SRI International

COARSE CODING FOR MATERIAL AND OBJECT IDENTIFICATION

Technical Note 442

July 1988

By: Kenneth I. Laws, Computer Scientist
Artificial Intelligence Center
Computer and Information Sciences Division

SRI Projects 2000 and 8388

The work reported herein was supported by the Defense Advanced Research Projects Agency under Contract Nos. MDA903-86-C-0084 and DACA76-85-C-0004.

SRI International®

333 Ravenswood Ave. • Menlo Park, CA 94025
(415) 326-6200 • TWX: 910-373-2046 • Telex: 334-486

Abstract

A new coarse-coding technique is presented for labeling image pixels and regions to match exemplars or multivariate material signatures. This multinomial classification method can be used for object cuing and tracking, as well as for material identification and image segmentation. Pixels are classified—and classification reliability can be estimated—with only single-band histograms and one pass through each image band. An example of four-class labeling illustrates the power of this two-level classification algorithm.

Contents

1 Introduction 1

2 Labeling Tasks 1

3 Information Gathering 2

4 Prototype Representation 5

5 Coarse Coding 7

6 Labeling Functions and Likelihood Tables 8

7 Region Extraction 10

8 Example 11

9 Summary 16

1 Introduction

Image analysis usually requires recognition of *objects*—scene regions with characteristic colors, textures, shapes, or motions. I have developed a *coarse-coding* method for finding and labeling objects of approximately known appearance. The objects need not be homogenous, complete, unoccluded, or of any fixed size or orientation. Any spectral or textural signature characteristics that differ from the background can be exploited.

My KNIFE digital-image-analysis system supports pixel classification followed by region extraction, as well as segmentation followed by region labeling. Pixel classification with extraction and noise cleaning is the faster of the two, but KNIFE's integrated split/merge segmentation [Laws 88a] followed by region-based classification tends to give better results. The latter approach could exploit region shape and context as well as multivariate pixel signatures [Wesley 86], although only shape-based postprocessing has been implemented to date [Fua 86, 87ab]. This paper describes the pixel classification technique, which is the same whether it is applied before or after region extraction.

I shall first list tasks for which classification is useful, then discuss the required processing steps. Briefly stated, derived data bands gather all relevant information into pixel feature vectors. Multinomial signatures (or histograms) are then gathered for each class prototype. Reference signatures are compiled into labeling functions that map gray levels into source-class assignments. Confusion matrices for the reference signatures help combine single-band class estimates into multiband labels. Connected components in the second-stage label map are then extracted as regions (or objects) for higher-level processing.

2 Labeling Tasks

There are tasks in which approximate foreground and background signatures are available, and in which target identification can be achieved by simple pixel classification or discriminant analysis. Cuing and counting tasks require that all scene objects of a certain class be identified after one or more prototypical examples have been provided [Touchberry 77; Conners 82, 84; Trivedi 84ab; Harlow 85; Laws 85; Lehrer 87]. Similar recognition problems are posed by Tomographic reconstruction of solid objects from slices. Tracking is also an important application: although often limited to cuing and extraction of a single object in a temporal sequence of images, tracking of all discriminable objects in a scene may be necessary for robot vision or autonomous vehicle navigation [Ariki 81; Corkill 82; Laws 88b].

Identification of material types (soil, asphalt, concrete, water, etc.) is similar, except that a stored database of signatures is used [Haralick 69, 74; Wacker 69; Kettig 76; Nagao 76; Wiersma 76; Narendra 77; Peich 77, 80; Parikh 78; Sadjadi 79; Matsumoto 81; Swain 81]. Since simple classification is seldom adequate, researchers have developed multistage or relaxation analyses that exploit spatial and semantic relationships among regions [Duda 70, 80; Milner 70; Bajcsy 73, 76; Yakimovsky 74; Barrow 76, 77; Bullock 76; Faugeras 79, 81, 82; Haralick 79; Price 79, 81, 82; Ohta 80; Parma 80; Browse 82; Goldberg 82, 83; Hwang 83, 85; Kitchen 84; Matsuyama 85; Belknap 86; Wesley 86; Bhanu 87].

Material labeling remains difficult, especially for uncalibrated imagery and rapidly changing scenes. Cluster analysis and spatial reasoning can sometimes extract objects,

but classification techniques are still necessary for their identification. The same is true for regions found by other segmentation methods [Laws 88a]. I have developed classification tools in the KNIFE package to perform this identificatory step by using any data bands and object signatures that are available.

Many advances have been made since the early days of pixel classification and crop acreage estimation in ERTS/Landsat imagery. We can now take advantage of better sensors (far superior to human vision), faster computers, and improved techniques of restoration, filtering, interpolation, enhancement, correlation matching, multivariate classification, range estimation, and shape from shading.

Nevertheless, progress in automated material and object identification has been marginal for aerial reconnaisance and almost nonexistent for low-angle or ground-based imagery (as is needed by an autonomous vehicle). Edge detection and shape analysis are useful in constrained industrial-inspection tasks, but have had very limited success in natural imagery. Segmentation techniques, even those specially designed for texture segmentation, are only beginning to approach human performance. Integrated approaches that utilize edge-based and area-based techniques in a pyramid of gradually improving image resolutions are still highly experimental, and artificial intelligence methods remain inapplicable until semantic features can be extracted more reliably.

In this paper I show the benefits of returning to pixel classification as an initial method of image partitioning and material identification.

3 Information Gathering

I assume that the initial information relating to a pixel's material type can be gathered into a vector of scalars stored (implicitly or explicitly) as the pixel's multivariate value. Intensity, color, infrared brightness, and radar reflectance are often available in this form, while many other point properties may be directly measurable in the industrial or medical domains. Derived data, such as hue, saturation, local texture, surface slope, albedo, and even optic flow, can also be associated with individual pixels. Similar treatment of region shape and semantic context may be possible, as described below.

I further assume that any available "normalizing" bands—such as range and surface slope—have been used to correct the other multivariate values, thus making position-independent material classification a reasonable approach. (The normalizing bands may also be useful during segmentation; however, they do generally exhibit smooth variations that are difficult to exploit.) In some cases it may be desirable to perform a preliminary segmentation, compensate for regional shading and hypothesized object properties, and then reanalyze certain image regions by using the methods described in this paper.

The important point is that all information relating to a pixel's scene label must be available as part of the descriptive vector. Useful properties of each neighborhood, such as local texture, should be computed and assigned to the central pixel. Classification can then be done without considering joint probability distributions over neighboring pixels. I am thus advocating the traditional feature extraction/classification paradigm, except that I employ new techniques of classification and spatial analysis. I also want to allow for different descriptive vectors in different image regions after the image has been partially segmented, as well as different vectors for different analysis tasks.

An unlimited number of texture measures are available (local gray-level statistics

[Piech 70; Haralick 73; Nagao 76]; edge properties [Rosenfeld 71; Pietikäinen 82; Kjell 84]; spot density [Zucker 75; Mitchell 77, 78, 79]; co-occurrence of texture elements [Davis 79, 81; Dyer 80; Hong 80; Terzopoulos 80; Voorhees 87ab]; texture energy [Laws 80]; fractal dimension [Pentland 84]; and others), each computed over a range of neighborhood sizes and shapes. My current practice is to precompute only local variance,then to compute more complex texture measures [Laws 88b] over highly textured areas if they must be further classified or segmented. Partitioning based on texture measures alone is very seldom required in natural imagery because regions that differ in texture will typically also differ in brightness, color, or variance. Computing texture measures after segmentation, when feasible, greatly reduces problems caused by regional edge effects.

Actually, any quantity closely related to local variance will work for classification and segmentation. I use the logarithm of local variance because it assigns a reasonable amount of weight to subtle variations, is relatively unaffected by illumination changes, and fits well within an eight-bit pixel descriptor. I compute the variance in small windows, using binomial (or approximate Gaussian) relative weighting patterns such as

$$\begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix} \qquad \begin{bmatrix} 1 & 4 & 6 & 4 & 1 \\ 4 & 16 & 24 & 16 & 4 \\ 6 & 24 & 36 & 24 & 6 \\ 4 & 16 & 24 & 16 & 4 \\ 1 & 4 & 6 & 4 & 1 \end{bmatrix}$$

to emphasize the central pixels. (Rectangular windows with elliptical or diagonal weighting patterns could also be used [Laws 88b].) Note that the weights fall off rapidly, giving the effect of even smaller measurement windows and helping to avoid regional border effects. Small variance operators are essential for identifying such details as fine tree branches, although large operators (or small operators applied to reduced images) may be more convenient for extracting whole trees.

Early researchers overlooked the power of such local data for classification of natural textures. Gray-level co-occurence statistics [Haralick 71, 73, 74] were found to be more powerful than Fourier measures [Weszka 75, 76; Dyer 76], especially for nearest-neighbor statistics, but the comparative studies did not make it clear that the most powerful Fourier features were also high-frequency measures. These were computed across large image windows and their coefficents were averaged over large regions in the Fourier domain;[1] although this reduced their power considerably, they still outperformed the low-frequency features.

Local measures generate bimodal histograms when computed over large-scale macro-textures: one peak for texture-element interiors and another for their borders. Although this created problems for Gaussian-based discriminant functions, it can be an advantage for a multinomial classifier. The KNIFE algorithm makes use of histogram shape rather than just mean and standard deviation.

The human eye is very sensitive to collinear edge alignments, even over large distances—something no local texture measure can capture. It would be useful if we could add mea-

[1] An oft-cited study of Fourier phase measures [Eklundh 79] was likewise flawed. It should be repeated by using Gabor filters to measure local phase relationships among individual Fourier frequencies [Laws 88b]. Such measure would be appropriate for recognizing blurred or noisy textures with unreliable nearest-neighbor statistics.

sures of such gestalt properties to each pixel in an image. Characteristics of local *shape environments* could also be exploited for cuing, counting, and tracking tasks requiring that objects of a particular shape be found. We might note any nearby discontinuities and surface maxima, then use classification to seek pixels with similar local contexts. Parallel hardware, such as the Connection Machine, might be ideal for broadcasting local feature positions and compiling the shape-environment descriptors [Zucker 78; Davis 79, 81; Dyer 80; Hong 80; Hillis 86].

Other contextual knowledge may be provided by distant regions, previous scenes, dynamic analysis goals, hypothesized interpretations, etc. Capturing such knowledge in a finite vector of pixel descriptors would be difficult, but there is a shortcut that is adequate for our purpose. We may be able to capture the combined effect of all such knowledge on a specific classification problem. It is as if we were asking an expert, or expert system, "Given all you know about this pixel and its environment, what are the relative likelihoods that it came from each material type?" Answers to this implicit question (from any number of evaluation functions) can then be combined with other descriptors to estimate material-class likelihood. This is similar to the approach used in the PROSPECTOR expert system [Hart 77; Duda 79; Reboh 81] for predicting mineral deposits at each position on a map.

This insight is the basis of my classifier. Each descriptor value gives evidence for its pixel's material class or object identity. Appropriate evaluation routines may examine the patterns of evidence and post their own opinions, in the manner of blackboard expert systems. A top-level evaluator then examines all of the evidence and makes the final judgment. *Coarse-coding*[2] means that class membership is determined from the pattern of evidence rather than by majority vote or by selection of a single most-reliable estimator. It would be possible, for instance, for the top-level evaluator to reject all lower-level judgments, as when it assigns pixels with "grass" and "soil" characteristics to a "pasture" category.

The classification method described below is quite tolerant of "garbage" data bands, but there are computational advantages to using only the bands that carry information for a given task. A useful method of band selection is to attempt traditional or classificatory segmentation of each single data band, keeping only those that produce reasonable partitions. (Useless bands are either unsegmentable or result in randomly interspersed pixel labels. For some tasks we can compute the accuracy over known training regions—or the degree of correlation with a reliable classification method—as a screening measure.) During interactive analysis, a user typically examines the data bands or their segmentation/classification results interactively to decide which contain useful information. Standard procedures develop quickly for any specific task.

The bands remaining after such screening provide a vector of numbers (or other codes) characterizing each pixel and its environment. Some of the descriptors measure inherent object properties, while others may be derived from sophisticated processing of the surrounding image data. The data vectors in a region will all have the same structure, so that the elements form two-dimensional image bands. (Some of the data may be missing, as when specular reflection prevents measurement. Special placeholder codes should mark

[2]The term comes from the field of neural networks, or parallel distributed processing [Hinton 86]. Related multistage models, such as Samuel's signature tables [Samuel 67; Thosar 73], have been around for a long time.

this fact.) If sufficiently refined measures were available, the classification task would now be trivial. The rest of this paper discusses the more typical case in which sophisticated classification and grouping techniques must still be employed.

4 Prototype Representation

Once data bands have been computed, we can extract training exemplars for compilation into labeling functions. The user might, for instance, outline a few regions interactively (or point to regions of a segmented image) and supply labels for them. All representative appearances of a material type or semantic object should be included in the training set so that nearest-neighbor classification can be used. (Such approaches, also referred to as *memory-based reasoning* [Stanfill 86; Waltz 87], permit classification decisions to be "explained" to the user by displaying the appropriate prototypes. Incorrect assignments can be remedied by including unrecognized regions as new prototypes. The coarse-coded classification proposed in this paper is not quite a nearest-neighbor technique, but the philosophy is similar.)

A material type can be characterized by its *signature*, or probability distribution, over the possible multivariate pixel values. Any one material type may have several signatures (depending, for instance, on illumination or scene distance); for simplicity, I shall treat these as separate classes that happen to share a single semantic label. Different material types (e.g., different types of vegetation) may be mapped to different labels for some purposes and to a single label for others. Pixels will be classified under the distinct signature classes, then grouped into regions according to the associated semantic labels.

Signatures have commonly been represented by Gaussian distributions in order to reduce the probability estimates to a manageable number of parameters.[3] Such parametric distributions yield elegant discriminant functions, but seldom model image data realistically. Consider the trivial task of discriminating a two-valued salt-and-pepper distribution from a Gaussian with the same mean, standard deviation, and [zero] skewness. The two signatures differ only in their fourth and higher-order moments and cannot be separated by quadratic discriminant analysis, yet almost perfect pixel classification can be achieved with other techniques (including human vision).

The traditional approach can be salvaged if signatures can be decomposed into sums of multivariate Gaussian distributions. Pixels can then be assigned to the subpopulations and hence to an overall class [Sclove 80]. Other parametric *mixture densities* can be handled similarly. Unfortunately, this decomposition is quite difficult for naturally occurring material signatures—even in the one-dimensional case. Signatures of simple material types may be quite irregular (especially after multiband transformation [Kender 76, 77]), whereas histograms of mixed terrain and vegetation may so closely approximate a broad Gaussian as to defy meaningful decomposition. We may also have to deal with nominal, ordinal, or nonnumeric band codes for which parametric methods are inappropriate.

Traditional multivariate classification compares each pixel vector with each material signature and assigns the pixel to the most similar (or least distant) class. Similarity metrics have been based on the probability of an observed pixel value, given the material

[3] Even a three-dimensional signature of 256 gray levels per dimension would be awkward to represent as a multivariate histogram. We may therefore have to deal with scores of signatures having a dozen or more data dimensions.

class, or, via Bayes' theorem, on the probability of a material class, given the observed value. The latter is preferable, but requires estimates of a priori class probabilities. (Any classification technique assumes some model of these a priori probabilities; it is the explicit treatment of them that makes Bayesian classification preferable when it can be used. Human judgment, however, often deviates from the Bayesian model even when the necessary information is available.)

I propose the following approach. Suppose we accept observed histograms as the best available estimates of material and object signatures. Histogram prototypes are equivalent to modeling each texture class as a multinomial process with an [almost] independent probability of producing each possible multivariate value. We can use ratios of matching pixel probabilities for two source classes to estimate whether a particular descriptor vector is more likely from one signature class than from another [Laws 85].

Smoothing the histograms of continuous (i.e., interval or ratio) numeric measures introduces a desirable correlation between nearby bin values.[4] A variety of *smoothing kernals* has been used for estimating probability distributions from histograms. I use Gaussian smoothing (and folding back of off-scale energy) with good results, but almost any moderate smoothing process would be acceptable. The optimal amount of smoothing depends on the expected variability of observed gray levels.

Multidimensional histograms are awkward to use, especially when different subsets of the data bands are to be used for different tasks. Sparse storage techniques do exist [O'Rourke 84], but histogram smoothing degrades their effectiveness. I have chosen to store only the single-band (or *marginal*) histograms, which are easy to compute, manipulate, display, and interpret. If the univariate histograms are inadequate, multivariate transformations can be employed to compute additional pixel descriptors that summarize the multiband information with respect to a particular goal. When, for instance, color images are segmented, an intensity-hue-saturation representation is sufficiently decoupled that original red-green-blue measures may usually be discarded [Laws 88a]. This is similar to using redundant multivariate transformations to search for axes along which multiband distributions are separable [Ohlander 78].

Note that I am not proposing univariate histogram representations as an approximation to multidimensional parametric signatures.[5] Parametric representations simply do not capture the full complexity of multimodal real-world data. Multivariate histograms do capture this complexity—but only too well, which is why smoothing is required. Carefully chosen univariate histograms not only capture most of the signature information, but also simplify the problem of combining data bands to compute similarity or distance functions. I propose that the weighting or feature extraction problem be confronted at this point because the resulting data bands and univariate histograms are in a form suitable for human understanding. This permits an "expert system" approach to system development, as well as providing very fast techniques for image segmentation and object identification.

[4] Smoothing does lose information when gray levels from one class are interdigitated with gray levels from another. Such interleaved *picket fence* effects sometimes occur in derived data bands when two distinct populations are mapped to a single interval.

[5] I do take such a stand elsewhere [Laws 88a], viewing segmentation that uses univariate histograms as a heuristic shortcut to full multidimensional cluster analysis.

5 Coarse Coding

Even a classification procedure that employs just the marginal histograms presents some data-handling difficulties. Assume that we have as many as sixteen data bands and 128 sixteen-band signatures (representing 128 source classes or possible appearances of materials). It would be inefficient to store 128 class likelihoods for each pixel, updating each sixteen times as the data bands are processed. Performing multiband classification pixel by pixel would seem more reasonable, but requires moving all 128 sixteen-band signatures through the computer for each pixel processed. Special parallel hardware would be needed to make such approaches practical.

Fortunately there is an efficient way to classify the pixels in a single pass through the data bands and signature histograms. Some accuracy may be sacrificed, but the processing effort and intermediate storage are greatly reduced. The key is a coarse-coded representation that encodes approximate likelihoods for all signature classes in a single integer or bit pattern. The final bit pattern for each pixel can then be decoded to provide fairly reliable likelihood estimates for each signature class. The more data bands used, the less effect misclassification on any band will have. Adding more data bands can only improve the classification results as long as the stored signatures for those bands are truly representative.

The essence of coarse coding is to let each bit or group of bits in a pattern encode independent information about the pixel's material class. The overall sequence of bits is then a more reliable indicator of material class than are the individual bit groups [Hinton 86]. This can be regarded as a two-step classification procedure: first each pixel vector is summarized by a coarse-coded bit pattern, after which the bit pattern is expanded to a vector of signature likelihoods or other outputs. (It can also be regarded as the feature extraction and classification steps of traditional pattern recognition, performed after the preliminary feature extraction that generated the data bands.)

The quality of the final classification obviously depends on the method of encoding and decoding these bit patterns. Bit groups are similar to the *hidden units* of connectionist pattern recognition [Rumelhart 86b]. That approach would use a gradient-descent algorithm to evolve a set of codes with satisfactory classificatory power on some training set; with luck, the codes would also generalize to additional classification problems. I have developed a more structured approach that uses image statistics and signature characteristics to select the codes dynamically for each task. Each relevant data band is reduced to a few bits in the code; the bit string is then decoded in the manner that best preserves the discriminability of the reference signatures.

I choose a coding scheme in which each group of descriptor bits represents a pixel's most likely signature class, as estimated from a single data band. The single-band codes are concatenated to form a full coarse-coded pixel descriptor. (This is equivalent to storing each bit group in a separate data band. Many operating systems, though, limit the number of images or data files that can be open at one time.) The required number of bits per pixel depends on the number of data bands used and the degree of accuracy that each requires. Seven bits per band, for instance, are sufficient to designate one of 128 signature classes uniquely.

Alternatively, the available bits per pixel can be partitioned optimally among the data bands. Source classes can be clustered beforehand into groups with similar signatures for

a particular band; only enough bits to represent the equivalence sets are then needed. Such partitioning can even increase overall classification accuracy [Hinton 86]. The set partitioning should be different for each data band, with more bits used for the more informative bands. Signature clusters should maximize similarity within a cluster and discriminability among clusters while maximizing classification accuracy across all bands. It may also be advantageous to group semantically related material classes, such as all of the vegetation signatures, for tasks in which such confusion is relatively unimportant.

Additional bits for each band could be allocated to record the second (or even third) best class, as well as an estimate of classification reliability. As this would make the decoding more difficult, it is worthwhile only if one of the data bands contains significant information that cannot be captured by the pattern across all bands. On the other hand, there little penalty is incurred for applying independent classification algorithms to one data band (and to the previously computed bit patterns), recording their opinions as if additional data bands had been employed.

The remainder of this paper treats only the signature classification problem and not the allocation of coarse-coding bits or the optimization of cluster assignments.

6 Labeling Functions and Likelihood Tables

Given a data band, we need to transform the pixel values to bit codes that can be appended to the coarse-coded descriptor band. This is just classification of observed gray levels into an a posteriori most likely signature class.

I start with a set of single-band histograms representing important objects and expected background signatures. Signatures may come from a database (suitably corrected for scene and sensor characteristics) or from labeled image regions. If the backgound statistics are unknown, they can be estimated from a full-image histogram (if the target objects are small) or from ensemble statistics of typical backgrounds.

The first step is to smooth any continuous-valued reference signatures. This spreads each bin probability over several bins in a manner that models the uncertainties of gray-level reproducibility. Considerable smoothing is needed for object recognition, much less for object tracking under uniform illumination conditions. A certain minimum amount of smoothing is needed to account for random sampling effects in the original signature histograms [Laws 85].

The next step is to estimate source class likelihoods for each possible gray level. The class code (i.e., bit pattern) for the most likely signature class can then be stored in a single-band lookup table for rapid pixel labeling. I call these lookup tables *labeling functions* to avoid confusion with other lookup tables described in this paper. The mechanics of actual classification depend on the available hardware, but lookup table transformations are typically quite efficient.

The selected source class for a given gray level could be just the signature having the highest probability for that bin, but we can use a priori source probabilities and Bayes' rule (as well as utility functions or error penalties) to make a better selection. The prior class probabilities can be estimated from historical frequencies or from an analysis of the data band histograms (as described below).

Once we have the labeling functions, we can compute expected confusion matrices (or, stated differently, a priori signature discriminability). Because this can be done before

applying the labeling functions to the input data bands, it can be used for task-dependent band selection. The confusion matrices also provide probabilities and likelihoods needed to decode the coarse-coded pixel descriptors.

The trick is to pass each single-band reference signature through the band labeling function. Some of the gray levels recorded for that class will be correctly labeled, while others will be attributed to the incorrect signature classes. The relative frequencies of the different labels, normalized to unit sum, indicate the probability of each class assignment, given the source class; this forms one row of a single-band confusion matrix. The process is repeated for each signature to fill out a confusion matrix for each band.

Given an assigned label, we can now use Bayes' theorem to derive the single-band posterior probability of each source class. If the source classes are equally likely, the relative likelihoods can be read from the columns of the appropriate confusion matrix. If not, weighting factors proportional to source probabilities adjust the column entries to yield the likelihoods; normalization to unit sum converts these to posterior probabilities.

One estimate of the source probabilities[6] can be obtained by passing the image data band histograms through the labeling functions, possibly combining the resulting label frequencies across bands. (These label frequencies can also be used for band selection; rates of target detection, for instance, would seldom be improved by using data bands in which the target label is never assigned.) Note that passing a reference signature through the labeling functions is generally much faster than passing image data through and then histogramming the result.

Combining all of the above, we can now get label probabilities for each band. These can be combined to get an estimate of the multiband class probabilities for any given pattern of coarse-coded descriptor bits. The probability of any bit pattern for a given source class is the product of the probabilities of the individual single-band labels. (I assume, as discussed above, that the selected data bands are sufficiently independent that we can ignore pairwise and higher-order band interactions. If this is not true, fuzzy combining functions might be more appropriate [Zadeh 74; Salton 83; Laws 85].) We can do this multiplication for each of the source classes, perform the Bayes' inversion to get the posterior class probabilities, and assign the most likely class label to the bit pattern.

The labels for all possible bit patterns can be precomputed and stored in a *classification lookup table* if the number of bits is small. Similar lookup tables can store the second-best label, the ratio of best to second-best posterior probabilities, the entropy (i.e., information-theoretic uncertainty), or any other function of the posterior probabilities.

Longer descriptors, or those physically stored in more than one intermedate band, may be more efficiently decoded by cached lookup. (I use software lookup tables for descriptors of up to 12 bits, dynamic decoding with cached lookup for longer patterns.) Patterns found in the image are decoded by using the same formula as before to select the most likely source class. The pattern and its label are then attached to a list (or a pair of corresponding lists). Each computed pixel descriptor is sought in the cached list and is expanded to its vector of class likelihoods only if the pattern has not been seen previously. There are typically only a few hundred distinct patterns in an image. Linear

[6]These estimates are intermediate between a priori and a posteriori estimates, thus arguably closer to the methods of human perception. The prior probability of having a zebra in my office is infinitesimal, but having seen a black-and-white striped animate object there, I should use an increased "zebra probability" in trying to identify it.

search that commences with the most recently seen pattern (or the most recently created) is satisfactory, although a hashed storage scheme would be faster.

The classification procedure has been described. Each whole-image or regional data band is passed through its corresponding labeling function to produce a source class estimate for each pixel. These estimates, or code bits, are appended to a band of coarse-coded pixel descriptors. The final coarse-coded descriptors are then passed through the classification lookup table or are dynamically decoded to obtain a consensus label for each pixel. These labels constitute a *label map*. The next step is to extract the connected components and instantiate them as regions in a knowledge base.

7 Region Extraction

The new region-extraction algorithm in the KNIFE package is greatly improved over one reported earlier [Laws 82]. An initial scan through the label map locates each connected component; a second pass then renumbers the pixels to form a *region map*. Region descriptors computed during this process can be used for identifying small *noise regions* that should be merged with neighbors [Laws 88a].

I use a statistical noise-cleaning technique. Each small connected component is considered for merger with the neighbor having the most similar multivariate regional histogram. The test for histogram similarity allows for the possibility that the small region's histogram matches only a portion (e.g., one tail) of its larger neighbor's histogram. (Most statistical goodness-of-fit tests assume random sampling and so require a full match.) A pseudo-F test for collinear surface fit then determines whether the merge is acceptable.

A rather complex problem arises when hierarchical signature classes are available. Suppose, for instance, that several kinds of grass are known to the analysis system. All of the signatures are likely to be similar, even though sufficiently distinct to form separate signature sets. In labeling a grassy field, the classifier is now likely to assign different grass labels to neighboring pixels. Where large clumps of one type occur we would like the classifier to report them, but where labels are intermixed we would like the classifer to group them all under a generic "grass" label. Similarly, interspersed grass and soil should be labeled "field."

We cannot search for the composites, then subdivide them into more specific signature classes: the ensemble signature for a mixture of unknown proportions is often unknowable or too broad to be useful. A better solution is to extract homogeneous regions from a fully labeled image, then replace intermixed labels with appropriate generic ones and extract any new homogeneous regions. This may have to be repeated with several different label generalizations, but the area to be reprocessed shrinks each time an identifiable region is found. I have worked out a way to do this during connected-component extraction, but have not yet implemented it in the KNIFE package.

The initial image or image region is thus partitioned into labeled regions. Selected regions can be further partitioned, if necessary, either by pixel labeling and grouping or by segmenting and then labeling. It is often effective to alternate the two techniques, since classification can break up complex imagery that stymies histogram-based segmentation, while spatial segmentation can identify subregions that match separate modes of a multimodal signature.

8 Example

Ground truth is difficult to obtain for natural imagery. Rather than present tables of classification accuracies, I am going to offer a visual presentation of classification on the training set. Success at such a task is not sufficient to prove the usefulness of my image analysis approach, but it is essential. A tracking or labeling system that cannot identify its training regions is hardly worth building.

I start with the color image in Figure 1(a) through (c), transformed to my VHS (vividness-hue-saturation) representation [Laws 88b]. From the vividness (or intensity) band I compute the 3×3 log-variance band in Figure 1(d). I would prefer a texture measure that responds less strongly to object edges, but have not yet developed a suitable normalization.

I then trace the four major scene objects—road, sky, trees, and ground. I haven't printed this training image, but it includes all image pixels except for rather thick strips along boundaries of the four regions. Tracing with a mouse takes only a few seconds for most large regions, although I admit to taking a couple of minutes to extract the grass and mountains as a single gerrymandered "ground" region. (KNIFE could handle multiple exemplars per semantic class, but its display and editing tools for composite signatures are rather primitive.) Training signatures can come from previous images or from a database, although crude tracing of large regions is a good strategy for acquiring new materials.

Signatures, or histograms, for the four material classes are shown in Figure 2. No one band is adequate for discriminating all four textures, but the patterns of confusion differ from one band to another. This is critical if coarse coding is to be effective, since I do not exploit interband correlations.

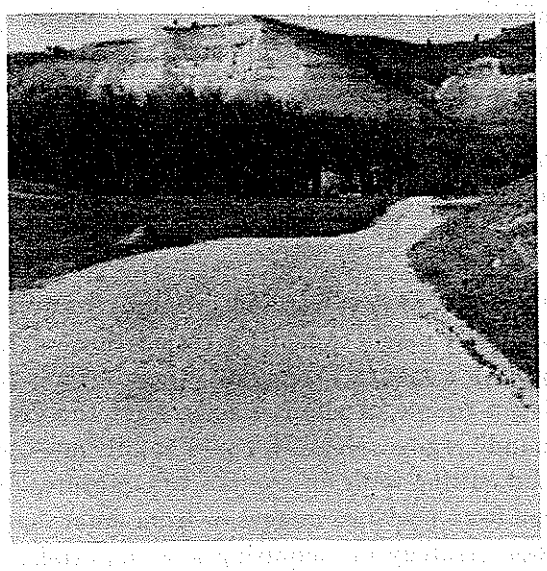
Figure 3 shows the single-band pixels labels computed with my method. Trees are marked with the darkest gray levels, then ground, road, and sky. The vividness band mislabels much of the sky, the mountain face, and the road. Hue mislabels much of the ground area. Saturation labels almost everything tree or road, while the *LV3* measure produces noisy patches of tree and sky labels. None of these classifiers can reconstruct the training set, but at least they make different patterns of errors.

Figure 4(b) shows how this trait can be exploited. A second-level classifier is applied to the four labels at each pixel. This operator, constructed to optimize labeling of the training signatures, "second guesses" the first-level classifiers and assigns a final pixel label. The result is still somewhat noisy, but most pixels have been classified correctly.

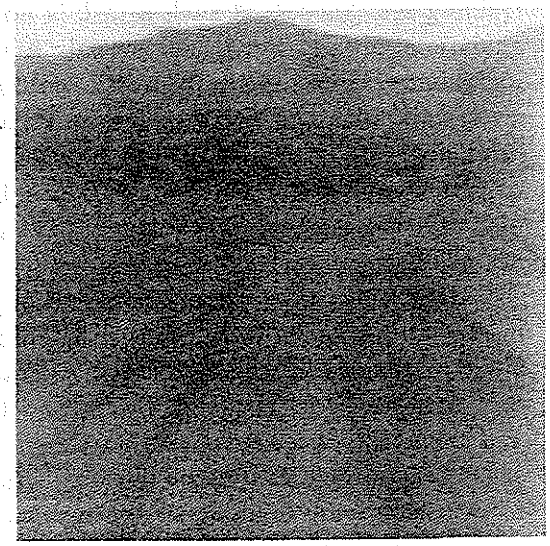
Since KNIFE is also a segmentation program, I can use its connected-component extraction routine to consolidate labeled pixels into regions and build corresponding data structures. Figure 4(c) shows the extracted regions when KNIFE's seglevel parameter is set to 1 (the coarsest setting for normal use). Figure 4(d) goes one step further, merging any region smaller than 200 pixels into its most similar neighbor.[7]

The final result is a clean segmentation in about a tenth of the time that KNIFE's integrated split/merge partitioning algorithm would require. Classification-based segmentation of a 256×256 region may take from one to ten minutes on a VAX 11/780, depending on the number of bands and the number of regions formed. The illustrated spectral/spatial labeling process has not only recovered its training set, but has done a good job of labeling

[7] KNIFE's region-growing operator would have much the same effect if applied to each of the major regions.

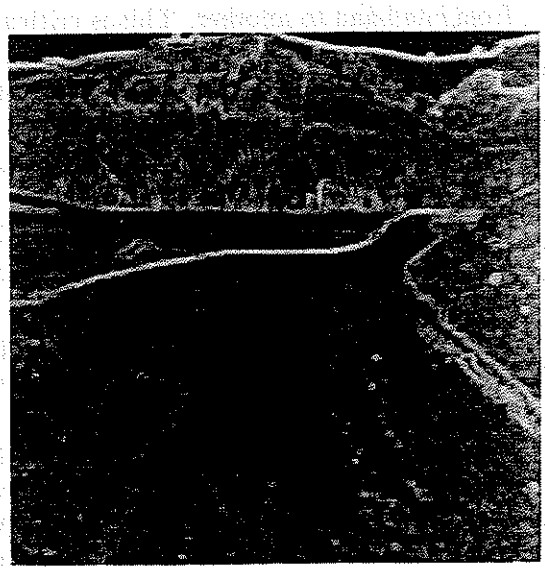


(a) Vividness (V) Band



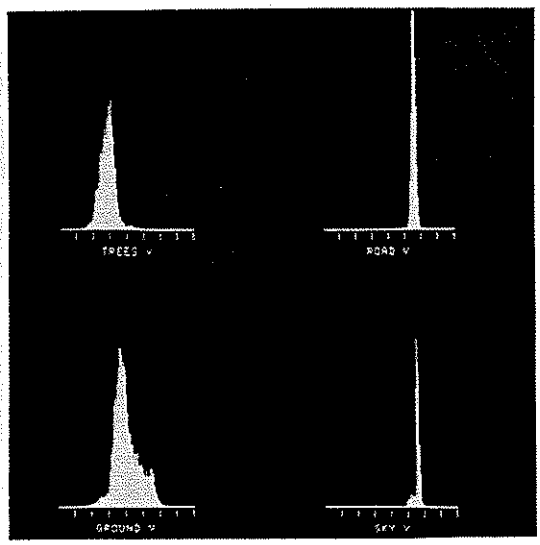
(b) Hue (H) Band

(c) Saturation (S) Band

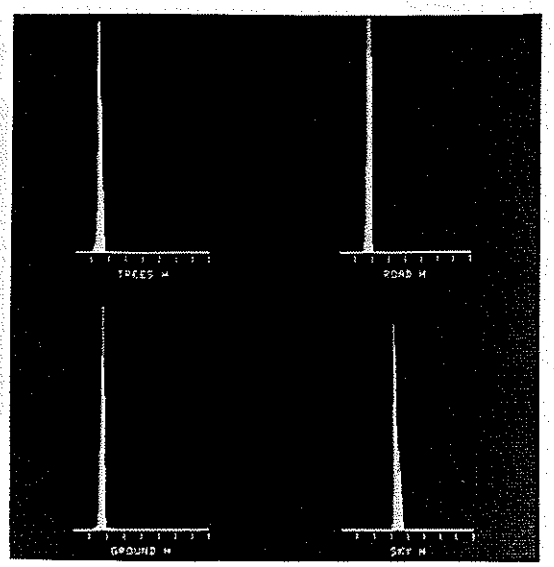


(d) Log-Variance ($LV3$) Band

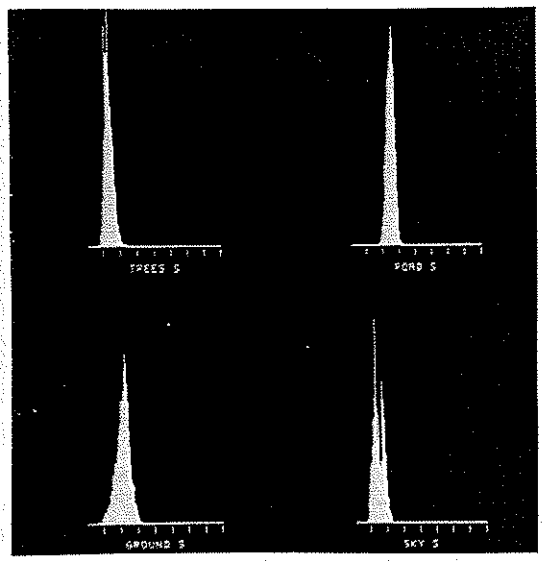
Figure 1: Multiband ALV519 Image



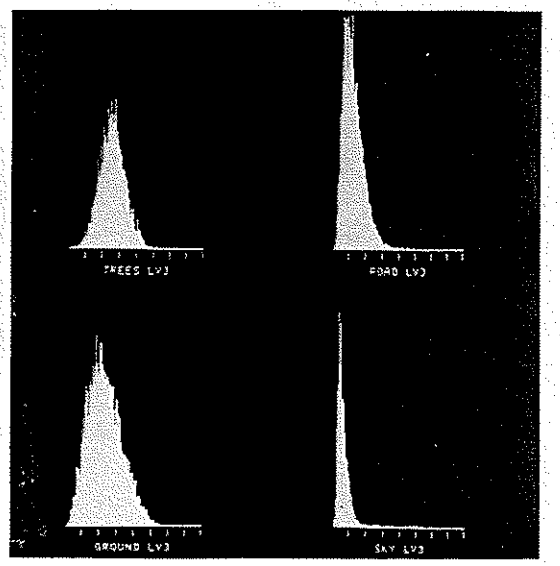
(a) Vividness Histograms



(b) Hue Histograms



(c) Saturation Histograms



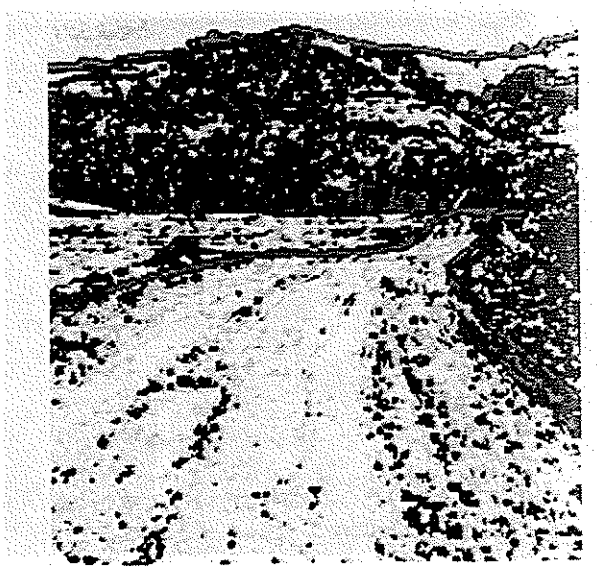
(d) Log-Variance Histograms

Figure 2: Prototype Signatures

(a) Vividness Classification

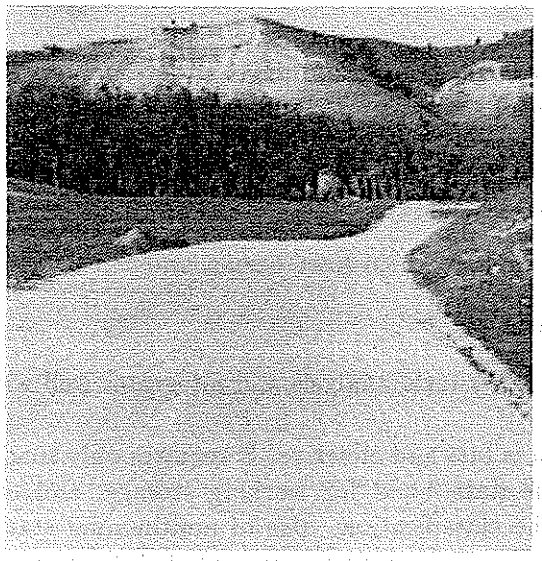
(b) Hue Classification

(c) Saturation Classification

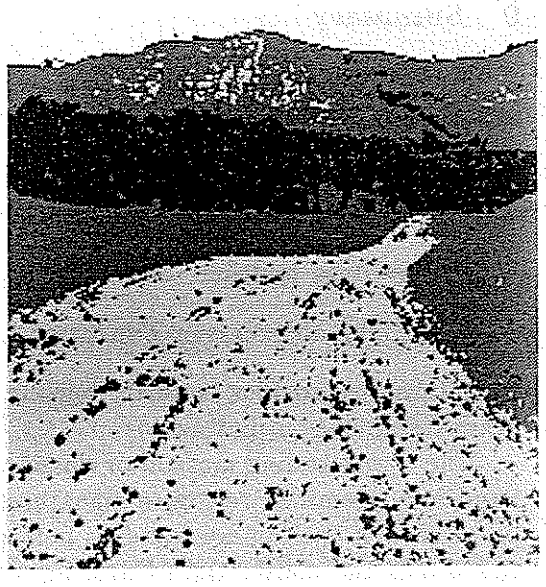


(d) Log-Variance Classification

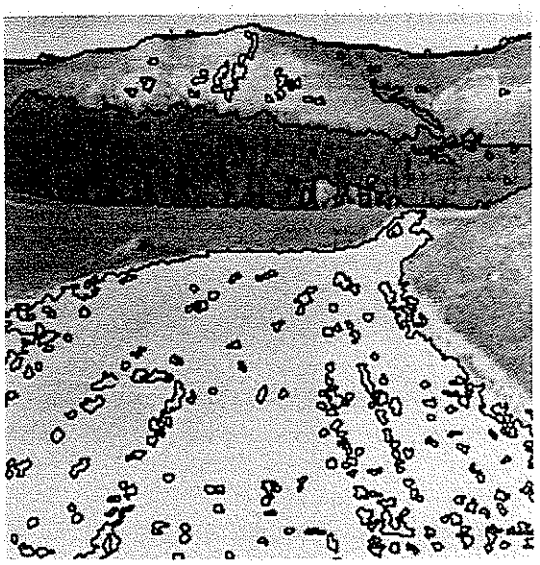
Figure 3: Single-Band Classification Maps



(a) Vividness (V) Band



(b) Multiband Classification



(c) Extracted Regions

(d) Merged Regions

Figure 4: Classification Results

the regional boundary pixels as well.

9 Summary

Many problems in image analysis can be solved by labeling pixels and grouping them into regions, or by partitioning pixels into regions and assigning labels based on regional pixel properties. Three such problems are

- Identifying materials with known brightness, color, or texture distributions
- Identifying multiple scene objects once some of them have been found
- Tracking objects from one image to another.

Pixel labeling provides tentative regions for higher-level analysis and integrates well with other segmentation methods.

The coarse-coding method of classification is fast and effective. It requires only single-band histograms as reference signatures, one data band for working storage, and one pass through each image band to perform the classification. Use of multinomial statistics avoids the multivariate Gaussian assumption built into traditional classification approaches. Needed probabilities can be estimated from the reference signatures and data bands, while missing data bands (in either the signatures or in areas of the image) can be handled with minimal difficulty. The approach is fairly intuitive and shares many of the benefits of blackboard-style expert-system development.

Acknowledgment

This work was supported by Defense Advanced Research Projects Agency under Contract Nos. MDA903-86-C-0084 and DACA76-85-C-0004.

References

[Ariki 81] Y. Ariki, *Top-Down Approach and Bottom-Up Approach for Moving Picture Analysis*, Ph.D. Thesis, Kyoto University, Japan (December 1981).

[Bajcsy 73] R. Bajcsy and M. Tavakoli, "A Computer Recognition of Bridges, Island, Rivers and Lakes from Satellite Pictures," *Proc. Symp. on Machine Processing of Remotely Sensed Data*, pp. 2A-54–2A-68 (October 1973).

[Bajcsy 76] R. Bajcsy and M. Tavakoli, "Computer Recognition of Roads from Satellite Images," *IEEE Trans. on Systems, Man and Cybernetics*, Vol. SMC-6, No. 9, pp. 623–637 (September 1976).

[Barrow 76] H.G. Barrow and J.M. Tenenbaum, *MSYS: A System for Reasoning about Scenes*, Technical Note 121, Artificial Intelligence Center, Stanford Research Institute, Menlo Park, California (April 1976).

[Barrow 77] H.G. Barrow, T.D. Garvey, J. Kremers, J.M. Tenenbaum, and H.C. Wolf, *Interactive Aids for Cartography and Photo Interpretation*, Technical Note 137, Artificial Intelligence Center, Stanford Research Institute, Menlo Park, California (January 1977).

[Belknap 86] R. Belknap, E. Riseman, and A. Hanson, "The Information Fusion Problem and Rule-Based Hypotheses Applied to Complex Aggregations of Image Events," *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, Miami Beach, Florida, pp. 227–234 (22–26 June 1986). Also published in *Proc. DARPA Image Understanding Workshop*, Miami Beach, Florida, pp. 279–292 (9–10 December 1985).

[Bhanu 87] B. Bhanu and P. Symosek, "Interpretation of Terrain Using Hierarchical Symbolic Grouping for Multi-Spectral Images, *Proc. DARPA Image Understanding Workshop*, Los Angeles, California, pp. 466–474 (23–25 February 1987).

[Browse 82] R.A. Browse, *Knowledge-Based Visual Interpretation using Declarative Schemata*, Technical Report TN82-12, Department of Computer Science, University of British Columbia, Vancouver, B.C., Canada (November 1982).

[Bryant 79] J. Bryant, "On the Clustering of Multidimensional Pictorial Data," *Pattern Recognition*, Vol. 11, pp. 115–125 (1979).

[Bullock 76] B.L. Bullock, "Finding Structure in Outdoor Scenes," in C.H. Chen (ed.), *Pattern Recognition and Artificial Intelligence*, Academic Press, New York (1976).

[Conners 82] R.W. Conners, M.M. Trivedi, and C.A. Harlow, *The Segmentation of a Complex Urban Scene Using Texture Operators*, Remote Sensing and Image Processing Technical Report 401.82, Dept. of Electrical and Computer Engineering, Louisiana State University (1982).

[Conners 84] R.W. Conners, M.M. Trivedi, and C.A. Harlow, "Segmentation of a High-Resolution Urban Scene Using Texture Operators," *Computer Vision, Graphics, and Image Processing*, Vol. 25, pp. 273–310 (March 1984).

[Corkill 82] D.D. Corkill, V.R. Lesser, and E. Hudlicka, "Unifying Data-Directed and Goal-Directed Control: An Example and Experiments," *Proc. National Conf. on Artificial Intelligence*, Pittsburgh, Pennsylvania, pp. 143–147 (18–20 August 1982).

[Davis 79] L.S. Davis, S. Johns, and J.K. Aggarwal, "Texture Analysis Using Generalized Co-Occurrence Matrices," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. PAMI-1, No. 3, pp. 251–259 (July 1979). Also published as L.S. Davis and J.K. Aggarwal, *Pattern Recognition and AI*, Princeton, p. 185, (April 1978), and in *Pattern Recognition and Image Processing*, Chicago, pp. 313–318. (May 1978).

[Davis 81] L.S. Davis, M. Clearman, and J.K. Aggarwal, "An Empirical Evaluation of Generalized Cooccurrence Matrices," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. PAMI-3, No. 2, pp. 214–221 (1981).

[Duda 70] R.O. Duda, *Some Current Techniques for Scene Analysis*, Technical Note 46, Artificial Intelligence Center, Stanford Research Institute, Menlo Park, California (October 1970).

[Duda 79] R.O. Duda, J. Gaschnig, and P.E. Hart, "Model Design in the Prospector System for Mineral Exploration," D. Michie (ed.), *Expert Systems in the Micro-Electronic Age*, Edinburgh University Press, University of Edinburgh, Scotland, pp. 153–167 (1979).

[Duda 80] R.O. Duda and T.D. Garvey, *A Study of Knowledge-Based Systems for Photo Interpretation*, Report ETL-0235, Artificial Intelligence Center, SRI International, Menlo Park, California (June 1980).

[Dyer 76] C.R. Dyer and A. Rosenfeld, "Fourier Texture Features: Suppression of Aperture Effects," *IEEE Trans. on Systems, Man and Cybernetics*, Vol. SMC-6, No. 10, pp. 703–705 (October 1976). Also published as TR-391, Computer Vision Laboratory, Computer Science Center, University of Maryland, College Park, Maryland (July 1975).

[Dyer 80] C.R. Dyer, T. Hong, and A. Rosenfeld, "Texture Classification Using Gray Level Cooccurrence Based on Edge Maxima," *IEEE Trans. on Systems, Man and Cybernetics*, Vol. SMC-10, No. 3, pp. 158–163 (1980). Also published as TR-731, Computer Science Center, University of Maryland, College Park, Maryland (March 1979).

[Eklundh 79] J.O. Eklundh, "On the Use of Fourier Phase Features for Texture Discrimination," *Computer Graphics and Image Processing*, Vol. 9, pp. 199–201 (1979).

[Faugeras 79] O.D. Faugeras and M. Berthod, "Scene Labeling: An Optimization Approach," *Proc. IEEE Conf. on Pattern Recognition and Image Processing*, Chicago, Illinois, pp. 318–326 (August 1979).

[Faugeras 81] O.D. Faugeras and K. Price, "Semantic Description of Aerial Images Using Stochastic Labeling," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. PAMI-3, No. 6, pp. 633–642 (November 1981). Also in *Proc. DARPA Image Understanding Workshop*, College Park, Maryland, pp. 89–94 (30 April 1980), and in *Proc. 5th Int. Jnt. Conf. on Pattern Recognition*, Miami Beach, Florida, pp. 352–357 (1–4 December 1980).

[Faugeras 82] O.D. Faugeras, "Relaxation Labeling and Evidence Gathering," *Proc. IEEE Conf. on Pattern Recognition and Image Processing*, Las Vegas, Nevada, pp. 672–677 (14–17 June 1982).

[Flinchbaugh 81] B.E. Flinchbaugh and B. Chandrasekaran, "A Theory of Spatio-Temporal Aggregation for Vision," *Artificial Intelligence*, Vol. 17, No. 1–3, pp. 387–407 (1981).

[Freuder 77] E.C. Freuder, "A Computer System for Visual Recognition Using Active Knowledge," *Proc. 5th Int. Jnt. Conf. on Artificial Intelligence*, Cambridge, Massachusetts, pp. 671–677 (1977). Reprinted in Y.-H. Pao and G.W. Ernst (eds.), *Context-Directed Pattern Recognition and Machine Intelligence Techniques for Information Processing*, IEEE Computer Society Press, Silver Spring, Maryland, pp. 511–517. (1982).

[Fua 86] P. Fua and A.J. Hanson, *Using Generic Geometric Knowledge to Delineate Cultural Objects in Aerial Imagery*, Technical Note 378, Artificial Intelligence Center, SRI International, Menlo Park, California (March 1986).

[Fua 87a] P. Fua and A.J. Hanson, "Resegmentation Using Generic Shape: Locating General Cultural Objects," *Pattern Recognition Letters*, Vol. 5, pp. 243–252 (1987).

[Fua 87b] P. Fua and A.J. Hanson, "Using Generic Geometric Models for Intelligent Shape Extraction," *Proc. DARPA Image Understanding Workshop*, Los Angeles, California, pp. 227–233 (23–25 February 1987). Also in *Proc. 6th National Conf. on Artificial Intelligence*, Seattle, Washington, pp. 706–711 (13–17 July 1987).

[Goldberg 82] M. Goldberg, D. Schlaps, M. Alvo, and G. Karam, "Monitoring and Change Detection with Landsat Imagery," *Proc. 6th Int. Jnt. Conf. on Pattern Recognition,* Munich, Germany, pp. 523–526 (19–22 October 1982).

[Goldberg 83] M. Goldberg, G. Karam, and M. Alvo, "A Production Rule-Based Expert System for Interpreting Multi-Temporal LANDSAT Imagery," *Proc. IEEE Conf. on Computer Vision and Pattern Recognition,* Washington, D.C., pp. 77–82 (19–23 June 1983).

[Haralick 69] R.M. Haralick and G.L. Kelly, "Pattern Recognition with Measurement Space and Spatial Clustering for Multiple Images," *Proc. IEEE,* Vol. 57, No. 4, pp. 654–665 (April 1969).

[Haralick 71] R.M. Haralick, "A Texture-Context Feature Extraction Algorithm for Remotely Sensed Imagery," *Proc. IEEE Decision and Control Conf.,* Gainsville, Florida, pp. 650–657 (December 1971).

[Haralick 73] R.M. Haralick, K. Shanmugam, and I. Dinstein, "Textural Features for Image Classification," *IEEE Trans. on Systems, Man and Cybernetics,* Vol. SMC–3, No. 6, pp. 610–621 (November 1973).

[Haralick 74] R.M. Haralick and K.S. Shanmugam, "Combined Spectral and Spatial Processing of ERTS Imagery Data," *J. of Remote Sensing of the Environment,* Vol. 3, No. 1, pp. 3–13 (1974).

[Haralick 79] R.M. Haralick and L.G. Shapiro, "The Consistent Labeling Problem: Part 1," *IEEE Trans. on Pattern Analysis and Machine Intelligence,* Vol. PAMI–1, No. 2, pp. 173–184 (April 1979).

[Harlow 85] C.A. Harlow, M. Trivedi, and R.W. Conners, "Use of Texture Operators in Segmentation," *Proc. SPIE Conf. on Applications of Artificial Intelligence,* Vol. 548, pp. 10–15+ (1985).

[Hart 77] P.E. Hart and R.O. Duda, *PROSPECTOR—A Computer Based Consultation System for Mineral Exploration,* Technical Note 155, Artificial Intelligence Center, SRI International, Menlo Park, California (October 1977).

[Hillis 86] W.D. Hillis and G.L. Steele, Jr., "Data Parallel Algorithms," *Communications of the ACM,* Vol. 29, No. 12, pp. 1170–1183 (December 1986).

[Hinton 86] G.E. Hinton, J.L. McClelland, and D.E. Rumelhart, "Distributed Representations," in [Rumelhart 86a], pp. 77–109.

[Hong 80] T.H. Hong, C.R. Dyer, and A. Rosenfeld, "Texture Primitive Extraction Using an Edge-Based Approach," *IEEE Trans. on Systems, Man and Cybernetics,* Vol. SMC–10, No. 10, pp. 659–675 (1980).

[Hwang 83] V. Hwang, T. Matsuyama, L.S. Davis, and A. Rosenfeld, *Evidence Accumulation for Spatial Reasoning in Aerial Image Understanding,* Technical Report CAR–TR–28, Center for Automation Research, University of Maryland, College Park, Maryland (1983).

[Hwang 85] V. Hwang, L.S. Davis, and T. Matsuyama, *Hypothesis Integration in Image Understanding,* Technical Report CAR–TR–130, Center for Automation Research, University of Maryland, College Park, Maryland (1985).

[Kahneman 82a] D. Kahneman and A. Tversky, "On the Study of Statistical Intuitions," *Cognition,* Vol. 11, No. 2, pp. 123–141 (March 1982).

[Kahneman 82b] D. Kahneman and A. Tversky, "Variants of Uncertainty," *ibid.,* pp. 143–157.

[Kender 76] J.R. Kender, *Saturation, Hue, and Normalized Color: Calculation, Digitization Effects, and Use,* Dept. of Computer Science, Carnegie-Mellon University, Pittsburgh, Pennsylvania (November 1976).

[Kender 77] J.R. Kender, "Instabilities in Color Transformations," *Proc. IEEE Conf. on Pattern Recognition and Image Processing*, Troy, New York, pp. 266–274 (6–8 June 1977).

[Kettig 76] R.L. Kettig and D.A. Landgrebe, "Classification of Multispectral Image Data by Extraction and Classification of Homogeneous Objects," *IEEE Trans. on Geoscience Electronics*, Vol. GE-14, No. 1, pp. 19–25 (January 1976). Also published as LARS Information Note 050975, The Laboratory for Application of Remote Sensing, Purdue University, West Lafayette, Indiana (1975).

[Kitchen 84] L.J. Kitchen and A. Rosenfeld, "Scene Analysis Using Region-Based Constraint Filtering," *Pattern Recognition*, Vol. 17, No. 2, pp. 189–203 (1984).

[Kjell 84] B. Kjell and C.R. Dyer, *Edge Separation and Orientation Texture Measures*, Technical Report 559, Computer Science Department, University of Wisconsin, Madison, Wisconsin (October 1984).

[Laws 80] K.I. Laws, "Rapid Texture Identification," *Proc. SPIE Conf. on Image Processing for Missile Guidance*, San Diego, California, Vol. 238, pp. 376–380 (29 July – 1 August 1980).

[Laws 82] K.I. Laws, *The PHOENIX Image Segmentation System: Description and Evaluation*, Technical Note 289, Artificial Intelligence Center, SRI International, Menlo Park, California (December 1982).

[Laws 85] K.I. Laws, "Goal-Directed Textured-Image Segmentation," *Proc. SPIE Conf. on Applications of Artificial Intelligence II*, Vol. 548, Arlington, Virginia, (9–11 April 1985). Summarizes Technical Note 334, Artificial Intelligence Center, SRI International, Menlo Park, California (September 1984).

[Laws 88a] K.I. Laws, *Integrated Split/Merge Image Segmentation*, Technical Note 441, Artificial Intelligence Center, SRI International, Menlo Park, California (July 1988).

[Laws 88b] K.I. Laws, *Classification-Based Tracking of Objects and Materials*, Technical Note 443, Artificial Intelligence Center, SRI International, Menlo Park, California (July 1988).

[Lehrer 87] N.B. Lehrer, G. Reynolds, and J. Griffith, "Initial Hypothesis Formation in Image Understanding Using an Automatically Generated Knowledge Base," *Proc. DARPA Image Understanding Workshop*, Los Angeles, California, pp. 521–537 (23–25 February 1987).

[Matsumoto 81] K. Matsumoto, M. Naka, and H. Yamanoto, "A New Clustering Method for LANDSAT Images Using Local Maximums of a Multidimensional Histogram," *Machine Processing of Remotely Sensed Data*, Purdue University, West Lafayette, Indiana, pp. 321–325 (23–26 June 1981).

[Matsuyama 85] T. Matsuyama and V. Hwang, "SIGMA: A Framework for Image Understanding—Integration of Bottom-Up and Top-Down Analyses," *Proc. 9th Int. Jnt. Conf. on Artificial Intelligence*, Los Angeles, pp. 908–915 (18–23 August 1985).

[Milner 70] D. Milner, *A Hierarchical Picture Interpretation System Utilising Contextual Information*, Dept. of Machine Intelligence and Perception, University of Edinburgh (1970).

[Mitchell 77] O.R. Mitchell, C.R. Myers, and W. Boyne, "A Max-Min Measure for Image Texture Analysis," *IEEE Trans. on Computers*, Vol. C-26, No. 4, pp. 408–414 (April 1977).

[Mitchell 78] O.R. Mitchell and S.G. Carlton, "Image Segmentation Using a Local Extrema Texture Measure," *Pattern Recognition*, Vol. 10, No. 3, pp. 205–210 (1978).

[Mitchell 79] O.R. Mitchell, S.P. Lutton, and S.P. Su, "Texture Image Segmentation Using Local Extrema," *Proc. IEEE Conf. on Pattern Recognition and Image Processing*, Chicago, Illinois, pp. 508–511 (August 1979).

[Nagao 76] M. Nagao, H. Tanabe, and K. Ito, "Agricultural Land Use Classification of Aerial Photographs by Histogram Similarity Method," *Proc. 3rd Int. Jnt. Conf. on Pattern Recognition*, Coronado, California, pp. 669–672 (8–11 November 1976).

[Narendra 77] P.M. Narendra and M. Goldberg, "A Non-Parametric Clustering Scheme for LANDSAT," *Pattern Recognition*, Vol. 9, pp. 207–215 (1977).

[Ohlander 78] R. Ohlander, K. Price, and D.R. Reddy, "Picture Segmentation Using a Recursive Region Splitting Method," *Computer Graphics and Image Processing*, Vol. 8, No. 3, pp. 313–333 (December 1978).

[Ohta 80] Y. Ohta, *A Region-Oriented Image-Analysis System by Computer*, Ph.D. Thesis, Dept. of Information Sciences, Kyoto University, Kyoto, Japan (March 1980).

[O'Rourke 84] J. O'Rourke and K.R. Sloan, Jr., "Dynamic Quantization: Two Adaptive Data Structures for Multidimensional Spaces," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. PAMI–6, No. 3, pp. 266–280 (May 1984).

[Parikh 78] J.A. Parikh and A. Rosenfeld, "Automatic Segmentation and Classification of Infrared Meteorological Satellite Data," *IEEE Trans. on Systems, Man and Cybernetics*, Vol. SMC–8, No. 10, pp. 736–743 (October 1978).

[Parma 80] C.C. Parma, A.R. Hanson, and E.M. Riseman, *Experiments in Schema-Driven Interpretation of a Natural Scene*, COINS Technical Report 80–10, University of Massachusetts at Amherst, Amherst, Massachusetts (April 1980). Also in R. Haralick and J.C. Simon (eds.), *NATO Advanced Study Institute on Digital Image Processing*, Bonas, France (1980).

[Pentland 84] A.P. Pentland, "Fractal-Based Description of Natural Scenes," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. PAMI–6, No. 6, pp. 661–674 (November 1984). Also published in *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, Washington, D.C., pp. 201–209 (19–23 June 1983) and in *Proc. DARPA Image Understanding Workshop*, Arlington, Virginia, pp. 184–192 (23 June 1983).

[Price 79] K. Price and R. Reddy, "Matching Segments of Images," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. PAMI–1, No. 1, pp. 110–116 (January 1979).

[Price 81] K.E. Price, "Relaxation Matching Applied to Aerial Images," *Proc. DARPA Image Understanding Workshop*, Washington, D.C., pp. 22–25 (23 April 1981).

[Price 82] K.E. Price, "Symbolic Matching of Images and Scene Models," *Proc. Workshop on Computer Vision: Representation and Control*, Rindge, New Hampshire, pp. 105–112 (23–25 August 1982). Also published in *Proc. DARPA Image Understanding Workshop*, Palo Alto, California, pp. 299–308 (15–16 September 1982).

[Piech 77] K.R. Piech, D.W. Gaucher, J.R. Schott, and P.G. Smith, "Terrain Classification Using Special Color Analysis Techniques," *Photogrammetric Engineering and Remote Sensing*, Vol. 43, pp. 507–513 (1977).

[Piech 80] K.R. Piech, "Material Identification Using Broad Band Visible Data," *Proc. SPIE Conf. on Image Processing for Missile Guidance*, San Diego, California, pp. 370–375 (29 July–1 August 1980).

[Pietikäinen 82] M. Pietikäinen and A. Rosenfeld, "Edge-Based Texture Measures," *Proc. 6th Int. Jnt. Conf. on Pattern Recognition*, Munich, Germany, pp. 298–300 (19–22 October 1982).

[Reboh 81] R. Reboh, *Knowledge Engineering Techniques and Tools in the Prospector Environment*, Technical Note 243, Artificial Intelligence Center, SRI International, Menlo Park, California (May 1981).

[Rosenfeld 71] A. Rosenfeld and M. Thurston, "Edge and Curve Detection for Visual Scene Analysis," *IEEE Trans. on Computers,* Vol. C-20, No. 5, pp. 562-569 (May 1971).

[Rumelhart 86a] D.E. Rumelhart, J.L. McClelland, and the PDP Research Group, *Parallel Distributed Processing: Explorations in the Microstructure of Cognition. Volume 1: Foundations,* Bradford Books/MIT Press, Cambridge, Massachusetts (1986).

[Rumelhart 86b] D.E. Rumelhart, G.E. Hinton, and R.J. Williams, "Learning Internal Representations by Error Propagation," in [Rumelhart 86a], pp. 318-362. Also published as ICS Report 8506, Institute for Cognitive Science, University of California, San Diego (September 1985).

[Sadjadi 79] F.A. Sadjadi and E.L. Hall, "Segmentation Using Material Information," *Proc. SPIE Conf. on Digital Processing of Aerial Images,* Vol. 186, pp. 254-259 (1979).

[Salton 83] G. Salton, E.A. Fox, and H. Wu, "Extended Boolean Information Retrieval," *Communications of the ACM,* Vol. 26, No. 12, pp. 1022-1036 (December 1983).

[Samuel 67] A.L. Samuel, "Some Studies in Machine Learning Using the Game of Checkers. II—Recent Progress," *IBM Journal of Research and Development,* Vol. 11, pp. 601-617 (November 1967).

[Sclove 80] S.L. Sclove, "Application of the Conditional Population-Mixture Model to Image Segmentation," *Proc. 5th Int. Jnt. Conf. on Pattern Recognition,* Miami Beach, Florida, pp. 518-521 (1-4 December 1980).

[Stanfill 86] C. Stanfill and D. Waltz, "Toward Memory-Based Reasoning," *Communications of the ACM,* Vol. 29, No. 12, pp. 1213-1228 (December 1986).

[Swain 81] P.H. Swain, S.B. Vardeman and J.C. Tilton, "Contextual Classification of Multispectral Image Data," *Pattern Recognition,* Vol. 13, No. 6, pp. 429-441 (1981).

[Terzopoulos 80] D. Terzopoulos and S.W. Zucker, "Texture Discrimination Using Intensity and Edge Co-Occurrence," *Proc. 5th Int. Jnt. Conf. on Pattern Recognition,* Miami Beach, Florida, pp. 565-569 (1-4 December 1980).

[Thosar 73] R.B. Thosar, *Estimation of Probability Density Using Signature Tables for Application to Pattern Recognition,* Memo AIM-198, STAN-CS-73-364, Stanford Artificial Intelligence Laboratory, Stanford, California (May 1973).

[Touchberry 77] R. Touchberry and R. Larson, "Symbolic Analysis of Images Using Prototype Similarity," *Proc. DARPA Image Understanding Workshop,* Minneapolis, Minnesota, pp. 78-82 (April 1977).

[Trivedi 84a] M.M. Trivedi, C.A. Harlow, R.W. Conners, and S. Goh, "Object Detection Based on Gray Level Cooccurrence," *Computer Vision, Graphics, and Image Processing,* Vol. 28, pp. 199-219 (1984).

[Trivedi 84b] M.M. Trivedi, C.A. Harlow, and S. Goh, "An Optimal Approach for the Incorporation of Context," Technical Report, Remote Sensing and Image Processing Laboratory, Dept. of Electrical Engineering, Louisiana State University, Baton Rouge, Louisiana (September 1984).

[Voorhees 87a] H. Voorhees and T. Poggio, "Detecting Blobs as Textons in Natural Images," *Proc. DARPA Image Understanding Workshop,* Los Angeles, California, pp. 892-899 (23-25 February 1987).

[Voorhees 87b] H. Voorhees and T. Poggio, "Detecting Textons and Texture Boundaries in Natural Images," *Proc. 1st Int. Conf. on Computer Vision,* London, England, pp. 250-258 (8-11 June 1987).

[Wacker 69] A.G. Wacker, *A Cluster Approach to Finding Spatial Boundaries in Multispectral Imagery*, Information Note 122969, Laboratory for Applications of Remote Sensing, Purdue University, West Lafayette, Indiana (1969).

[Waltz 87] D.L. Waltz, "Applications of the Connection Machine," *Computer*, Vol. 20, No. 1, pp. 85–97 (January 1987).

[Wesley 86] L.P. Wesley, "Evidential Knowledge-Based Computer Vision," *Optical Engineering*, Vol. 25, No. 3, pp. 363–379 (March 1986). Also published as Technical Note 374, Artificial Intelligence Center, SRI International, Menlo Park, California (January 1986).

[Weszka 75] J.S. Weszka, A. Rosenfeld, E.J. Carton, R.L. Kirby, and J.M. Mohr, "A Comparative Study of Texture Measures for Terrain Classification," TR–361, Computer Vision Laboratory, Computer Science Center, University of Maryland, College Park, Maryland (March 1975).

[Weszka 76] J.S. Weszka, C.R. Dyer, and A. Rosenfeld, "A Comparative Study of Texture Measures for Terrain Classification," *IEEE Trans. on Systems, Man and Cybernetics*, Vol. SMC–6, No. 4, pp. 269–285 (April 1976).

[Wiersma 76] D.J. Wiersma and D. Landgrebe, "Use of Spatial Characteristics for the Improvement of Multispectral Classification of Remotely Sensed Data," *Symp. on Machine Processing of Remotely Sensed Data*, Lafayette, Indiana, pp. 18–25 (June 1976).

[Yakimovsky 74] Y.Y. Yakimovsky, "On the Recognition of Complex Structures: Computer Software Using Artificial Intelligence Applied to Pattern Recognition," *Proc. 2nd Int. Jnt. Conf. on Pattern Recognition*, Copenhagen, Denmark, pp. 345–353 (13–15 August 1974). Also published as Technical Memo 33–693, Jet Propulsion Laboratory (1 May 1974).

[Zadeh 74] L. Zadeh, *Fuzzy Logic and Approximate Reasoning*, Electronics Research Lab., University of California, Berkeley (November 1974).

[Zucker 75] S.W. Zucker, A. Rosenfeld, and L.S. Davis, "Picture Segmentation by Texture Discrimination," *IEEE Trans. on Computers*, Vol. C–24, No. 12, pp. 1228–1233 (December 1975).

[Zucker 78] S.W. Zucker, "Vertical and Horizontal Processes in Low Level Vision," in A.R. Hanson and E.M. Riseman (eds.), *Computer Vision Systems*, Academic Press, New York, pp. 187–195 (1978). Also published as TR 77–4 (Revised), Computer Vision and Graphics Laboratory, Department of Electrical Engineering, McGill University, Montreal, Canada (1977).