MULTISPECTRAL IMAGE PARTITIONING*

Thomas V. Robertson
Laboratory for Applications of Remote Sensing
Purdue University

1. INTRODUCTION

A multispectral image is considered here to be a digitized representation of the reflectance of a target in several regions of the electromagnetic spectrum. Such images can be produced by multispectral scanners [1], or by digitizing photographs [2]. Multispectral scanners have been flown in aircraft [3] and the ERTS-1 satellite [4] for remote sensing of the earth's surface.

In a multispectral image, target area is represented by a finite number of image points. For each image point there are reflectance measurements (gray levels) in several spectral bands (channels). The gray levels are quantized, often to either 64 or 256 values. Figure 1 shows an example of a multispectral image.

Processing a multispectral image ultimately leads to an understanding of the target being sensed; for example an image may be processed to determine the population or crop yield of a geographical area. Such processing usually involves several steps. In each step an image processor manipulates a description of the image to develop a new image description for use by a succeeding processor. Some processing steps seem to be best handled by

*This research was sponsored by NASA Contract NAS 52-1773.

Telecommunications Conference, Nov. 26-28, 1973 Atlanta, GA

people, but computer-implemented algorithms have been successful in steps such as image coding [5] and classification [6].

In Figure 1 it is apparent to the eye that there are regions of internal regularity in the image that differ from surrounding regions. Such regions, which we will call objects, occur when a forest, body of water, or other part of the target corresponds to more than one image point. Although the human eye percieves most images as an arrangement of objects, most automatic coding and classification algorithms work with a point-by-point description of the image as an input. The purpose of the image partitioning technique described in this paper is to take a "point-by-point" description of a multispectral image and produce an "arrangement-of-objects" description. This is done by dividing the image into pieces, where the pieces approximate objects.

An arrangement-of-objects description is often better than a point-by-point description as an input to coding or classification algorithms. In image coding it is desirable to remove redundancy from an image. If an image is divided into objects, information about an object can replace a larger amount of information about many image points. Classification of objects instead of individual image points allows the measurement and use of texture and other spatial characteristics, thus potentially improving accuracy [7, 8, 9].

Much of the previous work in image partitioning has employed algorithms that search for points on the boundaries between

adjacent objects [10, 11, 12, 13]. These methods work well if the probability is large that two image points in the same object are similar, and two image points in different objects are dissimilar. Individual edge points are hard to find if the image is noisy, or if adjacent objects differ mainly in texture [14]. Another difficulty is using edge points is that further processing must be done to link the edge points into closed boundaries so that objects can be defined [15, 16].

Another basic approach to image partitioning, which we will call the LBLOCK algorithm, has two steps [17, 18].

(1) The image is divided into blocks of image points by a regular rectangular grid.

(2) Objects are built up from these blocks by merging adjacent blocks if they are "similar".

The main difficulty in using LBLOCK algorithms is choosing the block size. If blocks are too large, small objects will be missed, and irregular object boundaries will be poorly approximated by block boundaries. If blocks are too small, texture patterns larger than the block size will not be detected, and errors in estimating the similarity between blocks will be increased because of noise.

We present here a Recursive Image Partitioning (RIMPAR) algorithm that divides an image into successively smaller rectangular blocks. At each step in the algorithm, a block is subdivided unless

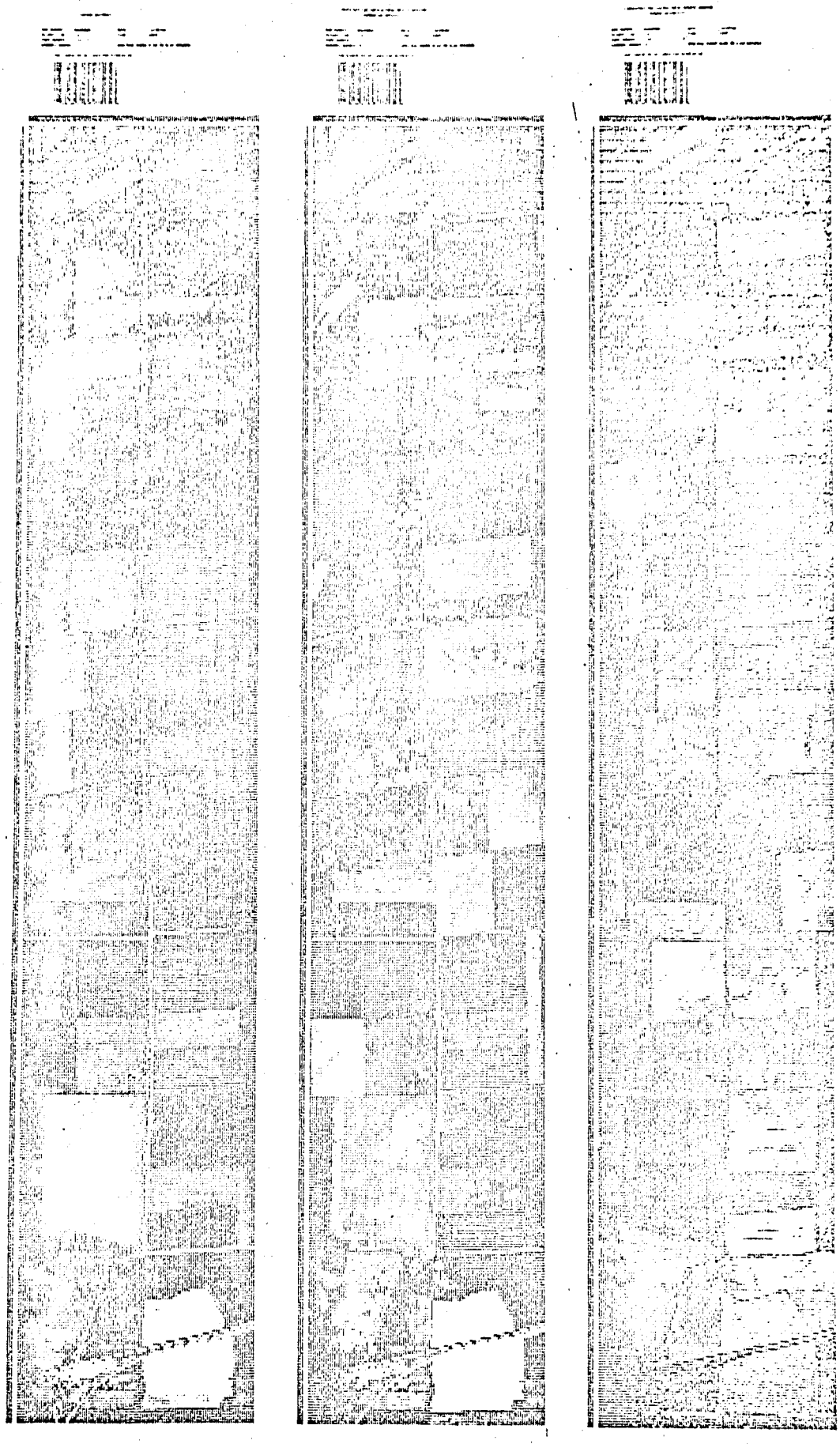


Figure 1. A multispectral image.

(1) the block's size (largest side) is less than a parameter MINSIZE; or

(2) it is likely that the image points within the block are from a single object.

Thus the RIMPAR algorithm partitions an image into blocks so that each object is approximated by a set of adjacent blocks. This algorithm works well with noisy or gradual object boundaries, and with objects of varying texture element size.

To present the basic properties of the RIMPAR algorithm, we will first consider an ideal, continuous image with infinite resolution. Next we will discuss the impact of finite resolution on the algorithm. Finally some results of applying the algorithm to various multispectral images will be presented.

2. PARTITIONING IDEAL IMAGES

The continuous image model is ideal in the sense that any part of an object, no matter how small, retains the statistical properties of the object. We assume that in this ideal image, we can estimate with zero error the mean gray-level value of any part of the image. A RIMPAR algorithm designed to partition this ideal image will be presented in this section, and in the next section we will discuss applications to non-ideal images.

2.1 Notation and Definitions

An image I is a set in the real plane that is surrounded by a closed curve C of finite length, such that any set J surrounded

by C satisfies $J \subseteq I$. Note that in our definition "image" refers to a set of points. The gray levels associated with an image will be discussed separately below. The essential properties of our "image" are its infinite resolution and the absence of "holes" in the point set. A subimage of I is an image J such that $J \subseteq I$. From this point on in this paper, we will assume the all point sets under discussion are images.

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 $j \neq i$,

$$I_j \cap I_i = \emptyset,$$

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 $|J|$. Two subimages of I, J_1 and J_2, are said to be adjacent if $J_1 \cup J_2$ is an image, and $J_1 \cap J_2 = \emptyset$.

Horizontal and vertical will refer to a set of perpendicular coordinate axes that are fixed with respect to I. The size of an image $J \subseteq I$ is the maximum 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 will 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:

$$M_{g_i}(J) = E[g_i(X) \mid X \varepsilon J]$$

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

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

$$S_{g_i}^2(J) = E[(g_i(X) - M_{g_i}(X))^2 \mid X \varepsilon J]$$

$$Z_{g_i}^2(J) = E[g_i(X)^2 \mid X \varepsilon J].$$

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, $K \cup J$ 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 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.

2.2 Properties of G-Regular Images

Theorem 2.1 If $K \subseteq J$ and J is G-regular, then K is G-regular.

Proof: Assume K is not G-regular. Then there is a $B \subseteq K$ such that

$$M_G(B) \neq M_G(K) \tag{2.2}$$

But $B \subseteq K \subseteq J$ and J is G-regular, so

$$M_G(B) = M_G(K) = M_G(J). \tag{2.3}$$

Since Eqn. 2.2 contradicts Eqn. 2.3, K must be G-regular.

Theorem 2.2 If J_1 and J_2 are adjacent and G-regular, and $M_G(J_1) = M_G(J_2)$, then $J = J_1 \cup J_2$ is G-regular.

Proof: Consider any subimage $K \subseteq J$. Then

$$\begin{aligned} M_G(K) &= \frac{|J_1 \cap K|}{|K|} M_G(J_1) + \frac{|J_2 \cap K|}{|K|} M_G(J_2) \\ &= M_G(J_1) \left\{ \frac{|J_1 \cap K|}{|K|} + \frac{|J_2 \cap K|}{|K|} \right\} \\ &= M_G(J_1) \end{aligned}$$

Therefore J is G-regular.

2.3 The Mean Test for G-Regularity

In this subsection we will show that to determine if an image J is G-regular, we need only to test for the equality of the mean vectors of L subimages of J, $L \geq 2$.

Theorem 2.3 Consider an image J and any partition of J, $P = \{J_1, J_2, \ldots, J_L\}$, $L \geq 2$. If G^* is a constant vector and $M_G(J_i) = G^*$, $1 \leq i \leq L$, then $\text{Prob}(J \text{ is G-regular}) = 1$.

Proof: Let GJ be the event "$M_G(J_i) = G^*$, $1 \leq i \leq L$", and RJ be the event "J is G-regular". $\overline{RJ}$ will stand for "J is not G-regular". The theorem can now be restated as:

$$\text{Prob}(RJ \mid GJ) = 1 \tag{2.4}$$

Since $\text{Prob}(RJ \mid GJ) + \text{Prob}(\overline{RJ} \mid GJ) = 1$, the theorem holds if and only if

$$\text{Prob}(\overline{RJ} \mid GJ) = \frac{\text{Prob}(\overline{RJ}, GJ)}{\text{Prob}(GJ)} = 0 \tag{2.5}$$

$\text{Prob}(GJ) \neq 0$ because we assume the existance of G-regular images, so the theorem holds if and only if $\text{Prob}(\overline{RJ}, GJ) = 0$. Now we will derive a matrix equation expressing the joint event $(\overline{RJ}, GJ)$, and show that the probability of satisfying this equation is zero.

If J is not G-regular, then there must be a partition P' of J, $P' = \{O_1, O_2, \ldots, O_M\}$, such that for $i \neq j$

$$M_G(O_i) \neq M_G(O_j).$$

If we let

$$c_{ij} = \frac{|J_i \cap O_j|}{|J_i|}$$

then we can express the mean vectors of the blocks of P in terms of the mean vectors of the blocks of P':

$$M_G(J_i) = \sum_{j=1}^{M} c_{ij} M_G(O_j)$$

Now if $M_G(J_i) = G^*$, $1 \leq i \leq L$, then $M_G(J) = G^*$, so if we let

$$d_j = \frac{|O_j|}{|J|}$$

then

$$\sum_{j=1}^{M} d_j \, M_G(O_j) = G^*.$$

Consider the matrix E with elements e_{ij} defined by

$$e_{ij} = c_{ij} - d_j \quad 1 \leq i \leq L, \; 1 \leq j \leq M.$$

Let $G^* = (g_1^*, g_2^*, \ldots, g_N^*)$. For every element $g_k(\cdot)$ of $G(\cdot)$ we can write

$$\sum_{j=1}^{M} e_{ij} \, M_{g_k}(O_j) = \sum_{j=1}^{M} c_{ij}(O_j) - \sum_{j=1}^{M} d_j \, M_{g_k}(O_j) = g_k^* - g_k^* = 0$$

if and only if GJ. Thus GJ can be expressed by

$$\begin{bmatrix} E \end{bmatrix} \quad \begin{bmatrix} M_{g_k}(O_1) \\ M_{g_k}(O_2) \\ \vdots \\ M_{g_k}(O_M) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \qquad 1 \leq k \leq N \qquad (2.6)$$

We now consider, for any k, solution vectors of the form $(M_{g_k}(O_1), M_{g_k}(O_2), \ldots, M_{g_k}(O_M))$ that satisfy Eqn. 2.6. The set of solution vectors for Eqn. 2.6 is called the "null space" of the matrix E [19]. The dimension of the null space of E is the number of columns of E (=M) minus the rank of E. Since the rank of E is greater than

zero, the dimension q of the null space of E is less than M.

For a given E, we assume that $(M_{g_k}(O_1), M_{g_k}(O_2), \ldots, M_{g_k}(O_M))$ is random vector, continuously distributed in a bounded region of M-space. (This assumption holds only if $\overline{RJ}$, for if J is G-regular, then all the vector elements are equal, and the equation is satisfied). Since the solutions of Eqn. 2.6 occupy a subspace of dimension q<M, the solution vectors occur with zero probability. Thus Eqn. 2.6 holds with zero probability, and the theorem is proved.

2.4 Uniqueness of the G-optimal Partition

Theorem 2.4 The G-optimal partition P* of I is unique.

Proof: Assume there are two partitions $P^* = \{I_1, I_2, \ldots, I_M\}$ and $P' = \{J_1, J_2, \ldots, J_L\}$ that are both G-optimal and $P^* \neq P'$. Since $P^* \neq P'$, for some k there are adjacent blocks J_i and J_j such that

$$I_k = (J_i \cap I_k) \cup (J_j \cap I_k)$$

Since I_k is G-regular

$$M_G(I_k) = M_G(J_i \cap I_k) = M_G(J_j \cap I_k). \tag{2.7}$$

J_i and J_j are also G-regular so

$$M_G(J_i \cap I_k) = M_G(J_i)$$

$$M_G(J_j \cap I_k) = M_G(J_j)$$

But if P' is G-optimal then $M_G(J_i) \neq M_G(J_j)$. This contradicts Eqn. 2.7, so there can be only one G-optimal partition.

2.5 Partition Criterion

We assume that the blocks in the G-optimal partition P* of I

correspond to the objects in I. Therefore a good partition of I is one that closely approximates P*. Figure 2 shows examples of several partitions. In this subsection we present a criterion function that is minimized by good partitions.

2.5.1 Criterion Definition

Consider the G-optimal partition of I, $P^* = \{O_1, O_2, \ldots, O_M\}$, an arbitrary partition $P = \{I_1, I_2, \ldots, I_L\}$, and a gray-level function $g(\cdot)$.

We first define a criterion $V_g(P)$ for the single gray-level function $g(\cdot)$:

$$V_g(P) = \sum_{i=1}^{L} \frac{|I_i|}{|I|} S_g^2(I_i) \tag{2.8}$$

For a gray-level vector $G(\cdot)$ we define

$$V_G(P) = \sum_{j=1}^{N} V_{g_j}(P) \tag{2.9}$$

We also define a <u>partition error</u> $\Delta V_g(P)$ for $g(\cdot)$,

$$\Delta V_g(P) = V_g(P) - V_g(P^*) \tag{2.10}$$

and a partition error for $G(\cdot)$,

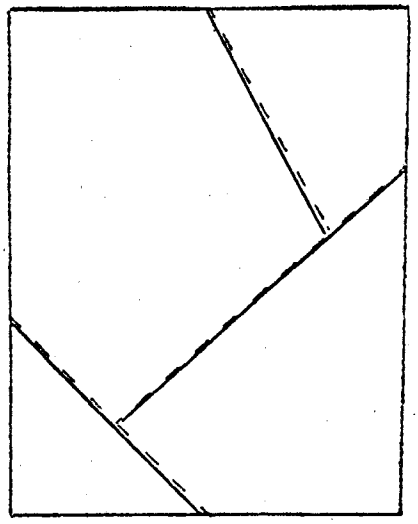
$$\Delta V_G(P) = V_G(P) - V_G(P^*) = \sum_{j=1}^{N} \Delta V_{g_i}(P) \tag{2.11}$$

2.5.2 Criterion Properties

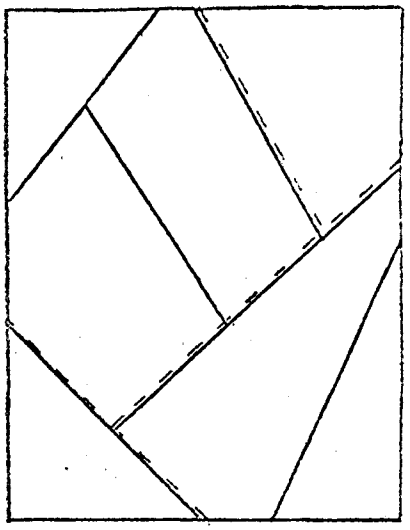
<u>Theorem 2.5</u>: $\Delta V_g(P) = \sum_{i=1}^{M} \sum_{j=1}^{L} \frac{|O_i \cap I_j|}{|I|} (M_g(O_i) - M_g(I_j))^2$.

<u>Proof</u>: From set theory,

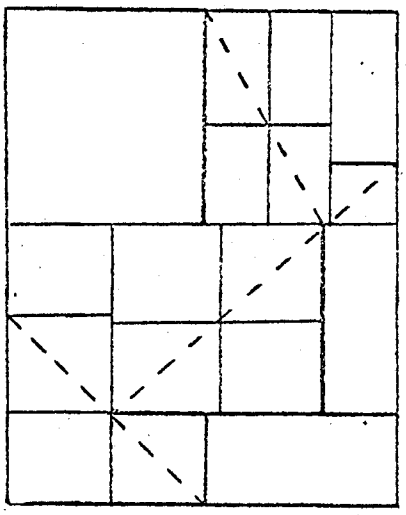
$$O_i = \bigcup_{j=1}^{L} (I_j \cap O_i).$$



G-Optimal Partition



G-Regular Partition



Approximately G-Regular Partition

Figure 2. Partition types.

Since $I_j \cap I_k = \emptyset$, $j \neq k$,

$$(I_j \cap O_i) \cap (I_k \cap O_i) = \emptyset, \quad j \neq k.$$

Therefore

$$|O_i| = \sum_{j=1}^{L} |I_j \cap O_i|$$

and

$$V_g(P^*) = \sum_{i=1}^{M} \frac{|O_i|}{|I|} S_g^2 (O_i)$$

$$= \sum_{i=1}^{M} \sum_{j=1}^{L} \frac{|O_i \cap I_j|}{|I|} S_g^2 (O_i).$$

Now

$$V_g(P) = \sum_{j=1}^{L} \frac{|I_j|}{|I|} S_g^2 (I_j)$$

so

$$V_g(P) - V_g(P^*) = \sum_{j=1}^{L} \left(\frac{|I_j|}{|I|} S_g^2 (I_j) - \sum_{i=1}^{M} \frac{|O_i \cap I_j|}{|I|} S_g^2 (O_i) \right)$$

$$= \sum_{j=1}^{L} \frac{|I_j|}{|I|} \left(S_g^2 (I_j) - \sum_{i=1}^{M} \frac{|O_i \cap I_j|}{|I_j|} S_g^2 (O_i) \right). \quad (2.13)$$

The probability density function $f(g(X) | X \epsilon I_j)$ is given by

$$f(g(X) | X \epsilon I_j) = \sum_{i=1}^{M} \frac{|O_i \cap I_j|}{|I_j|} f(g(X) | X \epsilon O_i). \quad (2.14)$$

We can write

$$S_g^2 (I_j) = Z_g^2 (I_j) - M_g^2 (I_j)$$

$$Z_g^2 (I_j) = \sum_{i=1}^{M} \frac{|O_i \cap I_j|}{|I_j|} Z_g^2 (O_i) \quad (2.15)$$

$$S_g^2(O_i) = Z_g^2(O_i) - M_g^2(O_i) \tag{2.16}$$

Now from Eqns. 2.13, 2.15, and 2.16

$$V_g(P) - V_g(P^*) = \sum_{j=1}^{L} \frac{|I_j|}{|I|} \left\{ \sum_{i=1}^{M} \frac{|O_i \cap I_j|}{|I_j|} Z_g^2(O_i) - M_g^2(I_j) - \sum_{i=1}^{M} \frac{|O_i \cap I_j|}{|I_j|} (Z_g^2(O_i) - M_g^2(O_i)) \right\}$$

$$= \sum_{j=1}^{L} \frac{|I_j|}{|I|} \left\{ \sum_{i=1}^{M} \frac{|O_i \cap I_j|}{|I_j|} M_g^2(O_i) - M_g^2(I_j) \right\} \tag{2.17}$$

From Eqn. 2.14 we have

$$M_g(I_j) = \sum_{i=1}^{M} \frac{|O_i \cap I_j|}{|I_j|} M_g(O_i) \tag{2.18}$$

and so

$$\sum_{i=1}^{M} \frac{|O_i \cap I_j|}{|I_j|} M_g^2(O_i) - M_g^2(I_j) = \sum_{i=1}^{M} \frac{|O_i \cap I_j|}{|I_j|} M_g^2(O_i) - \left(\sum_{i=1}^{M} \frac{|O_i \cap I_j|}{|I_j|} M_g(O_i) \right)^2$$

$$= \sum_{i=1}^{M} \frac{|O_i \cap I_j|}{|I_j|} M_g^2(O_i) - 2 \left(\sum_{i=1}^{M} \frac{|O_i \cap I_j|}{|I_j|} M_g(O_i) \right) \cdot \left(\sum_{k=1}^{M} \frac{|O_k \cap I_j|}{|I_j|} M_g(O_k) \right) + \left(\sum_{k=1}^{M} \frac{|O_k \cap I_j|}{|I_j|} M_g(O_k) \right)^2 \cdot \left(\sum_{i=1}^{M} \frac{|O_i \cap I_j|}{|I_j|} \right) \tag{2.19}$$

The last equation follows from

$$\sum_{i=1}^{M} \frac{|O_i \cap I_j|}{|I_j|} = 1$$

Using Eqn. 2.18 we simplify Eqn. 2.19 to

$$\sum_{i=1}^{M} \left\{ \frac{|O_i \cap I_j|}{|I_j|} M_g^2 (O_i) \right\} - M_g^2 (I_j)$$

$$= \sum_{i=1}^{M} \left\{ \frac{|O_i \cap I_j|}{|I_j|} M_g^2 (O_i) - 2 M_g (O_i) M_g (I_j) + M_g (I_j)^2 \right\}$$

$$= \sum_{i=1}^{M} \frac{|O_i \cap I_j|}{|I_j|} (M_g (O_i) - M_g (I_j))^2.$$

Now we can rewrite Eqn. 2.17

$$V_g(P) - V_g(P^*) = \sum_{j=1}^{L} \frac{|I_j|}{|I|} \sum_{i=1}^{M} \frac{|O_i \cap I_j|}{|I_j|} (M_g (O_i) - M_g (I_j))^2$$

$$= \sum_{i=1}^{M} \sum_{j=1}^{L} \frac{|O_i \cap I_j|}{|I|} (M_g (O_i) - M_g (I_j))^2.$$

Theorem 2.6: $\Delta V_g(P) = 0$ if and only if P is g-regular.

Proof: First assume $\Delta V_g(P) = 0$. From Eqn. 2.12 this implies $|O_i \cap I_j| (M_g(O_i) - M_g (I_j))^2 = 0$, $1 \le i \le m$, $1 \le j \le L$. Therefore if $(O_i \cap I_j) \neq \emptyset$, we must have $M_g(O_i) = M_g(I_j)$.

Consider a block I_j of P, and any subimage $K \subseteq I_j$. We can write

$$K = \bigcup_{i=1}^{M} O_i \cap K \tag{2.20}$$

For each i, $O_i \cap K \neq \emptyset$ implies $O_i \cap I_j \neq \emptyset$ and $M_g (O_i) = M_g (I_j)$. Since $(O_i \cap K) \subseteq O_i$, $M_g (O_i \cap K) = M_g (O_i) = M_g (I_j)$, $1 \le i \le M$, when $O_i \cap K \neq \emptyset$.

Therefore each term on the right side of Eqn. 2.20 has the same mean $M_g(I_j)$, so $M_g(K) = M_g(I_j)$, and I_j is g-regular. Since every block I_j of P is g-regular, P is g-regular.

Now assume P is g-regular. Consider each $|O_i \cap I_j| (M_g(O_i) - M_g(I_j))^2$, $1 \leq i \leq M$, $1 \leq j \leq L$. We will show if $|O_i \cap I_j| \neq 0$, then $M_g(O_i) = M_g(I_j)$.

If $|O_i \cap I_j| \neq 0$ then $(O_i \cap I_j) \neq \emptyset$. We can write

$$O_i = \bigcup_{j=1}^{L} O_i \cap I_j \qquad 1 \leq i \leq m \tag{2.21}$$

For every nonempty term on the right side of Eqn. 2.21 we have $(O_i \cap I_j) \subseteq O_i$ so

$$M_g(O_i \cap I_j) = M_g(O_i) \; 1 \leq i \leq M, \; 1 \leq j \leq L. \tag{2.22}$$

We can also write

$$I_j = \bigcup_{i=1}^{M} I_j \cap O_i \qquad 1 \leq j \leq L$$

and since each I_j is g-regular, if $(I_j \cap O_i) \neq \emptyset$,

$$M_g(O_j \cap I_j) = M_g(O_i), \quad 1 \leq i \leq M, \quad 1 \leq j \leq L. \tag{2.23}$$

From Eqns 2.22 and 2.23 we have if $(O_i \cap I_j) \neq \emptyset$. then

$$M_g(O_i) = M_g(I_j) \quad 1 \leq i \leq M, \; 1 \leq j \leq L.$$

Therefore $\Delta V_g(P) = 0$.

Theorem 2.7: $\Delta V_G(P) = 0$ if and only if P is G-regular.

Proof: From Eqn. 2.11 we see that $\Delta V_G(P) = \sum_{j=1}^{N} \Delta V_{g_j}(P) = 0$ if and only if $\Delta V_{g_j}(P) = 0$, $1 \leq j \leq N$. (Note from Eqn. 2.8 that each $\Delta V_{g_j}(P) \geq 0$). From Theorem 2.6, $\Delta V_{g_j}(P) = 0$ if and only if P is g_j-regular, $1 \leq j \leq N$. Therefore $\Delta V_G(P) = 0$ if and only if P is

G-regular.

From Eqn.2.12 it is clear that $\Delta V_g(P) \geq 0$, and so $\Delta V_G(P) \geq 0$. Therefore from the preceeding two theorems it follows that $V_G(\cdot)$ is minimized by the G-optimal partition P*, and also by any G-regular partition. The algorithm presented in the next section produces a partition P_f that tends to minimize $V_G(P_f)$, so P_f is approximately G-regular. A G-regular partition is sub-optimal, but is a simple matter (in the ideal case) to transform a G-regular partition to the G-optimal partition, by merging adjacent blocks that have identical mean vectors. Therefore a partition that is approximately G-regular can be made approximately G-optimal by merging adjacent blocks having identical mean vectors.

2.6 PARTITIONING ALGORITHM FOR IDEAL IMAGES

Fig. 3 is a flow chart of the basic RIMPAR algorithm. The algorithm is recursive in the sense that every image J (except when J=I) that is divided into two parts, is a subimage of a previously divided image.

The partitioning of each J into $\{J_1, J_2\}$ is assumed to be carried out under the following constraints.

(1) If the vertical extent of J is greater than MINSIZE, partition J with a horizontal line. Otherwise partition J with a vertical line. Recall that for the size of J to be greater than MINSIZE, either the horizontal or vertical extent of J must be greater than MINSIZE.

(2) The partition of J is made so that

$$\text{MIN}\left(\frac{|J_1|}{|J|}, \frac{|J_2|}{|J|}\right) \geq K_D.$$

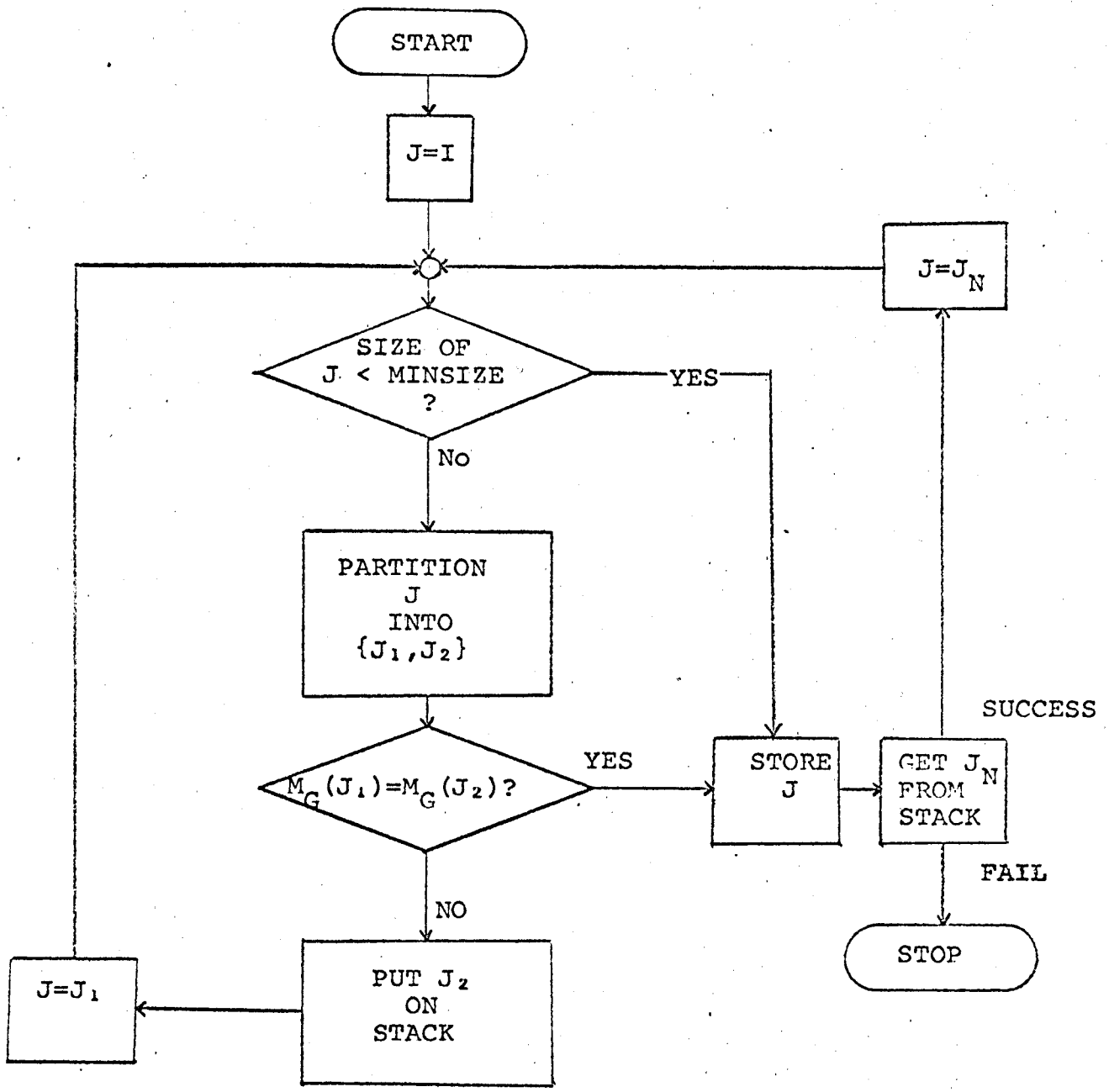


Figure 3. Basic RIMPAR flow chart.

Thus the area of each block of the partition of J is at least $100K_D\%$ of the area of J.

RIMPAR generates a sequence of partitions $P_0, P_1, \ldots, P_f$ of I. In Fig. 4 we show an example of such a sequence of partitions. The following theorem shows that each member P_{i+1} of the sequence of partitions is an improvement over the preceeding member P_i.

Theorem 2.8: $V_G(P_{i+1}) < V_G(P_i)$, $0 \le i \le f-1$.

Proof: Consider a partition of I, $P_i=\{I_1, I_2, \ldots, I_{L-1}, J\}$, where J is the block that is partitioned into $\{J_1, J_2\}$ to produce P_{i+1} from P_i. Let $g(\cdot)$ be an arbitrary gray-level function of $G(\cdot)$.

$$V_g(P_i) = \sum_{i=1}^{L-1} \frac{|I_i|}{|I|} S_g^2(I_i) + \frac{|J|}{|I|} S_g^2(J)$$

and

$$V_g(P_{i+1}) = \sum_{i=1}^{L-1} \frac{|I_i|}{|I|} S_g^2(I_i) + \frac{|J_1|}{|I|} S_g^2(J_1) + \frac{|J_2|}{|I|} S_g^2(J_2).$$

Therefore

$$V_g(P_i)-V_g(P_{i+1}) = \frac{|J|}{|I|} S_g^2(J)-\frac{|J_1|}{|I|} S_g^2(J_1)-\frac{|J_2|}{|I|} S_g^2(J_2)$$

$$=\frac{1}{|I|}\left\{|J_1|Z_g^2(J_1)+|J_2|Z_g^2(J_2)-(|J_1|M_g(J_1)+|J_2|M_g(J_2))^2\left(\frac{I}{|J|}\right)-\right.$$

$$\left.-|J_1|Z_g^2(J_1)+|J_1|M_g^2(J_1)-|J_2|Z_g^2(J_2)+|J_2|M_g^2(J_2)\right\}$$

$$=\frac{1}{|I||J|}\left\{|J|(|J_1|M_g^2(J_1)+|J_2|\ M_g^2(J_2))-(|J_1|M_g(J_1)+|J_2|M_g(J_2))^2\right\}$$

By using $|J| = |J_1|+|J_2|$, the above reduces to

$$V_g(P_i)-V_g(P_{i+1})=\frac{|J_1||J_2|}{|I|\ |J|}(M_g(J_1)-M_g(J_2))^2. \qquad (2.25)$$

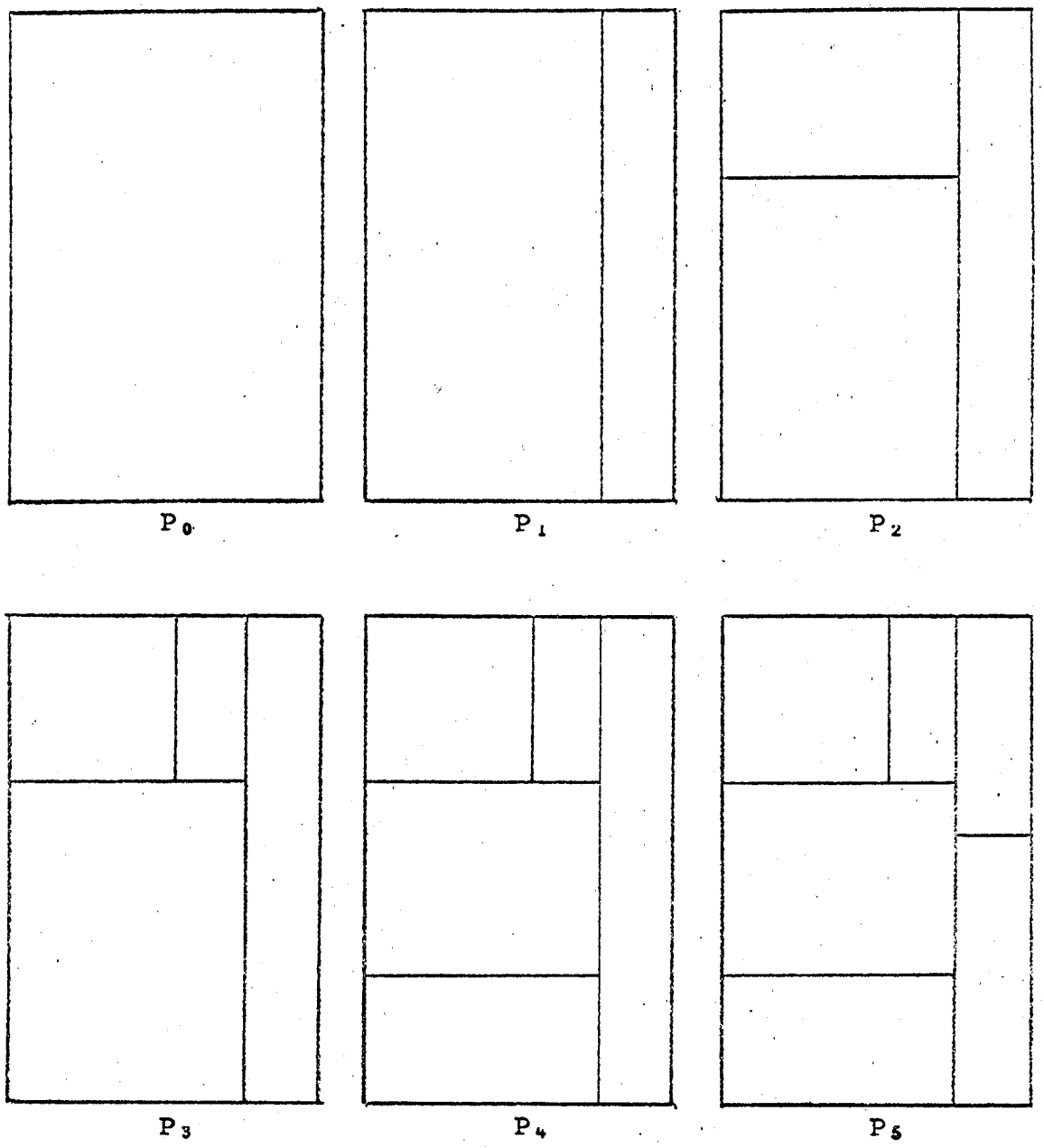


Figure 4. Sequence of partitions generated by RIMPAR

Therefore for every $g(\cdot)$ of $G(\cdot)$, $V_g(P_i)-V_g(P_{i+1}) \geq 0$; and therefore $V_G(P_i)-V_G(P_{i+1}) \geq 0$. $V_G(P_i)=V_G(P_{i+1})$ if and only if $M_G(J_1)=M_G(J_2)$. Therefore, since we assume that RIMPAR decides to partition J, $M_G(J_1)\neq M_G(J_2)$ and $V_G(P_i)>V_G(P_{i+1})$.

When RIMPAR reaches the box labeled "STORE J" in Fig. 3, J is accepted as one of the blocks of the final partition P_f. From the flow chart it is evident that J is accepted for P_f only if the size of J is less than MINSIZE, or if $M_G(J_1)=M_G(J_2)$. The algorithm is structured so that at any point in the sequence of algorithm steps, all parts of I that are not included in the already-accepted blocks of P_f are either in J, or on the stack. Therefore no part of I is excluded from testing by the algorithm, and every block I_j in P_f satisfies either

(1) the size of I_j is less than MINSIZE; or

(2) I_j is G-regular.

The following theorem shows how these properties allow partition error to be controlled by MINSIZE.

Theorem 2.9: Consider any image I and any $\varepsilon>0$. There exists a value of MINSIZE such that RIMPAR produces a final partition P_f satisfying $\Delta V_G(P_f)<\varepsilon$.

Proof: From Eqn. 2.11 we see that $\Delta V_G(P_f)=\sum_{j=1}^{N} \Delta V_{g_j}(P_f)$ so the theorem holds if $\Delta V_{g_j}(P_f)<\varepsilon/N$, $1\leq j\leq N$. For an arbitrary $g(\cdot)$ in $G(\cdot)$ we have

$$\Delta V_g(P_f)=\sum_{i=1}^{M}\sum_{j=1}^{L}\frac{|O_i\cap I_j|}{|I|}(M_g(O_i)-M_g(I_j))^2.$$

We assume that for $1\leq i\leq m$ and $1\leq j\leq L$

$$(M_g(O_i)-M_g(I_j))^2<D$$

where D is some constant. Therefore

$$V_g(P_f)<\frac{D}{|I|}\sum_{i=1}^{M}\sum_{j=1}^{L}|O_i\cap I_j|.$$

Now $\sum_{i=1}^{M}\sum_{j=1}^{L}|O_i\cap I_j|$ is just the area of the blocks of P_f that overlap boundaries of objects (objects are blocks in P*). The object boundaries have a finite total arc length S_T. The smallest side of a block in P_f is greater than K_D·MINSIZE (see the algorithm constraints). Therefore the number of overlapping blocks is less than $S_T/(K_D\cdot\text{MINSIZE})$. Since the area of each overlapping block is less than $(\text{MINSIZE})^2$,

$$\sum_{i=1}^{M}\sum_{j=1}^{L}|O_i\cap I_j| < \frac{S_T\cdot(\text{MINSIZE})^2}{K_D\cdot(\text{MINSIZE})} \tag{2.24}$$

and

$$\Delta V_g(P_f)<\frac{DS_T}{|I|K_D}(\text{MINSIZE}).$$

Therefore if we set

$$\text{MINSIZE} = \frac{|I|K_D\varepsilon}{DS_TN} \tag{2.26}$$

then $\Delta V_g(P_f)<\varepsilon/N$ and therefore $\Delta V_G(P_f)<\varepsilon$.

It is interesting to note that to insure that the right hand side of Eqn. 2.24 decreases with MINSIZE, the <u>size</u> of overlapping blocks must decrease more slowly with MINSIZE than the <u>area</u> of overlapping blocks. Therefore RIMPAR must limit block size rather than block area.

3. PARTITIONING REAL IMAGES

We now consider the application of RIMPAR to images with a finite number of points.

3.1 Sufficient Gray-Level Vectors

In order for RIMPAR to produce a partition that is meaningful, the blocks of the G-optimal partition P* should correspond to areas of the target we wish to consider as single entities. A gray-level vector G for which P* satisfies this goal is called sufficient. In other words, G is sufficient if for any adjacent areas H_1 and H_2 in I such that H_1 and H_2 represent different target entities, $M_G(H_1) \neq M_G(H_2)$.

The first logical candidate for gray-level functions is the set $G_1(\cdot)=(g_1(\cdot), g_2(\cdot), \ldots, g_n(\cdot))$ corresponding to the reflectance in the n channels of the multispectral image. In many applications $G_1(\cdot)$ or even a subset of $G_1(\cdot)$ may be sufficient. However, if for adjacent areas H_1 and H_2, $M_{g_i}(H_1)=M_{g_i}(H_2)$, $1 \leq i \leq n$, then $G_1(\cdot)$ is not sufficient. This would be the case if H_1 and H_2 had identical means in all channels, but different variances in some channels. To handle such situations we propose sets of gray-level functions $G_k(\cdot)=(g_1^k(\cdot), g_2^k(\cdot), \ldots, g_n^k(\cdot))$ where $g_i^k(X)$ is an unbiased estimate of $E\left[(g_i(Y))^k / Y \in N(X)\right]$ and N(X) is a small neighborhood of X, consisting perhaps of X and the eight nearest neighbors of X. If any H_1 and H_2 had gray-level distributions that differ in at least one of the first r moments of at least one of the first n channels, then $G(\cdot)=(G_1(\cdot), G_2(\cdot), \ldots, G_r(\cdot))$ is sufficient.

The gray-level vector can be further generalized to distinguish between adjacent target areas that differ in either pattern or dependence among neighboring points. The procedure in these cases is to define gray-level functions with mean values that reflect these differences.

3.2 Testing G- Regularity

Since a real image has a finite number of image points, we can only estimate the mean vector for any subimage J. Therefore testing G-regularity cannot be done with zero error, and it is desirable to find a test that tends to minimize testing errors.

Recall that to determine if J is G-regular, we partition J into $\{J_1, J_2\}$ and ask "$M_G(J_1)=M_G(J_2)$?" In this subsection we discuss the implementation of this test.

3.2.1 Partitioning the Gray-Level Vector

To determine $M_G(J_1) \neq M_G(J_2)$ it is sufficient to determine $M_{g_i}(J_1) \neq M_{g_i}(J_2)$ for some $g_i(\cdot) \varepsilon G(\cdot)$. Therefore it is not necessary to use all gray-level functions to decide $M_G(J_1) \neq M_G(J_2)$. If we partition $G(\cdot)$ into subvectors $(G_1(\cdot), G_2(\cdot), \ldots, G_p(\cdot))$, then we can test for G-regularity as shown in Fig. 5. This testing scheme has two advantages:

(1) All gray level functions need not be evaluated for each J.

(2) Testing subvectors instead of individual gray-level functions allows the use of redundancy to smooth out noisy estimates.

3.2.2 A Test Statistic

We calculate an estimate $\hat{M}_{G_i}(J_k)$ of $M_{G_i}(J_k)$ by

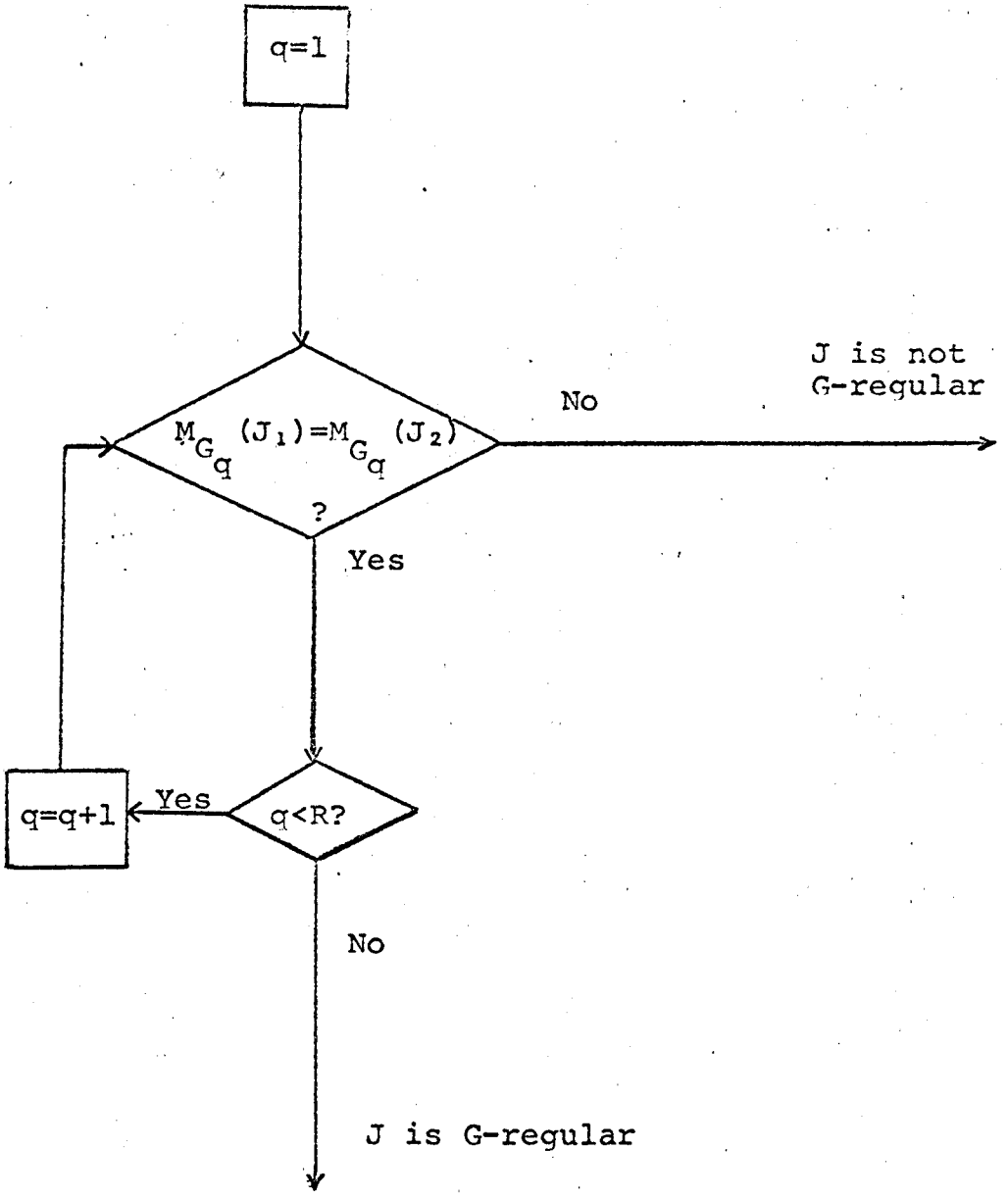


Figure 5. G-Regularity Test

$$\hat{M}_{G_i}(J_k) = \frac{1}{|J_k|} \sum_{X \in J_k} G_i(X) \qquad k=1,2 \tag{3.1}$$

Throughout this section $|K|$ stands for the number of points in any image K. Let

$$S = \frac{1}{|J|-2} \{ \sum_{X \in J_1} (G_i(X) - \hat{M}_{G_i}(J_1))(G_i(X) - \hat{M}_{G_i}(J_1))^T + \sum_{X \in J_2} (G_i(X) - \hat{M}_{G_i}(J_2))$$

$$(G_i(X) - \hat{M}_{G_i}(J_2))^T \}$$

and

$$T^2 = \frac{|J_1||J_2|}{|J|} (\hat{M}_{G_i}(J_1) - \hat{M}_{G_i}(J_2))^T S^{-1} \cdot (\hat{M}_{G_i}(J_1) - \hat{M}_{G_i}(J_2)) \tag{3.2}$$

The statistic T^2 has been used [20] to test the equality of the mean vectors of two multivariate Gaussian distributions in the following way: the two mean vectors are considered equal (at significance level α) if

$$T < \frac{(|J|-2)r}{(|J|-r-1)} F_{r,\,|J|-r-1}(\alpha) \tag{3.3}$$

where r is the number of gray-level functions in $G_i(\cdot)$, and $F_{r,\,|J|-r-1}(\alpha)$ is the upper $100\alpha\%$ point of the F distribution with r and $|J|-r-1$ degrees of freedom. In our application we have no guarantee that the distributions $f(G_i(X)/X \in J_1)$ and $f(G_i(X)/X \in J_2)$ are multivariate Gaussian, but the distributions are probably close enough to Gaussian to allow the use of T^2 with some success. We will approximate Eqn. 3.3 by first noting that since $|J| >> r$ [21],

$$\frac{(|J|-2)r}{(|J|-r-1)} F_{r,|J|-r-1}(\alpha) \approx \chi^2_r(\alpha).$$

where $\chi_r^2(\alpha)$ is the upper $100\alpha\%$ of the χ^2 distribution with r degrees of freedom. For a given α, $\chi_r^2(\alpha)$ increases with r. Therefore we use the following test: decide $M_{G_i}(J_1)=M_{G_i}(J_2)$

$$T^2 < rT_m$$

where T_m is an experimentally determined constant.

3.2.2 Efficiency of Partitions

If it is found that $M_{G_i}(J_1) \neq M_{G_i}(J_2)$, then J is partitioned into J_1 and J_2. Although any partition $\{J_1, J_2\}$ is suitable for the mean test, some partitions of J have advantages to offer in later steps of RIMPAR. For example, if J_1 happens to be G-regular, then J_1 will not be partitioned later. To evaluate various partitions of J (with respect to $G_i(\cdot)$) we define the <u>partition efficiency</u> $\eta_{G_i}(J_1, J_2)$ by

$$\eta_{G_i}(J_1,J_2)=\frac{|J_1||J_2|}{|J|}(M_{G_i}(J_1)-M_{G_i}(J_2))^T(M_{G_i}(J_1)-M_{G_i}(J_2)). \qquad (3.4)$$

The partition efficiency indicates the improvement of the partition criterion due to partitioning J into $\{J_1, J_2\}$. This can be shown by referring to Eqn. 2.25 and noting that

$$V_{G_i}(P_j)-V_{G_i}(P_{j+1})=\frac{1}{|I|}\ \eta_{G_i}(J_1,J_2)\ ,$$

when P_{j+1} is derived from P_j by replacing **J** by $\{J_1, J_2\}$.

To achieve partitions with high efficiency, RIMPAR tries several partitions of each J, and picks the partition with the highest efficiency. The trial partitions are generated by (K_D-1) equally-spaced horizontal lines, and (K_D-1) equally-spaced vertical lines. Here K_D is the same as in Eqn. 2.26.

The test for G-regularity is applied only to the most efficient partition of J.

3.2.4 Storage and Retrieval of Subimages

In the computer implementation of the G-regularity test, much processing time is used to retrieve the gray levels of J, and various subimages of J, for use in calculating the $\eta_{G_i}(J_1, J_2)$ and T^2. The following steps are taken to minimize retrieval time.

(1) The total image I is initially stored on tape. Before partitioning, I is transferred to a disc that allows almost direct access to any line of gray-levels.

(2) If at any point in RIMPAR we have $|J| <$ MAXCOR, then the gray-levels of J are transferred to core. All subimages of J can then be retrieved from core instead of disc. Arrays in the FORTRAN program are dynamically allocated to maximize MAXCOR, the amount of core available for image storage.

(3) As J is retrieved from either core or disc, the sums and sums of squares of the gray-levels of blocks of J formed by $2(K_D-1)$ equally-spaced horizontal and vertical lines are calculated. These partial sums are stored in core and are used to calculate the $\eta_{G_i}(J_1, J_2)$'s for the $2(K_D-1)$ trial partitions, and T^2 for the most efficient partition.

3.3 Choice of MINSIZE

In an ideal image, MINSIZE can be made arbitrarly small to limit partition errors due to blocks that overlap object boundaries. In a real image, it is still desirable to minimize overlap errors, but if MINSIZE becomes too small, the mean estimates necessary for deciding G-regularity become poor. Consider a partition block J. For any gray-level function $g(\cdot)$ we can form an estimate $\hat{M}_g(J)$ of $M_g(J)$ as follows:

$$\hat{M}_g(J) = \frac{1}{|J|} \sum_{X \in J} g(X).$$

Here we use $|J|$ to stand for the number of points in J. We define a block error by

$$E_g(J)=E\left[(\hat{M}_g(J)-M'_g(J))^2\right], \tag{3.5}$$

where $E[\cdot]$ is expected value and $M'_g(J)=M_g(O_k)$ where k satisfies

$$|O_k \cap J| = \underset{i \le j \le m}{MAX} |O_j \cap J|. \tag{3.6}$$

We now derive a rough approximation $\hat{E}_g(J)$ to $E_g(J)$, and we will select MINSIZE to minimize $\hat{E}_g(J)$.

The value of $E_g(J)$ for a particular J is highly dependent on whether J is g-regular. Therefore to find an expression for $E_g(J)$ for a "typical" J, we consider two errors: $E_g(J/R)$, the error for a g-regular J, and $E_g(J/\overline{R})$, the error for a non-g-regular J. Thus we write

$$E_g(J)=E_g(J/R)\,\mathrm{Prob}(R)+E_g(J/\overline{R})\,\mathrm{Prob}(\overline{R})=(E_g(J/\overline{R})-E_g(J/R))\,\mathrm{Prob}(\overline{R})+E_g(J/R) \tag{3.7}$$

where R and $\bar{R}$ stand for the events that J is g-regular and J is not g-regular respectively.

Expanding Eqn. 3.5 we get

$$E_g(J)=E\left[\left((M_g(J)-\hat{M}_g(J))+(M_g(J)-M'_g(J))\right)^2\right]=E\left[(\hat{M}_g(J)-M_g(J))^2+(M_g(J)-M'_g(J))^2\right]$$

From the well-known variance of the sample mean [22] we get

$$E_g(J)=\frac{S_g^2(J)}{|J|}+(M_g(J)-M'_g(J))^2 \tag{3.8}$$

Now if J is g-regular, $M_g(J)=M'_g(J)$, so

$$E_g(J/R)=\frac{S_g^2(O_k)}{|J|} \tag{3.9}$$

where k is defined in Eqn. 3.6. To form a rough approximation to Eqn. 3.9, we substitute Q for $S_g^2(O_k)$, where Q is a "typical" object variance. For |J| we substitute the smallest possible |J|, $(\text{MINSIZE}/K_D)^2$. This results in

$$\hat{E}_g(J/R)=\frac{K_D^2 Q}{(\text{MINSIZE})^2} \tag{3.10}$$

If J is not g-regular, then we assume that J overlaps two objects O_k (see Eqn. 3.6) and O_q. Therefore

$$M_g(J)=\frac{|O_k\cap J|}{|J|}M_g(O_k)+\left(1-\frac{|O_k\cap J|}{|J|}\right)M_g(O_q) \text{ and } M'_q(J)=M_g(O_k)$$

so

$$(M_g(J)-M'_g(J))^2=\left(1-\frac{|O_k\cap J|}{|J|}\right)^2(M_g(O_k)-M_g(O_q))^2$$

In order to roughly approximate Eqn. 3.8 when J is not g-regular, we substitute c Q for $S_g^2(J)$, where c is called the <u>contrast ratio</u> and is defined by

$$c=\frac{S_g^2(J)}{S_g^2(O_k)}$$

where J is a typical non g-regular block and O_k is a typical object. The parameter C depends basically on how different neighboring objects are. We also assume that $(M_g(O_k)-M_g(O_q))^2 \approx S_g(J)$, so cQ is also substituted for $(M_g(O_k)-M_g(O_q))^2$. For $|J|$ we again substitute the smallest $|J|$, $(MINSIZE/K_D)^2$. Thus we have

$$\hat{E}_g(J/\bar{R}) = \frac{cK_D^2Q}{(MINSIZE)^2} + \frac{cQ}{4} \tag{3.11}$$

where for $(1-\frac{|O_k \cap J|}{|J|})^2$ we have substituted its greatest possible value, 1/4.

To approximate Prob$(\bar{R})$ we consider a simple model in which a non-g-regular block J is its largest possible size, a square MINSIZE x MINSIZE. We consider a typical object O to be a B x B square. Assuming the upper left hand corner of J is equally likely to coincide with any point of O,

$$\text{Prob}(\bar{R}) \approx \frac{2\ MINSIZE}{B} \tag{3.12}$$

Now we can approximate Eqn. 3.7 using Eqns. 3.10, 3.11, and 3.12:

$$\hat{E}_g(J) = \left(\frac{cK_D^2Q}{(MINSIZE)^2} + \frac{cQ}{4} - \frac{K_D^2Q}{(MINSIZE)^2}\right)\left(\frac{2\ MINSIZE}{B}\right) + \frac{K_D^2Q}{(MINSIZE)^2}$$

$$= \frac{Q}{2B}\ C(MINSIZE) + \frac{4(c-1)K_D^2}{MINSIZE} + \frac{2K_DB}{(MINSIZE)^2} \tag{3.13}$$

To choose a value for MINSIZE, values for c and B are specified using knowledge of the image under consideration and K_D is chosen to give efficient partitions with reasonable computation time. Next $\hat{E}_g(J)$ is calculated for MINSIZE=1,2,...,B; and the value for which Eqn. 3.13 is a minimum is chosen.

4. APPLICATION RESULTS

We have presented an algorithm for dividing a multispectral image into parts, where each part is likely to correspond to a single target entity. The algorithm has been implemented on an IBM 360/67 general-purpose digital computer at the Laboratory for Applications of Remote Sensing, Purdue University. RIMPAR has been applied to several types of multispectral images, and details of these experiments are described elsewhere [23].

An indication of the quality of the partitions produced by RIMPAR is given by the following results. RIMPAR partitions were compared to partitions produced by equally-spaced horizontal and vertical lines for two images. Image I_1 was scanned from an airplane and consists of 40,000 image points. Image I_2 was scanned from the ERTS-1 satellite and has 12,000 image points. Both images depict agricultural scenes, but the ground resolution of I_1 is about $10m^2$, and for I_2 it is about $6000m^2$.

We approximate the partition criterion $V_g(P)$ (Eqn. 2.8) by using number of points for area, and the sample variance of each block I_j for $S_g^2(I_j)$. The approximate criterion $\hat{V}_g(P)$ was calculated for a RIMPAR partition and several regular-grid partitions of I_1 and I_2. The two partition types were compared by observing the number of partition blocks required by each to give a particular approximate criterion value.

For I_1 the RIMPAR partition P_R achieved $\hat{V}_g(P_R)=324.0$ with

595 blocks, and the regular-grid partition P_G achieved $\hat{V}_q(P_G)=324.1$ with 1932 blocks. For I_2, $\hat{V}_q(P_R)=6.21$ with 340 blocks, and $\hat{V}_q(P_G)=6.20$ with 1034 blocks. The RIMPAR processing time for I_1 was 3.89 CPU minutes, and for I_2 it was 1.0 CPU minute.

Nearly equal criterion values indicates that for both images, P_R and P_G represent about equally the information contained in the images. Since P_R has many fewer blocks, the average block size is much less in P_R than in P_G. Any processing task for which a partitioned image is an input works better with larger blocks, assuming the blocks preserve the information in the image. Therefore we conclude that for this data RIMPAR forms better partitions than the regular-grid algorithm.

The regular-grid partition P_G can be used as an input to the LBLOCK algorithm described in the INTRODUCTION. This will produce a new partition P'_G such that $V_q(P'_G)$ may be only slightly higher than $V_q(P_G)$. However, the small size of the blocks in P_G will probably lead to many LBLOCK errors; that is, errors in deciding the similarity of adjacent blocks. LBLOCK could also be applied to P_R, and the larger average block size would tend to minimize LBLOCK errors.

REFERENCES

[1] R. A. Holmes and R. B. MacDonald, "The Physical Basis of System Design for Remote Sensing in Agriculture", Proc. IEEE, Vol. 57, No. 4, April 1969, pp. 629-639.

[2] P. E. Anuta and R. B. MacDonald, "Crop Surveys from Multiband Satellite Photography Using Digital Techniques", Remote Sensing of Environment 2, 1971, pp. 53-67.

[3] K. S. Fu, D. A. Landgrebe, and T. L. Phillips, "Information Processing of Remotely Sensed Agricultural Data", Proc. IEEE, Vol. 57, No. 4, April 1969, pp. 639-653.

[4] ERTS Data Users Handbook, Document No. 71SD4249, NASA Goddard Space Flight Center, Greenbelt, Maryland.

[5] P. J. Ready, "Multispectral Data Compression Through Transform Coding and Block Quantization", TR-EE 72-2, Purdue University, Lafayette, Indiana, 1972.

[6] T. V. Robertson, "Comparison of Gaussian and Linear Classifiers in Multispectral Image Processing", Technical Memorandum T-4, Laboratory for Applications of Remote Sensing, Purdue University, Lafayette, Indiana.

[7] K. Shanmugam, R. M. Haralick, and R. Bosley, "Land Use Classification Using Texture Information in ERTS MSS Imagery", Tech. Rept. No. 2262-1, University of Kansas Center for Research, Inc., January, 1973.

[8] L. Sayn-Wittgenstein, "Patterns of Spatial Variation in Forests and other Natural Populations", Pattern Recognition, Vol. 2, No. 4, December 1970, pp. 245-253.

[9] A. G. Wacker and D. A. Landgrebe, "The Minimum Distance Approach to Classification", LARS Information Note 100771, Purdue University, Lafayette, Indiana.

[10] A. Rosenfeld, Picture Processing by Computer, Academic Press, New York, 1969, pp. 127-133.

[11] K. K. Pingle, "Visual Perception By a Computer", in Automatic Interpretation and Classification of Images, A. Grasselli (ed), Academic Press, New York, 1969, pp. 277-284.

[12] M. H. Hueckel, "An Operator Which Locates Edges in Digital Pictures", J. ACM, Vol. 18, No. 1, January 1971, pp. 113-125.

[13] A. G. Wacker, "A Cluster Approach to Finding Spatial Boundaries in Multispectral Imagery", LARS Information Note 122969, Purdue University, Lafayette, Indiana.

[14] A. Rosenfeld, M. Thurston, Y-H Lee, "Edge and Curve Detection: Further Experiments", IEEE Trans. Comp., Vol. C-21, No. 7, July 1972, pp. 677-715.

[15] C. T. Zahn, "A Formal Description for Two-Dimensional Patterns", Proc. First Joint International Conf. on Artificial Intelligence, 1969, pp. 621-628.

[16] H. Freeman, "Boundary Encoding and Processing", in Picture Processing and Psychopictorics, B. S. Lipkin and A. Rosenfeld (eds.), Academic Press, New York, 1970, pp. 241-266.

[17] E. M. Rodd, "Closed Boundary Field Selection in Multispectral Digital Images", IBM Publication No. 320. 2420, January, 1972.

[18] C. R. Brice and C. L. Fennema, "Scene Analysis Using Regions", Artifical Intelligence I, 1970, pp. 205-266.

[19] G. Hadley, Linear Algebra, Addison Wesley, 1961, pp. 175-176.

[20] T. W. Anderson, An Introduction to Multivariate Statistical Analysis, John Wiley and Sons, New York, 1958, pp. 108-109.

[21] A Hald, Statistical Theory with Engineering Applications, John Wiley and Sons, New York, 1952, p. 384.

[22] A. Papoulis, Probability, Random Variables, and Stochastic Processes, McGraw Hill, New York, 1965, p. 246.

[23] T. V. Robertson, "Extraction and Classification of Objects in Multispectral Images" , in preparation.