Fractal Analysis of a Classified Landsat Scene

Lee De Cola
Department of Geography, University of Vermont, Burlington, VT 05405

ABSTRACT: Remotely sensed images tend to be spatially very complicated, revealing regions of homogeneously classified pixels with quite convoluted perimeters. Fractal analysis, the study of complicated phenomena manifesting self-similarity at many scales, is suited to the description of the form and sizes of these regions. I apply fractal analysis to the patterns created by eight land-cover classes from a Landsat TM image of northwest Vermont. The results suggest that forests manifest high fractal dimension and large regions, agricultural activities have large regions with fractal dimension inversely related to the intensity of cultivation, and urban land cover yields small regions with relatively high fractal dimension. Analysis of the individual urban regions provides a data structure in the form of a raster-based GIS which can be used to investigate the location and description of individual regions and to diagnose the reliability of the classification and labeling processes.

INTRODUCTION

THE OBJECTIVE OF THIS PAPER is to demonstrate the applicability of fractal concepts to the spatial analysis of classified digital imagery. The paper is organized as follows. In the next section I use an example of an ideal fractal process to illustrate the theory underlying the analytical technique to be developed. Next, I introduce an algorithmic approach to three analytical steps: the segmentation of a classified image into homogeneous regions, the estimation for each class of the fractal dimension of regional perimeters, and the determination of a parameter describing regional sizes. Section four presents the empirical analysis of the eight land-cover classes from a Landsat Thematic Mapper image of northwest Vermont. This discussion provides fractal parameters which can be related to the classes, as well as a database containing locations and descriptions of the individual regions (clusters of pixels) from each class. I conclude by outlining a complete image segmentation/fractal analysis/GIS scheme that could profit from such advanced computing techniques as parallel processing.

FRACTALS AND REMOTE SENSING

REGIONS IN THE SPATIAL ANALYSIS OF DIGITAL DATA

The processing of digital image sensed data relies heavily on the analysis of energy spectra in order to classify pixels and to make judgments about the appropriateness of class labels. A great deal of attention in the literature is therefore given to quite sophisticated multidimensional techniques for the examination of patterns in the *spectral* space spanned by the various sensor bands (Richards, 1986). The pattern recognition literature, however, concentrates more on the investigation of the *spatial* patterns made by the classes themselves in the Euclidean space ($\mathcal{R}^2$ or the real plane) in which the image lies (NASA, 1988, chapter 7). This distinction is not meant to deny the existence in remote sensing analysis of such approaches as filtering and fourier techniques, which examine both neighborhood patterns at the small scale as well as texture and orientation characteristics at the large (Whalley and Orford, 1982). But it must be admitted that comparatively little use is made of much of the rich information that can be extracted from the spatial as opposed to the spectral characteristics of a scene.

An important reason for the relatively restricted extent of explicitly spatial techniques is the daunting complexity and apparent noisiness of images. Figure 1, for example, shows the patterns made by each of the eight classes of 256^2 pixels centered around St Albans, Vermont (these images are subscenes from the full study area discussed below). The human eye has little trouble distinguishing among the various textures, shapes, and forms in these pictures; yet it is at present a major challenge to bestow even rudimentary visual sophistication upon machines (Hillis, 1985). The first characteristic to appreciate about such patterns, however, is that they are formed of a collection of distinct regions of different sizes and circumscribed by complicated perimeters of various lengths. Let us focus, therefore, on the numbers, sizes, and perimeters of regions formed by clusters of homogeneous pixels in the digital (or map) space $I \times I = I^2$ (a subset of $\mathcal{R}^2$).

THE FRACTAL DIMENSION OF REGIONAL PERIMETERS

A fractal has been defined recently by Benoit Mandelbrot as a pattern "made of parts similar to the whole in some way" (quoted in Feder (1988), p. 11). Phenomena displaying apparent self-similarity are abundant in the real world. Observation reveals that large waves are made up of smaller waves and ripples, clouds can be decomposed into smaller clouds and wisps, trees are made up of hierarchical systems of trunks-branches-twigs-leaves-veins, and so forth. Current fractal research in the spatial sciences makes use of the fact that complicatedness is a direct consequence of the operation of many different spatial processes at a wide range of scales. Mark and Aronson (1984), for example, use fractal dimension to describe terrains; Arlinghaus (1985) demonstrates the fractal nature of central place systems; and Plotnick (1986) points out the fractal characteristics of sedimentation.

In what follows I shall explore the application of fractal concepts to regions, which may be defined as connected sets of points, e.g., pixel locations, for which area is an appropriate measure. A natural way to describe and compare regions is by measuring area, which in the case of a digital image would be a count of the number of pixels in the region multiplied by the area of each pixel. Another regional characteristic is shape, which may be defined as the degree of similarity between the region and some simple geometric figure. Indices of shape have a venerable geographic history (Haggett *et al.*, 1977, section 9.6), but shape is by no means a simple concept. For digital spatial data the concept of shape requires a measurement of regional perimeter, which may be defined as the number of line segments bounding the region.

A common example of the investigation of regions is the fractal analysis of islands and continents, which are bounded by coastlines having a fractal dimension D, defined below (Mandelbrot, 1982, chapter 5). In fact, a coastline is simply one kind of perimeter — in this case an isoline at sea level — having fractal dimension ranging from $D = 1$ (a straight line) to $D = 2$ (an infinitely complicated and therefore plane-filling curve) (Goodchild and Dubuc, 1987; Burrough, 1986). When the scale of analysis is increased, coastlines reveal three important and

St. Albans study area
CLASS=URBAN

St. Albans study area
CLASS=SOFTWOODS

St. Albans study area
CLASS=HAYLAND

St. Albans study area
CLASS=HARDWOODS

St. Albans study area
CLASS=WATER

St. Albans study area
CLASS=BARE,CORN

St. Albans study area
CLASS=BRUSH,WETLANDS

St. Albans study area
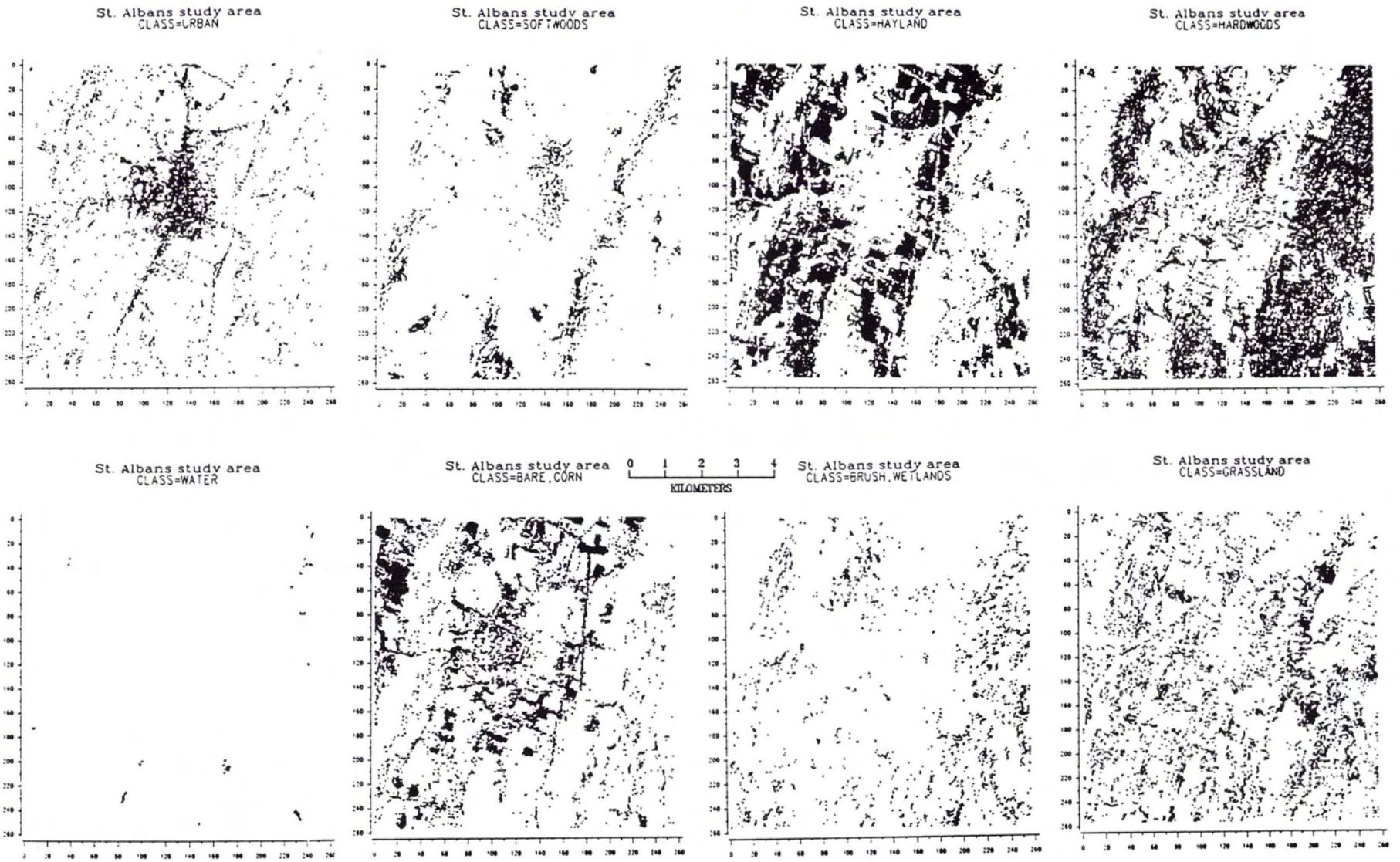CLASS=GRASSLAND

0   1   2   3   4
KILOMETERS

FIG. 1. Regions generated by eight land-cover classes in a 256- by 256-pixel window around St Albans, Vermont, from the full study area.

related characteristics. First, as Mandelbrot observes above, these curves are self-similar: large scale patterns look very much like small scale patterns. Second, these curves are quite complicated, so that lengths measured at large scales are greater than lengths measured at small scales. Third, these regions tend to have no characteristic size. I shall discuss this third characteristic below.

Consider for the moment the two sets of regions shown in Figure 2: squares and so-called Koch quadratic figure (Mandelbrot, 1982, chapter 12; Feder, 1988, chapter 12). These "regions" may be measured in terms of AREA and PERIMETER. On the one hand, note that for squares -- regions with smooth perimeters -- it is the case that PERIMETER $\propto \sqrt{\text{AREA}}^1$. This relationship holds for all regions of fractal dimension $D = 1$, and for squares in particular it is the case that PERIMETER $= 4\sqrt{\text{AREA}}^1$.

Now consider, on the other hand, the Koch quadratic regions, which have been constructed to have the same AREAs as the squares — 16 and 256, respectively — but with significantly more convoluted perimeters, so that PERIMETER $\propto \sqrt{\text{AREA}}^{3/2}$. Consequently, large Koch regions reveal more detail than do small regions, even though the former are constructed from the latter. In other words, if we think of the larger figures as "growing" out of the smaller figures (perhaps as a specific type of land cover spreads), then for the Koch model perimeter grows faster than it does in the case of growing squares. In fact, these Koch regions have been constructed with PERIMETER $= 4\sqrt{\text{AREA}}^{3/2}$.

To make this argument more formal, for any region let $s =$ AREA (or, more generally, "size") and $p =$ PERIMETER. If a given region is bounded by a smooth perimeter of fractal dimension $D = 1$, then it is the case that

$$p = c\sqrt{s} \qquad (1)$$

where $c = 4$ for a square, $c = 2\sqrt{\pi}$ for a disk, and so forth. In such cases — where we assume that any imaging system of however small a resolution would continue to show a smooth perimeter — c is an index of shape.

| AREA: | 16 | 256 |
| PERIMETER: | 16 | 64 |

| AREA: | 16 | 256 |
| PERIMETER: | 32 | 256 |

FIG. 2. Comparisons of measurements of squares ($D = 1$) and Koch regions ($D = 3/2$).

For a *fractal* phenomenon, however, it is the case that

$$p = c\sqrt{s}^D \qquad (2)$$

where $1 \leq D \leq 2$. For smooth shapes such as circles and squares (which might be called "degenerate" fractals) $D = 1$, as in Equation 1, while $D = 3/2$ for the Koch model, and $D = 2$ in the case of infinitely complicated feature whose perimeter fills the plane. Mandelbrot (1982, chapter 12) uses the term "fractal dimension" to characterize the parameter D.

It should be emphasized that two spatial indices are therefore being distinguished here: c, a measure of overall regional *shape* (which, because $c = 4$ in both cases, obviously does *not* distinguish between the square and the Koch regions!) and fractal dimension D, an index of what I call regional *form* or perimeter complicatedness. The Koch model illustrates the essential self-similarity of ideal fractals: each increase in map scale (in this case by a linear factor of 4) reveals a longer perimeter (here by a factor of $(4)^{3/2} = 8$) in the case of complicated figures than in the case of simple figures. In other words, fractal phenomena for which $D > 1$ have the cartographically commonplace characteristic of revealing more detail at a larger scale. Moreover, at a *given* scale, larger features manifest significantly longer perimeters (by the exponent D) than do smaller. It is these inter-feature comparisons of area and perimeter that I shall use below to estimate regional fractal dimension D according to Equation 2.

TECHNIQUE

IMAGE SEGMENTATION AND THE CREATION OF REGIONS

Perhaps the clearest empirical example of the kind of argument to be used here is Lovejoy's (1982) fractal analysis of cloud images. Clouds tend to be quite complicated, self-similar phenomena. Lovejoy studied digital Geostationary Operational Environmental Satellite (GOES) and radar images of clouds and rain areas using the perimeter/area relationship of Equation 2 and determined that the fractal dimension of the perimeter of these cloud/regions was $D = 1.73$ over a range of cloud sizes from less than 1 km² to over 10^6 km². Lovejoy's technique is formalized in this section and applied below to the analysis of land-cover classes from a Thematic Mapper image.

Let a digital space be $I^2 = I \times I$, where I is the set of integers, and let $W \subset I^2$ be a square image or "window" composed of N^2 pixels $x \in W$. Next, associate with each of these pixels an integer-valued index $i \in [0, m] \subset I$ representing one of m land cover classes (class 0 can be regarded as residual or unclassified). In the simplest case $m = 1$ so that the image is a black-and-white scene such as the 6 by 6 image shown in Figure 3; class 0 is colored white and class 1 colored black.

Each class i, therefore, imposes a *regionalization* on the image corresponding to the $m+1$-valued function $f: x \rightarrow [0, m]$. For any i a subset of this regionalization is the set $F_i = \{x:f(x) = i\}$, the set of all pixels to which the class i is ascribed. The first step in image description is usually the examination of the set $\{N_i\}$, the histogram of the image given by the number of pixels in each class i, and the corresponding fraction of such pixels $\{f_i\}$, where $f_i = N_i/N_i^2$. For the simple scene, $f_i = 9/36$.

Labeling may be thought of as the association of some spatial/spectral pattern with class i, but for the moment assume that the pixels are merely assigned to distinct categories: spectral, functional, or otherwise. Let us therefore focus on a single land-cover class and, for convenience, drop references to the index i. We shall return later to inter-class comparisons, once class i has been spatially characterized.

The following image segmentation algorithm partitions the regionalization F (referred to as F_i above) into n distinct and disjoint *regions* E_j such that $\cup E_j = F$, where $j = 1, \ldots, n$. Each
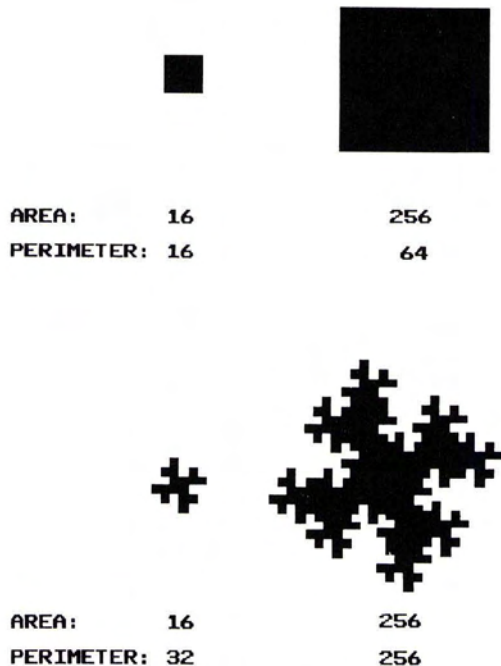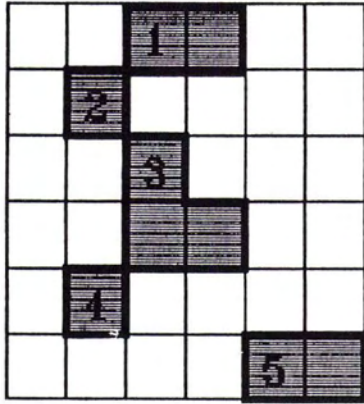
FIG. 3. A simple 6 by 6 two-class image with five regions.

region is grown using the recursive "rook's" rule that for any $x \in E_j$ then $y \in E_j$ if both $f(y) = i$ (the class index) and x and y are either row or column adjacent, i.e., y is a north, south, east, or west neighbor of x (Rosenfeld and Kak, 1982). In the present case I begin at x, the upper left pixel (leftmost pixel of uppermost row) of region E_j, and then recursively add pixels y to the set, sweeping down the image and across each row until all adjacent pixels have been accounted for. Note that the pixel x may be used to stand for unique location, which is of value in associating the region with some geographic feature such as place, link, physical region, etc.

Figure 3 illustrates this approach, providing $n = 5$ distinct regions for the 6×6 simple image. The algorithm produces a set of regions $\{E_j: j = 1, \ldots, n\}$ in which each E_j is a distinct "island" such that, for any $k \neq j$, E_j and E_k are not only disjoint (they share no pixels) but also unconnected (they share no edge). Obviously, the number of regions n is an important descriptor of the regionalization F.

DESCRIPTION OF THE REGIONS

Having identified the unique collection of pixels associated with each region E_j, we next determine the area and perimeter of each region. In the case of a specific imaging system, the area of each pixel (the instantaneous field-of-view of the delivered data product) may be standardized to unity. Then let the *area* or size of any region E_j be s_j, the number of pixels in the region. Next, the perimeter of a region E_j is a subset of the edges of its constituent pixels. Consider two adjacent pixels x and y and define a unit of perimeter to exist when $f(x) \neq f(y)$: pixels x and y are obviously not in the same region; therefore, an "active" edge exists between them. Let the *perimeter* p_j of region j be the number of active pixel edges bounding the region. For example, for a single-pixel region $p = 4$, while a one-class image will have a single region for which $p = 4N$. In the case of the simple image of Figure 3, E_3 is the largest region, $s_3 = 3$, and $p_3 = 8$: i.e.,

REGION j	ROW	COLUMN	AREA s_j	PERIMETER p_j
1	1	3	2	6
2	2	2	1	4
3	3	3	3	8
4	5	2	1	4
5	6	5	2	6

FRACTAL ANALYSIS OF REGIONAL PERIMETERS

Following from the discussion in the previous section, if we assume that digital regions, like the Koch regions of Figure 2,

have fractal characteristics over a wide range of scales, then it is possible to determine D for the collection of regions $\{E_j\}$ associated with a given class. The two parameters from Equation 2 may be estimated from the linear expression

$$\ln(p_j) = \ln(c) + D \ln(\sqrt{s_j}) + \epsilon \qquad (3)$$

where ϵ is an error term reflecting the homogeneity of the regions as well as the consistency of the classification process. This equation yields D for a specific class. I focus on D and not c becuase the former should reflect spatial pattern, in the sense that land covers having smooth perimeters should have $D \sim 1$ and covers with complicated perimeters should have $D > 1$. Two further details should be noted here. First, natural logarithms will be used in the analysis for computational speed: Figure 4 shows the natural logarithm of regional perimeter plotted against the log of regional area, as well as the regression line, for regions from the URBAN class in the Landsat image discussed below. Second, as in all regression modeling, residuals estimating ϵ will be useful in describing observations (the n regions) and diagnosing the effectiveness of the regression model (Equation 3).

PARETO ANALYSIS OF REGIONAL SIZE

Fractal analysis therefore provides a practical measure of the forms of regions generated by classification, but what of the sizes or areas of these regions $\{s_j\}$? In addition to self-similarity, another important characteristic of fractal phenomena is *scaling*, or the tendency of sets of fractals to have no characteristic or typical size (Feder, 1988, chapter 2). For fractal regions in particular, this means that increasing the size of a window W is likely to bring into view a larger region, while decreasing pixel size (or increasing the number of pixels in a given window) will always reveal smaller regions. The phenomenon of scaling allows us to describe the size distribution of regions.

Let the regions be ordered from largest to smallest by area s, and define rank

$$r_j = Nr(E:s > s_j) + 1 \qquad (4)$$

which is the number of regions larger than E_j. One statistic characterizing $\{s_j\}$ is the median $\tilde{s} = \{s: r_j = n/2\}$. But a typical problem with digital images is that $\tilde{s} = 1$ in many cases — i.e.,
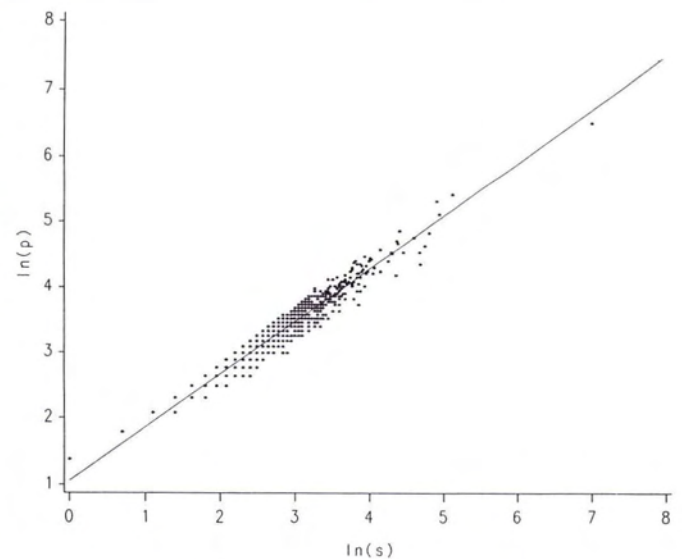

FIG. 4. Plot of natural logarithm of PERIMETER$_j$ by natural logarithm of AREA$_j$ for the URBAN class, with regression line.

more than half the regions are single pixels — so clearly $\bar{s}$ is likely to have little descriptive value.

Another way to characterize the size distribution of the n regions is $\max\{s_j\}$, the size of the region of rank $r_j = 1$, which statistic may be descriptive of the underlying pattern but obviously will be strongly determined as well by the specific window W chosen. A more effective tactic is to use the hyperbolic form of Equation 4, reflecting the size distribution of $\{s\}$, to examine the Pareto function describing the sizes of the regions from their unique ranks (Goodchild and Mark, 1987; De Cola, in press).

To do this, we capitalize on a consequence of scaling: that the largest region in W generally bears the same relationship in size to the next largest region as the 10th largest region does to the 20th: i.e., ratios of ranks tend to be similar to some power of the inverse ratios of sizes. More generally, this gives a hyperbolic relationship

$$s_j = A(r_j)^{-b} \qquad (5)$$

whose parameters may be estimated from the linear model

$$\ln(s_j) = \ln(A) - b \ln(r_j) + \epsilon \qquad (6)$$

which provides least-squares estimates both of the size parameter $a = \ln(A)$ (a prediction of $\max\{s_j\}$) as well as of b, the Pareto index of inequality. The Pareto parameter a is a convenient measure of the size of the regions $\{E_j\}$ and less sensitive (more statistically robust) than $\max\{s_j\}$. (See De Cola (1985) for empirical results from a rank-size analysis of city populations, and Arnold (1983, pp. 44, 202) for the details of the technique as well as the properties of the estimators.)

It should be noted that the regression Equation 6 is fitted to unique size groups rather than to each region; this avoids the problem of estimating a line through data which are heavily dominated by small (single- or few-pixel) regions, so that we are able to describe the larger regions of inherently greater interest. Figure 5 illustrates the data and regression line for the URBAN class from the Landsat image discussed below.

Spatial Statistics Describing Classes

We now have an ensemble of six spatial statistics describing the image W. Spectral image classification produces $i = 1, \ldots, m$ classes, whi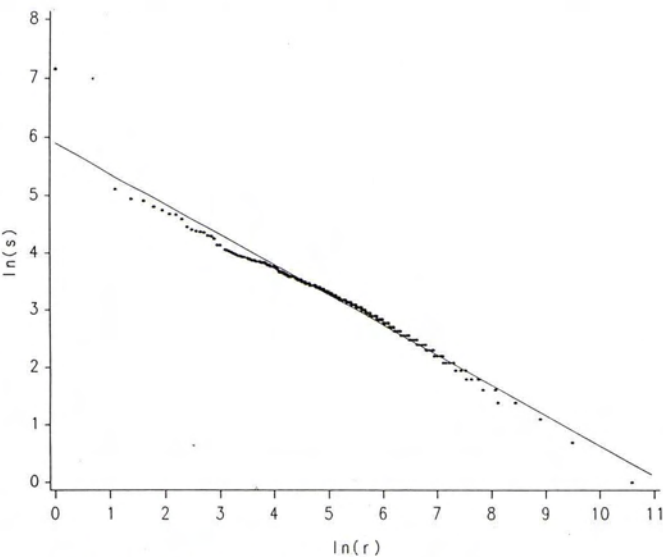ch can be described in terms of both (1) the number of pixels N_i and (2) the fraction f_i of the scene classified in class i. Image segmentation provides information about (3) the number of connected regions n_i and (4) the size of the largest region $\max\{s\}_i$. Finally, fractal analysis yields (5) fractal dimension D_i and (6) the Pareto size parameter a_i reflecting the areas of the regions generated by class i.

ANALYSIS OF THE LANDSAT IMAGE

The Northwest Vermont Study Area

The methodology described above was applied to a 26 October 1982 Landsat Thematic Mapper image of a 4200 km² area of northwest Vermont (Figure 6). The data, consisting of 2048^2 pixels each of size $(31.81 \text{ m})^2$, were geometrically corrected and subjected to supervised classification using the Interactive Digital Image Manipulation System (IDIMS) at the NASA Goddard Space Flight Center. Eight land-cover classes were ultimately created, shown in Plate 1. The image is spatially quite complicated, yet the analytical system outlined above is able to resolve this spatial pattern into distinct regional measurements.

Description of the Regions

Table 1 presents the ensemble of key indices for each of the eight land-cover classes in the image. The 3.7 million pixels of

FIG. 5. Plot of natural logarithm of AREAⱼ by natural logarithm of RANKⱼ for the URBAN class, with regression line.
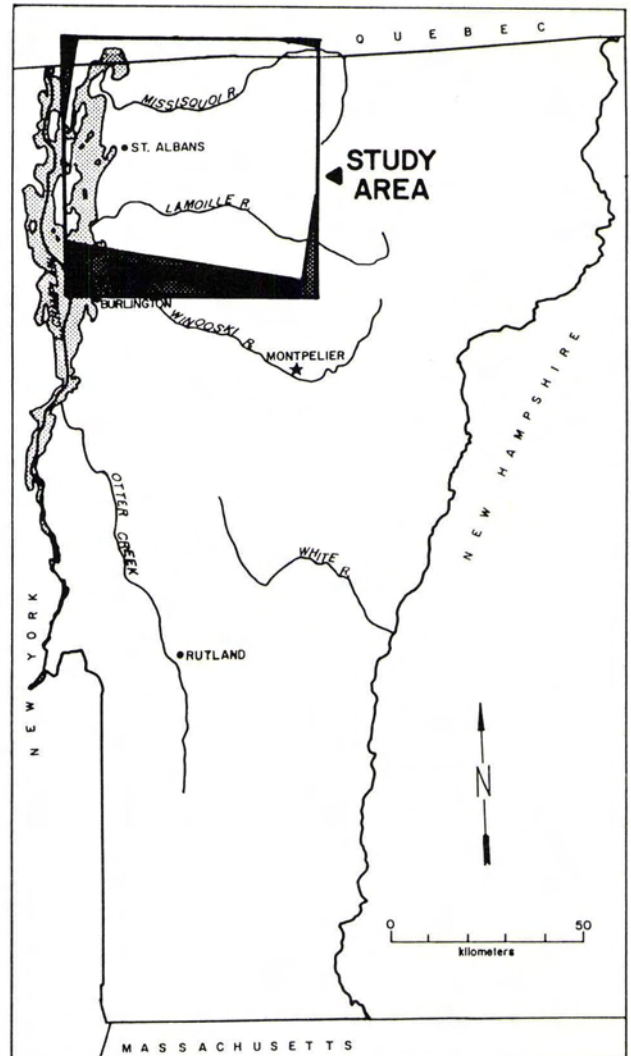

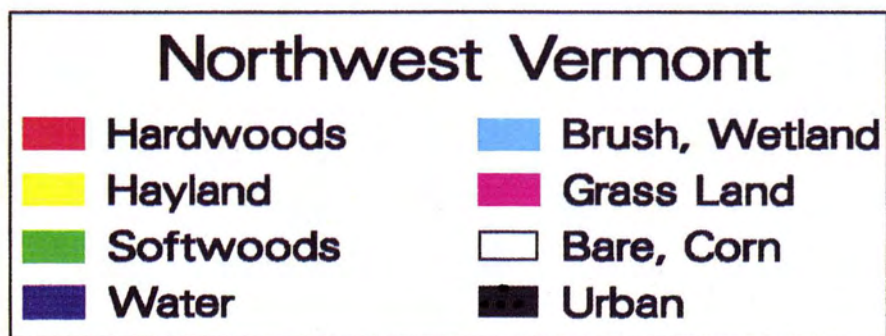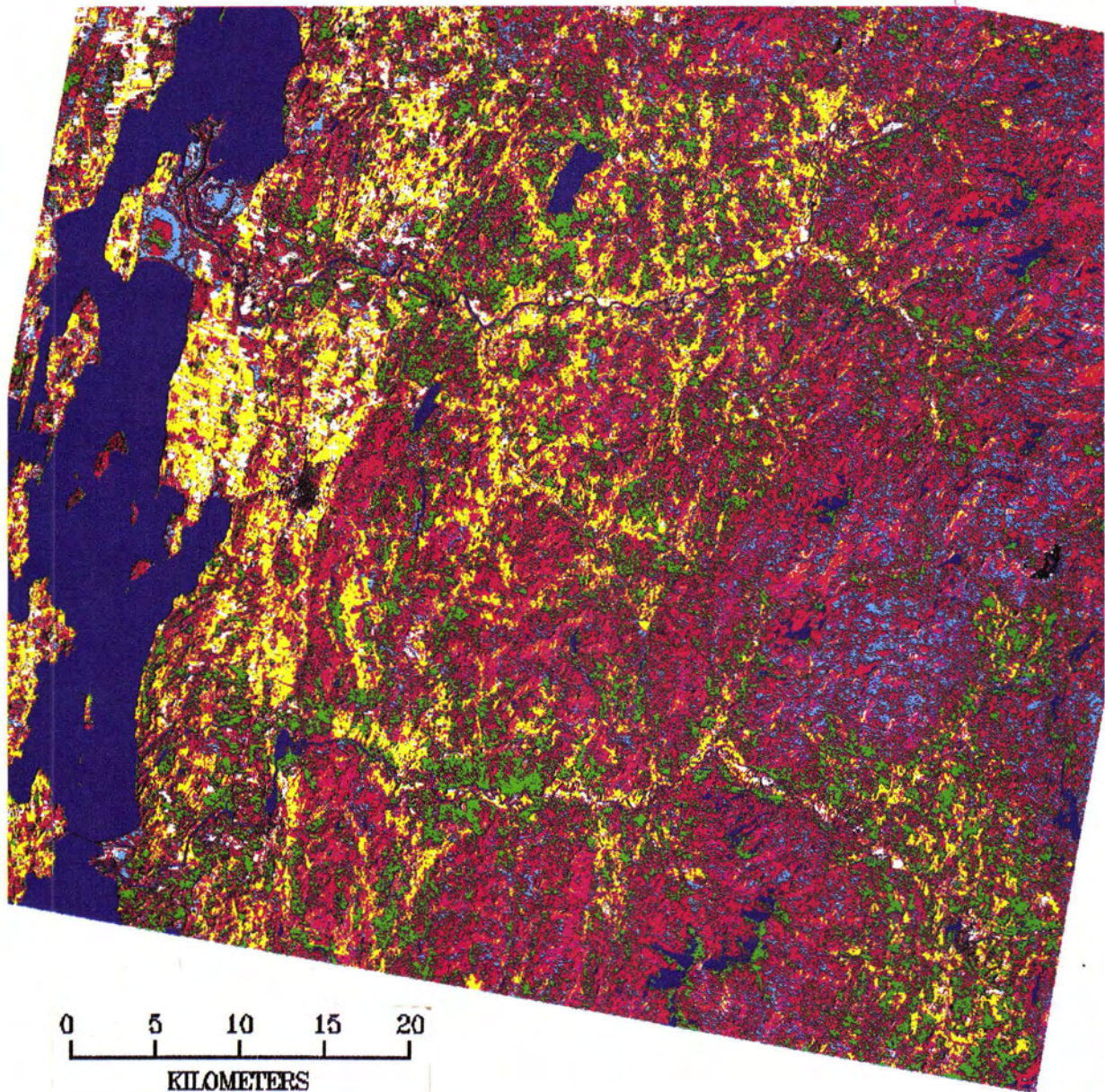FIG. 6. Location of the classified Landsat TM scene.

PLATE 1. Classified Landsat TM image for northwest Vermont.

the image have been aggregated into 480,000 regions whose number, size, and form describe the land-cover classes. Almost 40 percent of the pixels in the scene (but only 13 percent of the regions) are classified as HARDWOODS, largely representing the maple forests associated with much of the Vermont landscape. The HAYLAND class constitutes the second dominant land cover, which along with SOFTWOODS and HARDWOODS account for over 72 percent of the dry land cover in the scene. At the other

TABLE 1. FRACTAL ANALYSIS OF EIGHT LAND-COVER CLASSES.

i	Land Cover Label	Number of Pixels N_i	% of Image f_i	Number of Regions n_i	Pixels in Largest Region $max\{s\}_i$	Fractal Dimension D_i	Pareto a_i
1	HARDWOODS	1,445,027	39.5	98,061	205,188	1.812	11.370
2	HAYLAND	466,865	12.8	34,372	6,468	1.597	10.728
3	SOFTWOODS	447,906	12.2	56,442	4,067	1.617	10.484
4	WATER	398,640	10.9	5,775	335,472	1.354	9.495
5	BRUSH, WETLAND	377,652	10.3	104,216	2,936	1.687	8.154
6	GRASSLAND	234,685	6.4	91,884	380	1.646	6.278
7	BARE, CORN	199,221	5.4	46,476	2,483	1.447	8.955
8	URBAN	87,334	2.4	39,497	1,284	1.614	5.901
	TOTALS:	3,657,330	100.0	476,723			

extreme, only 2.4 percent of the scene is classified as URBAN, yet this class accounts for 8.3 percent of the regions. The largest number of regions, one-fifth of the total, are BRUSH/WETLAND, whose moderately sized patches tend to be quite dispersed within the scene, as are GRASSLAND regions, which tend to be small (the largest is only about 37 hectares). Still focusing on regional description, the WATER class manifests the fewest number of regions but accounts for the largest region: the 339 km² "blob" of Lake Champlain. Obviously, image regionalization reveals much about the spatial organization of the scene.

FRACTAL ANALYSIS

These observations are reinforced by the fractal results of Table 1, which provide insights about the classes and their labels. First, it should be noted that these eight sets of class estimates are all reasonably good: the minimum $R^2 = 0.91$ for the regression Equation 3 estimating D, and minimum $R^2 = 0.96$ for Equation 6 estimating a (D'Agostino and Stephens, 1986). Second, the two key parameters have a wide range of values: while D is constrained to be in the range $1 \leq D \leq 2$, actual values of D span almost half of this interval: $1.35 \leq D_i \leq 1.85$, and the range of the logarithmic size parameter a reflects a broad range of regional sizes from less than 0.4 km² to almost 90 km². Third, the fractal D and Pareto a parameters can be regarded as arraying the classes in a two-dimensional space "form" × "size," as in Figure 7, which can be used as a technique for evaluating the spatial structure of the classes as well as the appropriateness of the labels.

Figure 7 shows not only that the HARDWOODS class dominates the scene, but also that these regions tend to be largest and to

have the highest fractal dimension ($D \sim 1.8$); i.e., hardwood forests are usually extensive and have extremely complicated perimeters. In this respect, they contrast strikingly with the generally smaller (less than half as large) and significantly smoother SOFTWOODS regions, for which $D \sim 1.6$. HAYLAND regions have a relatively high D, as these fields are not intensively cultivated and therefore manifest more complicated perimeters than do BARE, CORN regions, which tend to have low D, reflecting the perimeter-conserving tendencies of agricultural development: farmers create rows and consequently blocks. The few WATER regions show the lowest D, associated with the lakes of northwest Vermont, and GRASSLAND regions are very much smaller than the other natural (non-agricultural, non-urban) regions, with BRUSH, WETLAND regions of intermediate size but quite high D. Finally, the smallest regions tend to be URBAN, which nevertheless have rather high values of D.

It should, of course, be borne in mind that this analysis specifically applies to this particular window, in this particular portion of the Vermont landscape, at this particular time, etc. But, in general, the results suggest that the most complicated (high-D) land-cover regions are formed by hardwood forests and the least complicated by agricultural activity (e.g., hay) and especially by intensively cultivated farming (crops). Urban land cover, while generally yielding the smallest regions, has a relatively high fractal dimension, strictly exceeded by only two other (and much larger) land covers. It is therefore clearly possible to associate land-cover patterns with fractal measurements.

REGIONAL ANALYSIS OF THE URBAN CLASS

Not only has fractal analysis yielded descriptive parameters for the aggregate classes, but image segmentation has also generated a spatial database which for the Vermont scene is a table of the locational and spatial characteristics of each region, a key element of any raster-based GIS (Lam et al., 1987). It is feasible to perform many of the usual GIS operations such as labeling, describing, overlaying, merging, etc. on these regions (Burrough 1986, chapter 5).

Worthy of detailed examination, therefore, is the set of regions labeled URBAN, which was formed from pixels showing the bare surfaces characteristic of urban areas (Lo and Welch, 1980; Haack, 1984). This class constitutes the smallest proportion of image pixels, yet gives rise to the third fewest number of regions. Figure 7 shows that D_{URBAN} is quite high, suggesting that URBAN regions have relatively complicated perimeters, while the URBAN regions are generally among the smallest as measured by a. The URBAN pixels are also interesting because this class can be expected to reflect intense human activity, making it possible to associate its regions with features identifiable in the landscape.

The largest 31 of the URBAN regions are listed along with the UTM coordinates of their locational pixels (minimum column within minimum row), their sizes, perimeters, and residual ratios
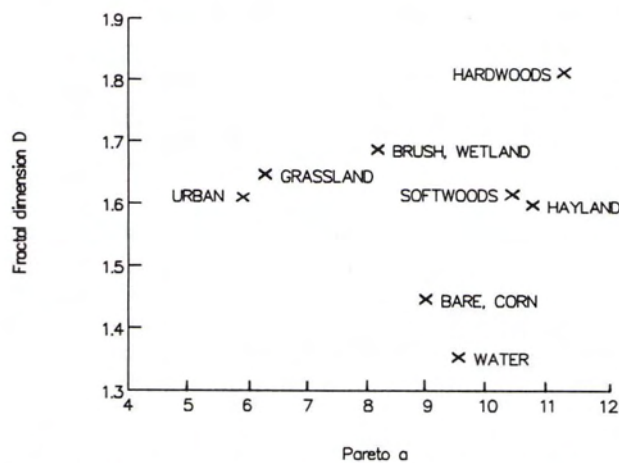

FIG. 7. Plot of fractal dimension D by Pareto scale parameter a for eight land-cover classes.

TABLE 2. URBAN REGIONS LARGER THAN 5 HECTARES IN THE DATABASE.

Region Rank r_j	UTM North	UTM East	Number of Pixels s_j	Perimeter p_j	Residual Ratio ρ_j	Place or Feature*
1	4964771	651591	1284	1470	1.58	VT St Albans
2	4960572	695966	1093	678	0.83	VT asbestos quarry
3	4982680	636768	167	226	1.26	VT Alberg
4	4937350	690781	140	168	1.08	VT Morrisville
5	4976095	648283	135	204	1.35	VT Swanton
6	4990155	683211	122	126	0.91	PQ field
7	4988437	642335	115	104	0.78	PQ field
8	4990378	654327	108	78	0.62	PQ field
9	4973773	635973	107	94	0.75	VT field
10	4982425	645643	99	118	1.00	VT lakeshore
11	4965375	651719	86	94	0.90	VT St Albans
12	4975872	644975	82	130	1.29	VT lakeshore
13	4986497	641189	80	108	1.09	PQ Desrauleau
14	4965884	643289	79	112	1.14	VT lakeshore
15	4987133	642685	78	66	0.68	PQ field
16	4938273	690877	74	92	0.99	VT Morrisville
17	4983698	645834	73	94	1.02	VT lakeshore
18	4962830	651973	70	82	0.92	VT St Albans
19	4945112	684133	63	98	1.20	VT Johnson
20	4990696	653913	63	70	0.86	PQ field
21	4966902	643480	58	74	0.97	VT lakeshore
22	4964166	650446	57	68	0.90	VT St Albans
23	4942058	640426	56	84	1.13	VT lakeshore
24	4988374	640108	55	86	1.18	PQ field
25	4991046	651401	54	82	1.14	PQ lakeshore
26	4990728	662534	53	70	0.99	PQ quarry
27	4971928	644466	52	68	0.97	VT lakeshore
28	4966997	651719	52	76	1.09	VT St Albans
29	4959458	694153	51	56	0.81	VT quarry
30	4985352	683688	51	88	1.28	VT Richford
31	4963848	650287	50	72	1.06	VT St Albans

*PQ = Québec.

TABLE 3. CROSSTABULATION OF URBAN REGION IDENTIFICATION AND RESIDUAL RATIO.

	FIELDS/ QUARRIES	LAKESHORE	SETTLEMENTS	TOTAL
$\rho \geq 1$	1	6	9	16
$\rho < 1$	9	2	4	15
TOTAL	10	8	13	31

$\chi^2 = 10.30$, $p < 0.01$.

(see below) in Table 2. Using 1979 Vermont Mapping Agency 1:5000-scale aerial photo maps, I have identified the features associated with each of the URBAN regions having an area $s_j >$ 50 hectares. Two regions stand out in the database: Region 1 is a large area representing 1.3 km² of adjacent pixels at St Albans, Vermont, the area shown earlier in Figure 1. Region 2, however, is a large asbestos quarry in the east central part of Plate 1. Even though these very different features occupy similar locations in spectral space, their spatial structure is radically different, and fractal analysis makes this contrast quantitative.

In fact, there are only 13 genuine urban regions identified in Table 2: six are associated with the settlement of St Albans, whose 1980 population was about 7000; six regions can be associated with other Vermont places and one with a town in Québec. But the other regions are *not* settlements: ten are fields and quarries of various sorts and eight are segments of the Lake Champlain shoreline. Clearly, we need a way to distinguish these regions spatially from one another, and fractal analysis suggests a technique.

An important bonus from fractal analysis is the set of residuals $\{e_j\}$ estimating the errors ϵ from Equation 3. These residuals may

be used to assess the degree to which a given region j belongs in class i. Using primes to denote predictions, for any region j let

$$e_j = \ln p_j - (\ln p_j)' \qquad (7)$$

and define a residual ratio

$$\rho_j = \exp(e_j) = p_j/p_j' \qquad (8)$$

reflecting the extent to which region j has a perimeter that is longer ($\rho_j > 1$) or shorter ($\rho_j < 1$) than expected. The extent to which $|\rho| > 1$ provides ancillary information about whether region j "belongs" in class i: if the class is regarded as a preliminary unlabeled collection of regions, then region j might be an "outlier" whereas, if the class is a labeled land cover category, then either class i or region j may be mislabeled (Richards et al., 1982).

These residual ratios ρ_j (which for each of the eight classes are normally distributed) are shown in Table 2 for the largest URBAN regions from the image. A perusal of the data suggests that there are important differences among the various URBAN regions identified, and this is reinforced in Table 3. Settlements (with typically very convoluted perimeters) and lakeshore regions (themselves quite complicated) tend to have large values of ρ while the more compact bare fields and quarries have lower values of ρ — in particular, the large asbestos quarry has the highest value of ρ in the database. Overall, this diagnostic index does a fairly good job of distinguishing the first two categories from the last, as illustrated by the value of the χ^2 statistic shown in Table 3.

CONCLUSION

This paper has demonstrated a methodology for segmentation of a classified digital image, description of a class-induced

regionalization, fractal analysis of perimeters and areas, and detailed examination of individual features. The concepts of self-similarity, fractal dimension D, and Pareto size parameter a have been shown to be of descriptive value in the analysis of digital images. Fractal analysis of the forms of perimeters from regions in northwest Vermont suggests that urbanization gives rise to more complicated forms than does intensive agriculture, with hardwood forests and water spanning the range of extreme D values. As for regional sizes, I find that certain natural land covers (e.g., forests) tend to produce larger regions than others (say grass), while agricultural activity produces larger regions than do urban processes. The technique also produces a grid-based GIS data structure based on regions, which may be used for the location and detailed examination and identification of important features. Individual regions themselves may provide important statistics for the diagnosis of mislabeling or misclassification.

The present system is a collection of image analysis algorithms, statistical procedures, and databases which requires a classified image; Figure 8 outlines how an integrated spatial analysis system works. First, the image is classified using conventional unsupervised techniques of clustering in spectral space. An initial image segmentation module produces from these classes training fields having desired size and form characteristics (Buchheim and Lillesand, 1987). The training fields are then used as initial input for supervised classification based on labeled land-cover classes which impose a partition of the image into second-round clusters to be subjected to fractal analysis of size and form. At this point the clusters, their locational, spectral, and spatial characteristics (size, perimeter) form a GIS data structure that can be supplemented with ancillary land-use information (Goodenough, 1988).

While the process can be regarded as complete, it is probable that critical evaluations of the first-round classification made by the GIS will call for further classification (dashed line), in the spirit of Bayesian analysis (Richards *et al.*, 1982; Coulson *et al.*, 1987). When the spatial/land-use classification is deemed sufficiently accurate and reliable for the purposes at hand, the system can be used for the production of the usual products expected from a grid-based GIS. A further improvement worthy of exploration is the incorporation of parallel processing into the region-growing and fractal algorithms, in which all pixels can associate simultaneously with their neighbors for the rapid analysis of a scene (Tofolli and Margolus, 1987).

In short, the system developed here makes practical use of a new and promising geometric concept for the following purposes: (1) to improve the spatial analysis, spectral classification, and functional labeling of remotely sensed data; (2) to address the task of making inferences about spatial process from spatial pattern, and (3) to link remote sensing with GIS through regions with fractal perimeters imaged at a specific scale.

ACKNOWLEDGMENTS

REFERENCES

Arlinghaus, S. L., 1985. Fractals Take a Central Place, *Geografiska Annaler*, 67:83–88.

Arnold, Barry C., 1983. *Pareto Distributions.* Fairland, Maryland: International Cooperative Publishing House.

Buchheim, Martin P., and Thomas M. Lillesand, 1987. A semi-automated training sample selector for multispectral land cover classification, *Technical Papers: 1987 ASPRS-ACSM Annual Convention*, Vol. 6:23–33.

Burrough, P. A., 1986. *Principles of Geographical Information Systems for Land Resources Assessment.* Oxford: Clarendon Press.

Coulson, Robert N., L. Joseph Folse, and Douglas K. Loh, 1987. Artificial intelligence and natural resource management, *Science*, 237:262–267.

D'Agostino, Ralph B., and Michael A. Stephens, 1986. *Goodness of Fit Techniques.* New York: Marcel Dekker.

De Cola, Lee, 1985. Lognormal estimates of macroregional city size distributions, 1950-1980, *Environment and Planning A*, 17:1637–1652.

———, 1989. Pareto and fractal description of regions from a binomial lattice, *Geographical Analysis.* 21(1):74–81.

Feder, Jens, 1988. *Fractals.* New York: Springer-Verlag.

Goodchild, Michael F., and D. M. Mark, 1987. The Fractal Nature of Geographic Phenomena, *Annals of the Association of American Geographers* 77:265–278.
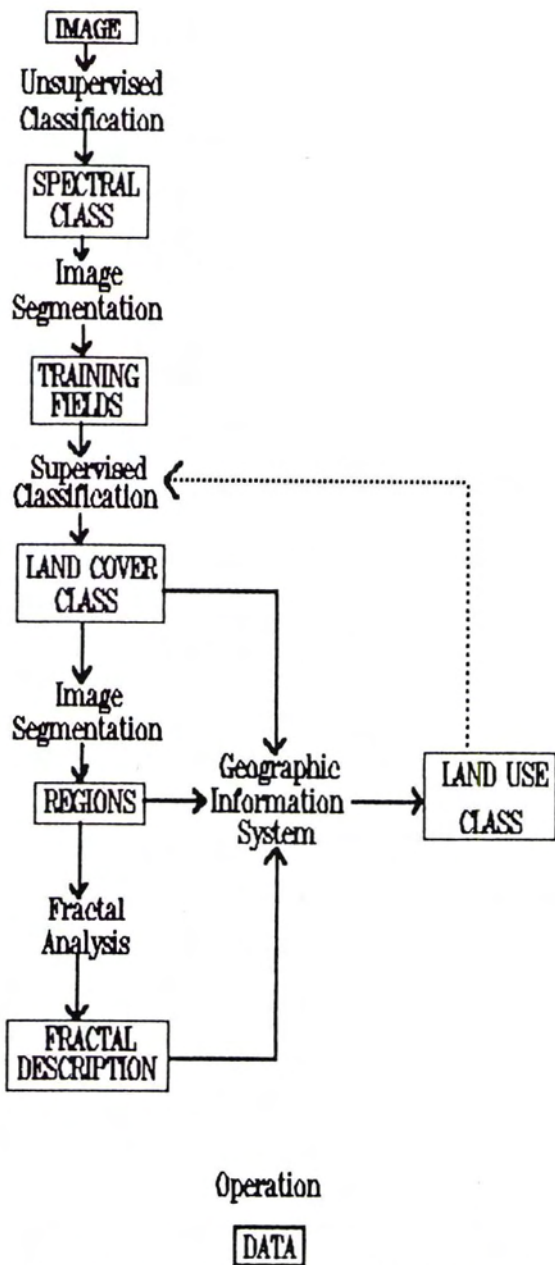

FIG. 8. Outline of a complete image classification, segmentation, fractal analysis, and geographic database system.

Goodchild, Michael F., and Odette Dubuc, 1987. A model of error for choropleth maps, with applications to GIS, *Autocarto* 8:165–174.

Goodenough, David D., 1988. Thematic mapper and SPOT integration with a geographic information system, *Photogrammetric Engineering and Remote Sensing*, 54(2):167–176.

Haack, Barry N, 1984. Multisensor Data Analysis of Urban Environments, *Photogrammetric Engineering and Remote Sensing*, 50(10):1471–1477.

Haggett, Peter, Andrew D. Cliff, and Allan Frey, 1977. *Locational Analysis in Human Geography*. London: Edward Arnold.

Hillis, W. Daniel, 1985. *The Connection Machine*. Cambridge, Massachusetts: MIT Press.

Lam, Nina Siu-Ngam, Paul J. Grim, and Farrell Jones, 1987. Data integration in geographic information systems: an experiment, *ASPRS/ACSM Proceedings GIS/LIS*, Vol. 5:53–62.

Lo, C. P., and R. Welch, 1980. Chinese urban population estimations, *Photogrammetric Engineering and Remote Sensing*, 46:246–253.

Lovejoy, S., 1982 Area-perimeter relation for rain and cloud areas, *Science*, 216:185–187.

Mandelbrot, Benoit, 1982. *The Fractal Geometry of Nature*. San Francisco: Freeman.

Mark, D. M., and P. B. Aronson, 1984. Scale dependent fractal dimensions of topographic surfaces with applications to computer mapping, *Mathematical Geology* 16/7:671–683.

NASA, 1988. *Earth Systems Science: A Closer View*. Washington: NASA.

Plotnick, Roy E., 1986. A fractal model for the distribution of stratigraphic hiatuses, *Journal of Geology*, 94:885–890.

Richards, J. A., 1986. *Remote Sensing Digital Image Analysis*. New York: Springer-Verlag.

Richards, J. A., D. A. Landgrebe, and P. H. Swain, 1982. A means for utilizing ancillary information in multispectral classification, *Remote Sensing of Environment* 12:463–477.

Rosenfeld, Azriel, and Arinash C. Kak, 1982. *Digital Picture Processing*, 2nd ed. New York: Academic Press.

Toffoli, Tommaso and Norman Margolus, 1987. *Cellular Automata Machines*. Cambridge, Massachusetts: MIT.

Whalley, W. B., and J. D. Orford, 1982. "Analysis of scanning electron microscopy images of particles by fractal dimension and Fourier analysis," *Scanning Electron Microscopy,*, 1982(2):639–647.

New Sustaining Members

Aerial Cartographics of America (ACA), Inc.
7000 Lake Ellenor Drive, Orlando, FL 32809
Telephone 407-851-7880

AERIAL CARTOGRAPHICS OF AMERICA (ACA), INC. is an Orlando, Florida-based company, originally established in 1973. After being sold and reorganized in 1983, ACA has grown to approximately 25 employees. Its PHOTOGRAPHIC DIVISION specializes in aerial photography and color and black-and-white photographic reproductions. The PHOTOGRAMMETRIC DIVISION produces all types of mapping, specializing in digital deliverables.

AERIAL PHOTOGRAPHY
The "in-house" flight department utilizes single- and multi-engine Cessna aircraft combined with Wild and Zeiss 6″ and 3 1/2″ camera systems for its mapping projects. Photographic assignments are accomplished with 6″ and 12″ systems supported by an array of specialized cameras for oblique assignments.

PHOTO LABORATORY
Color and black-and-white photographic reproductions are produced utilizing ten (10) dark rooms using six (6) 10″ x 10″ precision enlargers, three (3) processors, two (2) LogEtronic printers and the normal assortment of support photographic equipment.

In addition to the standard photographic reproductions, rectified mylars and mosaics, ACA specializes in color rectified enlargements to 50″ x 100″ for their engineering clients and for the photogrammetric industry.

TOPOGRAPHIC MAPPING
Multi-shifted, Kern DSR 11 plotting systems with MAPS 300 interfaced with CADMAP edit stations, form the nucleus of the ACA mapping department. Flat bed and drum plotters supported by miscellaneous pc's complete the hardware necessary for producing a variety of mapping deliverables.

Packaged and specially designed software permits transfer and conversion of data for delivery of many clients CAD systems, including but not limited to, Intergraph, AutoCAD, VANGO and Standard Interchange Format.

PHILOSOPHY
According to Charles (Chuck) Woodward, ACA's president, the success of the company is mainly a result of good client liaison and a quality product delivered in a timely manner.

Trimble Navigation, Ltd.
645 North Mary Avenue, PO Box 3642, Sunnyvale, CA 94088-3642
Telephone 408-730-2900; fax 408-730-2997

TRIMBLE NAVIGATION, LTD. designs, manufactures, and sells a wide selection of Global Positioning System (GPS) receivers for marine, airborne, and land applications. These field-proven GPS receivers range from 12-channel precision survey systems with millimeter accuracy to low-cost 2-channel receivers for vehicle tracking and navigation. Trimble GPS equipment is sold worldwide in 40 countries.

For GIS applications in building a GPS control network, the GPS survey grade receivers will be an indispensable tool. The lead survey product is the model 4000ST GPS Field Surveyor. For other GIS applications in database building and maintenance or ground referencing of satellite imagery, the navigation-accuracy GPS Pathfinder receivers will be most appropriate.

The picture shows the GPS Pathfinder in a field application. This portable, light-weight GPS receiver has software allowing position data to be collected in the field data logger and transferred into a GIS database. Most popular GIS data formats are supported, including ARC/INFO, GRASS, DLG-3, MOSS, and AutoCAD.

Trimble Navigation has a GPS primer called "GPS — A Guide to the Next Utility" which is available free of charge. Contact your nearest Trimble Corporate office or William F. Dupin, Vice President, Sales.