithm stops.

4. If there is a change in class membership, calculate a new estimate of the $M$ class distributions based on the new members of each class, then return to step 2.

The details of the classification algorithms are discussed in Ref. 1.

B. CLASSIFYING AGRICULTURAL AREAS

In the first set of experiments supervised classification is used to identify crop types in 5 images. The distributions of the classes of interest are estimated before classification using training fields. The characteristics of these 5 images are summarized in Table 1. In Table 2 we compare RIMPAR classification (classifying an image partitioned by RIMPAR) with per-point classification (classifying individual image points). Classification accuracy is calculated by comparing the classification results with test fields that contain points of known classification. These test fields are distinct from the fields used to estimate distributions used by the classifiers. The processing time reported is in seconds of virtual CPU time on an IBM 360/67 time shared computer. Results storage is in bytes, and is calculated assuming one byte for each class label and 4 bytes to specify a partition block location. The channels used for partitioning and classification are, in general, different for each image. For the aircraft images, wavelengths from 0.40 to 1.7 microns are used; and for the satellite images, wavelengths from 0.6 to 0.8 microns are used.

From the results shown in Table 2 we conclude that in comparing per-point and RIMPAR classification, the latter technique gives comparable accuracy (an average of 1% improvement in these experiments), less results storage (24% - 42% in these experiments), and larger processing time (900% - 1250%) compared to the former technique.

C. CLASSIFYING URBAN AREAS

In the next set of experiments, a 93,000 point image from the ERTS-1 satellite is used to investigate the classification of urban areas. This image contains 5 relatively large cities. From top to bottom, the three largest cities are (see Figure 2) Janesville, Wisconsin; Beloit, Wisconsin; and Rockford, Illinois. A
smaller city, Belvedere, Illinois, appears to the right of Rockford, and above Belvedere is Poplar Grove, Illinois. The goal of these experiments is to isolate these cities from the rest of the image. This isolation is accomplished by performing unsupervised classification (clustering) of the image and displaying the cluster classes as different gray levels. The cities are considered to be effectively isolated if they are represented exclusively by a single cluster class. Two methods using clustering are compared: clustering the individual image points and clustering the partition blocks produced by RIMPAR.

In Figure 2 we show the results of clustering the individual points of the image into 5 classes using channels 2 (0.6 - 0.7 microns) and 4 (0.8 - 1.1 microns). Visually this clustered image seems to be a good representation of the cities in the image. However, the human visual system does a lot of spatial integration in viewing such a picture. As shown in the right side of Figure 2, the cluster class most nearly representing the cities consists of (1) separated points within the cities, and (2) many superfluous points outside the cities. Thus the image description stored in the computer, represented by Figure 2, does not specify 5 major objects that represent cities. The cities are not found as distinct objects when individual points are clustered because cities are characterized by texture as well as the reflectance of individual image points.

In Figure 3 we show the results of clustering the image using channels 2 and 4 after the image was first partitioned by RIMPAR. From the figure it is clear that the cities have been approximately isolated. Although the boundaries of the cities are not precise, the image of Figure 3 is a useful input to more detailed processing.

V. SUMMARY

An image partitioning algorithm is presented and applied to the classification of agricultural and urban areas. This method of classification is shown to give small classification results storage at the expense of large computation time. The technique is also shown to be clearly superior to a per-point method in isolating cities in an IRS3-1 image.

VI. REFERENCES


Table 1. Image Characteristics

<table>
<thead>
<tr>
<th>Image</th>
<th>Source</th>
<th>Classes of Interest</th>
<th>Average Field Size (Points)</th>
<th>No. Training Field Points</th>
<th>No. Test Field Points</th>
</tr>
</thead>
<tbody>
<tr>
<td>69002901</td>
<td>Aircraft 2400 feet Forage, Forest, Water</td>
<td>Corn, Soybeans, Wheat</td>
<td>221</td>
<td>2727</td>
<td>5237</td>
</tr>
<tr>
<td>66000600</td>
<td>Aircraft 2600 feet Clover, Alfalfa, Bare Soil</td>
<td>Corn, Soybeans, Wheat, Oats</td>
<td>410</td>
<td>4459</td>
<td>13562</td>
</tr>
<tr>
<td>71053900</td>
<td>Aircraft 5000 feet Forest, Water</td>
<td>Corn, Soybeans, Wheat</td>
<td>64</td>
<td>1387</td>
<td>6410</td>
</tr>
<tr>
<td>7203280A</td>
<td>Satellite 580 miles (Other Vegetation)</td>
<td>Corn, Soybeans, Other Vegetation</td>
<td>18</td>
<td>850</td>
<td>4842</td>
</tr>
<tr>
<td>7203280B</td>
<td>Satellite 580 miles (Other Vegetation)</td>
<td>Corn, Soybeans, Other Vegetation</td>
<td>18</td>
<td>1309</td>
<td>1409</td>
</tr>
</tbody>
</table>

Table 2. Comparison of RIMPAR and Per-Point Classifiers

<table>
<thead>
<tr>
<th>Image</th>
<th>No. Channels to Partition/Classify</th>
<th>Accuracy*</th>
<th>Time RIMPAR/Per-Point</th>
<th>Time RIMPAR/Per-Point</th>
<th>Results Storage RIMPAR/Per-Point</th>
</tr>
</thead>
<tbody>
<tr>
<td>69002901</td>
<td>2/4</td>
<td>76.7/78.5</td>
<td>1214/100</td>
<td>15630/44000</td>
<td></td>
</tr>
<tr>
<td>66000600</td>
<td>2/4</td>
<td>79.9/78.5</td>
<td>967/95</td>
<td>9535/40280</td>
<td></td>
</tr>
<tr>
<td>71053900</td>
<td>2/4</td>
<td>95.4/93.2</td>
<td>1145/105</td>
<td>13950/46509</td>
<td></td>
</tr>
<tr>
<td>7203280A</td>
<td>2/2</td>
<td>82.6/81.3</td>
<td>753/81</td>
<td>14125/36000</td>
<td></td>
</tr>
<tr>
<td>7203280B</td>
<td>2/2</td>
<td>74.0/71.8</td>
<td>615/67</td>
<td>11635/27900</td>
<td></td>
</tr>
</tbody>
</table>

* Accuracy calculated as 100X(Number of correctly classified points)/(Number of test field points).
Figure 1. Basic RIMPAR Flow Chart
Figure 2. Per-Point Clustered Satellite Image

Figure 3. Clustered Partitioned Image